UNIVERSITAT
JAUME•I

DOCTORAL THESIS

Application and Use of Quality Metrics for the Prediction of Grasp Success and Evaluation of Artificial Hands

---

Author:
Carlos Rubert Escuder

Supervisor:
Dr. Antonio Morales Escrig

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science

in the

Doctoral School
Universitat Jaume I

June 14, 2018

Programa de Doctorat en Informática / Programa de Doctorado en Informática

Escola de Doctorat de la Universitat Jaume I / Escuela de Doctorado de la Universitat Jaume I

Application and Use of Quality Metrics for the Prediction of Grasp Success and Evaluation of Artificial Hands

Memòria presentada per Carlos Rubert Escuder per a optar al grau de doctor per la Universitat Jaume I
Memoria presentada por Carlos Rubert Escuder para optar al grado de doctor por la Universitat Jaume I

Doctorando: Carlos Rubert Escuder **Director: Antonio Morales Escrig**

Castelló de la Plana, Juny de 2018 / Castellón de la Plana, Junio de 2018

Finançament rebut
Agències finançadores del doctorand:

- A la Fundació Balaguer Gonel Hermanos l'ajuda concedida per a la realització d'una estada predoctoral al Max-Planck Institute que ha possibilitat la sol.licitud de la "Menció de Doctorat Internacional" per a la present tesi doctoral.

Agències finançadores del projecte de recerca o dels recursos materials específics del grup de recerca.

- Ministerio de Educacion, Ciencia y Tecnología: R31-2008-000-10062-0.
- Ministerio de Economía y Competitividad: DPI2014-60635-R y DPI2011-27846.
- Generalitat Valenciana: PROMETE0/2009/052 and PROMETEOII/2014/028.
- Fundació Caixa Castelló-Bancaixa: P1-1B2011-54 i PI-1B2011-25.
- Universitat Jaume I: P1-1B2011-54.
- Max-Planck Society.

Financiación recibida
Agencias financiadoras del doctorando:

- A la Fundación Balaguer Gonel Hermanos la ayuda concedida para la realización de una estancia predoctoral en el Max-Planck Institute que ha posibilitado la solicitud de la "Mención de Doctorado Internacional" para la presente tesis doctoral.

Agencias financiadoras del proyecto de investigación o de los recursos materiales específicos del grupo de investigación:

- Ministerio de Educacion, Ciencia y Tecnología: R31-2008-000-10062-0.
- Ministerio de Economía y Competitividad: DPI2014-60635-R and DPI2011-27846.
- Generalitat Valenciana: PROMETE0/2009/052 and PROMETEOII/2014/028.
- Fundació Caixa Castelló-Bancaixa: P1-1B2011-54 and PI-1B2011-25.
- Universitat Jaume I: P1-1B2011-54.
- Max-Planck Society.

"First, let me dispose of Socrates because I am sick and tired of this pretense that knowing you know nothing is a mark of wisdom."

Isaac Asimov, *The Relativity of Wrong*.

Universitat Jaume I

Abstract

Universitat Jaume I
Doctoral School

Doctor of Philosophy in Computer Science

Application and Use of Quality Metrics for the Prediction of Grasp Success and Evaluation of Artificial Hands

by Carlos Rubert Escuder

Artificial manipulation has been one of the great areas of interest in robotics for decades. Find a proper grasp to seize objects and design robotic hands capable of such grips are two of the main issues in this field.

First, grasp synthesis is focused in generate a proper hand configuration to resolve a manipulation task. This derives in finding a good grasp among the infinite set of candidates. The selection of this configuration implies to use a methodology to evaluate grasps hypotheses. Quality metrics have been widely used in the robotics field to evaluate grasp candidates. However, experiments showed these metrics are not capable to classify properly these grasps as *good* or *bad*.

On the other hand, finding a *good* grasp is also dependent on the gripper capabilities. For this purpose, different robotic grippers and artificial hands have been developed. Although there are notable differences between grippers, there is no proper way to evaluate their performance. Different approaches have been defined to solve this problem, but none of them seemed to give a convincing evaluation. We consider quality metrics can be applied in this field to provide an evaluation on different gripper capabilities and also, generate a benchmark to evaluate artificial hands.

The work described in this thesis has three objectives: first, study the characteristics of grasp quality metrics. Second, find a combination of metrics capable of predict the success of a grasp hypotheses. And third, the appliance of these measures to evaluate the performance of artificial hands.

In order to achieve these objectives, we perform an exhaustive statistical study on the characteristics of the most common quality metrics. Then, different classification methods are trained to generate a model capable to predict the outcome of real grasp executions. Finally, quality metrics are considered as evaluators of different manipulator *properties*. A methodology is proposed in order to evaluate different hands and design improvements.

The results of this thesis will provide a better understanding of quality metrics. A model to predict grasp success. And a methodology to evaluate the functionality of artificial hands.

Universitat Jaume I

Resumen

Universitat Jaume I
Doctoral School

Doctor of Philosophy in Computer Science

Application and Use of Quality Metrics for the Prediction of Grasp Success and Evaluation of Artificial Hands

by Carlos Rubert Escuder

La manipulación artificial ha sido una de las grandes áreas de interes en robótica durante décadas. Encontrar un agarre adecuado para sostener objetos y deñar manos robóticas capades de dichos agarres son dos de los principales problemas en este campo.

Primero, la síntesis del agarre está centrada en generar una configuración de la mano adecuada para resolver una tarea de manipulación. Esto deriva en encontrar un buen agarre entre las infinitas posiblidades. La selección de esta configuración implica el uso de una metodología para evaluar las hipótesis de agarre. Las métricas de calidad han sido ampliamente utilizadas en el campo de la robótica para evaluar agarres candidatos. Sin embargo, los experimentos han mostrado que estas métricas no son capaces de clasificar de forma adecuada estos agarres como *buenos* o *malos*.

Por otra parte, encontrar un *buen* agarre depende también de las capacidades del manipulador. Con este propósito, distintos manipuladores robóticos y manos artificiales han sido desarrollados. Aunque hay diferencias notables entre manipuladores, no existe una forma adecuada de evaluar su rendimiento. Distintos intentos se han realizado para resolver este problema, pero ningun de ellos parece porporcionar una evaluación convincente. Nostros consideramos que las métricas de calidad pueden ser aplicadas en este campo para proporcionar una evaluación de las distintas capacidades de los manipuladores, y también proporcionar unos puntos de referencia para evaluar manos artificiales.

El trabajo descrito en esta tesis tiene tres objetivos: primero, estudiar las características de las métricas de calidad del agarre. Segundo, encontrar una combinación de métricas capaz de predecir el éxito de una hipótesis de agarre. Y tercero, la aplicación de estas medidas para evaluar el rendimiento de manos artificiales.

Con el objetivo de lograr estos objetivos, se realiza un exahustivo estudio estadístico sobre las características de las métricas de calidad más comunes. A continuación, se entrenan distintos métodos de clasificación para generar un modelo capaz de predecir el resultado de ejecuciones de agarre reales. Finalmente, las métricas de caldiad son consideradas como evaluadores de distintas propiedades en los manipuladores. Se propone una metodología con el objetivo de evaluar distintas manos y mejoras de diseño.

Los resultados de esta tesis proporcionarán un mejor entendimiento de las métricas de calidad. Un modelo para predecir el éxito del agarre. Y una metodología para evaluar la funcionalidad de manos artificiales.

Acknowledgements

First, I want to thank my supervisor, Antonio Morales, for all the advices, corrections and knowledge during all these years. First, as my professor in the bachelor degree in Computer Science and later, for giving me the chance to start an internship at the Robotic Intelligence Lab during my Master thesis.

I want to thank the members, present and past, of the Robotic Intelligence Lab and IRSLab at Universitat Jaume I, for sharing their knowledge and expertise. But also, for being not only workmates, but great friends and supporters during all this years. I want to thank specially Beatriz León, my advisor during my Master thesis, whose knowledge and support persuade me to start this Doctoral Thesis.

I want to thank Ángel Durán and Javier Felip for all his support and help with the Tombatossals Robot. And specially Ángel, for all his help and advices during the long time it took me to write this thesis.

I also would like to thank the members of the Devalhand Project in the Biomechanical and Ergonomics group at the Universitat Jaume I. Whose knowledge in hand design and prosthesis allowed me to expand my mind and work in new ideas for the experiments.

Next, I want to thank Jeannette Bohg for her support and confidence in my work, and for allowing me to join her research group at the Max-Planck Institute in Tübingen (Germany). This was a great step for my growing as researcher and for fulfill this thesis. I want to thank also the members at the MPI-IS, for her knowledge and help during my research stays. Specially Vincent, Daniel and Jan for all the troubleshooting help with the Apollo Robot.

Finally, I want to thanks all my friends, family and rugby teammates. Their friendship, support and empathy were a way of scape from daily stress. Without them, I wouldn't be able to continue in this hard road and make me strive to achieve my goals.

The research done in this thesis wouldn't been possible without the financial support to the author and research centres from:

- Ministerio de Educacion, Ciencia y Tecnología: R31-2008-000-10062-0.
- Ministerio de Economía y Competitividad: DPI2014-60635-R and DPI2011-27846.
- Generalitat Valenciana: PROMETE0/2009/052 and PROMETEOII/2014/028.
- Fundació Caixa Castelló-Bancaixa: P1-1B2011-54 and PI-1B2011-25.
- Universitat Jaume I: P1-1B2011-54.
- Max-Planck Society.
- To Balaguer Gonel Hermanos Foundation, the award granted for the completion of a predoctoral stay at the Max-Planck Institute, that has enabled the application to the "Mention of International Doctorate" for this doctoral thesis.

Contents

Abstract **vii**

Acknowledgements **xi**

1 Introduction **1**
- 1.1 Motivation 2
- 1.2 Objectives 8
- 1.3 Outline 8

2 Basic foundations **11**
- 2.1 Grasp theory 11
- 2.2 Grasp analysis 14
- 2.3 Grasp planning 17
- 2.4 Grasp quality metrics 18
  - 2.4.1 Group A: Algebraic properties of G 19
  - 2.4.2 Group B: Distribution of contact points 20
  - 2.4.3 Group C: Magnitude of forces 22
  - 2.4.4 Group D: Configuration of the manipulator 23

3 Experimental platform **25**
- 3.1 Simulation environment 25
  - 3.1.1 OpenHand 26
- 3.2 Artificial hands 28
- 3.3 Object models 32
- 3.4 Grasp generation 32
- 3.5 Robotic platforms 35
  - 3.5.1 Robotic Grippers 36
- 3.6 Real grasp evaluation 37
- 3.7 Robotic hands evaluation 38
- 3.8 Grasp databases 39

4 Quality metrics analysis **41**
- 4.1 Grasp generation study 42
  - 4.1.1 Results 43
- 4.2 Variability 45
  - 4.2.1 Probability density function 46
  - 4.2.2 Statistic values and normalization 47
  - 4.2.3 Results 48
    - $\mathbf{Q_{A1}}$ - Smallest singular value of $\mathbf{G}$ 49
    - $\mathbf{Q_{A2}}$ - Volume of $\mathbf{G}$ in the wrench space 50
    - $\mathbf{Q_{A3}}$ - Grasp Isotropy Index 51
    - $\mathbf{Q_{B1}}$ - Distance between the centroid of the contact polygon and the center of mass of the object 52

$\mathbf{Q_{B2}}$ - Area of the grasp polygon . . . 53
$\mathbf{Q_{B3}}$ - Shape of the grasp polygon . . . 54
$\mathbf{Q_{C1}}$ - Smallest maximum wrench to be resisted . . . 55
$\mathbf{Q_{C2}}$ - Volume of the convex hull . . . 56
$\mathbf{Q_{D1}}$ - Posture of manipulator joints . . . 57
$\mathbf{Q_{D2}}$ - Inverse of the condition number of $\mathbf{G_J}$ . . . 58
Normalization . . . 59
4.3 Sensitivity . . . 65
4.3.1 Methods . . . 65
4.3.2 Results . . . 67
Preliminary study . . . 67
$\mathbf{Q_{A1}}$ - Smallest singular value of $\mathbf{G}$ . . . 68
$\mathbf{Q_{A2}}$ - Volume of $\mathbf{G}$ in the wrench space . . . 69
$\mathbf{Q_{A3}}$ - Grasp Isotropy Index . . . 70
$\mathbf{Q_{B1}}$ - Distance between the centroid of the contact polygon and the center of mass of the object . . . 71
$\mathbf{Q_{B2}}$ - Area of the grasp polygon . . . 72
$\mathbf{Q_{B3}}$ - Shape of the grasp polygon . . . 73
$\mathbf{Q_{C1}}$ - Smallest maximum wrench to be resisted . . . 74
$\mathbf{Q_{C2}}$ - Volume of the convex hull . . . 75
$\mathbf{Q_{D1}}$ - Posture of manipulator joints . . . 76
$\mathbf{Q_{D2}}$ - Inverse of the condition number of $\mathbf{G_J}$ . . . 77
Summary . . . 78
4.4 Correlation . . . 79
4.4.1 Methods . . . 79
4.4.2 Results . . . 79
4.5 Discussion & conclusions . . . 82
Conclusions . . . 83
Limitations . . . 84
Future & related work . . . 84

5 Grasp success prediction **87**
5.1 Introduction . . . 88
5.2 Foundations and related work . . . 89
5.2.1 Grasps database . . . 89
5.2.2 A physics metric . . . 90
5.2.3 Selection of quality metrics . . . 90
5.2.4 Human oracle . . . 91
5.2.5 Related work . . . 91
5.3 Binary classifiers . . . 92
Logistic regression . . . 92
Naive bayes . . . 93
k-Nearest neighbors . . . 93
Classification trees . . . 94
Support vector machines . . . 95
Gaussian processes . . . 95
Neural networks . . . 96
Evaluation metrics for classifiers . . . 96
5.4 Case study 1: Human oracle . . . 97
5.4.1 Feature vector and label . . . 97
5.4.2 Methodology . . . 97

Dataset 97
Training, validation and testing 97
5.4.3 Results 98
Comparison of classification methods 98
Quality metrics predictive capability 99
Discussion 101
5.4.4 Conclusions 104
5.5 Case study 2: Real grasp execution 105
5.5.1 Notation 105
5.5.2 Methodology 107
Experimental protocol 107
Experimental score 110
Input feature vector and dataset 111
Classification methods 112
5.5.3 Results 113
Experimental score vs. human labels vs. physics score 113
Classification algorithms 114
5.5.4 Discussion 115
Are humans good grasp predictors? 115
Can grasp quality metrics predict the outcome of real grasp executions? 116
Can contextual information help improve the grasp prediction? 117
5.6 Case study 3: Metrics limitations and contextual information 121
5.6.1 Methodology 121
Grasp evaluation 122
Classification models 122
Object models 122
Grasps dataset 122
5.6.2 Results 123
Tombatossals grasp dataset 123
Apollo grasp dataset 124
Individual metrics 124
5.6.3 Discussion 125
5.6.4 Prediction model 125
Discussion 127
5.7 Generalization of the prediction model 127
5.7.1 Methodology 128
5.7.2 Results 128
5.7.3 Discussion 128
5.8 Conclusions 129
Case study 1: human oracle 129
Case study 2: real grasp execution 129
Case study 3: metrics limitations and 3-categories scale 130
5.8.1 Prediction model 130
5.8.2 Limitations 131
5.8.3 Future & related work 131

6 Artificial hands evaluation **133**
6.1 Introduction 134
6.2 Artificial hand properties 135
  6.2.1 Methodology 135
  6.2.2 Results 136
    Variability analysis 136
    Sensitivity analysis 136
  6.2.3 Discussion 140
6.3 Quality metrics and anthropomorphic index 141
  6.3.1 Anthropomorphic index 141
  6.3.2 Methodology 142
    Anthropomorphic hands 142
    Objects 142
    Experiments 143
  6.3.3 Results 144
  6.3.4 Discussion 147
6.4 Grasp success prediction and evaluation of artificial hands 149
  6.4.1 Methodology 149
  6.4.2 Results 150
  6.4.3 Discussion 150
6.5 Hand design improvement 151
  6.5.1 Methodology 151
  6.5.2 Results 152
  6.5.3 Discussion 152
6.6 Conclusions 154
  6.6.1 Artificial hand properties 154
  6.6.2 Comparison with hand indexes 155
  6.6.3 Grasp success prediction and evaluation of artificial hands . . . 155
  6.6.4 Limitations 155
  6.6.5 Future & Related Work 156

7 Conclusions **157**
7.1 Contributions 157
  7.1.1 Quality metrics analysis 157
  7.1.2 Grasp success prediction 157
  7.1.3 Artificial hands evaluation 158
7.2 Future work 158
7.3 Related publications 158
7.4 Contribuciones 161
  7.4.1 Análisis de métricas de calidad 161
  7.4.2 Predicción del éxito del agarre 161
  7.4.3 Evaluación de manos artificiales 162
7.5 Trabajo Futuro 162
7.6 Publicaciones relacionadas 162

A Quality Metrics Analysis **165**
A.1 Numeric Results from Sensitivity Analysis 165

B Grasp prediction for different combinations of metrics **169**
B.1 Precission results for best combinations of metrics 169

Bibliography **171**

List of Figures

1.1 Example showing different robots in industrial environments. 2
1.2 Example of robots working in stochastic environments. In a) the *UJI Baxter* robot has to pick objects in a shelf system during the Amazon Picking Challenge 2015. In b) the *UJI librarian* robot is picking books from a library. 3
1.3 Example of a metric misjudging a grasp. The quality of grasp A is worst than B. However the grasp B fails when applied in real world. . 4
1.4 Some of the most known robotic grippers. 5
1.5 Extract of the taxonomy of human grasps, full table can be found in Feix et al., 2016. This tables shows most of the common grasp poses employed by humans in his ADL. 7

2.1 Notation for a contact between two rigid bodies. 12
2.2 Relationship between grasp matrix, hand Jacobian and grasp Jacobian (**Leon2013d**). 17

3.1 The OpenRAVE simulator. 26
3.2 The OpenHand interface. 27
3.3 Robotic grippers. 28
3.4 Prosthetic hands. 30
3.5 Sample of the different objects from the KIT Database. 32
3.6 Virtual object models used for the real experiments from the 3DNet database. 33
3.7 Example of the approach-rays generation process used by OpenRAVE: (left) square grid of points around the object and (right) the projection of this box over the object surface. 34
3.8 Example showing the parameters needed for grasp generation. 34
3.9 The Apollo robot system. 35
3.10 The Tombatossals robot system. 36
3.11 Real models of the robotic grippers. 37
3.12 Example of the experimental setup with the Tombatossals robot. 38

4.1 Example showing the variations in δ and α studied to generate approach rays 42
4.2 Results of the grasp quality metrics for the different approach-rays generation methods detailed in Table 4.1. The grasp quality scale varies to better how the results. 44
4.3 Histogram and Probability Density Function for metric Q_{A1} 49
4.4 Histogram and Probability Density Function for metric Q_{A2}. Detailed graph for *PDFs* in range $[0, 3]$. 50
4.5 Histogram and Probability Density Function for metric Q_{A3} 51
4.6 Histogram and Probability Density Function for metric Q_{B1} 52
4.7 Histogram and Probability Density Function for metric Q_{B2} 53

4.8 Histogram and Probability Density Function for metric Q_{B3} 54
4.9 Histogram and Probability Density Function for metric Q_{C1} 55
4.10 Histogram and Probability Density Function for metric Q_{C2} 56
4.11 Histogram and Probability Density Function for metric Q_{D1} 57
4.12 Histogram and Probability Density Function for metric Q_{D2} 58
4.13 Histograms showing the comparison of different normalization methods for each metric. Red line points out the *Percentile 50*. 60
4.10 Variations in the hand posture along their axes. 65
4.11 Graph comparing the Global Sensitivity Index between different set sizes for each quality metric. 67
4.12 GSI_N for metric Q_{A1}. 68
4.13 Variation in the metric Q_{A1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 68
4.14 GSI_N for metric Q_{A2}. 69
4.15 Variation in the metric Q_{A2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 69
4.16 GSI_N for metric Q_{A3}. 70
4.17 Variation in the metric Q_{A3} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 70
4.18 GSI_N for metric Q_{B1}. 71
4.19 Variation in the metric Q_{B1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 71
4.20 GSI_N for metric Q_{B2}. 72
4.21 Variation in the metric Q_{B2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 72
4.22 GSI_N for metric Q_{B3}. 73
4.23 Variation in the metric Q_{B3} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 73
4.24 GSI_N for metric Q_{C1}. 74
4.25 Variation in the metric Q_{C1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 74
4.26 GSI_N for metric Q_{C2}. 75
4.27 Variation in the metric Q_{C2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 75
4.28 GSI_N for metric Q_{D1}. 76

4.29 Variation in the metric Q_{D1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 76
4.30 GSI_N for metric Q_{D2}. 77
4.31 Variation in the metric Q_{D2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps. . . . 77
4.32 Comparison between highest correlated metrics. Outlier values are discarded. 81

5.1 Examples of *Successful* (green mark) and *Unsuccessful* (red mark) grasps. 89
5.2 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{A1}. Histogram with log scale of y-axis. 99
5.3 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{B1}. Histogram with log scale of y-axis. 99
5.4 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{B2}. Histogram with log scale of y-axis. 100
5.5 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{B3}. Histogram with log scale of y-axis. 101
5.6 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{C2}. Histogram with log scale of y-axis. 102
5.7 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{D1}. Histogram with log scale of y-axis. 103
5.8 Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{D2}. Histogram with log scale of y-axis. 104
5.9 Example of an object with three different gravity vectors. World axis is showed as reference for the normal gravity vector (0,0,-1). 106
5.10 Relationships between candidate grasps, experimental grasps, grasp clusters and average grasps. 107
5.11 Object models used for the experiments with the Apollo robot system. Objects are: (a)Bottle 1, (b)Bottle 2, (c)Camera, (d)Lemon, (e)Bowl 1, (f)Bowl 2, (g)Toaster, (h)Jar 1 and (i)Jar 2. Subcaptions show the object name in the dataset. 108
5.12 Example of a successful real grasp execution 109
5.13 Example of an unsuccessful real grasp execution 110
5.14 Example of a failed try. The gripper throws the object when approaching to the grasp pose. 110
5.15 Distribution of grasps evaluated with quality metrics, physics metric and human label. 111
5.16 Maps of values for metrics QA2 vs QB1, QB2, QB3. Graphs on left show the FS dataset, Graphs on the right column show the SS dataset. Graphs for the *FS dataset* contain 1106 datapoints, graphs in the *SS dataset* contain 343 datapoints. 120
5.17 Object models used for the experiments with Tombatossals. Upper row 3D Models. Lower row 3D printed objects. Left column: Bottle 1, Right column: Toaster. 123

6.1 Hand models analyzed with metrics and the *AI* 142

6.2 Sample of the different objects used for calculate the Anthropomorphic Index. 143
6.3 Percentiles of Quality Metrics for Artificial Hands. 145
6.4 2D latent space of artificial hands projected over the human space. Red lines/dots denotes the space covered by the artificial hand. Latent spaces are showed for both sampling methods: grasp and random 146
6.5 Comparative between AI and Grasp Quality. 148
6.6 Distribution of grasps on each category for different artificial hand models. Green: Robust, Yellow: Fragile, Red: Futile. 151
6.7 CMC and MCP joints to be modified in the Imma Hand. 152
6.8 Sample of the different objects used from the YCB Database. 153

List of Tables

2.1 Notations 13
2.2 Summary of the selected quality metrics 19

3.1 Robot Hand Models 29
3.2 Prosthetic Hand Models 31

4.1 Parameters used in each approach-rays generation method 43
4.2 Parameters selected to generate grasp hypotheses 45
4.3 Summary of statistic values and percentiles for each metric. 48
4.4 Parameters for best fits of distribution functions and statistical values on metric $_{QA1}$ 49
4.5 Parameters for best fits of distribution functions and statistical values on metric Q_{A2} 50
4.6 Parameters for best fits of distribution functions and statistical values on metric Q_{A3} 51
4.7 Parameters for best fits of distribution functions and statistical values on metric Q_{B1} 52
4.8 Parameters for best fits of distribution functions and statistical values on metric Q_{B2} 53
4.9 Parameters for best fits of distribution functions on metric Q_{B3} 54
4.10 Parameters for best fits of distribution functions on metric Q_{A1} 55
4.11 Parameters for best fits of distribution functions on metric Q_{C2} 56
4.12 Parameters for best fits of distribution functions on metric Q_{D1} 57
4.13 Parameters for best fits of distribution functions on metric Q_{D2} 58
4.14 Statistic thresholds of the grasp quality values for each metric 59
4.15 Ranges to classify the quality of a grasp for each metric 59
4.16 Percentage of grasps accomplishing force-closure 67
4.17 GSI_N(%) for every metric 78
4.18 Correlation coefficients 80

5.1 Summary of the selected independent quality metrics 91
5.2 Comparison of classification models. We report mean/std of classification accuracy using cross validation as well as accuracy on a test set. Results are computed for two dataset splits: random and held-out objects. As baseline, the Null Error rate on each dataset is showed. . . 98
5.3 Comparison of classification methods using different metrics to evaluate models performance. 98
5.4 Performance of individual Quality Metrics to predict grasp stability. . 100
5.5 Ablation analysis on the use of Quality Metrics to predict the stability of grasps. 101

5.6 Distribution of *Stable-Unstable* grasps among the different filtered datasets. For the OGS the name indicates the gravity axis. OGSX: gravity only in axis X. OGSX-Y gravity in axis X and -Y. Only OGS models which contain grasp tries are included in the table. 112
5.7 Predictive ability between Human labels, Physics Metric and Real Experiments. Results shown are accuracy score. 113
5.8 Score report on the use of humans and Physics to predict grasp success on real experiments. Upper table shows results for human labels. Lower table shows results for physics metric. 113
5.9 Predictive ability of Human Labels and Physics metric according to specific gravity vectors. Number of averaged grasps evaluated by humans and physics metric is also provided. Only two gravity vectors (green) have enough data points to provide reliable results. 114
5.10 Predictive ability of individual quality metrics. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar). 114
5.11 Classification methods trained with the *FS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar). . . 115
5.12 Classifiers trained with the *CS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar). 116
5.13 Classifiers trained with the *SS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar). 117
5.14 Classifiers trained with the *GS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar). 118
5.15 Summary of the training with the *OGS datasets*. Reported best results with each variation of the gravity vector. For *OGSY and OGSXY datasets* there are not enough datapoints to perform a proper 10-fold CV. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar). 119
5.16 Summary of best training results obtained using different combinations of metrics as input feature vector. The training was done using the *FS model* and only with *KNN* classifier 119
5.17 Object weights (grams) of different printed models 122
5.18 Dataset distribution of grasps and scores among different objects. . . . 123
5.19 Classification results using binary score and 3-categories scale 124
5.20 Classification results using held-out grasps method for testing the prediction models with new and unseen grasp types. 124
5.21 Classification results with the Apollo grasps dataset. First column shows classification using the Binary classification system. Second column shows the 3-categories scale. 124
5.22 Classification results for individual quality metrics. First column shows the 3-categories scale. Second column considers only *Robust/Futile* grasps. 125
5.23 Classification results with the Tombatossals dataset and the proposed prediction model. First column shows the performance for filter the *Fragile* grasps. Second column shows the performance when considering only *Robust* and *Futile* grasps. 126

5.24 Classification results with the Apollo dataset and the proposed prediction model. First column shows the performance for filter the *Fragile* grasps. Second column shows the performance when considering only *Robust* and *Futile* grasps. 126
5.25 Comparison between the classification models for the 3-categories scale and the Binary Classification method. Combined datasets from Apollo and Tombatossals experiments. 126
5.26 Classification results with the Tombatossals and Apollo datasets and the proposed prediction model. First column shows the performance for filter the *Fragile* grasps. Second column shows the performance when considering only *Robust* and *Futile* grasps. 127
5.27 Classification results or the classifiers trained with the *Tombatossals* dataset and tested with the *Apollo* dataset. First column shows the performance using the 3-categories scale. Second column shows the accuracy for filtering *Fragile* grasps. Third column shows the performance considering only *Robust/Futile* grasps. 128
5.28 Classification results or the classifiers trained with the *Apollo* dataset and tested with the *Tombatossals* dataset. First column shows the performance using the 3-categories scale. Second column shows the accuracy for filtering *Fragile* grasps. Third column shows the performance considering only *Robust/Futile* grasps. 128

6.1 Statistical Values for metric Q_{A1} per hand model. Green box shows percentiles 10-90. 136
6.2 Statistical Values for metric Q_{B1} per hand model. Green box shows percentiles 10-90. 137
6.3 Statistical Values for metric Q_{B2} per hand model. Green box shows percentiles 10-90. 137
6.4 Statistical Values for metric Q_{B3} per hand model. Green box shows percentiles 10-90. 138
6.5 Statistical Values for metric Q_{C2} per hand model. Green box shows percentiles 10-90. 138
6.6 Statistical Values for metric Q_{D1} per hand model. Green box shows percentiles 10-90. 139
6.7 Statistical Values for metric Q_{D2} per hand model. Green box shows percentiles 10-90. 139
6.8 Sensitivity Values per each Hand and Metric 140
6.9 Details of the Hand Models Used 142
6.10 Range of the DoFs used to generate hand configurations in the AI evaluation. 144
6.11 Quality Evaluation in Anthropomorphic Hands 144
6.12 Anthropomorphism Index Results 146
6.13 Quality Evaluation and AI in Hands models 147
6.14 Summary of grasps distribution per each hand model. Results shows percentage and number of grasps on each category. 150
6.15 Parameters selected to generate grasp hypotheses 152
6.16 Summary of grasps distribution for Joint CMC 0. Graph shows percentage of grasps on each category. 153
6.17 Summary of grasps distribution for Joint CMC 15. Graph shows percentage of grasps on each category. 153

6.18 Summary of grasps distribution for Joint CMC 30. Graph shows percentage of grasps on each category. 154
6.19 Summary of grasps distribution for Joint CMC 45. Graph shows percentage of grasps on each category. 154

A.1 Comparison of GSI_N for each hand and metric 165
A.2 GSI_N calculated for each quality metric distinguishing between good or bad reference grasps and variations that improve or worsen the metric values 166
A.3 Distribution of the variation of reference grasps (good or bad) according to whether the variation improve or worsen the reference quality value. 167

B.1 Best results on models trained using combinations of two different metrics. 169
B.2 Best results on models trained using combinations of three different metrics. 169
B.3 Best results on models trained using combinations of four different metrics. 169
B.4 Best results on models trained using combinations of five different metrics. 170
B.5 Best results on models trained using combinations of six different metrics. 170

List of Abbreviations

ADL	**A**ctivities of **D**aily **L**ive
AI	**A**bntrhopomorphic **I**ndex
APC	**A**mazon **P**icking **C**hallenge
ARC	**A**mazon **R**obotic **C**hallenge
CM	**C**enter of **M**ass
CS	**C**luster **S**et
CT	**C**lassification **T**rees
DBOT	**D**epth **B**ased **O**bject **T**racking
DoF	**D**egrees **o**f **F**reedom
FC	**F**orce **C**losure
FS	**F**ull **S**et
GP	**G**aussian **P**rocess
GS	**G**ravity **S**et
HL	**H**uman **L**abel
KIT	**K**arlsruhe **I**nstitute of **T**echnology
KNN	**k** **N**earest **N**eighbors
MPI	**M**ax **P**lanck **I**nstitute
NN	**N**eural **N**etworks
QH	(**Q**)onvex **H**ull
QM	**Q**uality **M**etrics
UI	**U**ser **I**nterface
UJI	**U**niversitat **J**aume **I**
OGS	**O**ne **G**ravity **S**et
PM	**P**hysics **M**etric
SAH	**S**chunk **A**nthropomorphic **H**and
SDH	**S**chunk **D**exterous **H**and
SI	**S**ensitivity **I**ndex
SS	**S**ynthetic **S**et
GSI	**G**lobal **S**ensitivity **I**ndex
YCB	**Y**ale-**C**MU-**B**erkeley

Chapter 1

Introduction

For decades, the way humans manipulate objects has served as inspiration for developing grasp techniques and artificial hands in the robotics field. However, there is still a huge gap between human and robotic manipulation. Humans start grasping since they are born and after several trial and error iterations, they learn to grasp (Oztop et al., 2004). This acquired knowledge make them capable of grasping new and unknown objects without difficulty. Even the lack of some primary sense, such as vision, can be overcome using other developed senses such as tactile (Castiello et al., 1993). Be on par with such *intelligent manipulation* skill acquired after years of evolution and learning is a great challenge in the field of robotic manipulation (Yoshida et al., 2007; Dang et al., 2011).

We suggest the human manipulation ability has two key factors: the shape and kinematics of the human hand and the developed intelligence to perform manipulation tasks. There are a lot of examples of different hands and paws in nature, but only the characteristics of the human hand allowed humans to increase their intelligence (Wilson, 2010). This made up a mutualistic relationship, the hand capability for manipulation incited a great rise in the human intelligence but also, with this great intelligence, new manipulation tasks are achieved (Wilson, 2010).

We define this *intelligent manipulation* skill as *the capability to choose wisely and arrange a grasp among the infinite number of possibilities*. Humans only need the object shape and an aim task, to decide the proper grasp type (Cutkosky and Howe, 1990). Thus, basic information as object geometry and expected contact points could be enough to select a good grasp among other possibilities in the robotic field.

Grasp quality metrics use mainly contact points and object shape information. These metrics, based in different principles and heuristics, evaluate aspects of the grasp similar to those a human will consider. The area of the grasp polygon or distance between the object center of mass and the centroid of the contact polygon are examples of measurable grasp properties. Therefore, a good combination of grasp quality metrics should be able to evaluate and predict the success of a grasp hypotheses. Furthermore, these aspects of the grasp evaluable with quality metrics could serve as desirable properties an artificial hand should achieve or, in other words, evaluable properties of artificial hands.

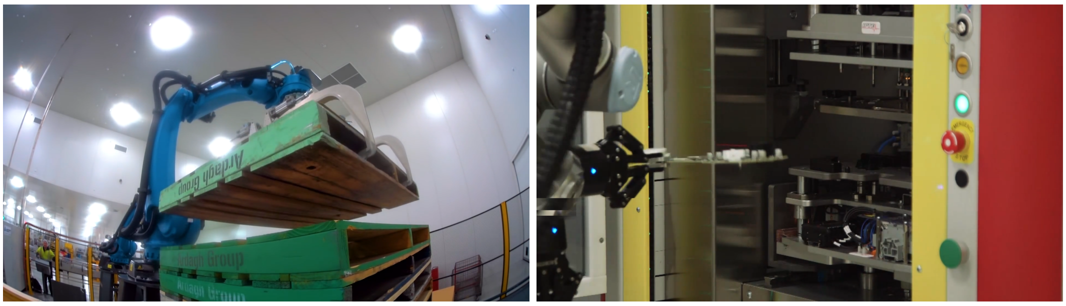
This thesis analyze two hypotheses: first, **grasp quality metrics** can be combined to predict the success of grasps candidates. And second, it is possible to use these quality metrics to define evaluable properties on artificial hands and evaluate their performance. On one hand, the characteristics of quality metrics and its predictive capabilities should be analyzed. On the other, a methodology based on a selection of metrics can be used to evaluate artificial hands performance.

1.1 Motivation

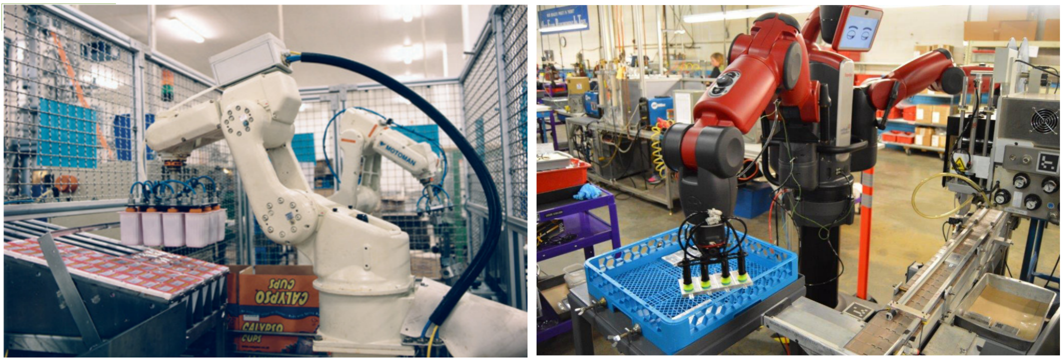
Robot applications

Currently, autonomous robotic systems are mainly used in industrial environments. Some examples are: the UR10 robot at Continental [1], the Baxter robot at Corner Dubilier [2], the Motoman SV-3 robots at Calypso Soft Drinks[3] or the KUKA KR 180 at Jamestrong Packaging[4] (Figure 1.1). The use of this robots present advantages and inconveniences. On the one hand, these environments are deterministic, robots do programmable and repetitive tasks where *artificial intelligence* and great manipulation ability is not needed.

On the other hand, traditionally this robots are designed and restraint to such environments. Adapt this robots to work in human industrial robot collaboration (HIRC) is an open problem in robotics. There are limitations in terms of automation constraints (Ore et al., 2016) and human safety (Zanchettin et al., 2016).



(1) Kuka KR 180 (2) UR 10



(3) Motoman SV-3 (4) Baxter

FIGURE 1.1: Example showing different robots in industrial environments.

The use of industrial robots has allowed to increase greatly the productivity besides reducing the production costs (Graetz and Michaels, 2015). In near future, industries will need more skilled autonomous robots. The factories of tomorrow will need robots who perform different tasks (Pedersen et al., 2016) or able to operate following *objectives* rather than orders or precise instructions (Rüßmann et al., 2015). One example of this request is the Amazon Robotic Challenge (Amazon Picking Challenge in the first edition). In this contest a robot has to serve the costumer orders, picking different objects from a shelf system and place them in a tote for delivery, Figure 1.21. This is an example of application where predefined grasps are useless, as the robot has to be able to select the proper grasp for known or unknown objects, which can also be in any (6D) pose.

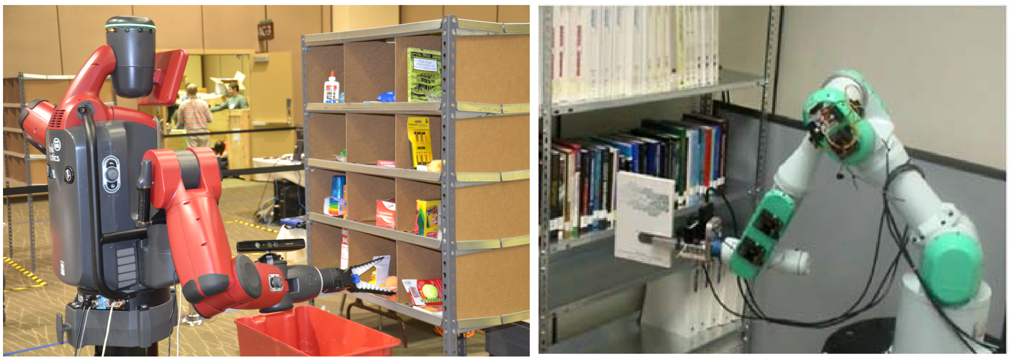
---

[1] Universal Robots https://www.universal-robots.com/

[2] Rethink Robotics http://www.rethinkrobotics.com/

[3] Yaskawa Motoman https://www.motoman.com/

[4] Kuka https://www.kuka.com/



(1) UJI Baxter in APC (2) Librarian UJI

FIGURE 1.2: Example of robots working in stochastic environments. In a) the *UJI Baxter* robot has to pick objects in a shelf system during the Amazon Picking Challenge 2015. In b) the *UJI librarian* robot is picking books from a library.

Other example is the UJI librarian robot (Figure 1.22). This robot is able to autonomously locate a book in an ordinary library, and grasp it from a bookshelf. Some of the difficulties it has to deal in this manipulation task are to perform a grasp without damaging the books, create gaps between books to perform the grasp or avoid grasping multiple books at once (Prats et al., 2008).

Another of the current aims in robotics is developing collaborative robots which should stand shoulder to shoulder with humans (Kemp et al., 2007). It is important then to develop robots capable not only of perform different manipulation task, but also to *decide* which grasp to perform in every situation. For example, a robot should be able to grasp a tool (e.g. a hammer) as a human will do (Laguillaumie et al., 2016).

This makes necessary to state a methodology that given a manipulator and an object, assess different grasp hypotheses and select, at least, one good grasp candidate among the infinite number of possibilities. Grasp synthesis is focused in solving this problem. However, the selection of the proper grasp for every situation is still an open problem.

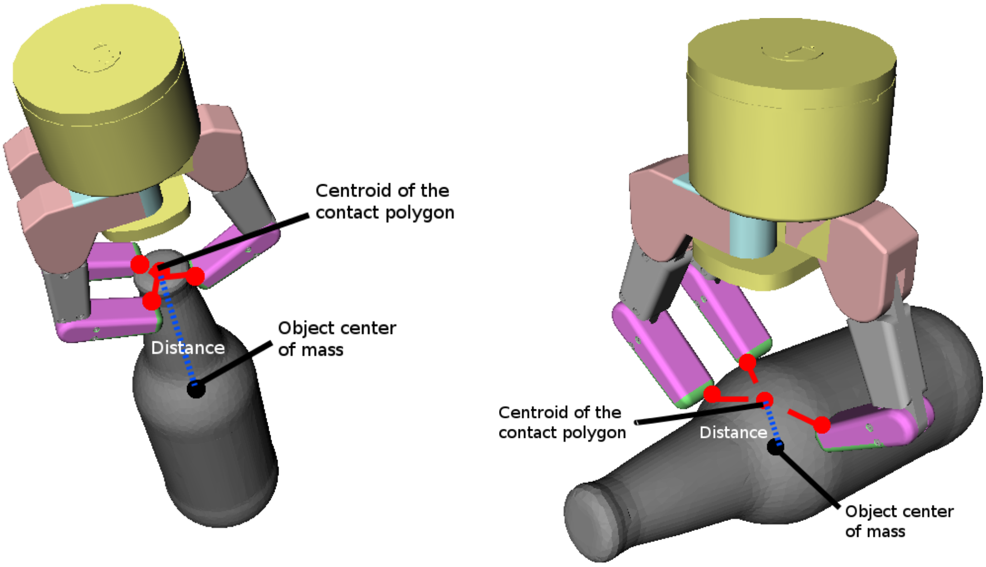
Grasp evaluation

Select a good grasp among other possibilities is not an easy task. Grasp planning aims to determine the contact points over the object surface to generate different grasp candidates. Then, this grasp hypotheses must be evaluated in order to select a good one. Grasp analysis consists on evaluate whether a grasp hypotheses is stable. For doing this evaluation, different grasp quality metrics have been developed during last decades. These metrics are based in different principles and heuristics. However, the predictive capability of the metrics should still be analyzed.

According to their theoretical definition, quality metrics evaluate different *logical* properties in a grasp. However, their practical use show these properties are not enough to classify a grasp as stable. Evaluate a grasp with quality metrics usually means, transfer it from simulated to real world, or viceversa. This process introduce some inaccuracies in the grasp itself. To know the exact location of the object or the contact points is not always possible. This means the evaluated grasp not being the real one. Taking in account dynamic simulation and pose uncertainty usually leads to more accurate evaluation of grasp success (Kim et al., 2013). However this dynamic simulations are computationally more expensive.

These metrics are usually applied to specific tasks, grippers or objects. This could introduce a bias, as a *good* value obtained using one specific gripper or object does not has to be satisfactory for other grippers or objects. The characteristics of these metrics are also quite unknown. There is no study about the behavior and limits of their values on practical cases. Neither how can be related one to other, if in case. Is assumed a better value in one metric result in a better grasp, but there is a lack of such ground truth. Even a simple binary classification of grasps as *good* or *bad* among the values of a metric doesn't show an uniform distribution.

As an example of misleading evaluation is the metric developed by Ding et al., 2001; Ponce et al., 1997. This metric measures the distance between the centroid of the contact polygon and the object's center of mass (*CM*). This principle has sense from a human perspective, but humans don't consider only one aspect of the grasp. Our knowledge make us take in account different constraints and preconceptions. Examples of this metric misjudging grasps are shown in Figure 1.3. In the first case the grasp will succeed although is far away from the object's CM. In the second case where the contact polygon is close to the object's CM, the object won't be grasped.



(1) Good grasp with poor quality value

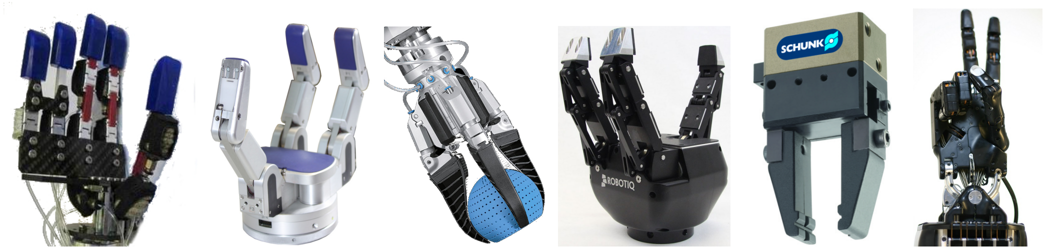
(2) Bad grasp with high quality value

FIGURE 1.3: Example of a metric misjudging a grasp. The quality of grasp *A* is worst than *B*. However the grasp *B* fails when applied in real world.

Hence, applying individual metrics could not be enough to evaluate grasps. A more complex combination of different metrics using machine learning methods showed to improve their predictive ability (Goins et al., 2014). Therefore, analyze which metrics are more accurate to predict the success of a grasps and combine them for the evaluation of a grasp can be a better approach to evaluate grasp hypotheses.

Robotic hands

The grasp planning not only depends in the selection of the grasp pose or the final contact points. It is important to consider the manipulator being used and its characteristics. Different studies aimed to classify artificial hands. The work of Kappassov et al., 2013 classify robotic hands in two categories according to its design principles: industrial grippers and anthropomorphic (or artificial) hands. According to its applications, the study of Tai et al., 2016 classify robotic hands as: industrial grippers, grippers for known or unknown environments, grippers for fragile objects, medical applications, micro and nano grippers and soft fabric grippers.



(1) ArmarIII hand (2) Barrett hand (3) Festo multichoice (4) Robotiq 3-fingers (5) Schunk MPG (6) Shadow Dexterous

FIGURE 1.4: Some of the most known robotic grippers.

Figure 1.4 show some of the most known robotic manipulators: the *Schunk MPG*[5] is an example of parallel gripper designed for industrial environments; the *Shadow Dexterous* and *Armar-III*[6] hands are anthropomorphic hands designed for humanoid robots, capable to perform a wide variety of different grasps; the other three grippers, *Barrett*[7], *Festo*[8] and *Robotiq*[9], are examples of industrial grippers with more DoF than the classical parallel grippers, which allow them to perform a wide variety of grasps.

The design and construction of suitable artificial grippers is one of the most active fields on the robotics realm. Through the years many designs have been proposed, with different degrees of success. Two trends have stood out in the approach to gripper design. On one side there is the construction of robot grippers that try to keep the trade-off between the functionality of the design and its simplicity and low cost. This has lead to devices composed of up to four fingers, where issues like robustness, under-actuated mechanisms, compliance and softness have been the most relevant features. This type of grippers are focused to be used by robots on a variety of manipulation tasks, and their aim is to provide robots with capable manipulation devices.

On the other side, all humanoid robots are equipped with robot hands that, in most cases, are intended to grasp and manipulate objects in the operating scenarios of the robots. These are, often, human-centered scenarios where the target objects are designed for human manipulation. It is important for this purpose to be able to measure the hand's capability to grasp the objects that will appear in the daily working routine. For this reason, methodologies to evaluate not only the mechanical structure of a robot hand, but also its grasping capabilities are extremely useful.

---

[5] Schunk https://www.schunk.com/

[6] Armar III http://www.kcist.kit.edu/

[7] Barrett Technologies http://www.barrett.com/

[8] Festo https://www.festo.com/

[9] Robotiq http://robotiq.com/

In industrial environments, task-oriented grippers are commonly used. This kind of grippers are usually conceived to perform a few predefined grasps, always under the same circumstances, in a full-restraint environment. This manipulators are suitable to work perfectly in specific tasks, but can be inadequate for other purposes. For example, the gripping force required in a car factory will be overwhelming for grasping eggs in a farm. Also, some industries require grippers to work in production lines with unknown objects or poses. On these cases, external sensory data, vision systems or flexible materials are used to solve the grasping problem (Lippiello et al., 2013).

The manipulation requirements for an industrial gripper are often simple. However, more complex task (e.g. manipulate fragile objects) are also required in production lines. Adding end-effector sensors in a gripper is a common solution for solving most of the more complex grasping tasks. This more complex grippers, as the *Barrett* or *Schunk SDH* hands, are often used also for academia and research purposes.

In the field of service robotics, there is a lot of effort in the development of anthropomorphic hands for humanoid robots. This kind of hands try to imitate the shape and performance of the human hand. There is a great interest in developing more *human* robots. Therefore, having an anthropomorphic hand is a requirement for this humanoid robots. However, in terms of grasping and manipulation, there is no proof more fingers means better functionality. The control and mechanisms in these hands is more complex than classic industrial grippers. For this reason, their use is usually limited to research purposes and their employment in other fields is scarce.

Having such amount of manipulator designs is good in terms of having a wide range of possibilities to select the proper tool to satisfy an specific task. However, there is no way to establish which designs are good or better than others. Or even if a hand design is really useful or just a cosmetic gadget, which is the case for most humanoid robots.

Selecting or designing a robotic hand, lay out some questions: How many DoF are needed? and fingers? Which type of control? Does it fit in my aim task? How *good* it is? Is it better than others? How can be compared?...Usually, the comparison between robotic manipulators is done in terms of weight, force, DoF or control, but there is no discernible relation between this *mechanical* properties and the real grasp performance.

Currently there is a lack of a methodology to evaluate the performance of robotic manipulators. Thus, it is necessary to define some desirable properties or capabilities a robotic hand should achieve, in order to be considered an useful tool. Our proposed solution to this problem is the use of quality metrics. The features evaluated in a grasp can be considered as measurable properties in a robotic hand and, consequently, be useful to evaluate their performance.

Anthropomorphic hands

The design and construction of artificial hands that resemble not only in the appearance but also the functionality of the human hand has been a driven force in robotics and biomechanics for a long time. Through the years, many designs have been proposed with different levels of success, but nowadays stills remains a challenge.

Having a functional prosthesis usually means to install an invasive, complex and unpleasant mechanism, which usefulness is limited. This make amputees tend to use more cosmetic prosthesis, usually latex gloves, with little or non-existent manipulation ability (**Jang_2011**). Developing attractive and functional prosthesis is an urge and, although there are a lot of different designs, these are usually highly expensive or their performance is far away from the real human hand.

In recent years the availability of cheap prototyping technologies like *3D Printing* and electronic platforms like *Arduino* has caused an explosion of artificial hand models that can be obtained, shared, built and operated at an affordable cost. This is definitively positive since it allows the participation of more agents and consequently increase the possibility of novel ideas coming up.

However, with so many design alternatives available, there is an urge for having objective procedures to evaluate these prototypes. Hand prosthesis are designed to achieve a series of predefined grasps following a taxonomy, Figure 1.5. This method provide desirable grasp poses an amputee would need to perform in his *activities of daily life (ADL)*. However, the process of manufacturing a hand prototype and evaluate these grasp poses is highly expensive, in time and materials. Thus, developing a methodology to evaluate more efficiently the grasp performance could be beneficial in the biomechanical field.

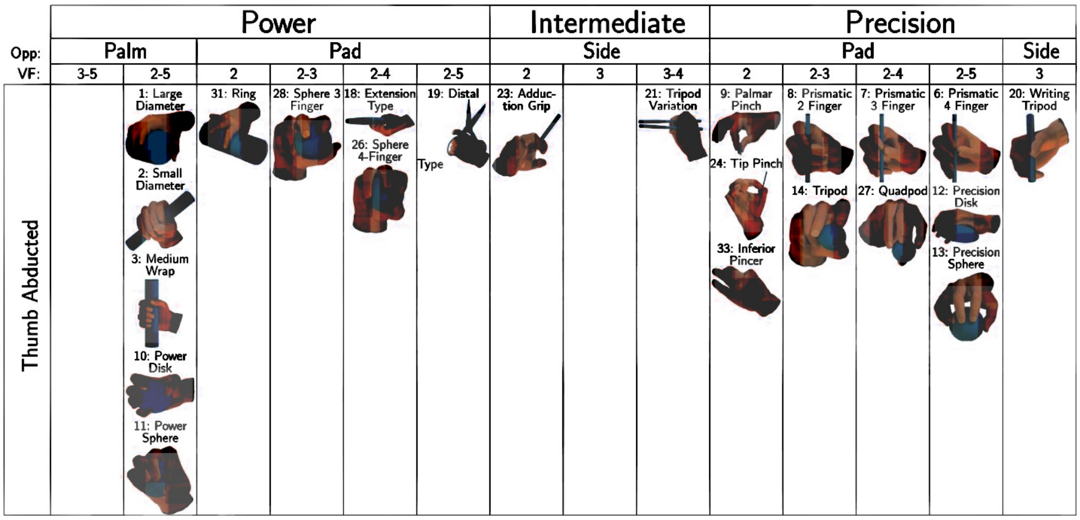

Power
Intermediate
Precision
Opp:
Palm
Pad
Side
Pad
Side
VF:
3-5
2-5
2
2-3
2-4
2-5
2
3
3-4
2
2-3
2-4
2-5
3
Thumb Abducted
1: Large Diameter
2: Small Diameter
3: Medium Wrap
10: Power Disk
11: Power Sphere
31: Ring
28: Sphere 3 Finger
18: Extension Type
26: Sphere 4-Finger
19: Distal Type
23: Adduction Grip
21: Tripod Variation
9: Palmar Pinch
24: Tip Pinch
33: Inferior Pincer
8: Prismatic 2 Finger
14: Tripod
7: Prismatic 3 Finger
27: Quadpod
6: Prismatic 4 Finger
12: Precision Disk
13: Precision Sphere
20: Writing Tripod

FIGURE 1.5: Extract of the taxonomy of human grasps, full table can be found in Feix et al., 2016. This tables shows most of the common grasp poses employed by humans in his ADL.

In our research we are specially interested in the *functional* evaluation of artificial hands. That is, measuring how much of the human hand functions, and more specifically the grasping functions, can be performed by an artificial hand. This can be done either by benchmarking real prototypes on predefined tests, or by measuring the virtual models using analytic tools. Analytic evaluation of virtual or simulated models allow a more exhaustive, fast and cheap study of a hand, though less reliable and realistic. In any case, both approaches have their pros an cons, but are complementary to obtain a full evaluation. The implementation of a methodology based in quality metrics to evaluate the prosthesis grasp capabilities and the use of simulated environments instead of real world grasps could improve the designing process of current hand prosthesis.

Simulation vs. real world

Evaluating grasp hypotheses and hand designs has some restrictions attached to the real world: the time needed to perform a grasp, the inaccuracies in the detection and tracking of objects, the cost of manufacture different hand prototypes, etc. Virtual environments can solve most of this problems, but introduce some other limitations.

In a simulated environment, there is no inaccuracy in the object or grasp poses, the cost over hand manufacture and the time needed to perform a grasp is trivial compared to real world. In contrast, including in such virtual models simulation of physics, forces or surface properties, such as friction coefficients, can be a problem and computational expensive. Therefore, virtual environments commonly provide unrealistic results.

Simulation environments are commonly used in the robotic field. Although they lack of some of the properties of the real world, and the expected behavior in the real world does not always match the simulated one, their advantages in terms of time and cost offset their scarcities. Therefore, using a simulated environment, although not faithful with respect to reality, can help to accelerate the study and evaluation of thousand of grasp hypotheses and hand designs.

1.2 Objectives

In this thesis we want to explore the applications of grasp quality metrics and how their use could lead to: first, improve the accuracy of robots to select good grasp hypotheses, second, provide a methodology to evaluate artificial hands, and third, assist in the designing process of an anthropomorphic hand. To do so, we will complete a series of goals:

1. The first goal is characterize the most commonly used quality metrics. Grasps will be generated and evaluated in a simulation environment. Different robotic hands and sets of objects will be used for this experiments. The results of these experiments will provide practical information about: the range of values of the metrics, how are distributed, the robustness against pose inaccuracies and finally, the correlation between metrics.

2. The second goal is provide a ground truth for the grasp evaluation. In this experiments different grasp evaluated in simulation will be performed in real environments and the predictability of the metrics will be studied. With the results of the real grasps, different machine learning methods will be applied in order to find a combination of metrics capable to predict the grasp success.

3. The third goal is to develop a methodology to evaluate artificial hands. Using the quality metrics previously studied, different grasps and hands will be evaluated in simulation. The result of these experiments will provide a protocol to evaluate the functionality of artificial hands using only a simulation environment.

4. Finally, we will demonstrate how to use this methodology to improve the design of a prosthetic hand.

1.3 Outline

This thesis is structured as follows:

- **Chapter 2 - Basic Foundations** Chapter 2 presents the basic background in grasping, involving the use of quality metrics, grasp prediction and hand evaluation methods. This will show the current methodologies, their problems and how we'll try to solve them.
- **Chapter 3 - Methodology** To accomplish our objectives, the work in this thesis will be done using both, simulation an real world experiments. This chapter will introduce the OpenHand simulator and the real robotic systems used to perform grasps and evaluate hands. It will describe also the software, materials, hands and objects models used in our experiments.
- **Chapter 4 - Quality metrics analysis** The experiments done in this chapter will fit with the first objective of the thesis, the results get there will provide a set of quality metrics evaluating different aspects of the grasp and its characteristics.
- **Chapter 5 - Grasp prediction** In this chapter different classification algorithms will be trained to develop a grasp predictor. To test this predictor, real grasp experiments will be performed using different objects and manipulators.
- **Chapter 6 - Artificial hand evaluation** Chapter 6 will present a methodology to evaluate and compare hand designs. This methodology will be applied in the design process of a low-cost prosthetic hand. This will point out how the use of combined quality metrics improve the design of artificial hands.
- **Chapter 7 - Conclusions** Chapter 7 will highlight the contributions of this thesis, and discuss the future lines of work to extend the results obtained here.

Chapter 2

Basic foundations

Disclaimer This Chapter presents a summary on the basic theoretical knowledge about grasping, contact points and quality metrics for evaluate grasps. Most of the definitions and formulas presented here have been widely discussed and studied in previous works (Prattichizzo and Trinkle, 2008; Roa, 2009; Leon et al., 2013a).

2.1 Grasp theory

Definition 2.1. *A **grasp** is defined as a set of contact points over the object surface, which purpose is to constrain the potential movements of the object in the event of external disturbances (Murray et al., 1994; Bicchi and Kumar, 2000; Prattichizzo and Trinkle, 2008; Roa, 2009).*

Assume there is a contact between two rigid bodies (Figure 2.1), let be $\{W\}$ the inertial frame fixed in the world and $\{O\}$ the the frame fixed to the object, whose origin is defined relative to $\{W\}$ by a vector $p \in \mathbb{R}^3$ and correspond to the object's center of mass. The position of the contact point i is defined by the vector $c_i \in \mathbb{R}^3$ relative to $\{W\}$. At the contact point i, a frame $\{C\}$ is defined with axes $\{\hat{n}_i, \hat{t}_i, \hat{o}_i\}$ with $\hat{n}_i$ containing c_i normal to the contact tangent plane and directed toward the object. The other two vectors are orthogonal and lie in the tangent plane of the contact. A list of notations is included in Table 2.1.

Definition 2.2. *A **contact point** can be defined as a joint between the finger and the object. The shape of the contacting surfaces and the stiffness and frictional characteristics of the contact bodies define the nature of this joint (Mason, 2001).*

Definition 2.3. *A **twist** is the representation of the spacial velocity of a rigid body and can be written as $t \in \mathbb{R}^6$:*

$$t = \begin{pmatrix} \upsilon \\ \omega \end{pmatrix} \tag{2.1}$$

where $\upsilon \in \mathbb{R}^3$ is the linear velocity of point p and $\omega \in \mathbb{R}^3$ is the angular velocity of the object in the world frame $\{W\}$. When a contact is produced, two velocities or *twists* are generated, $t_{i,hand}, t_{i,obj}$ referred to the contact i relative to the $hand$ and the $object$, respectively.

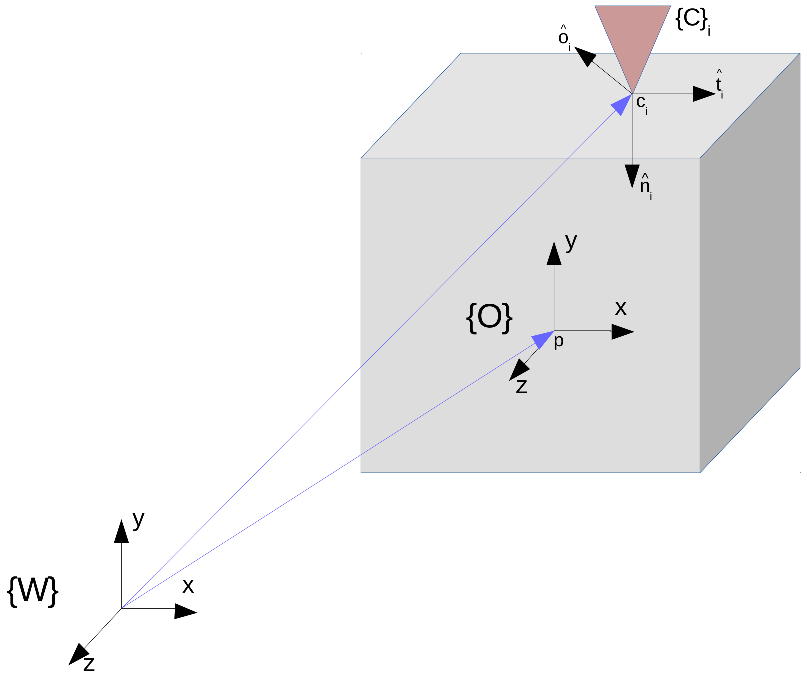


FIGURE 2.1: Notation for a contact between two rigid bodies.

The force applied by a finger at the contact point generates a *wrench* on the object with force and torque components, represented by vector $w_i \in \mathbb{R}^6$:

$$w_i = \begin{pmatrix} f_i \\ \tau_i \end{pmatrix} \tag{2.2}$$

where $f_i \in \mathbb{R}^3$ is the force applied to the object at the point c_i and $\tau_i \in \mathbb{R}^3$ the resulting moment at point p. As forces and torques are dimensionally different, a parameter ρ is introduced:

$$w_i = \begin{pmatrix} f_i \\ \tau_i/\rho \end{pmatrix} \tag{2.3}$$

selecting ρ to have *length* units, allows all the components of w to have units of *force*. If there are multiple contacts acting on an object, the total set of wrenches w_o that can be transmitted to the object through the n_c contacts is the linear combination of all individual wrenches:

$$w_o = \sum_{i=1}^{n_c} w_i \tag{2.4}$$

To prevent a grasp from slipping, the forces in the contacts (and their corresponding wrenches w_i) and any external disturbing force or wrench w_{ext} have to be in equilibrium:

$$w_o + w_{ext} = 0 \tag{2.5}$$

This equation is valid always that the contact forces satisfy the contact constrains described in the next section.

TABLE 2.1: Notations

$\{W\}$	Word coordinate frame
$\{O\}$	Object coordinate frame
n_c	Number of contact points
c_i	Contact point i relative to $\{W\}$
$\{C\}_i$	Contact point i coordinate frame with axis $\{\hat{n}_i, \hat{t}_i, \hat{o}_i\}$
$\hat{n}_i$	Unit normal to the contact tangent plane directed toward the object
p	Position of the object relative to $\{W\}$
v	Linear velocity of point p
ω	Angular velocity of the object relative to $\{W\}$
w_i	Generalized force acting on the object for a unit force along $\hat{n}_i$
f_i	Force applied to the object at the point c_i
τ_i	Resulting moment at point p
w_o	Total set of wrenches transmitted to the object through the n_c
w_{ext}	Disturbing external wrenches
μ	Friction coefficient of the contacting materials
β	Half-angle of the friction cone
m	Number of faces of discretized friction cone
B	Selection matrix
l	Total number of twist components transmitted
G	Grasp matrix
$\tilde{G}_i$	Partial grasp matrix
$\tilde{G}$	Complete grasp matrix
J	Hand Jacobian matrix
$\tilde{J}_i$	Partial Hand Jacobian matrix
$\tilde{J}$	Complete Hand Jacobian matrix
G_J	Grasp Jacobian matrix

Contact model

The set of contact points contained in a grasp is defined by the contact model. This models characterizes both the forces that can be transmitted through the contact as well as the allowed motions of the contact bodies. Two different contact models are used in robotics: *rigid-body models* and *compliant models*. Rigid-body models held two different types of point contact: *point contact without friction* and *point contact with friction*. Compliant models define *soft-finger contact*, a type of point contact with friction.

- **Contact without friction:** These contact points can only transmit forces through the normal to the object surface at the contact point. Contact force arise from the constraint of incompressibility and the impenetrability between rigid bodies. This model is used when the contact patch is very small and the surfaces of the hand and object are slippery (Prattichizzo and Trinkle, 2008). It does not represent the real contact situations that appear in robotic manufacturing operations (Cutkosky, 1989; Lin et al., 2000) and, when used, the machine accuracy is negatively affected. Moreover, they are not capable of predicting the individual contact forces of a multiple-contact fixture (Bicchi, 1994; Harada et al., 2000).
- **Contact with friction:** A contact point with friction can transmit forces in the normal and tangential directions to the surface at the contact point but non of

the moment components are transmitted. This is used when there is significant contact friction, but the contact patch is small so that no appreciable friction moment exists (Prattichizzo and Trinkle, 2008).

- **Soft-finger contact:** The soft contact model is used when the surface friction and the contact patch are large enough to generate significant friction forces and a friction moment about the contact normal. It is used to model the contact between a soft finger and a rigid object allowing the finger to apply an additional torsional moment with respect to the normal at the contact point (Ciocarlie et al., 2005; Ciocarlie et al., 2007; Howe et al., 1988; Kao and Cutkosky, 1992; Howe and Cutkosky, 1996; Kao and Yang, 2004).

The *Coloumb friction* is a classical model that defines the friction coefficient. It is based on the idea that friction opposes motion and that its magnitude is independent of the velocity and contact area. It is an empirical model that asserts that the allowed tangential force is proportional to the applied normal force by $f_t \leq \mu f_n$, where μ is called the friction coefficient of the contacting materials. The friction forces can be represented geometrically as a friction cone where the set of all forces that can be applied is constrained to be inside a cone centred about the surface normal with half-angle $\beta = tan^{-1}(\mu)$. In the spacial case, the friction cone is a circular cone, defined by

$$\sqrt{f_{t_i}^2 + f_{o_i}^2} \leq \mu f_{n_i}, f_{n_i} \geq 0. \tag{2.6}$$

A particular contact model is defined through the *selection matrix* $B_i \in \mathbb{R}^{l_i \times 6}$, which works like a filter selecting l_i components of the relative contact twist and sets them to zero.

$$B_i(t_{i,hand} - t_{i,obj}) = 0 \tag{2.7}$$

2.2 Grasp analysis

Definition 2.4. ***Grasp analysis*** *consists on finding whether the grasp is stable using common closure properties, given an object and a set of contacts. Then, quality measures can be evaluated in order to enable the robot to select the best grasp to execute (***Leon2013d***).*

The first test for evaluating a grasp consists of determining its ability to constrain the motions of the manipulated object and to apply arbitrary contact forces on the object without violating friction constraints at the contacts (Bicchi, 1995). Two commonly used properties have been proposed to ensure this condition: force and form closure.

Definition 2.5. *A grasp is in force-closure if the fingers can apply, through the set of contacts, arbitrary wrenches on the object, which means that any motion of the object is resisted by the contact forces (Nguyen, 1988).*

Definition 2.6. *A grasp is in form-closure if the location of the contact points on the object ensures its immobility (Bicchi, 1995).*

Form closure is a stronger condition than force closure and it is mostly used when executing power grasps (Prattichizzo and Trinkle, 2008). Force closure is possible with fewer contacts, making it suitable for executing precision grasps, but it requires the ability to control internal forces. The analysis of form closure is intrinsically geometric. A necessary and sufficient condition for form-closure is that the contact wrenches of the grasp positively span the whole wrench space (Salisbury and Roth, 1983).

Definition 2.7. *A grasp wrench space (GWS) is the space of wrenches that can be applied to the object at each contact point.*

The boundary of the wrench space can be calculated as a convex hull. Form-closure then can be equivalently determined verifying if the origin of the wrench space lies inside this convex hull (Mishra et al., 1987). Based on the above necessary and sufficient conditions, many tests that have been proposed in the literature (Nguyen, 1988; Hirai and Asada, 1993; Xiong, 1994). The test of Ferrari and Canny, 1992 is the most widely used. They proposed to calculate the radius of the largest ball inscribed in the convex hull centred in the origin and verify that it is larger than zero. Zhu and Wang, 2003 developed a numerical test which measures the scaling factor for the maximum compact set inscribed in the GWS with centre in the origin.

Assessing the force-closure property of a robotic grasp is much more difficult because of the nonlinear nature of the Coulomb friction cone (Zhu and Wang, 2003). Nakamura et al., 1989 formulated the force-closure test as 12 nonlinear programming problems. Trinkle, 1992 formalized the force closure condition as a linear programming problem. Bicchi, 1995 observed that the force-closure problem is equivalent to the stability of an ordinary differential equation. Liu, 1999 reformulated the force closure condition as a ray-shooting problem by linearizing the friction cones and proposed a clean-cut test for force closure grasps. Han et al., 2000 proposed a force-closure test representing the nonlinear friction cone constraints as linear matrix inequalities, for which efficient algorithms are now available. With the linearization of the friction cone, most of the existing form-closure tests can be generalized to force-closure analysis. Zhu et al., 2004 proposed a numerical criterion for 3-D grasps with frictional point contacts or soft contacts, formulated as a convex constrained optimization problem without linearization of the friction cone. More recently, Zheng and Chew, 2009 proposed an algorithm for computing the distance between a point and a convex cone in n-dimensional space that can be applied to force-closure test and improve their efficiency.

Grasp matrix G and hand jacobian J

There are two matrices of the utmost importance in grasp analysis: the grasp matrix G and the hand Jacobian J. They are introduced here, but a complete explanation can be found in Prattichizzo and Trinkle, 2008. The grasp matrix G and hand Jacobian J define the relevant velocity kinematics and force transmission properties of the contacts. The hand Jacobian maps the joint velocities to the twists of the hand expressed in the contact frames. The inverse of the grasp matrix refers the object twist to the contact frames.

For each contact frame i, the partial grasp matrix $\tilde{G}_i$ can be calculated as:

$$\tilde{G}_i = \begin{pmatrix} R_i & 0 \\ S(c_i - p)R_i & R_i \end{pmatrix} \tag{2.8}$$

where $R_i \in \mathbb{R}^{3\times3}$ represents the rotation matrix of the $\{C\}_i$ contact frame with respect to $\{W\}$, c_i the position of the contact point, p the position of the object and $S(c_i - p)$ is the cross-product matrix.

The complete transposed grasp matrix $\tilde{G}^T$ is the combination of the transposed grasp matrices for each of the n_i contact points:

$$\tilde{G}^T = \begin{pmatrix} \tilde{G}_1^T \\ \vdots \\ \tilde{G}_{n_i}^T \end{pmatrix} \tag{2.9}$$

The transpose of the complete *grasp matrix* $\tilde{G}^T \in \mathbb{R}^{6\times6}$ maps the object twist from W to the contact frame C:

$$t_{i,obj} = \tilde{G}_t^T \tag{2.10}$$

where t denote the object twist related to $\{W\}$ and and $t_{i,obj} \in \mathbb{R}^{6n_i}$ is a vector containing all the twist of the object in the contact frames:

$$t_{i,obj} = (t_{1,obj}^T \dots t_{n_i,obj}^T)^T \tag{2.11}$$

The complete hand Jacobian $\tilde{J}$ is the combination of the hand Jacobians for each of the n_i contact points:

$$\tilde{J} = \begin{pmatrix} \tilde{J}_1 \\ \vdots \\ \tilde{J}_{n_i} \end{pmatrix} \tag{2.12}$$

Each partial hand Jacobian, $\tilde{J}_i$ can be calculated as:

$$\tilde{J}_i = R_i \begin{pmatrix} d_{i,1} & \dots & d_{i,n_q} \\ l_{i,1} & \dots & l_{i,n_q} \end{pmatrix} \tag{2.13}$$

where:

$$\begin{aligned} d_{i,j} &= \begin{cases} 0_{3\times1} & \text{if contact i does not affect the joint j} \\ \hat{z}_j & \text{if joint j is prismatic} \\ S(c_i - \zeta_j)^T \hat{z}_j & \text{if joint j is revolute} \end{cases} \\ l_{i,j} &= \begin{cases} 0_{3\times1} & \text{if contact i does not affect the joint j} \\ 0_{3\times1} & \text{if joint j is prismatic} \\ \hat{z}_j & \text{if joint j is revolute} \end{cases} \end{aligned} \tag{2.14}$$

being ζ_j the origin of the coordinate frame associated with the j-th joint and $\hat{z}_j$ is the unit vector in the direction of z-axis in the same frame, expressed in $\{W\}$.

The complete hand Jacobian $\tilde{J}$ relates the joint velocities to the contact twists on the hand:

$$t_{i,hand} = \tilde{J}\dot{q} \tag{2.15}$$

where $\dot{q} = [q_1 \dots q_{n_q}]^T$ represents the vector of joint displacements and n_q the number of hand joints and $t_{i,hand} \in \mathbb{R}^{6n_i}$ is a vector containing all the twist of the hand in the contact frames:

$$t_{i,hand} = (t^T_{1,hand} \dots t^T_{n_i,hand})^T \tag{2.16}$$

Replacing equations 2.10 and 2.15 in the contact model equation 2.7, the contact constraint equations are defined as:

$$B(\tilde{J}\dot{q} - \tilde{G}^T t) = 0 \tag{2.17}$$

Therefore, the *grasp matrix* G and *hand Jacobian* J can be defined as:

$$G^T = B\tilde{G}^T \tag{2.18}$$

$$J = B\tilde{J}, \tag{2.19}$$

Grasp jacobian

The grasp Jacobian G_J is the transformation from the joint velocities to the velocity of the object being grasped (Shimoga, 1996):

$$t = G_J \dot{q}$$

It takes into account the transformations for each finger from joint velocities to fingertip Cartesian velocity (J), the contact relationships and the transformations from the contact frames of reference to the object frame of reference (G). Thus, it is a function of the hand posture and the lengths of the finger segments.

$$G_J = (G^+)^T J \tag{2.20}$$

with G^+ being the generalized inverse of G. Figure 2.2 summarizes the relationships between velocities in a multi-fingered grasp.

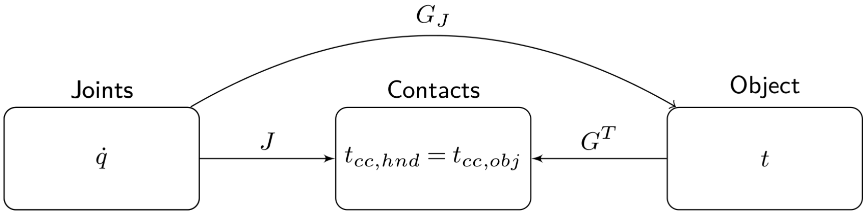


FIGURE 2.2: Relationship between grasp matrix, hand Jacobian and grasp Jacobian (**Leon2013d**).

2.3 Grasp planning

Definition 2.8. ***Grasp planning*** *consists in determining the finger contact locations on an object surface and the appropriate gripper configuration which allow the constraints, forces and torques produced on the object through the contact points to reach one or more desirable properties like dexterity, force-closure, stability or equilibrium and other which has been defined in the literature (Sahbani et al., 2012; Shimoga, 1996; Bicchi, 2000).*

Analytical approaches try to solve the grasp planning problem by using kinematics and dynamics formulations (Sahbani et al., 2012). There exists a vast literature focusing on the development of algorithms to find grasps under a variety of assumptions regarding dimensionality, contact models, object shapes and others (Sahbani et al., 2012; Shimoga, 1996; Bicchi, 2000). Often the proposed solutions produce a large number of possible grasps. Quality metrics are the tools that allow these sets of solutions to be ranked. Actually, metrics play a triple role: allowing relative comparisons between grasps, providing an absolute quantitative assessment of the goodness of a grasp, and, finally, serving as evaluation functions in grasp synthesis optimization algorithms (Mishra, 1995).

In recent years, though more attention has been paid to alternative data-driven and experimental approaches for grasp planning, many of them still rely on analytical algorithms as an important part of their pipelines. Bohg et al., 2014, point out that most methods assuming a previously known model of the objects perform an offline initial analysis of the possible grasps for each object using analytical techniques, being quality metrics a critical tool to rank grasps.

2.4 Grasp quality metrics

Definition 2.9. *A **grasp quality metric** is an algorithm based in different mathematical principles and heuristics that quantifies the goodness of a grasp. They measure the ability of a grasp to restrain an object and ensure his stability against external disturbances.*

Many grasp quality metrics have been proposed in the robot grasping research literature. Roa and Suárez, 2014 identify up to 24 different grasp quality metrics designed to quantify the goodness of a grasp. Quality metrics play a principal role in the so called analytical approach to the grasp planning problem, also often referred as *grasp synthesis*.

A limitation of the use of metrics is that each of them is designed to assess a specific aspect of a grasp. The work of Roa and Suárez, 2014 distinguish between metrics associated with the contact points and those associated with the hand configuration. The former ones are the most populated, being subdivided into metrics based on the properties of the *Grasp matrix*, metrics based on geometric relations, and those that consider limitations of the applied forces.

They also reviewed several attempts to define global metrics by combining existing metrics. Combined metrics can be a solution to overcome the specificity of each one. An early correlation study of several quality metrics showed the existence of at least five independence dimensions on the evaluation of human grasps (León et al., 2012). Roa and Suárez, 2014, identify two trends for building combined indexes, serial and parallel. On the serial approach one metric is used to generate and select a subset of grasp configurations, then a second metric is used to rank them (Hester et al., 1999). A most common approach is the parallel combination of metrics. In this approach every grasp is evaluated by a set of metrics, and the values obtained are added to produce a unique evaluation index (Boivin et al., 2004; Aleotti and Caselli, 2006). Different normalization procedures can be applied to every metric for their combination, as well as different weighting coefficients (Chinellato et al., 2005).

But combined indexes are not free of limitations. First, each combinatorial method is still based on a subset of quality metrics, and there are no clear criteria on how to choose them. Some metrics might be measuring the same property while others capture different underlying properties. And second, there is no obvious way of merging metrics since they do not provide values on the same range not even on the same physical magnitudes.

For our research, ten different quality metrics have been selected, these can be classified into four categories depending on which principles they obey: the algebraic properties of G, the distribution of the contact points, the magnitude of forces and the configuration of the manipulator. A summary of these metrics is presented in the following sections, for a more detailed and complete review the reader is referred to Leon et al., 2013b and Roa and Suárez, 2014. Table 2.2 summarizes the definition of these metrics.

TABLE 2.2: Summary of the selected quality metrics

Name		Formula	Min	Max
Group A: Algebraic properties of G				
Q_{A1}	Smallest singular value of G (Li and Sastry, 1987)	$\sigma_{min}(G)$	0	-
Q_{A2}	Volume of G in the wrench space (Li and Sastry, 1987)	$\prod_{i=1}^{r}\sigma_i$	0	-
Q_{A3}	Grasp Isotropy Index (Kim et al., 2001)	$\frac{\sigma_{min}(G)}{\sigma_{max}(G)}$	0	1
Group B: Distribution of contact points				
Q_{B1}	Distance between the centroid of the contact polygon and the center of mass of the object (Ding et al., 2001; Ponce et al., 1997)	$1-\frac{distance(p,p_c)}{distance_{max}}$	0	1
Q_{B2}	Area of the grasp polygon (Mirtich and Canny, 1994)	$\frac{Area(Polygon(p_1,...p_n))}{Area_{max}}$	0	1
Q_{B3}	Shape of the grasp polygon (Kim et al., 2001)	$1-\frac{1}{\theta_{max}}\sum_{i=1}^{n_f}\lvert\theta_i-\bar{\theta}\rvert$	0	1
Group C: Magnitude of Forces				
Q_{C1}	Smallest maximum wrench to be resisted (Ferrari and Canny, 1992; Kirkpatrick et al., 1990)	$\frac{\min_{w\in CW}\Vert w\Vert}{\sqrt{2}}$	0	1
Q_{C2}	Volume of the convex hull (Miller and Allen, 1999)	$\frac{Volume(CW)}{Volume_{max}}$	0	1
Group D: Configuration of the manipulator				
Q_{D1}	Posture of manipulator joints (Liegeois, 1977)	$1-\frac{1}{n_q}\sum_{i=1}^{n_q}(\frac{y_i-a_i}{a_i-y_{iM}})^2$	0	1
Q_{D2}	Inverse of the condition number of $\mathbf{G_J}$ (Salisbury and Craig, 1982; Kim and Khosla, 1991)	$\frac{\sigma_{min}(G_J)}{\sigma_{max}(G_J)}$	0	1

2.4.1 Group A: Algebraic properties of G

These are stability indicators that consider the algebraic properties of the grasp matrix G to measure the grasp capability of withstanding external wrenches; they use the contact points and normal directions. They do not consider any limitation on the finger forces, so that in some cases the fingers have to apply very large forces to resist small perturbations.

The grasp matrix G is the linear transformation between the contact force vector and the net wrench on the object. Geometrically, G maps the unitary sphere in the force domain of the contact points into an ellipsoid in the wrench space (Roa, 2009). The lengths of the principal axes of the ellipsoid are the singular values of G, which are the n nonnegative square roots σ_i of the eigenvalues of GG^T. The following quality measures use these singular values to define the quality of the grasp.

$\mathbf{Q_{A1}}$ - Smallest singular value of G

It measures how far the grasp configuration is from falling into a singular configuration (Li and Sastry, 1987). A grasp matrix $G \in \mathbb{R}^{6xr}$ has 6 singular values given by the positive square roots of the eigenvalues of GG^T, when a grasp is in a singular configuration, at least one of the singular values of G is zero. It is calculated as:

$$Q_{A1} = \sigma_{min}(G) \tag{2.21}$$

where $\sigma_{min}(G)$ is the smallest singular value of the matrix G. The lower limit is zero and the upper limit is not determined.

$\mathbf{Q_{A2}}$ - Volume of G in the wrench space

Considering the grasp matrix G as an ellipsoid in the wrench space, it measures the global contribution of all the contact forces as the volume of this ellipsoid (Li and Sastry, 1987). is calculated as:

$$Q_{A2} = v(G) = \prod_{i=1}^{r} \sigma_i \tag{2.22}$$

where r is the rank of G, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ denote the nonzero singular values of G. The lower limit is zero and the upper limit is not determined.

$\mathbf{Q_{A3}}$ - Grasp Isotropy Index

A grasp is considered isotropic when the magnitudes of the internal forces are similar (Kim et al., 2001). This index measures how uniform is the contribution of the contact forces to the total wrench over the object. It is calculated as:

$$Q_{A3} = \frac{\sigma_{min}(G)}{\sigma_{max}(G)} \tag{2.23}$$

It approaches to one at a desirable configuration (isotropic) and is equal to zero at the singular configuration. Therefore, it is already restraint into the range 0 to 1.

2.4.2 Group B: Distribution of contact points

These are indicators that use the distribution of the contact points. Better stability is assumed when contact points are distributed in a uniform way on the object surface and around the object centre of mass aiming to minimize the effect of gravitational and inertia forces.

$\mathbf{Q_{B1}}$ - Distance between the centroid of the contact polygon and the center of mass of the object

It estimates the effect of gravitational and inertia forces during the motion of the robot, which tends to zero when the distance between the centre of mass p of the grasped object and the centroid of the contact polygon p_c (Ding et al., 2001) is minimized. The centroid of the contact polygon is calculated as:

$$p_c = \frac{1}{n_c} \sum_{i=1}^{n_c} c_i \tag{2.24}$$

where n_c is the number of contact points and c_i is the location of each contact point.

We propose its lower limit to be zero, calculated as the maximum distance from the centre of mass of the object to any point in the object's contour, $distance_{max}$ can be obtained as the maximum distance from the center to any of the corners of the object bounding box. The center of mass of the object is assumed to be at the center of the bounding box. Thus, this metric is calculated as:

$$Q_{B1} = 1 - \frac{distance(p, p_c)}{distance_{max}} \tag{2.25}$$

$\mathbf{Q_{B2}}$ - Area of the grasp polygon

This measure is defined as the area of the polygon formed by the contact points. With the same finger force, a grasp with a larger area can resist larger external torques (Xiong et al., 1999). In 3-finger grasps a triangle is considered, for four or more contact points, the indicator is extended defining a contact plane as proposed by Supuk et al., 2005. The contact plane is generated by selecting three fingers and the remaining contacts are perpendicularly projected onto that plane.

We propose to take into account that its lower limit is zero and the upper limit ($Area_{max}$) can be calculated as the area of the polygon when the manipulator is open in a plane with the joints at their maximum aperture limits. Then, the measure can be calculated as:

$$Q_{B2} = \frac{Area(Polygon(p_1, p_2, ..., p_n))}{Area_{max}}. \tag{2.26}$$

where p_i are the projected on a plane contact points of the manipulator.

Limitations This metric is limited to grasps containing at least three contact points as it is not possible to have a polygon with only two contact points.

$\mathbf{Q_{B3}}$ - Shape of the grasp polygon

This measure compares how far the internal angles of the grasp polygon are from those of the corresponding regular polygon (Kim et al., 2001). For a planar grasp, is desirable that the contact points are distributed in a uniform way on the object surface to improve the grasp stability (Park and Starr, 1992). This index is calculated as:

$$Q_{B3} = \frac{1}{\theta_{max}} \sum_{i=1}^{n_c} |\theta_i - \bar{\theta}| \tag{2.27}$$

where n_c denotes the number of contacts, θ_i is the inner angle at the i^{th} vertex of the polygon, $\bar{\theta}$ denotes the average angle of all inner angles of the grasp polygon and θ_{max} is the sum of the differences between the internal angles when the polygon has the most ill-conditioned shape (degenerates into a line) and those of the regular polygon:

$$\theta_{max} = \sum_{i=1}^{n_c} |\theta_i - \bar{\theta}|_{\textit{ill conditioned}} = (n_f - 2)(180 - \bar{\theta}) + 2\bar{\theta} \tag{2.28}$$

For four or more contact points, the planar grasp polygon is obtained in the same way as for Q_{B2} and the contact points are set in groups of three to estimate the average inner angles. The lower limit is zero and his upper limit is 1, which correspond to an equilateral triangle.

Limitations As with metric Q_{B2}, this metric is limited to grasps with three or more contact points due to requirements of the grasp polygon.

2.4.3 Group C: Magnitude of forces

These are stability indicators that take into account the magnitude of forces applied at the contact points.

$\mathbf{Q_{C1}}$ - Smallest maximum wrench to be resisted

This index is defined as the largest perturbation wrench that the grasp can resist with independence of its direction (Ferrari and Canny, 1992). Only the directions of forces are used and their magnitudes are upper-bounded to 1. Defining GWS as the set of all possible wrenches w acting on the object, the maximum of $w \in GWS$ lies on the boundary approximated as the convex hull over the discretized frictions cones (CW). Then the quality metric is the radius of the largest sphere centred at the origin, which is contained in GWS:

$$Q_{C1} = \min_{w \in CW} \|w\| \tag{2.29}$$

The index depends on the choice of the origin of the reference system used to compute torques and the ρ parameter used to scale the torque to the force magnitude. In this work, we use the centre of mass of the object and limited the magnitude of the torques to 1 choosing ρ as $distance_{max}$ defined previously for the measure Q_{B1}. Then, the upper limit of the index is $\sqrt{2}$ and the lower limit is zero. Then, we can normalize this index between 0 and 1 as:

$$Q_{C1_N} = \frac{Q_{C1}}{\sqrt{2}} \tag{2.30}$$

Limitations The Convex Hull is calculated using the PyHull library [1]. In some specific cases, the distribution of the contact points cause the library to not be capable to estimate the CW. Also, as like metrics Q_{B2} and Q_{B3}, this metric is limited to grasps with at least three contact points.

---

[1] https://github.com/materialsvirtuallab/pyhull

$\mathbf{Q_{C2}}$ - Volume of the convex hull

This measure is defined to avoid the dependence of the previous index on the selection of the origin of the reference system. The measure calculates the volume of the boundary of the set of all possible wrenches acting on the object (Miller and Allen, 1999).

$$Q_{C2} = Volume(CW) \tag{2.31}$$

The reference system and ρ have been chosen as described in the previous measure. Lower limit is zero and upper limit is not determined so that it is initially not possible to normalize the index in the range 0 to 1. However, we used the Monte Carlo method (see **Leon2013d**) to estimate the upper limit in each manipulator. The normalized measure then can be calculated as:

$$Q_{C2_N} = \frac{Q_{C2}}{Volume_{max}} \tag{2.32}$$

Limitations The same limitations applied to Q_{C1} apply to this metric.

2.4.4 Group D: Configuration of the manipulator

The measures presented here are intended as manipulability indices, describing the ability to reach a certain position or to change the position or orientation at a given configuration.

$\mathbf{Q_{D1}}$ - Posture of manipulator joints

This metrics looks for configurations of the manipulator whose joints are far from its physical limits (Liegeois, 1977):

$$Q_{D1} = 1 - \frac{1}{n_q}\sum_{i=1}^{n_q}\left(\frac{y_i - a_i}{R_i}\right)^2 \tag{2.33}$$

where n_q is the number of manipulator joints, y_i the actual position of the joint i and R_i is the joint angle range between the middle-range position a_i and either the upper or lower angle limit, used to normalize the index. It is defined as:

$$R_i = \begin{cases} a_i - y_{im} & \text{if } y_i < a_i \\ y_{iM} - a_i & \text{if } y_i > a_i \end{cases}$$

where y_{iM} and y_{im} are the maximum and minimum angle limits of the i^{th} joint .

With this index, the grasp is optimal when all joints are at the middle-range position, having a quality measure of one, and it goes to zero when all its joints are at their maximum or minimum angle limits.

$\mathbf{Q_{D2}}$ - Inverse of the condition number of $\mathbf{G_J}$

This index measures the capability of a hand to move the object in any direction with the same gain. This is a dexterity measure which implies a good manipulation ability (Suárez et al., 2006). The ratio between the maximum singular value to its minimum singular value is defined as the condition number of a matrix. In the Jacobian, the inverse condition number gives a measure of the uniform transformation between the velocity in the finger joints and the velocity of the object. it is calculated as:

$$Q_{D2} = \frac{\sigma_{min}(G_J)}{\sigma_{max}(G_J)}$$

where σ_{min} and σ_{max} are the smallest and largest singular values of the grasp Jacobian matrix G_J (defined in Sec. 2.2). This measure has his lower limit in zero and upper limit is one, indicating a uniform transformation and a grasp with the maximum quality.

Chapter 3

Experimental platform

The experiments performed in this thesis will be done in both, simulation and real environments. In the first part of the experiments, simulation will be used to characterize quality metrics, obtain its statistical values and correlations. Then, real experiments will be performed to classify the grasp evaluated in simulation and to generate a prediction model for grasp success.

For the functional evaluation of robotic hands, simulation will be used to measure the grasping capabilities of different hand models. This properties will be compared to other indexes for measuring hands capabilities. Finally, this evaluation method will be applied to evaluate and improve the design of a prosthetic hand.

In this section a global view on the platforms, methods and materials will be explained. Details of the methodology applied in each experiment will be explained in the corresponding chapter.

3.1 Simulation environment

The first goal in this thesis is the study of quality metrics. In order to do that, different hands and objects will be used to generate and evaluate grasps in a simulated environment. We will use *OpenHand-Simulator* (Leon et al., 2013b) as our simulation environment. *OpenHand* is a simulation toolkit based in *OpenRAVE* (Diankov and Kuffner, 2008). *OpenRAVE* is an open architecture targeting a simple integration of simulation, visualization, planning, scripting and control of robot systems.

With *OpenRAVE* we are able to simulate different hands, objects and generate grasp hypotheses. *OpenHand* allows us to evaluate these grasp hypotheses using different quality metrics. Also, OpenHand, in contrast to OpenRAVE, presents an *User Interface* (UI) which allows to:

- Configure the parameters for grasp hypotheses generation
- Determine the number of hypotheses to be generated
- Set the grasp generation method, following an uniform or random distribution
- Manually generate grasp hypotheses over an object surface
- Evaluate grasp hypotheses using different quality metrics
- Compare the functional performance of robotic hand models

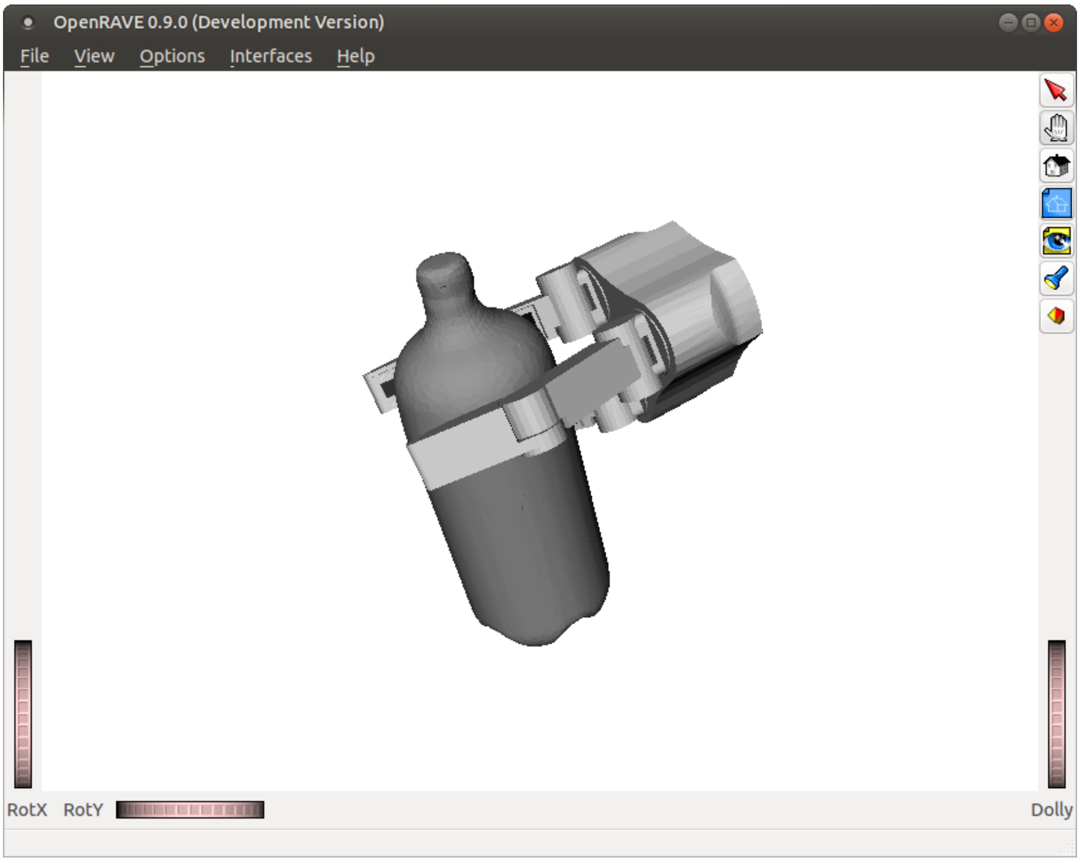


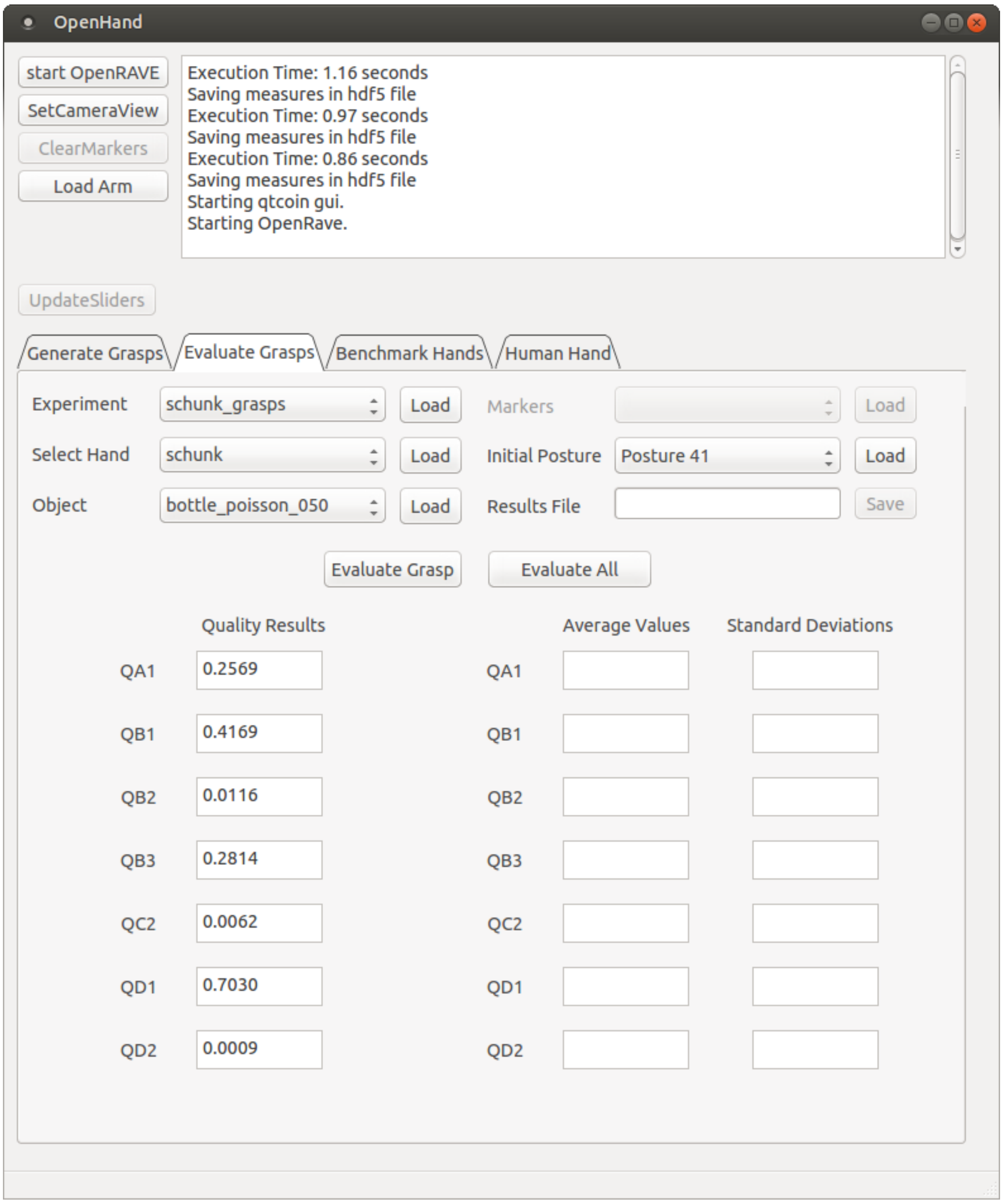
FIGURE 3.1: The OpenRAVE simulator.

3.1.1 OpenHand

OpenHand (Figure 3.2) presents a UI over *openrave*. It is based in *Python* and *Qt*. OpenHand uses static simulation, this is, the object and gripper pose never changes during the grasping process. OpenHand includes:

- User Interface: This interface allows the user to perform most of the common grasping tasks available in openrave through command-line functions
- Quality Metrics library: this python library has the implementation of 10 different common quality metrics used for grasping. This metrics can be used either trough the OpenHand UI or importing the library directly in the code.
- Utilities library: this python library implements some utilities needed by OpenHand to simulate and perform grasps. It includes also an upper layer over the some native functions of openrave.
- Human hand library: this is a C++ library which includes the simulation of a human hand. Parameters and characteristics of the hand can be configured through the UI or the code.

The OpenHand tool has 4 different main modules. With this modules, the user is capable of performing different experiments, and evaluate different grasp-related characteristics: grasp generation parameters, hand performances, object's graspability, grasp types, etc.


OpenHand
start OpenRAVE
SetCameraView
ClearMarkers
Load Arm
UpdateSliders
Generate Grasps
Evaluate Grasps
Benchmark Hands
Human Hand
Load
Markers
Load
Load
Load
Results File
Save
Evaluate Grasp
Evaluate All

```
Execution Time: 1.16 seconds
Saving measures in hdf5 file
Execution Time: 0.97 seconds
Saving measures in hdf5 file
Execution Time: 0.86 seconds
Saving measures in hdf5 file
Starting qtcoin gui.
Starting OpenRave.
```

Experiment: schunk_grasps

Select Hand: schunk

Object: bottle_poisson_050

Initial Posture: Posture 41

	Quality Results	Average Values	Standard Deviations
QA1	0.2569		
QB1	0.4169		
QB2	0.0116		
QB3	0.2814		
QC2	0.0062		
QD1	0.7030		
QD2	0.0009		

FIGURE 3.2: The OpenHand interface.

Human Hand This modules allows to load a virtual model of a human hand and arm, parameters such as hand length or width can be easily configured. With this module you can set the values for different joints of the hand and define different grasp postures. It allows you to load different objects models, perform and evaluate grasps. It is also possible to set different grasp types and configure which fingers will be closing for each grasp.

Generate Grasps This module allows the user to generate different sets of grasps using a wide variety of artificial hand models and objects. It allows to configure different parameters for the grasp generation and also to generate grasp poses manually. The grasps generated in this module are stored as *experiments* using the hdf5 file format.

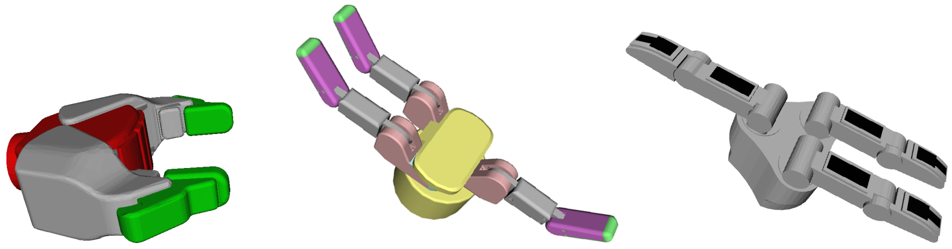
Evaluate Grasps With this module, you can load the different grasp experiments generated in the *generate grasps* module and evaluate them. It shows also the predicted successfulness of the grasp. The results of this grasp evaluation are shown in the UI but also stored in the hdf5 file of the experiment.

Benchmark Hands This module allows to evaluate the performance of different robotic hand models. The evaluation of a hand require a set of different grasp hypotheses and evaluates them in order to estimate the expected performance of the hand. It is possible to define the grasp generation method.

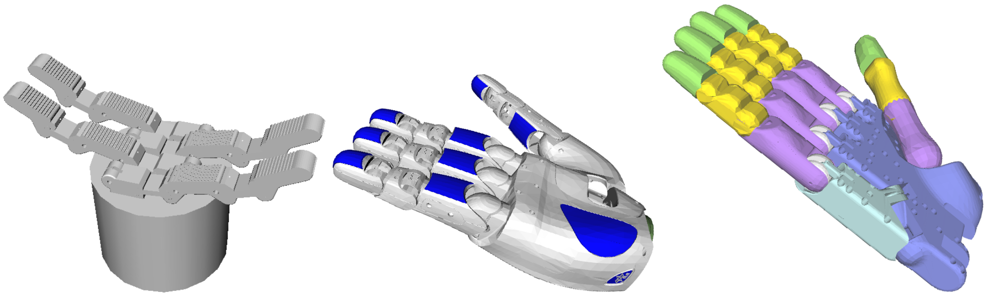
3.2 Artificial hands

The experiments in simulation are done using up to 11 different robotic hands, including common manipulators used in robotics and industrial applications, as well as prosthetic hands.

Four models correspond to classic robotic grippers: the *Barrett* hand, the *PR2* hand, the *Model-T* and the dexterous three-fingered Schunk *SDH* hand. Two models correspond to robotic anthropomorphic hands: the four-fingered Schunk *SAH* hand and the 5-fingered *Shadow* dexterous hand.



(1) PR2 hand (2) Barrett hand (3) Schunk SDH hand



(4) ModelT hand (5) Schunk SAH hand (6) Shadow hand

FIGURE 3.3: Robotic grippers.

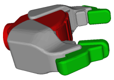
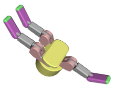
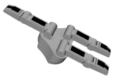
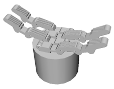
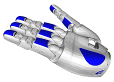
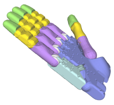
In our virtual models, all these grippers are assumed to have rigid bodies. In all the models, the finger joints close at the same rate. Table 3.1 summarize the details on each robotic hand model. Virtual models are shown in Figure 3.3.

The PR2 hand is a robotic gripper from the robot PR2 developed by Willow-Garage[1]. This gripper has two fingers with two joints each, controlled by the same DoF. For the virtual model, we added an extra DoF, making each finger independent.

---

[1] Willow Garage: `http://www.willowgarage.com/`

TABLE 3.1: Robot Hand Models

Hand	Name	Joints	Actuators	DOFs
	PR2 Gripper	4	2	2
	BarrettHand	6	4	8
	Schunk SDH	6	7	8
	Model T	8	4	8
	Schunk SAH	12	13	13
	Shadow Dexterous	15	7	17

The two links on the finger are still connected to the same DoF. When a collision is detected in a link the closure of the finger stops.

The Barrett hand is a robotic gripper developed by Barrett Technology[2]. It has 3 fingers with two joints each. Finger closure is controlled by strains. The closure starts in the lower link and when a collision is detected, it continues in the upper link. It has also an extra DoF for a symmetric abduction up to 180°of the two parallel fingers. For the virtual model, this abduction DoF is blocked and only varies to predefined positions prior the closure event.

The Schunk SDH hand is a robotic gripper developed by SCHUNK GmbH & Co. KG[3]. This gripper has 3 fingers with two joints each. Each joint has its own DoF. It also has an extra DoF for the counterwise rotation, up to 90°, of the two parallel fingers. In our virtual model this rotation is applied only in predefined positions prior the closure event.

The Model-T hand ((Ma et al., 2013; Dollar and Howe, 2010)), is an under-actuated robotic gripper designed for the Yale OpenHand Project [4]. This printable gripper has four fingers and two compliant flexure joints each. It has one actuator for all the fingers. During grasp acquisition, each finger will continue to move until the links make contact with the object. In our virtual model, we consider each joint independent, having its own DoF. In total this model has 8 DoFs.

The Schunk SAH hand is an anthropomorphic robot hand developed by SCHUNK GmbH & Co. KG. This hand has four fingers with three phalanxes each. Each joint has its own DoF, there is an additional DoF for the Thumb abduction. In total, this hand has 13 DoFs.

---

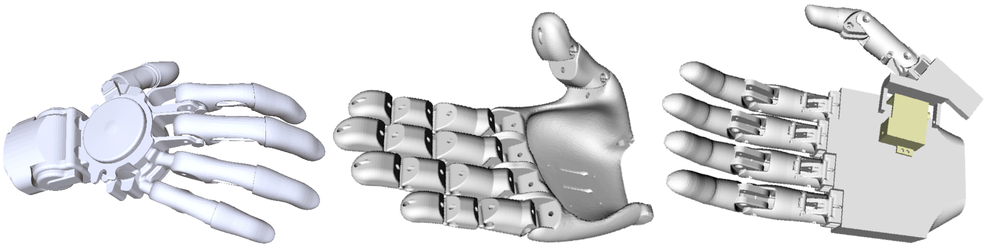
[2] Barrett Technologies: https://www.barrett.com/

[3] SCHUNK GmbH & Co. KG: https://schunk.com/us_en/homepage/

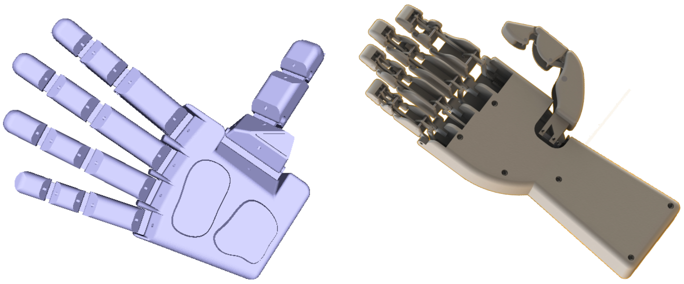
[4] Yale OpenHand Project: https://www.eng.yale.edu/grablab/openhand/

The Shadow dexterous hand is an anthropomorphic hand developed by Shadow Robot Company[5]. This hand has 5 fingers with three phalanxes each. Each joint has its own DoF. It also has two extra DoF for the abduction of the Thumb and the Little fingers. In our virtual model the abduction of the *little* is disabled. In total this hand has 16 DoFs.

Finally, the other five models correspond to hand prosthesis intended for humans: the *Michelangelo* hand, the *Fley-Hand* 2 prosthetic hand, the *BruJa* Model-B hand, the *Tact* hand and the *Imma* hand. These hands are designed to be used by amputees.



(1) Michelangelo hand (2) Flexy hand (3) Tact hand



(4) Imma hand (5) BruJa hand

FIGURE 3.4: Prosthetic hands.

These hand prosthesis are assumed to have rigid bodies. In all the models, the finger joints close at the same rate. Table 3.2 summarize the details on each prosthetic hand. Virtual models are shown in Figure 3.4.

The Michelangelo hand is a prosthetic hand developed by Ottobock[6]. this hand prosthesis has 5 fingers. It has one DoF for the closure of the Thumb and another for the other 4 fingers. It has an extra DoF for the abduction of the Thumb to up 6 pre-defined positions. In our virtual model this abduction is fixed. In total this model has 2 DoF.

The Flexy-hand 2 is a prosthetic hand developed by Gyrobot[7]. This hand has 5 fingers with three Phalanxes each. The thumb has only two phalanxes. There is

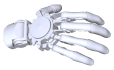
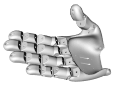
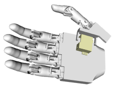
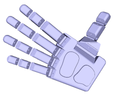
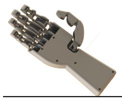
---

[5] Shadow Robot Company: `https://www.shadowrobot.com/`

[6] Ottobock: `https://www.ottobockus.com/`

[7] Gyrobot: `http://www.gyrobot.co.uk/`

TABLE 3.2: Prosthetic Hand Models

Hand	Name	Joints	Actuators	DOFs
	Michelangelo	6	2	3
	Flexyhand	14	5	14
	Tact hand	11	6	6
	Imma hand	15	6	15
	BruJa hand	15	6	6

no thumb abduction. The fingers are actuated by strains. In our virtual model each joint has its own DoF. Thus, this model has 14 DoFs.

The Tact hand is a prosthetic hand developed by Slade et al., 2015. It has 5 fingers with two phalanxes each. Fingers are linked by a bar mechanism. When a collision is detected in a phalanx the closure of its finger stops. It has an extra DOF for the Thumb abduction. In total this hand has 6 DoFs.

The Imma hand is a prosthetic hand developed in the Devalhand Project[8]. This hand has 5 fingers and three phalanx each. The Thumb has two phalanxes for the closure and another for the abduction. The joints are linked trough strains. It has one actuator for each finger and a last one for the Thumb Abduction. Our virtual model for this hand has 15 DoFs.

The BruJa hand is a prosthetic hand developed in the Devalhand Project. This hand has 5 fingers and three phalanxes each. It has one DoF only for the Thumb abduction. Joints are linked through a bar mechanism. In our experiments we use the original design of the BruJa hand Model-B inspired in the TBM hand (Dechev et al., 2001). This hand has 6 DoF.

Except for the Michelangelo hand, the other four models can be built using 3D printing technologies. Models for download and instructions for print and build are available on their web pages.

On chapter 4, the PR2, Barrett, Schunk SDH and SAH, Shadow, Michelangelo and Model-T hands are used. Chapter 5 uses the Barrett and Schunk SDH grippers. Finally the 11 hand models are used in Chapter 6.

---

[8] Devalhand Project: `https://sites.google.com/a/uji.es/devalhand/`

3.3 Object models

In order to perform the grasps, different sets of objects are used. First, we will explode the KIT objects Database (Kasper et al., 2012). This database contains virtual models of more than a hundred different objects commonly found in a domestic environment. Figure 3.5 shows a sample of some of the objects in the database.

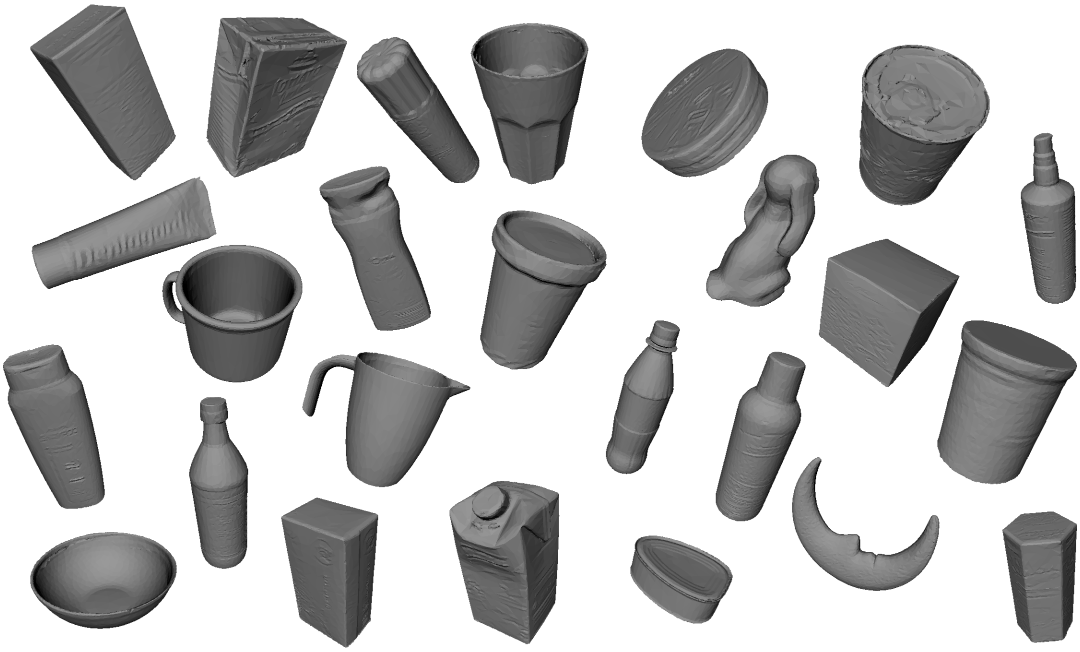


FIGURE 3.5: Sample of the different objects from the KIT Database.

The second set of objects are from the 3DNet database (Wohlkinger et al., 2012). This database contains more than 700 different objects. Some of the object models of the dataset are built physically using 3D printing technologies for the real experiments. Figure 3.6 shows the virtual models of the objects used for real experiments.

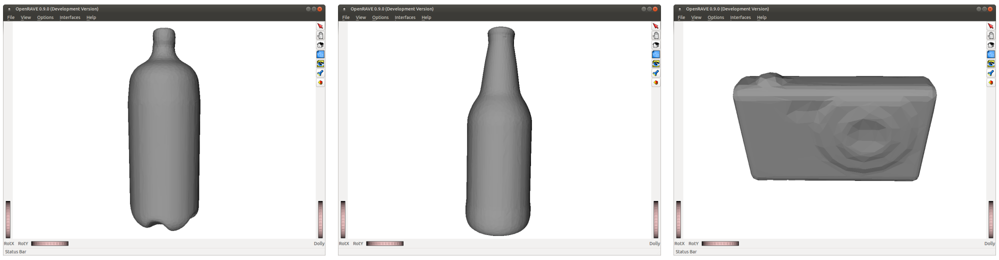
The third set of object models used are from the *Yale-CMU-Berkeley (YCB) Object and Model set* (Calli et al., 2015b; Calli et al., 2015a). This dataset contains more than 30 physical models of the objects. This objects were virtually modeled to be usable in the simulation environment.

Finally, we use 15 common objects from the *ADLs* (Feix et al., 2013). This objects are used for the computation of the *Anthropomorphic Index* of prosthetic hands. With this set of objects is possible to represent all the hand configurations of the taxonomy of human grasps.

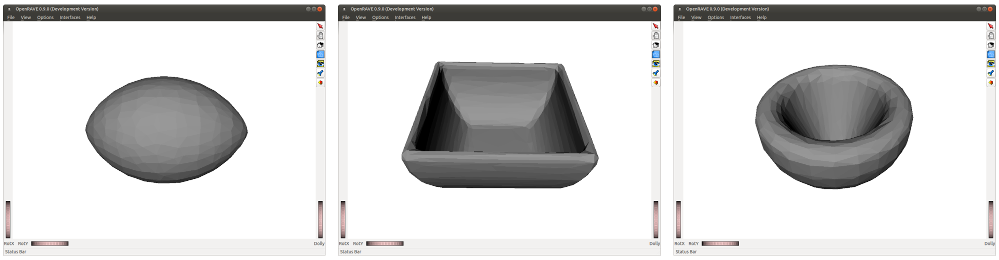
Objects from the KIT Database are used in chapter 4 and chapter 6. Objects from the 3DNet are used in chapter 5. Finally objects from the YCB dataset and ADLs are used in chapter 6.

3.4 Grasp generation

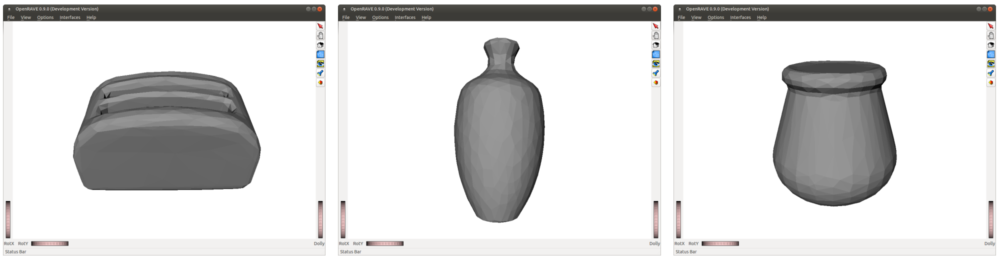
In pursuance of obtaining a significant set of candidate grasps applicable to an object, a large set of grasp hypotheses needs to be generated and evaluated. A grasp hypothesis is defined as the combination of an object, a hand, an initial position of the hand with respect to the object, and the initial configuration of the hand joints. Generation of a variety of grasp hypotheses for a given object with a given hand is



(1) Bottle_050 (2) Bottle_047 (3) Camera_015



(4) Lemon_003 (5) Bowl_022 (6) Bowl_025



(7) Toaster_001 (8) Jar_002 (9) Jar_004

FIGURE 3.6: Virtual object models used for the real experiments from the 3DNet database.

performed using the Grasping Module[9] from the Database Generators available in OpenRAVE.

This algorithm provides a parametric approach to generate a number of grasps distributed over the object surface. Among the different algorithms to generate candidate grasps we have chosen this because a number of reasons: it is able to produce an almost uniform distribution over the object surface; it can be tuned easily to produce more or less candidates; it starts from the bounding box of the shape, which limits the candidates to grasps to be able to approach the object with a gripper; and last but not least, its implementation is fully available.

In short, the algorithm generates a square grid of points uniformly distributed around the object bounding box and projects it over the object surface (Fig. 3.7). From each of these points, an approach ray is created pointing outwards the object surface. The hand is placed along the approach ray, facing the object at a given distance, and it can be rotated about the approach ray by a given angle.

The algorithm depends on five parameters (Fig. 3.8):

- δ**:** Distance between the points in the square grid around the surface of the bounding box to place the approach rays.

---

[9] OpenRAVE Grasping Module: http://openrave.org/docs/0.6.6/openravepy/databases.grasping/

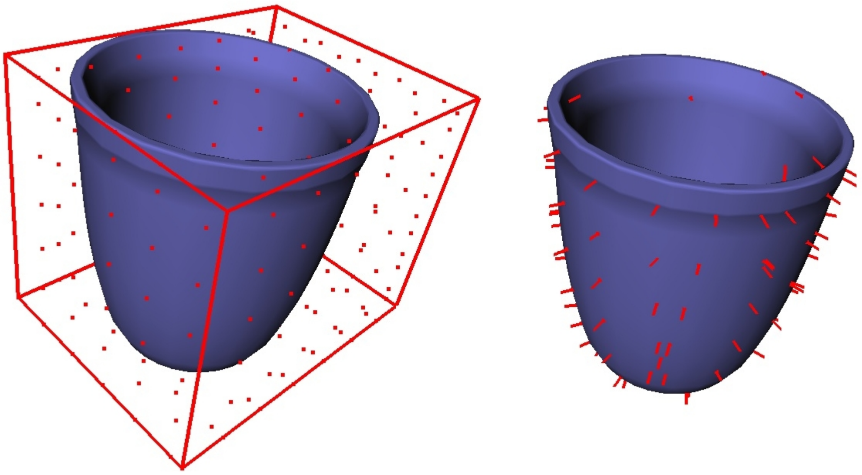


FIGURE 3.7: Example of the approach-rays generation process used by OpenRAVE: (left) square grid of points around the object and (right) the projection of this box over the object surface.

- α**:** Angle between the approach ray and the normal to the object surface (default value of zero). Depending on the value of α, a set of oblique approaching rays is generated.
- θ**:** Standoff distance along the approach ray where the reference point of the hand is placed.
- N_r: The angles to rotate the hand (roll) about the approach ray.
- H_J: Configuration of the hand joints (preshape) before the closing algorithm starts.

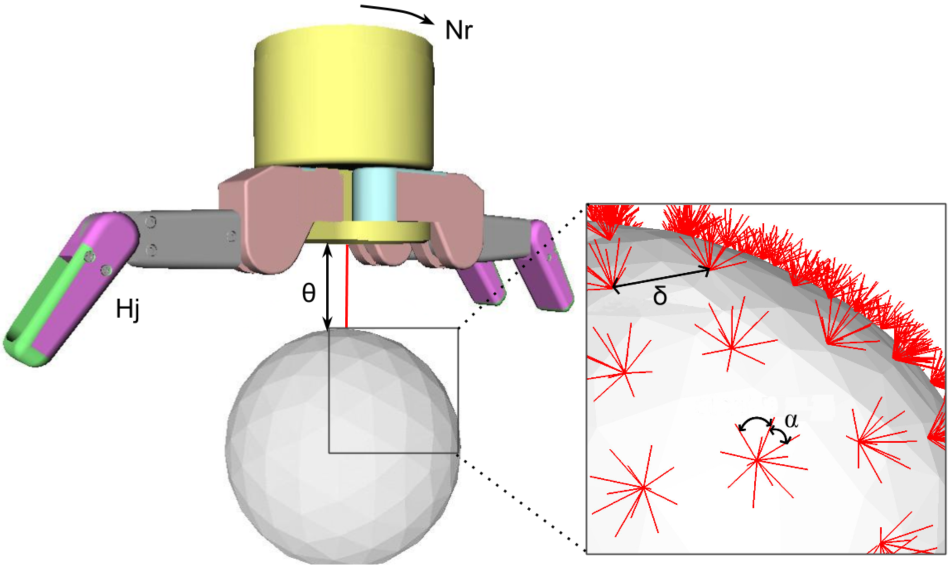


FIGURE 3.8: Example showing the parameters needed for grasp generation.

Using these parameters, a large variety of grasp hypotheses are generated, and subsequently evaluated with the methods described in the following chapters. The size of the objects also affects the number of hypotheses generated, varying from a few hundreds to dozen of thousands. Section 4.1 analyses the selection of these parameters.

This grasp generation method is applied in all the chapters. As the simulation and generation of grasps is computationally expensive, in some experiments there are variations in the parameters used. This will reduce the computational cost of the experiments.

3.5 Robotic platforms

The evaluation of real grasps (chapter 5) will be done using two different robotic platforms: the Apollo robot at the Max-Planck Institute for Intelligent Systems and the Tombatossals robot at the Robotic Intelligence Laboratory at the Universitat Jaume I.

Apollo Apollo[10] (Figure 3.9) is a dual-arm manipulation platform used to study active perception, grasping, and manipulation. It has two arms, hands with tactile sensors, and an active vision head. The robot has two KUKA lightweight robot arms (7 DOFs), two Barrett hands, and a Sarcos head featuring different vision sensors. This robot has also an *Asus Xtion PRO* for the vision system.

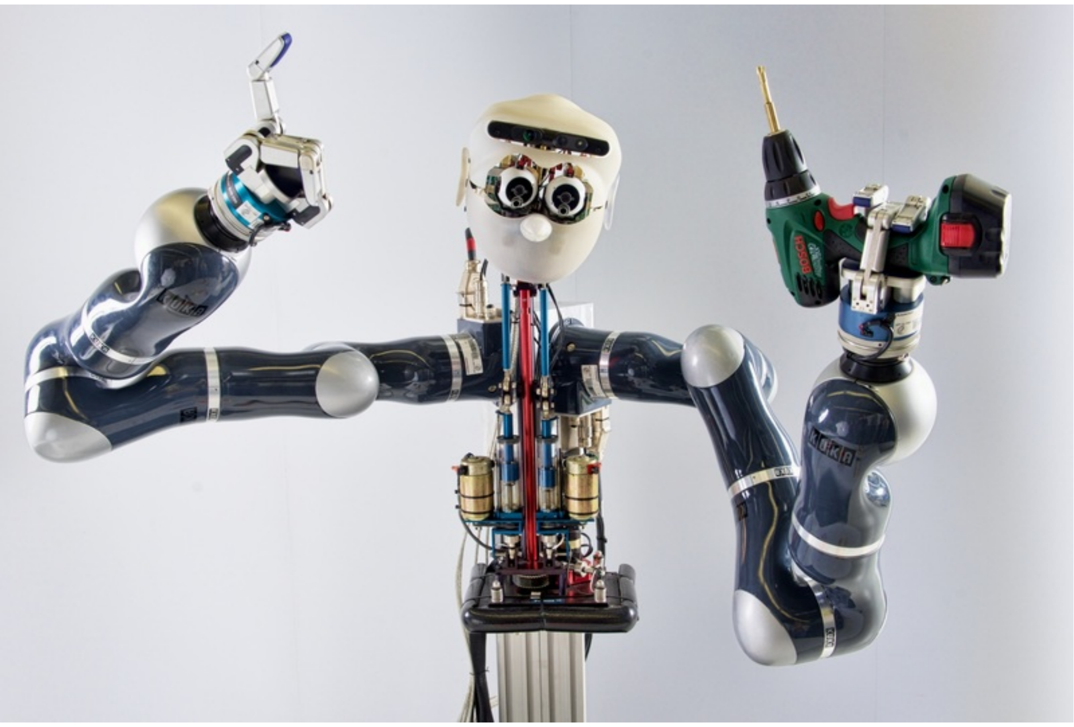


FIGURE 3.9: The Apollo robot system.

Tombatossals Tombatossals[11] or *tombato* (Figure 3.10) is a a multipurpose humanoid torso, for research of autonomous grasping and manipulation tasks in unstructured household scenarios. The humanoid torso is composed of two arms, two hands and a head for a total of 29 DOF. Both arms are *Mitsubishi PA10-7C*, 7 DOFs industrial manipulators with a position repeatability of 0.1mm. Each arm weights 40Kg and has 10Kg payload.

---

[10] Apollo Robot: `https://am.is.tuebingen.mpg.de/pages/robots`

[11] Tombatossals Robot: `http://robinlab.uji.es/our_robots`

It has a *Barrett Hand* in its right arm and a *Schunk SDH Hand* on the left arm. Both hands are equipped with tactile sensors from *Weiss Robotics* [12]. The head is composed of a *TO40* pan-tilt-vergence system and a *Kinect*. More details on Tombatossals are described in Felip et al., 2015.

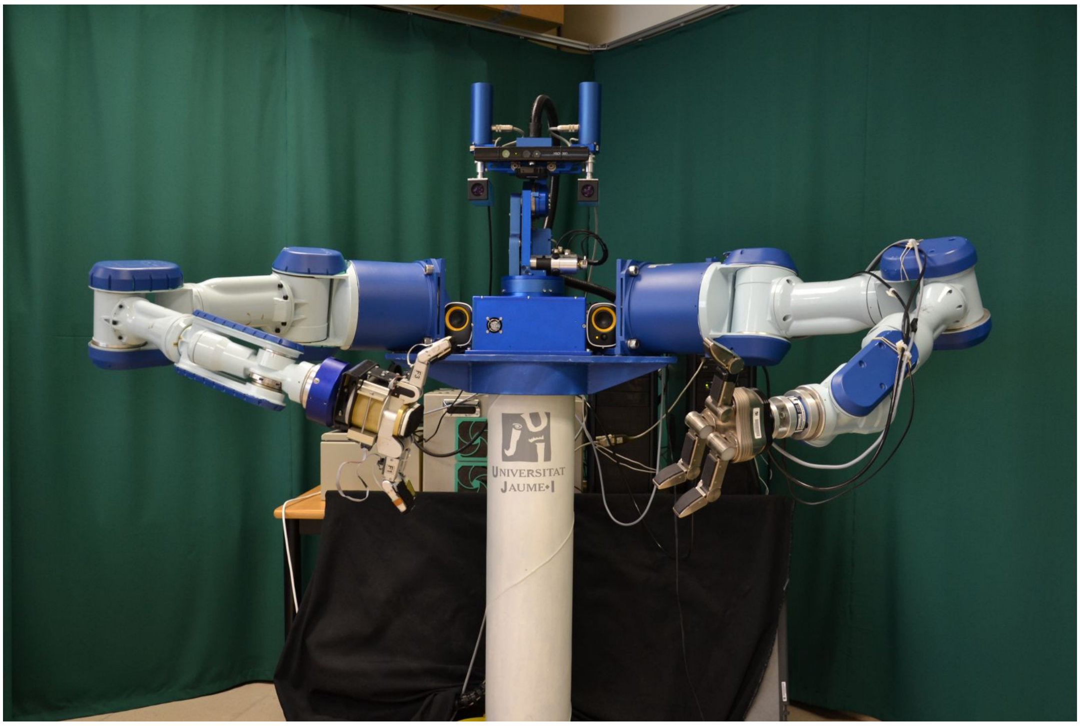


FIGURE 3.10: The Tombatossals robot system.

3.5.1 Robotic Grippers

The *Barrett* and *Schunk SDH* grippers (Figure 3.11) are used for performing the experiments with real robots. In both models the closure of the fingers is done until a contact is detected. Then, the joint is blocked and the closure continues for the distal joint. For the Barrett hand, a strain measurement detects the collision in the finger. In the case of the Schunk SDH, there are tactile sensors for detecting the contact on each link of the hand.

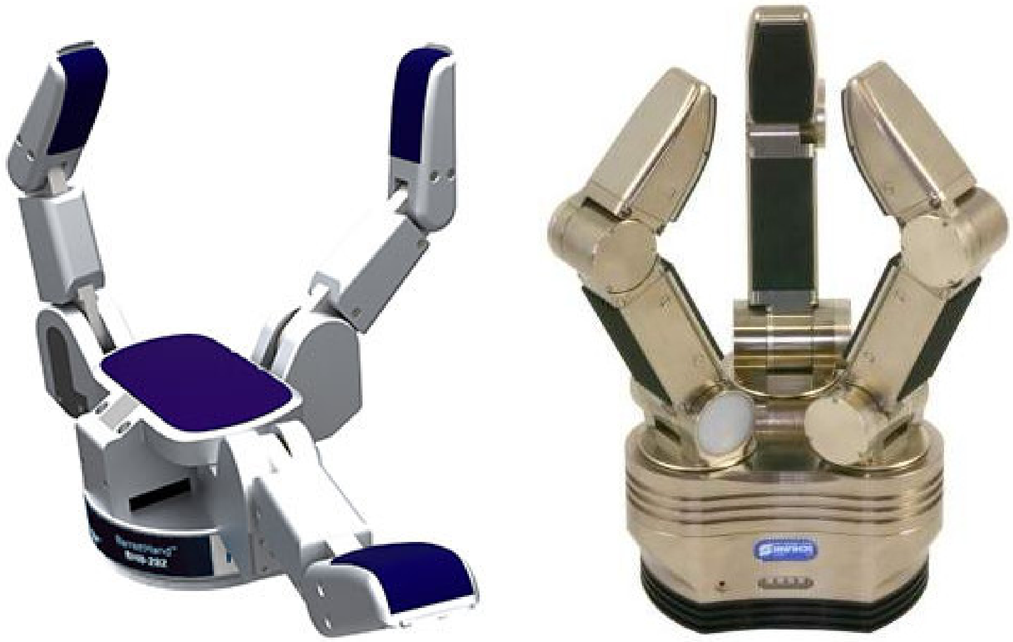
The *Barrett* hand has a weight of 980 g. Its payload is 6 kg. It has three fingers with two joints each. Two of them have an extra degree of freedom with 180 degrees of lateral mobility supporting a large variety of grasp types. All joints have high-precision position encoders. It has 3 fingertip torque sensors, one per finger.

The Schunk *SDH* hand weights 1.95 Kg. It has 3 fingers with two links each. Two fingers can contrariwise rotate up to 90°. It presents high flexibility in terms of shape, size and position of the objects to be gripped. It has 6 tactile sensors for pressure and surface recognition.

These two robotic grippers are used for performing the experiments on Chapter 5.

---

[12] Weiss Robotics sensors: `http://www.weiss-robotics.de/`



(1) Barrett hand (2) Schunk SDH hand

FIGURE 3.11: Real models of the robotic grippers.

3.6 Real grasp evaluation

Some of the grasps generated and evaluated in simulation, will also be tested in the real world. Two different experiments are performed for this evaluation. One with the Barrett hand and the Apollo robot. The other experiment will be done with the Schunk SDH Hand and the Tombatossals robot. The grasps evaluated in real world will be labeled as *stable* or *unstable*.

For the experiments it is necessary to take in account different constraints from real world, not present in the simulation environment. First, the influence of gravity can be critical for the success of a grasp. Second, the constraints of the workspace limits the number of reachable grasps. The kinematic of the robot and the presence of objects, as the table surface, are obstacles to perform grasps. Third, the vision system and object tracking is a critical point. Having a perfect estimation of the object pose is complicated.

It is necessary to assume there will be always a small error when tracking the object's pose. The movements of the robots could also be erratic. The grasp pose is highly dependant on the accuracy of the robot arms. As with the tracking of the object, a small error should be assumed for the gripper pose before grasping. The vision system used in the experiments is based in the *Depth Based Object Tracking Library (dbot)* (Wüthrich et al., 2013; Issac et al., 2016). This library only needs an MS Kinect or Asus XTION depth sensor, and the mesh model, to detect and track the pose of an object.

As setup for the experiments (Figure 3.12), objects are placed on a table. Objects in a table surface require to be stable prior to the grasp execution, which could discard several grasp hypotheses. In the experiments, the robot should reach the grasp pose, close the fingers and lift the object to ensure the stability of the grasp. If the object is lifted and does not slip from the fingers, the grasp will be considered as stable. For every candidate grasp 3 tries are performed in order to be sure the grasp is stable or not.

Specific details on the experimental protocol will be described in Chapter 5.5.

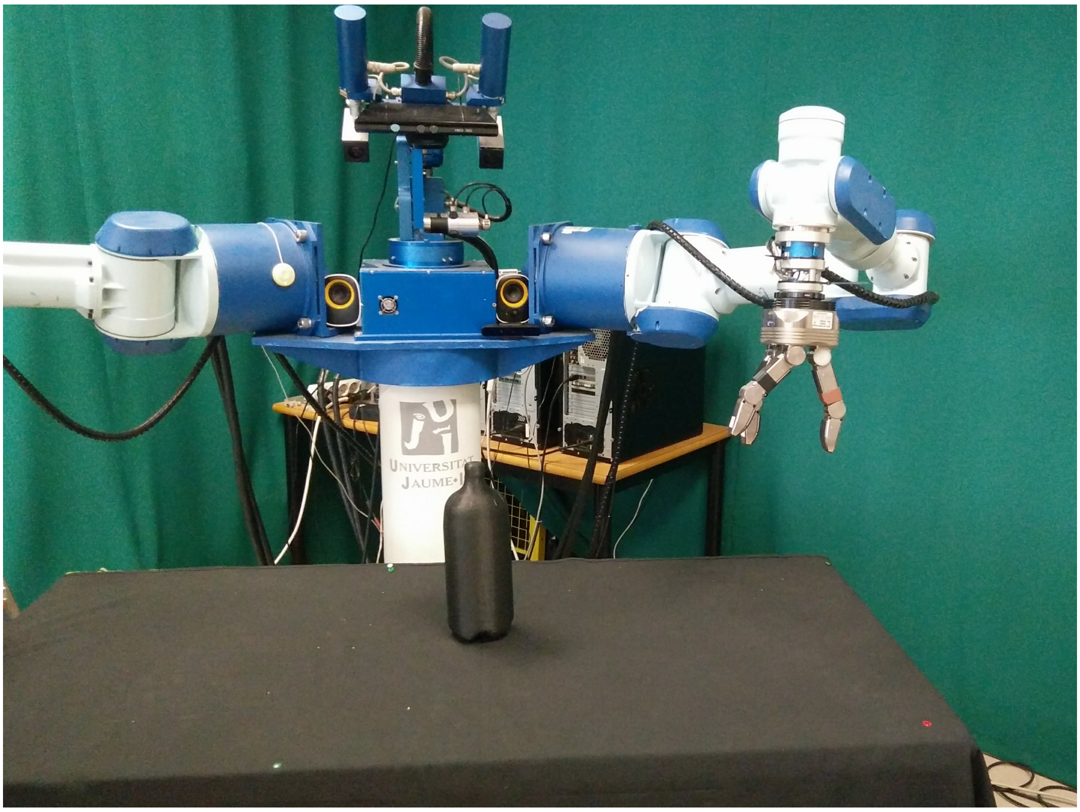


FIGURE 3.12: Example of the experimental setup with the Tombatossals robot.

3.7 Robotic hands evaluation

The evaluation of hands performance is done in simulation. Two different methodologies are applied to evaluate the design of robotic hands. On one hand, an statistical study on their properties when grasping is done. On the other hand, a grasp prediction model is used to evaluate the successfulness of a hand for grasping.

The statistical study on the hand properties is based in the use of quality metrics. Grasp metrics are considered as evaluators of different hand properties. In this study several grasps are evaluated among different objects. The values obtained from this experiment give us an estimation on the performance of the hand according to each property. This evaluation allows to establish similarities and difference between hand models.

Using a grasp success prediction model it is possible to estimate how likely is a prosthetic hand to successfully grasp an object. With this model we can estimate how often a grasp will we *robust*, *fragile* or *futile*. This evaluation gives an assessment on the grasp performance of artificial hands.

Specific details on the evaluation of artificial hands will be seen in chapter 6.

3.8 Grasp databases

Different databases of objects and grasps are used to perform the experiments. The first database contains objects from the *KIT object model* dataset. This database of grasps is used to analyze the relations and behavior of quality metrics. This database is used for the experiments on chapters 4 and 6.

The second database of grasps is from Kappler et al., 2015. This database contains objects from the 3DNet. Simulated grasps of this database are evaluated using physics simulation and human labeling. For the human labeling, different humans are sown grasps from three angles. These humans have to label the grasps as *stable*, *unstable* or *unknown*. Some of the grasps on this database will be evaluated with real world with the Apollo platform. More details of this database are explained in Chapter 5.

The third database contains grasp candidates evaluated both, in simulation and real world. This database contains objects from the 3DNet object model set. Grasps on this database are executed using the Tombatossals robot platform. These grasps are classified using a 3-grade scale: *Robust*, *Fragile* and *Futile*. This database is used for the experiments on chapter 5.

The fourth database contains grasps with objects from the ADLs. This database is used for estimating the *AI* on different anthropomorphic hands and compare it to the evaluation with quality metrics. This database is used in the experiments of chapter 6.

Finally, a database of grasps containing objects from the *YCB* model set is used. This database contains 24 different objects modeled in virtual environments. This database is used to improve the design of the *Imma* prosthetic hand. Details on this database are explained in chapter 6.

More details on the objects, hand models and grasp generation method for each database will be explained in the corresponding chapter.

Chapter 4

Quality metrics analysis

The main goal of the work described in this chapter is to answer the relevant questions about practical use of grasp quality metrics. For each metric, which are the practical ranges within a grasp that can be considered *good* or *bad*? How sensitive is the evaluation of a grasp with respect to contact location uncertainty? And, which quality metrics are similar or measure different grasp aspects?

To address these questions this chapter describes a numerical study of the characteristics of a set of ten metrics selected among the most commonly used, with the purpose of deepening the knowledge on them. This provides factual data to decide about which metrics to select and how to combine them. The assumed focus of this work is on metrics oriented to the analysis of static grasps performed by a robot gripper on rigid 3D objects for holding and transporting them.

The contribution of this study is a set of numerical ranges and thresholds which establish the practical parameters that can be used to interpret the grasp evaluation provided by the selected quality metrics. The methodology followed consists in the statistical analysis of the quality values obtained for 10 different metrics (Chapter 2.4) used to evaluate the grasps of a wide database, generated on a simulation environment using 126 different models of rigid objects (Chapter 3.3) and 7 models of robot hands (Chapter 3.2).

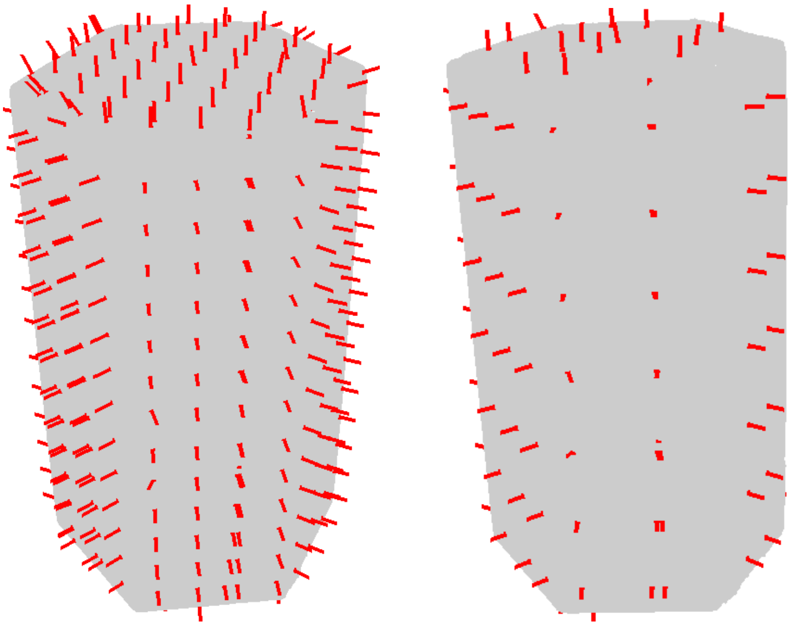
The study contains three different analyses:

- **Variability analysis:** The distribution of each metric values are studied when grasping different objects with different robotic hands and with different grasp configurations. The range of values for each metric is established and used to homogenize the scales of the different metrics and to set practical thresholds for *good* and *bad* grasps (see Section 4.2).
- **Sensitivity analysis:** Metrics are recomputed when displacements and rotations are applied to grasp postures. The robustness of each metric with respect to inaccuracies in the placement of the finger contact points is determined (see Section 4.3).
- **Correlation analysis:** The underlying relations between metrics are determined (see Section 4.4).

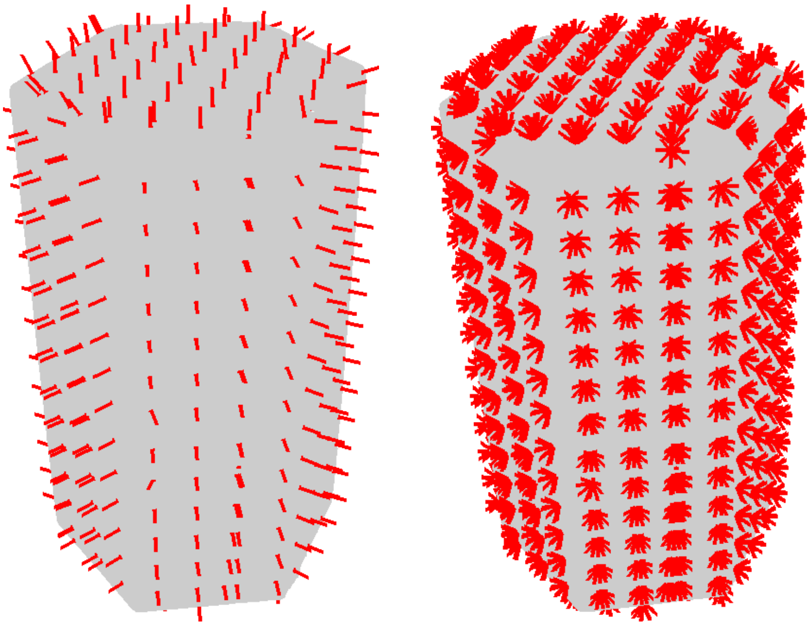
The section is structured as follows. Subsection 4.1 analyze the parameters used to generate grasp hypotheses. Subsections 4.2 to 4.3 explain the analyses on variability and sensitivity and show their results. Subsection 4.4 analyze the correlations between metrics.

4.1 Grasp generation study

An initial study on the parameters used to generate grasp candidates is done. This study aims to avoid the possible biases derived from generating grasps under the same circumstances, discarding other possibilities. The procedure to generate grasp candidates is explained in Chapter 3.4.



(1) Variations in δ (0.01 to 0.02 m)



(2) Variations in α (0 to $\pi/3$ rad)

FIGURE 4.1: Example showing the variations in δ and α studied to generate approach rays

A proper selection of the values for the parameters presented in Section 3.4, especially for δ and α, is required to generate an appropriate set of grasps hypotheses for every pair object-hand. Although Rombokas et al., 2012; Balasubramanian et al., 2012 showed a high efficiency when using orthogonal approach angles for grasp success, considering only orthogonal approach angles could introduce a bias on the grasp set.

A preliminary study on the the influence of the selection of δ and α values on the quality metrics is performed. A subset of 21 different objects from the *KIT database* and two hands (BarrettHand and Schunk SDH) is selected. Two δ values are tested: a small value (0.01 m) and a high value (0.02 m) (Fig. 4.11).

Additionally, three ways for generating the direction of the approach rays are tested (Fig. 4.12): orthogonal to the object surface ($\alpha = 0$), at a fixed angle ($\alpha = \pi/3$), and at a variable angle ($\alpha \in [0, \pi/3]$). To summarize, four methods are tested with a combination of the selected δ and α values, as presented in Table 4.1.

TABLE 4.1: Parameters used in each approach-rays generation method

Generation method		α (rad)	δ (m)	N° hypotheses
1	Small delta	0	0.01	84968
2	Big delta	0	0.02	24208
3	Fixed alpha	$\pi/3$	0.01	4078464
4	Variable alpha	[0, $\pi/3$]	0.01	4163432

The other parameters needed to define the grasp hypotheses are selected as follows, and are kept unchanged throughout the four methods:

$$\theta = [0.01\ \mathrm{m}, 0.02\ \mathrm{m}]$$
$$N_r = [0, \pi/2, \pi, 3\pi/2, 2\pi]$$
$$H_J = \begin{cases} [0,0,0,0,0,0] & \text{if hand != BarrettHand} \\ [0,0,0,0], [0,0,0,\pi/4], \\ [0,0,0,\pi/2], [0,0,0,3\pi/4] & \text{if hand = BarrettHand} \end{cases}$$

Using these parameters with the procedure described in Chapter 3.4 a set of grasp candidates is produced . The number of grasp hypotheses generated per object ranges from a few hundreds with the first two methods to hundred of thousands with the other methods. Thus, in order to have a tractable number of candidate grasps for these former methods, a 1% of the total amount of candidates are randomly selected in this preliminary study.

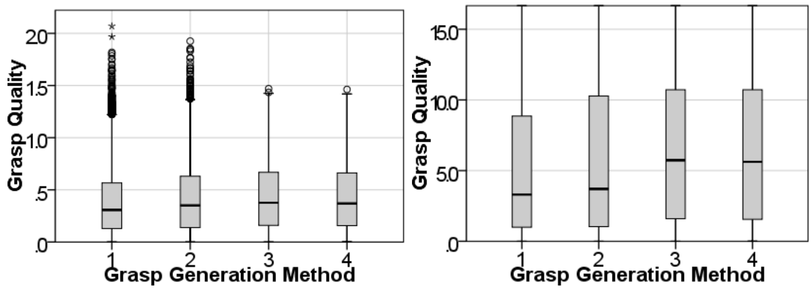
Next, the grasp quality values obtained using each metric for all the candidates are statistically analyzed; median and interquartile ranges of the different distributions obtained from the use of the four different methods are compared in order to select the most appropriate to generate the approach rays.

4.1.1 Results

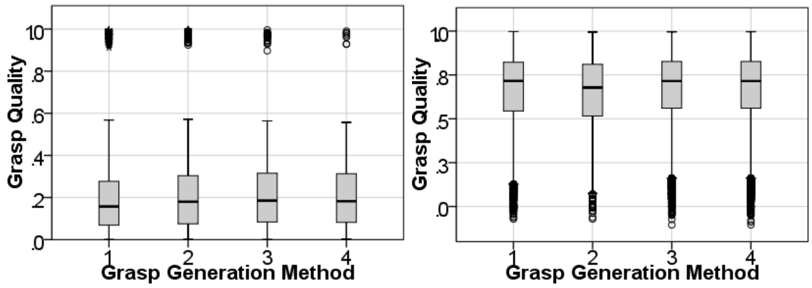
The median and interquartile ranges of the quality metrics obtained in the preliminary study are summarized in the box plots presented in Fig. 4.2. The dark line in the middle of the boxes is the median. The bottom of the box indicates the 25th percentile. The top of the box represents the 75th percentile. The *T-bars* that extend from the boxes extend to 1.5 times the height of the box or, if no case/row has a value in that range, to the minimum or maximum values. The points are outliers (extreme values). The asterisks or stars are extreme outliers.

No significant differences between the grasp qualities obtained using the different methods can be observed from the box plots. Furthermore, the results do not consistently show better grasp qualities for any of the methods.

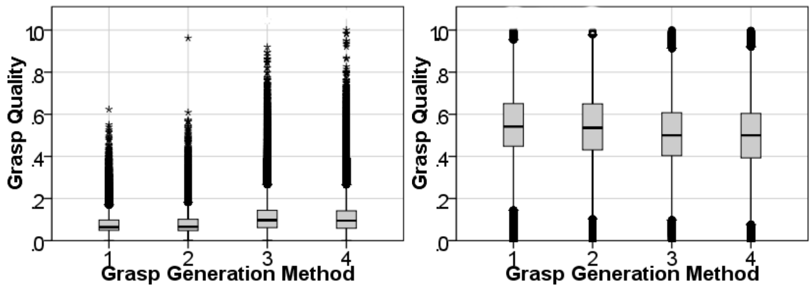
The variable method to generate approach rays provides a wider range of possibilities for grasping, and it is, thus, selected for analyzing the full set of grasps. A



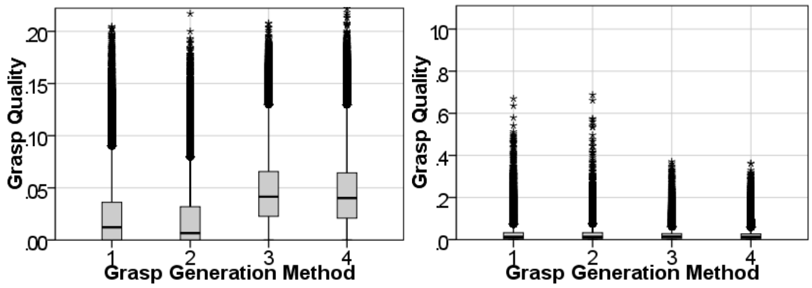
(1) Q_{A1} (2) Q_{A2}



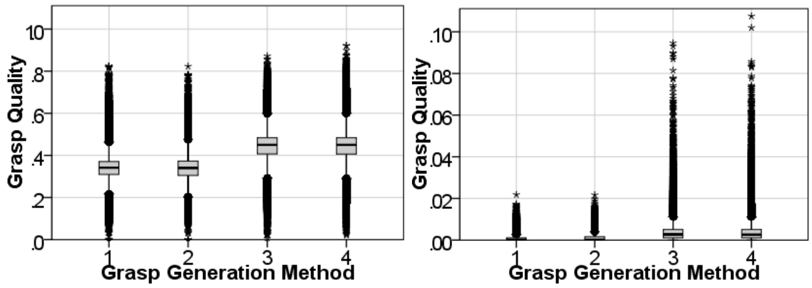
(3) Q_{A3} (4) Q_{B1}



(5) Q_{B2} (6) Q_{B3}



(7) Q_{C1} (8) Q_{C2}



(9) Q_{D1} (10) Q_{D2}

FIGURE 4.2: Results of the grasp quality metrics for the different approach-rays generation methods detailed in Table 4.1. The grasp quality scale varies to better how the results.

summary of the selected parameters used to generate the grasp hypotheses in the next subsections, is presented in Table 4.2.

TABLE 4.2: Parameters selected to generate grasp hypotheses

	Value	Hand
δ	0.01 m	All
α	$[0, \pi/3]$	All
θ	$[0.01\ \mathrm{m}, 0.02\ \mathrm{m}]$	All
N_r	$[0, \pi/2, \pi, 3\pi/2, 2\pi]$	All
H_J	[0,0]	PR2
	[0,0,0,0], [0,0,0,$\pi/4$], [0,0,0,$\pi/2$], [0,0,0,$3\pi/4$]	Barrett
	[0,0,0,0], [0,0,0,$\pi/4$], [0,0,0,$\pi/2$], [0,0,0,$3\pi/4$]	Schunk SDH
	[0,0,0,0,0,0,0,0,0,0,0,0,0]	Schunk SAH
	[0,0,0,0,0,0]	Shadow
	[0,0,0,0]	Model T
	[0,0,0]	Michelangelo

4.2 Variability

The distribution of the values for each of the selected metrics is studied in this section. This is relevant, as previous studies suggested that these values tend to be contained in very narrow ranges (León et al., 2013). In this section, the range of values for each metric is experimentally established and subsequently used for its normalization, so that all the metrics have a more homogeneous distribution in the range [0,1]. Also, the distribution is used to set practical thresholds for *good* and *bad* grasps.

The methodology consists in generating and evaluating a sufficiently large amount of grasp hypotheses for every pair object-hand on the simulation environment. For each of these selected grasps, the values for each quality metric are obtained and statistically analyzed. With these results, appropriate thresholds are selected to normalize each metric in the range of [0,1].

Once all the required parameters are selected from the results of the preliminary study, a set of grasp hypotheses is generated for each pair hand-object, using the selected parameters. With this method more than 25 million grasp hypotheses are generated. In order to make it computable, we decided to select only 1.000 random grasps per object and hand, being a total of 882.000 different grasps evaluated within the 7 hand models.

Each grasp is simulated and, using the contact information, the grasp quality is evaluated according to each of the 10 selected metrics. First, the *Probability Density Function* (*PDF*) for each metric is estimated. This provides information about the distribution of the data, its shape, variability and how is grouped. This information is useful, as it could be used as benchmark for analyze grasp generated and evaluated in other databases. Then, different methods are compared in order to select appropriate thresholds to normalize the data and make it comparable.

Different options are considered to select appropriate thresholds to normalize each quality metric: 1) *Full Range*, using the maximum and minimum values obtained for each metric; 2) *Soft Statistics*, using the mean ± the standard deviation to define the thresholds (to minimize the effect of atypical values that may appear in case of values grouped within smaller limits Leon et al., 2014); 3) *Percentiles*, using the 10^{th} and 90^{th} percentiles of the grasp quality values (to avoid negative values for the minimum threshold in case of high standard deviations).

Once the appropriate threshold values are selected for each metric, the following formula is used to normalize the values:

$$Q_N = \frac{Q - T_{min}}{T_{max} - T_{min}} \tag{4.1}$$

where Q is the grasp quality value and T_{max} and T_{min} the selected lower and upper thresholds for each metric.

Finally, the normalized quality values are graded using a four level scale: *bad*, *fair*, *good*, and *great* quality. The grasp falling under 0 are qualified as *bad*, those between 0 and the 50-percentile value are labeled as *fair*, those up to 1 are *good*, and finally those over 1 are considered *great*. In any case, this classification has no practical implication and is given as a suggestion to sort out grasps.

4.2.1 Probability density function

For each quality metric, several *Probability Density Functions (PDFs)* are tested in order to obtain the best fit for the distribution of data [1]. Results reported are the best 4 *PDFs* on each metric, and the corresponding parameters for fitting the data to the distribution function. The principal *PDF*s found were[2]:

- Exponential: The exponential distribution (also known as negative exponential distribution) is the probability distribution that describes the time between events in a Poisson process. The probability density function (pdf) of an exponential distribution is:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{4.2}$$

- Extreme Value: The extreme value distribution is appropriate for modeling the smallest value from a distribution whose tails decay exponentially fast, for example, the normal distribution. It can also model the largest value from a distribution, such as the normal or exponential distributions, by using the negative of the original values. The probability density function for the extreme value distribution with location parameter μ and scale parameter σ is

$$y = f(x|\mu,\sigma) = \sigma^{-1} exp(\frac{x-\mu}{\sigma}) exp(-exp(\frac{x-\mu}{\sigma})) \tag{4.3}$$

- Generalized Extreme Value: The generalized extreme value (GEV) distribution is the only possible limit distribution of properly normalized maxima of a sequence of independent and identically distributed random variables. The GEV distribution is often used as an approximation to model the maxima of long (finite) sequences of random variables. The probability density function is:

$$f(s;\sigma,\xi) = \frac{1}{\sigma} \begin{cases} (1+\xi s)^{(-1/\xi)-1} \exp(-(1+\xi s)^{-1/\xi}) & \xi \neq 0 \\ \exp(-s) \exp(-\exp(-s)) & \xi = 0 \end{cases} \tag{4.4}$$

---

[1] The best fit is measured using the Bayesian Information Criterion (*BIC*) https://en.wikipedia.org/wiki/Bayesian_information_criterion

[2] More details on the pdfs can be found in: https://es.mathworks.com/help/stats/fitdist.html

- Generalized Pareto: The generalized Pareto distribution (GPD) is a family of continuous probability distributions. It is often used to model the tails of another distribution. It is specified by three parameters: location μ, scale σ, and shape ξ. The probability density function (pdf) is:

$$f_{(\xi,\mu,\sigma)}(x) = \frac{\sigma^{\frac{1}{\xi}}}{(\sigma + \xi(x-\mu))^{\frac{1}{\xi}+1}} \tag{4.5}$$

- Normal: The normal (or Gaussian) distribution is a very common continuous probability distribution. Normal distributions are important in statistics and are often used in the natural and social sciences to represent real-valued random variables whose distributions are not known. The probability density of the normal distribution is:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.6}$$

  Where μ is the mean or expectation of the distribution (and also its median and mode), σ is the standard deviation and σ^2 is the variance.

- Rician: The Rice distribution, Rician distribution or Ricean distribution is the probability distribution of the magnitude of a circular bivariate normal random variable with potentially non-zero mean. The probability density function is:

$$f(x \mid \nu, \sigma) = \frac{x}{\sigma^2} \exp\left(\frac{-(x^2+\nu^2)}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right) \tag{4.7}$$

  where $I_0(z)$ is the modified Bessel function of the first kind with order zero.

- t Location-Scale: A location–scale family is a family of probability distributions parametrized by a location parameter and a non-negative scale parameter. The t-distribution is any member of a family of continuous probability distributions that arises when estimating the mean of a normally distributed population in situations where the sample size is small and population standard deviation is unknown. The probability density function is:

$$p(x \mid \nu, \mu, \sigma) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}\sigma} \left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}} \tag{4.8}$$

4.2.2 Statistic values and normalization

For the study of variability on each quality metric, different statistical values are obtained and compared: mean, median, std, max and min values. Also, percentiles 10 and 90 are obtained for each metric.

This values are later used for the normalize the metrics, comparing three different methods of normalization:

- Full Range: *Max* and *Min* values on each metric are used as thresholds for normalize the values.

- Soft Statistics: The thresholds for normalize the metrics are obtained using the *mean* and *std* values.
- Percentiles: *Percentile-10* and *Percentile-90* are considered as minimum and maximum values for normalize the metrics. The grasps below or over this thresholds are considered as outliers on the metric.

4.2.3 Results

First, we analyze the shape and distribution of values for each metric. Results provided are: *Probability Density Functions*, and *Statistic* values. For the *PDF*s it is detailed in some cases a narrow sample of the data to better show the results. Also in some cases, there are peks of values, these corresponding to specific hand models. A specific analysis on the variability of the hands is performed in Chapter 6.

Table 4.3 summarize the statistic values obtained per each metric.

TABLE 4.3: Summary of statistic values and percentiles for each metric.

	Min	Max	Mean	STD	Median	10th Pctl	90th Pctl
Q_{A1}	0.0000	1.8773	0.4218	0.3747	0.3137	0.0398	0.9820
Q_{A2}	0.0000	735.5761	12.9225	24.4199	3.4270	0.0671	35.6456
Q_{A3}	0.0000	0.9997	0.2104	0.2131	0.1486	0.0267	0.4323
Q_{B1}	-0.1365	0.9987	0.5808	0.2248	0.6175	0.2764	0.8423
Q_{B2}	0.0000	0.9992	0.0953	0.0732	0.0770	0.0237	0.1907
Q_{B3}	0.0001	0.9988	0.4756	0.1715	0.4653	0.2648	0.6892
Q_{C1}	0.0000	0.2804	0.0282	0.0364	0.0136	0.0000	0.0803
Q_{C2}	0.0000	0.9988	0.0200	0.0271	0.0104	0.0013	0.0499
Q_{D1}	0.0000	1.0000	0.4631	0.2355	0.3925	0.2116	0.8636
Q_{D2}	0.0000	0.9954	0.0528	0.1381	0.0021	0.0000	0.1839

Q_{A1} - Smallest singular value of G

Figure 4.3 shows the histogram of metric Q_{A1} along with the *PDFs* that best fit this metric. Table 4.4 shows the values of the *PDFs*, along with the statistical values of the metric.

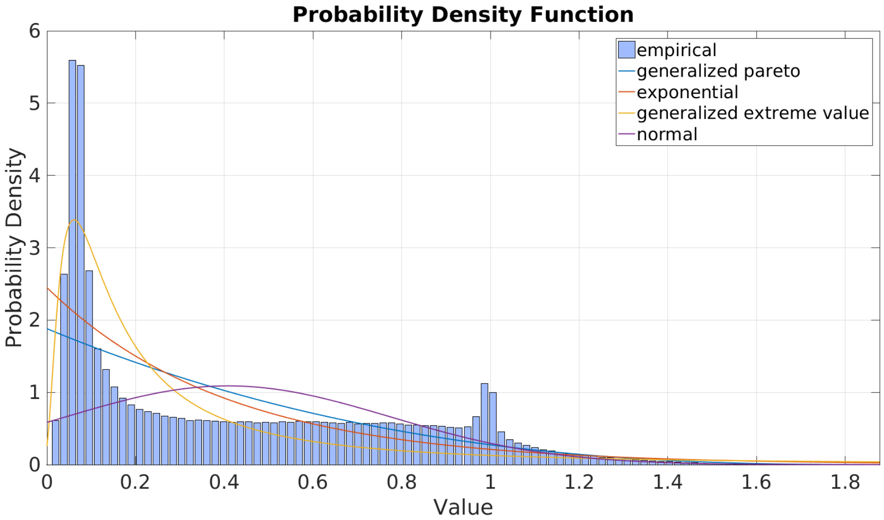


FIGURE 4.3: Histogram and Probability Density Function for metric Q_{A1}

TABLE 4.4: Parameters for best fits of distribution functions and statistical values on metric Q_{A1}

Distribution Name	k	sigma	theta	mu	nu
Generalized pareto	-0.2828	5.32E-01	-2.22E-15	NA	NA
Expponential	NA	NA	NA	0.4082	NA
Generalized Extreme Value	0.9455	0.1541	NA	0.1372	NA
Normal	NA	0.3747	NA	0.4218	NA

Statistics	Min	Max	Median	Mean	SD
Q_{A1}	0.0000	1.8773	0.3137	0.4218	0.3747

The results show the values on this metric are concentrated around two main points, 0.1 and 1, having an uniform distribution for the range of values: $[0.2, 1]$. This metric has also some values higher than 1, but these seem to be uncommon.

Q_{A2} - Volume of G in the wrench space

Figure 4.4 shows the histogram of metric Q_{A2} along with the *PDFs* that best fit this metric. Table 4.5 shows the values of the *PDFs*, along with the statistical values of the metric.

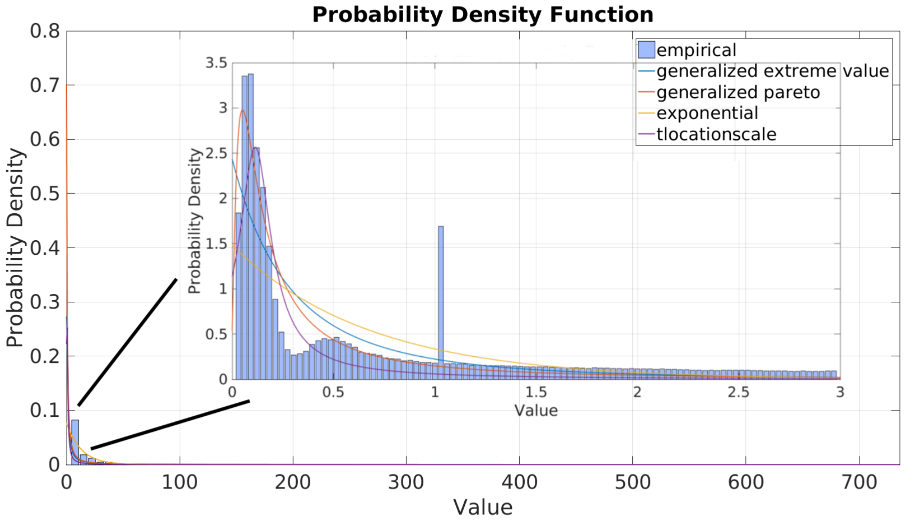


FIGURE 4.4: Histogram and Probability Density Function for metric Q_{A2}. Detailed graph for *PDFs* in range $[0, 3]$.

TABLE 4.5: Parameters for best fits of distribution functions and statistical values on metric Q_{A2}

Distribution Name	k	sigma	theta	mu	nu
Generalized Extreme Value	2.2156	1.589	NA	7.08E-01	NA
Generalized Pareto	1.8018	1.4243	-2.22E-15	NA	NA
Exponential	NA	NA	NA	12.9288	NA
T-Location Scale	NA	0.8355	NA	0.4203	0.4326

Statistics	Min	Max	Median	Mean	SD
Q_{A2}	0.0000	735.5761	3.4270	12.9225	24.4199

Although this metric has a wide range of values, from 0 to 700, half of them are concentrated in the range $[0, 3.427]$. As it shows the subfigure in 4.4 more than 50% of the grasps are concentrated in this narrow range. This graphs show better the *PDFs* on metric Q_{A2}.

Q_{A3} - Grasp Isotropy Index

Figure 4.5 shows the histogram of metric Q_{A3} along with the *PDFs* that best fit this metric. Table 4.6 shows the values of the *PDFs*, along with the statistical values of the metric.

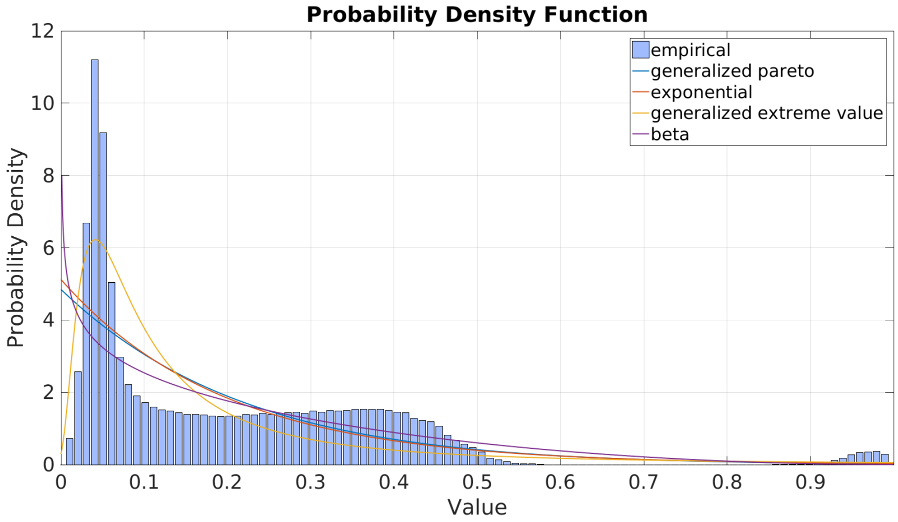


FIGURE 4.5: Histogram and Probability Density Function for metric Q_{A3}

TABLE 4.6: Parameters for best fits of distribution functions and statistical values on metric Q_{A3}

Distribution Name	k	sigma	theta	mu	a	b
Generalized Pareto	-0.0567	0.2062	-2.22E-15	NA	NA	NA
Exponential	NA	NA	NA	0.1952	NA	NA
Generalized Extreme Value	0.7812	0.0757	NA	0.0765	NA	NA
Beta	NA	NA	NA	NA	0.7936	2.8836

Statistics	Min	Max	Median	Mean	SD
Q_{A3}	0.0000	0.9997	0.1486	0.2104	0.2131

This metric has half of its values between 0 and 0.1486. Also it presents a gap between values 0.6 and 0.8. It has an uniform distribution of values of values in the range $[0.1, 0.5]$.

Q_{B1} - Distance between the centroid of the contact polygon and the center of mass of the object

Figure 4.6 shows the histogram of metric Q_{B1} along with the *PDFs* that best fit this metric. Table 4.7 shows the values of the *PDFs*, along with the statistical values of the metric.

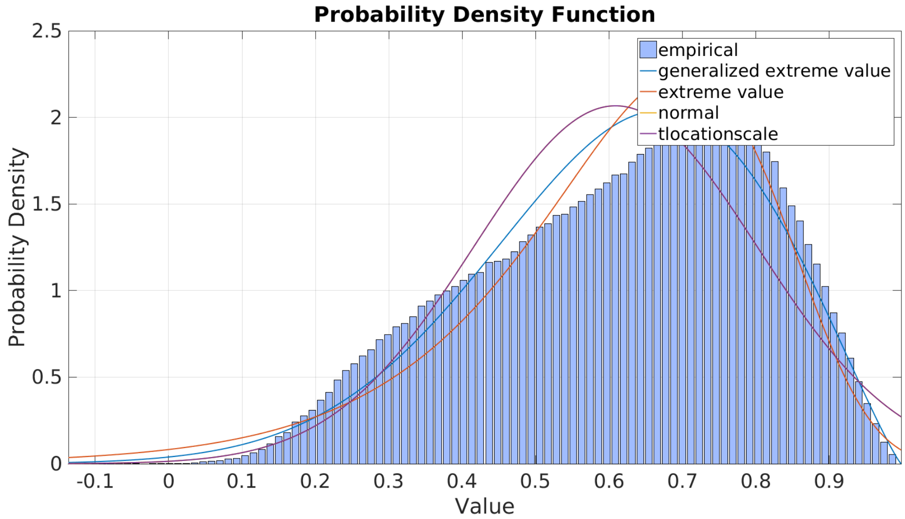


FIGURE 4.6: Histogram and Probability Density Function for metric Q_{B1}

TABLE 4.7: Parameters for best fits of distribution functions and statistical values on metric Q_{B1}

Distribution Name	k	sigma	theta	mu	nu
Generalized Extreme Value	-0.4679	0.2058	NA	0.5579	NA
Extreme Value	NA	0.1631	NA	0.7012	NA
Normal	NA	0.1930	NA	0.6088	NA
T-Location Scale	NA	0.1930	NA	0.6088	4.61E+06

Statistics	Min	Max	Median	Mean	SD
Q_{B1}	-0.1365	0.9987	0.6175	0.6088	0.1930

This metric shows a normal distribution of values centered in 0.80 for the range $[0.1, 1]$. Values under 0.1 are rarely shown for this metric.

Note: this metric has a minimum value below zero. Although theoretically its minimum value should be zero, the computation of the metric depends on the estimation of the max distance between the object's center of mass and its surface, which could be inaccurate in some objects.

Q_{B2} - Area of the grasp polygon

Figure 4.7 shows the histogram of metric Q_{B2} along with the *PDFs* that best fit this metric. Table 4.8 shows the values of the *PDFs*, along with the statistical values of the metric.

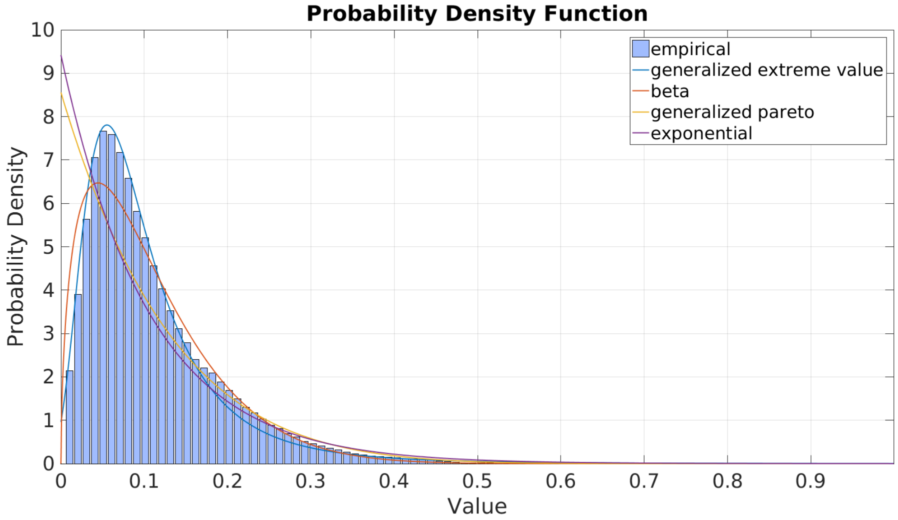


FIGURE 4.7: Histogram and Probability Density Function for metric Q_{B2}

TABLE 4.8: Parameters for best fits of distribution functions and statistical values on metric Q_{B2}

Distribution Name	k	sigma	theta	mu	a	b
Generalized Extreme Value	0.2363	0.0484	NA	6.52E-02	NA	NA
Beta	NA	NA	NA	NA	1.5629	13.099
Generalized Pareto	-0.1169	0.1169	-2.22E-15	NA	NA	NA
Exponential	NA	NA	NA	0.1062	NA	NA

Statistics	Min	Max	Median	Mean	SD
Q_{B2}	0.0000	0.9992	0.0770	0.0953	0.0732

The distribution of values on Q_{B2} shows a clear *PDF* of generalized extreme values. Most of its values are concentrated in a narrow lower range between 0 and 0.2, having a long tail until its maximum values, close to 1.

Q_{B3} - Shape of the grasp polygon

Figure 4.8 shows the histogram of metric Q_{B3} along with the *PDFs* that best fit this metric. Table 4.9 shows the values of the *PDFs*, along with the statistical values of the metric.

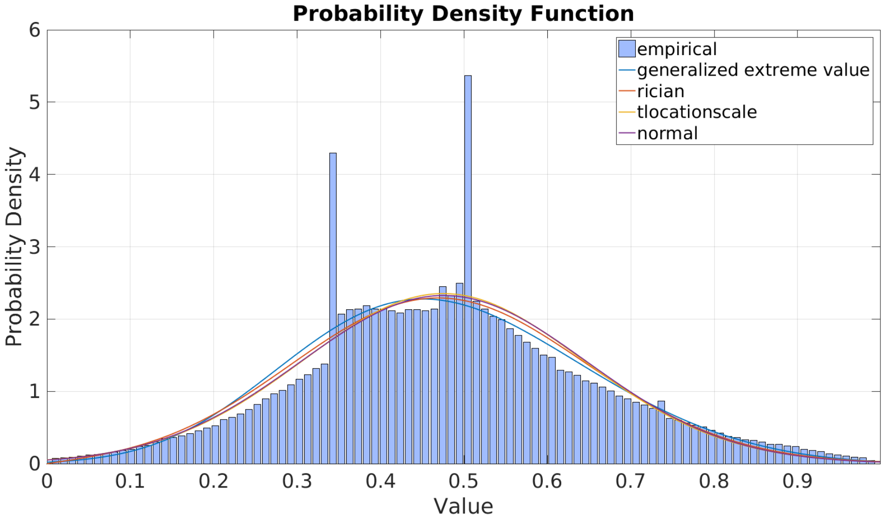


FIGURE 4.8: Histogram and Probability Density Function for metric Q_{B3}

TABLE 4.9: Parameters for best fits of distribution functions on metric Q_{B3}

Distribution Name	k	sigma	theta	mu	s	nu
Generalized Extreme Value	-0.2204	0.1658	NA	0.4106	NA	NA
Rician	NA	0.1816	NA	NA	0.4355	NA
T-Location Scale	NA	0.1690	NA	0.4751	NA	67.4197
Normal	NA	0.1715	NA	0.4756	NA	NA

Statistics	Min	Max	Median	Mean	SD
Q_{B3}	0.0001	0.9988	0.4653	0.4756	0.1715

This metric presents a normal distribution along its full range $[0, 1]$ with two peaks on values 0.35 and 0.5.

Q_{C1} - Smallest maximum wrench to be resisted

Figure 4.9 shows the histogram of metric Q_{C1} along with the *PDFs* that best fit this metric. Table 4.10 shows the values of the *PDFs*, along with the statistical values of the metric

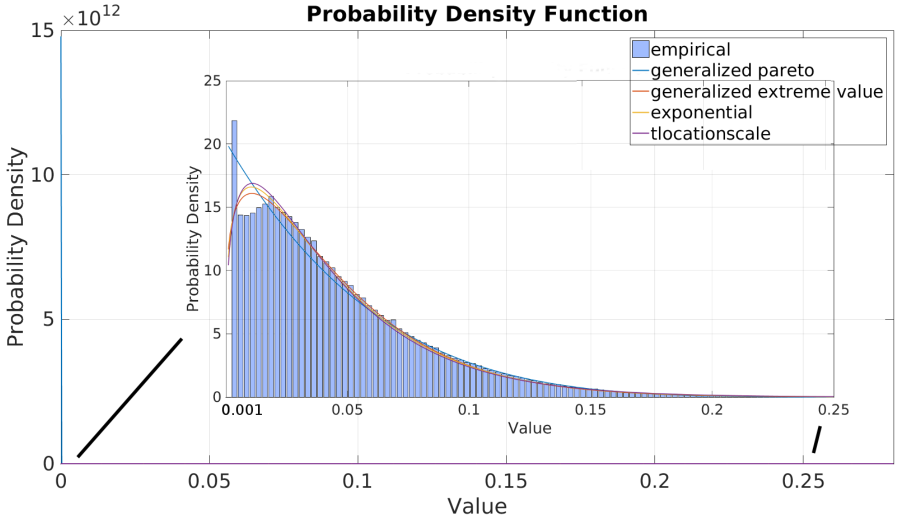


FIGURE 4.9: Histogram and Probability Density Function for metric Q_{C1}

TABLE 4.10: Parameters for best fits of distribution functions on metric Q_{A1}

Distribution Name	k	sigma	theta	mu	nu
Generalized Pareto	23.7244	9.88E-15	-2.22E-15	NA	NA
Generalized Extreme Value	3.6031	1.02E-04	NA	2.80E-05	NA
Exponential	NA	NA	NA	2.86E-02	NA
T-Location Scale	NA	0.0227	NA	0.0176	2.7422

Statistics	Min	Max	Median	Mean	SD
Q_{C1}	0.0000	0.2804	0.0136	0.0282	0.0364

This metric shows a short range of values, from 0 to 0.25, having most of them concentrated around 0.02.

Q_{C2} - Volume of the convex hull

Figure 4.10 shows the histogram of metric Q_{C2} along with the *PDFs* that best fit this metric. Table 4.11 shows the values of the *PDFs*, along with the statistical values of the metric.

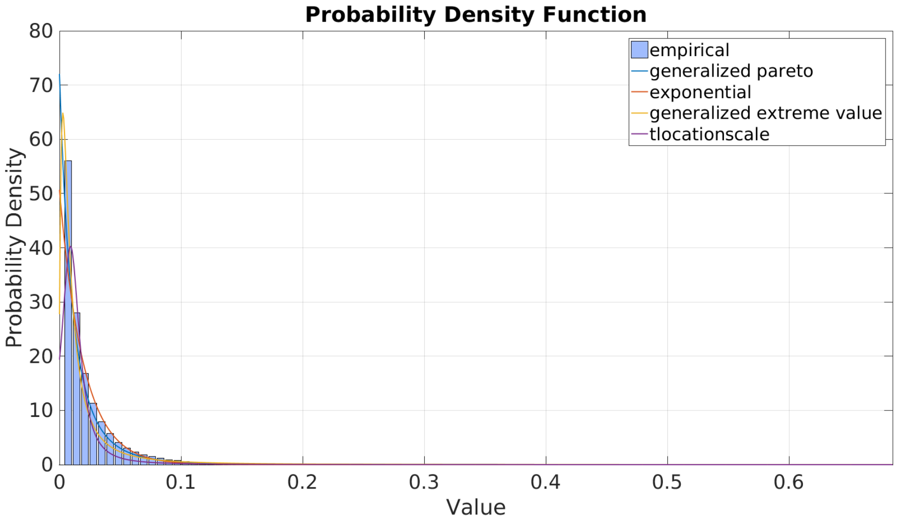


FIGURE 4.10: Histogram and Probability Density Function for metric Q_{C2}

TABLE 4.11: Parameters for best fits of distribution functions on metric Q_{C2}

Distribution Name	k	sigma	theta	mu	nu
Generalized Pareto	0.3093	0.0139	-2.22E-15	NA	NA
Exponential	NA	NA	NA	0.0197	NA
Generalized Extreme Value	0.7938	0.0073	NA	0.0062	NA
T-Location Scale	NA	0.0084	NA	0.0091	1.3899

Statistics	Min	Max	Median	Mean	SD
Q_{C2}	0.0000	0.9988	0.0104	0.0200	0.0271

Although this metric has its values along its theoretical range $[0, 1]$, most of them are concentrated in a narrow range from 0 to 0.1.

Q_{D1} - Posture of manipulator joints

Figure 4.11 shows the histogram of metric Q_{D1} along with the *PDFs* that best fit this metric. Table 4.12 shows the values of this *PDFs*, along with the statistical values of the metric.

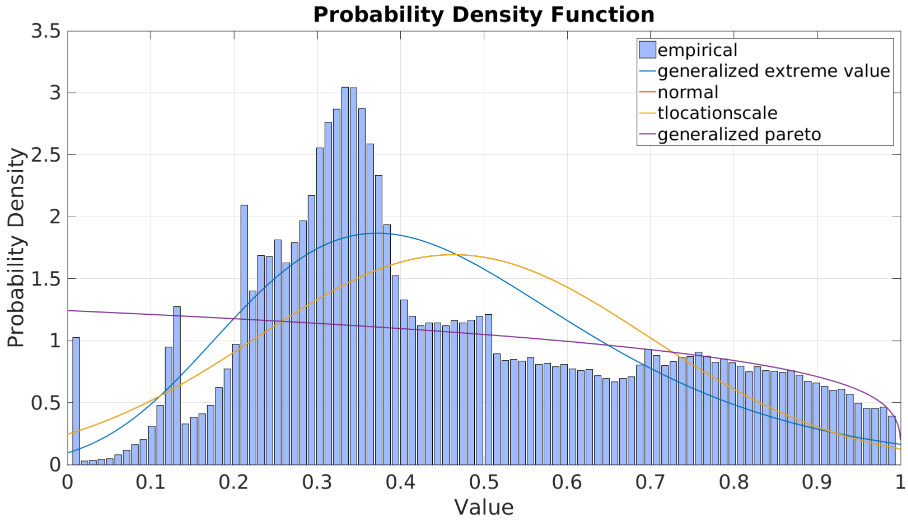


FIGURE 4.11: Histogram and Probability Density Function for metric Q_{D1}

TABLE 4.12: Parameters for best fits of distribution functions on metric Q_{D1}

Distribution Name	k	sigma	theta	mu	nu
Generalized Extreme Value	-0.0632	0.1974	NA	0.3587	NA
Normal	NA	0.2355	NA	0.4631	NA
T-Location Scale	NA	0.2355	NA	0.4631	6.35E+06
Generalized Pareto	-0.8040	0.8044	-2.22E-15	NA	NA

Statistics	Min	Max	Median	Mean	SD
Q_{D1}	0.0000	1.0000	0.3925	0.4631	0.2355

This metric doesn't present a clear distribution of values along its range. Thus, it's difficult to classify it with the classic parametric distribution functions.

Q_{D2} - Inverse of the condition number of G_J

Figure 4.12 shows the histogram of metric Q_{D2} along with the *PDFs* that best fit this metric. Table 4.13 shows the values of this *PDFs*, along with the statistical values of the metric.

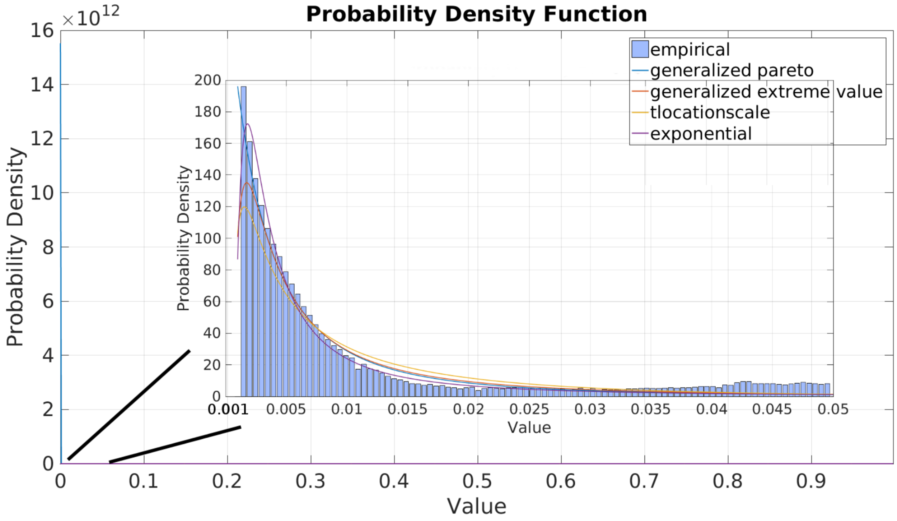


FIGURE 4.12: Histogram and Probability Density Function for metric Q_{D2}

TABLE 4.13: Parameters for best fits of distribution functions on metric Q_{D2}

Distribution Name	k	sigma	theta	mu	nu
Generalized Pareto	22.6694	8.93E-15	-2.22E-15	NA	NA
Generalized Extreme Value	5.1432	1.08E-04	NA	2.10E-05	NA
T-Location Scale	NA	3.34E-06	NA	3.79E-08	0.1458
Exponential	NA	NA	NA	0.0541	NA

Statistics	Min	Max	Median	Mean	SD
Q_{D2}	0.0000	0.9954	0.0021	0.0528	0.1381

As it is with metric Q_{C1}, metric Q_{D2} has most of its values around 0. However it has values along its theoretical range from 0 to 1.

Normalization

The information required to obtain all different thresholds proposed are detailed in Table 4.14. Results corroborate that the *Full Range* thresholds are much higher than the *Soft Statistics* and *Percentiles* thresholds, because of the high amount of atypical values, especially for metric Q_{A2}. Also, the high standard deviations of metrics Q_{A2}, Q_{A3}, Q_{C1}, Q_{C2}, and Q_{D2} result in negative minimum thresholds for the *Soft Statistics* thresholds.

TABLE 4.14: Statistic thresholds of the grasp quality values for each metric

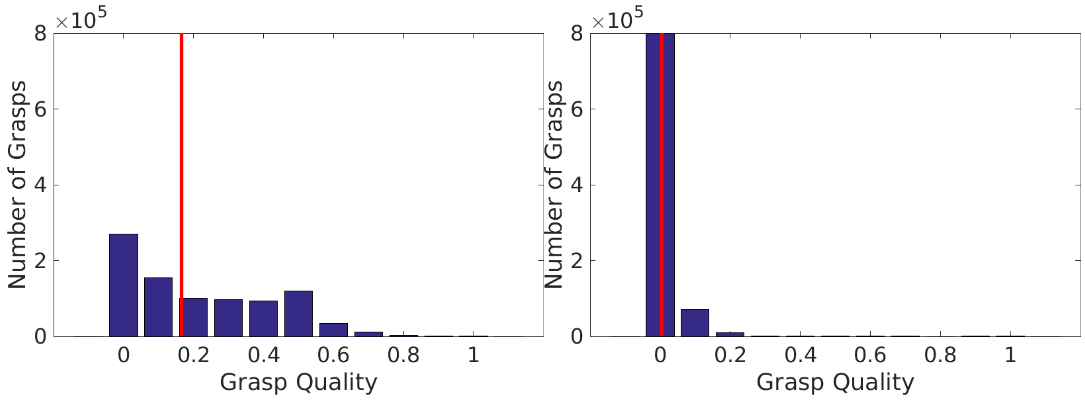
	Full Range		Soft Statistics		Percentiles		Nº of
	Min	Max	Mean-SD	Mean+SD	10th Pctl	90th Pctl	Grasps
QA1	0.0000	1.8773	0.0471	0.7965	0.0398	0.9820	882000
QA2	0.0000	735.5761	-11.4975	37.3424	0.0671	35.6456	882000
QA3	0.0000	0.9997	-0.0026	0.4235	0.0267	0.4323	882000
QB1	-0.1365	0.9987	0.4158	0.8018	0.2764	0.8423	882000
QB2	0.0000	0.9992	0.0220	0.1685	0.0237	0.1907	634446
QB3	0.0001	0.9988	0.3041	0.6471	0.2648	0.6892	634446
QC1	0.0000	0.2804	-0.0082	0.0647	0.0000	0.0803	634446
QC2	0.0000	0.9988	-0.0071	0.0471	0.0013	0.0499	634446
QD1	0.0000	1.0000	0.2276	0.6986	0.2116	0.8636	882000
QD2	0.0000	0.9954	-0.0853	0.1910	0.0000	0.1839	882000

Histograms on Figure 4.13 shows the distribution of each of the grasp quality metrics using the three different normalization methods. As is shown, the *Percentile* normalization shows an even distribution of the values for each metric. It can be seen that most of the metrics are distributed along the [0,1] range, with the notorious exception of metric Q_{D2}, which presents 50% of the grasp quality values within the range [0, 0.1]. Also note Q_{C2} and Q_{D1} have a lot of values under 0, due to their *Percentile 10* coincide with *Percentile 30*.

Finally, the values of metrics, normalized with the *Percentiles* method, are classified into grades according to the 50th percentile. Ranges for each grade are presented in Table 4.15.

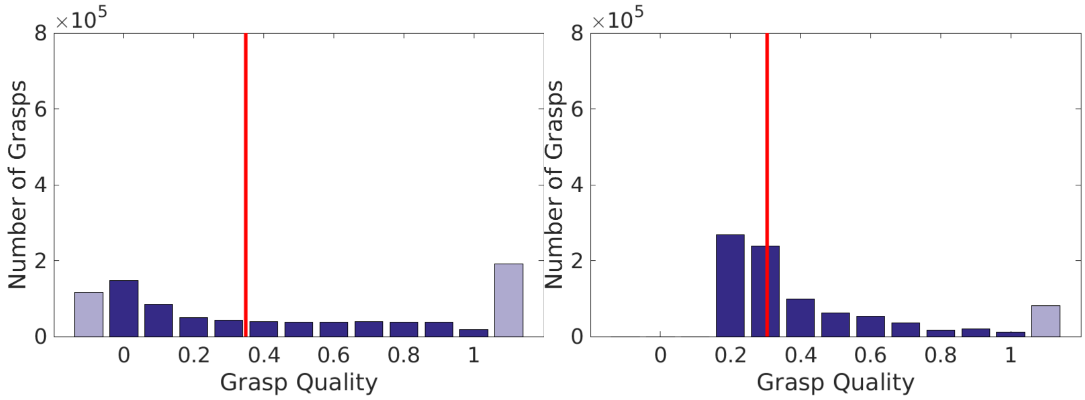
TABLE 4.15: Ranges to classify the quality of a grasp for each metric

	Bad Quality	Fair Quality	Good Quality	Great Quality
Q_{A1}	[-∞, 0]	[0, 0.2887]	[0.2887, 1]	[1, ∞]
Q_{A2}		[0, 0.0944]	[0.0944, 1]	
Q_{A3}		[0, 0.3018]	[0.3018, 1]	
Q_{B1}		[0, 0.6103]	[0.6103, 1]	
Q_{B2}		[0, 0.3092]	[0.3092, 1]	
Q_{B3}		[0, 0.4679]	[0.4679, 1]	
Q_{C1}		[0, 0.1684]	[0.1684, 1]	
Q_{C2}		[0, 0.1886]	[0.1886, 1]	
Q_{D1}		[0, 0.3015]	[0.3015, 1]	
Q_{D2}		[0, 0.0147]	[0.0147, 1]	



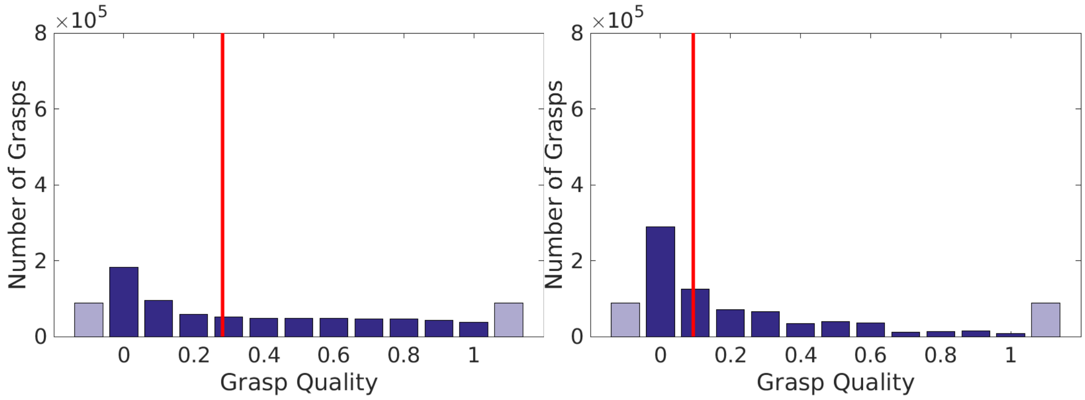
(1) Full Range normalization for metric Q_{A1}

(2) Full Range normalization for metric Q_{A2}



(3) Soft Statistics normalization for metric Q_{A1}

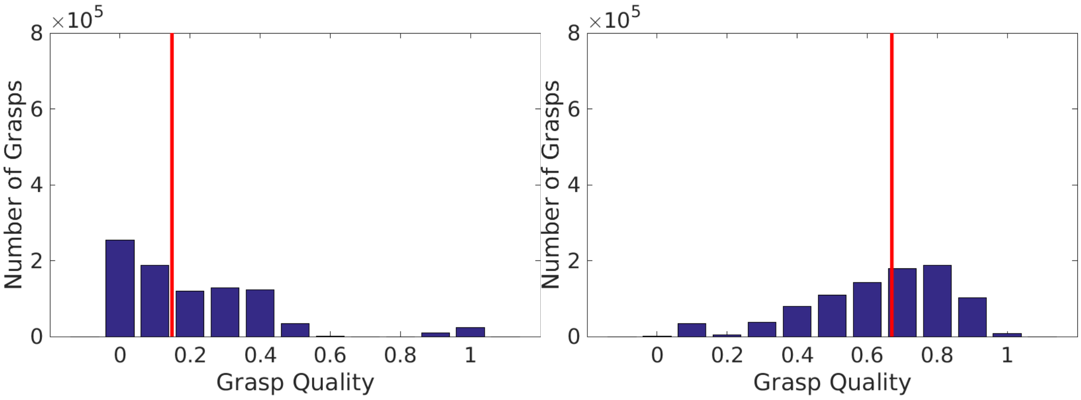
(4) Soft Statistics normalization for metric Q_{A2}



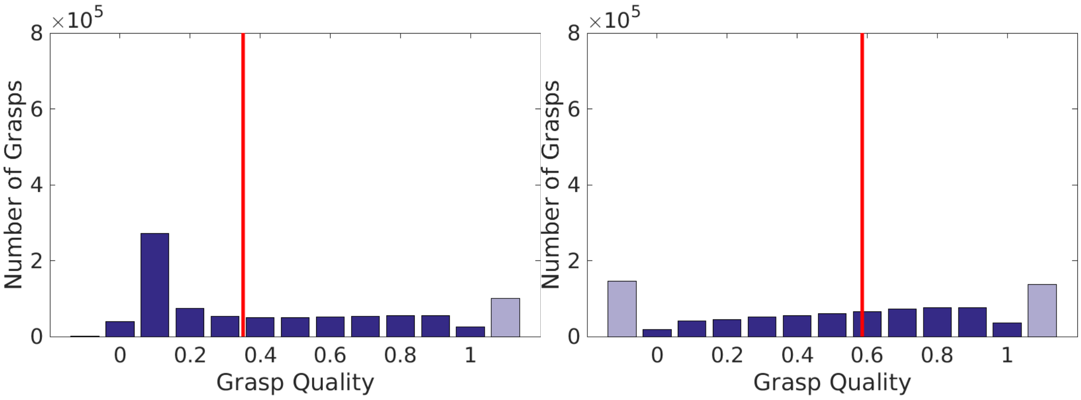
(5) Percentile normalization for metric Q_{A1}

(6) Percentile normalization for metric Q_{A2}

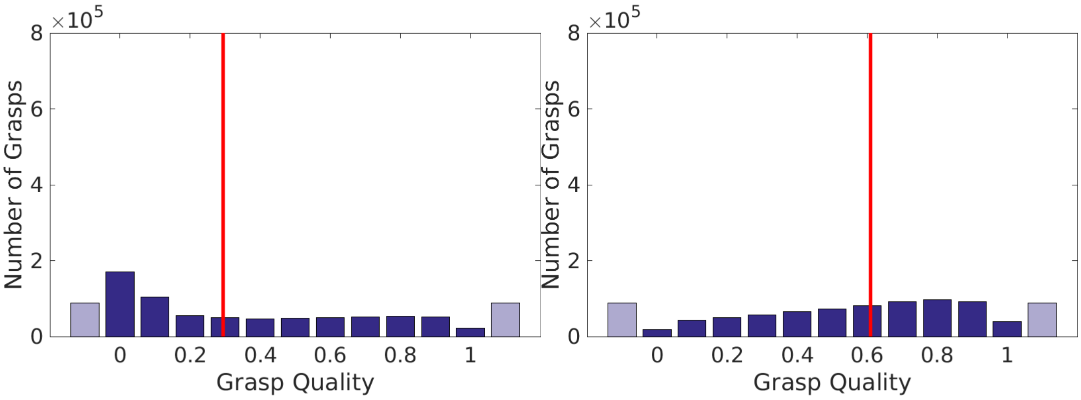
FIGURE 4.13: Histograms showing the comparison of different normalization methods for each metric. Red line points out the *Percentile 50*.



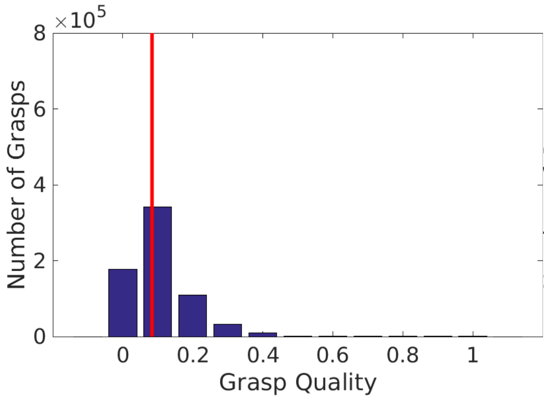
(7) Full Range normalization for metric Q_{B1} (8) Full Range normalization for metric Q_{B2}



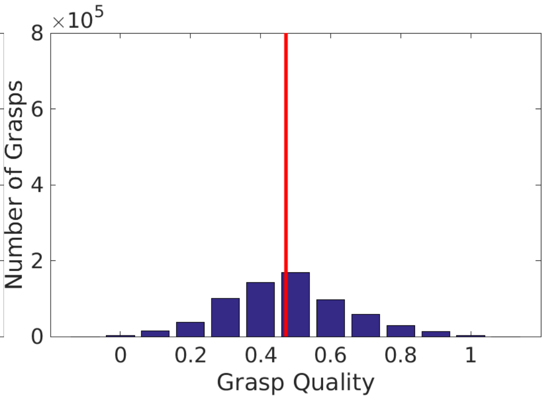
(9) Soft Statistics normalization for metric Q_{B1} (10) Soft Statistics normalization for metric Q_{B2}



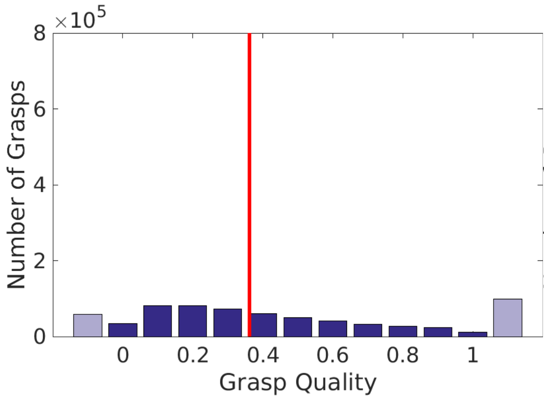
(11) Percentile normalization for metric Q_{B1} (12) Percentile normalization for metric Q_{B2}



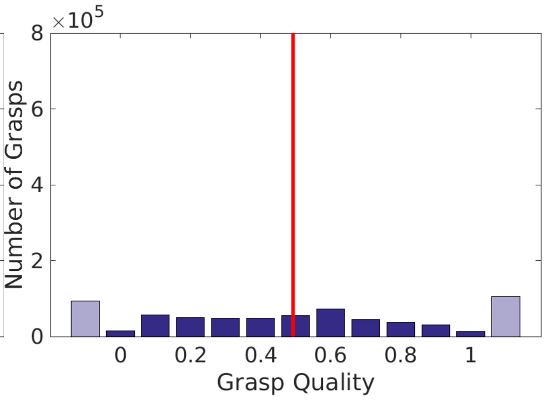
(13) Full Range normalization for metric Q_{B2}



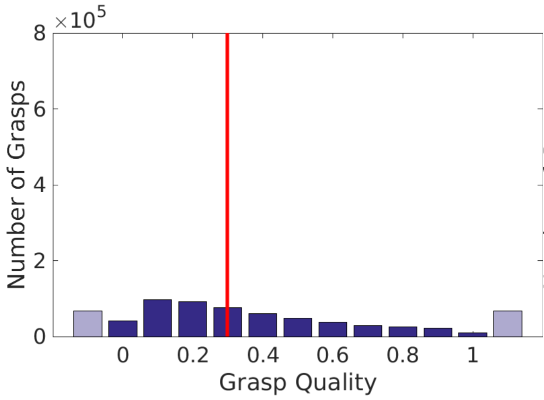
(14) Full Range normalization for metric Q_{B3}



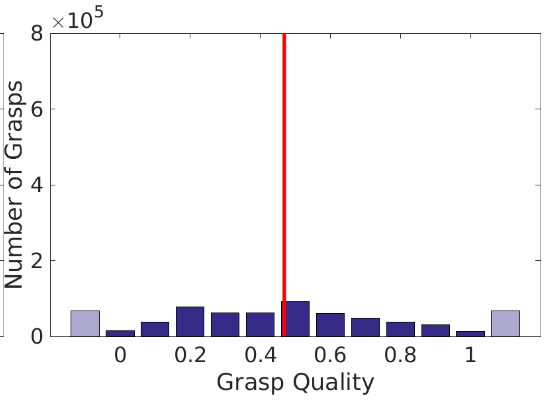
(15) Soft Statistics normalization for metric Q_{B2}



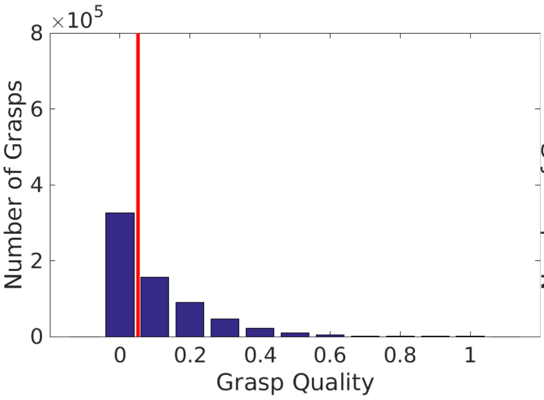
(16) Soft Statistics normalization for metric Q_{B3}



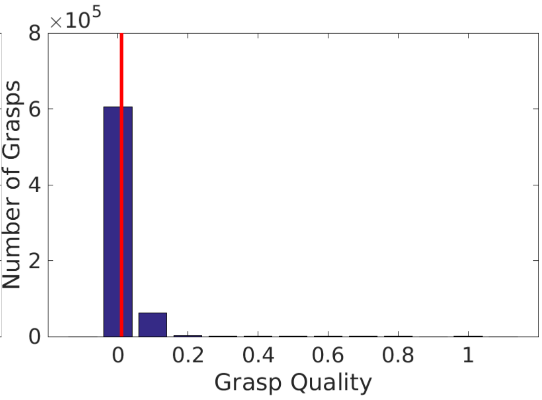
(17) Percentile normalization for metric Q_{B2}



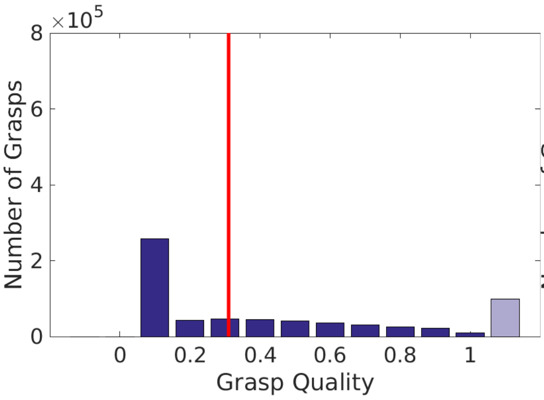
(18) Percentile normalization for metric Q_{B3}



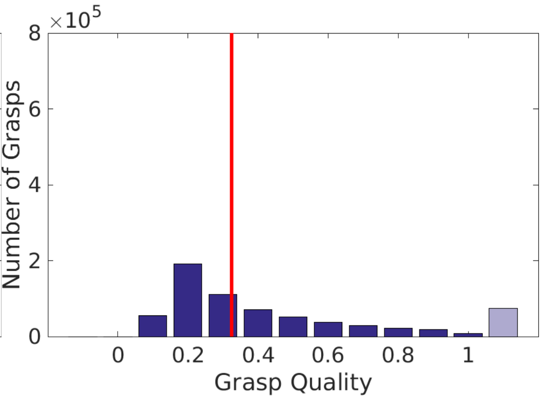
(19) Full Range normalization for metric Q_{C1}



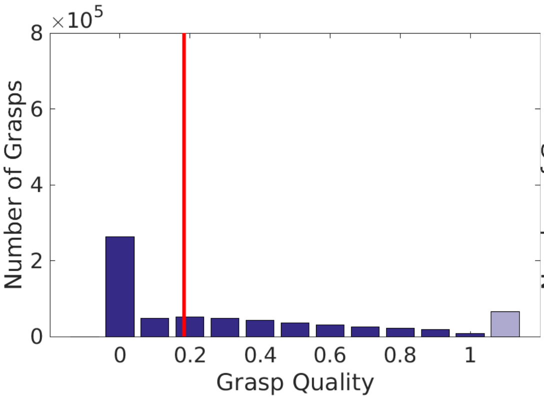
(20) Full Range normalization for metric Q_{C2}



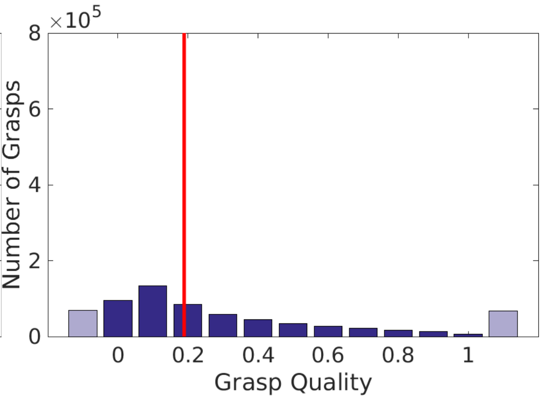
(21) Soft Statistics normalization for metric Q_{C1}



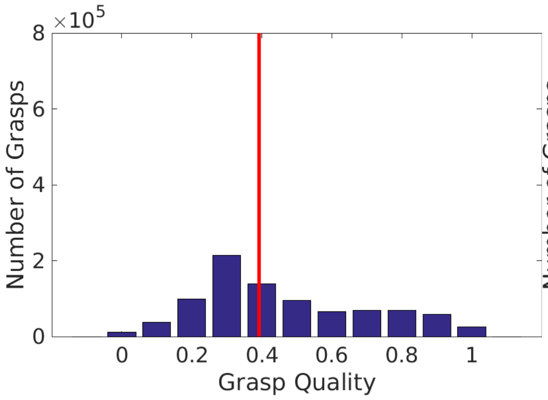
(22) Soft Statistics normalization for metric Q_{C2}



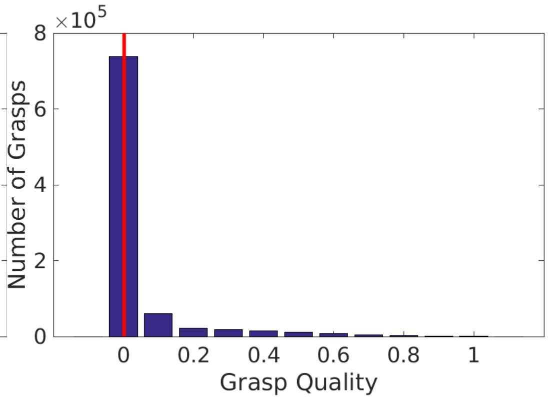
(23) Percentile normalization for metric Q_{C1}



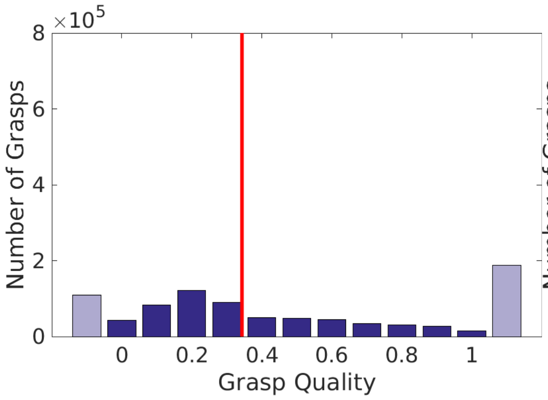
(24) Percentile normalization for metric Q_{C2}



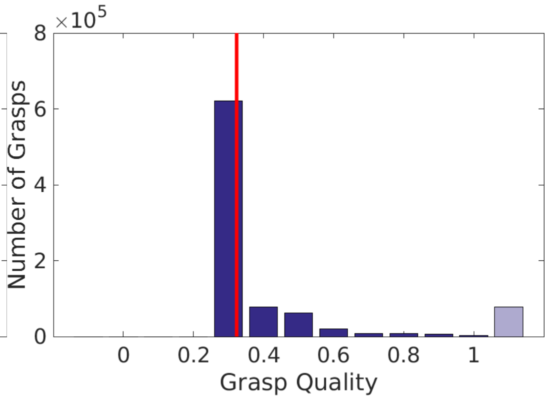
(25) Full Range normalization for metric Q_{D1}



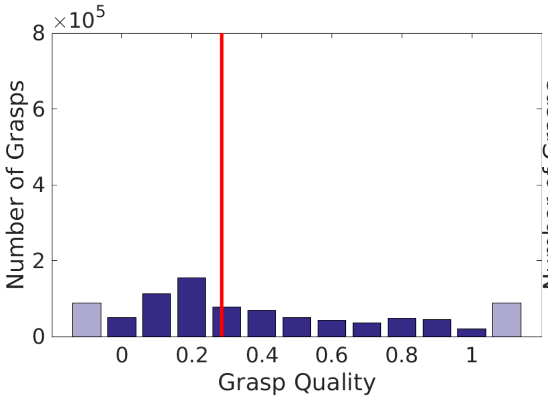
(26) Full Range normalization for metric Q_{D2}



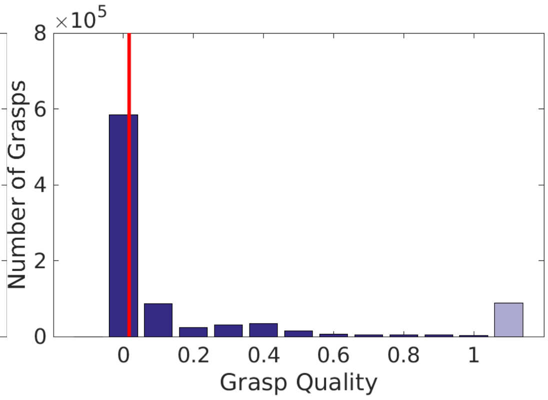
(27) Soft Statistics normalization for metric Q_{D1}



(28) Soft Statistics normalization for metric Q_{D2}



(29) Percentile normalization for metric Q_{D1}



(30) Percentile normalization for metric Q_{AD2}

4.3 Sensitivity

A sensitivity analysis is conducted to assess the robustness of each quality metric with respect to the uncertainty in the object's pose. The effect of inaccuracies in the positioning of the real robot hand on the quality of the grasp is investigated. This problem may arise when the final execution of a selected grasp does not succeed in placing the contacts in the desired grasp locations.

4.3.1 Methods

An initial set of grasp configurations are selected as reference. Several modifications of the hand position are computed for each reference grasp configuration by introducing a random translation and rotation on each of the hand position axes (see Fig. 4.10).

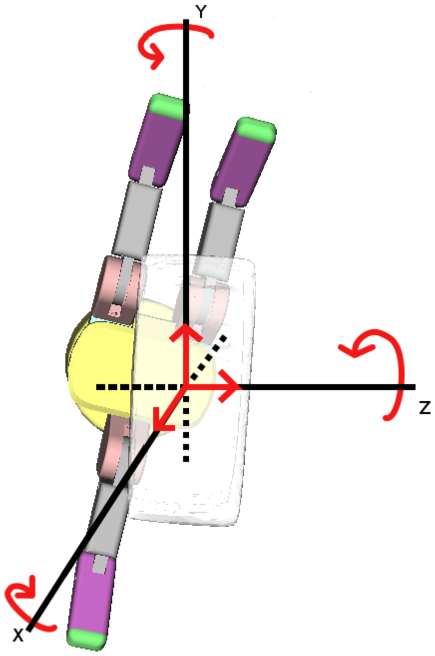


FIGURE 4.10: Variations in the hand posture along their axes.

The measurement variation of the metric values for a given object is computed using the Sensitivity Index (SI), which is defined as the mean value of the standard deviations with respect to the metric calculated for the reference grasp:

$$SI = \frac{1}{n}\sum_{x=1}^{n}\sigma_x$$

where n is the number of grasps for each object and σ_x the standard deviation calculated as:

$$\sigma_x = \frac{1}{n_v}\sqrt{\sum_{i=1}^{n_v}(x_i - x_0)^2}$$

where n_v is the number of variations of the reference grasp, x_0 the value of the metric of the reference grasp and x_i the value of the metric calculated for each variation. Finally, a Global Sensitivity Index (GSI) has been calculated for each metric, as the mean value of the SI per object previously calculated:

$$GSI = \frac{1}{n_o}\sum_{i=1}^{n}(SI_i)$$

where n_o is the number of objects and SI_i is the Sensitivity Index calculated for each object.

GSI is expressed in percentage (GSI_N) by using the metric ranges obtained in the variability analysis (see Table 4.3):

$$GSI_N = \frac{GSI}{max - min}100\%$$

where min y max are the proposed minimum and maximum thresholds, respectively.

A preliminary study is carried out to determine the number of variations required for the analysis. Two hands (Barrett and Schunk SDH and 10 objects are chosen for this study. For every pair hand-object, 10 reference grasps are selected randomly from those generated in the experiments described in Section **??**. Grasp variations are generated by introducing random translations and rotations on each of the hand position axes obtained from Gaussian distributions with standard deviation of 1 cm and 0.1 rad for translations and rotations, respectively (these Gaussian distributions provide 68% of the deviations inside the bell curve). Set sizes from 10 to 300 variations are investigated (with increments of 10 variations until set size of 50 variations, and with increments of 50 variations afterwards). Grasp variations not meeting the force-closure condition were discarded, as well as variations resulting in an initial collision between the hand and the object in order to avoid unreachable grasps.

The metrics' values are computed for all variations, and the GSI_N is calculated for each metric, using the 10%-90% percentile thresholds determined in Section **??**. The number of variations from which the GSI_N gets stabilized is used as the set size N_{var} to be considered in the full analysis. The full analysis considers 126 objects and 7 hand models. A set of 20 reference grasps is randomly selected for every pair hand-object from those generated in the experiments described in Section **??**.

For each reference grasp, N_{var} grasp variations are generated analogously as in the preliminary study. GSI_N values are computed using the thresholds defined by the *Percentiles* method. For each metric, a distinction between *good* and *bad* reference grasps is considered as those for which the metric value is over or below the 50^{th} percentile threshold (thresholds taken from the Table 4.15).

Several analysis are conducted regarding the different distribution of variations that improve or worsen the value of the metric from their respective reference grasp. First a comparison on the GSI_N for each metric grouping by *good* and *bad* quality grasps and differentiating between variations which improve or worsen the quality. A similar study is conducted analyzing the number of grasps which accomplish these characteristics. Finally an study on the influence of each hand model to the sensitivity of the metrics is conducted. Results are reported individually per each metric.

4.3.2 Results

Preliminary study

Figure 4.11 shows the GSI_N values calculated in the preliminary study for each metric versus the different set sizes considered. GSI_N gets stable for sets equal or larger than 50 variations. Thus, the number of variations N_{var} to be used for the full analysis was set to 50. This preliminary study also revealed that only 20-23% of the variations randomly generated meet the necessary force-closure condition (see Tab. 4.16).

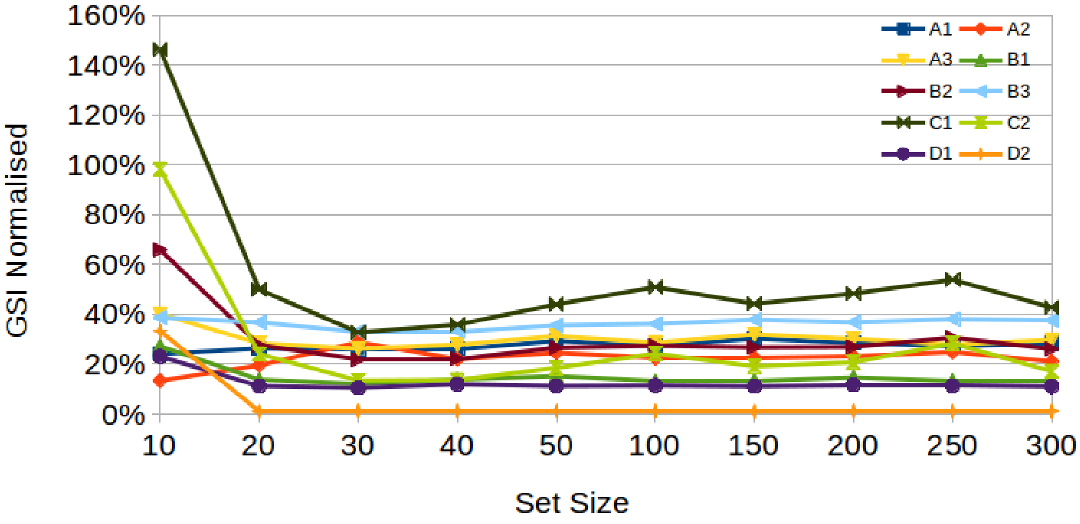


FIGURE 4.11: Graph comparing the Global Sensitivity Index between different set sizes for each quality metric.

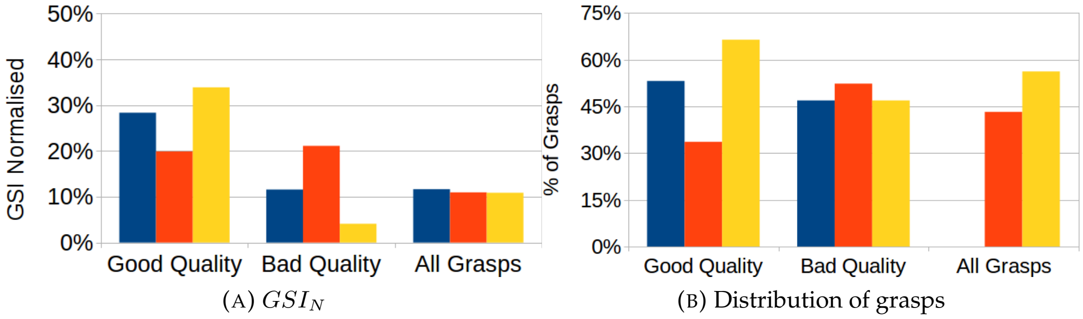
TABLE 4.16: Percentage of grasps accomplishing force-closure

Set Size	10	20	30	40	50	100	150	200	250	300
Valid Grasps (%)	22.8	21.2	22.8	20.1	20.2	22.4	22.5	21.4	21.6	22.0

The full analysis considers, then, a subset of 50 variations for every reference grasp, which makes a total amount of 1020 grasps (references plus variations) per object, 882000 total variations, and 128520 different grasps for each hand. The force-closure analysis performed to discard unstable grasps provides, in average, only 20.06% of feasible grasps (176,900 from the initial set of 882,000 variations).

Q_{A1} - Smallest singular value of G

Graphs on Figure 4.12a show the GSI_N for metric Q_{A1}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.12b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N

(B) Distribution of grasps

FIGURE 4.12: GSI_N for metric Q_{A1}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.13 shows the comparison of GSI_N between different hand models.

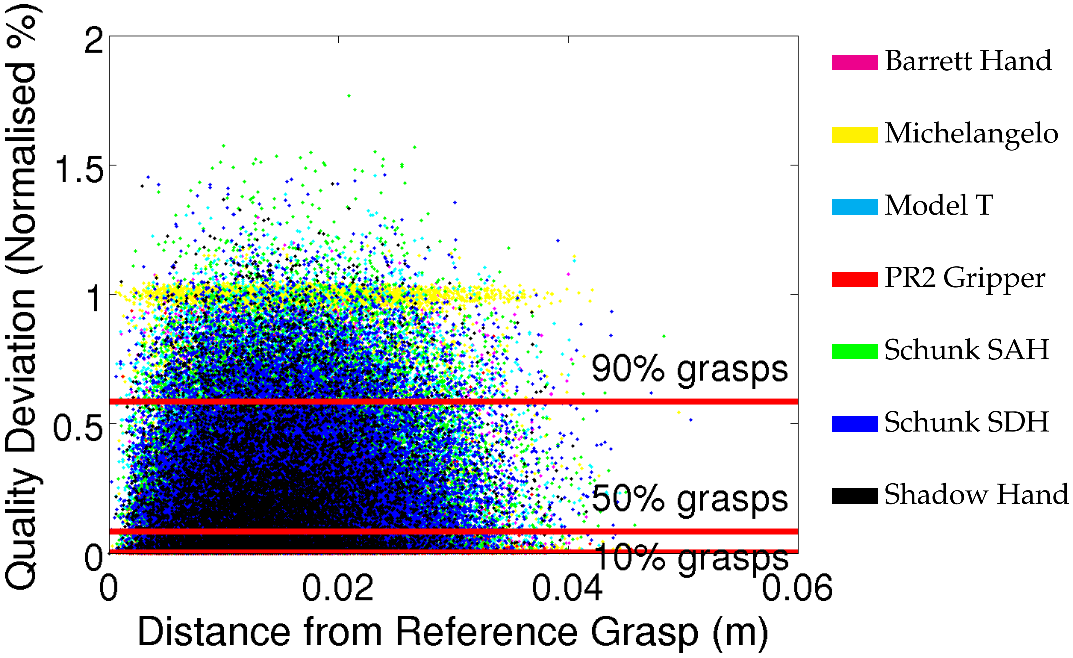
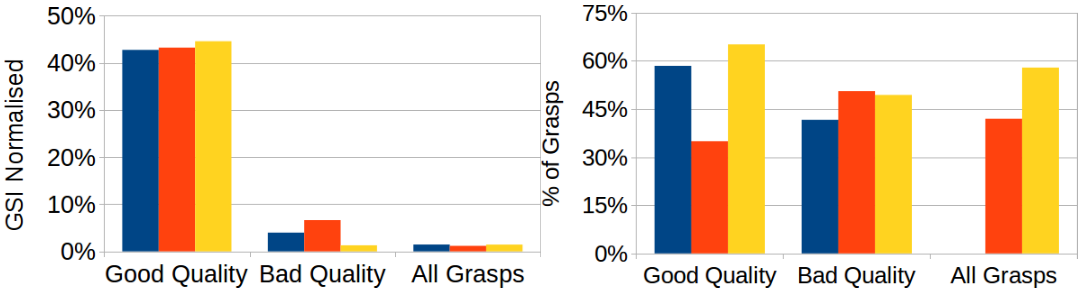


FIGURE 4.13: Variation in the metric Q_{A1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{A2} - Volume of G in the wrench space

Graphs on Figure 4.14a show the GSI_N for metric Q_{A2}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.14b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N (B) Distribution of grasps

FIGURE 4.14: GSI_N for metric Q_{A2}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.15 shows the comparison of GSI_N between different hand models.

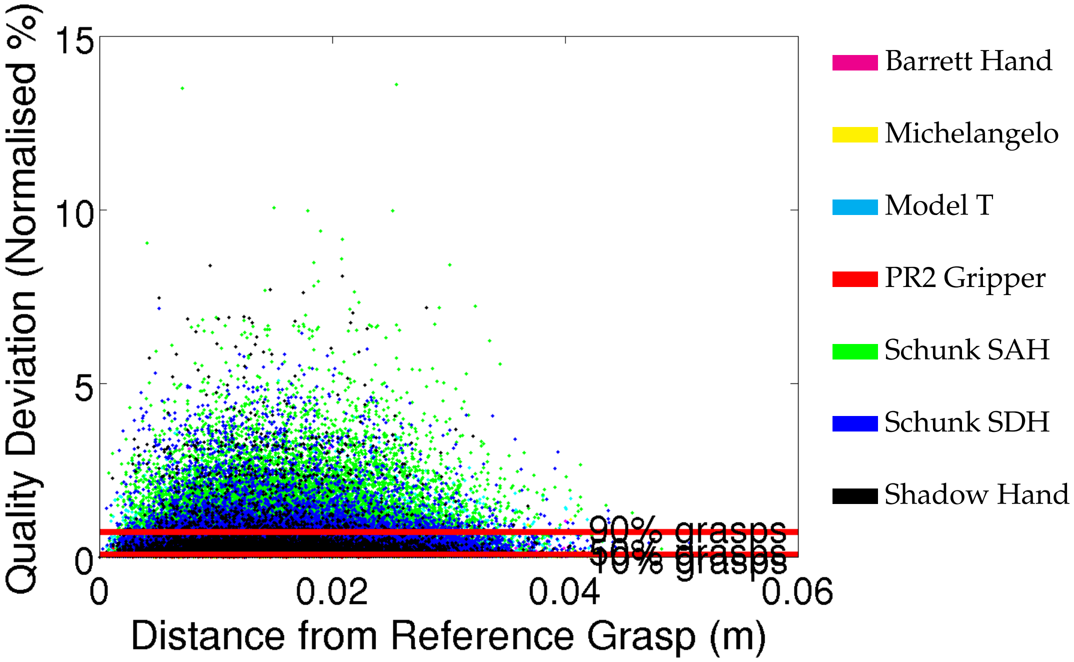
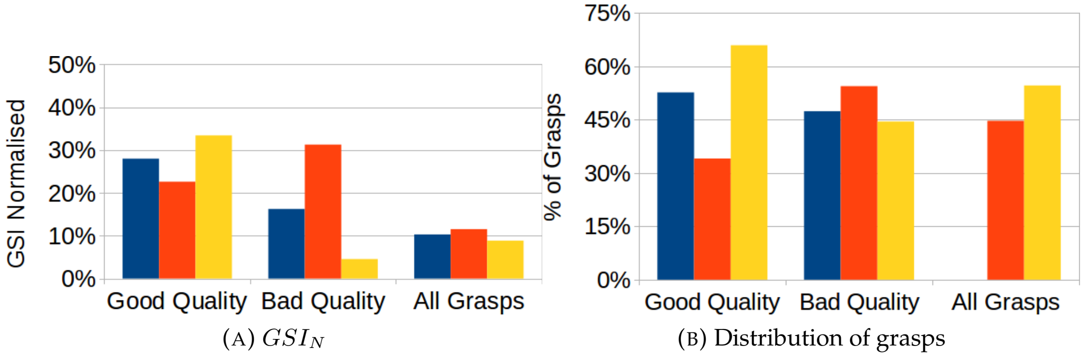


FIGURE 4.15: Variation in the metric Q_{A2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{A3} - Grasp Isotropy Index

Graphs on Figure 4.16a show the GSI_N for metric Q_{A3}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.16b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N (B) Distribution of grasps

FIGURE 4.16: GSI_N for metric Q_{A3}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.17 shows the comparison of GSI_N between different hand models.

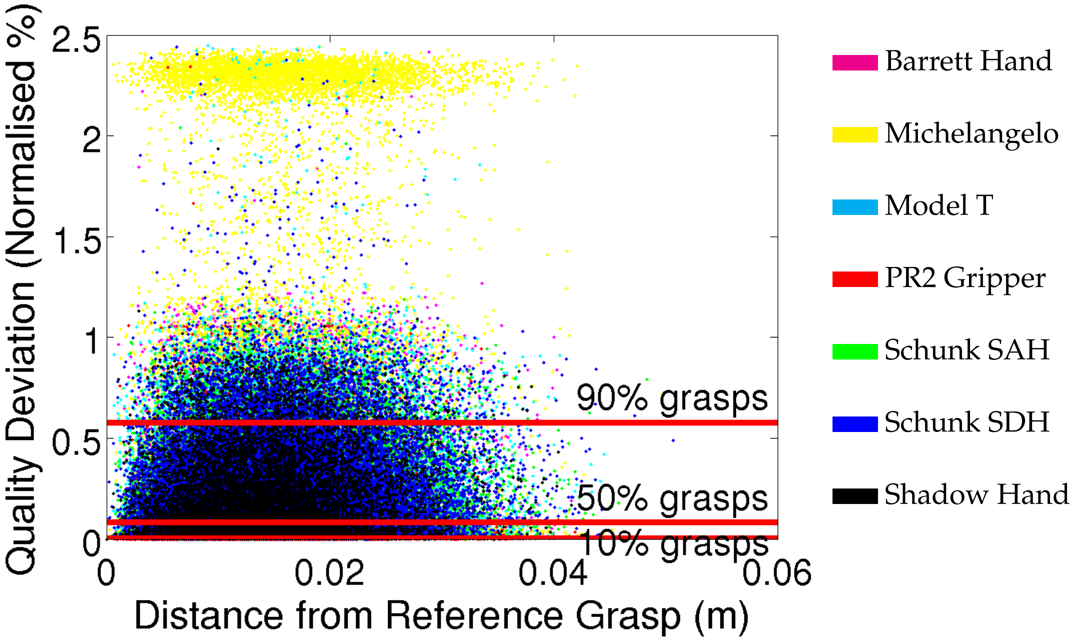
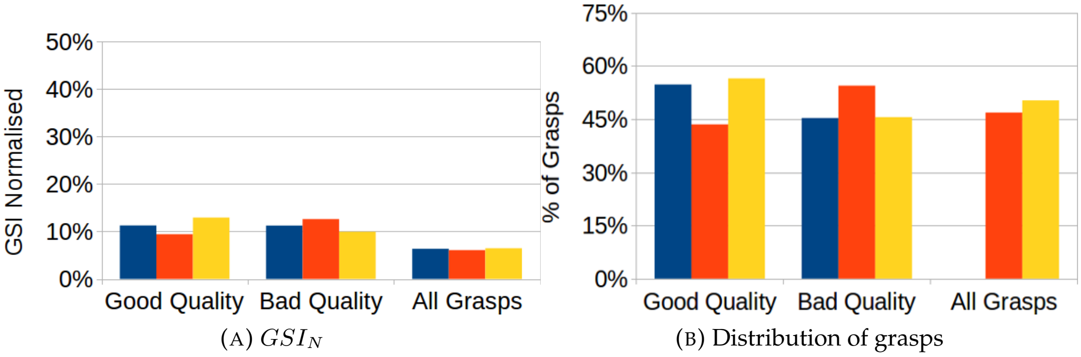


FIGURE 4.17: Variation in the metric Q_{A3} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q$_{B1}$ - Distance between the centroid of the contact polygon and the center of mass of the object

Graphs on Figure 4.18a show the GSI_N for metric Q_{B1}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.18b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N (B) Distribution of grasps

FIGURE 4.18: GSI_N for metric Q_{B1}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.19 shows the comparison of GSI_N between different hand models.

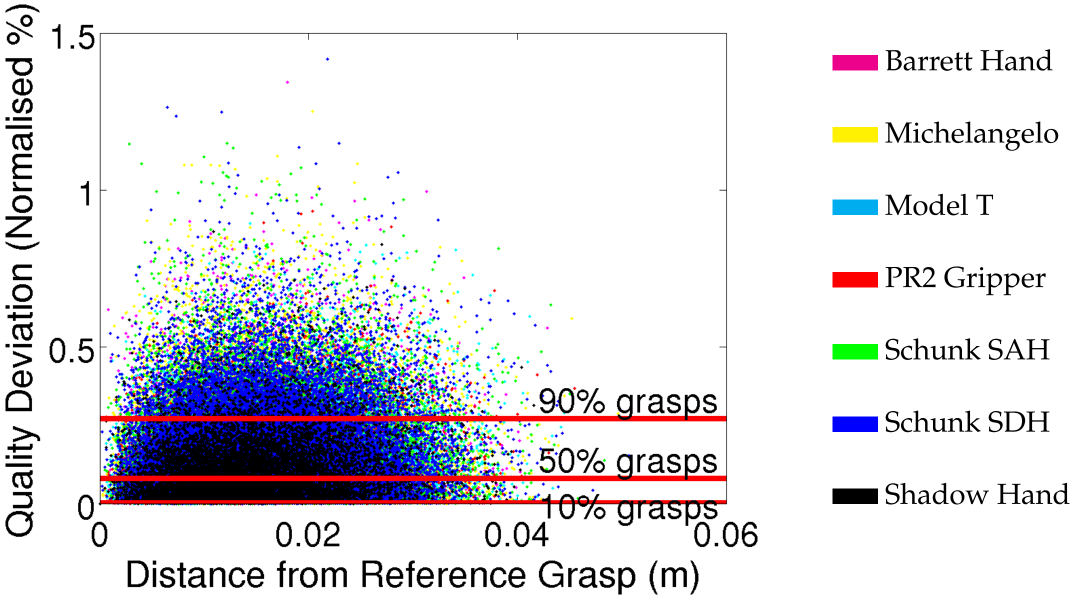
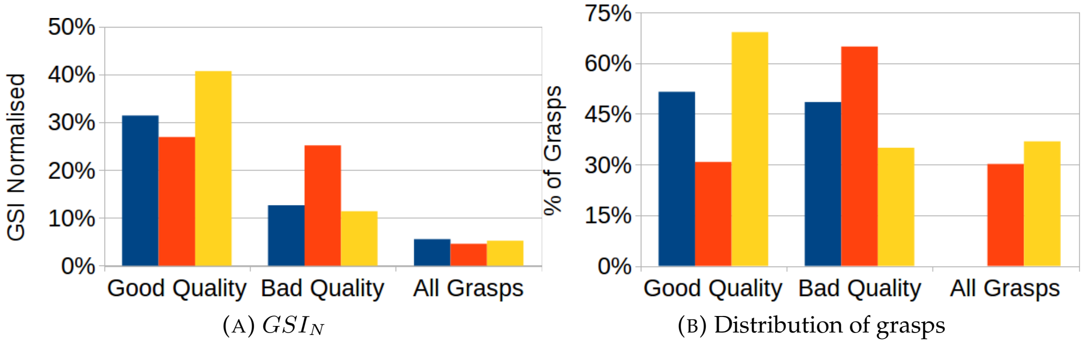


FIGURE 4.19: Variation in the metric Q_{B1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{B2} - Area of the grasp polygon

Graphs on Figure 4.20a show the GSI_N for metric Q_{B2}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.20b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N

(B) Distribution of grasps

FIGURE 4.20: GSI_N for metric Q_{B2}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.21 shows the comparison of GSI_N between different hand models.

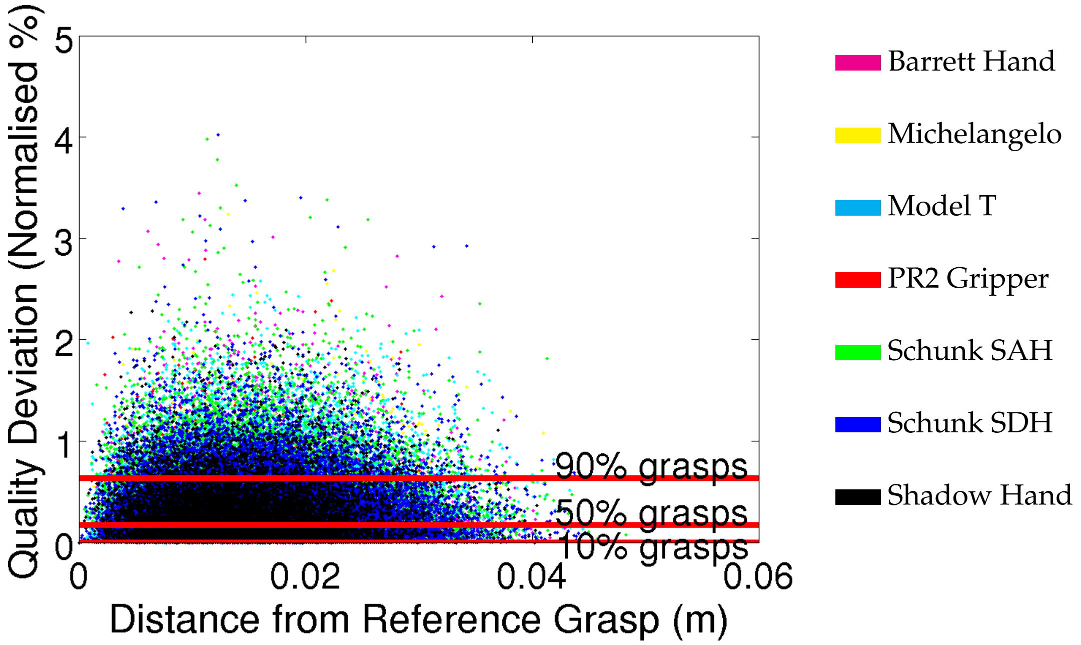
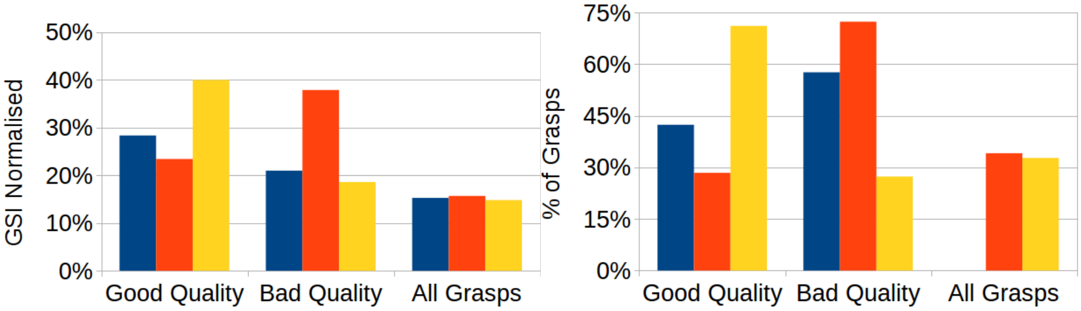


FIGURE 4.21: Variation in the metric Q_{B2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{B3} - Shape of the grasp polygon

Graphs on Figure 4.22a show the GSI_N for metric Q_{B3}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.22b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N (B) Distribution of grasps

FIGURE 4.22: GSI_N for metric Q_{B3}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.23 shows the comparison of GSI_N between different hand models.

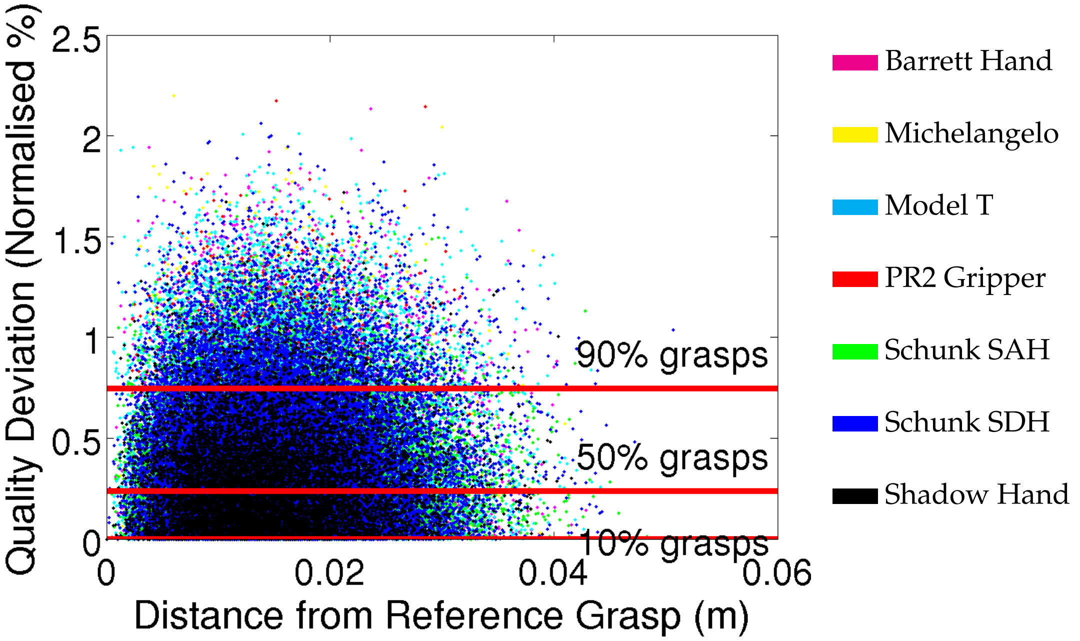
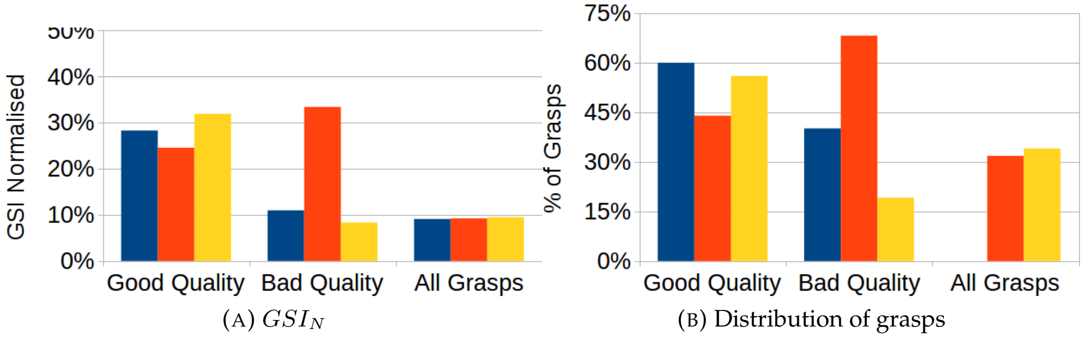


FIGURE 4.23: Variation in the metric Q_{B3} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{C1} - Smallest maximum wrench to be resisted

Graphs on Figure 4.24a show the GSI_N for metric Q_{C1}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.24b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N

(B) Distribution of grasps

FIGURE 4.24: GSI_N for metric Q_{C1}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.25 shows the comparison of GSI_N between different hand models.

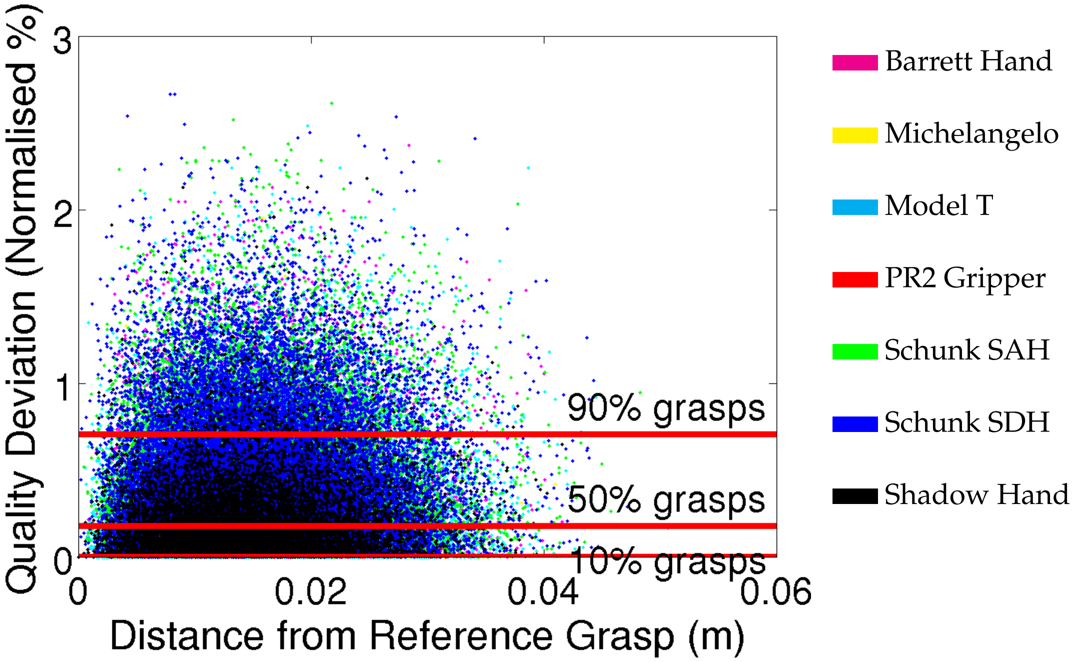
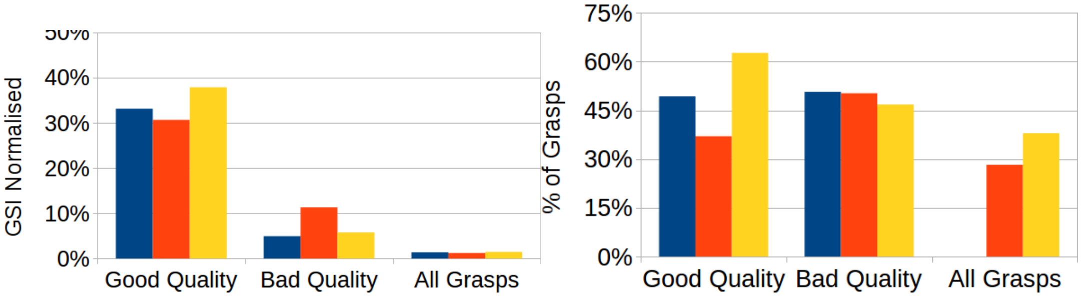


FIGURE 4.25: Variation in the metric Q_{C1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{C2} - Volume of the convex hull

Graphs on Figure 4.26a show the GSI_N for metric Q_{C2}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.26b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N

(B) Distribution of grasps

FIGURE 4.26: GSI_N for metric Q_{C2}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.27 shows the comparison of GSI_N between different hand models.

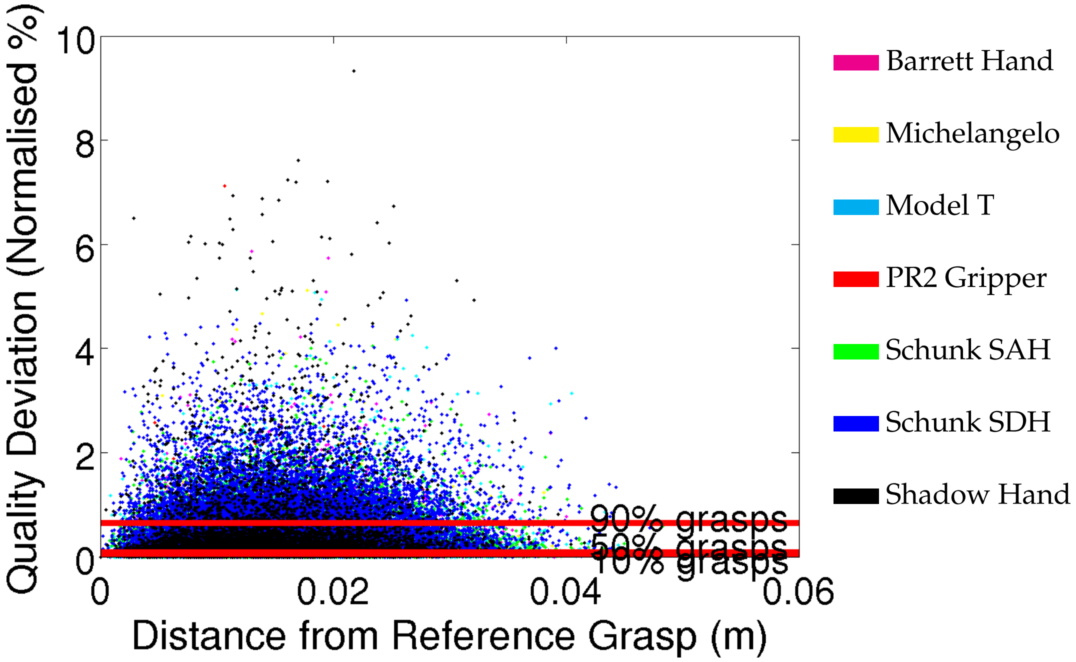
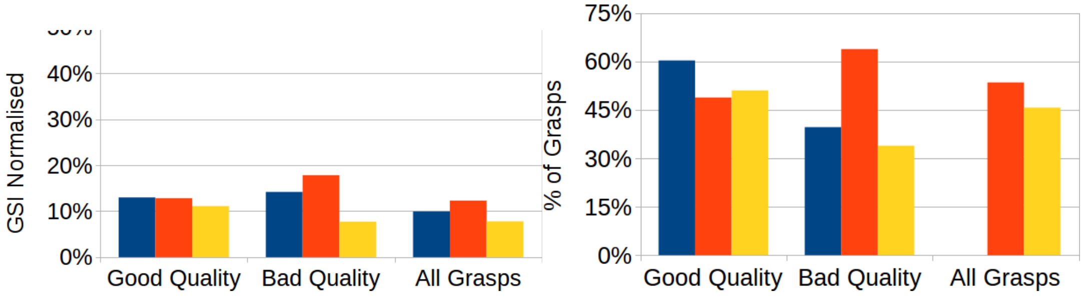


FIGURE 4.27: Variation in the metric Q_{C2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{D1} - Posture of manipulator joints

Graphs on Figure 4.28a show the GSI_N for metric Q_{D1}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.28b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N (B) Distribution of grasps

FIGURE 4.28: GSI_N for metric Q_{D1}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.29 shows the comparison of GSI_N between different hand models.

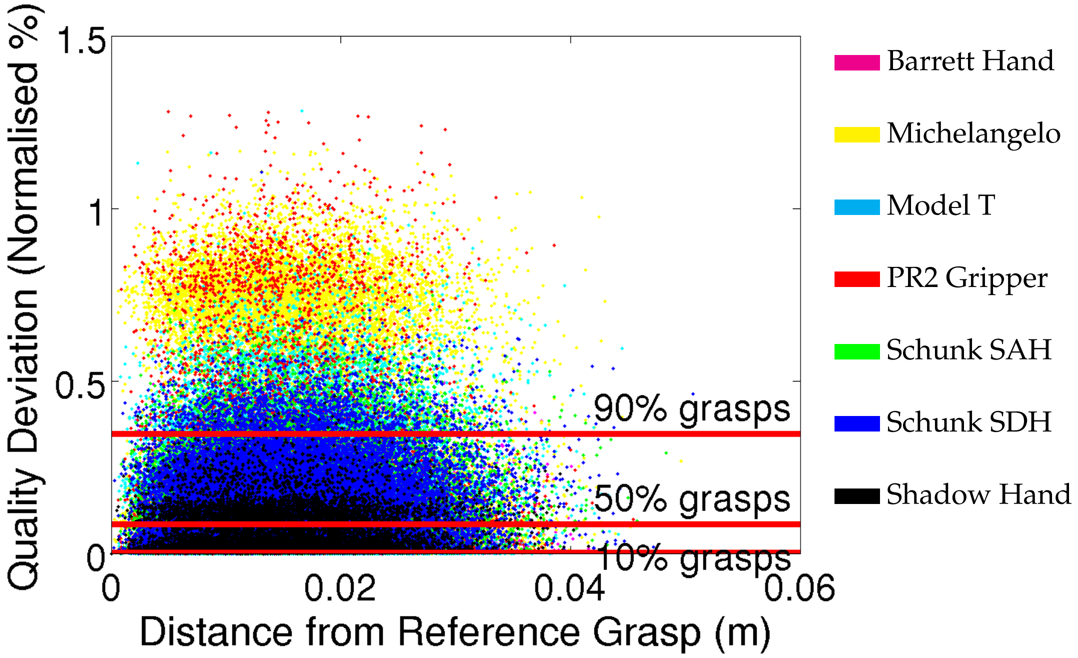
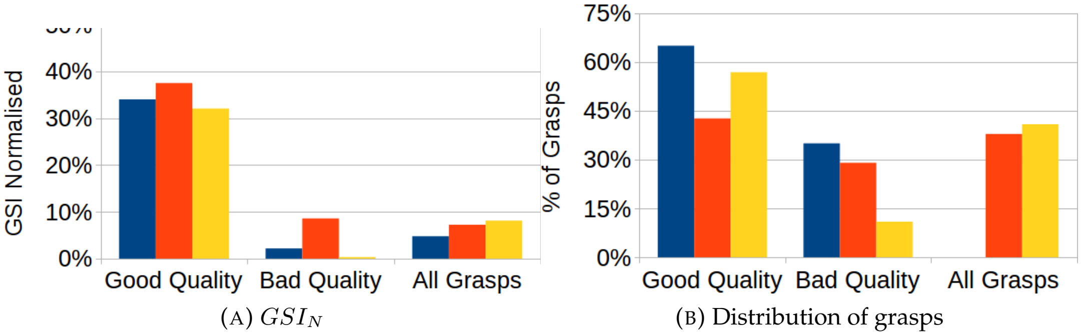


FIGURE 4.29: Variation in the metric Q_{D1} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Q_{D2} - Inverse of the condition number of G_J

Graphs on Figure 4.30a show the GSI_N for metric Q_{D2}. This graphs differentiate between *good* and *bad* quality grasps, it also compares the sensitivity in variations which improve or worsen the quality of the grasp. The distribution of grasps in these categories are shown in Figure 4.30b. It compares how many grasps were for good/bad quality grasps, and how many variations improved or worsen the reference grasp.



(A) GSI_N (B) Distribution of grasps

FIGURE 4.30: GSI_N for metric Q_{D2}.

In each graph there are three groups of columns: the left one shows the results for the *good* references grasps (value over the 50% threshold); the middle group for the *bad* reference grasps; and the third group for all reference grasps. In each group *blue* column represents all the variations in the category; the *orange*, those which improve the metric values; and the *yellow* column which worsen the metric.

Finally, Figure 4.31 shows the comparison of GSI_N between different hand models.

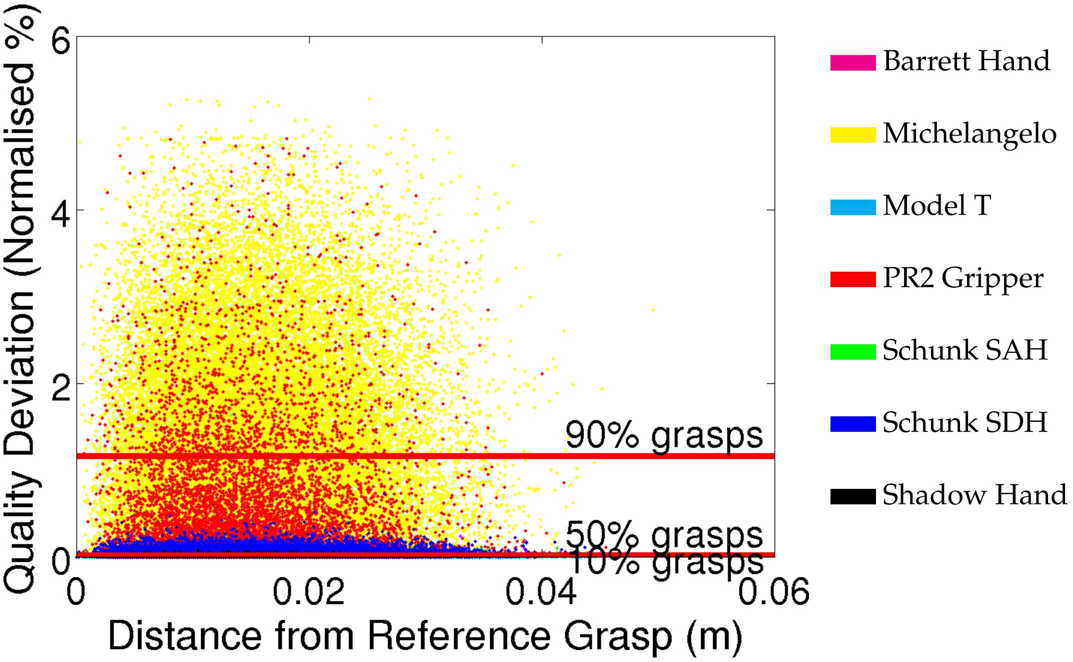


FIGURE 4.31: Variation in the metric Q_{D2} for each grasp with respect to the distance to its reference grasp. Each point represents a grasp, and grasps on different hands are coloured differently. Horizontal lines are drawn to indicate the limits containing 10%, 50% and 90% of the grasps.

Summary

The GSI_N for each metric is detailed in Table 4.17

TABLE 4.17: GSI_N(%) for every metric

Q_{A1}	Q_{A2}	Q_{A3}	Q_{B1}	Q_{B2}	Q_{B3}	Q_{C1}	Q_{C2}	Q_{D1}	Q_{D2}
23.19	30.12	25.39	12.76	33.23	35.91	31.76	27.98	15.25	25.88

In general, higher GSI_N values are observed for $good$ reference grasps than for bad ones, especially for Q_{A2}, Q_{C2}, and Q_{D2} metrics; very low differences between $good$ and bad reference grasps are found for Q_{B1} and Q_{D1}.

If we analyze the distribution of grasps and GSI_N for variations that improve or worsen the quality of the reference grasp, is shown, in general, the sensitivity is higher in *good* grasps which are worsen and *bad* grasps which improve. This suggest is easier to improve, and at a bigger rate, a bad grasp than to worsen it, and backwards, it's easier to make worst a grasp which is actually good. According to the distribution of values seen in the previous section 4.2, values of metrics tend to concentrate in specific ranges, which agree to the fact that its easier to worsen a good grasp and improve a bad one. The values corresponding to these graphs are detailed in Appendix A.1.

Finally, in the relations between the quality variation versus the distance from a grasp to its reference grasp, deviations are clearly different for some metrics depending on the hand considered; lowest deviations correspond to the Shadow Hand for all metrics, while the Michelangelo prosthetic hand produce high deviations for the metrics Q_{A1}, Q_{A3}, Q_{D1}, and Q_{D1}, as well as PR2 Gripper for metrics Q_{D1} and Q_{D2}.

This differences in GSI_N with different manipulators, suggest each manipulator could have their own properties and some could perform better than others according to specific metrics. Details in this properties and some others will be discussed later, on section 6.

4.4 Correlation

The purpose of this study is to find out which metrics provide similar evaluations for the wide set of grasps considered in this chapter. This would indicate that such metrics are basically representing the same underlying property and, consequently, that one of them is sufficient to evaluate that property, while the others are redundant. As a final consequence, this would allow the reduction of the number of metrics to be computed without losing any relevant information.

4.4.1 Methods

For this study, the same set of grasp generated in Section 4.2 is used, since it is sufficiently exhaustive and representative of grasping. The grasps are compared without any normalization method. This is done to avoid distorting the numeric values, as the correlation coefficients could be influenced by this variations. The Pearson correlation coefficient is calculated for each combination of metrics. However, this method is highly sensitive to outliers and, as it was shown in Section 4.2, this is quite common in data from quality metrics. Thus, the Spearman correlation coefficient is also calculated, as it has been shown to be more robust in front of outliers Fieller et al., 1957. Spearman correlation is also interesting since it can measure non-linear mappings between metrics.

4.4.2 Results

Table 4.18 shows the results for the Pearson and Spearman correlation coefficients for each pair of metrics. Cells with weak correlation values (≥ 0.4 and ≤ 0.7) are colored in yellow; cells with strong correlation values (≥ 0.7) are colored in green and full correlation cells (≥ 0.95) are colored in blue. When using the Pearson coefficient, only a strong case of correlation is found, Q_{A1} with Q_{A3}; when the effect of outliers is mitigated using the Spearman coefficient, three strong binary correlations are found, the ones between all metrics of group A.

Figure 4.32 plots the relation between metrics with highest correlation values: Q_{A1} vs. Q_{A2}, Q_{A1} vs. Q_{A3}, Q_{A2} vs. Q_{A3}, Q_{A2} vs. Q_{C2} and Q_{C1} vs. Q_{C2}.

TABLE 4.18: Correlation coefficients

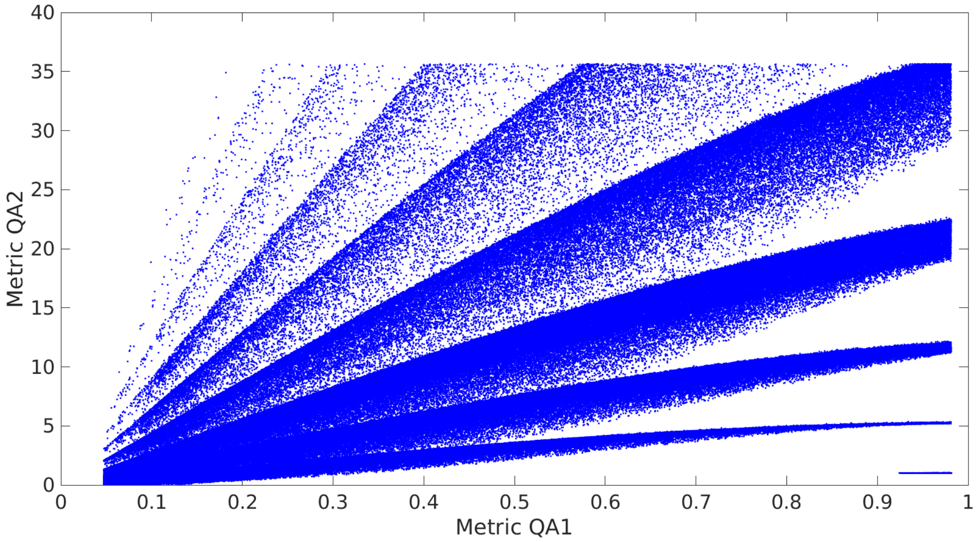
(A) Pearson correlation coefficient

	Q_{A1}	Q_{A2}	Q_{A3}	Q_{B1}	Q_{B2}	Q_{B3}	Q_{C1}	Q_{C2}	Q_{D1}	Q_{D2}
Q_{A1}	1									
Q_{A2}	0.65	1								
Q_{A3}	0.86	0.31	1							
Q_{B1}	-0.13	0.07	-0.36	1						
Q_{B2}	0.21	0.26	0.16	0.07	1					
Q_{B3}	-0.12	-0.15	-0.06	0.03	0.01	1				
Q_{C1}	0.34	0.43	0.24	0.23	0.22	-0.04	1			
Q_{C2}	0.38	0.35	0.33	0.23	0.22	-0.02	0.66	1		
Q_{D1}	-0.20	-0.16	-0.09	-0.05	0.17	0.16	0.16	0.22	1	
Q_{D2}	-0.30	-0.17	-0.24	0.00	0.15	0.18	-0.07	-0.01	0.40	1

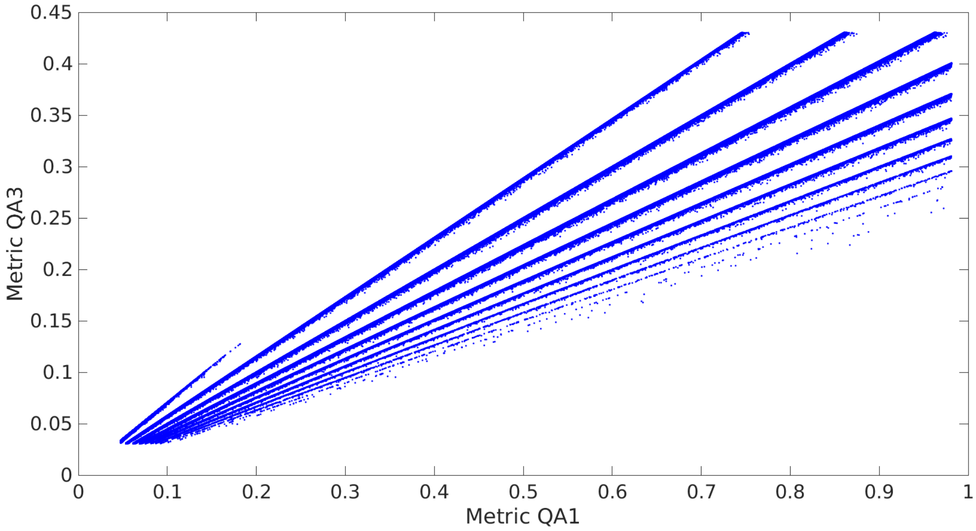
(B) Spearman correlation coefficient

	Q_{A1}	Q_{A2}	Q_{A3}	Q_{B1}	Q_{B2}	Q_{B3}	Q_{C1}	Q_{C2}	Q_{D1}	Q_{D2}
Q_{A1}	1									
Q_{A2}	0.86	1								
Q_{A3}	0.98	0.79	1							
Q_{B1}	-0.04	0.09	-0.05	1						
Q_{B2}	0.26	0.34	0.21	0.08	1					
Q_{B3}	-0.13	-0.22	-0.07	0.03	0.00	1				
Q_{C1}	0.34	0.46	0.26	0.19	0.27	-0.02	1			
Q_{C2}	0.49	0.56	0.44	0.23	0.39	-0.02	0.59	1		
Q_{D1}	-0.22	-0.31	-0.19	-0.01	0.18	0.11	0.19	0.31	1	
Q_{D2}	-0.28	-0.24	-0.28	0.05	0.08	-0.06	0.06	0.32	0.48	1

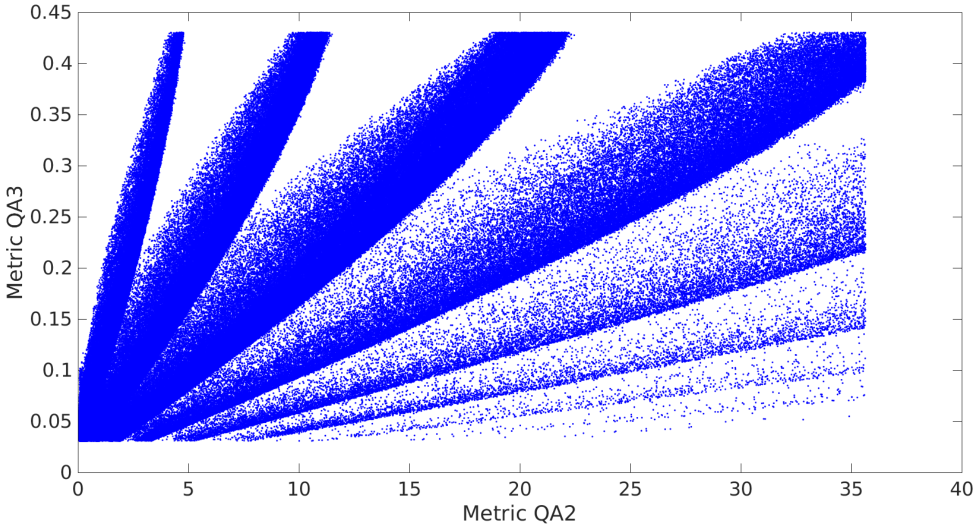
Yellow: soft correlation (≥ 0.4), Green: strong correlation (≥ 0.7) and Blue: full correlation (≥ 0.95). In both tables, the significance values are less than 0,05.



(A) Q_{A1} vs. Q_{A2}



(B) Q_{A1} vs. Q_{A3}



(C) Q_{A2} vs. Q_{A3}

FIGURE 4.32: Comparison between highest correlated metrics. Outlier values are discarded.

4.5 Discussion & conclusions

Sections 4.2 to 4.4 describe the three analyses that have been carried out on an extensive set of simulated data. The amount of results presented in the form of figures and tables is large and a careful review of them is necessary in order to reach useful conclusions. The most relevant of them are discussed in this section.

First, Section 4.2 describes the variability analysis, and the first question addressed is how to generate an unbiased set of grasps. Four alternatives are proposed and compared. The results shown in Fig. 4.2 indicate that the four methods produce indistinguishable sets of grasps, as long as these sets are sufficiently large.

The analysis of PDFs show the shape of metric values on the dataset. With is information, it is possible to establish benchmarks to study if other datasets of grasps are biased for any metric. There are some *peaks* of values for metrics Q_{A1}, Q_{A2}, Q_{B3}, Q_{C1} and Q_{D1}. These peaks correspond to specific hand models. Chapter 6 will present a more deep study in the characteristics and performance of different hand models according to specific quality metrics.

The results of the complete variability study are summarized in Table 4.3. These results are relevant, as they can be used to assess the room for improvement of a specific grasp, regarding a given quality metric. An important phenomenon observed on these results, also, is the characteristic distribution of most of the values of the metrics which tend to concentrate in narrow ranges with an important number of atypical values above and below these ranges. This phenomenon complicates the selection of upper and lower thresholds for normalization if the aim is to obtain an almost uniform distribution of values within the proposed limits.

Several approaches to set those range limits have been tested, and finally the use of 10^{th} and 90^{th} percentiles seem to be a good compromised option for all metrics. Some of them, however, still show a non uniform distribution, especially Q_{D2} and Q_{C1}, and more refined approaches could be used for these specific metrics. In practical terms a normalization based on the 10/90 percentiles and the thresholds derived form them seem to be most advisable since it provides the most uniform distributions, but with the limitations discussed above.

The second study is the sensitivity study detailed in Section 4.3, which is aimed to determine the behavior of the metrics in front of uncertainty in hand positioning. According to Table 4.17, all metrics present a GSI_N over 10%. Globally, the most robust metrics are Q_{B1} and Q_{D1}, while Q_{B2}, Q_{B3}, Q_{C1}, and Q_{A2} are the more fragile ones. More detailed information can be obtained from figures 4.12a to 4.30a, that shows the GSI_N distinguishing between *good* and *bad* reference grasps.

The obvious conclusion from these figures is that reference grasps whose values are over 50^{th} percentile are more fragile than those under 50^{th} percentile for most of the quality metrics, and especially for the cases of Q_{A2}, Q_{C2}, and Q_{D2}. Metrics Q_{D1} and Q_{B1}, which are the most robust ones globally, have a similar robustness independently of the goodness of their quality value. Also, for *good* reference grasps, the metrics are more fragile for those grasps variations that worsen the quality metric, except for metrics Q_{D1} and Q_{D2}; while for *bad* reference grasps, the metrics are more fragile for those grasp variations that improve the quality metric.

It is also interesting to note on figures from 4.13 to 4.31 that quality deviations are significantly different depending on the robotic hand considered. The Shadow Hand presents the lowest deviations for all metrics, while the PR2 Gripper and the Michelangelo prosthesis present the highest deviations for several metrics, which could be related with the low number of actuators of both hands, that greatly reduces their versatility. Grasp quality is, thus, highly dependent on the robotic hand design. Therefore, the design of robotic hands and prostheses could benefit from the grasp simulation and computation of grasp quality metrics if improvement of these magnitudes were used to guide design modifications.

Finally, an important result of this second analysis that should not be neglected is the high rate of grasp variations that do not meet the force-closure condition. Almost four fifths of them are discarded for this reason. We have decided to do so because some of the metrics require the force-closure condition to be met so the metric values make sense. It should be considered if some more relaxed condition could be chosen and how this could affect the results, and more importantly whether results sensitivity are relevant with such amount of variations discarded.

The third and final analysis is the correlation analysis presented in Section 4.4, aimed to determine relations between metrics. The results from this analysis are again deeply affected by the presence of atypical values in the metrics. This problem has been addressed by using the Spearmam correlation coefficient. Table 4.18 shows the results using both the Pearson and Spearman correlation rules. The latter table shows higher correlations than the former.

These results indicate that all the metrics based on the algebraic properties of G matrix (Q_{A1}, Q_{A2}, and Q_{A3}) represent the same underlying principle. Metrics of *Group C*, shows a weak correlation with this group. Finally, Q_{C2} also shows a weak correlation with metric Q_{C1}, from its same group.

In short, the ten metrics could be reduced to seven groups composed of five independent metrics Q_{B1}, Q_{B2}, Q_{B3}, Q_{D1}, and Q_{D2}, a pair of metrics, Q_{C1} and Q_{C2}, and a triplet of metrics, Q_{A1}, Q_{A2}, and Q_{A3}. Because of the weak correlations between the metrics from groups C and A (both groups based on algebraic properties of G), all these metrics could be even reduced to the same group without losing a significant amount of assessment information.

Once analyzed the robustness of the different metrics, we can postulate that Q_{A1} could be a good choice for representing the triplet of metrics: Q_{A1}, Q_{A2}, and Q_{A3}, as it is the one among them that has the lowest GSI_N.

Conclusions

To conclude, this chapter presents a numerical study on the practical parameters that characterize the use of ten quality metrics for evaluating grasps performed by robot hands on rigid objects. In the first place, the thresholds that indicate the practical upper and lower limits of the quality metrics have been established. These thresholds allow a more realistic normalization of the metrics values. However, the influence of object's size on these thresholds should be addressed in future work.

Second, the sensitivity analysis has shown the fragility or robustness of the metrics in front of positioning uncertainty. Finally, a correlation analysis has revealed the underlying relations between some metrics and the independence of others, which leads to conclude the existence of at least seven relevant grasp aspects to be taken into account when evaluating a grasp.

The results presented in this work may provide several benefits. First, they provide parameters regarding quality metrics which allow a better use of them in grasp planners, by the setting of data-based thresholds for evaluating and ranking of grasps. Second, the resulting data are a guidance in the selection of metrics in terms of robustness and equivalence between different alternatives. And finally, the results also give good hints in order to build combined metrics, which can capture more relevant aspects of a grasp and, thus, provide a better prediction of the grasp performance.

Limitations

However, this study has also some limitations that should not be disregarded. First, this work only considers ten quality metrics, while at least 24 have been surveyed Roa and Suárez, 2014. Though some of them were not compatible with the present study, i.e., biologically oriented, combined, and task oriented metrics, others could have been included, but were discarded.

Second, this study is focused on static grasp for holding objects, consequently, no task-oriented grasp metrics or analyses have been considered. An obvious complementary extension of this work would be to include task oriented analyses.

Third, only static contacts and states were simulated. Dynamic behavior of objects during grasping would provide a much more realistic study of the metrics as quality predictors. This should be possible using physics engines in order to simulate dynamic effects of grasps, though the unreliable physical fidelity that kind of engines are able to achieve poses some shadows on this approach.

Fourth, the selection of hand models may introduce a bias, as the metrics could be giving better or worst values depending on the hand used, as will be demonstrated in Chapter 6. In any case, we believe that the wide variety of hands and models used is enough the reduce if not eliminate this possible bias. And fifth, the election of OpenRave as simulation tool can also affect the results as information of contact points may differ from those that could be obtained using other simulators. Replicating the study, using other simulators and grasp generation procedures could clarify the concerns.

Finally, the friction coefficient through this analysis has a constant value of 0.4, which adds another limitation. The friction cones produced at the resulting contact points are affected when changing the value of the friction coefficient, which could lead to different evaluations of the quality measures for the same grasp pose. Therefore, the friction coefficient has a major role as a filter to discard more or less candidates. The sensitivity of the metrics to friction coefficient changes has been previously studied to some extentZheng and Qian, 2005; Hang et al., 2013 and it is not discussed in this thesis.

And last but not least, the most obvious limitation is the problem of transferring results from simulation to real robots. There are some studies that compare and analyze the performance of metrics for real applications. In fact, recent experiments have questioned the utility of analytical metrics. In Diankov and Kuffner, 2008; Balasubramanian et al., 2012 the authors show that grasps highly ranked by the ϵ-metric (Q_{C1} in this thesis, see Table 2.2) perform poorly on real robots. It has also been demonstrated that this metric has a high sensitivity to positioning errors Weisz and Allen, 2012.

There can be several reasons for this fragility of the grasp quality metrics. On one hand, the metric used in the above works, Q_{C1}, has shown to be very sensitive to uncertainty, a key aspect on the execution on real robots. On the second hand, the correlation analysis has shown that there are at least seven independent dimensions for evaluating a grasp. Grasp planners which rely on a single metric and ignore the other dimensions are probably providing incomplete grasp quality predictions.

Future & related work

The results on this chapter are exclusively based on static simulation. They might be taken with cautions until a proper validation and correlation with results obtained from a real-world scenario. However, the preparation and realization of such experimentation is complex and out of the scope of this chapter. We see the results presented here as the first and necessary phase of such a work which will be discussed next, in Chapter (Chapter 5).

We foresee applications which could take profitable use of the results and methodology proposed in this chapter. The results on this chapter constitute a huge background database to be used as reference for evaluating hand designs. The methodology would consist on testing a hand on a sufficiently large database of objects, and employing a large number of grasps in each case. The results could be statistically compared against the ones shown in this work, and evidence-based conclusions could be reached regarding the capabilities of the novel design. The methodology and results described in this chapter could be easily adapted for this application. The results of such application would be seen in Chapter 6.

The motivation, studies and experiments described in the present chapter have been partial or totally published in Leon et al., 2014 and Rubert et al., 2018.

Chapter 5

Grasp success prediction

In this chapter, we analyze how well different, commonly-used grasp metrics (Chapter 2.4) are able to predict grasp success, either individually or in combination. We hypothesize that these metrics capture different aspects of precision grasp stability and thereby have different biases and generalization characteristics. There seems to be a non-trivial mapping from each of these metrics or a combination of them to the binary decision of grasp success.

Our aim is to find a classifier that takes one or all of these metrics as input and outputs a success label. We test a set of classification methods that covers different potential characteristics of the function of interest (linear versus non-linear; parametric versus non-parametric). In this way, we aim to understand the underlying structure of the data and whether these commonly used metrics are expressive enough.

As ground-truth for training these classifiers, three different studies are applied. First, we use human labeled data from Kappler et al., 2015. Second, we compare these human labels with real grasp executions and this data is used for training the classifiers. Details on the dataset used for these first two experiments are detailed in Section 5.2.1. Third, we propose a model to predict the outcome of real grasp execution using machine learning and a combination of grasp quality metrics.

We also analyze how much the context of a grasp influences its outcome. With context we refer to object properties or the gravity orientation. This is, how object properties or gravity orientations can influence the success or failure of a grasp. We also analyze whether this information can enhance the grasp success prediction in addition to the metrics.

With the experiments and results obtained we propose a design of classifier able to predict the outcome of real grasp executions. This classifier employs a two-step classification to label grasp candidates as *Robust*, *Futile* or *Fragile*. The success rate of the designed classifier is over 90%.

Finally, we test whether this classification model could be generalized to other grippers and objects.

The contributions of this chapter are: (i) an analysis of classic grasp metrics given a much larger data set of human annotations (4699) on a much larger variation of objects (616) than previous studies; (ii) we use a wide set of classification methods to analyze the underlying structure of the function mapping; (iii) we quantitatively and qualitatively compare the outcome of our analysis with previous similar studies (Leon et al., 2014; Rubert et al., 2018; Goins et al., 2014; Kappler et al., 2015); (iv) we study how reliable humans are as oracle for grasp prediction; (v) we analyze the outcome of real grasp executions using two different grippers and varying the weight of the objects (vi) we analyze how relevant contextual information is for the success of a grasp and (vii) a classifier for label grasp candidates is presented.

Outline This chapter is structured as follows: first, we introduce a review of previous and related work which inspired and motivated this study. Next we present the basic foundations of the machine learning algorithms used in our experiments and finally, the results. We split our work in three different studies. The first study is a comparison between the predictive capability of grasp metrics with a physic metric generated in dynamic simulation. In the second experiment, real grasp executions are used as ground-truth to analyze the predictive ability of quality metrics, physics metric and human labels. In the third and last experiment, we study grasp executions in objects with different weights and propose a grasp prediction model based on a 3-categories scale.

5.1 Introduction

Definition 5.1. *The **prediction** of a grasp consist in determining whether a simulated grasp will be successful in real world. A grasp is considered **successful** or stable, if it is able to hold and restrain an object, constraining its potential movements.*

An open problem in autonomous robotic manipulation is the prediction of grasp success *before* executing it in the real world. This is especially the case for precision grasps. They are characterized by few contact points between hand and object which makes them more brittle than power grasps. Precision grasps allow for deliberate object motion within the hand (Feix et al., 2016) but are also more brittle than power grasps where the hand envelopes the object.

Many analytic metrics have been developed to reliably predict precision grasp success. They usually assume known object and robot hand models along with their relative pose to compute precise contact points between object and hand. Some of these metrics have been carefully evaluated by comparing their prediction with the outcome of real robot experiments (Balasubramanian et al., 2012; Weisz and Allen, 2012). The results of these studies showed that the most popular grasp metrics are not very good predictors of grasp success in the real world. This insight lead to the development of data driven techniques for finding metrics to predict grasp success even for the cases where significant uncertainty is introduced through unknown object shape, uncertain object pose estimation and inaccurate actuation (see Bohg et al., 2014 for an extensive review).

Data-driven approaches require ground truth annotations of grasp success, of the kind shown in Fig. 5.1. One way to generate these annotations is to execute grasps with a real robot through an existing grasp strategy and observing their outcome. For instance, Levine et al., 2016; Pinto et al., 2016 approach the objects always from the top using a simple gripper. In both works, convolutional neural network are trained they predict the success of a grasp execution. large-scale dataset are used for the training and the data is obtained through hour and months of experimentation. Detry et al., 2011 sample not only top grasps but consider all 6 DOF for generating grasp candidates. In his work the robot follows a sequence on grasp-and-drop actions with the object to generate a *grasp density*. It links object-relative grasp poses to their success probability.

In other approaches, humans demonstrate a grasp to the robot (Balasubramanian et al., 2012; Goins et al., 2014; Herzog et al., 2014) or manually label images of objects with good grasps (Lenz et al., 2015; Redmon and Angelova, 2015). In these works the grasp candidates relies in the human knowledge and expertise, which could discard successful grasp candidates. All of these approaches are expensive, tedious and may require an entire robot farm (Levine et al., 2016) to sample sufficiently many annotated data points.

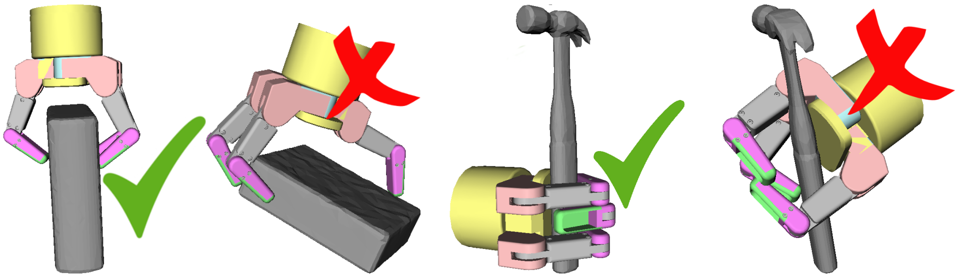


FIGURE 5.1: Examples of *Successful* (green mark) and *Unsuccessful* (red mark) grasps.

Another possibility is to execute and automatically label grasp candidates in simulation (Kappler et al., 2015; Mahler et al., 2016; Popovic et al., 2011). This way, several hundred thousand data points can be generated without any manual labeling or time-consuming execution in the real world. However, this leads to a *chicken-and-egg* situation: We want automatically label training data for the purpose of learning more informative grasp metrics. However, existing metrics do not lend themselves well to the automatic labeling task (Balasubramanian et al., 2012; Weisz and Allen, 2012). Kappler et al., 2015 propose to use physics simulation for this purpose. By comparing to the answers of a human oracle, the authors show how this leads to more realistic labels than when using a classic grasp stability metric. However, a full physics simulation is orders of magnitude more expensive than to compute the classic grasp metrics.

5.2 Foundations and related work

5.2.1 Grasps database

There are several publicly-available databases that contain grasps with ground truth annotations (Pinto et al., 2016; Levine et al., 2016; Lenz et al., 2015; Mahler et al., 2016; Goins et al., 2014). However, only (Mahler et al., 2016; Kappler et al., 2015; Goins et al., 2014) include the ground truth object models and relative pose between hand and object. This is necessary to compute the classic grasp metrics which we study in this chapter.

There are two ways to obtain ground-truth annotations. One is using real grasp execution and label it as Stable or Unstable. The other way is using showing grasp candidates to humans which will label them as Stable or Unstable.

In some databases, ground truth labels of grasp success are generated by executing grasps in the real world (Pinto et al., 2016; Levine et al., 2016; Goins et al., 2014). These are much more informative labels than can be generated by human subjects just looking at pictures of grasps. However, they are also much more time-consuming and expensive to generate. Consequently, the resulting data sets are often much smaller than data sets generated in simulation or may require an entire robot farm (Levine et al., 2016) to generate sufficiently many grasps.

In previous work, Kappler et al. (2015) generated a large-scale grasp database on a wide variety of objects in simulation using *OpenRave*. The work of (Kappler et al.,

2015) serve as precursor of the studies and experiments performed in this chapter. In this work, the grasps are generated by sampling grasp candidates uniformly around the object surface; thereby introducing less bias. In total, it contains approximately half a million grasps generated on more than 600 different object models with the BarrettHand.

Kappler et al. (2015) aimed at automatically generating training data with ground truth labels of grasp stability. The authors assumed that human subjects are perfect oracles for this prediction task and their judgment could therefore be used to validate the proposed automatic labeling process. Therefore, a subset of the grasps from the database were presented to human subjects using *Amazon Mechanical Turk* (AMT). They labeled them as either *Stable*, *Unstable* or *Unknown*. Each grasp was judged by at least 5 different subjects. The final label per grasp was averaged over their responses to normalize for the remaining noise.

In this chapter, we are using this extensive database as a basis to analyze an entire suite of grasp metrics (Section 2.4) and whether they are useful by themselves or in combination to predict grasp stability. As a ground truth label of this data, we will use the human labels, only for *Stable* and *Unstable* grasps. The database is the only one that has *human* grasp annotation for a *large* set of grasps (4752) and applied to a wide variety of different objects (616) with a known 3D shape model. It also considers *noise* in the object pose for each candidate grasp.

5.2.2 A physics metric

The central hypothesis by Kappler et al. (2015) is that physical forward simulation of grasps is more suitable for automatic labeling of data than the classical and most common Ferrari & Canny metric (Q_{C1} in this thesis). The final labels are computed by averaging over the physically-simulated outcomes of 30 grasps on a slightly perturbed object pose around a reference pose. This mimics noise in perception and actuation of a robot.

The hypothesis is tested against the answers of human subjects on a subset of the generated grasps. The results show that the labels based on the physics simulation are more realistic and therefore more suitable to learn from. However, it is also orders of magnitudes more expensive to compute than the classic metrics.

For reference, we also report this physics metric in the results section. It indicates the prediction performance of grasp stability that can be achieved when investing into this computationally expensive physics simulation.

5.2.3 Selection of quality metrics

In Chapter 4 is performed a statistical analysis of the values produced by ten selected quality metrics on a database that included grasps for seven different hands and more than hundred objects, resulting in around 900.000 different grasp configurations. The analysis is performed exclusively in simulation and consisted in establishing upper and lower thresholds for the normalization of the metrics; measuring the stability of the metrics in the presence of small disturbances; and more importantly, visualizing the correlation between different metrics. This last result allowed to discard three metrics reducing the initial set to seven. These are the seven independent metrics which will be used in this chapter (see Table 5.1).

The results of these works offer numerical and practical information about the use of the metrics. The main motivation of this chapter is to address the limitation of Leon et al. (2014) and Rubert et al. (2018) that no relation between highly ranked

TABLE 5.1: Summary of the selected independent quality metrics

Name		Formula
Q_{A1}	Smallest singular value of G (Li and Sastry, 1987)	$\sigma_{min}(G)$
Q_{B1}	Distance between the centroid of the contact polygon and the center of mass of the object (Ding et al., 2001; Ponce et al., 1997)	$1 - \frac{distance(p, p_c)}{distance_{max}}$
Q_{B2}	Area of the grasp polygon (Mirtich and Canny, 1994)	$\frac{Area(Polygon(p_1, ...p_n))}{Area_{max}}$
Q_{B3}	Shape of the grasp polygon (Kim et al., 2001)	$1 - \frac{1}{\theta_{max}} \sum_{i=1}^{n_f} \lvert\theta_i - \bar{\theta}\rvert$
Q_{C2}	Volume of the convex hull (Miller and Allen, 1999)	$\frac{Volume(CW)}{Volume_{max}}$
Q_{D1}	Posture of manipulator joints (Liegeois, 1977)	$1 - \frac{1}{n_q} \sum_{i=1}^{n_q} (\frac{y_i - a_i}{a_i - y_{iM}})^2$
Q_{D2}	Inverse of the condition number of $\mathbf{G_J}$ (Salisbury and Craig, 1982; Kim and Khosla, 1991)	$\frac{\sigma_{min}(G_J)}{\sigma_{max}(G_J)}$

grasps and the success of their execution is presented. The human labels from Kappler et al., 2015 and the real experiments performed in this chapter, provide a valuable ground truth to fulfill this purpose.

5.2.4 Human oracle

The work in Kappler et al., 2015 used humans as ground-truth to evaluate the grasp success. In this study several images of a grasp are shown to a human subject who evaluates it as *Stable*, *Unstable*, or *Unknown*. Each grasp is evaluated by at least 5 humans and the average of the evaluations is used to set a grasp as successful or not.

To filter spamming during the human experiments a rejection ratio is applied to the evaluations. This *rejection ratio* discards the evaluations of subjects who failed to correctly classify the grasp at least *n* times. In their work, Kappler et al., 2015 applied a rejection ratio of 0.3.

However, the use of humans as ground-truth is still questionable. There are no report or study on the accuracy of human knowledge to evaluate simulated grasps. Is thus, part of this chapter study the predictive capability of humans.

5.2.5 Related work

Goins et al. (2014) present very relevant work to this chapter. The authors address the problem of learning a predictor for successful grasps. In their approach, candidate grasps are instructed by humans and tested on a real robot. Up to 13 metrics are calculated for each grasp. Dimensionality reduction and a Gaussian Process are used to produce a predictor that demonstrates better performance than when using individual metrics.

There are similarities with our approach but there are also relevant differences. First of all, we are using 7 metrics instead of 13, but our subset has been previously reduced by discarding highly correlated metrics (Rubert et al., 2018). Although the classifiers may be able to cope with 13 metrics, we simplify the learning problem and the subsequent analysis by reducing the dimensionality of the input space. An obvious strength of (Goins et al., 2014) is that grasps are tested on a real robot, which is a more valuable ground-truth.

However, all of these grasps have been demonstrated by humans. In our opinion, this introduces a bias in the selection of grasps that could affect the learned models. This process potentially discards grasp configurations which could be successful but are not intuitive for humans. Furthermore, the dataset in their work is much smaller than the dataset used in this chapter. Also fewer objects are considered.

5.3 Binary classifiers

We aim to find a model $y = f(\mathbf{x}; \mathbf{w})$ that can predict binary grasp success y given an input feature vector $\mathbf{x}$ consisting of different grasp quality metrics. Our approach is to learn a binary classifier from the data in (Kappler et al., 2015) that ideally minimizes the following equation:

$$\min_{\mathbf{w}} \sum_{(\mathbf{x},y)\in\mathcal{D}} 1 - l(f(\mathbf{x}; \mathbf{w}), y) \tag{5.1}$$

where $\mathbf{w}$ denotes the parameter vector of the classifier and

$$l(f(\mathbf{x}; \mathbf{w}), y) = \begin{cases} 1, & \text{if } f(\mathbf{x}; \mathbf{w}) = y \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

Every classifier considered in this chapter minimizes a loss in this setting. For various reasons, the exact loss formulations may vary per method, e.g. through different regularizers or by dropping the indicator function in order to get gradients.

Given this data set $\mathcal{D}$, we train the following classifiers that have different characteristics (linear versus non-linear, non-parametric versus parametric) using SciKit-Learn (Pedregosa et al., 2011) or Matlab (MATLAB, 2010).

The optimization for each method was done using grid search and cross-validation. Details are presented in each subsection. A more in-depth description of the classification methods used in this chapter can be found in (Bishop, 2006).

Logistic regression

Logistic regression is a linear model for classification. The logistic function $\sigma(\mathbf{w}^T\mathbf{x})$ models the probability $P(y = 1|\mathbf{x})$ of feature vector $\mathbf{x}$ describing a stable grasp $y = 1$ with

$$\sigma(a) = 1/(1 + \exp(-a)). \tag{5.3}$$

To find the weight vector $\mathbf{w}$, we optimize the following regularized loss function:

$$\min_{\mathbf{w}} \|\mathbf{w}\|_1 - C \sum_{(\mathbf{x},y)\in\mathcal{D}} y \ln t + (1 - y) \ln(1 - t) \text{ with} \tag{5.4}$$

with $t = \sigma(\mathbf{w}^T\mathbf{x})$, $\|\cdot\|_1$ denotes the L1-norm and encourages a sparse weight vector. C balances between the classification loss and the regularizer. The key assumption in logistic regression is linearity between the input features and the log odds, i.e. the log of ratios of binary class probability. If the data is linearly separable this method can recover the global minima of (5.1). We optimized the parameters using two loss functions (*L1* and *L2*) and two values for C: [1,2]. The loss functions *L1* and $C = 1$ showed the best performance.

The logistic function $\sigma(\mathbf{w}^T\mathbf{x})$ is defined as:

$$f(\mathbf{x};\mathbf{w}) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}} \tag{5.5}$$

Naive bayes

Naive Bayes models the probability of a grasp being either *stable* or not using Bayes theorem.

$$P(y|\mathbf{x}) = \frac{P(y)P(\mathbf{x}|y)}{P(\mathbf{x})}. \tag{5.6}$$

Its key assumption is that given the class label y, the distributions of the input variables $\mathbf{x} = (x_1, \dots, x_7)^T$ are independent. Thus, we can rewrite Bayes theorem as:

Formally, this means

$$P(x_i|y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_7) = P(x_i|y) \tag{5.7}$$

leading to the following simplification of (5.6):

$$P(y \mid x_1, ..., x_7) = \frac{P(y)\prod_{i=1}^{7} P(x_i \mid y)}{P(x_1, \dots, x_7)} \tag{5.8}$$

To finally classify a data point, we have

$$f(\mathbf{x};\mathbf{w}) = \hat{y} = \arg\max_y P(y)\prod_{i=1}^{7} P(x_i \mid y) \tag{5.9}$$

The parameters of this Naive Bayes model are $P(y)$ (the relative frequency of class y in the training set) and $P(x_i \mid y)$ (the distribution of a grasp quality metric given a class label; modeled as Gaussian in this chapter). We use *Maximum A Posteriori* (MAP) estimation to determine these parameters.

k-Nearest neighbors

K-Nearest Neighbors (KNN) is a non-parametric approach. It makes no assumptions about the distribution that generates the data. Instead, it classifies a test data point based on the class membership of its K nearest neighbors in feature space. To define distances, we assume that this space is Euclidean.

Formally, KNN uses Bayes theorem to model the posterior probability of a grasp being successful or not

$$P(y = c|\mathbf{x}) = K_c/K \tag{5.10}$$

where K_c is the number of data points belonging to class c among the K nearest neighbors of $\mathbf{x}$ and c corresponds to either *Stable* or *Unstable*. To classify a data point, we maximize this posterior distribution over the binary class labels.

K is the hyperparameter in this classification method. In general a larger K suppresses the effects of noise, but makes the classification boundaries less distinct. We used a validation set to find the optimal K for our data set. KNN is well suited for our relatively low-dimensional feature space. However, in its most basic form inference for this method does not scale well with the number of data points and dimension of feature space.

Two different weight functions were compared during the training: *uniform* and *distance*. The $weights = uniform$, assigns uniform weights to each neighbor. $weights = distance$ assigns weights proportional to the inverse of the distance from the query point. The K value was analyzed from 1 to 20. In our results the *distance* function showed a better performance with a K value of 5.

The distance function is:

$$f(\mathbf{x};\mathbf{w}) = \sqrt{\sum_{1}^{k}(x_i - y_i)^2} \tag{5.11}$$

Classification trees

Classification Trees (CTs) are another non-parametric approach. A CT is a binary tree that divides the input feature space at each node j into two regions according to whether $x_i \leq \theta_j$ or $x_i > \theta_j$. θ_j is a parameter of the model. These sub-regions are independently subdivided further by moving down in the tree until a leaf node is reached.

Each leaf node of the tree encodes a region $\mathcal{R}_\tau$ in the input space. This region contains N_τ training data points associated with class labels y_n. Let us assume that a test data point with features $\mathbf{x}$ ended up in the leaf node corresponding to $\mathcal{R}_\tau$. The probability $P(y = c|\mathbf{x})$ is then the fraction of data points labeled with c in that region.

$$P(y = c|\mathbf{x}) = m_{c,\tau} \text{ with } m_{c,\tau} = \frac{1}{N_\tau} \sum_{(\mathbf{x}_n, y_n) \in \mathcal{R}_\tau} 1(y_n, c) \tag{5.12}$$

For the final classification of a test data point, we again maximize $P(y = c|\mathbf{x})$ over the class labels c.

To learn a CT, we need to find the optimal split parameters θ_j at each node of the tree. This is done by a standard greedy strategy where at each new node, we optimize a certain criteria over a set $\mathcal{S}$ of candidate pairs of input features x_i and thresholds θ_j. The criteria has to capture the purity of class labels within the sub-region $\mathcal{R}_\tau$ resulting from a candidate split of the input region. In this chapter, we chose the Gini impurity $g(\mathcal{R}_\tau)$:

$$\min_{s_\tau \in \mathcal{S}} g(\mathcal{R}_\tau) \text{ with } g(\mathcal{R}_\tau) = 1 - \sum_{c \in \{0,1\}} m_{c,\tau}^2 \tag{5.13}$$

The more nodes are added, the deeper the tree and therefore also the more complex the decision rule. To avoid overfitting, we have found a maximum tree $depth = 5$ using the validation set. We have further investigated using entropy as an impurity measure with no significant difference in the results.

Support vector machines

In this chapter, we use Radial Basis Function *Support Vector Machines* (RBF SVMs) as a non-probabilistic binary classifier. It is characterized by defining a max-margin hyperplane that separates the two classes such that the distance between the closest data point and the hyperplane is maximized. Given a data point with $\mathbf{x}$ as feature vector, we classify it as 1 if $f(\mathbf{x};\mathbf{w}) > 0$ or 0 otherwise

$$y = \begin{cases} 1, & \text{if } f(\mathbf{x};\mathbf{w}) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{5.14}$$

To find the optimal parameters $\mathbf{w}$ and b of the classifier given the dataset, we maximize the dual problem formulation

$$\max_{\alpha} \sum_{i=1}^{\|\mathcal{D}\|} \alpha_i - \sum_{i=1}^{\|\mathcal{D}\|}\sum_{j=1}^{\|\mathcal{D}\|} y_i \alpha_i k(\mathbf{x}_i, \mathbf{x}_j) y_j \alpha_j \tag{5.15}$$

$$st. \sum_{i=1}^{\|\mathcal{D}\|} \alpha_i y_i = 0 \text{ and } 0 \leq \alpha_i \leq C \tag{5.16}$$

where C trades off between margin size and correct classification of the data points and $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$ is the RBF kernel with the hyperparameter γ. The hinge loss defines a soft margin that is beneficial in case the data is not separable but otherwise reduces to the hard margin SVM. In order to infer the class label we replace $f(\mathbf{x};\mathbf{w})$ with the result of the dual optimization problem,

$$f(\mathbf{x};\alpha) = \sum_{\alpha_i:\alpha_i>0} \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b. \tag{5.17}$$

Here the weight is determined based on the kernel evaluation and the dual variable.

Gaussian processes

A *Gaussian Process* (GP) is a probabilistic, non-parametric method. More specifically, it is defined as a collection of a finite number of random variables with a joint Gaussian distribution. In our case, this collection consists of the grasps in the training data $\mathcal{D} = \{\mathbf{x}, y\}_{\|\mathcal{D}\|}$. In the following, we summarize the data in the matrix $\mathbf{X}$ and vector $\mathbf{y}$. A GP can be seen as a distribution over functions with a mean μ and covariance Σ. If we want to query the label y_0 of a test data point with feature vector $\mathbf{x}_0$, we have $f(\mathbf{x}_0) \sim N(\mu, \Sigma)$ with

$$\mu = k(\mathbf{x}_0, \mathbf{X})^T [K(\mathbf{X}, \mathbf{X}) + \sigma_M^2 I]^{-1} \mathbf{y} \tag{5.18}$$

$$\Sigma = k(\mathbf{x}_0, \mathbf{x}_0) - k(\mathbf{x}_0, \mathbf{X})^T [K(\mathbf{X}, \mathbf{X}) + \sigma_M^2 I]^{-1} k(\mathbf{x}_0, \mathbf{X}). \tag{5.19}$$

σ_M^2 is the variance of the noise on the target values.

The entries of the covariance matrix $K(\mathbf{X}, \mathbf{X})_{p,q}$ at row p and column q are defined based on a covariance function $k(\mathbf{x}_p, \mathbf{x}_q)$ with some hyperparameters θ. We use the squared exponential covariance function

$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_l^2 \exp(-((\mathbf{x}_p - \mathbf{x}_q)^T L^{-1} (\mathbf{x}_p - \mathbf{x}_q))/2) \tag{5.20}$$

where the hyperparameters are σ_l, the signal variance, and L, the identity matrix multiplied with the length scale l. We optimize these hyperparameters in the standard way by maximizing the marginal likelihood. To compute $P(y_0 = c|\mathbf{x}_0)$, we *squash* $f(\mathbf{x}_0)$ through the logit function (5.3).

Neural networks

Neural networks are nonlinear functions that we use in this chapter to map from our input feature vector $\mathbf{x}$ to the binary labels y. They are organized in multiple layers. Except from the input layer, each layer takes the output of the previous layer as input, potentially transformed with some non-linear function. For a binary classification, the output of the last layer is transformed using the logistic function (5.3). To classify a test data point with for instance a two-layer network, we need to compute

$$P(y = c|\mathbf{x}) = \sigma\left(\sum j = 0^M w_{cj}^{(2)} h\left(\sum i = 1^7 w_{ji}^{(1)} x_i\right)\right) \tag{5.21}$$

where $h(\cdot)$ refer to the logsig transfer function of the output from the first layer. The superscript (1) and (2) indicates the different weights of each layer.

Up to 12 different algorithms and 100 hidden layers for training the *Neural Network* were tested using the Matlab Neural Network Toolbox[1]. The best one was Bayesian Regularization (Foresee and Hagan, 1997), with layer size 17. It relies on the Levenberg-Marquardt algorithm to learn the optimal weights $\mathbf{w}^{(l)}$ for each layer. The validation set is used for early stopping and manual network structure optimization.

Evaluation metrics for classifiers

For comparing the different classification methods, we report accuracy, precision, recall and f1-score.

- The accuracy score correspond to the count of correct predictions, i.e. the count of *true positives* (tp) and *true negatives* (tp).
- The precision is the ratio $tp/(tp + fp)$ where fp is the number of false positives. Intuitively, it measures the capability of the classifier to avoid labeling a negative as a positive sample.
- The recall is the ratio $tp/(tp + fn)$ where fn the number of false negatives. Intuitively, recall measures the capability of a classifier to find all the positive samples.
- The f1-score is the harmonic mean of precision and recall: $F1 = 2 * (precision * recall)/(precision + recall)$. It reaches its maximum at 1 and minimum at 0.

---

[1] MATLAB, 2010

5.4 Case study 1: Human oracle

In this study we aim to find a model that predicts grasp success given an input feature vector of seven grasp metrics. Our approach is to learn this model from a large scale dataset with human annotations. In the following, we first describe the input features and the output label.

Then the methodology applied for the training procedure is explained and finally, the results are presented. We present results from two sets of experiments. First, we compare different classification methods to analyze their predictive capability. Second, we analyze the relevance of each grasp metric.

5.4.1 Feature vector and label

We compute the input feature vector $\mathbf{x}$ for each grasp in Kappler et al., 2015 with OpenHand (3.1). This input feature vector is composed by the evaluation with quality metrics on each grasp. Most of these metrics use the contact point information provided by the simulator, along with different information about the object and gripper geometry. The implementation and details of the quality metrics used is widely discussed in the literature: Roa and Suárez (2014), Leon et al. (2014), and Rubert et al. (2018) and the previous chapter (Chapter 4).

For contact detection, we use ODE. The friction coefficient of the contact points is 0.4. Each grasp has a ground truth success label provided by a human oracle where $y = 1$ refers to a *Stable* and $y = 0$ to an *Unstable* grasp. Our data set $\mathcal{D}$ is comprised of tuples $(\mathbf{x}, y)$, a feature vector and the corresponding label.

5.4.2 Methodology

Dataset

The *null error rate* of the dataset is 57%. This rate measures the accuracy of a classifier that would label all grasps in the dataset as *Unstable*. We use this as a lower baseline to study the effectiveness of the different predictors.

The dataset of grasps is split in the two following ways. First, we perform a *random* split where the *Train* set contains 80% and the *Test* set 20% of the grasps.

Second, we split the dataset such that the *Train* set only contains grasps from 80% of the objects and the *Test* set grasps from the remaining 20%. In this way, the trained models will be tested on a set of *held-out objects* they have not previously seen.

Training, validation and testing

For both split methods (random and *held-out*) the *Train* dataset is used to perform a 10-fold cross-validation. This method splits the train dataset into 10 smaller subsets (folds). It trains 10 models, each using one of the folds for validation and the remaining folds for training. We report the *mean* classification accuracy and its standard deviation (*std*) over these 10 models. We optimized the hyperparameter of the classification methods using this 10-fold cross-validation. Once each classifier is cross validated, a new training is done using the *Train* dataset and the model's performance is tested on the *Test* set. The results of the cross-validation and testing are reported for the two types of data set splits: *random* and *held-out objects*. Results on the *held-out objects* are only shown for the comparison of the classification methods as scores remained similar to the *random* split in the second set of experiments.

5.4.3 Results

In this section is shown first, the results comparing different machine learning algorithms for training our prediction model. Next, are shown the results on the performance of individual metrics to predict grasp success.

TABLE 5.2: Comparison of classification models. We report mean/std of classification accuracy using cross validation as well as accuracy on a test set. Results are computed for two dataset splits: random and held-out objects. As baseline, the Null Error rate on each dataset is showed.

Classifier	Random Split		Held-out Objects	
	CrossVal	Test	CrossVal	Test
Logistic Regression	0.73 ± 0.03	0.72	0.73 ± 0.05	0.72
Naive Bayes	0.55 ± 0.02	0.55	0.55 ± 0.06	0.58
K-Nearest neighbors	**0.80 ± 0.02**	**0.78**	0.76 ± 0.02	0.74
Classification trees	**0.80 ± 0.01**	**0.78**	0.75 ± 0.01	0.72
SVM	0.70 ± 0.02	0.70	0.69 ± 0.07	0.70
Gaussian Process	0.75 ± 0.03	0.75	**0.74 ± 0.04**	**0.76**
Neural Networks	0.80 ± 0.05	0.75	0.76 ± 0.05	0.73
Null Error Rate	0.55	0.55	0.55	0.58

Comparison of classification methods

Table 5.2 presents the mean classification accuracy and its standard deviation per algorithm using cross-validation for the two different splits of the dataset. It also reports classification accuracy on the test sets.

The results shown in this table are achieved with models whose hyperparameters are already optimized on a validation set.

A deeper analysis of the performance of different algorithm is shown in Table 5.3. Additionally to the accuracy already reported in Table 5.2, it also report precision, recall and F1-score. As a reference, we also calculate these metrics for the binary labels generated through physics simulation of the grasps (Kappler et al., 2015).

TABLE 5.3: Comparison of classification methods using different metrics to evaluate models performance.

	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.72	0.75	0.72	0.72
Naive Bayes	0.55	0.73	0.55	0.50
K-Nearest Neighbors	0.78	0.77	0.77	0.77
Classification Trees	**0.78**	**0.78**	**0.78**	**0.78**
SVM	0.70	0.72	0.70	0.70
Gaussian Process	0.75	0.76	0.76	0.76
Neural Networks	0.75	0.75	0.75	0.75
Physics Metric	0.87	0.87	0.87	0.87

Quality metrics predictive capability

In this second set of experiments, we analyze the performance of individual quality metrics to predict the stability of a grasp. This analysis is done in two ways, first a study on the individual performance of each metric and then, an ablation study on the influence of each metric for the predictors. To support this analysis, we visualize the distribution of *Stable/Unstable* grasps for each metric. Next figures, from fig. 5.2 to fig. 5.8 show the histograms of this data (with normal or log scale).

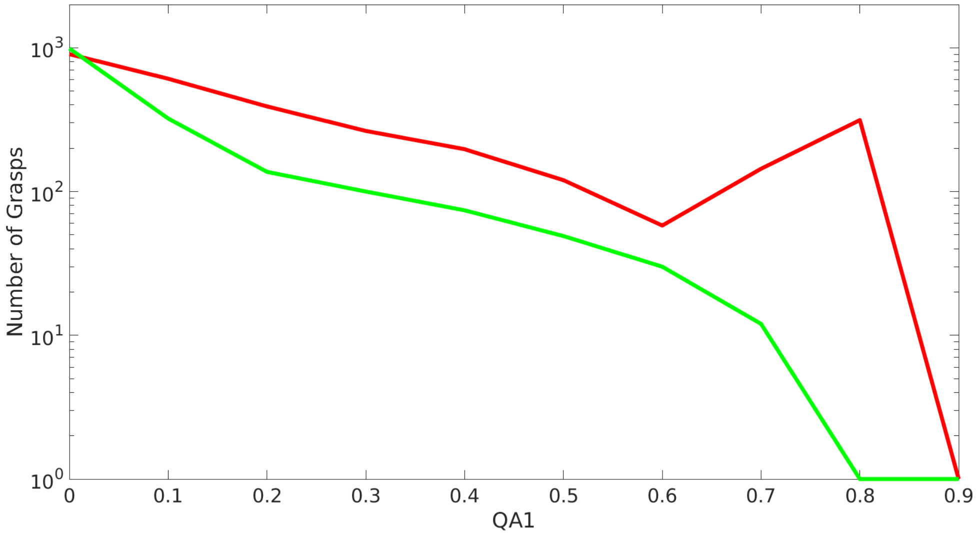


FIGURE 5.2: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{A1}. Histogram with log scale of y-axis.

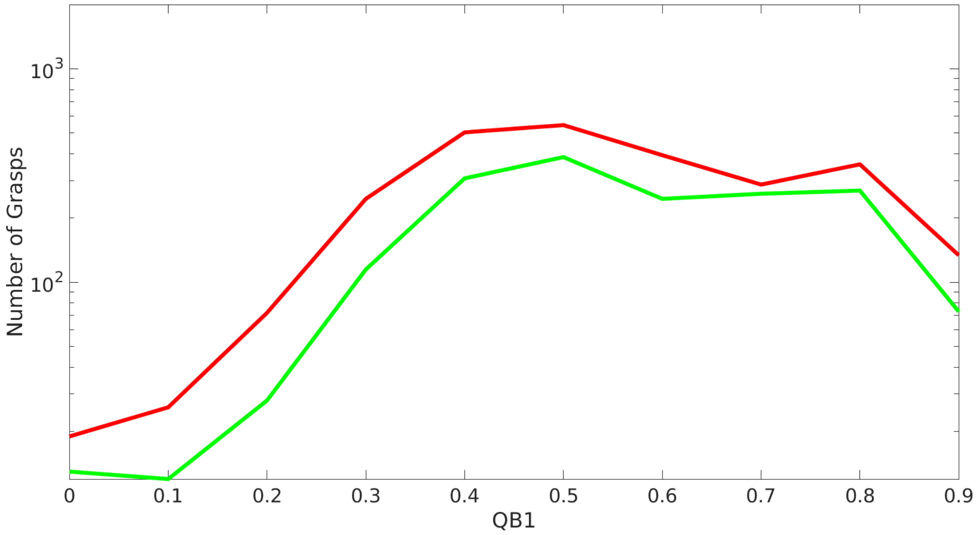


FIGURE 5.3: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{B1}. Histogram with log scale of y-axis.

Given the results from the comparison of classification models, we selected Classification Trees and KNN to study the capability of individual quality metrics to

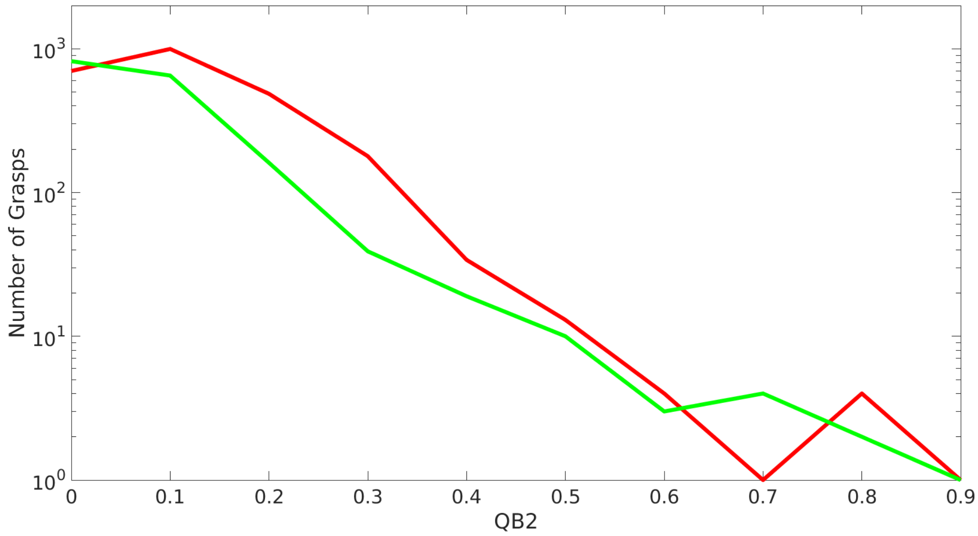


FIGURE 5.4: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{B2}. Histogram with log scale of y-axis.

predict grasp success. Table 5.4 reports classification accuracy. Additionally, we analyzed the performance of these classifiers when they use as input features all but one of the quality metrics. Table 5.5 shows the results of this ablation analysis.

TABLE 5.4: Performance of individual Quality Metrics to predict grasp stability.

Selected	Classification trees		K-Nearest neighbors	
Metric	CrossVal	Test	CrossVal	Test
Q_{A1}	0.60 ± 0.03	0.62	0.61 ± 0.02	0.61
Q_{B1}	0.57 ± 0.00	0.59	0.57 ± 0.02	0.58
Q_{B2}	0.57 ± 0.00	0.59	0.57 ± 0.01	0.56
Q_{B3}	0.57 ± 0.00	0.59	0.63 ± 0.02	0.64
Q_{C2}	**0.75 ± 0.03**	**0.72**	**0.73 ± 0.02**	**0.71**
Q_{D1}	0.57 ± 0.00	0.59	0.57 ± 0.03	0.58
Q_{D2}	0.57 ± 0.00	0.59	0.57 ± 0.03	0.59

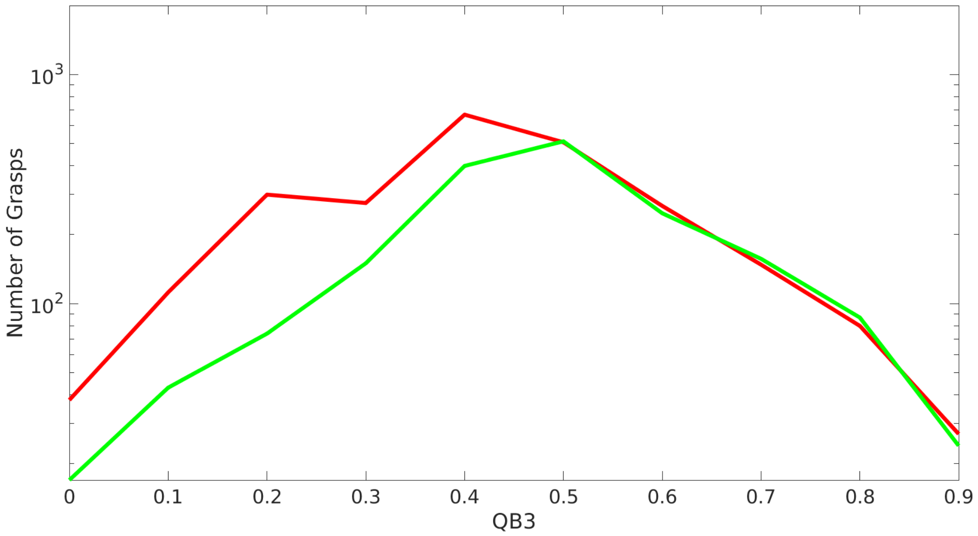


FIGURE 5.5: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{B3}. Histogram with log scale of y-axis.

TABLE 5.5: Ablation analysis on the use of Quality Metrics to predict the stability of grasps.

Discarded Metric	Classification trees CrossVal	Classification trees Test	K-Nearest neighbors CrossVal	K-Nearest neighbors Test
Q_{A1}	0.73 ± 0.02	0.74	0.75 ± 0.02	0.76
Q_{B1}	0.74 ± 0.02	0.75	0.78 ± 0.02	0.76
Q_{B2}	0.74 ± 0.02	0.74	0.76 ± 0.02	0.77
$\mathbf{Q_{B3}}$	0.72 ± 0.02	0.73	**0.76 ± 0.03**	**0.80**
Q_{C2}	0.69 ± 0.03	0.70	0.72 ± 0.02	0.72
Q_{D1}	0.74 ± 0.02	0.74	0.75 ± 0.02	0.76
Q_{D2}	0.74 ± 0.02	0.76	0.77 ± 0.02	0.77

Discussion

In this section has been studied the reliability of different machine learning algorithms for predicting the stability of a grasp. A database with more than 4699 grasps labeled by humans as *Stable* or *Unstable* has been evaluated using the quality metrics listed in Table 5.1. With this data, we analyzed the direct relation between independent *Quality Metrics* and the human-evaluated stability of the grasp.

Table 5.2 shows the results for the different learning models using the seven quality metrics. Four types of models obtain the best results (*K-Nearest neighbors*, *Classification trees*, *Gaussian Process* and *Neural Networks*). The column presenting results when the data is split by objects is a good indicator of the generalization capability of the models. Here, a *Gaussian Process* performs best. An additional comment comes from the comparison between the cross-validation performance on the training set and the performance on the test set. The values are close in general, indicating that overfitting is well controlled. In general, non-parametric and non-linear methods seem to capture the structure of the data best. As expected from figures 5.2 to 5.8, assuming conditional independence of the input quality metrics (Naive Bayes) leads to the worst performance.

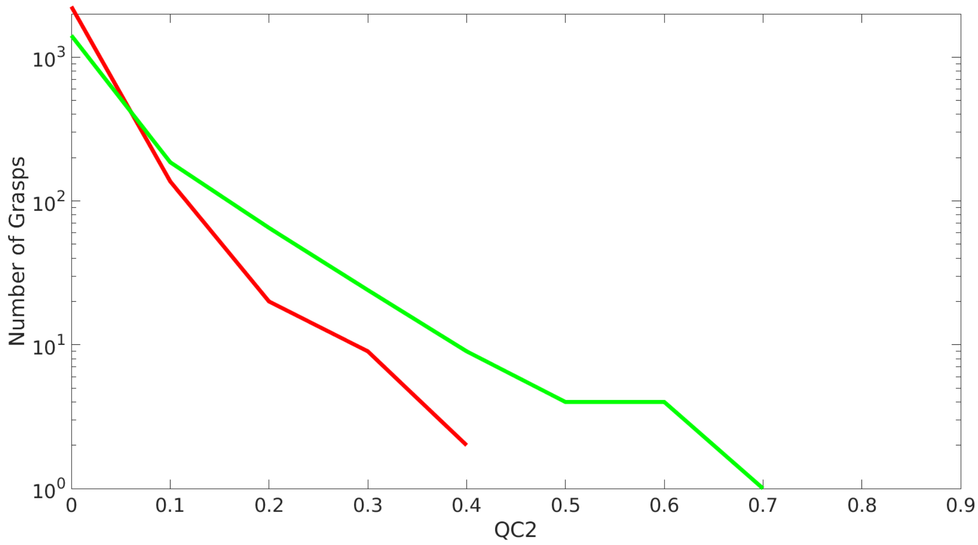


FIGURE 5.6: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{C2}. Histogram with log scale of y-axis.

Table 5.3 reports similar results as Table 5.2. Interestingly, this table also includes the performance of the *Physics Metric* (Kappler et al., 2015) on the test dataset. It clearly outperforms the models learned using the combination of the other seven quality metrics, and as we show below the performance of any individual metric. The drawback is that this metric is based on a dynamic simulation and is computationally much more expensive than computing the grasp quality metrics. For real time grasp prediction, the computation of a few quality metrics will be always faster than dynamic simulation.

Between the four more successful models some practical considerations in terms of computational effort can be made. Especially, prediction time is of interest in this context. Except from *K-Nearest Neighbors* and *Gaussian Processes*, all classification methods scale well with an increasing number of training data points.

More interesting is the analysis of the contributions of each individual metric in the prediction. This is shown in Tables 5.4 and 5.5. Two main results can be observed. First, only metric Q_{C2} shows a clear discriminative capability (72% using *Classification trees*) while the remaining metrics remain at the *null error rate*, 57%, or slightly above. These results highlight the complexity of the classification problem. Evidence of this can be found in figures 5.2 to 5.8 where, for each metric, the distribution of *Stable* and *Unstable* labels is represented. There is a strong overlap of these distributions making it hard to establish clearly differentiated clusters on any of the metrics.

Second, however, when combined, the predictive capability of the quality metrics increases to 12% above *null error rate* in the worst case (70% when only discarding Q_{C2} in Table 5.5). This suggests that the data is easier separable when represented in the combined space. This structure can be exploited by the classifiers (except by Naive Bayes).

To relate to results in similar chapters (Goins et al., 2014), we also report the results from a *Gaussian Process* model in Table 5.2. It shows similar performance to other learned models. However, we cannot draw further conclusions as the dataset used in (Goins et al., 2014) is very different (in terms of size, features and labels) and the results are reported differently.

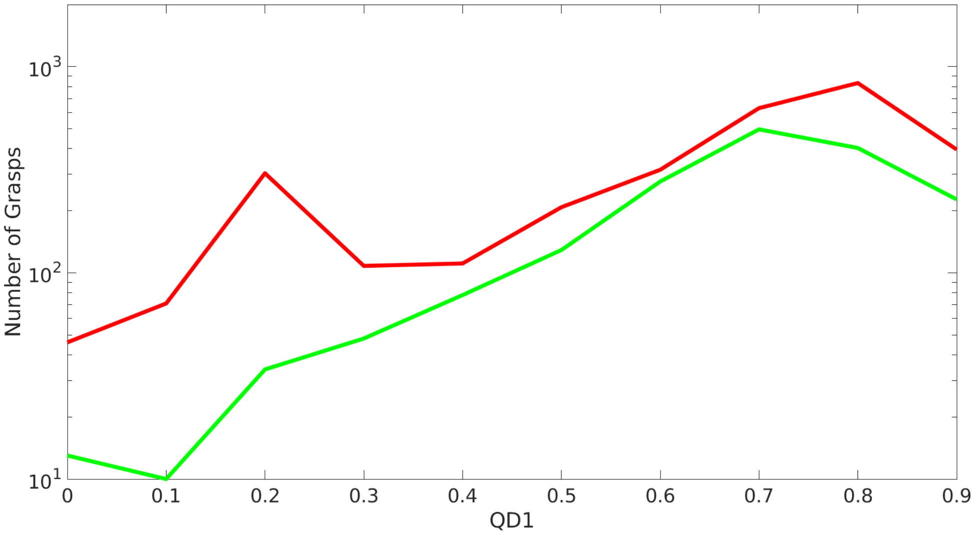


FIGURE 5.7: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{D1}. Histogram with log scale of y-axis.

The evaluation through metrics lays in the ability to obtain accurate data from simulation, that will be noisy in real world. Although the robustness of metrics is studied in other works Rubert et al., 2018; Leon et al., 2014, direct application of these results to real experiments should be taken carefully.

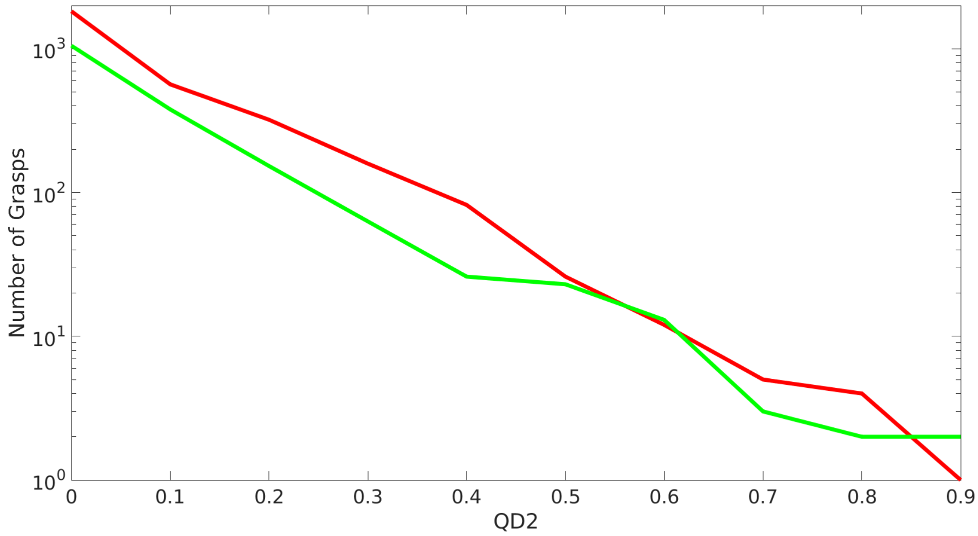


FIGURE 5.8: Distribution of *Stable* (green) and *Unstable* (red) grasps for metric Q_{D2}. Histogram with log scale of y-axis.

5.4.4 Conclusions

This section presented a study on the capability and relevance of quality metrics to predict the stability of a given grasp. A large database of human labeled samples has been employed to train several binary classifiers. The results have demonstrated the complexity of the problem but also that a combination of different metrics can deliver a performance up to 0.78; more than 0.20 over the *null error rate* and higher that any individual metric.

The contributions of the section are therefore: we analyzed classic quality metrics on a database of 4699 human labeled samples, more that in any other study; a set of classification methods has been used and compared, showing which of them performs best; and, the results have been quantitatively and qualitatively compared with previous results.

There are several future directions. In the current data set, human subjects have judged the stability of a grasp by looking at pictures of it that are taken from different viewpoints. Although this allows a very large amount of labels, they have not been verified by executing the grasps on a real robot and checking the outcome. A further study with samples executed on real robots is performed in the following section 5.5, to confirm the results of this work. Two main questions should be addressed: *Do the results presented here translate to real robot executions?*, and *How good are humans as predictors of grasp success?*.

5.5 Case study 2: Real grasp execution

In this study we address three main questions: first, are humans good grasp predictors? Second, can we find a combination of metrics able to predict the outcome of real grasp executions? And third, can we use *contextual information* to improve the prediction of grasp executions?

To answer these questions we perform a series of grasp using a real robot system. We score these grasp executions as *Stable* or *Unstable*. These results are compared to the evaluation using dynamic simulation and human labeled data. Finally we train and compare four classifiers, non-linear and non-parametric, using as input feature vector the evaluation with grasp metrics and *contextual information*.

This section is structured as follows, first a description on the notation and semantics used to detail the experiments, second the experiments and evaluation of grasps with a real robotic system are explained, and third we output the results of this grasp executions and the analysis with other grasp evaluation methods and machine learning algorithms.

5.5.1 Notation

For a better understanding of the experiments and results, the notation and definitions used are detailed next:

- **Pose**: A pose is defined as a 7 elements array: translation + rotation quaternion (x, y, z, w, ax, ay, az).
- **Grasp pose**: A grasp is defined by two different poses, a hand and an object. First, the pose of the gripper prior closing its fingers. Second, the pose of the object. A grasp is assumed to have been evaluated using real experiments, QM and the Physics Metric.
- **Experimental score**: A *experimental score* is the result of a grasp preformed in real world. Such score can be *Stable*, *Unstable* or *Unknown*[2].
- **Candidate grasp**: A candidate grasp is a grasp generated and simulated in OpenRave. Such grasps are evaluated using quality metrics and some have been evaluated by a physics metric and/or labeled by humans.
- **Quality metrics (QM)**: The quality metrics are a set of 7 independent metrics used to evaluate the quality of a grasp.
- **Physics metric**: The physics metric (Kappler et al., 2015) is a metric for evaluate grasps based in dynamic simulation. It is a binary metric: 1-Stable, 0-Unstable.
- **Human grasp label**: Candidate grasps are labeled by a human subject who predicts whether this grasp will be successful or not. Each candidate grasp is evaluated by multiple subjects and their scores are averaged. We call the final binary success label a human grasp label.
- **Experimental grasp**: An experimental grasp is a grasp performed in the real robot. Each experimental grasp has an associated candidate grasp and gravity

---

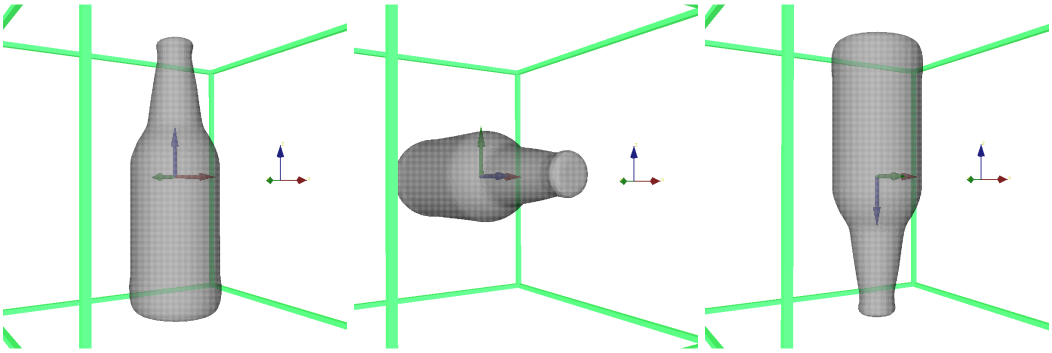
[2] Unknown grasps are discarded from the experimental results as they do not provide relevant information

vector. The presence of noise in the real robot, makes the candidate and experimental grasp poses slightly different. Each experimental grasp is evaluated in simulation using quality metrics and has an associated experimental score.

On each grasp execution, the gripper and object poses are recorded, and are later evaluated with *QM* using OpenHand (see Section 3.1). The experimental score associated to the experimental grasp is annotated by the human operator.

- **Grasp cluster**: A cluster of grasps is defined as a set of different experimental grasps which have the same candidate grasp and the same gravity vector. A cluster of grasps contains from 3 to 5 experimental grasps.

- **Gravity Vector**: The gravity vector represents the direction of gravity relative to the local object coordinate system. For example, a bottle in its canonical pose (Figure 5.9a) will have a gravity vector [0,0,-1], if it's upside down it will be [0,0,1]. Each experimental grasp has an associated gravity vector. Figure 5.9 shows different gravity vectors over a bottle.



(A) Gravity 0,0,-1 (x,y,z) (B) Gravity 0,-1,0 (x,y,z) (C) Gravity 0,0,1 (x,y,z)

FIGURE 5.9: Example of an object with three different gravity vectors. World axis is showed as reference for the normal gravity vector (0,0,-1).

- **Average Grasp**: For each cluster of grasps we define an average grasp. We use this average grasp as representative of the cluster. We also assume that it is similar to the candidate grasp that generated the cluster. Its quality is the averaged QM for each of the experimental grasps within the cluster. It has an experimental score associated as the most frequent value of the experimental scores of the experimental grasps. An average grasps has the same physics metric, human label and experimental score as its candidate grasp. The values of the QM are different but still similar.

  Figure 5.10 illustrates the relations between candidate grasps, experimental grasps, grasp clusters and average grasps.

- **Object Properties**: For each object we consider two properties: object volume and object weight.

- **Contextual Information**: We consider *Contextual Information* as external data related to the grasp, object or gripper, but not directly related to the contact points or quality metrics. In this study, we analyze how the *Gravity Vector* or the *Object Properties* could influence the outcome of grasp executions.

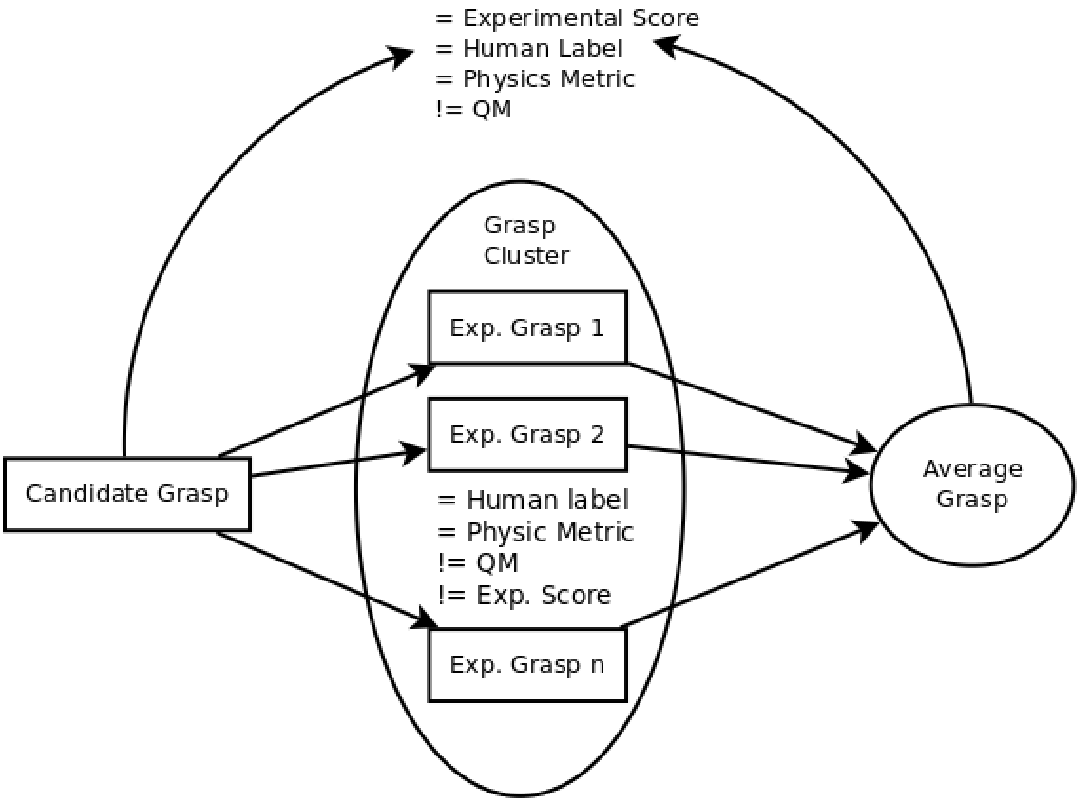

= Experimental Score
= Human Label
= Physics Metric
!= QM
Grasp
Cluster
Exp. Grasp 1
Exp. Grasp 2
= Human label
= Physic Metric
!= QM
!= Exp. Score
Exp. Grasp n
Candidate Grasp
Average
Grasp

FIGURE 5.10: Relationships between candidate grasps, experimental grasps, grasp clusters and average grasps.

- **Input Signal**: We define the input signal as a set or subset of QM, object properties and gravity vector related to an experimental grasp. This input signal is used by the classification methods to train or test different prediction models.

5.5.2 Methodology

To obtain the *experimental score* of a candidate grasp, it is necessary to evaluate it on a real robotic platform. In this experiment we use the *Apollo Robot* (Section 3.5). For the purpose of the experiments we will use only one arm and one hand of the Apollo system.

For performing the experiments we selected 9 different objects from the dataset of the previous study (Section 5.4). These objects are printed using 3D printers and cover different weights, dimensions and shapes. The objects are: bottle 1, bottle 2, toaster, camera, lemon, bowl 1, bowl 2, jar 1 and jar 2. Figure 5.11 shows the object models.

Experimental protocol

To perform the experiments, we apply the following experimental protocol:

Step 1: Initial setup. Move the arm/gripper to an initial pose. The gripper is placed initially in a top left position. This way the robot is able to plan and move the arm to a variety of poses over the table for grasping an object.

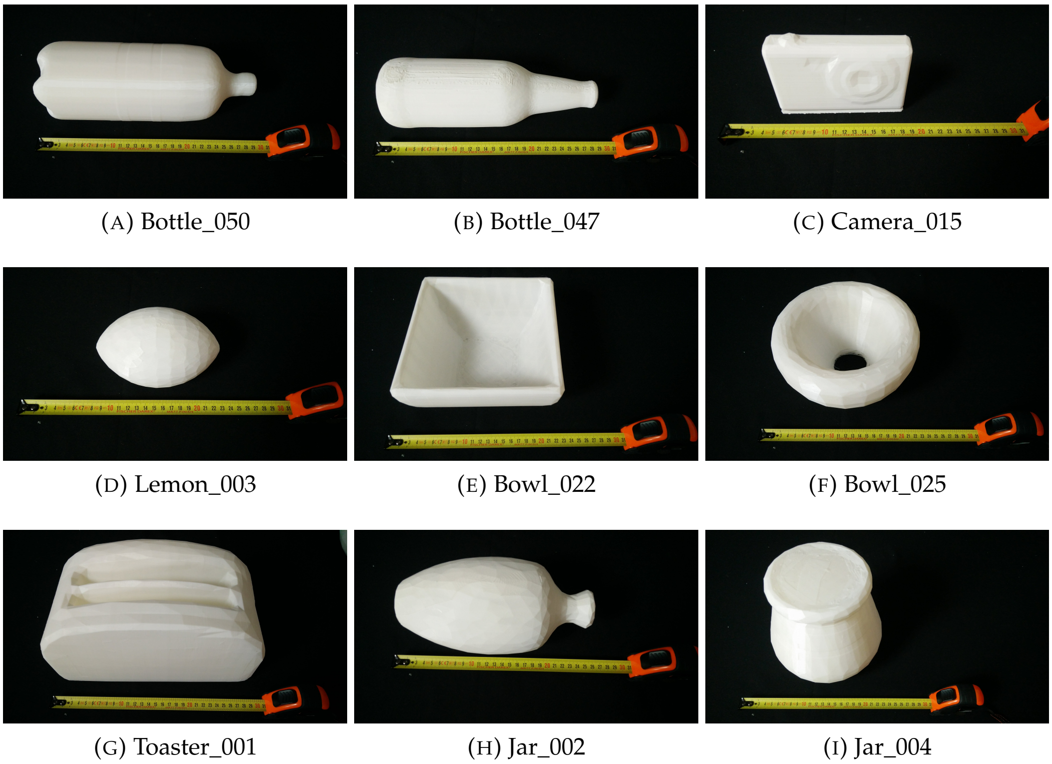

(A) Bottle_050 (B) Bottle_047 (C) Camera_015
(D) Lemon_003 (E) Bowl_022 (F) Bowl_025
(G) Toaster_001 (H) Jar_002 (I) Jar_004

FIGURE 5.11: Object models used for the experiments with the Apollo robot system. Objects are: (a)Bottle 1, (b)Bottle 2, (c)Camera, (d)Lemon, (e)Bowl 1, (f)Bowl 2, (g)Toaster, (h)Jar 1 and (i)Jar 2. Subcaptions show the object name in the dataset.

Step 2: Detect and track the object pose. To recognize and track the object we use the *Depth-Based Bayesian Object Tracking Library* (Wüthrich et al., 2013; Issac et al., 2016). This library implements two different algorithms for object tracking: a particle filter and a Gaussian filter. In our experiment the particle filter is used. This library allows to automatically detect the desired object and obtain its pose w.r.t. the robot platform.

Step 3: Select a candidate grasp to be tested. Among the different candidate grasps generated and evaluated with OpenHand, the human operator selects one to be tested. A candidate grasp has to be feasible to be selected. This means the robot or hand won't collide with the table when trying to acquire the grasp pose and the arm/planner is capable to move to that pose.

Also, when possible, the candidate grasps are evaluated using different object poses or orientations. In these cases, although it is the same candidate grasp, it is considered as a new *grasp cluster*, performing a new series of experimental grasps. This way, a candidate grasp could be evaluated in up to 6 different object poses (*gravity vectors*), each executed for at least three different trials.

Depending on the object shape and environment restrictions, some candidate grasps are evaluated with only one gravity vector, while others are evaluated with different gravity vectors. There is no specific protocol to set whether a grasp is evaluated with different gravity orientations. It depends on the human operator decisions whether a candidate grasp is evaluated with different gravity vectors.

Step 4: Move the arm/gripper to the grasp target pose. Once we have selected a feasible grasp, the planner moves the arm/gripper to desired grasp target pose. In case the arm hits the object during its movement or the plan fails to achieve the goal gripper pose, the grasp try is considered as *Not Valid* or *Unknown* and a new experimental grasp is attempted.

Step 5: Close the hand. If the gripper achieves the desired grasp target pose, the robot starts closing its fingers until a minimum strain is detected for each finger joint on the gripper. Each joint has a strain threshold to ensure the gripper applies enough pressure over the object and not just touches it.

Step 6: Move the gripper up: 15cm for small/medium objects, 25 cm for larger objects. Once the minimum strain on each joint is achieved, the joints of the fingers are blocked and the arm starts moving the gripper between 15 to 25cm up in the air. If the grasp is stable the object will be lifted. For unstable grasps, the object will not be lifted or will slip during this lifting event.

Step 7: Hold the hand in the lift pose for three seconds. Once the lift pose is achieved. The arm keeps this pose for 3 seconds, if the object remains in the gripper for this time, the grasp is considered *Stable*. If this time has expired and the object is not in the gripper, the grasp is considered *Unstable*. Figures 5.12 and 5.13 illustrates two examples of successful and unsuccessful experimental grasps. Both experimental grasps correspond to the same candidate grasp.

Step 8: Place the object on the table. After holding the object, the gripper is placed again on the desired grasp pose. A margin of +2cm is applied in the *Z axis* to avoid the object hitting directly the table surface.

Step 9: Release the object. After achieving the releasing pose, the fingers are opened and the object is finally released (in case the grasp was Stable).

Step 10: Move the arm/gripper to the initial pose. Finally the arm/gripper are moved to its initial pose (step 1).

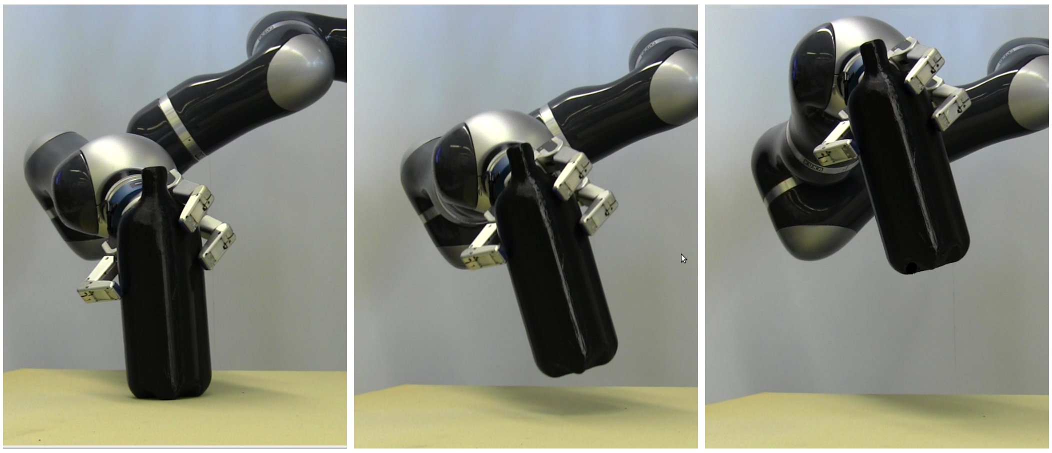


FIGURE 5.12: Example of a successful real grasp execution

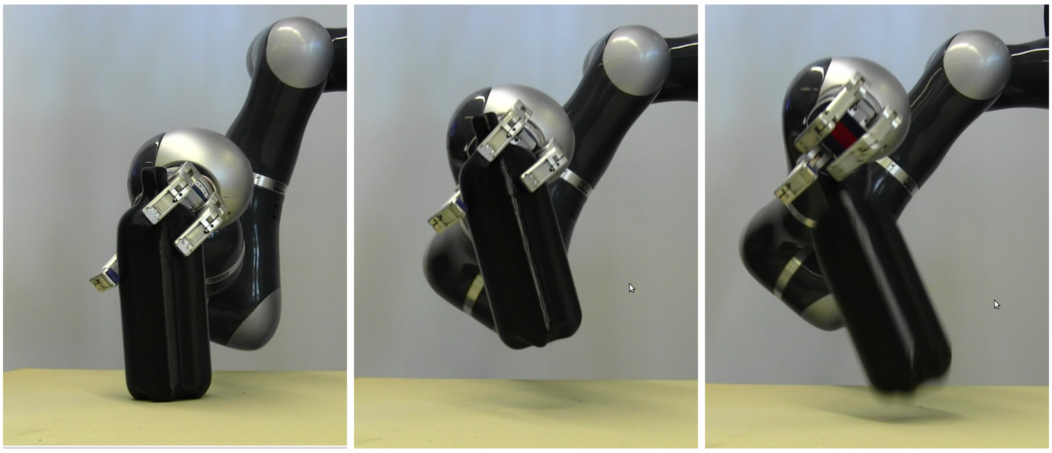


FIGURE 5.13: Example of an unsuccessful real grasp execution

Experimental score

For each experimental grasp, we score it as *success(1)*, *fail(0)* or *unknown(-1)*. If we get three trials with the same result we stop the series of trials and move to the next candidate grasp to be evaluated. In total, 1349 experimental grasps are performed. On average, 50 candidate grasps are tested and 150 different experimental grasps are performed with each object.

Exceptions: There are some circumstances which may diverge the experimental protocol: the arm/gripper hits the object during the movement, the object pose is unstable prior to grasping, the object falls while the hand closes, the planner moves to a wrong pose or the object tracking fails/loses the object. Figure 5.14 illustrates an example of a *Not Valid* try. In this case the motion planner moves the arm too close to the object pose and the gripper hits the object, making it fall before closing its fingers.

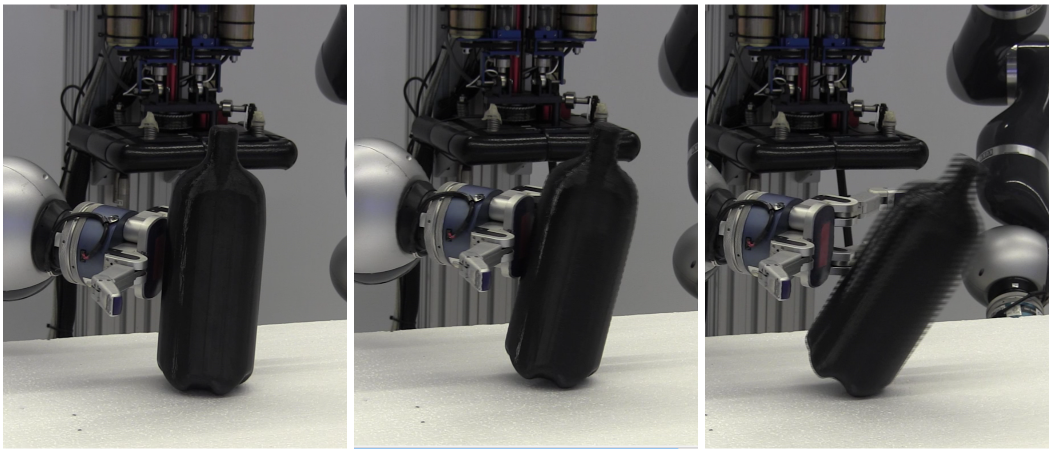


FIGURE 5.14: Example of a failed try. The gripper throws the object when approaching to the grasp pose.

In all these cases, the experimental grasp was considered as *Not valid* and thus, discarded. From the 1349 experimental grasps performed, 243 (18%) were discarded. Our final dataset contains 1106 *Stable/Unstable* experimental grasps. All of them are evaluated with *QM*, 830 were evaluated with the physics metric, 600 have the human label and 324 have both, the human label and the physics metric (see Figure 5.15).

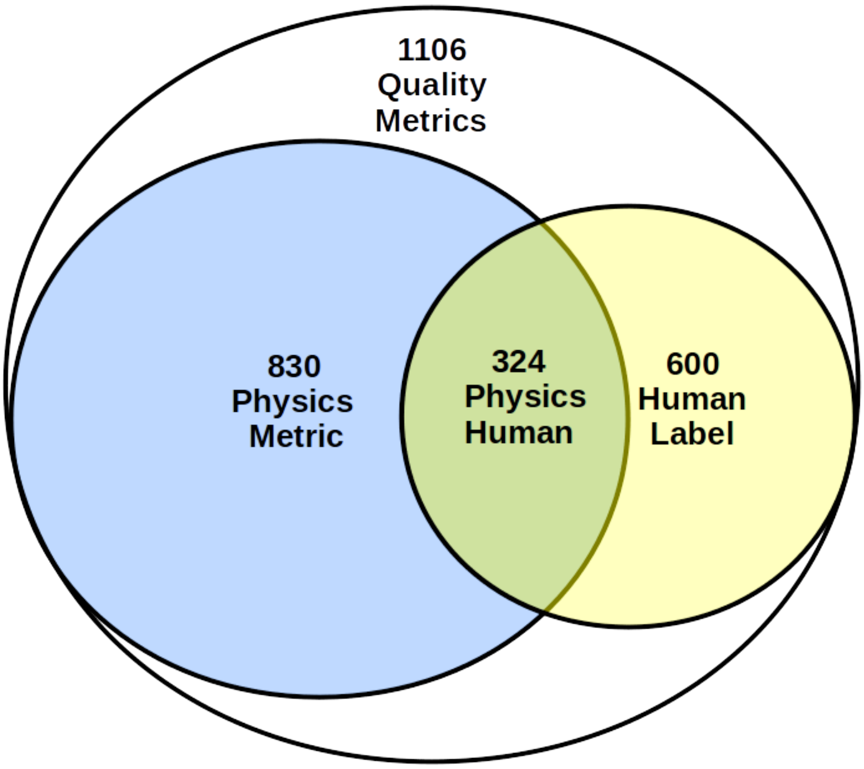


FIGURE 5.15: Distribution of grasps evaluated with quality metrics, physics metric and human label.

Input feature vector and dataset

For defining the input signals used by the classification methods, we will use different data sets. These data sets depend on the *type of grasp* used: experimental grasp, grasp clusters or average grasps; and how we select or filter them. In all our datasets we split the data in Train and Test with 80% and 20% of the grasps respectively. We will define 5 different data sets:

- **FS** (*full set*): This dataset uses the experimental grasps and the data is split randomly between train and test.
- **CS** (*cluster set*): Experimental grasps are split according to different clusters of grasps. This is, *held-out* clusters in the test test. Thus, the prediction models are tested with unseen clusters of experimental grasps.
- **SS** (*synthetic set*): Only the averaged grasps are used. It is split randomly between train and test as in FS.
- **GS** (*gravity set*): This set contains only clusters of grasps with same candidate grasp but different gravity vectors. I.e. lets say we have five clusters of grasps (A.B,C,D,E) with candidate grasp (1,2,1,3,3) respectively, Only clusters A,C,D and E are selected for this data mode. The data is splitted randomly among different tries as in *FS*.

- **OGS**(*XYZ*) (*one gravity set*): This sets contains only experimental grasps with the same gravity vector. I.e. if we have only five grasps (A,B,C,D,E) with 1D gravity vectors [1,0,1,-1,1]. We'll have three different OGS: Gravity 1 with grasps (A,C,E), Gravity 0 with grasp (B) and Gravity -1 with grasp (D). As our gravity vectors are 3D and have 3 different values, this generates could generate to 27 seven different OGS modes of training. In our dataset, there are only 10 different *OGS* models. Data is split randomly between train and test.

Table 5.6 shows the distribution of grasps on each dataset.

TABLE 5.6: Distribution of *Stable-Unstable* grasps among the different filtered datasets. For the OGS the name indicates the gravity axis. OGSX: gravity only in axis X. OGSX-Y gravity in axis X and -Y. Only OGS models which contain grasp tries are included in the table.

	Stable	Unstable	N° of Grasps
FS	582	524	1106
CS	582	524	1106
SS	128	215	343
GS	194	181	375
OGS-X-Y	9	14	23
OGS-X	14	32	46
OGS-XY	6	15	21
OGS-Y	11	21	32
OGS-Z	272	202	474
OGSZ	103	120	223
OGSY	10	8	18
OGSX-Y	3	12	15
OGSX	18	32	50
OGSXY	11	6	17

Finally, for each one of the data sets 4 different types of input signals are considered:

- **QM**: 7 Dimensional input signal containing the evaluation of grasps with quality metrics.
- **QMO**: 9-D input signal with the evaluation with QM and the Object Properties (object's volume and weight) .
- **QMG**: 10-D input signal with the evaluation with QM and the Gravity Vector.
- **QMOG**: 12-D input signal with the evaluation with QM, the Object Properties and the Gravity Vector.

Classification methods

From the results obtained in the study with Human labels, four different classifiers are tested to generate the prediction model:

- Classification trees
- Neural networks

- k-Nearest neighbors
- Gaussian process

These classifiers correspond to those non-linear and non-parametric algorithms which showed better prediction results in the previous case study, Section 5.4.

5.5.3 Results

Experimental score vs. human labels vs. physics score

First, we will compare the capability of the physics metric and the human labeling to predict the outcome of real executions of grasps. As the evaluation with human labels and physics metric was done with the specific grasp poses of the candidate grasps, we use only the averaged grasps (Figure 5.10). This data set contains 343 grasps.

These results are shown in Table 5.7. A more detailed analysis on the predictive capability of Humans and Physics metric is shown in Table 5.8. It shows different metrics to report the score of a prediction model. These tables show for the *success* and *fail* labels, the scores on precision, recall and f1-score.

TABLE 5.7: Predictive ability between Human labels, Physics Metric and Real Experiments. Results shown are accuracy score.

	Humans	Physics	Real Exp.
Humans	1.00		
Physics	**0.85**	1.00	
Real	0.61	0.64	1.00

TABLE 5.8: Score report on the use of humans and Physics to predict grasp success on real experiments. Upper table shows results for human labels. Lower table shows results for physics metric.

	Human Labels		
Real Score	Precision	Recall	f1-Score
0	**1.00**	0.03	0.05
1	0.60	**1.00**	**0.75**
avg/total	**0.76**	0.61	0.47

	Physics Metric		
Real Score	Precision	Recall	f1-Score
0	**0.78**	0.53	0.63
1	0.55	**0.80**	0.65
avg/total	0.68	0.64	0.64

Table 5.9 shows the capability of humans and the physics metric to predict the outcome of real grasp executions according to specific gravity orientations. it employs the *OGS* datasets, we selected only averaged grasps with specific gravity vectors.

TABLE 5.9: Predictive ability of Human Labels and Physics metric according to specific gravity vectors. Number of averaged grasps evaluated by humans and physics metric is also provided. Only two gravity vectors (green) have enough data points to provide reliable results.

Real	vs. Human	vs. Physics	N° of Grasps
OGS-X-Y	0.33	**0.75**	3 / 4
OGS-X	**0.86**	0.54	7 / 13
OGS-XY	0	0.6	2 / 5
OGS-Y	0.5	**0.88**	2 / 8
OGS-Z	0.66	0.64	**56 / 146**
OGSZ	0.54	0.68	**35 / 77**
OGSY	0.67	0.2	3 / 5
OGSX-Y	0	0.67	2 / 3
OGSX	**0.71**	0.45	7 / 11
OGSXY	1	1	2 / 3

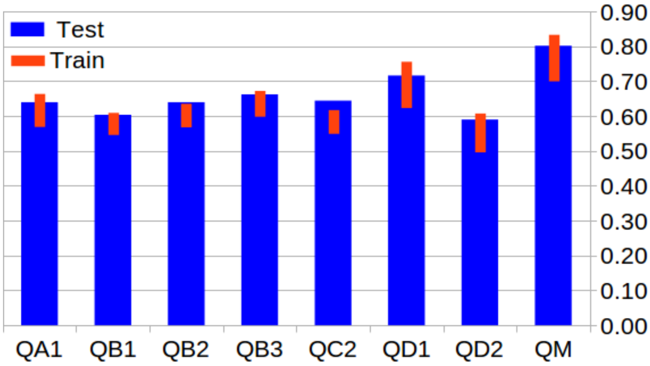
This table compares if the predictive capability improves to specific gravity vectors (or *object poses*). The *N°of Grasps* shows how many averaged grasps were evaluated by humans and the physics metric. As is shown in this table, only two gravity vectors (OGS-Z, OGSZ) have enough grasps to provide reliable results.

Classification algorithms

First, the predictive capability of individual QM is tested as benchmark for the models trained in next experiments. Table 5.10 shows the performance of each metric. The dataset used for this experiment is *FS* and *QM* as input signal.

TABLE 5.10: Predictive ability of individual quality metrics. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar).

Metric	CrossVal	Test
Q_{A1}	0.62 ± 0.05	0.64
Q_{B1}	0.58 ± 0.03	0.60
Q_{B2}	0.60 ± 0.03	0.64
Q_{B3}	0.64 ± 0.04	0.66
Q_{C2}	0.58 ± 0.03	0.64
$\mathbf{Q_{D1}}$	$\mathbf{0.69 \pm 0.07}$	**0.72**
Q_{D2}	0.55 ± 0.06	0.59
QM	0.77 ± 0.07	0.80

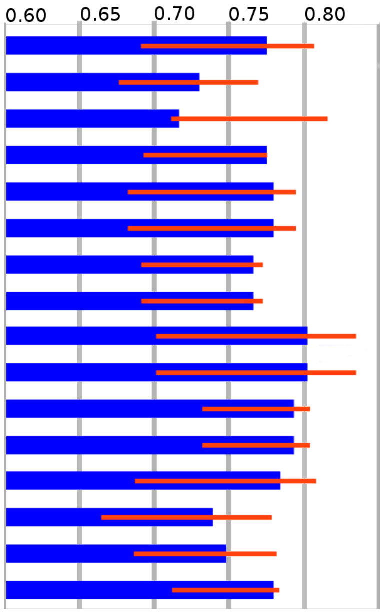


Next, Table 5.11 shows the results training the classifiers using the *FS dataset*. The 4 types of input signals are also tested. The *FS dataset* uses the experimental grasps and the data is splitted randomly between train and test.

Table 5.12 shows the classifiers trained with the *CS dataset* and different Input Signals. In the *CS dataset* the experimental grasps are split according to different clusters of grasps. The prediction models are tested with unseen clusters of experimental grasps.

TABLE 5.11: Classification methods trained with the *FS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar).

Classifier	Input	CrossVal	Test
Classification Trees	qm	0.75 ± 0.06	0.77
Classification Trees	qmo	0.72 ± 0.05	0.73
Classification Trees	qmg	0.76 ± 0.05	0.72
Classification Trees	qmog	0.73 ± 0.04	0.77
Gaussian Process	qm	0.74 ± 0.06	0.78
Gaussian Process	qmo	0.74 ± 0.06	0.78
Gaussian Process	qmg	0.73 ± 0.04	0.77
Gaussian Process	qmog	0.73 ± 0.04	0.77
K-Nearest Neighbors	qm	0.77 ± 0.07	0.80
K-Nearest Neighbors	qmo	0.77 ± 0.07	0.80
K-Nearest Neighbors	qmg	0.77 ± 0.04	0.79
K-Nearest Neighbors	qmog	0.77 ± 0.04	0.79
Neural Networks	qm	0.75 ± 0.06	0.78
Neural Networks	qmo	0.72 ± 0.06	0.74
Neural Networks	qmg	0.73 ± 0.05	0.75
Neural Networks	qmog	0.75 ± 0.04	0.78



Results of classification using the *SS dataset* are shown in Table 5.13. In the *SS dataset* only the averaged grasps are used. It is split randomly between train and test as in the *FS dataset*.

Table 5.14 shows the classification using the *GS dataset*. This dataset contains only clusters of grasps with same candidate grasp but different gravity vectors. The data is split randomly among different tries as in *FS*.

For the *OGS dataset* Table 5.15 summarize the results using different classifiers and input signals with the specific gravity vectors. This datasets contains only grasp trials with the same gravity vector. The data is split randomly between train and test.

Figure 5.16 shows an example of the distribution of grasps for metrics Q_{A3}, Q_{B1}, Q_{B2} and Q_{B3} for the *FS dataset* and *SS dataset*. These graphs illustrate the dispersion and space of the grasps within the metrics.

Finally, a wide study trying to find the best combination of *QM* is performed. In this study we vary our input signal selecting all possible combinations of quality metrics. Table 5.16 shows the best results obtained selecting different combinations of metrics.

Appendix B.1 presents more results on the study of different combinations of metrics to predict the outcome of real grasp executions.

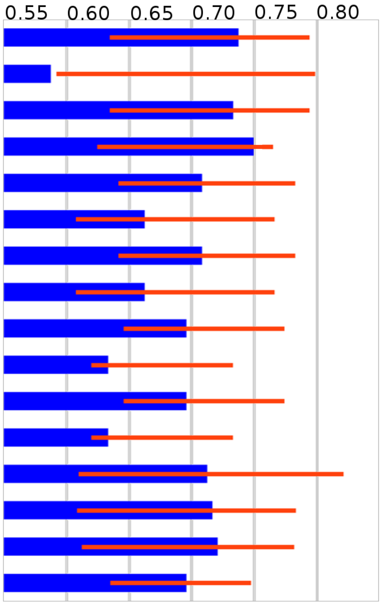
5.5.4 Discussion

Are humans good grasp predictors?

Results in Tables 5.7 and 5.9 showed neither humans nor physics metric are good choices for predicting the outcome of real grasp executions. Both methods showed good performance ro predict the failure of a grasp. Humans are perfect at this with a precision score of 100%. However, they lack of reliability for predicting the success of a grasp as they have a lot of false positives.

TABLE 5.12: Classifiers trained with the *CS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar).

Classifier	Input	CrossVal	Test
Classification Trees	qm	0.71 ± 0.08	0.74
Classification Trees	qmg	0.70 ± 0.10	0.59
Classification Trees	qmo	0.71 ± 0.08	0.73
Classification Trees	qmog	0.70 ± 0.07	0.75
Gaussian Process	qm	0.71 ± 0.07	0.71
Gaussian Process	qmg	0.69 ± 0.08	0.66
Gaussian Process	qmo	0.71 ± 0.07	0.71
Gaussian Process	qmog	0.69 ± 0.08	0.66
K-Nearest Neighbors	qm	0.71 ± 0.06	0.70
K-Nearest Neighbors	qmg	0.68 ± 0.06	0.63
K-Nearest Neighbors	qmo	0.71 ± 0.06	0.70
K-Nearest Neighbors	qmog	0.68 ± 0.06	0.63
Neural Networks	qm	0.72 ± 0.11	0.71
Neural Networks	qmg	0.70 ± 0.09	0.72
Neural Networks	qmo	0.70 ± 0.08	0.72
Neural Networks	qmog	0.69 ± 0.06	0.70



The analysis on specific gravity orientations didn't prove humans seems to evaluate better when specific conditions are applied. It is important to note, there was no specific gravity orientation when humans had to label the grasps. So it is unclear which assumptions they made during the experiments. Only two orientations showed good results *OGS-X* and *OGSXY* but there are not enough data points in these models to provide strong results.

Physics metric seem to have a better behavior for specific gravity orientations. It is important to consider, the physics metric didn't take into account possible gravity forces when calculated. Alongside with the favorable orientations for humans, with *OGS-Y* and *OGS-X-*, the physics metric showed good performance. However, as with humans, there are not enough data to make convincing assumptions.

Can grasp quality metrics predict the outcome of real grasp executions?

Results on Table 5.10 showed individually, metrics are not good grasp predictors, only one metric Q_{D1} showed a good performance (72%). According to results on Table 5.11, it is possible to find a proper combination of metrics to predict with high precision (up to 80%) the outcome of real grasps executions.

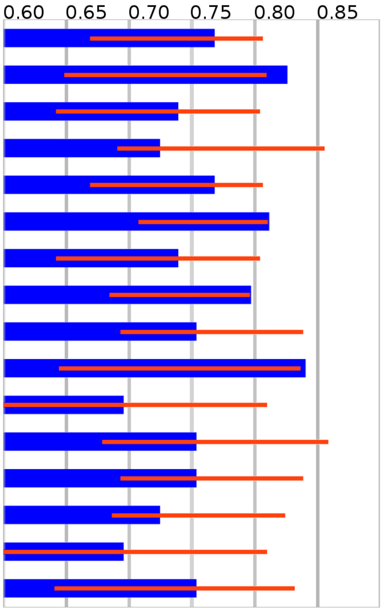
The study with the *CS dataset* showed lower values than the *FS dataset*. This could mean these experiments could be biased by the similarity between grasps within a cluster which leads to an over-fitting of the classification methods.

However, according to the graphs showed in Figure 5.16 for the *FS dataset*, is shown different cluster of grasps share the same dimensional space. Thus, two different clusters of grasps are not necessarily in different dimensional spaces and there is not such bias in the dataset. Although prediction results with the *CS dataset* are a bit worse, they are still inside the range of their standard deviation.

It is important to consider the possibility of not having enough candidate grasps. If we compare results with the *FS dataset* (Table 5.11) and *SS dataset* (Table 5.13), it is shown a model trained with the averaged grasps will also provide good results.

TABLE 5.13: Classifiers trained with the *SS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar).

Classifier	Input	CrossVal	Test
Classification Trees	qm	0.73 ± 0.08	0.83
Classification Trees	qmo	0.76 ± 0.05	0.81
Classification Trees	qmg	0.74 ± 0.10	0.84
Classification Trees	qmog	0.76 ± 0.07	0.72
Gaussian Process	qm	0.72 ± 0.08	0.74
Gaussian Process	qmo	0.72 ± 0.08	0.74
Gaussian Process	qmg	0.70 ± 0.11	0.70
Gaussian Process	qmog	0.70 ± 0.11	0.70
K-Nearest Neighbors	qm	0.74 ± 0.07	0.77
K-Nearest Neighbors	qmo	0.74 ± 0.07	0.77
K-Nearest Neighbors	qmg	0.77 ± 0.07	0.75
K-Nearest Neighbors	qmog	0.77 ± 0.07	0.75
Neural Networks	qm	0.77 ± 0.08	0.72
Neural Networks	qmo	0.74 ± 0.06	0.80
Neural Networks	qmog	0.74 ± 0.10	0.75
Neural Networks	qmg	0.77 ± 0.09	0.75



Graphs on Figure 5.16 showed although the *SS dataset* has less datapoints than the *FS dataset* (see Table 5.6) the distribution of points in its dimensional space is still wide and dispersed. This suggest the dataset is large enough and the selection of candidate grasps is representative of the population of feasible grasps.

Regarding the selection of quality metrics, if we compare the results on predictive capability of metrics from Tables 5.10 and 5.16, it is shown individual metrics have poor performance for predicting the stability of a grasp. However, it is easy to find a combination with high accuracy (>80%). These results showed also this combination does not rely in one metric, and different combinations of metrics are able to capture different and independent aspects of the grasp which are relevant to predict its success. An extend version of these tables is given in Appendix B.1.

Finally, results showed an upper bound in the prediction for the *FS dataset* with a 82% of maximum accuracy. This could be either a limit in the predictive capability of metrics, meaning that almost 20% of grasps cannot be explained by quality metrics. Or with this selection of candidate grasps it is not possible to have a better accuracy. To test which hypothesis is true, one would need to collect a larger dataset.

Can contextual information help improve the grasp prediction?

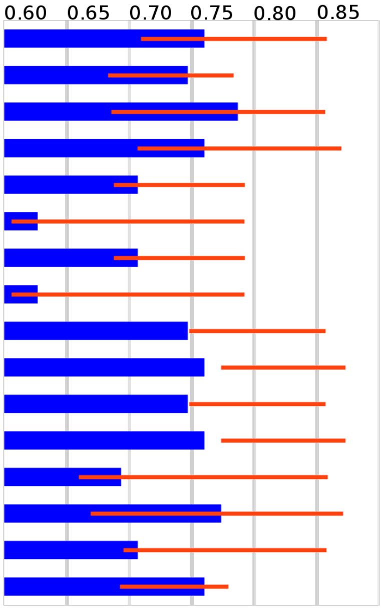
Our analysis with different classifiers and data models included also the comparison of the training when adding contextual information to the quality metrics. Experiments showed the grasp success varies with different gravity vectors.

Regarding the object properties, heavier objects seemed to be more difficult to grasp, but this could be also due to the selection of grasp tested. Another study addressing this feature should be done to analyze the influence of different object weights in the successfulness of a grasp.

A deeper analysis on the influence on different gravity orientations is shown in Table 5.15. In this case, classifiers trained with specific gravity orientations can better fit a model able to predict the outcome of real grasp executions. However, as stated

TABLE 5.14: Classifiers trained with the *GS dataset* and different Input Signals. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar).

Classifier	Input	CrossVal	Test
Classification Trees	qm	0.78 ± 0.07	0.76
Classification Trees	qmg	0.73 ± 0.05	0.75
Classification Trees	qmo	0.77 ± 0.09	0.79
Classification Trees	qmog	0.79 ± 0.08	0.76
Gaussian Process	qm	0.74 ± 0.05	0.71
Gaussian Process	qmg	0.70 ± 0.09	0.63
Gaussian Process	qmo	0.74 ± 0.05	0.71
Gaussian Process	qmog	0.70 ± 0.09	0.63
K-Nearest Neighbors	qm	0.80 ± 0.05	0.75
K-Nearest Neighbors	qmg	0.82 ± 0.05	0.76
K-Nearest Neighbors	qmo	0.80 ± 0.05	0.75
K-Nearest Neighbors	qmog	0.82 ± 0.05	0.76
Neural Networks	qm	0.76 ± 0.10	0.69
Neural Networks	qmg	0.77 ± 0.10	0.77
Neural Networks	qmo	0.78 ± 0.08	0.71
Neural Networks	qmog	0.74 ± 0.04	0.76



for the human labeling and physics metric, the lack of enough datapoints in some gravity orientations does not allow to make valid conclusions.

Finally, in our experiments it was shown the same candidate grasp can succeed or fail depending on the gravity orientation. However, this information seems to be useless in combination with quality metric and classification methods.

TABLE 5.15: Summary of the training with the *OGS datasets*. Reported best results with each variation of the gravity vector. For *OGSY and OGSXY datasets* there are not enough datapoints to perform a proper 10-fold CV. Table shows the classification accuracy. Graphs show the training ± std (red line) and the accuracy on the Test set (blue bar).

Gravity	CrossVal	Test
OGS-X-Y	0.87 ± 0.31	1.00
OGS-X	0.96 ± 0.12	1.00
OGS-XY	1.00 ± 0.00	1.00
OGS-Y	0.80 ± 0.31	0.86
OGS-Z	0.77 ± 0.07	0.80
OGSZ	0.79 ± 0.10	0.80
OGSY	NA	NA
OGSX-Y	0.85 ± 0.32	1
OGSX	0.83 ± 0.15	0.7
OGSXY	NA	NA

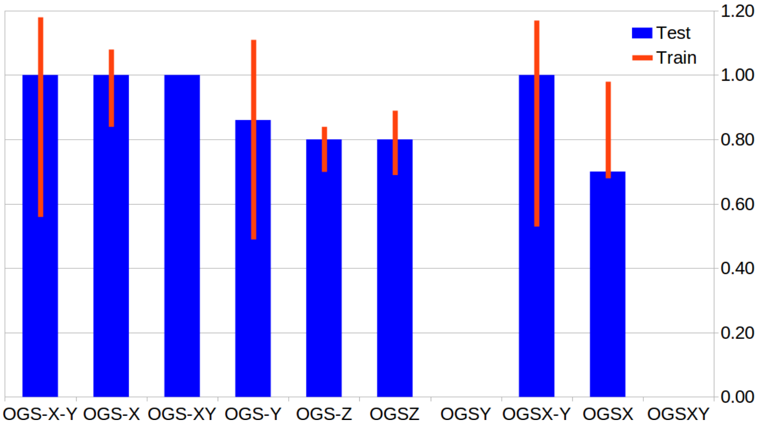
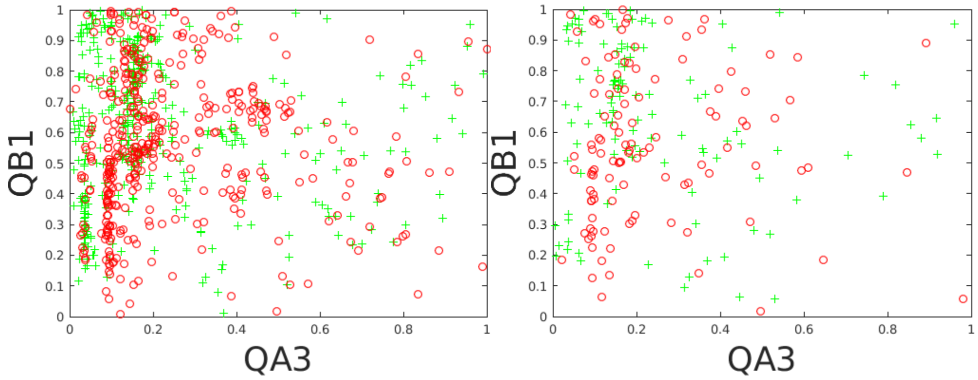
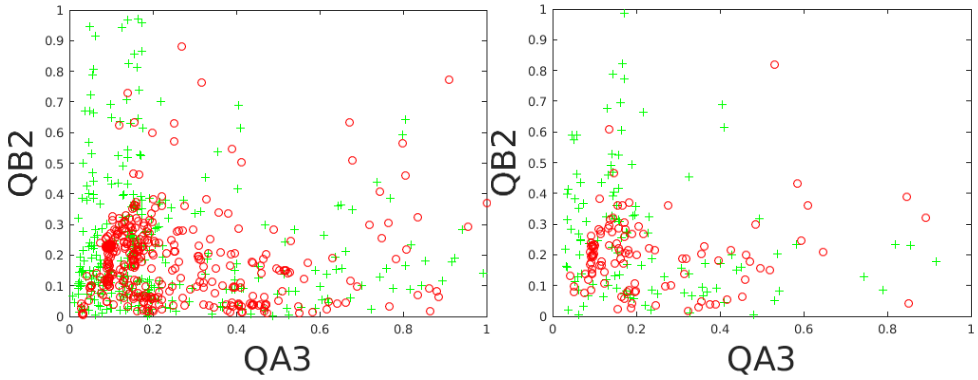


TABLE 5.16: Summary of best training results obtained using different combinations of metrics as input feature vector. The training was done using the *FS model* and only with *KNN* classifier

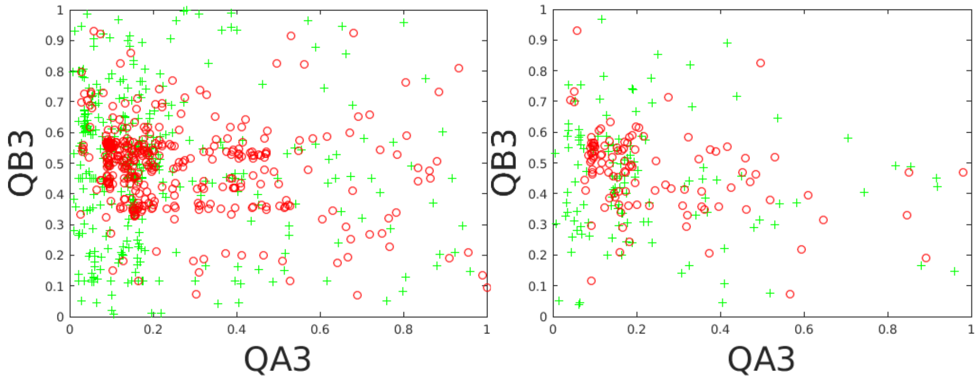
Classifier	Metrics Used	CrossVal	Test
K-Nearest Neighbors	(Q_{A1}, Q_{D1})	0.73 ± 0.07	0.78
K-Nearest Neighbors	(Q_{A1}, Q_{C2}, Q_{D1})	0.74 ± 0.07	0.81
K-Nearest Neighbors	(Q_{B1}, Q_{B3}, Q_{D1})	0.75 ± 0.03	0.80
K-Nearest Neighbors	$(Q_{A1}, Q_{C2}, Q_{D1}, Q_{D2})$	0.75 ± 0.05	0.81
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B3}, Q_{D1}, Q_{D2})$	0.76 ± 0.06	0.81
K-Nearest Neighbors	$(Q_{B1}, Q_{B2}, Q_{B3}, Q_{C2}, Q_{D1})$	0.77 ± 0.04	0.81
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B2}, Q_{C2}, Q_{D1}, Q_{D2})$	0.78 ± 0.04	0.81
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B2}, Q_{B3}, Q_{C2}, Q_{D1}, Q_{D2})$	0.77 ± 0.07	0.80



(A) Map of Values in *FS model* (B) Map of Values in *SS model*



(C) Map of Values in *FS model* (D) Map of Values in *SS model*



(E) Map of Values in *FS model* (F) Map of Values in *SS model*

FIGURE 5.16: Maps of values for metrics QA2 vs QB1, QB2, QB3. Graphs on left show the FS dataset, Graphs on the right column show the SS dataset. Graphs for the *FS dataset* contain 1106 datapoints, graphs in the *SS dataset* contain 343 datapoints.

5.6 Case study 3: Metrics limitations and contextual information

On the previous study, it has been shown metrics alone are not enough to predict the success of a grasp. Contextual information as gravity orientation or object properties could provide relevant information to help improve the prediction of grasp executions in the real world. Also, experiments showed a candidate grasp not always fails or succeeds, the same grasp could sometimes work and next fail. Its success could depend on the gravity orientation or an accurate gripper pose.

In this study we analyze the influence of the object weight and gravity orientation in the success of a grasp. For analyzing the influence of the object weight we will perform new experiments with the *Tombatossals* (section 3.5) robot platform and two different object models with different weights: light and heavy.

We will consider a 3-categories scale for the evaluation of a grasp execution: *Robust*, *Fragile* and *Futile*.

A grasp is considered **Fragile** if its success or failure is dependent in the gravity orientation, object's weight or is highly dependent in the specific contact points.

A grasp is considered **Robust** if it always succeeds, independently of the gravity orientation, object properties or having accurate contact points.

A grasp is considered **Futile** if it always fails, independently of the gravity orientation, object properties or having accurate contact points.

This study has two main objectives. First, establish if a grasp is influenced by the object weight or gravity orientation and how to use this information in combination with quality metrics. And Second, analyze whether this 3-categories scale is a better approach to label grasp executions.

5.6.1 Methodology

This study will consider two different datasets of grasps. First, the dataset with grasps from the Apollo system obtained in the previous study. This dataset contains a large database of grasps executed with different gravity orientations. Second we will use the *Tombatossals* robot platform to generate a new dataset of grasps. This dataset will contain grasps from two different object models with two different weights each.

Although results are shown mainly with the *Tombatossals* platform, the *Apollo* grasp dataset is used as benchmark to evaluate the performance of the trained models and the new proposed prediction model.

The methodology and experimental protocol applied will be the same described in section 5.5.2. There are also some differences in this study, first, the experimental platform will be Tombatossals(3.5). Second, only two different object models are used: *Bottle_050* and *Toaster_001*. These objects will have two versions, varying its weight: light and heavy.

Third, we consider a different scale to evaluate real grasp executions:*Robust*, *Fragile* and *Futile*. The real executed grasps are initially labeled as in the previous study: *Stable* or *Unstable*, and a post-processing is done later in this study to identify the *Robust*, *Fragile* and *Futile* grasps.

Unlike in the previous study, where clusters of grasps are divided according to different gravity orientations, in this study a cluster of experimental grasps only require to have the same candidate grasp. Thus, experimental grasps with different object weights are considered also belonging to the same cluster of grasps.

Grasp evaluation

Grasp executions on the real robot are initially scored as *Stable* or *Unstable*. After all the experiments are done, a post-processing is done for evaluating grasps using the 3-categories scale. If for a *cluster of grasps* there are different experimental scores, all the grasps in the cluster are labeled as *Fragile*, otherwise they are considered as *homogeneous* and we keep their original score.

A *homogeneous* grasp could be either *Robust* or *Futile*. The term *homogeneous* denotes a candidate grasp will always succeed or fail independent of the object's weight or the gravity orientation.

According to this, the classifiers and training applied consider a 3-Dimensional space where grasp trials will be scored as *Robust*, *Futile* and *Fragile*.

Classification models

As in the previous study, we will formulate the problem as a classification problem and train a classifier to predict the outcome of real grasp executions. In this experiment we will use only *k-NN* and *Classification Trees*, as they showed the best performance in previous studies. Initially, we consider three categories to train our models: *Robust*, *Fragile* and *Futile* grasps.

Object models

Two object models are used (Figure 5.17): *Bottle 1* and *Toaster*. These object models are made using 3D printing technology with two different infill values. This allows to make objects with different weights and robustness. Table 5.17 summarize the properties of the objects used in this study.

TABLE 5.17: Object weights (grams) of different printed models

Object	Light	Heavy
Bottle 1	181.5	726.5
Toaster	527	885.5

Grasps dataset

First, we generate randomly 100 grasps per object model, following the methodology described in section 3.4. These grasps candidates are executed with the real platform using the *light* and *heavy* object models. The same grasp candidates are evaluated using the two weights of the objects.

Then, these grasps tries are post-processed and the final *Robust*, *Futile* or *Fragile* label is settle. Our final dataset contains more than 600 experimental grasps distributed as shown in Table 5.18.

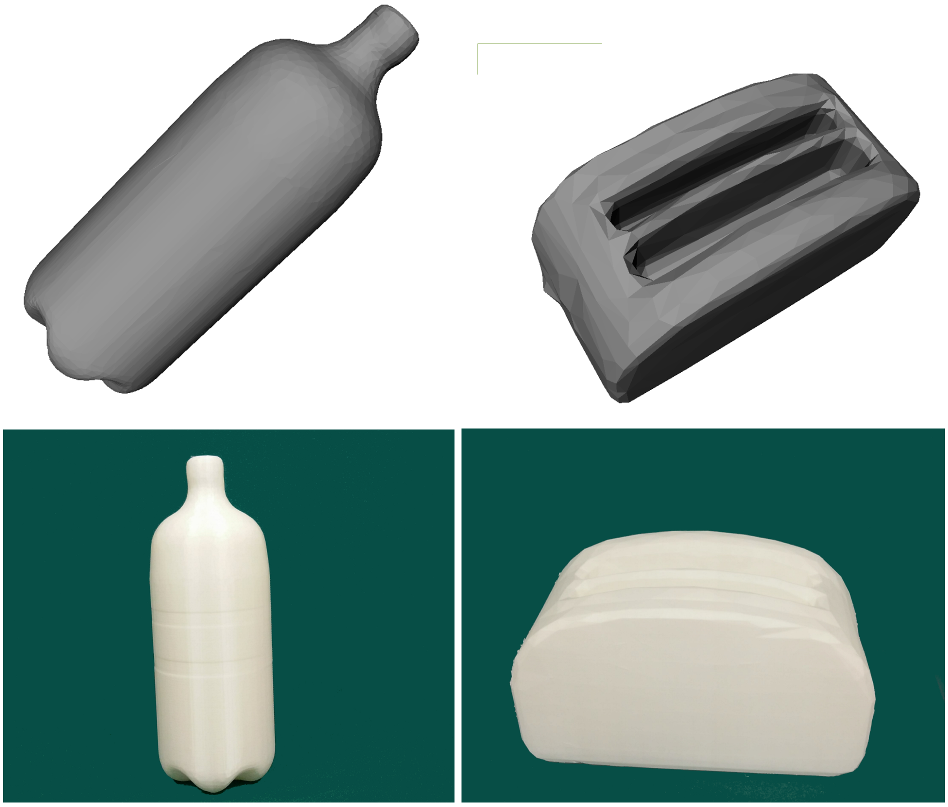


FIGURE 5.17: Object models used for the experiments with Tombatossals. Upper row 3D Models. Lower row 3D printed objects. Left column: Bottle 1, Right column: Toaster.

5.6.2 Results

First, we will compare the classic binary classification against the 3-categories scale. We will use both, the grasp dataset with Tombatossals and the grasps with Apollo. Second, we will study the grasp capability of individual metrics using this 3-categories scale.

Tombatossals grasp dataset

First, we will train a prediction model using the original methodology detailed in the previous study (section 5.5) and compare it with a 3-categories scoring method. Results are shown in Table 5.19. Results shown the accuracy of the classifier with the train dataset following a 10 fold cross-validation, and the results against the test set.

TABLE 5.18: Dataset distribution of grasps and scores among different objects.

	Binary classification		3-categories scale		
	Stable	Unstable	Robust	Futile	Fragile
Toaster	127	204	48	132	151
Bottle 1	156	129	78	39	168
Total	283	333	126	171	319

Grasps on the dataset are randomly split between train and test, as in the *FS dataset* from the previous study (Section 5.5).

TABLE 5.19: Classification results using binary score and 3-categories scale

	Binary Score		3-categories scale	
Classifier	Train ± Std	Test	Train ± Std	Test
K-Nearest Neighbours	0.70 ± 0.08	0.72	0.83 ± 0.04	0.83
Classification Trees	0.68 ± 0.05	0.65	0.86 ± 0.04	0.86

Table 5.20 shows the results of classifiers using a test data set *held out* during training. In this case 20% of the candidate grasps for each object are used as test. With this method, the trained classifiers are tested with new grasp types.

TABLE 5.20: Classification results using held-out grasps method for testing the prediction models with new and unseen grasp types.

Classifier	Train ± Std	Test
K-Nearest Neighbors	0.75 ± 0.15	0.75
Classification Trees	0.78 ± 0.11	0.81

Apollo grasp dataset

First, we show the performance of models trained with the 3-categories scale and compare it with the original binary classification system. Results are shown on Table 5.21. In the Tombatossals data set, the object weight could turn a robust/futile grasp into fragile. In the Apollo data set, the gravity orientation is the key feature which changes the outcome of grasp executions.

TABLE 5.21: Classification results with the Apollo grasps dataset. First column shows classification using the Binary classification system. Second column shows the 3-categories scale.

	Binary Classification		3-categories scale	
Classifier	Train ± Std	Test	Train ± Std	Test
K-Nearest Neighbors	0.77 ± 0.07	0.80	0.76 ± 0.04	0.78
Classification Trees	0.75 ± 0.06	0.77	0.73 ± 0.04	0.79

Individual metrics

Next, Table 5.22 shows the results of the training when considering only one metric as the input feature vector and the 3-categories score. Results are shown for the *Tombatossals* grasp dataset the data is split randomly between train and test sets.

This table shows first, the prediction with the three categories model. Second column shows the performance considering only *Robust/Futile* grasps.

TABLE 5.22: Classification results for individual quality metrics. First column shows the 3-categories scale. Second column considers only *Robust/Futile* grasps.

	3-categories scale		Robust vs. Fragile	
Metric	Train ± Std	Test	Train ± Std	Test
Q_{A1}	0.67 ± 0.09	0.68	0.92 ± 0.06	0.85
Q_{B1}	0.55 ± 0.04	0.56	0.72 ± 0.10	0.60
Q_{B2}	0.54 ± 0.04	0.52	0.61 ± 0.07	0.63
Q_{B3}	0.53 ± 0.01	0.48	0.57 ± 0.10	0.53
Q_{C2}	0.55 ± 0.03	0.52	0.62 ± 0.07	0.53
Q_{D1}	0.72 ± 0.04	0.73	0.90 ± 0.05	0.85
Q_{D2}	0.51 ± 0.02	0.50	0.65 ± 0.10	0.60

5.6.3 Discussion

In our experiments, we showed that the object weight or gravity vector could vary the outcome of real grasp executions. These are two variables which are independent from the grasp evaluation with quality metrics and should be considered when predicting the success of a grasp.

Results in real experiments showed that there is a clear difference between *Robust*, *Futile* and *Fragile* grasps. First, our results show adding a new category for scoring the grasps: *Fragile*, could improve the classification methods. Table 5.19 showed this method is better than the binary scoring system for the Tombatossals dataset.

With this scheme of three labels for a grasp type, our models are more accurate for predicting the outcome of real grasp executions. This scoring method improves the classifiers between a 11% and 21% as is shown in Table 5.19. Although for the Apollo dataset, the 3-categories scale did not improve the performance of the predictors, classification accuracy was also not decreasing. As was discussed in the previous case study, this could be due to a limitation of the dataset itself.

Next, the analysis on individual metrics showed that this score method fits better for evaluating grasp executions. Metrics Q_{A1} and Q_{D1} showed a great performance(90%) for evaluating only *Robust/Fragile* grasps. Although the other metrics showed better results with this 3-categories scale model, they are still not sufficient individually for predicting the success of a grasp.

From these results, we can conclude this 3-categories scale for scoring grasp executions could be a better option than a classic binary classification.

The training using the held-out grasps method, showed our trained methods can be useful for new grasps using these objects and manipulator, as the performance of the trained classifiers is similar between tables 5.19 and 5.20.

5.6.4 Prediction model

With the experiments and results obtained in this study, we propose a new prediction model. To predict the outcome of real grasp executions we will consider a two-step prediction model. First, we will train our predictors to differentiate between *Fragile* and *homogeneous* grasps. Second, we will generate a prediction model for the homogeneous grasps, differentiating between *Robust* and *Futile* grasps.

Table 5.23 compares the accuracy of classifiers for filtering *Fragile* grasps. First column shows the performance to classify a grasp as *homogeneous* or *Fragile* (binary classification), second column shows the performance of classifiers trained and tested only with *homogeneous* grasps (binary classification: *Robust* vs. *Futile*).

TABLE 5.23: Classification results with the Tombatossals dataset and the proposed prediction model. First column shows the performance for filter the *Fragile* grasps. Second column shows the performance when considering only *Robust* and *Futile* grasps.

	Filter Fragile		Robust vs. Futile	
Classifier	Train ± Std	Test	Train ± Std	Test
K-Nearest Neighbors	0.85 ± 0.05	0.89	0.90 ± 0.05	0.85
Classification Trees	0.89 ± 0.06	0.91	0.91 ± 0.04	0.97

Table 5.24 shows the proposed prediction model using the grasp dataset from the experiments with the Apollo robot (Section 5.5). For this dataset we selected the *FS dataset* with the 7 *QM* as input feature vector. As in the previous table, results show first the performance for differentiating between *Fragile* and homogeneous grasps. Second column shows the classification models trained and tested only with the *Robust/Futile* grasps on this dataset.

TABLE 5.24: Classification results with the Apollo dataset and the proposed prediction model. First column shows the performance for filter the *Fragile* grasps. Second column shows the performance when considering only *Robust* and *Futile* grasps.

	Filter Fragile		Robust vs. Futile	
Classifier	Train ± Std	Test	Train ± Std	Test
K-Nearest Neighbors	0.76 ± 0.06	0.71	0.80 ± 0.04	0.90
Classification Trees	0.77 ± 0.06	0.75	0.84 ± 0.03	0.87

Finally, we train our prediction models combining both, Apollo and Tombatossals, datasets. Table 5.25 shows the performance comparing the 3-categories scale and the binary classification. Table 5.26 shows the performance with the proposed prediction model. First column shows the capability of the model to filter *Fragile* grasps. Second column shows the performance of the trained models when considering only *Robust/Futile* grasps.

TABLE 5.25: Comparison between the classification models for the 3-categories scale and the Binary Classification method. Combined datasets from Apollo and Tombatossals experiments.

	Binary Classification		3-categories scale	
Classifier	Train±Std	Test	Train±Std	Test
K-Nearest Neighbors	0.74±0.03	0.71	0.74±0.04	0.74
Classification Trees	0.72±0.05	0.69	0.76±0.04	0.76

TABLE 5.26: Classification results with the Tombatossals and Apollo datasets and the proposed prediction model. First column shows the performance for filter the *Fragile* grasps. Second column shows the performance when considering only *Robust* and *Futile* grasps.

	Filer Fragile		Robust vs. Futile	
Classifier	Train ± Std	Test	Train ± Std	Test
K-Nearest Neighbors	0.79±0.02	0.81	0.82±0.04	0.82
Classification Trees	0.79±0.04	0.81	0.85±0.04	0.84

Discussion

Results on Table 5.23 showed quality metrics have high prediction capability (>90%) with *Robust* and *Futile* grasps. This shows that it is possible to first, filter *Fragile* grasps with high accuracy (~90%), and then classify the remaining *homogeneous* grasps as *Robust* or *Futile*. This reinforces the idea of using quality metrics as grasp predictors, but filtering them first, in order to discard *Fragile* grasps.

Table 5.24 showed the methods applied in this study can be used with other grasps dataset. Although in this case the performance of the classifiers is similar between the binary scoring and the 3-dimensional method (Table 5.21), results considering only *Robust* and *Futile* grasps showed again a high performance (~90%).

Finally, tables 5.25 and 5.26 showed first, the 3-categories scale is a better option to classify grasps than a binary method. And Second, the proposed prediction model has also a high performance (>80%) with different gripper designs (Barrett and SDH hands).

From the results of this study, we can conclude first, a 3-categories scale system provides a more reliable approach to classify real grasps. Second, classifiers trained and tested with only *Robust* and *Fragile* grasps have a high rate of success (>90%). And third, it's possible to use quality metrics to filter *Fragile* grasps with great accuracy (between $75\% - 90\%$)

The applications and results obtained in this study are limited to only two different grippers and a few objects. Although the held-out method proved the prediction models could work well with new grasps, apply the classifiers trained with this datasets to new objects or robotic manipulators is not advisable. A more in-deep study using other object models of different characteristics and manipulators could provide a more robust background for the classification methods described in this study.

5.7 Generalization of the prediction model

One important remark about the prediction model proposed is the generalization itself of the predictor. Can we use it with other objects and hands? Or is it restricted to be used only with these two grippers and this limited set of objects? For studying this possible generalization, we will exploit our databases of real grasps and train different classifiers using different combinations of train/test sets.

5.7.1 Methodology

Using the methodologies described in the previous case study, we will compare the predictive ability of a model trained with one gripper and tested with another different. We will study its capabilities using the 3-categories scale and the the binary classification for filtering *Fragile* grasps and identify *Robust/Futile* grasps.

We will consider the *Tombtossals* and *Apollo* datasets independently. Our models will be trained with the *Barrett* hand, and tested with the *Schunk SDH* or viceversa.

5.7.2 Results

Table 5.27 shows the results of a classifier trained with the *Tombatossals* database and tested with the *Apollo* database. Table shows the results using the 3-categories scale, the filtering of *Fragile* grasps and the classification with only *Robust/Futile* grasps.

TABLE 5.27: Classification results or the classifiers trained with the *Tombatossals* dataset and tested with the *Apollo* dataset. First column shows the performance using the 3-categories scale. Second column shows the accuracy for filtering *Fragile* grasps. Third column shows the performance considering only *Robust/Futile* grasps.

Classifier	3-categories scale		Filter Fragile		Robust vs. Futile	
	Train±Std	Test	Train±Std	Test	Train±Std	Test
K-Nearest Neighbors	0.78±0.09	0.41	0.93±0.06	0.58	1.00±0.00	0.78
Classification Trees	0.82±0.13	0.38	0.90±0.05	0.48	1.00±0.00	0.83

Table 5.28 shows the results of a classifier trained with the *Apollo* database and tested with the *Tombato* database. Table shows the results using the 3-categories scale, the filtering of *Fragile* grasps and the classification with only *Robust/Futile* grasps.

TABLE 5.28: Classification results or the classifiers trained with the *Apollo* dataset and tested with the *Tombatossals* dataset. First column shows the performance using the 3-categories scale. Second column shows the accuracy for filtering *Fragile* grasps. Third column shows the performance considering only *Robust/Futile* grasps.

Classifier	3-categories scale		Filter Fragile		Robust vs. Futile	
	Train±Std	Test	Train±Std	Test	Train±Std	Test
K-Nearest Neighbors	0.61±0.13	0.31	0.70±0.26	0.52	0.80±0.16	0.81
Classification Trees	0.57±0.10	0.47	0.67±0.26	0.39	0.91±0.15	0.33

5.7.3 Discussion

Results show that generalize a predictor trained with one gripper to another is not initially possible. Results using the 3-categories scale and the filtering of futile grasps have a low accuracy. However, what seems promising, is the fact that the classifiers keep good performance (>80%) when we consider only the *Robust* and *Futile* grasps.

A model trained only with the *Barrett* hand is capable to keep its success rate (~80%) with the *Schunk SDH* hand. In the case of the model trained with the *Schunk SDH* and tested with the *Barrett* hand, the success rate lowers (from 100% to 80%), but it is still high (>80%). It is important to note the *Barrett* hand uses 9 objects while the *Schunk SDH* only 2 object models.

These results suggest first, metrics are good evaluators for *Robust* and *Futile* grasps. Second, it is important to find a methodology capable to filter or discard such *Fragile* grasps. And third, it is possible to generalize a classifier trained with one gripper to another different, but its performance require a previous filtering of those *Fragile* grasps.

5.8 Conclusions

In this chapter we presented the results of three different studies to analyze the predictive capability of grasp quality metrics and to propose a model to predict the outcome of real grasp executions. In the first study this predictive capability was compared to human labeled data. The second study compared this predictive capability using real grasp executions. The third study introduce a new category to score real grasp executions. With this method it is possible to filter ambiguous grasps and improve the predictive capability of combined grasp metrics.

Case study 1: human oracle

Results on the first study showed individually, metrics are not good enough to predict the outcome of grasp executions. This study also pointed out that it is possible to combine different metrics using classification methods to improve their performance. This study showed finally, using non-linear and non-parametric methods are better options for finding an underlying structure of combined metrics to predict the outcome of grasp executions.

Comparing the results of combined quality metrics against dynamic simulation, we showed that the physics metric is a better approach to predict the human labeled data for grasp executions. However, this study presented a huge disadvantage: all experiments and studies were based first, in simulated grasp which may differ from conditions in the real world (for example gravity is active).

Second, the ground-truth for grasp success is based in using humans as *oracles* which reliability must be proven. Humans have to judge the grasp in a physical environment that is different from the usual one (vacuum + three-fingered robotic hand). Our results in the next case study showed humans are not reliable as grasp oracles.

Case study 2: real grasp execution

The second study aimed to solve these problems from the previous study. In this study grasp candidates are evaluated on a real robot system. This provided a reliable ground-truth for comparing the physics metric, human labeling, and the predictive capability of grasp quality metrics. Results on this study showed first, humans and physics metric are not as good predictors as it culd be expected. Although their predictions are highly accurate to identify unsuccessful grasps, they fail to predict the success of a grasp.

Second, this study showed that individually, metrics are not good predictors, but they can be combined to improve their performance, up to an 80% of accuracy. This study showed also *contextual information* such as gravity orientation can change the outcome of real grasp executions, but do not provide relevant information to predict the outcome of grasp executions.

Case study 3: metrics limitations and 3-categories scale

The third study analyzed the limitations of grasp metrics and how the object's weight can affect the outcome of grasp executions. This study showed the same grasp can succeed or fail depending on the object's weight, as it occurs with different gravity orientations. This is an object property not considered by the grasp metrics, and connote a limitation on their predictive capabilities.

In our study, we assume we only have the information provided from the simulation for the classifiers. This is, we only have information regarding the contact points, quality metrics and object shape. Object's weight and gravity orientation is unknown and is not included as input for the classification models.

Under these assumptions, a 3-categories scale system for grasp executions was proposed. With this scoring method, grasps can be classified as *Robust*, *Futile* or *Fragile*, depending if the execution of the candidate grasp always *succeeds*, *fails* or both.

This 3-categories scale method improved the classification and predictive capability of the models using quality metrics. Results showed this 3-categories scale reflects better the outcome of real grasp executions. The use of different gravity orientations and object weights proved to be useful, not for the classification of *Stable/Unstable* grasps, but for determining which grasps are *Robust*, *Futile* or *Fragile*.

5.8.1 Prediction model

Using our proposed prediction model is shown metrics are capable to first, filter *Fragile* grasps. And second, classify grasp candidates as *Robust* or *Futile* with a high success rate (>90%). This suggests metrics are good solutions to identify and classify *Robust* and *Futile* grasps, while the characteristics and qualities of *Fragile* grasps are fuzzier.

Although classifiers are able to identify *Fragile* grasps with good performance (>70%), it is important to consider using the contextual information as a factor which could help to identify these cases.

With respect to the generalization of the prediction model, it is shown first, metrics are good evaluators for *Robust* and *Futile* grasps. Second, it is important to find a methodology capable to filter or discard such *Fragile* grasps. And third, it is possible to generalize a classifier trained with one gripper to another different, but its performance require a previous filtering of those *Fragile* grasps.

Finally, from the studies performed in this chapter we can conclude:

- Human labeled data, dynamic simulation and grasp quality metrics alone, are not the best options to predict the success of a grasp in real world.
- There seems to be some underlying information, not intuitive for humans, which helps to better understand the outcome of grasp executions.
- Combinations of different quality metrics seem to enable finding this underlying structure to predict grasp success.
- *Contextual information*: geometric properties (volume, weight) or the gravity approach vector affect to the outcome of real grasp executions, but the prediction with quality metrics cannot benefit from it.
- Real grasp executions can be classified using either, a binary model (*stable* vs. *unstable*) or a 3-categories scale: *Robust*, *Futile* and *Fragile* grasps.

- Classifiers that use *QM* as input, perform better when using this 3-categories scale than the classic binary classification.
- Quality metrics prediction performs better when considering only *Robust* and *Futile* grasps.
- It is possible to combine different grasp metrics using classification methods to generate a model able to predict the outcome of real grasp executions with high precision (90%).
- The proposed prediction models could be generalized to other objects and grippers, but first it is necessary to find a better method to discard/filter fragile grasps.

5.8.2 Limitations

This chapter presented an extensive study on the performance of different quality metrics for predicting grasp success, but it had also some limitations. First, only two manipulators were used to perform the experiments and evaluate grasps. Extending the results on prediction models to other grippers should be done carefully, as results may vary. Second, we used a reduced number of objects to perform the experiments, although different shapes and weights were tested, an extended study with more objects should be done.

Third, we considered different geometric characteristics of the objects, but there are other properties that should be taken in account: materials, elasticity, friction coefficient, etc. Fourth, the real grasps executions were restricted to an environment with a table holding the object prior the grasp. Repeating these experiments in other environments with different restrictions or without restrictions is advisable. Finally, the prediction models were generated using few different types of classification methods. A wider study with this data can be done using other algorithms or methods, as it could provide better results.

5.8.3 Future & related work

As future work, is suggested to extend this study using other grippers and object models. It will provide a wider view on the predictive capability of combined or independent quality metrics. As showed the results for the generalization of the model, there is not enough data to reliably apply the models trained here to other grippers or objects. Try to find a prediction model generalizable to new data should be a next step.

Our studies showed the outcome of grasp executions is dependent on object geometric properties but the prediction based on quality metrics does not benefit from it. Therefore a study using the same object models but with different materials and size could provide new relevant information to improve the prediction of grasp executions.

For comparing the influence of different gravity directions relative to the local object frame, we did not have enough data and therefore could not reach convincing conclusions. It is recommendable to make a more deep study varying the orientation of the objects prior grasping and compare these results to human labels, physics metric and grasp metrics.

The 3-categories scale showed great results for *Robust* and *Futile* grasps, but is important to find a better method to filter first those *Fragile* grasps. Using different quality metrics or contextual information could be a key factor for being able to filter those *fragile* grasps.

The studies, methodology and results presented in this chapter were total or partially published in Rubert et al., 2017.

Chapter 6

Artificial hands evaluation

In this chapter, we use grasp quality metrics to evaluate artificial hands. We propose to use grasp quality metrics as evaluators of different hand properties. We want to build a benchmark procedure that produces numeric evaluation of artificial hands.

In Chapter 4, we have studied the quality of a grasp using several quality metrics. Also, we have developed a methodology to characterize the practical parameters of the different quality metrics. These results showed that the average values of a metric greatly depends on the hand that was being used. A first approach in benchmarking hand models was made in Leon et al., 2013a comparing the quality of different grasps when using two different models of hands: a simulated model of a human hand and a prosthetic hand.

In Chapter 5 we have developed a methodology to classify grasp hypotheses. Making use of machine learning algorithms and quality metrics, we developed a model capable to predict the successfulness of a grasp hypotheses. Following this methodology we are able to classify a grasp hypotheses as *Robust*, *Fragile* or *Futile*. These results allowed to classify grasps with the *Barrett* and the *SDH Schunk* hand with more than a 80% of success rate.

In this chapter we want to exploit both methodologies in order to evaluate artificial hands. First we'll analyze quantitatively the characteristics of different hand models. We hypothesize quality metrics evaluate not only independent aspects of the grasp, but also different properties of hand models. We will make use of seven quality metrics to assess the performance of each hand according to specific properties. Results obtained with grasp metrics will be compared to other indexes used to evaluate hands.

Second, we will use the prediction models developed in Chapter 5 to evaluate the grasping capabilities of different artificial hands. With the trained models we are be able to predict the outcome of different grasp hypotheses with different object models and hands. We will propose a methodology which evaluates the likelihood of a hand to produce successful grasps.

Finally, we intend to exploit this methodology by defining a standard that can be used not only to analyze different robot hands but also to allow comparison between different designs of the same hand. This evaluation is based on the independent grasp aspects evaluated by each selected quality metric and the prediction models.

With these studies we want to answer some questions: Can quality metrics reveal differences between artificial hands? Can these quality metrics be related to other indexes used to evaluate hands? Can we evaluate the likelihood of a hand to produce successful grasps? Can we apply these methodologies to improve hand designs?

This chapter presents three studies that try to establish: (i) differences and similarities between artificial hands, (ii) the relations between the grasp metrics and the anthropomorphism index (AI) developed by Feix et al., 2013, (iii) a methodology to evaluate the successfulness of hand models and (iv) how to apply this methodology to improve the design of a prosthetic hand. To achieve this, several virtual models of robotic grippers and anthropomorphic hands are used. All the experiments are performed in simulation and their results are compared statistically.

Outline The chapter is structured as follows: first, an introduction and review of different studies which aim to evaluate artificial hands. And second, we expose four different studies which evaluate the properties of different hand models. The first study uses quality metrics to establish benchmarks between artificial hand models. The second study aims to find a correlation between the quality values and the anthropomorphism of artificial hands. The third study uses a classification model to evaluate the likelihood of a hand to produce successful grasps. An last, the fourth study evaluates different designs of the same prosthetic hand.

6.1 Introduction

The design of robot hands usually has been focused on three main aspects: dexterity in manipulation, robustness in the grasp and human operability (Bicchi and Kumar, 2000; Bicchi, 2000). Additionally, there are two ways of design: emulate human anatomy or focus on specific tasks (Kappassov et al., 2013; Tai et al., 2016). The survey of Tai et al., 2016, reviews the state of the art in robotic gripper designs and applications. Several authors (Butterfass et al., 2001; Dollar and Howe, 2010; Birglen et al., 2007; Kappassov et al., 2013) have analyzed different robot hands mainly studying the hands' physical properties such as their sensors, degrees of freedom or skeleton.

The work of Dollar and Howe, 2010 evaluates experimentally the grasping capabilities of the *SDM Hand* under unstructured scenarios. The analysis of Butterfass et al., 2001 is focused in the physical characteristics for designing the *DLR hand*. The book of Birglen et al., 2007 present a full study on the design of underactuated robotic hands and the properties of their fingers. Kappassov et al., 2013 compares the grasp patterns, force capabilities, maximum payloads and control of different different robotic grippers and multigrasp robotic hands.

Physical characteristics on the hands as fingers, DOF, motors or control are the most common properties compared in robotic hands. However, there is no direct relation between this properties and their grasping capabilities. It is needed to find a methodology capable to evaluate the grasping capabilities of different artificial hands.

Different studies have addressed the problem of how to evaluate prosthesis. Bullock et al., 2015 and Feix et al., 2014 perform a deep analysis in how humans grasp objects, analyzing hours of record of humans during activities of daily living and how to replicate these grasps in prosthesis. Kargov et al., 2004 measure and compare the grip force distribution between human and artificial hand grasps, defining this as a metric for evaluating hand prosthesis. Matheus and Dollar, 2010 expose a series of benchmarks for evaluating manipulator grasp contacts based on common objects and its physical properties, coefficient frictions and contact forces. Belter et al., 2013 compared the mechanical designs and performance specifications, in terms of forces, velocities and kinematic characteristics, of different prosthetic hands.

Feix et al., 2013 propose an anthropomorphism index (AI) which summarizes which percentage of the configuration space of the human hand can be replicated by an artificial hand while grasping objects. This configuration space is defined by the positions of the fingertips. Thus, a high percentage on this index indicates the artificial hand is able to replicate a large number of human hand shapes. The hypothesis behind this metric is that artificial hands with a high index have a similar working space to that of the human hand and consequently are more functional.

All these studies focus the evaluation of artificial hands in their ability to replicate the behavior and shape of human hands or the kinematic characteristics of the fingers. However, this has not been proved to be a good benchmark for evaluating hand prosthesis. As shown by Jang et al., 2011, amputees tend to prefer cosmetic prostheses over such more functional, due to the poor performance on the functional ones. This exposes that actual prosthetic hands are far from provide a good substitution of the human hand. Therefore, it is needed to focus in develop functional tests to evaluate the grasping capabilities in artificial hands during its design phase instead of only seek replicate the human hand behavior.

6.2 Artificial hand properties

In this first study we will analyze the properties of different artificial hands. These hand properties are defined as the 7 independent quality metrics (see Table 5.1) obtained in Chapter 4. Results on this chapter showed hand models present different statistical values for each metric. We consider each one of this metrics evaluates a property in a hand and thus, can be used for establish benchmarks between artificial hands.

In this section we will exploit the grasp database generated in Chapter 4 in order to evaluate different hand models. These grasps have been already evaluated with quality metrics and we will analyze these values to compare different hand models. Results will provide a qualitative analysis on different hand models. With the quality values on each hand we will be able to asses differences and similarities between hand models according to specific grasp aspects.

6.2.1 Methodology

Our database contains 9 different hand models, 126 different objects and 1.000 grasps per hand and object. The hand models correspond to the 7 hands used in Chapter 4 plus two more hand models: *FlexyHand* and *Tact hand*. For this two new models, grasps are generated using the same methodology and parameters applied in Chapter 4.1. Grasps are also evaluated using OpenHand (section 3.1). Objects used are from the KIT Database (see Chapter 3.3.

We will analyze the variability and sensitivity of each hand for performing grasps and compare them. First, a variability analysis on the properties evaluated by different metrics for each robot hand. This will establish the statistical values and range of variation of different hands for each metric. For each metric we obtain their max and min values, the mean and standard deviation, and their percentiles 10, 50 and 90. We compare in tables and figures this values for each hand model and metric.

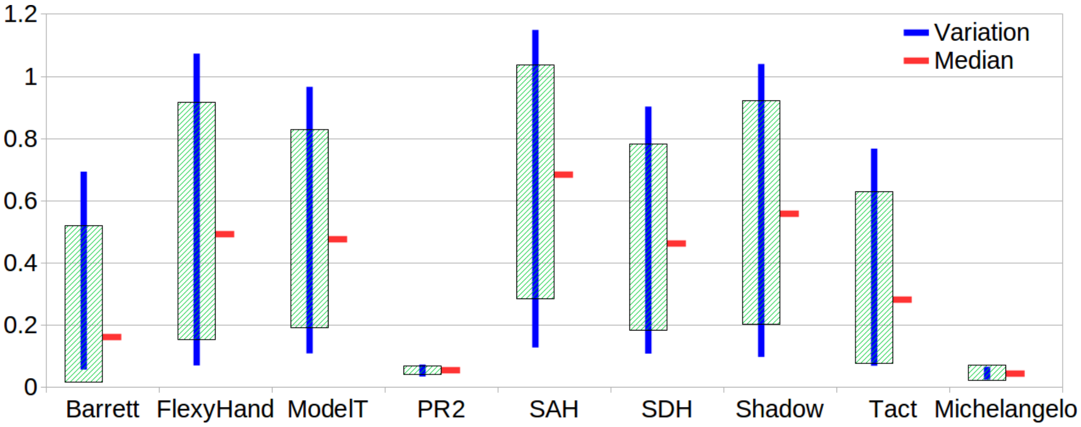
Secondly, we will focus on analyze the sensitivity of each hand against positioning errors. This will provide a sensitivity index of the hand and how reliable it is to replicate the same grasp with the same quality. We will introduce uncertainty in the grasp pose for 20 random grasps from the database for each hand model and object. We will measure, for a reference grasp pose, 50 small variations in translation and rotation. Results will provide a Sensitivity Index (*SI*) on each hand model for each metric. Details on the variations for each grasp pose and how to obtain the *SI* are described in section 4.3.

To demonstrate the methodology we compare nine hand models: the *Barrett* hand, the *FleyHand* prosthetic hand, the *Model-T*, the *PR2* hand, the four-fingered Schunk *SAH* hand, the dexterous three-fingered Schunk *SDH* hand, the 5-fingered *Shadow* hand and the prosthetics *Tact* and *Michelangelo* hands. See section 3.2 for details of each hand model.

6.2.2 Results

Variability analysis

Tables 6.1 to 6.7 show the statistical values on each hand for the 7 independent quality metrics. The blue line shows the variation of values between its mean±std. The green box represents the values between percentiles 10 and 90. Finally, the red line correspond to the median value (percentile 50). Tables show these statistical values and the number of grasps generated and evaluated with each hand model. This is relevant, as not all the hand models were capable of generating the minimum of 1.000 different grasps per object.

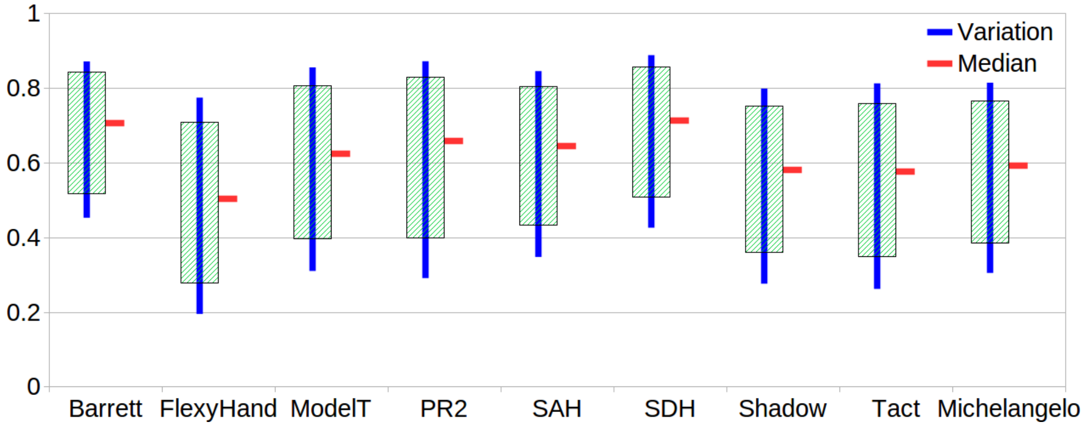


Q_{A1}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.0082	1.3749	0.2691	0.2525	0.0576	0.1608	0.6941	126000
FlexyHand	0.0002	1.7885	0.5348	0.3813	0.0710	0.4926	1.0732	84744
ModelT	0.0075	1.5440	0.5105	0.3180	0.1098	0.4765	0.9665	126000
PR2	0.0147	0.4448	0.0555	0.0139	0.0354	0.0559	0.0739	126000
SAH	0.0034	1.7993	0.6606	0.3760	0.1287	0.6830	1.1494	126000
SDH	0.0071	1.3961	0.4838	0.2988	0.1091	0.4617	0.9031	126000
Shadow	0.0000	1.9408	0.5633	0.3587	0.0983	0.5589	1.0400	126000
Tact	0.0008	1.4629	0.3536	0.2767	0.0700	0.2820	0.7681	125319
Michelangelo	0.0004	0.9484	0.0469	0.0250	0.0257	0.0460	0.0667	126000

TABLE 6.1: Statistical Values for metric Q_{A1} per hand model. Green box shows percentiles 10-90.

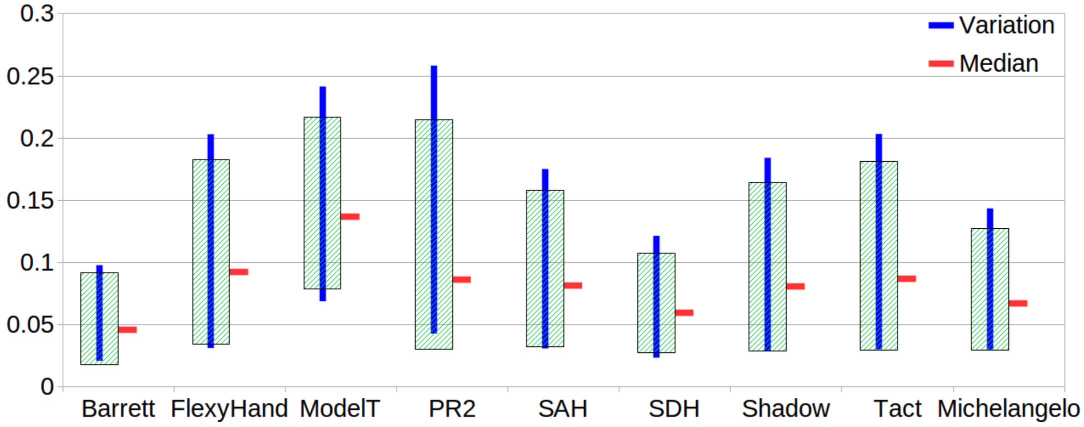
Sensitivity analysis

Table 6.8 shows the robustness of each hand for the different metrics. Values shown are GSI_N per metric and hand. Values are normalized over 10/90th percentiles from Table 4.3 in Chapter 4.2.



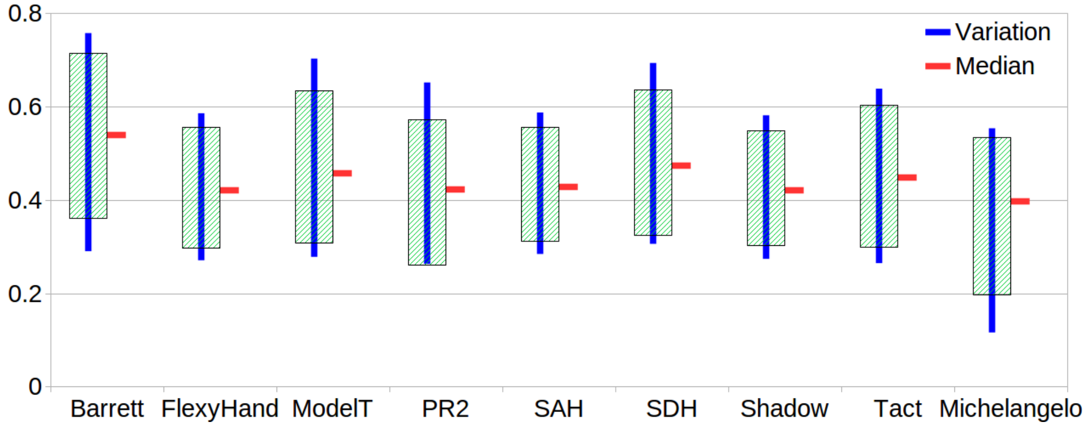
Q_{B1}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.0634	0.9995	0.6805	0.1624	0.4535	0.7071	0.8722	126000
FlexyHand	0.0000	0.9825	0.4937	0.2151	0.1958	0.5039	0.7753	84744
ModelT	0.0285	0.9945	0.6020	0.2044	0.3108	0.6254	0.8560	126000
PR2	0.0015	0.9974	0.6148	0.2162	0.2917	0.6585	0.8726	126000
SAH	0.0428	0.9959	0.6189	0.1855	0.3485	0.6447	0.8462	126000
SDH	0.0317	0.9973	0.6830	0.1745	0.4267	0.7134	0.8888	126000
Shadow	0.0016	0.9933	0.5569	0.1952	0.2768	0.5825	0.7995	126000
Tact	0.0033	0.9915	0.5540	0.2052	0.2629	0.5760	0.8134	125319
Michelangelo	0.0084	0.9972	0.5750	0.1901	0.3056	0.5926	0.8152	126000

TABLE 6.2: Statistical Values for metric Q_{B1} per hand model. Green box shows percentiles 10-90.



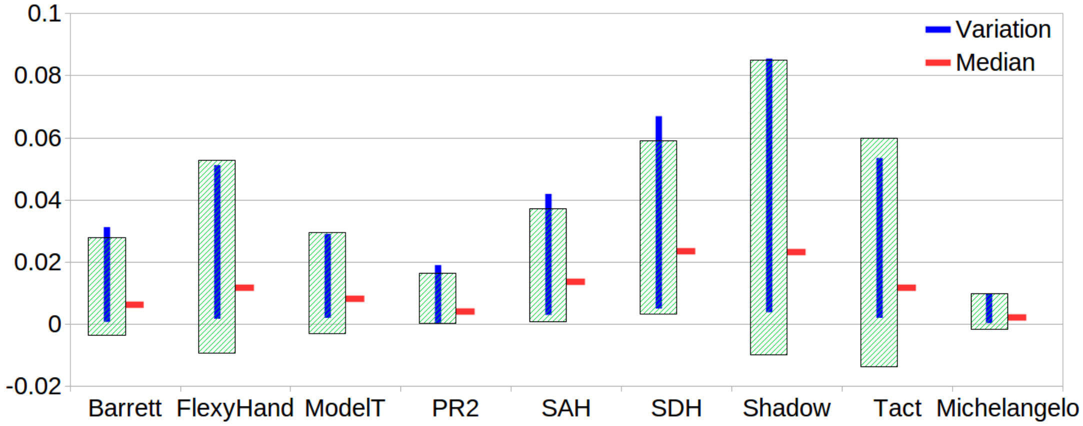
Q_{B2}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.0022	0.4063	0.0550	0.0371	0.0210	0.0462	0.0979	119277
FlexyHand	0.0001	0.9595	0.1085	0.0740	0.0314	0.0928	0.2032	84074
ModelT	0.0002	0.6834	0.1479	0.0691	0.0690	0.1372	0.2415	125201
PR2	0.0341	0.3403	0.1225	0.0922	0.0429	0.0860	0.2583	63
SAH	0.0001	0.8518	0.0951	0.0629	0.0310	0.0814	0.1752	125781
SDH	0.0000	0.3928	0.0674	0.0402	0.0236	0.0597	0.1215	125718
Shadow	0.0000	0.8700	0.0964	0.0676	0.0287	0.0808	0.1842	124745
Tact	0.0000	0.9598	0.1053	0.0758	0.0303	0.0869	0.2034	124947
Michelangelo	0.0083	0.3195	0.0785	0.0487	0.0302	0.0675	0.1436	543

TABLE 6.3: Statistical Values for metric Q_{B2} per hand model. Green box shows percentiles 10-90.



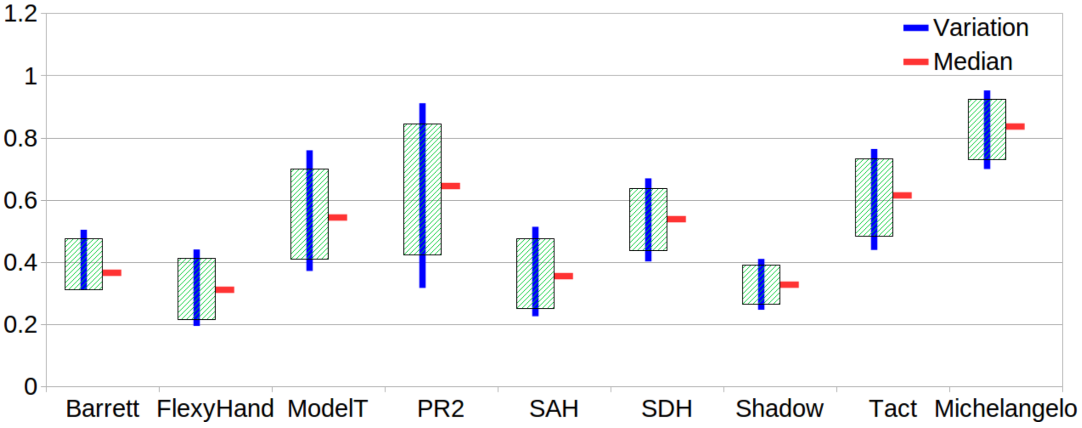
Q_{B3}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.0000	0.9987	0.5378	0.1771	0.2910	0.5401	0.7584	119277
FlexyHand	0.0006	0.9891	0.4268	0.1295	0.2717	0.4207	0.5867	84074
ModelT	0.0055	0.9973	0.4715	0.1629	0.2790	0.4579	0.7038	125201
PR2	0.0021	0.7311	0.4176	0.1555	0.2639	0.4230	0.6526	63
SAH	0.0029	0.9932	0.4338	0.1224	0.2851	0.4291	0.5881	125781
SDH	0.0001	0.9956	0.4801	0.1559	0.3068	0.4748	0.6944	125718
Shadow	0.0001	0.9842	0.4260	0.1238	0.2747	0.4222	0.5822	124743
Tact	0.0001	0.9960	0.4518	0.1518	0.2656	0.4488	0.6394	124947
Michelangelo	0.0006	0.9336	0.3655	0.1685	0.1169	0.3972	0.5544	543

TABLE 6.4: Statistical Values for metric Q_{B3} per hand model. Green box shows percentiles 10-90.



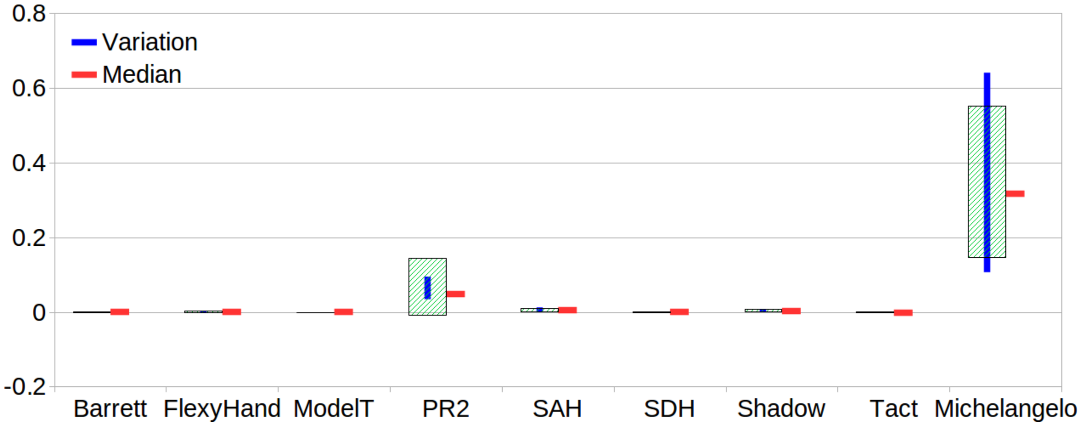
Q_{C2}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.0000	0.1902	0.0122	0.0157	0.0008	0.0064	0.0313	119277
FlexyHand	0.0000	0.6846	0.0218	0.0310	0.0018	0.0118	0.0513	84073
ModelT	0.0000	0.2786	0.0133	0.0162	0.0021	0.0082	0.0291	125201
PR2	0.0003	0.0289	0.0084	0.0080	0.0005	0.0042	0.0191	63
SAH	0.0000	0.2396	0.0191	0.0181	0.0031	0.0138	0.0420	125769
SDH	0.0000	0.3170	0.0311	0.0279	0.0051	0.0236	0.0670	125718
Shadow	0.0000	0.9213	0.0377	0.0474	0.0039	0.0233	0.0855	124736
Tact	0.0000	1.1636	0.0232	0.0368	0.0021	0.0118	0.0536	124943
Michelangelo	0.0000	0.0608	0.0041	0.0057	0.0004	0.0022	0.0099	543

TABLE 6.5: Statistical Values for metric Q_{C2} per hand model. Green box shows percentiles 10-90.



Q_{D1}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.1174	0.8259	0.3938	0.0823	0.3136	0.3672	0.5054	126000
FlexyHand	0.0000	0.7964	0.3155	0.0980	0.1963	0.3124	0.4420	84744
ModelT	0.1238	0.9651	0.5565	0.1456	0.3730	0.5458	0.7609	126000
PR2	0.0057	0.9999	0.6353	0.2105	0.3185	0.6461	0.9121	126000
SAH	0.0238	0.8619	0.3650	0.1121	0.2273	0.3561	0.5149	126000
SDH	0.2405	0.8216	0.5379	0.1007	0.4034	0.5390	0.6708	126000
Shadow	0.0791	0.6433	0.3294	0.0635	0.2482	0.3281	0.4116	126000
Tact	0.1783	0.9950	0.6097	0.1249	0.4403	0.6166	0.7651	125319
Michelangelo	0.2072	0.9990	0.8289	0.0971	0.7007	0.8361	0.9533	126000

TABLE 6.6: Statistical Values for metric Q_{D1} per hand model. Green box shows percentiles 10-90.



Q_{D2}	Min	Max	Mean	Std	P10	P50	P90	N°Grasps
Barrett	0.0000	0.0178	0.0007	0.0011	0.0000	0.0002	0.0019	126000
FlexyHand	0.0000	0.1602	0.0021	0.0024	0.0002	0.0017	0.0044	84744
ModelT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	126000
PR2	0.0000	0.9391	0.0680	0.0758	0.0352	0.0493	0.0957	126000
SAH	0.0000	0.0480	0.0068	0.0045	0.0016	0.0060	0.0131	126000
SDH	0.0000	0.0188	0.0008	0.0014	0.0000	0.0002	0.0024	126000
Shadow	0.0000	0.0219	0.0048	0.0031	0.0011	0.0045	0.0091	126000
Tact	0.0000	0.0253	0.0005	0.0012	0.0000	0.0000	0.0018	125319
Michelangelo	0.0000	0.9980	0.3492	0.2029	0.1074	0.3170	0.6419	126000

TABLE 6.7: Statistical Values for metric Q_{D2} per hand model. Green box shows percentiles 10-90.

TABLE 6.8: Sensitivity Values per each Hand and Metric

	Q_{A1}	Q_{B1}	Q_{B2}	Q_{B3}	Q_{C2}	Q_{D1}	Q_{D2}
Barrett	15.72%	14.57%	12.33%	27.94%	14.16%	7.76%	0.43%
Flexyhand	24.13%	7.93%	31.71%	24.91%	19.68%	10.62%	0.59%
ModelT	23.04%	10.46%	34.29%	34.54%	13.26%	16.72%	0.00%
PR2	0.29%	10.74%	16.43%	7.76%	7.03%	4.80%	39.76%
SAH	30.55%	13.40%	34.22%	28.33%	29.69%	15.19%	1.92%
SDH	20.99%	13.06%	17.00%	34.47%	27.83%	11.78%	0.52%
Shadow	25.50%	9.85%	33.90%	23.72%	42.50%	8.09%	0.97%
Tact	14.69%	9.40%	28.48%	28.81%	20.06%	9.29%	0.40%
Michelangelo	20.25%	11.52%	50.22%	36.41%	30.10%	14.03%	59.69%

6.2.3 Discussion

The quantitative evaluation with metrics allows us to establish similarities and differences between hand models. According to metrics Q_{A1} and Q_{C2}, *PR2* hand and *Barrett* hand, have a similar low performance. *Shadow*, *Tact* and *Michelangelo* showed also similar values for metric Q_{B1}. Same similarities with this metric are found for hands *Barrett* and *SDH*. According to metric Q_{B2} hands *Tact* and *Flexyhand* are the most similar.

For the metric Q_{B3}, there are similarities between hands *SAH*, *Shadow* and *Flexyhand*. *Tact* and *Flexyhand* showed also similar values in metric Q_{C2}. In the case of metric Q_{D1}, none of the hands seemed to be similar to the others. Finally for metric Q_{D2}, most of the hand models showed low values, only *PR2* and *Michelangelo* presented a huge difference in terms of range of values for this metric.

It is shown each hand model has different values for the metrics. Results did not show a hand clearly overwhelming the other models in terms of variability. Regarding the sensitivity of hand models, the *PR2* hand shows the lowest *SI* in most of the metrics.

As show results on Chapter 5, there is no direct relation between metric values and grasp successfulness. In this terms, we can hypothesize a hand with high variability between metrics will have more chances of finding a *stable* grasp hypotheses. In the other hand, having a low *SI* implies the hand will be more robust against inaccuracies and noise in the grasp pose.

It is also important to remark not all the hand models were capable to generate the 126.000 initial grasp hypotheses. This is a value to take in account, as it implies some hand models have more chances to achieve a *favorable* grasp hypotheses. Also, some of the metrics require a minimum number of contact points, which is difficult to achieve with some hand models (*PR2* and *Michelangelo*).

Although it is not possible to state which hand is better than others. Is possible to use quality metrics to find similarities between different hand models. This could be beneficial in the industrial environment, as it will be possible to replace robotic grippers with cheaper models with the same grasp properties.

Also, it could be useful in the case of prosthetic hands, which aim to replicate the behavior of the human hand. Thus, with this methodology it is possible to evaluate a model of the human hand in order to establish a benchmark for comparing anthropomorphic hand models.

Results reported are purely statistical and doesn't imply one hand model is better or worse than others. An study relating the values in quality metrics and the real grasping capabilities of each hand model is needed to validate the results obtained in this analysis.

6.3 Quality metrics and anthropomorphic index

From previous section, it was shown each hand has its own values according to different grasp aspects. But, can we asses one hand model is better than other according to this values? It is necessary to determine whether specific metric values are related to grasp success or particular hand characteristics. Regarding the first, studies on chapter 5 showed there is no direct relation between having higher or lower values in a metric and providing successful grasps. Thus, in this study we want to determine if it is possible to relate quality metrics to other hand indexes.

The *Anthropomorphic Index (AI)* (Feix et al., 2013) is a metric which evaluates the similarity between a prosthetic hand and the human hand. It is assumed, the human hand is the perfect manipulator. Therefore, an artificial hand capable of achieving the same workspace and movements will be considered as good as the human hand.

In this chapter we aim to establish a relation between quality metrics evaluation and the Anthropomorphic Index (AI). We will evaluate prosthetic hands using a similar procedure done in the previous section to establish the quality values of hands. Then, we will evaluate these same hands with the *AI*. With the results obtained we will try to find a correlation between the quality values and their anthropomorphism.

6.3.1 Anthropomorphic index

Feix et al., 2013 defines a metric for comparing the anthropomorphic capability of artificial hands. This metric calculates an Anthropomorphism Index (*AI*) which gives a value of similarity between human and artificial hands. It measures if an artificial hand is capable of producing the same movements and postures of the human hand.

The workspace of a hand is determined using a 2D latent space. This space is generated with the recording of the finger tips of the hand model for different grasps or hand postures. Then it is compared the latent space of the artificial hand to the space the human hand (base space).

The space for the human hand is generated using 31 different grasp types based on an extensive grasp taxonomy (Figure 1.5). For each grasp event, the finger tips of the hand are recorded. This produces a high dimensional workspace, then, with non-linear dimensionality reduction techniques and kinematic structures, it is possible to generate the 2D latent space.

Over this latent space, different configurations of artificial hands can be projected, which gives an estimation of the anthropomorphic index of a given hand as the percentage surface covered by the artificial hand on the human hand space.

To generate a latent space for an artificial hand requires different hand configurations. This could be done either from different grasp poses or random configurations of the hand model. A problem that arises is the huge computational cost in terms of memory, related to the size of the matrix that defines the set of different configurations a hand can achieve. This matrix depends on the number of DoF, the range of movement for every DoF and the step chosen for the joint movement on each DoF.

The *Grade your Hand Toolbox*[1] (Feix et al., 2013) provides the latent space for the human hand and the tools necessary to calculate the latent space of any other hand.

---

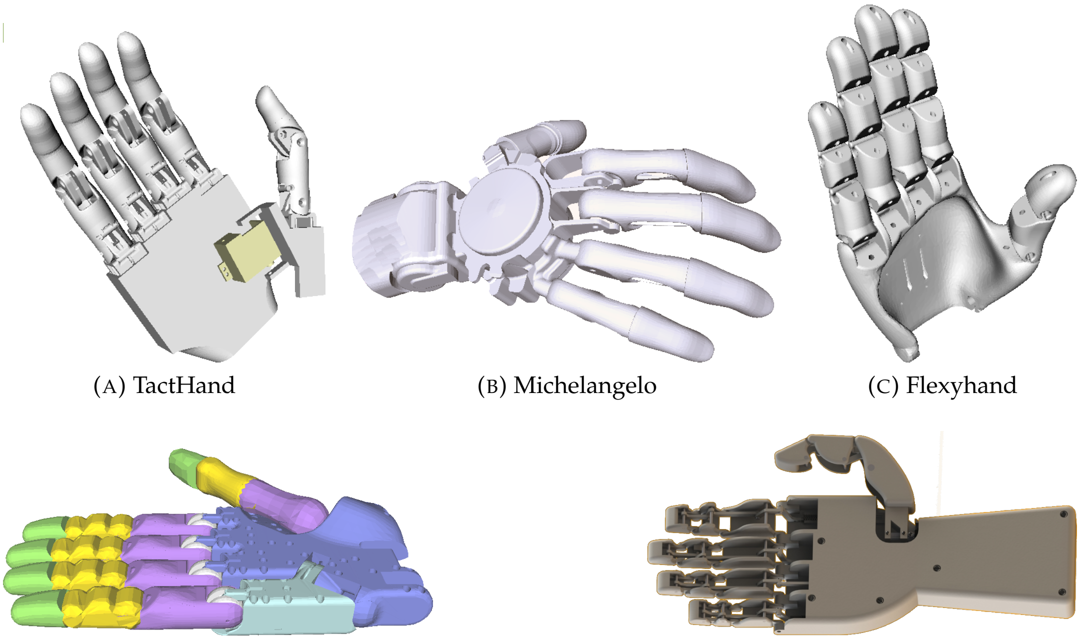
[1] `http://grasp.xief.net/`

For this computation only the hand kinematics are necessary. It is important to note this metric is intended for anthropomorphic hands and thus, classic grippers or robotic manipulators are impaired.

6.3.2 Methodology

Anthropomorphic hands

We have selected 5 artificial anthropomorphic hand models (Figure 6.1): *TactHand*, *Michelangelo*, *FlexyHand*, *ShadowHand* and *BruJa*. More details of this hand are presented in section 3.2. Table 6.9 summarizes the specifications for the whole set of hand models. All of them are assumed to have rigid bodies.



(A) TactHand (B) Michelangelo (C) Flexyhand

(D) Shadow Hand (E) BruJa

FIGURE 6.1: Hand models analyzed with metrics and the *AI*

TABLE 6.9: Details of the Hand Models Used

Hand	Fingers	Joints	Actuators	Description
Flexy	5	14	5	FlexyHand model
Mich	5	6	2	Michelangelo Hand model
Tact	5	11	6	TactHand model
BruJa	5	15	6	BruJa Model-B hand model
Shadow	5	17	7	ShadowHand model

Objects

Different object models are selected depending on the index. For the AI we selected 15 objects models of common objects. These object models are the same used by

Feix et al. Feix et al., 2013 in their calculation of the latent space of the human hand, which is the base to calculate the Anthropomorphism Index. This ensures that both approaches are based on the same set of objects. Figure 6.2 shows the whole set of objects.

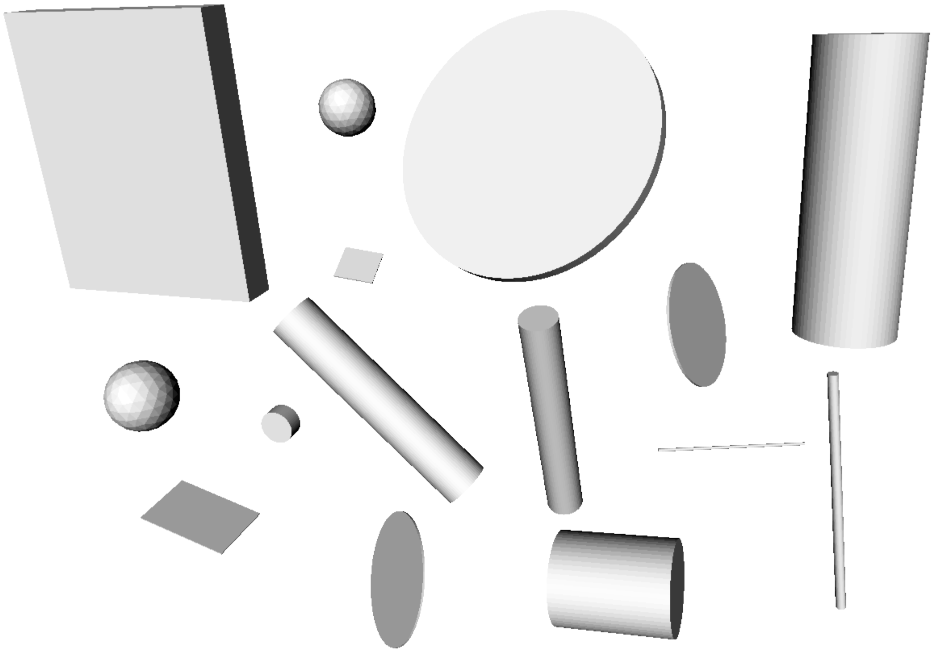


FIGURE 6.2: Sample of the different objects used for calculate the Anthropomorphic Index.

Experiments

First, we will generate 1000 random grasps with the 15 objects models for each hand. Each one of these grasps are generated and evaluated with the seven independent quality metrics using *OpenHand*.

The second part of the analysis computes the AI for each hand model. The computation of the AI is dependent on the number of DoF of the hand and the range of movement of each joint. For hands with many DoFs or large edges of movement, it is not feasible to take in account all the configurations, so a subset of grasp hand configurations has to be sampled. For the sampling, two different methods have been used: a *random* sub-sampling and a *grasp configuration* based selection:

Random sub-sampling A step size of 1°is used and 100.000 random configurations of each hand are selected randomly. This sub-sampling has been done ten times per hand, in order to get a standard deviation of the measure. In the case of the *Michelangelo* the sub-sampling is not necessary, since the total amount of hand configurations is 420, (it has only 2 DoF).

Grasp Configuration sub-sampling 1000 random grasps from the quality metrics analysis for every hand are selected. With this data a series of configurations going from initial common position of the hand to these selected grasp configurations is generated using steps of 1°. This method gives hand configurations representing real grasp postures, in opposite to the first method which can give any type of hand posture.

The range of movement of the different DoF in every hand is detailed in Table 6.10.

TABLE 6.10: Range of the DoFs used to generate hand configurations in the AI evaluation.

Hand	Thumb	Thumb Abduction	Index/Middle Ring/Small	Small Abduction
Flexy	[0°- 70°]	NA	[0 70°]	NA
Mich	[0°- 70°]	[0,4,8,12,16,20]	[0°- 70°]	NA
Tact	[0°- 90°]	[0°- 90°]	[0°- 90°]	NA
BruJa	[0°- 90°]	[0°- 90°]	[0°- 90°]	NA
Shadow	[0°- 45°]	[0°- 90°]	[0°- 90°]	[0°- 45°]

6.3.3 Results

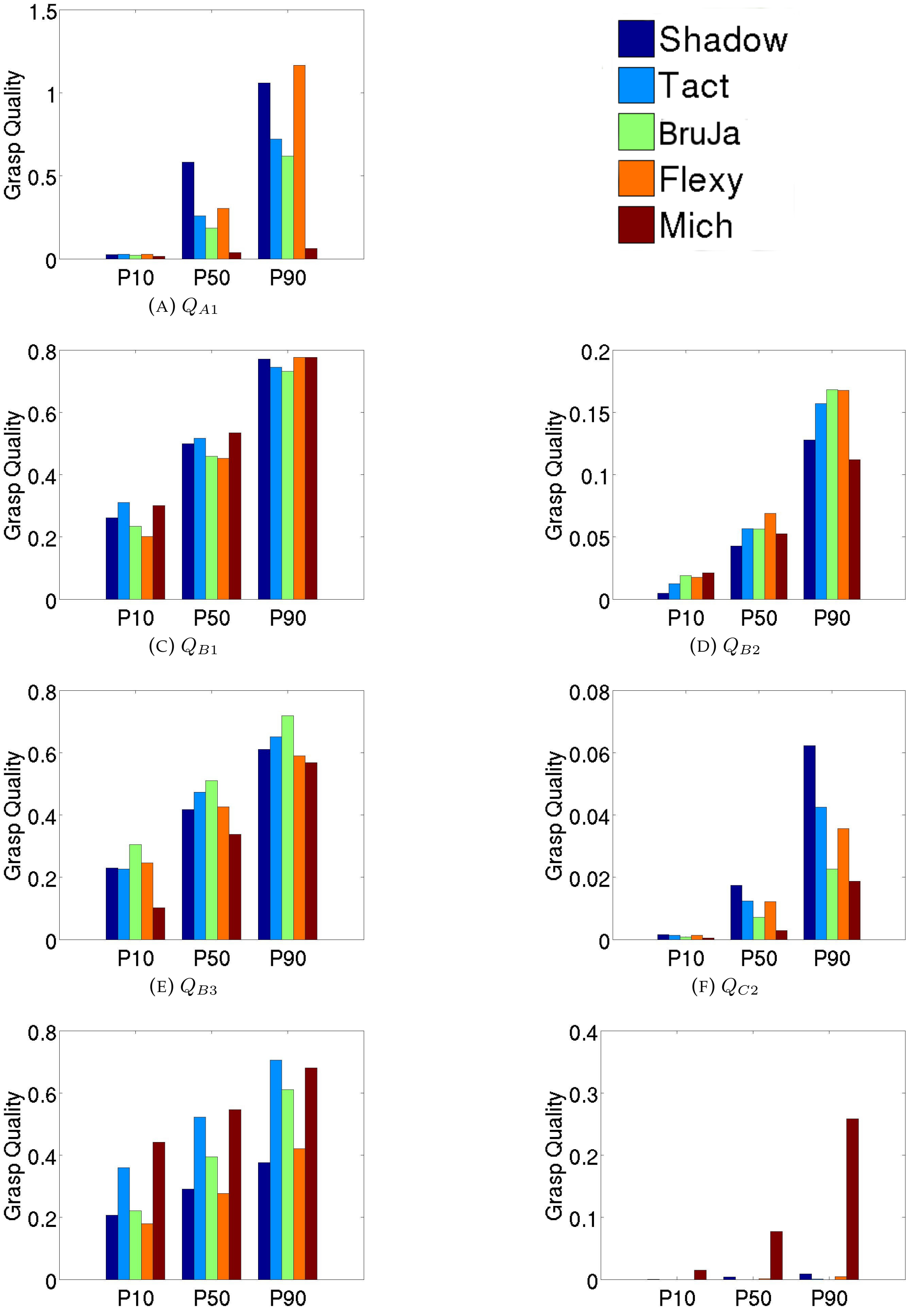
Table 6.11 shows the values of the metrics for each hand. The ranges are selected using the Percentiles 10 and 90. These limits indicate typical ranges for a hand with respect to a specific metric. As showed results in previous study (section 6.2), the ranges of values on metrics typically present many outliers. These outliers would distort the intervals in which most of the values lay if maximum and minimum values were chosen as range limits.

TABLE 6.11: Quality Evaluation in Anthropomorphic Hands

		A1	B1	B2	B3	C2	D1	D2	Grasps
Shadow	Perc10	0.02	0.26	0.00	0.23	0.00	0.21	0.00	13504
	Perc90	1.06	0.77	0.13	0.61	0.06	0.38	0.01	
Tact	Perc10	0.03	0.31	0.01	0.23	0.00	0.36	0.00	9561
	Perc90	0.72	0.75	0.16	0.65	0.04	0.71	0.00	
BruJa	Perc10	0.02	0.23	0.02	0.30	0.00	0.22	0.00	6884
	Perc90	0.62	0.73	0.17	0.72	0.02	0.61	0.00	
Flexy	Perc10	0.03	0.20	0.02	0.25	0.00	0.18	0.00	8099
	Perc90	1.17	0.78	0.17	0.59	0.04	0.42	0.00	
Mich	Perc10	0.02	0.30	0.02	0.10	0.00	0.44	0.02	10994
	Perc90	0.06	0.78	0.11	0.57	0.02	0.68	0.26	

This table shows also the number of grasp candidates generated per each hand model. Figure 6.3 shows graphically a comparative between grasp quality metrics for different hand models.

Table 6.12 shows the AI for every hand, using both methods for hand configuration sampling, for the *Michelangelo* hand this study was not needed as in this hand model all possible configurations can be calculated. This values are shown in Figure 6.4. This figures show the 2D latent space of the hand models projected over the human hand.



(A) Q_{A1}

(C) Q_{B1}

(D) Q_{B2}

(E) Q_{B3}

(F) Q_{C2}

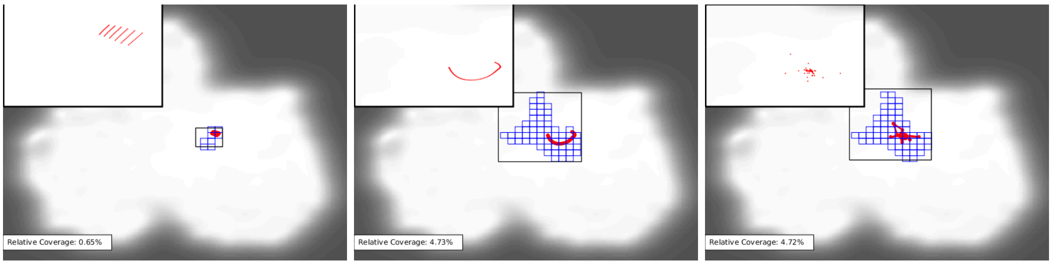
(G) Q_{D1}

(H) Q_{D2}

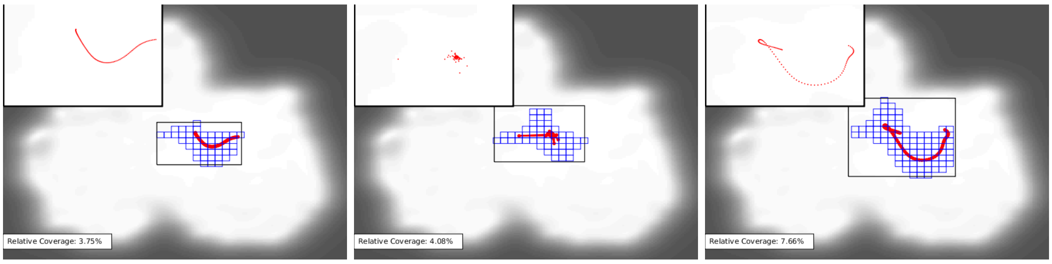
FIGURE 6.3: Percentiles of Quality Metrics for Artificial Hands.

TABLE 6.12: Anthropomorphism Index Results

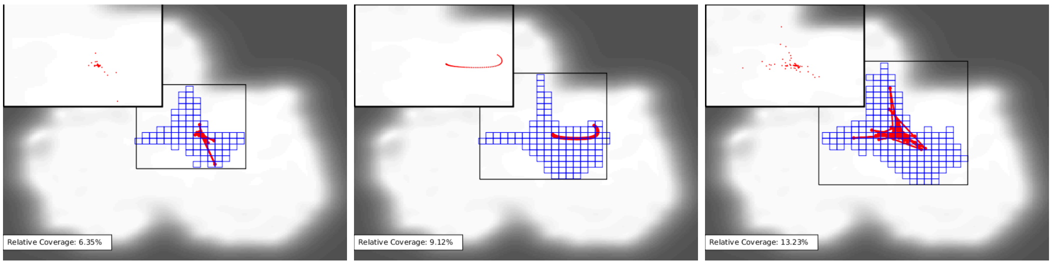
Hand	Random Sampling	Grasp Sampling	
Model	AI ± std	AI	Set Size
Mich	0.65 ± 0.00	NA	NA
Flexy	4.65 ± 0.11	4.73	69821
Tact	3.91 ± 0.13	3.75	87346
BruJa	6.61 ± 0.28	7.66	88720
Shadow	14.10 ± 0.32	9.12	82718



(A) Michelangelo Grasp (B) Flexyhand Grasp (C) Flexyhand Random



(D) Tacthand Grasp (E) Tacthand Random (F) BruJa Grasp



(G) BruJa Random (H) Shadowhand Grasp (I) Shadowhand Random

FIGURE 6.4: 2D latent space of artificial hands projected over the human space. Red lines/dots denotes the space covered by the artificial hand. Latent spaces are showed for both sampling methods: grasp and random

Finally graphs in Figure 6.5 show the comparative between the AI of each hand, using the *Random* sampling method, and its mean quality for each quality metric. The data plot on these graphs are shown in Table 6.13.

TABLE 6.13: Quality Evaluation and AI in Hands models

Median	Shadow	Tact	BruJa	Flexy	Mich
Q_{A1}	0.582	0.259	0.185	0.304	0.037
Q_{B1}	0.499	0.517	0.459	0.453	0.534
Q_{B2}	0.043	0.057	0.056	0.069	0.053
Q_{B3}	0.418	0.473	0.510	0.426	0.338
Q_{C2}	0.017	0.012	0.007	0.012	0.003
Q_{D1}	0.291	0.522	0.395	0.277	0.546
Q_{D2}	0.004	0.000	0.000	0.001	0.077
AI	14.100	3.910	6.610	4.650	0.650

6.3.4 Discussion

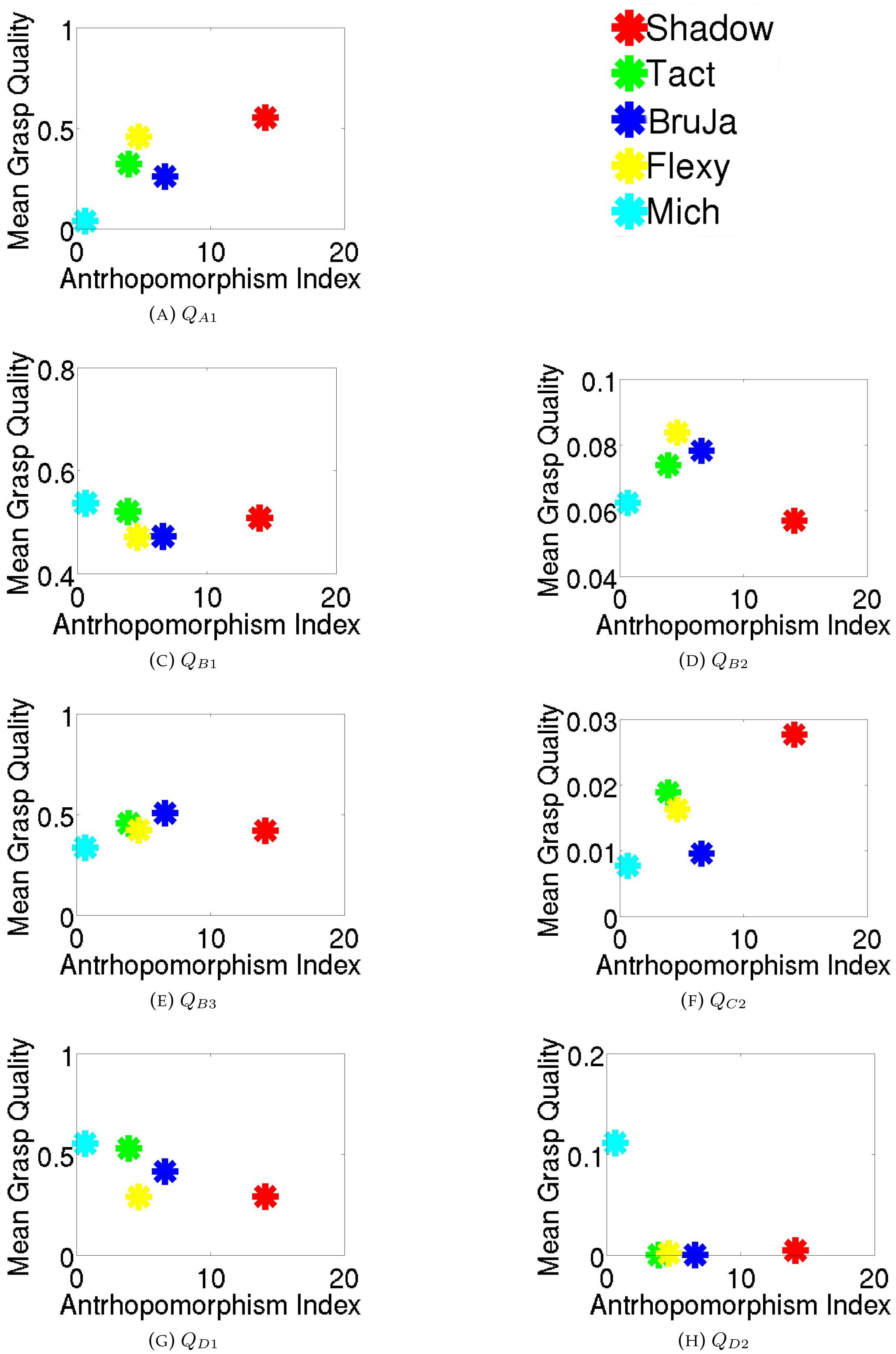
The results shown in Figure 6.3 and Table 6.11 with the distribution of each quality metric for every hand are relevant because indicate the capability of each hand to produce good grasping configurations. In general there are no significant differences between hands, with only small variations.

The AI has shown to be strongly dependent on two factors (Table 6.12), the latent space used for making comparisons and the parameters used for generating the different configurations of the hand. Two different methods for sampling hand configurations have been tested. The grasp based method, which is more related to real grasp configurations, only gives better results in a few models, being the *random* sampling method also a good choice for generating hand configurations. A more deep analysis of this metric should be made, in order to more accurately define the parameters needed for the step size in the range of the DoFs and the set size for generating configurations.

For this study, the latent space used in the AI was related to evaluate configurations used to make certain type of grasps, which means there are a lot of different human poses that are not taken in account. This makes it unable to evaluate other common human hand configurations, such from a person typewriting or playing piano. Thus, is really important to define a full latent space not only limited to grasps in Activities of Daily Life, but also common hand gestures and movements from humans.

The AI measure gives an assessment of how *human* the configurations of a artificial hand can be, even though it is dependent of the latent space defined to make the comparison, it is a good metric for evaluating the anthropomorphism and the behavior of a hand in different situations. So, in spite of its limitations it can be used to optimize the design of artificial hands , focusing in mimic only certain configurations of the human hand or making them task oriented.

Finally the comparison between both methodologies for evaluating hands shows there is not a clear relation between the anthropomorphism of a hand and its grasp qualities. Although the AI could be used to design more human hands, the grasping ability related to this index should still be proven. Similar real experiments as the



(A) Q_{A1}

(C) Q_{B1}

(D) Q_{B2}

(E) Q_{B3}

(F) Q_{C2}

(G) Q_{D1}

(H) Q_{D2}

FIGURE 6.5: Comparative between AI and Grasp Quality.

performed in Chapter 5 will provide a better insight on whether more human-like artificial hands are better manipulators.

6.4 Grasp success prediction and evaluation of artificial hands

In this study we propose a methodology capable to evaluate the expected performance of an artificial hand, using a grasp success prediction model. From the previous chapter (Chapter 5) it was shown it is possible to train and generate a prediction model capable to classify, with high success rate (>80%), the outcome of real grasp executions.

This prediction model uses a 3-categories scale to classify grasps as *Robust* (always succeeds), *Futile* (never works) and *Fragile*(Its success depends on precise contact points, object characteristics or gravity orientation). This prediction model had its limitations: it was tested with only two different grippers and a small set of objects. More important, results showed generalize this model to other grippers is still not advisable. But, what if we consider a prediction model trained with more gripper/hand designs and a huge set of objects? If this is the case, we could assume, this prediction model would be capable of evaluate grasp candidates with any kind of artificial hand and object.

In this section we propose to apply the grasp prediction model trained in the previous chapter as an evaluator for artificial hand designs. We will assume it has been trained and tested with a variety of robotic hands and objects, and the success rate remains as in the previous chapter (>80%). The objective in this study is not to provide a specific evaluation of different artificial hands, but to present a methodology capable of doing such evaluation.

Disclaimer The study done in this section is purely hypothetical. We assume there is a grasp prediction model capable to classify the outcome of grasp executions with an accuracy higher than 80%. Results obtained for each hand model are not reliable and conclusions about them should not be done.

6.4.1 Methodology

First, we will simulate and evaluate a wide variety of grasp poses and objects with an artificial hand. For this purpose we will use the grasps database generated in Chapter 4 and extended in the first study of this chapter (see Section 6.2). We have added also to this study the *BruJa* used in the previous study (see Section 6.3) with the anthropomorphism index.

Then, we will classify these simulated grasps with our trained model. Results will provide an estimation on the number of grasps which will be *Robust*, *Fragile* or *Futile* with each artificial hand. With this method we will be able to evaluate independently different artificial hands and compare them.

Our grasp database contains 10 different hand models and 126 objects. For each hand an object we selected up to 1.000 random grasps to be evaluated. In some models (tacthand, flexyhand and bruja) it was not possible to obtain this thousand random grasps.

The hand models to be evaluated in this section are: *PR2* hand, *Barrett* hand, Schunk *SDH* hand, *Model-T*, Schunk *SAH* hand, *Michelangelo* hand, *Shadow* hand, *Flexy*-hand, *BruJa*-hand and *Tact* hand. The first four models correspond to *classic* robotic grippers. The other six models correspond to anthropomorphic hands. See section 3.2 for more details of the hands.

To predict the outcome of real grasp executions, we will use the classification model generated in Chapter 5.6. This prediction model uses a 3-categories scale to classify grasp candidates as *Robust*, *Futile* and *Fragile*.

6.4.2 Results

Table 6.14 shows the distribution of *Robust*, *Fragile* and *Futile* grasps among different hand models. Results are reported in percentage of grasps per category. Also the total number of generated grasps per hand model is reported.

Graphs in Figure 6.6 show the distribution of grasps on each category and hand model.

TABLE 6.14: Summary of grasps distribution per each hand model. Results shows percentage and number of grasps on each category.

Hand	Distribution of Grasps (%)			Number
Model	Futile	Fragile	Robust	of grasps
PR2	0.08	0.56	0.36	126000
Barrett	0.16	0.29	0.55	126000
SDH	0.44	0.30	0.26	126000
ModelT	0.13	0.32	0.56	126000
SAH	0.13	0.32	0.55	126000
Michelangelo	0.15	0.69	0.17	126000
Shadow	0.05	0.30	0.66	126000
Flexy	0.08	0.30	0.62	84744
Deval	0.20	0.35	0.45	105911
Tact	0.26	0.34	0.40	125319

6.4.3 Discussion

As shown results obtained in this study. The *Shadow* hand seems to be the best model for grasping as it has the highest success rate (66%) and lowest failure rate (5%). The *Flexy*-hand model showed also similar values to the *Shadow hand*. These two models of anthropomorphic hands seemed to be the best options for grasping.

On the other hand, the Schunk *SDH* hand model showed the worst performance, with a high failure rate (44%) and a low success rate (26%). The *Michelangelo* hand model showed a lower success rate (17%), but it has a high number of *Fragile* grasps (69%) meaning most of these grasp candidates could be good options for grasping.

It is important to take in account in this analysis, fragile grasps could easily turn into successful grasps. This would depend in first, the error detection and grasp reconfiguration for robotic grippers and second, the human reaction and manipulability with the prosthetic hands.

In the case of prosthetic hands, we can assume humans would react and adapt easily to overcome these fragile grasps and turn them into successful grasps. For robotic grippers, it would depend in the *intelligence* implement in the robotic system.

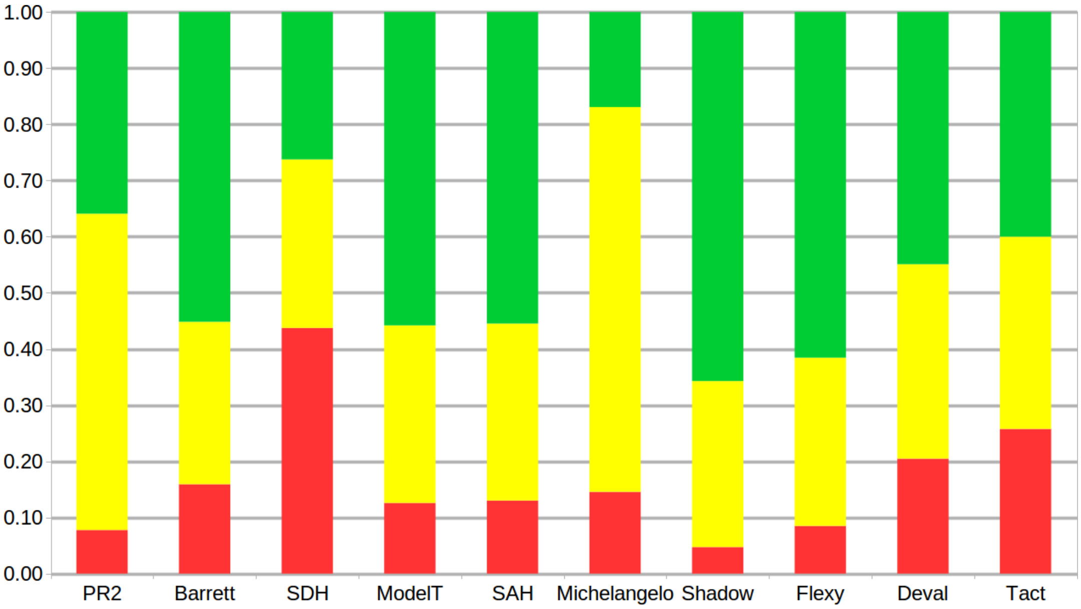


FIGURE 6.6: Distribution of grasps on each category for different artificial hand models. Green: Robust, Yellow: Fragile, Red: Futile.

6.5 Hand design improvement

The previous study presented a methodology to evaluate the grasping capabilities of artificial hands. In this study we aim to exploit this methodology to evaluate the design of a low-cost prosthetic hand.

Prosthetic hands are intended for human use. They are designed to help amputees and one armed people on their activities of daily life (ADL). The YCB-Dataset (see Section 3.3) contains up to 77 different object models of daily life with different shapes, sizes, textures, weight and rigidity. In this study we will simulate and evaluate different grasps with this object set.

Disclaimer The study done in this section is purely hypothetical. We assume there is a grasp prediction model capable to classify the outcome of grasp executions with an accuracy higher than 80%. Results obtained for each hand design are not reliable and conclusions about them should be done carefully.

6.5.1 Methodology

In this study we will evaluate different designs of the Imma Hand 3.2. Our objective is to improve the original design of the Imma hand. We will evaluate the performance of the hand with different orientations of the thumb joints. We will change the orientation of the thumb *cmc-joint* from 0 to 45 degrees and the *mcp-joint* from 0 to 90 degrees, with a step size of 15 on each joint (Figure 6.7).

As this is a prosthetic hand, intended to be used by humans, we have selected common objects from the *ADLs*. We have selected 24 different objects from the *YCB Dataset* (Figure 6.8). Following the grasp generation method described in Chapter 3.4, we have used a simple procedure to generate up to 100 different grasp candidates per object. Details on the parameters used for generating the grasps are shown in Table 6.15.

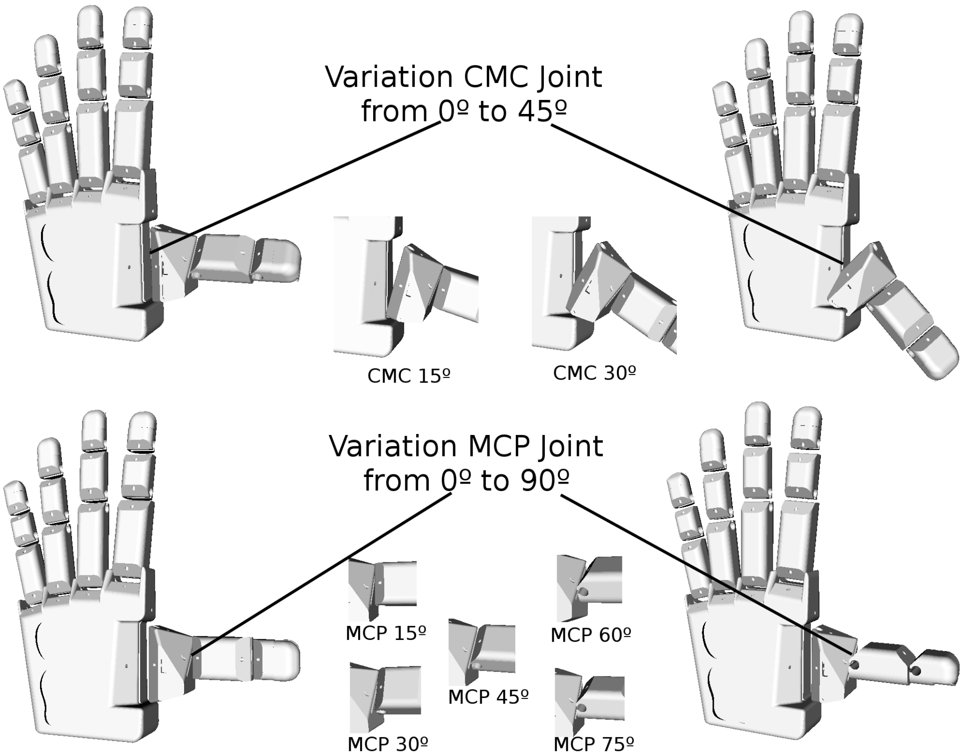


FIGURE 6.7: CMC and MCP joints to be modified in the Imma Hand.

In total, 28 different designs for the same hand model and up to 2400 grasp candidates per design will be evaluated.

TABLE 6.15: Parameters selected to generate grasp hypotheses

	Parameter	Value
δ	Delta Space	0.01 m
α	Alpha Angle	$[0, \pi/3]$
θ	Standoff Distance	$[0.01\ \mathrm{m}, 0.02\ \mathrm{m}]$
N_r	Gripper Rotations	$[0, \pi/2, \pi, 3\pi/2, 2\pi]$

6.5.2 Results

We have split our results among the different orientations of the thumb cmc joint. Table 6.16 shows the values and comparison of grasp successfulness for the thumb joint cmc 0°and the mcp from 0°to 90°. Table 6.17 shows the values and comparison of grasp successfulness for the thumb joint cmc 15°and the mcp from 0°to 90°.

Table 6.18 shows the values and comparison of grasp successfulness for the thumb joint cmc 30°and the mcp from 0°to 90°. Table 6.19 shows the values and comparison of grasp successfulness for the thumb joint cmc 45°and the mcp from 0°to 90°.

Tables show the number of grasps on each category and the total number of grasp candidates generated with the params specified on Table 6.15.

6.5.3 Discussion

Results on tables 6.16 to 6.19 showed our hand evaluation methodology could help to improve the design of a hand prosthesis. On each table we remarked the best

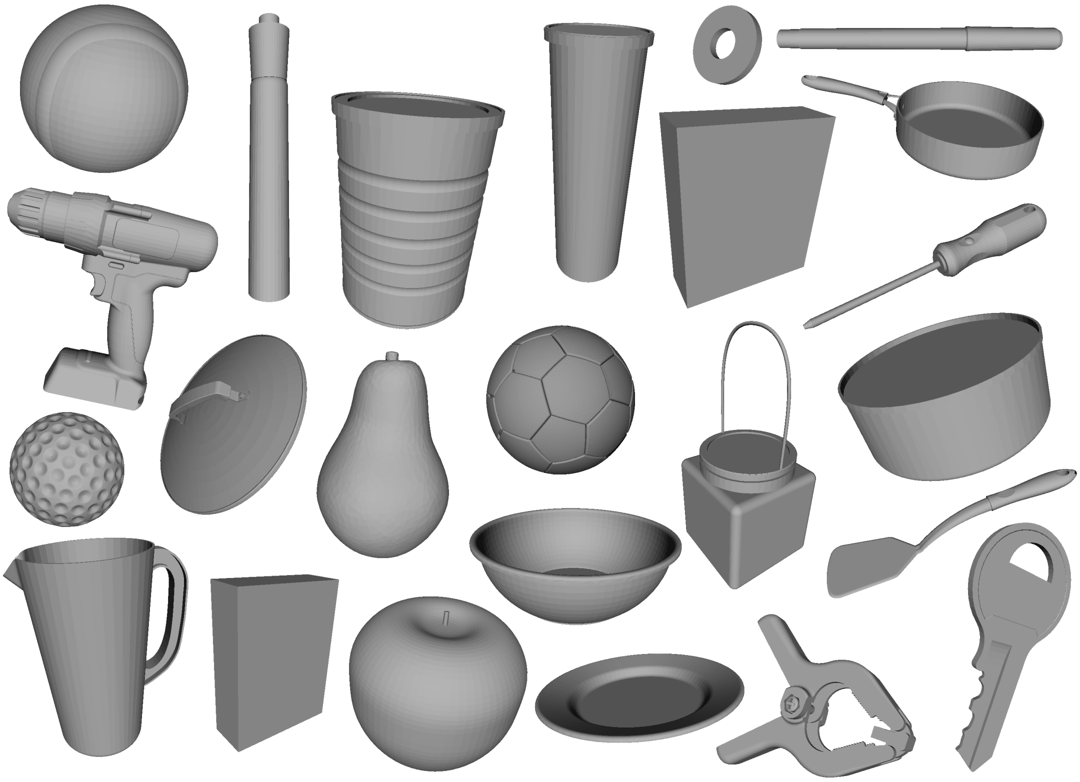


FIGURE 6.8: Sample of the different objects used from the YCB Database.

TABLE 6.16: Summary of grasps distribution for Joint CMC 0. Graph shows percentage of grasps on each category.

MCP	Futile	Fragile	Robust	Total
0°	0.09	0.28	0.63	1991
15°	0.08	0.32	0.60	2010
30°	0.10	0.46	0.44	2040
45°	0.11	0.43	0.46	2038
60°	0.11	0.46	0.43	2032
75°	0.08	0.31	0.61	2046
90°	0.09	0.32	0.59	2010

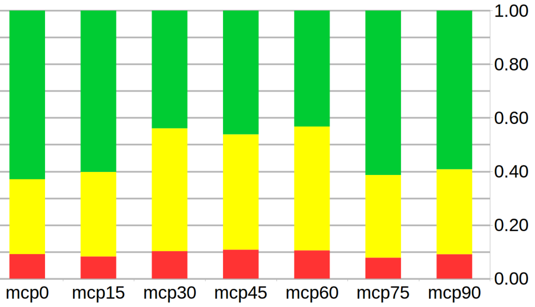


TABLE 6.17: Summary of grasps distribution for Joint CMC 15. Graph shows percentage of grasps on each category.

MCP	Futile	Fragile	Robust	Total
0°	0.08	0.32	0.60	2043
15°	0.09	0.30	0.61	2063
30°	0.08	0.30	0.62	2071
45°	0.09	0.29	0.63	2068
60°	0.07	0.32	0.61	2076
75°	0.08	0.32	0.60	2034
90°	0.06	0.36	0.59	2039

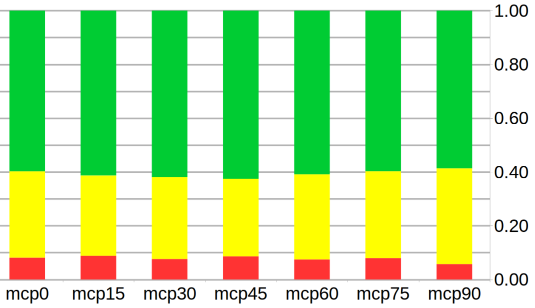


TABLE 6.18: Summary of grasps distribution for Joint CMC 30. Graph shows percentage of grasps on each category.

MCP	Futile	Fragile	Robust	Total
0°	0.06	0.33	0.61	2027
15°	0.05	0.32	0.62	2003
30°	0.06	0.32	0.62	1975
45°	0.07	0.32	0.62	1944
60°	0.07	0.34	0.59	2044
75°	0.06	0.35	0.59	2050
90°	0.06	0.37	0.57	2006

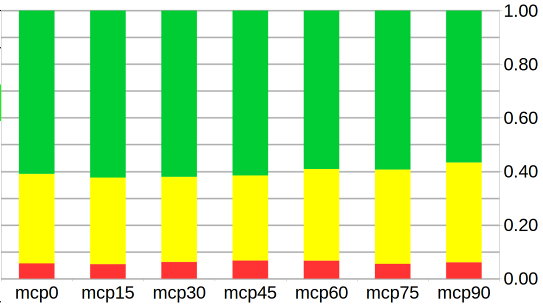
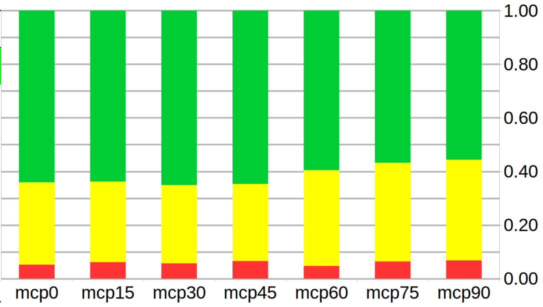


TABLE 6.19: Summary of grasps distribution for Joint CMC 45. Graph shows percentage of grasps on each category.

MCP	Futile	Fragile	Robust	Total
0°	0.05	0.31	0.64	2008
15°	0.06	0.30	0.64	1990
30°	0.06	0.29	0.65	1955
45°	0.07	0.29	0.65	1931
60°	0.05	0.36	0.60	2060
75°	0.06	0.37	0.57	2009
90°	0.07	0.38	0.56	2001



design for the different orientation of the Thumb.

The criteria applied to select the best design was, first, we compared the rate of *Robust* and *Futile* grasps. Then we checked the number of grasp hypotheses generated. It is important to note we can not rely in only one criteria for selecting the best design.

In the case of Thumb Joint CMC 45 (Table 6.19) it was not very clear the best design, as four different models (mcp0, mcp15, mcp35 and mcp45) had a high success rate ($> 60\%$). In this case we considered as criteria for selecting the best design the number of grasp hypotheses generated.

In this study we applied a methodology to evaluate artificial hands to different designs of the same prosthetic hand. We considered as the main criteria for selecting a design the success rate of robust grasps. However, other criteria such as the number of grasp hypotheses generated or the rates of fragile and futile grasps could be taken in account.

Our results rely on a prediction model which has to be tested with more objects and hands, thus, conclusions on which is the best hand design could not be valid. This study aimed to show how our methodology to evaluate artificial hand can help to improve the design of prosthetic hands.

6.6 Conclusions

6.6.1 Artificial hand properties

In this experiment we applied and exhaustive analysis on the characteristics of different artificial hands. We analyzed the quality of different hands according to specific grasp metrics. These metrics evaluate independent aspects of the grasp which we consider as hand properties.

The solely evaluation with quality metrics can't provide information about the performance of an artificial hand, or how good it is for grasping. However, it could be used as benchmark for comparing different hand models. Moreover, it could be used as a method to find similarities in the performance between hand models.

The evaluation of grasp candidates with quality metrics showed different combinations of metric values could provide *good* grasp solutions. Thus, a hand capable of achieve a large range of different quality values in different metrics is more suitable to make *successful* grasps.

6.6.2 Comparison with hand indexes

In the analysis with the Anthropomorphic Index, two different approaches to evaluate the functional performance of anthropomorphic artificial hands have been studied and compared. The anthropomorphism index (AI) focused on how an artificial hand can replicate the workspace of a human hand. On a different trend, robotic quality metrics (QM) offer different alternatives to asses how good is a grasp performed by a robot hand on an object. Our results show that the assessments provided by both approaches are basically independent. That is, it does not matter how anthropomorphic is the design of an artificial hand, in the sense defined by the AI, in order to obtain more stable grasps.

A final remark on this study is that it has been constrained to simulation. The quality assessments provided by either the quality metrics or the anthropomorphism index has not been validated nor compared against assessments obtained with real hands on real scenarios. Therefore the results obtained should be taken with caution. Preparing proper experimental scenario and carrying out exhaustive prototype experiments should be definitively the next step on this research.

6.6.3 Grasp success prediction and evaluation of artificial hands

In the last two studies we presented a methodology to evaluate hand models based on its capability to achieve successful grasps. The methodology relies on a prediction model capable to evaluate grasp candidates using a 3-categories scale to classify them as *Robust*, *Futile* and *Fragile*.

All the experiments performed in this studies are based on a prediction model tested only with two hand models and a few objects. Thus, results obtained are not conclusive and the evaluation and comparison between different hand models should be tested with real experiments.

The main objective of this section was not to provide an evaluation of different hand models, but to present a methodology capable of this procedure. As show the experiment comparing different hand designs, this methodology could be applied in the process of designing hand prosthesis. This methodology, with the proper prediction model, is also suitable to different sets of objects and task oriented designs of artificial hands.

6.6.4 Limitations

The main limitation on this Chapter is the lack of a ground-truth of real experiments for the evaluation of artificial hands. The evaluation with metrics and the anthropomorphic index require real experimentation to relate the evaluation in simulation and the real performance of artificial hands.

The evaluation using the classification model for predicting the grasp success also require a wider ground-truth. The real experiments in which rely this evaluation method were done with only two different grippers and a few objects. Thus, is needed a larger set of real experiments with different artificial hand and object models.

6.6.5 Future & Related Work

As future work, it is necessary to evaluate the artificial hands analyzed in this chapter in real world. Thus, we should define a methodology to evaluate the grasping capabilities of different hand models and compare them to the evaluation done with quality metrics and the grasp success prediction model.

This applies also to evaluate the different designs of the *Imma* hand. As this is a prosthetic hand intended to be built using 3D-printed materials, construct and test in real world the different designs evaluated in this chapter should be the next step to do.

The studies, methodology and results presented in this chapter were total or partially published in Rubert et al., 2014; Rubert and Morales, 2016; Rubert et al., 2018[2].

Also, the use of OpenHand and the methodologies described in this Chapter have been used to evaluate prosthetic hand designs in: Llop-Harillo et al., 2017; Llop et al., 2017; Cardín-Catalán et al., 2016; Llop-Harillo et al., 2018; Andrés et al., 2018.

---

[2] In Rubert and Morales, 2016 the BruJa hand is referred as *Devalhand*

Chapter 7

Conclusions

This chapter highlights the main contributions of the thesis. As well as a list of publications and some ideas on future lines of research.

7.1 Contributions

7.1.1 Quality metrics analysis

In this chapter we realized a deep analysis on the characteristic of 10 different grasp quality metrics. The main contributions are:

1. We analyzed and defined a method to generate grasp hypotheses to avoid biases.
2. We provided an statistical study to obtain thresholds and ranges of values in different metrics.
3. We obtained a sensitivity index of different metrics against pose errors.
4. We established correlations between metrics.
5. **We obtained 7 independent grasp metrics for evaluate grasp hypotheses.**

7.1.2 Grasp success prediction

In this chapter we analyzed the prediction capabilities of different quality metrics. The contributions on this chapter are:

1. We studied the prediction capabilities of humans and dynamic simulation.
2. We analyzed different machine learning algorithms to define a grasp prediction model using quality metrics as input feature vector.
3. We found non-linear and non-parametric algorithms work better for combining grasp metrics and predict grasp success.
4. We analyzed how the object's weight and gravity orientation influence the grasp success.
5. **We propose a grasp prediction model able to classify grasp candidates as *Robust, Fragile* and *Futile*.**

7.1.3 Artificial hands evaluation

In this chapter we used quality metrics as evaluators of hand models. The main contributions of this chapter are:

1. We defined grasp metrics as evaluators of hand properties.
2. We proposed a method for benchmarking and compare hand models.
3. We compared the evaluation with metrics to the anthropomorphism of hand prosthesis.
4. **We defined a methodology to evaluate the grasp capability of different hands.**
5. We applied this methodology to improve the design of a hand prosthesis.

7.2 Future work

As future work, we should consider three lines of research.

Grasp metrics Our studies considered only 10 grasp quality metrics, but there are more in the literature. The studies on Chapter 4 should be repeated considering new metrics and analyze possible correlations between them.

Grasp prediction The ground-truth to define a prediction model considered only two grippers and a few objects. More real experiments with other gripper and objects should be definitely done. Also, it should be considered a more deep analysis on the object characteristics and influence of the gravity on the grasp success.

Hand evaluation This study relied in simulation. It is necessary to perform real experiments to evaluate the characteristics and grasping capabilities of different hand models, and how are related to quality metrics and methodologies proposed.

7.3 Related publications

The studies and experiment exposed in this thesis were total or partially published in:

Journal papers

1. C. Rubert, B. León, A. Morales, J. Sancho-Bru (2018). "Characterisation of grasp quality metrics". *Journal of Intelligent & Robotic Systems*.
2. J. Andrés, A. Pérez-González, C. Rubert, J. Fuentes, B. Sospedra (2018). "Comparison of grasping performance of tendon and linkage transmission systems in an electric powered low cost hand prosthesis". *Journal of mechanisms and Robotics*. [1]

---

[1] Accepted. Pendent publishing.

Conference Papers

1. B. Leon, C. Rubert, J. Sancho-Bru, and A. Morales (2013). "Evaluation of prosthetic hands prehension using grasp quality measures". *IEEE International Conference on Intelligent Robots and Systems (2013).*

2. B. Leon, C. Rubert, A. Morales, and J. Sancho-Bru (2014). "Characterization of Grasp Quality Measures for Evaluating Robotic Hands Prehension". *IEEE International Conference on Robotics and Automation (2014).*

3. C. Rubert, B. León, A. Morales (2014)."Grasp Quality Metrics for Robot Hands Benchmarking". *2014 IEEE/RSJ International Conference on Humanoid Robots*

4. C. Rubert, A. Morales (2016). "Comparison between grasp quality metrics and the anthropomorphism index for the evaluation of artificial hands". *IEEE - BioRob 2016. Singapur.*

5. D. Cardín-Catalán, M.C. Mora, J. Andrés, B. Sospedra, C. Rubert (2016). "Diseño y simulación de un prototipo de mano protésica de bajo coste". *XXI Congreso Nacional de Ingeniería Mecánica, Elche.*

6. C. Rubert, A. Morales (2016). "Evaluación de manipuladores robóticos mediante métricas de calidad del agarre". *Jornadas de Automática 2016, Madrid.*

7. I. Llop-Harillo, A. Pérez-González, V. Gracia-Ibañez, C. Rubert (2017). "Design and evaluation of anthropomorphic hands". *Summer School on soft manipulation. July, 17th-21st 2017 Lake Chiemsee, Germany.*

8. I. Llop, A. Pérez, C. Rubert (2017). "Anthropomorphism index of opposition for hand prostheses". *23rd Congress of the European Society of Biomechanics, Jul 2017, Sevilla.*

9. C. Rubert, D. Kappler, A. Morales, S. Schaal, J. Bohg (2017). "On the relevance of grasp metrics for predicting grasp success." *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017), Sep 2017, Vancouver.*

10. I. Llop-Harillo, C. Rubert, A. Pérez-González (2018). "Computation of grasp quality metrics in OpenHand simulator to improve a 3-D printed prosthetic hand". *Computer Methods in Biomechanics and Biomedical Engineering Congress, Lisboa.*

Conclusiones

Este capítulo destaca las principales contribuciones de la tesis. Así como una lista de publicaciones y algunas ideas para futuras líneas de investigación.

7.4 Contribuciones

7.4.1 Análisis de métricas de calidad

En este capítulo hemos realizado un exhaustivo análisis sobre las características de 10 métricas de calidad del agarre distintas. Las contribuciones principales son:

1. Hemos analizado y definido un método para generar hipótesis de agarre para evitar sesgos.
2. Se ha proporcionado un estudio estadístico para obtener los umbrales y rangos de valores en distintas métricas.
3. Hemos calculado un índice de sensibilidad de las distintas métricas ante la presencia de errores de posición.
4. Se han establecido correlaciones entre métricas.
5. **Hemos obtenido 7 métricas independientes para evaluar las hipótesis de agarre.**

7.4.2 Predicción del éxito del agarre

En este capítulo hemos analizado la capacidad de predicción de distintas métricas de calidad. Las contribuciones del capítulo son:

1. Hemos estudiado la capacidad de predicción de humanos y la *métrica física* en simulación.
2. Hemos comparado distintos métodos de clasificación para definir un modelo de predición del agarre, utilizando métricas como vector de características de entrada.
3. Hemos comprobado que los algoritmos no lineales y no paramétricos funcionan mejor para combinar métricas y predecir el éxito del agarre.
4. Hemos analizado como el peso del objeto y la orientación de la gravedad influyen en el éxito del agarre
5. **Hemos propuesto un modelo de predicción del agarre capaz de clasificar agarres candidatos como *Robustos*, *Frágiles* y *Futiles*.**

7.4.3 Evaluación de manos artificiales

En este capítulo hemos utilizado las métricas de calidad como evaluadores de modelos de manos. Las contribuciones principales del capítulo son:

1. Hemos definido las métricas de agarre como evaluadores de distintas propiedades de las manos.
2. Hemos propuesto un método para establecer valores de referencia y comparar modelos de mano.
3. Hemos comparado la evaluación con métricas con el antropomorfismo de manos protésicas.
4. **Hemos definido una metodología para evaluar las capacidades de agarre de distintas manos.**
5. Hemos aplicado dicha metodología para mejorar el diseño de una mano protésica.

7.5 Trabajo Futuro

Como trabajo futuro, deberíamos considerar tres líneas de trabajo:

Métricas del agarre Nuestros estudios han considerado sólo 10 métricas de calidad del agarre, pero hay muchas más en la literatura. Los estudios en el Capítulo 4 podrían ser repetidos considerando nuevas métricas y analizando posibles correlaciones entre ellas.

Predicción del agarre Los datos de base para definir un modelo de predición utilizan sólo dos manipuladores y unos pocos objetos. Más experimentación real con otros manipuladores y objetos debe ser realizada. Se debe considerar también, realizar un análisis más profundo en las características de los objetos y la influencia de la gravedad en el éxito del agarre.

Evaluación de manos Este estudio está basado en simulación. Es necesario realizar experimentos reales para evaluar las características y capacidades de agarre de los distintos modelos de manos, y como están relacionadas con las métricas de calidad y las metodologías propuestas.

7.6 Publicaciones relacionadas

Los estudios y experimentos propuestos en esta tesis han sido total o parcialmente publicados en:

1. B. Leon, C. Rubert, J. Sancho-Bru, and A. Morales (2013). "Evaluation of prosthetic hands prehension using grasp quality measures". *IEEE International Conference on Intelligent Robots and Systems (2013).*
2. B. Leon, C. Rubert, A. Morales, and J. Sancho-Bru (2014). "Characterization of Grasp Quality Measures for Evaluating Robotic Hands Prehension". *IEEE International Conference on Robotics and Automation (2014).*

3. C. Rubert, B. León, A. Morales (2014)."Grasp Quality Metrics for Robot Hands Benchmarking". *2014 IEEE/RSJ International Conference on Humanoid Robots*
4. C. Rubert, A. Morales (2016). "Comparison between grasp quality metrics and the anthropomorphism index for the evaluation of artificial hands". *IEEE - BioRob 2016. Singapur.*
5. D. Cardín-Catalán, M.C. Mora, J. Andrés, B. Sospedra, C. Rubert (2016). "Diseño y simulación de un prototipo de mano protésica de bajo coste". *XXI Congreso Nacional de Ingeniería Mecánica, Elche.*
6. C. Rubert, A. Morales (2016). "Evaluación de manipuladores robóticos mediante métricas de calidad del agarre". *Jornadas de Automática 2016, Madrid.*
7. I. Llop-Harillo, A. Pérez-González, V. Gracia-Ibañez, C. Rubert (2017). "Design and evaluation of anthropomorphic hands". *Summer School on soft manipulation. July, 17th-21st 2017 Lake Chiemsee, Germany.*
8. I. Llop, A. Pérez, C. Rubert (2017). "Anthropomorphism index of opposition for hand prostheses". *23rd Congress of the European Society of Biomechanics, Jul 2017, Sevilla.*
9. C. Rubert, D. Kappler, A. Morales, S. Schaal, J. Bohg (2017). "On the relevance of grasp metrics for predicting grasp success." *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017), Sep 2017, Vancouver.*
10. C. Rubert, B. León, A. Morales, J. Sancho-Bru (2018). "Characterisation of grasp quality metrics". *Journal of Intelligent & Robotic Systems.*
11. I. Llop-Harillo, C. Rubert, A. Pérez-González (2018). "Computation of grasp quality metrics in OpenHand simulator to improve a 3-D printed prosthetic hand". *Computer Methods in Biomechanics and Biomedical Engineering Congress, Lisboa.*

Appendix A

Quality Metrics Analysis

A.1 Numeric Results from Sensitivity Analysis

Table A.1 details the GSI_N values for every metric computed for the reference grasps, distinguishing between hand models. This table is related to the grapsh on section 4.3.

TABLE A.1: Comparison of GSI_N for each hand and metric

	PR2	Barrett	Schunk	Mich	SAH	ModelT	Shadow
Q_{A1}	18.28%	17.90%	23.54%	3.19%	25.14%	22.89%	18.33%
Q_{A2}	8.55%	2.68%	28.36%	1.09%	72.69%	32.76%	21.59%
Q_{A3}	20.63%	33.49%	22.93%	3.72%	20.19%	21.83%	17.54%
Q_{B1}	12.63%	11.93%	10.31%	8.95%	11.99%	13.32%	8.39%
Q_{B2}	19.24%	17.36%	27.74%	13.10%	29.87%	26.20%	24.41%
Q_{B3}	29.44%	20.28%	36.37%	14.91%	27.57%	35.67%	23.94%
Q_{C1}	27.10%	6.27%	23.37%	4.00%	34.24%	35.62%	18.38%
Q_{C2}	16.76%	8.70%	12.37%	5.00%	23.07%	39.63%	24.45%
Q_{D1}	5.18%	21.32%	15.38%	13.30%	12.68%	12.36%	4.92%
Q_{D2}	1.60%	112.34%	0.00%	29.00%	1.80%	2.57%	0.62%

Table A.2 details the GSI_N values for every metric computed for the reference grasps, distinguishing between good or bad reference grasps and between those variations that improve or worsen the metric values from the reference grasp. This table is related to the grapsh on section 4.3.

TABLE A.2: GSI_N calculated for each quality metric distinguishing between good or bad reference grasps and variations that improve or worsen the metric values

		Good Quality	Bad Quality	All Grasps
Q_{A1}	All Variations	28.36%	11.57%	11.64%
	Improve	19.89%	21.11%	10.95%
	Worsen	33.89%	4.11%	10.87%
Q_{A2}	All Variations	42.72%	3.98%	1.46%
	Improve	43.21%	6.62%	1.18%
	Worsen	44.57%	1.28%	1.44%
Q_{A3}	All Variations	28.01%	16.25%	10.30%
	Improve	22.63%	31.29%	11.54%
	Worsen	33.43%	4.56%	8.86%
Q_{B1}	All Variations	11.27%	11.24%	6.36%
	Improve	9.44%	12.62%	6.08%
	Worsen	12.94%	9.95%	6.47%
Q_{B2}	All Variations	31.38%	12.63%	5.55%
	Improve	26.89%	25.15%	4.59%
	Worsen	40.66%	11.38%	5.23%
Q_{B3}	All Variations	28.32%	20.95%	15.26%
	Improve	23.40%	37.84%	15.66%
	Worsen	39.94%	18.57%	14.79%
Q_{C1}	All Variations	28.27%	10.96%	9.09%
	Improve	24.53%	33.37%	9.24%
	Worsen	31.88%	8.34%	9.49%
Q_{C2}	All Variations	33.13%	4.94%	1.36%
	Improve	30.66%	11.32%	1.22%
	Worsen	37.86%	5.76%	1.45%
Q_{D1}	All Variations	12.99%	14.19%	9.94%
	Improve	12.82%	17.82%	12.30%
	Worsen	11.10%	7.73%	7.78%
Q_{D2}	All Variations	34.04%	2.18%	4.78%
	Improve	37.52%	8.59%	7.24%
	Worsen	32.08%	0.33%	8.12%

Table A.3 shows the distribution of the variations of good and bad reference grasps according to whether the variation improve or worsen the reference quality value. This table is related to graphs on section 4.3.

TABLE A.3: Distribution of the variation of reference grasps (good or bad) according to whether the variation improve or worsen the reference quality value.

		Good Quality	Bad Quality	All Grasps
Q_{A1}	All Variations	53.12%	46.88%	
	Improve	33.60%	52.29%	43.22%
	Worsen	66.39%	46.88%	56.20%
Q_{A2}	All Variations	58.40%	41.60%	
	Improve	34.92%	50.54%	41.95%
	Worsen	65.08%	49.35%	57.86%
Q_{A3}	All Variations	52.64%	47.36%	
	Improve	34.07%	54.41%	44.66%
	Worsen	65.90%	44.48%	54.60%
Q_{B1}	All Variations	54.71%	45.29%	
	Improve	43.48%	54.41%	46.81%
	Worsen	56.45%	45.53%	50.25%
Q_{B2}	All Variations	51.51%	48.49%	
	Improve	30.74%	64.92%	30.16%
	Worsen	69.18%	34.97%	36.86%
Q_{B3}	All Variations	42.37%	57.63%	
	Improve	28.39%	72.34%	34.09%
	Worsen	71.11%	27.31%	32.71%
Q_{C1}	All Variations	59.91%	40.09%	
	Improve	43.91%	68.10%	31.77%
	Worsen	55.93%	19.13%	34.00%
Q_{C2}	All Variations	49.31%	50.69%	
	Improve	37.03%	50.27%	28.25%
	Worsen	62.70%	46.78%	37.97%
Q_{D1}	All Variations	60.33%	39.67%	
	Improve	48.85%	63.84%	53.51%
	Worsen	51.05%	33.90%	45.72%
Q_{D2}	All Variations	64.99%	35.01%	
	Improve	42.65%	29.02%	37.88%
	Worsen	56.84%	10.94%	40.90%

Appendix B

Grasp prediction for different combinations of metrics

B.1 Precission results for best combinations of metrics

This appendix shows the results on prediction models using different combinations of grasp metrics. Except for combinations of two metrics, only models with high precision ($>= 80\%$) are reported.

TABLE B.1: Best results on models trained using combinations of two different metrics.

Classifier	Combined Metrics	CrossVal	Test
K-Nearest Neighbors	(Q_{A1}, Q_{D1})	0.73 ± 0.07	0.78
K-Nearest Neighbors	(Q_{B1}, Q_{D1})	0.72 ± 0.03	0.76
K-Nearest Neighbors	(Q_{B2}, Q_{D1})	0.72 ± 0.05	0.77

TABLE B.2: Best results on models trained using combinations of three different metrics.

Classifier	Combined Metrics	CrossVal	Test
K-Nearest Neighbors	(Q_{B1}, Q_{B3}, Q_{D1})	0.75 ± 0.03	0.80

TABLE B.3: Best results on models trained using combinations of four different metrics.

Classifier	Combined Metrics	CrossVal	Test
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B3}, Q_{D1})$	0.75 ± 0.06	0.81
K-Nearest Neighbors	$(Q_{B1}, Q_{B2}, Q_{C2}, Q_{D1})$	0.76 ± 0.03	0.80

TABLE B.4: Best results on models trained using combinations of five different metrics.

Classifier	Combined Metrics	CrossVal	Test
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B2}, Q_{B3}, Q_{D1})$	0.76 ± 0.07	0.80
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B3}, Q_{D1}, Q_{D2})$	0.76 ± 0.06	0.81
K-Nearest Neighbors	$(Q_{B1}, Q_{B2}, Q_{B3}, Q_{C2}, Q_{D1})$	0.77 ± 0.04	0.81
K-Nearest Neighbors	$(Q_{B1}, Q_{B3}, Q_{C2}, Q_{D1}, Q_{D2})$	0.76 ± 0.03	0.80

TABLE B.5: Best results on models trained using combinations of six different metrics.

Classifier	Combined Metrics	CrossVal	Test
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B2}, Q_{B3}, Q_{D1}, Q_{D2})$	0.77 ± 0.06	0.80
K-Nearest Neighbors	$(Q_{A1}, Q_{B1}, Q_{B3}, Q_{C2}, Q_{D1}, Q_{D2})$	0.75 ± 0.06	0.80
K-Nearest Neighbors	$(Q_{B1}, Q_{B2}, Q_{B3}, Q_{C2}, Q_{D1}, Q_{D2})$	0.77 ± 0.04	0.80

Bibliography

Aleotti, A. and S. Caselli. “Grasp recognition in virtual reality for robot pregrasp planning by demonstration”. In: *Proceedings - IEEE International Conference on Robotics and Automation* 2006 (2006), p. 2801.

Andrés, Javier, Antonio Pérez-González, Carlos Rubert, José Fuentes, and Bruno Sospedra. “Comparison of grasping performance of tendon and linkage transmission systems in an electric powered low cost hand prosthesis”. In: *Journal of Mechanisms and Robotics* (2018).

Balasubramanian, R., L. Xu, P. D. Brook, J. R. Smith, and Y. Matsuoka. “Physical Human Interactive Guidance: Identifying Grasping Principles From Human-Planned Grasps”. In: *IEEE Trans. on Robotics* 28.4 (2012), pp. 899–910.

Belter, Joseph T., Jacob L. Segil, Aaron M. Dollar, and Richard F. Weir. “Mechanical design and performance specifications of anthropomorphic prosthetic hands: A review”. In: *The Journal of Rehabilitation Research and Development* 50.5 (2013), p. 599.

Bicchi, A. and V. Kumar. “Robotic grasping and contact: a review”. In: *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*. IEEE, 2000, pp. 348–353.

Bicchi, Antonio. “On the Problem of Decomposing Grasp and Manipulation Forces in Multiple Whole-limb Manipulation”. In: *International Journal of Robotics and Autonomous Systems* 13 (1994), pp. 127–147.

Bicchi, Antonio. “On the Closure Properties of Robotic Grasping”. In: *International Journal of Robotics Research* 14 (1995), pp. 319–334.

Bicchi, Antonio. “Hands for dexterous manipulation and robust grasping: a difficult road toward simplicity.” In: *IEEE Transactions on Robotics and Automation* 16.6 (2000), pp. 652–662.

Birglen, Lionel, Thierry Laliberté, and Clément M Gosselin. *Underactuated robotic hands*. Vol. 40. Springer, 2007.

Bishop, Christopher M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

Bohg, Jeannette, Antoinio Morales, Tamim Asfour, and Danica Kragic. “Data-Driven grasps synthesis - A survey”. In: *IEEE Transactions on Robotics* 30.2 (2014), pp. 289 –309.

Boivin, E., I. Sharf, and M. Doyon. “Optimum grasp of planar and revolute objects with gripper geometry constraints”. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. Vol. 1. 2004, pp. 326 –332.

Bullock, Ian M., Thomas Feix, and Aaron M. Dollar. “The Yale Human Grasping Data Set: Grasp, Object and Task Data in Household and Machine Shop Environments”. In: *The International Journal of Robotics Research* 34 (2015), pp. 251–255.

Butterfass, J., M. Grebenstein, H. Liu, and G. Hirzinger. “DLR-Hand II: next generation of a dextrous robot hand”. In: *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*. Vol. 1. 2001, 109–114 vol.1.

Calli, B., A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar. "Benchmarking in Manipulation Research: Using the Yale-CMU-Berkeley Object and Model Set". In: *IEEE Robotics Automation Magazine* 22.3 (2015), pp. 36–52.

Calli, B., A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar. "The YCB object and Model set: Towards common benchmarks for manipulation research". In: *2015 International Conference on Advanced Robotics (ICAR)*. 2015, pp. 510–517.

Cardín-Catalán, Daniel, Marta Mora, Javier Andrés, Bruno Sospedra, and Carlos Rubert. "Diseño y simulación de un prototipo de mano protésica de bajo coste". In: *XXI Congreso Nacional de Ingeniería Mecánica* (2016).

Castiello, U., K. M. B. Bennett, and C. Mucignat. "The reach to grasp movement of blind subjects". In: *Experimental Brain Research* 96.1 (1993), pp. 152–162.

Chinellato, E., A. Morales, R.B. Fisher, and A.P. del Pobil. "Visual quality measures for Characterizing Planar robot grasps". In: *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 35.1 (2005), pp. 30 –41.

Ciocarlie, M., A. Miller, and P. Allen. *Grasp analysis using deformable fingers*. 2005, p. 3585.

Ciocarlie, Matei, Claire Lackner, and Peter Allen. "Soft Finger Model with Adaptive Contact Geometry for Grasping and Manipulation Tasks". In: *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (Mar. 2007), pp. 219–224.

Cutkosky, M. R. and R. D. Howe. "Human Grasp Choice and Robotic Grasp Analysis". In: ed. by S. T. Venkataraman and T. Iberall. Dexterous Robot Hands. New-York: Springer-Verlag, 1990. Chap. 1, pp. 5–31.

Cutkosky, M.R. "On grasp choice, grasp models, and the design of hands for manufacturing tasks". In: *IEEE Transactions on Robotics and Automation* 5.3 (June 1989), pp. 269–279.

Dang, H., J. Weisz, and P. K. Allen. "Blind grasping: Stable robotic grasping using tactile feedback and hand kinematics". In: *2011 IEEE International Conference on Robotics and Automation*. 2011, pp. 5917–5922.

Dechev, N, W.L Cleghorn, and S Naumann. "Multiple finger, passive adaptive grasp prosthetic hand". In: *Mechanism and Machine Theory* 36.10 (2001), pp. 1157 –1173.

Detry, R., D. Kraft, O. Kroemer, L. Bodenhagen, J. Peters, N. Krüger, and J. Piater. "Learning Grasp Affordance Densities". In: *Paladyn. Journal of Behavioral Robotics* 2.1 (2011), pp. 1–17.

Diankov, Rosen and James Kuffner. *OpenRAVE: A Planning Architecture for Autonomous Robotics*. Tech. rep. CMU-RI-TR-08-34. Pittsburgh, PA: Robotics Institute, 2008.

Ding, Dan, Yun-Hui Lee, and Shuguo Wang. "Computation of 3-D form-closure grasps". In: *IEEE Transactions on Robotics and Automation* 17.4 (Aug. 2001), pp. 515 –522.

Dollar, Aaron M. and Robert D. Howe. "The Highly Adaptive SDM Hand: Design and Performance Evaluation". In: *The International Journal of Robotics Research* 29.5 (2010), pp. 585–597.

Feix, T., J. Romero, H. B. Schmiedmayer, A. M. Dollar, and D. Kragic. "The GRASP Taxonomy of Human Grasp Types". In: *IEEE Transactions on Human-Machine Systems* 46.1 (2016), pp. 66–77.

Feix, Thomas, Javier Romero, Carl Henrik Ek, Heinz Bodo Schmiedmayer, and Danica Kragic. "A metric for comparing the anthropomorphic motion capability of artificial hands". In: *IEEE Transactions on Robotics* 29.1 (2013), pp. 82–93.

Feix, Thomas, Ian M. Bullock, and Aaron M. Dollar. "Analysis of Human Grasping Behavior: Object Characteristics and Grasp Type". In: *IEEE Transactions on Haptics* 7.3 (2014), 311–323.

Felip, Javier, Angel J Durán, Marco Antonelli, Antonio Morales, and Angel P Del Pobil. “Tombatossals: A humanoid torso for autonomous sensor-based tasks”. In: *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*. IEEE. 2015, pp. 475–481.

Ferrari, C. and J. Canny. “Planning optimal grasps”. In: *Proceedings 1992 IEEE International Conference on Robotics and Automation* (1992), pp. 2290–2295.

Fieller, E. C., H. O. Hartley, and E. S. Pearson. “TESTS FOR RANK CORRELATION COEFFICIENTS. I”. In: *Biometrika* 44.3-4 (1957), pp. 470–481.

Foresee, F. Dan and M. T. Hagan. “Gauss-Newton approximation to Bayesian learning”. In: *Neural Networks,1997., International Conference on*. Vol. 3. 1997, 1930–1935 vol.3.

Goins, Alex K., Ryan Carpenter, Weng-Keen Wong, and Ravi Balasubramanian. “Evaluating the Efficacy of Grasp Metrics for Utilization in a Gaussian Process-Based Grasp Predictor”. In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014) September 14-18, 2014, Chicago, IL, USA*. IEEE, 2014.

Graetz, Georg and Guy Michaels. “Robots at work”. In: *CEPR Discussion Paper No. DP10477* (2015).

Han, Li, Jeffrey C. Trinkle, and Z. X. Li. “Grasp analysis as linear matrix inequality problems.” In: *IEEE Transactions on Robotics* 16.6 (2000), pp. 663–674.

Hang, Kaiyu, Florian T. Pokorny, and Danica Kragic. “Friction Coefficients and Grasp Synthesis”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013.

Harada, K., M. Kaneko, and T. Tsuji. “Rolling based manipulation for multiple objects”. In: *Proceedings. ICRA ’00. IEEE International Conference on Robotics and Automation*. Vol. 4. 2000, pp. 3887–3894.

Herzog, Alexander, Peter Pastor, Mrinal Kalakrishnan, Ludovic Righetti, Jeannette Bohg, Tamim Asfour, and Stefan Schaal. “Learning of Grasp Selection based on Shape-Templates”. In: *Autonomous Robots* 36.1-2 (Jan. 2014), pp. 51–65.

Hester, R.D., M. Cetin, C. Kapoor, and D. Tesar. “A criteria-based approach to grasp synthesis”. In: *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*. Vol. 2. 1999, 1255–1260 vol.2.

Hirai, Shinichi and Haruhiko Asada. “Kinematics and Statics of Manipulation Using the theory of Polyhedral Convex Cones”. In: *International Journal of Robotics Research* 12.5 (1993), pp. 434–447.

Howe, R., I. Kao, and M. Cutkosky. “Sliding of robot fingers under combined tornsion and shear loading”. In: 1988, pp. 103–105.

Howe, RD and MR Cutkosky. “Practical force-motion models for sliding manipulation”. In: *International Journal of Robotics Research* 15.6 (1996), pp. 557–572.

Issac, Jan, Manuel Wüthrich, Cristina Garcia Cifuentes, Jeannette Bohg, Sebastian Trimpe, and Stefan Schaal. “Depth-based Object Tracking Using a Robust Gaussian Filter”. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) 2016*. IEEE, May 2016.

A Survey on Activities of Daily Living and Occupations of Upper Extremity Amputees. Vol. 35. 6. Korean Academy of Rehabilitation Medicine (KAMJE), 2011, p. 907.

Kao, I. and M. R. Cutkosky. “Quasi-Static Manipulation with Compliance and Sliding”. In: *International Journal of Robotics Research* 11.1 (1992), pp. 20–40.

Kao, I. and FQ Yang. “Stiffness and contact mechanics for soft fingers in grasping and manipulation”. In: *IEEE Transactions on Robotics and Automation* 20.1 (2004), pp. 132–135.

Kappassov, Zhanat, Yerbolat Khassanov, Artur Saudabayev, Almas Shintemirov, and Huseyin Atakan Varol. “Semi-anthropomorphic 3D printed multigrasp hand for industrial and service robots". In: *2013 IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2013* c (2013), pp. 1697–1702.

Kappler, Daniel, Jeannette Bohg, and Stefan Schaal. "Leveraging Big Data for Grasp Planning". In: *2015 IEEE International Conference on Robotics and Automation (ICRA) Washington State Convention Center Seattle, Washington, May 26-30, 2015*. IEEE, 2015.

A comparison of the grip force distribution in natural hands and in prosthetic hands. Vol. 26. 12. Informa UK Limited, 2004, 705–711.

Kasper, Alexander, Zhixing Xue, and Rüdiger Dillmann. "The KIT object models database: An object model database for object recognition, localization and manipulation in service robotics". In: *The International Journal of Robotics Research* 31.8 (2012), pp. 927 –934.

Kemp, C. C., A. Edsinger, and E. Torres-Jara. "Challenges for robot manipulation in human environments [Grand Challenges of Robotics]". In: *IEEE Robotics Automation Magazine* 14.1 (2007), pp. 20–29.

Kim, Byoung-Ho, Sang-Rok Oh, Byung-Ju Yi, and Il Hong Suh. "Optimal grasping based on non-dimensionalized performance indices". In: *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*. Vol. 2. 2001, pp. 949 –956.

Kim, J., K. Iwamoto, J. J. Kuffner, Y. Ota, and N. S. Pollard. "Physically Based Grasp Quality Evaluation Under Pose Uncertainty". In: *IEEE Transactions on Robotics* 29.6 (2013), pp. 1424–1439.

Kim, Jin-Oh and Pradeep Khosla. "Dexterity measures for design and control of manipulators". In: *Proceedings IROS Workshop on Intelligent Robots and Systems* (1991), pp. 758–763.

Kirkpatrick, D. G., B. Mishra, and C. K. Yap. "Quantitative Steinitz's theorems with applications to multifingered grasping". In: *Proceedings of the twenty-second annual ACM symposium on Theory of computing*. STOC '90. Baltimore, Maryland, United States: ACM, 1990, pp. 341–351.

Laguillaumie, P., M. A. Laribi, P. Seguin, P. Vulliez, A. Decatoire, and S. Zeghloul. "From Human Motion Capture to Industrial Robot Imitation". In: *Robotics and Mechatronics: Proceedings of the 4th IFToMM International Symposium on Robotics and Mechatronics*. Ed. by Saïd Zeghloul, Med Amine Laribi, and Jean-Pierre Gazeau. Cham: Springer International Publishing, 2016, pp. 301–312.

Lenz, Ian, Honglak Lee, and Ashutosh Saxena. "Deep Learning for Detecting Robotic Grasps". In: *IJRR* (2015).

León, Beatriz, Joaquín Sancho-Bru, Néstor Jarque-Bou, Antonio Morales, and Máximo Roa. "Evaluation of Human Prehension Using Grasp Quality Measures". In: *International Journal of Advanced Robotic Systems* (2012).

León, Beatriz, Antonio Morales, and J.L. Sancho-Bru. "A Framework for Robot and Human Grasping Simulation". In: *2013 IEEE International Conference on Robotics and Automation (ICRA). Workshop on Developments of Simulation Tools for Robotics and Biomechanics*. 2013.

Leon, Beatriz, Carlos Rubert, Joaquin Sancho-Bru, and Antonio Morales. "Evaluation of prosthetic hands prehension using grasp quality measures". In: *IEEE International Conference on Intelligent Robots and Systems* (2013), pp. 3501–3506.

Leon, Beatriz, Antonio Morales, and Joaquin Sancho-Bru. *From Robot to Human Grasping Simulation*. Vol. 19. Cognitive Systems Monographs. Springer International Publishing, 2013, p. 261.

Leon, Beatriz, Carlos Rubert, Antonio Morales, and Joaquin Sancho-Bru. "Characterization of Grasp Quality Measures for Evaluating Robotic Hands Prehension". In: *IEEE International Conference on Robotics and Automation* (2014), pp. 3688–3693.

Levine, Sergey, Peter Pastor, Alex Krizhevsky, and Deirdre Quillen. “Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection”. In: 2016.

Li, Z. and S.S. Sastry. “Task-oriented optimal grasping by multifingered robot hands”. In: *IEEE Journal of Robotics and Automation,* 4.1 (Feb. 1987), pp. 32 –44.

Liegeois, A. “Automatic Supervisory Control of the Configuration and Behavior of Multibody Mechanisms”. In: *IEEE Trans. Systems, Man, and Cybernetics* 7.12 (1977), pp. 842–868.

Lin, Q., J. M. Burdick, and E Rimon. “A stiffness-based quality measure for compliant grasps and fixture”. In: *IEEE Transactions on Robotics and Automation* 16.6 (2000), pp. 675–688.

Lippiello, Vincenzo, Fabio Ruggiero, Bruno Siciliano, and Luigi Villani. “Visual grasp planning for unknown objects using a multifingered robotic hand”. In: *IEEE/ASME Transactions on Mechatronics* 18.3 (2013), pp. 1050–1059.

Liu, Yun-Hui. “Qualitative test and force optimization of 3-D frictional form-closure grasps using linear programming”. In: *Robotics and Automation, IEEE Transactions on* 15.1 (Feb. 1999), pp. 163 –173.

Llop, Immaculada, Antonio Pérez, and Carlos Rubert. “Anthropomorphism index of opposition for hand prostheses.” In: *23rd Congress of the European Society of Biomechanics* (2017).

Llop-Harillo, Imma, Carlos Rubert, and Antonio Pérez-González. “Computation of grasp quality metrics in OpenHand simulator to improve a 3-D printed prosthetic hand”. In: *Computer Methods in Biomechanics and Biomedical Engineering Congress* (2018).

Llop-Harillo, Immaculada, Antonio Pérez-González, Verónica Gracia-Ibañez, and Carlos Rubert. “Design and evaluation of anthropomorphic hands”. In: *Summer School on soft manipulation* (2017).

Ma, Raymond R., Student Member, IEEE, Lael U. Odhner, Member, IEEE, and Aaron M. Dollar. “A Modular, Open-Source 3D Printed Underactuated Hand”. In: *2013 IEEE International Conference on Robotics and Automation (ICRA) Karlsruhe, Germany, May 6-10, 2013*. IEEE, 2013.

Dex-Net 1.0: A Cloud-Based Network of 3D Objects for Robust Grasp Planning Using a Multi-Armed Bandit Model with Correlated Rewards. IEEE, 2016.

Mason, Matthew T. *Mechanics of robotic manipulation*. Cambridge, MA, USA: The MIT Press, 2001, p. 265.

Matheus, Kayla and Aaron M. Dollar. “Benchmarking grasping and manipulation: Properties of the objects of daily living”. In: *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings* (2010), pp. 5020–5027.

MATLAB. *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.

Miller, Andrew T and Peter K Allen. “Examples of 3D grasp quality computations”. In: *Proceedingsof the IEEE International Conference on Robotics and Automation*. Vol. 2. IEEE, 1999, pp. 1240–1246.

Mirtich, B. and J. Canny. “Easily computable optimum grasps in 2-D and 3-D”. In: *Proceedings IEEE International Conference on Robotics and Automation*. May 1994, pp. 739–747.

Mishra, B. “Grasp Metrics: Optimality and Complexity”. In: *Proceedings of the Workshop on Algorithmic Foundations of Robotics*. WAFR. San Francisco, California, USA: A. K. Peters, Ltd., 1995, pp. 137–165.

Mishra, B., J. T. Schwartz, and M. Sharir. “On the existence and synthesis of multifinger positive grips”. In: *Algorithmica* 2.1-4 (Nov. 1987), pp. 541–558.

Murray, Richard N., Zexiang Li, and Shankar Sastry. *A mathematical introduction to robotics manipulation.* CRC Press, 1994, pp. I–XIX, 1–456.

Nakamura, Yoshihiko, Kiyoshi Nagai, and Tsuneo Yoshikawa. "Dynamics and Stability in Coordination of Multiple Robotic Mechanisms". In: *The International Journal of Robotics Research* 8.2 (1989), pp. 44–61.

Nguyen, V.-D. *Constructing force-closure grasps*. Institute of Electrical and Electronics Engineers, 1988, pp. 1368–1373.

Ore, Fredrik, Lars Hanson, Magnus Wiktorsson, and Yvonne Eriksson. "Automation constraints in human industrial robot collaborative workstation design". In: *Swedish Production Symposium 2016 SPS 2016, 25 Oct 2016, Lund, Sweden*. 2016.

Oztop, Erhan, Nina S. Bradley, and Michael A. Arbib. "Infant grasp learning: a computational model". In: *Experimental Brain Research* 158.4 (2004), pp. 480–503.

Park, Young C. and Gregory P. Starr. "Grasp Synthesis of Polygonal Objects Using a Three-Fingered Robot Hand". In: *The International Journal of Robotics Research* 11.3 (1992), pp. 163–184.

Pedersen, Mikkel Rath, Lazaros Nalpantidis, Rasmus Skovgaard Andersen, Casper Schou, Simon Bøgh, Volker Krüger, and Ole Madsen. "Robot skills for manufacturing: From concept to industrial deployment". In: *Robotics and Computer-Integrated Manufacturing* 37 (2016), pp. 282 –291.

Pedregosa, F. et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

Pinto, Lerrel, Abhinav Gupta, The Robotics Institute, and Carnegie Mellon University. "Supersizing Self-supervision: Learning to Grasp from 50K Tries and 700 Robot Hours". In: *2016 IEEE International Conference on Robotics and Automation (ICRA) Stockholm, Sweden, May 16-21, 2016*. IEEE, 2016.

Ponce, Jean, Steve Sullivan, Attawith Sudsang, Jean-Daniel Boissonnat, and Jean-Pierre Merlet. "On Computing Four-Finger Equilibrium and Force-Closure Grasps of Polyhedral Objects". In: *The International Journal of Robotics Research* 16.1 (1997), pp. 11–35.

Popovic, M., G. Kootstra, J. A. Jorgensen, D. Kragic, and N. Krüger. "Grasping unknown objects using an Early Cognitive Vision system for general scene understanding". In: 2011, pp. 987–994.

Prats, Mario, Ester Martínez, Pedro J. Sanz, and Angel P. del Pobil. "The UJI librarian robot". In: *Intelligent Service Robotics* 1.4 (2008), p. 321.

Prattichizzo, Domenico and Jeffrey Trinkle. "Grasping". In: *Springer Handbook of Robotics* (Sept. 2008). Ed. by Bruno Siciliano and Oussama Khatib, pp. 671–700.

Redmon, Joseph and Anelia Angelova. "Real-Time Grasp Detection Using Convolutional Neural Networks". In: *2015 IEEE International Conference on Robotics and Automation (ICRA) Washington State Convention Center Seattle, Washington, May 26-30, 2015*. IEEE, 2015.

Roa, Máximo. "Grasp Planning Methodology for 3D Arbitrary Shaped Objects". Ph. D. thesis. Universidad Politécnica de Cataluña, 2009.

Roa, Máximo A. and Raúl Suárez. "Grasp quality measures: review and performance". English. In: *Autonomous Robots* (2014), pp. 1–24.

Rombokas, Eric, Peter Brook, Joshua R Smith, and Yoky Matsuoka. "Biologically inspired grasp planning using only orthogonal approach angles". In: *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on*. 2012, pp. 1656–1661.

Rubert, Carlos and Antonio Morales. "Comparison between grasp quality metrics and the anthropomorphism index for the evaluation of artificial hands". In: 2016, pp. 1352–1357.

Rubert, Carlos, Beatriz Leon, and Antonio Morales. "Grasp Quality Metrics for Robot Hands Benchmarking". In: *Humanoid Robots, 2014 IEEE/RSJ International Conference on* (Nov. 2014), pp. 761–766.

Rubert, Carlos, Daniel Kappler, Antonio Morales, Stefan Schaal, and Jeannette Bohg. "On the relevance of grasp metrics for predicting grasp success". In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2017).

Rubert, Carlos, Beatriz León, Antonio Morales, and Joaquín Sancho-Bru. "Characterisation of Grasp Quality Metrics". In: *Journal of Intelligent & Robotic Systems* 89.3 (2018), pp. 319–342.

Rüßmann, Michael, Markus Lorenz, Philipp Gerbert, Manuela Waldner, Jan Justus, Pascal Engel, and Michael Harnisch. "Industry 4.0: The future of productivity and growth in manufacturing industries". In: *Boston Consulting Group* (2015), p. 14.

Sahbani, A., S. El-Khoury, and P. Bidaud. "An overview of 3D object grasp synthesis algorithms". In: *Robotics and Autonomous Systems* 60.3 (2012), pp. 326 –336.

Salisbury, J K and J J Craig. "Articulated Hands: Force Control and Kinematic Issues". In: *The International Journal of Robotics Research* 1.1 (1982), pp. 4–17.

Salisbury, J. K. and B. Roth. "Kinematic and Force Analysis of Articulated Mechanical Hands". In: *Journal of Mechanisms Transmissions and Automation in Design* 105.1 (1983), pp. 35–41.

Shimoga, Karun B. "Robot Grasp Synthesis Algorithms: A Survey." In: *International Journal of Robotic Research* 15.3 (1996), pp. 230–266.

Slade, Patrick, Aadeel Akhtar, Mary Nguyen, and Timothy Bretl. "Tact: Design and performance of an open-source, affordable, myoelectric prosthetic hand". In: *2015 IEEE International Conference on Robotics and Automation (ICRA)* (May 2015), pp. 6451–6456.

Suárez, Raúl, Máximo Roa, and Jordi Cornella. *Grasp quality measures*. Tech. rep. Technical University of Catalonia, 2006.

Supuk, Tamara, Timotej Kodek, and Tadej Bajd. "Estimation of hand preshaping during human grasping." In: *Medical engineering & physics* 27.9 (Nov. 2005), pp. 790–7.

Tai, Kevin, Abdul-Rahman El-Sayed, Mohammadali Shahriari, Mohammad Biglarbegian, and Shohel Mahmud. "State of the Art Robotic Grippers and Applications". In: *Robotics* 5.2 (2016), p. 11.

Trinkle, J.C. "On the stability and instantaneous velocity of grasped frictionless objects". In: *Robotics and Automation, IEEE Transactions on* 8.5 (Oct. 1992), pp. 560 –572.

Weisz, Jonathan and Peter K Allen. "Pose Error Robust Grasping from Contact Wrench Space Metrics". In: *IEEE Int. Conf. on Robotics and Automation (ICRA)*. 2012, pp. 557–562.

Wilson, Frank R. *The hand: How its use shapes the brain, language, and human culture*. Vintage, 2010.

Wohlkinger, W., A. Aldoma, R. B. Rusu, and M. Vincze. "3DNet: Large-scale object class recognition from CAD models". In: *2012 IEEE International Conference on Robotics and Automation*. 2012, pp. 5384–5391.

Wüthrich, M., P. Pastor, M. Kalakrishnan, J. Bohg, and S. Schaal. "Probabilistic Object Tracking Using a Range Camera". In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Nov. 2013, pp. 3195–3202.

Xiong, Caihua, Youfu Li, Han Ding, and You-Lun Xiong. "On the Dynamic Stability of Grasping." In: *I. J. Robotic Res.* 18.9 (1999), pp. 951–958.

Xiong, Y.-L. "Theory of point contact restraint and qualitative analysis of robot grasping". In: *Science in China (Scientia Sinica) Series A* 37.5 (1994), pp. 629–640.

Yoshida, Morio, Suguru Arimoto, and Ji-Hun Bae. "Blind Grasp and Manipulation of a Rigid Object by a Pair of Robot Fingers with Soft Tips". In: (2007).

Zanchettin, Andrea Maria, Nicola Maria Ceriani, Paolo Rocco, Hao Ding, and Björn Matthias. "Safety in human-robot collaborative manufacturing environments: Metrics and control". In: *IEEE Transactions on Automation Science and Engineering* 13.2 (2016), pp. 882–893.

Zheng, Yu and Chee-Meng Chew. "Distance Between a Point and a Convex Cone in -Dimensional Space: Computation and Applications". In: *Robotics, IEEE Transactions on* 25.6 (2009), pp. 1397 –1412.

Zheng, Yu and Wen-Han Qian. "Coping with the Grasping Uncertainties in Force-closure Analysis". In: vol. 24. 4. SAGE Publications, The International Journal of Robotics Research, 2005, 311–327.

Zhu, X. and J. Wang. "Synthesis of force-closure grasps on 3-D objects based on the Q distance". In: *IEEE Transactions on Robotics and Automation* 19.4 (2003), p. 669.

Zhu, Xiangyang, Han Ding, and M.Y. Wang. "A numerical test for the closure properties of 3-D grasps". In: *Robotics and Automation, IEEE Transactions on* 20.3 (2004), pp. 543 –549.