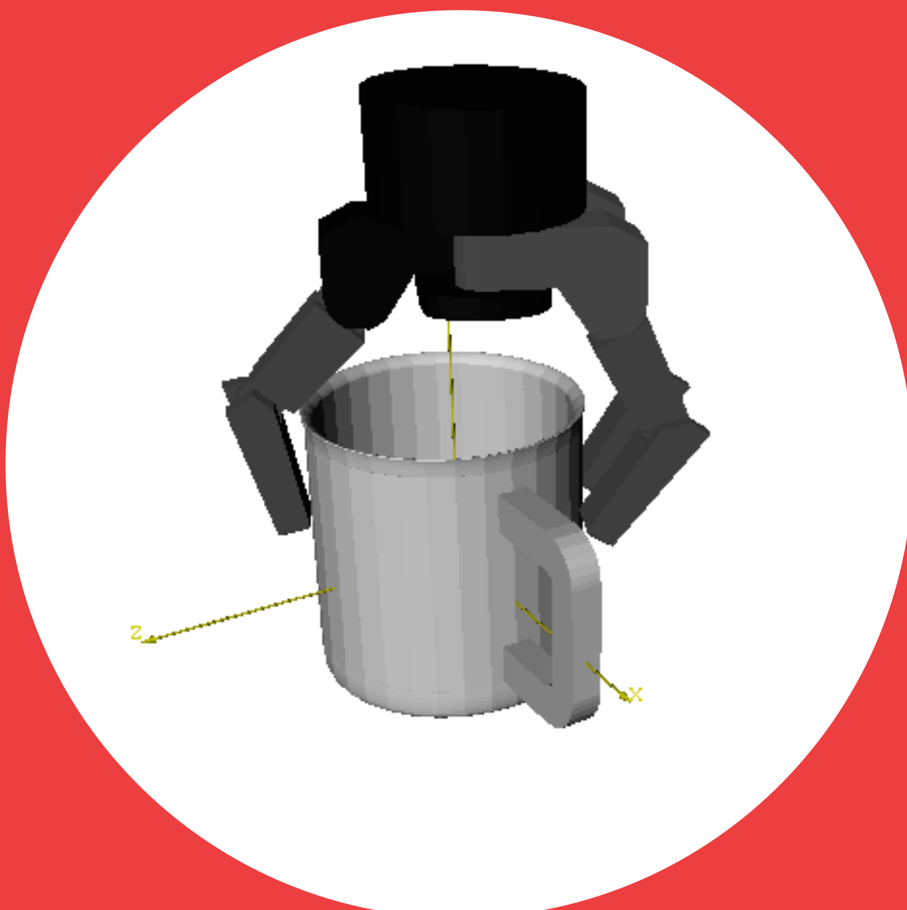
Grasp planning under uncertainty

Ekaterina Kolycheva (née Nikandrova)



Aalto University

DOCTORAL
DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 9/2016

Grasp planning under uncertainty

Ekaterina Kolycheva (née Nikandrova)

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Electrical Engineering, at a public examination held at the lecture hall AS1 of the school on 22 January 2016 at 12.

Aalto University
School of Electrical Engineering
Department of Electrical Engineering and Automation
Intelligent Robotics group

Supervising professor
Prof. Ville Kyrki

Thesis advisor
Prof. Ville Kyrki

Preliminary examiners
Dr. Renaud Detry, Montefiore Institute University of Liege, Belgium
Prof. Matei Ciocarlie, Columbia University, USA

Opponents
Prof. Carl Henrik Ek, University of Bristol, UK

Aalto University publication series
DOCTORAL DISSERTATIONS 9/2016

ISBN 978-952-60-6619-6 (printed)
ISBN 978-952-60-6620-2 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-6620-2

Unigrafia Oy
Helsinki 2016

Finland

441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Ekaterina Kolycheva (née Nikandrova)

Name of the doctoral dissertation
Grasp planning under uncertainty

Publisher School of Electrical Engineering

Unit Department of Electrical Engineering and Automation

Series Aalto University publication series DOCTORAL DISSERTATIONS 9/2016

Field of research Automation Technology

Manuscript submitted 10 August 2015 **Date of the defence** 22 January 2016

Permission to publish granted (date) 16 December 2015 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

Advanced robots such as mobile manipulators offer nowadays great opportunities for realistic manipulators. Physical interaction with the environment is an essential capability for service robots when acting in unstructured environments such as homes. Thus, manipulation and grasping under uncertainty has become a critical research area within robotics research.

This thesis explores techniques for a robot to plan grasps in presence of uncertainty in knowledge about objects such as their pose and shape. First, the question how much information about the graspable object the robot can perceive from a single tactile exploration attempt is considered. Next, a tactile-based probabilistic approach for grasping which aims to maximize the probability of a successful grasp is presented. The approach is further extended to include information gathering actions based on maximal entropy reduction. The combined framework unifies ideas behind planning for maximally stable grasps, the possibilities of sensor-based grasping and exploration.

Another line of research is focused on grasping familiar object belonging to a specific category. Moreover, the task is also included in the planning process as in many applications the resulting grasp should be not only stable but task compatible. The vision-based framework takes the idea of maximizing grasp stability in the novel context to cover shape uncertainty. Finally, the RGB-D vision-based probabilistic approach is extended to include tactile sensor feedback in the control loop to incrementally improve estimates about object shape and pose and then generate more stable task compatible grasps.

The results of the studies demonstrate the benefits of applying probabilistic models and using different sensor measurements in grasp planning and prove that this is a promising direction of study and research. Development of such approaches, first of all, contributes to the rapidly developing area of household applications and service robotics.

Keywords Grasp planning, probabilistic models, MCMC, GPR, entropy, PSO, optimization

ISBN (printed) 978-952-60-6619-6 **ISBN (pdf)** 978-952-60-6620-2

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2016

Pages 167 **urn** http://urn.fi/URN:ISBN:978-952-60-6620-2

Preface

The work for this dissertation has been started at the Machine Vision and Pattern Recognition Laboratory of the Faculty of Technology Management in Lappeenranta University of Technology in 2011. Already in my Master's studies I have become very interested in robotics and especially in manipulation planning under uncertainty. In my Master's thesis I developed a method, which allows updating a simulation object model based on actual measurements to achieve a success of a planned task. In 2013 I moved to the new Intelligent Robotics research group created by my supervisor Prof. Ville Kyrki in the Department of Electrical Engineering and Automation at the Aalto University. Last year of my doctoral studies I were working remotely. The department and Aalto ELEC Doctoral School were responsible for most of the funding.

First of all, I am very thankful to my instructor and supervisor Prof. Ville Kyrki for his valuable advices and financial support. He introduced me the world of intelligent robotics. His continuous support and guidance has taught me a lot about research and science. He encouraged me to go further in my research which finally allowed me to write this dissertation even working remotely. I want to also thank the pre-examiners of my dissertation, Dr. Renaud Detry and Prof. Matei Ciocarlie, for their valuable comments.

This work would not have been possible without all my colleagues and friends both in the MVPR laboratory and in the Intelligent Robotics group. I am particularly indebted to Jonna Laaksonen, Nataliya Strokina, Andrey Maglyas, Joni Pajarinen, Polychronis Kondaxakis, Alberto Montebelli, Rajkumar Muthusamy and many others for numerous inspiring discussions about robotics, machine learning and life in general.

My special thanks go to my mother for her love and support, and especially to my grandmother who encouraged the best in me.

Last but not least, I express my deepest gratitude to my dearest husband Aleksandr for his love, patience and support during the preparation of this work.

Saint Petersburg, Russian Federation, December 16, 2015,

Ekaterina Kolycheva (née Nikandrova)

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I E. Nikandrova and V. Kyrki. What do contacts tell about an object?. In *Proceedings of the 2012 4th IEEE RAS & EMBS, International Conference on Biomedical Robotics and Biomechatronics*, pages 1895-1900, Roma, Italy, June 2012.

II J. Laaksonen, E. Nikandrova and V. Kyrki. Probabilistic Sensor-based Grasping. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2019-2026, Vilamoura, Algarve, Portugal, October 2012.

III E. Nikandrova and V. Kyrki. Explorative sensor-based grasp planning. In *Towards Autonomous Robotic Systems (TAROS)*, pages 197-208, Bristol, UK, August 2012.

IV E. Nikandrova, J. Laaksonen and V. Kyrki. Towards informative sensor-based grasp planning. *Robotics and Autonomous Systems*, Volume 62, Issue 3, pages 340-354, March 2014.

V E. Nikandrova and V. Kyrki. Category-based task specific grasping. *Robotics and Autonomous Systems*, Volume 70,pages 25-35, August 2015.

VI E. Kolycheva and V. Kyrki. Task-specific Grasping of Similar Objects by Probabilistic Fusion of Vision and Tactile Measurements. Accepted for publication in *IEEE-RAS International Conference on Humanoid Robots (Humanoids 2015)*, November 2015.

Author's Contribution

Publication I: "What do contacts tell about an object?"

The background review about learning from haptic exploration was done by the author. Together with the second author she developed the idea of the simulation-based approach in order to answer the question how much information about the environment can be obtained from a single tactile exploration attempt. The technique implementation based on minimizing the difference between predicted and real sensor measurements and conduction of experiments were also done by the author. The author studied several optimization algorithms and made their performance comparison. She had the main responsibility of the manuscript writing.

Publication II: "Probabilistic Sensor-based Grasping"

The author participated in improving the performance of stability maximization approach for grasping proposed and implemented by the first author by modifying the way of representation the grasp stability probability using Particle Swarm Optimization approach. She also contributed to writing of the publication.

Publication III: "Explorative sensor-based grasp planning"

The author proposed the idea of explorative entropy-based probabilistic framework for grasp planning. She extended the initial stability maximizing approach implemented by the second author. She proposed and implemented the entropy-based model and carried out the experiments in simulation. The article was mostly written by the author. V. Kyrki contributed actively to revising of the publication, provided discussion and comments on the publication.

Publication IV: "Towards informative sensor-based grasp planning"

The background research, the design, detailed analysis and implementation of methods at the base of the exploration stage was done by the author. She also ran the simulation experiments. J. Laaksonen developed the stability maximizing part of the framework and ran experiments on a real platform. The writing of the article was a collaborative effort of the first and the third authors. J. Laaksonen participated in the manuscript revision and provided comments.

Publication V: "Category-based task specific grasping"

The author together with V. Kyrki designed the probabilistic framework for sensor-based task-specific grasping of objects with shape variations inside the category. She co-developed the probabilistic models and implemented them. She was responsible for carrying out the experiments both in simulation and on a real robot. The article was written by the author, with cooperation with the co-author.

Publication VI: "Task-specific Grasping of Similar Objects by Probabilistic Fusion of Vision and Tactile Measurements"

The author extended the vision-based framework from the Publication V by including tactile feedback in optimization in order to improve the stability quality of the grasps, implemented the general methods, run experiments in simulation and on a real robot and performed the comparison between the new and earlier approach. She was mainly responsible for article writing. V. Kyrki actively provided discussion and comments on the publication and contributed to writing of the publication

List of Abbreviations

CGDB	Columbia grasp database
DOF	Degree of freedom
FITC	Fully independent training conditional
GP	Gaussian process
GPR	Gaussian process regression
GRAB	Guaranteed recursive adapting bounding
MCMC	Markov chain Monte Carlo
PF	Particle filter
POMDP	Partially observable Markov decision process
PSO	Particle swarm optimization
QM	Stability quality metric
RBPF	Rao–Blackwellized particle filter
RGB-D	Red, green, blue plus depth
SLAM	Simultaneous localization and mapping

List of Symbols

A	Action attributes
$E[.]$	Expected value
G	Grasp configuration
i	Index variable
k	Index variable
$\hat{m}$	Fixed known measurement value
M	Sensor measurements
O	Object attributes
$\vec{p}$	Object pose
$P(X)$	Probability or probability density of X
$P(X\|Y)$	Conditional probability of X given Y
$P(X, Y)$	Joint probability of X and Y
q	Stability metric
S	Success metric (grasp stability)
t	Time counter
T	Task constraint compatibility
w	Shape goodness-of-fit weight
x	Scalar coordinate
z	Scalar coordinate
α	Angle of rotation

δ	Fitting error
ϵ	Epsilon grasp quality measure
θ	Model grasp location
$\phi(\delta)$	Fitting error
$\psi(q)$	Stability weight
$\wedge$	AND logical operator

1. Introduction

Robotics is a rapidly developing field with a large number of open research problems. Manipulation and especially grasping under uncertainty is one of the currently critical research questions for future service robots, which need to act in unstructured and uncertain environments such as homes. To cope with the uncertainty, the environment needs to be perceived using sensors while acting in it. Tactile sensors are particularly valuable sources of information in manipulation. Moreover, vision sensors are widely used in robotics to estimate object pose or shape. Nevertheless, modern sensors are far from ideal and their measurements are noisy and uncertain. Thus, development of grasp planning algorithms contributes to traditional and service industries, health sector and especially to the rapidly developing field of household applications.

The main emphasis of the research presented in this dissertation is on developing methods that cope with imperfect knowledge and uncertain senses in robotics using probabilistic mathematical models. Such grasp planning approaches will be able to reduce costs, enhance the safety, efficiency and productivity.

1.1 Background

Grasp planning is a fundamental problem in the field of robotics that has been attracting an increasing number of researchers during recent decades. Grasping has been traditionally studied in the physical context of attaining a form or force closure. The quality of a grasp is usually measured using quality metrics [1]. The principles of kinematics and dynamics are used to determine the contact locations on the object and the hand configurations. The computation of stable grasp using screw theory [2, 3], potential function method [4] was extensively studied already in late seventies - eighties. Even in these early studies [5] it was concluded that the choice of a *good* grasp does not depend only on object attributes but mostly on the task that needs to be performed with this object. Thus, a *good* grasp should be task-oriented.

Nowadays, the study of force closure determination is developed by taking task specificity in consideration [6]. However, physics-based methods require perfect knowledge of object models and robot poses in order to guarantee a resulting stable grasp.

Another group of methods in robotic grasping are data-driven approaches. They do not assume perfect knowledge of robot and object parameters. In contrast, in order to find relatively good grasp for some object they require a comprehensive database of objects and appropriate grasp configurations.

The development of data-driven approaches started with availability of grasp simulation environments like Graspit! [7] and OpenRave [8]. They allow to generate thousands of grasp candidates, fully control the environment and its attributes without needing to build expensive hardware. For instance, Columbia Grasp Database [9] was created using GraspIt!. It takes a grasp planning algorithm based on the data in the database and computes best grasps for a set of hundreds of object models. Recently, Kappler et al. in [10] proposed a new large-scale simulation-based database containing hundreds of thousands of grasps annotated with different stability metrics generated for a large set of objects from numerous categories. The dataset was constructed using OpenRave simulation environment. However, the collection of big amount of data, needed for good generalization, is a time and resource-consuming process.

Contrary to physical (analytical) methods, data-driven approaches depend more on graspable object representation and perceptual data processing, like object recognition and classification based on similarity metrics or pose estimation. Based on a-priori knowledge about an object, data-driven approaches can be divided into three categories: grasping of known objects, familiar objects and unknown objects [11].

Grasping of known objects usually assumes the existence of an experience database consisting of different object models and associated grasp configurations. A query object belongs to the database. The goal is to recognize an object, define its pose and retrieve a suitable grasp.

Grasping of familiar objects considers the problem of finding grasps for objects similar to previously encountered ones. The similarity can be defined differently, either simply by shape or texture, or on higher level, for example, by object category. The main problem of such approaches is to measure the similarity between objects and transfer grasps accordingly.

For unknown objects there are no grasp models or grasp experience available. These approaches extract various features from sensory data. The grasps are generated and ranked based on these features.

An important aspect for grasp generation is a source of information from which the robot acquires knowledge about its environment. The control of the robot system given a complete world model is a well-defined and solvable task. However, such models are usually unavailable. Perception through sensors allows to compensate the lack of prior information. There exist different types of sensors, which are used in robotics for the extraction of meaningful world features. Vision-based and tactile sensors can be highlighted among others as the most significant.

Vision is a powerful sense, which provides an enormous amount of information in order to intelligently interact with an environment. In robotics, vision is used for locating objects [12], estimating the shape of the object [13], object recognition [14] and classification [15].

Vision is a rich source of information about the environment and tasks. However, visual sensors are not perfect, they incorporate some noise. The sense of touch is the only one without which humans are not able to hold or safely manipulate objects [16]. Tactile sensors allow to compensate the lack of information about objects in unstructured environments. Tactile sense is particularly important in robotic manipulation as it provides data to estimate object properties like stiffness, geometry, and contact characteristics.

Tactile sensors are widely used in robotics. Initially, they were mainly applied for object recognition [17, 18]. Over the past years, they have been extensively used for solving problems like object classification [19], pose estimation [20, 21] and dexterous grasping [22].

Tactile feedback can be also used to improve the initial estimate of the grasp obtained, for example from vision, and to perform necessary robot hand adjustments [23]. Both tactile and visual information was used for estimating stability of the grasps planned onto novel objects using part-based grasp planner, composed of grasp prototypes learned from experience [24].

1.2 Objectives and scope

This thesis is focused on grasp planning for known or similar objects under uncertainty in location or shape. In industrial robotics the problem of object manipulation and grasping is solved. Robots are tuned to the particular tasks and there is no uncertainty in the environment. However, in many real-world scenarios, like household environments, knowledge about world state cannot be exact. The errors in object's and robot's attributes can be crucial while performing a manipulation task.

The objective of this thesis is to show how to deal with different types of uncertainty using tactile and visual sensor data in order to find stable grasp configurations, which take into account given manipulation tasks constraints.

1.3 Contributions

The core contributions of this dissertation are worked out in the six author's publications. This thesis starts by looking at how much information about an object a robot can optimally learn from a single tactile exploration attempt and goes further by developing a probabilistic approach for grasp planning under object pose uncertainty unifying the ideas of stability maximization, information gathering by minimizing the entropy and using sensor's feedback. Next step is to look at uncertainty in object shape and consider the term "category" in the scope of grasping as well as taking into account task applicability. Finally, the idea of on-line sensory information for grasp planning is applied for the case of shape uncertainty. The major scientific contributions in this thesis are summarised in the following list:

- A simulation-based approach that allows to study how much information a robot can optimally learn from a single tactile exploration attempt. The approach is based on minimizing the difference between predicted and measured sensor readings. To avoid bias in the results due to the concrete implementation the algorithms used include both directional methods (Steepest descent) and metaheuristics (Simulated Annealing, Particle Swarm optimization and Firefly algorithm). An important finding is that the inference is surprisingly difficult from a single explorative action. The level of difficulty varies a lot for different object attributes.

- A probabilistic approach combining information collection (exploration based on entropy prediction) and stability maximization under uncertainty. For the entropy calculations a computationally efficient discrete entropy estimate that uses only particle weights was developed. The experimental results demonstrated that the approach allows to accomplish statistically optimal grasp planning, while simultaneously reducing uncertainty about an object's pose. It was shown that each successive grasp reduces the probability of an unstable grasp by refining the belief of the object pose. Moreover, a combined approach with entropy-based exploration stage outperforms the results over only maximally stable action selection.

- A probabilistic approach for task-specific stable grasping of objects with shape variations inside the category. The idea of maximizing grasp stability is taken in the novel context to cover shape uncertainty. It was shown that by combining information over multiple grasps and multiple objects, the proposed approach results in more stable grasps compared to the classical approach of using the most similar model's grasp. In addition, the technique can cope with a sparse training set in contrast to most data-driven approaches. Moreover, the method does not require large amount of data. It requires only an incomplete point cloud obtained from a single RGB-D image.

- An extended probabilistic framework, which combines the ideas of planning for the maximally stable grasp, using vision as a source for initial guess and online sensory-based grasping. It was shown that combination of vision and tactile sensors performed better than solely vision-based technique. Moreover, it was demonstrated that the iterative nature of the method allowed to finally succeed and find a stable grasp even if first rounds were unsuccessful by collecting more information and using previous results as starting values for the next round of optimization.

1.4 Structure of the thesis

The remainder of this thesis consists of two parts: an introductory part of four more chapters and publications. Chapter 2 presents an overview of related work about robot learning from touch in uncertain environments as well as author's contributions in determining what can be learned from haptic exploration. Chapter 3 starts with the related background in the area of grasping known objects under pose uncertainty. It continues by author's contributions to planning for the maximally stable and the most informative grasps. The existing approaches of category-based grasping are presented at the beginning of Chapter 4. The second part of the chapter is focused on the author's probabilistic approaches for task-specific stable grasping based on vision information and extended by incorporating tactile feedback are studied. Finally, the thesis is concluded and summarized in Chapter 5.

2. Learning from touch in uncertain environments

Robotic manipulation, especially grasping, is a research area highly affected by different types of uncertainties. Nowadays robots are increasingly being used to act in unstructured environments. The sense of touch is the one, which is the most important to safely manipulate objects [25]. People with impaired tactile sensibility have significant difficulties with in-hand manipulation because a brain lacks the contact information needed to control manipulation activities [26].

Similar to humans, tactile sensing in robotics can help to organize the interaction with objects. Despite the great number of recent works about using tactile sensors in solving robotic manipulation problems, the learning aspect is not thoroughly investigated. This chapter covers the research question: How much and what information can be learned about the robot environment from a haptic exploration attempt?

2.1 Related work

Early works about using tactile sensors in robotics were concentrated on industrial applications and especially workpiece localization problem [27, 28]. Hillis in [18] proposed a robot manipulator incorporating a tactile sensor, which was able to recognise various fastening devices, like nuts and bolts. One of the first overview papers about the perspectives of robotic tactile sensing appeared already in 1984 [29]. After the creation of first advanced multifingered robot hands in 1980's, the tactile sensors began to be used for control of dexterous manipulation [30].

More recent works about tactile perception are focused on object classification, recognition and localization problems. A range of works addresses the task of recognition various object properties, such as ridges and bumps on surface of an object [31], object materials [32] or even internal states of an object such as presence of liquid or open/closed state [33].

Traditionally, contact information has been used to recover 3D object models by constructing a convex hull of collected contact points [34], creating volumetric models [35] or using superquadrics [36]. Alternative approaches directly use the haptic sensor data to classify an object without building a 3D model of the object [19]. Several recent works adapt the "bag-of-features" approach from vision-based object recognition for object identification using low-resolution intensity images from tactile sensors [37, 38].

Another important application area for tactile sensors is grasp stability estimation. For example, Bierbaum et al. in [39] presented a method to generate grasp affordances based on reconstructed from extracted features faces of an object through tactile exploration. Dang et al. in [22] proposed a machine learning approach called *blind grasping* to predict stable grasps on an unknown object based on tactile feedback and hand kinematic data. Bekiroglu et al. [40] presented a probabilistic learning framework to assess grasp stability while grasping an object using information from tactile sensors.

Object localization using tactile measurements can be performed using two different approaches. First is based on cost function minimization in order to find a solution, which best fits the measurements [41, 42]. Alternatively, probabilistic formulations can be used to represent current belief about object location, which allows to incorporate different types of uncertainty [43]. In [21], the authors presented a decision-theoretic approach which minimizes the uncertainty in the relative pose between the robot and objects using arm trajectories to enable task specific grasps on objects. Tactile sensors were used to detect contacts between the hand and objects. Petrovskaya et al. in [20, 44] proposed a Bayesian approach for 6DOF global object localization via touch. This is a Monte-Carlo approach, which performs a series of refinements using annealing. Recently, Chebotar et al. in [45] address the problem of in-hand object localization and object manipulation with tactile feedback.

Despite the great number of recent researches which use tactile information for solving various robotics tasks the question to what extent properties of the robot environment can be inferred from the tactile sense was not previously considered. Thus, an approach presented in Publication I first among other related studies tries to answer the question to what degree a robot can use tactile sense to learn about its environment.

2.2 Learning about objects from haptic exploration

A novel simulation-based approach that allows to study how much information a robot can optimally learn from a single tactile exploration attempt was presented in Publication I. The term “information” in this context denotes different attributes of a graspable object. The distinguishing feature of the approach is that the simulator [8] is used both as an internal model of the environment for the robot and as substitution for the real world. Thus, the robot is able to try out actions in simulation before executing them for real. More than that, simulation provides an ideal environment without measurement errors. This made it possible to focus on studying how much can be inferred using the tactile sensors in the ideal case, rather than how good sensors are used in experiments.

The method proposed in Publication I is based on optimization, as the robot environment does not contain measurement uncertainty. The goal is to minimize a difference between predicted and measured sensor readings. To ensure that the robot will succeed in its task the simulation model is updated based on error minimization. The general scenario of robot learning through simulation is shown in Figure 2.1.

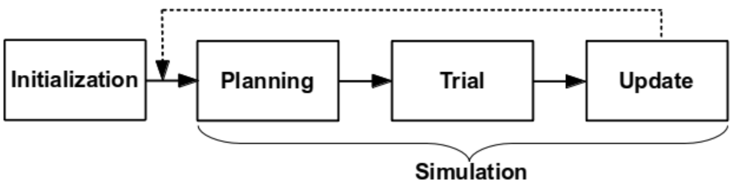


Figure 2.1. Learning process: *Initialization*: robot obtains an initial guess about object attributes (e.g. location); *Planning*: using the predicted values robot plans to complete the task; *Trial*: robot tries the trajectory and obtains teal sensor measurements; *Update*: robot searches for the new state of its world model by minimizing the difference between planned and real values. Adapted from Publication I.

Traditionally, simulation is used in robotics for planning or trying some actions. However, in the proposed approach simulation is applied to update the world model and change the action plan before its execution on a real platform. This new idea transforms simulation into the internal mental view of the robot.

The approach presented in the Publication I is based on optimization. Thus, the choice of the fitness function affects dramatically the method convergence properties and, thereby, the whole approach performance. The form of the fitness function was obtained after several trials. Final version includes information about contacts between robot and object and information about finger joint angles. The function typically has only one global minimum and few flat areas. A more detailed description of error function can be found in Publication I.

Due to multimodality and non-linearity of the fitness function the error minimization was not a trivial task. Moreover, such factors as accuracy and computational complexity had to be accounted in the choice of the optimization method. Thus, to avoid a bias in the results because of the particular algorithm, several approaches including directional methods and metaheuristics were also studied. As a directional method, Steepest Descent approach, widely used in optimization, was implemented. To cope with problems of several local optima and deceptivity in some regions of the fitness function metaheuristics approaches were studied. First, Particle Swarm Optimization approach (PSO), which is an efficient non-derivative simple to implement global search algorithm with low number of parameters [46]. However, its weak points are slow convergence and weak local search ability. Second, Simulated Annealing, which is a robust general technique which can deal with highly nonlinear models, noisy data and many constraints [47]. On the other hand, it requires to tune a lot of parameters to convert into an actual algorithm. The last is Firefly algorithm, which as PSO considers particles moving in the search space according to specific dependencies [48]. In contrast to PSO, Firefly algorithm deals more efficiently with multimodal functions.

For the experiments an object transportation task was chosen. Geometric attributes namely an object pose and size were set as unknown parameters. Contact sensors which detect the presence or absence of contacts between the robot hand and a target object together with information about joint angles of the fingers after closing the hand were used as sources of information. All experiments were done in OpenRave simulator [8] using Barrett WAM arm with the Barrett hand models. For generalization purpose the set of testing objects included objects with simple geometrical shape as well as mostly symmetrical entities and more complex examples.

The experimental results of performing optimization after single grasping attempt for 3DOF case (uncertainty in location and orientation) showed that tactile readings received during a single manipulation action carry quite limited amount of information about an object. Even PSO, which performed consistently well for all testing objects, resulted in significant angular error values. Thus, the possibility to learn about the object orientation from a single grasp is highly object and hand dependent. For symmetrical objects it is impossible to find a correct orientation. There is a complex interplay between Cartesian and angular position variables affecting the fitness function value. Given exact Cartesian position, angular position can be predicted quite precisely. However, object location is usually uncertain. The experimental results for 4DOF case (3DOF + size) showed that size information can be estimated very precisely and fitness error values are small.

2.3 Discussion

Publication I introduced a simulation-based approach to find the limitations of tactile sensing in optimal conditions from single exploration attempt. The studies conducted in the paper relate mostly to pose estimation, although the scale of the object is considered in addition. The technique is based on minimizing the fitness function, which models the difference between predicted and real sensor measurements, including contact and joint angles components.

The main conclusion of the study is that learning about the environment from a single tactile exploration attempt is surprisingly difficult, even in optimal conditions without any sensor noise. Moreover, the difficulty of estimating different attributes varies significantly: experimental results showed that object location and scale factor can often be estimated relatively well, but the accuracy of orientation estimation is very object and robot hand dependent. More than that, the estimation problem can be ambiguous, for example, due to symmetries in object shape.

In order to increase the accuracy in predicting all parameters an initial planning for object exploration to maximize information gain could be done. This step can be especially benefiting for objects with asymmetries on a small area. However, such attempts seldom result in stable grasps and further exploration is needed. Planning of exploration in uncertain conditions is a non-trivial task, which may require construction of probabilistic models. In contrast, in Publication I the exploration is a by-product of a grasp attempt.

The work in Publication I is restricted to ideal sensing modeled in simulation without contact uncertainty. However, the information from real sensors could be applied in the approach. Moreover, more complex error measurements could be used instead of single Euclidean distance. One more possible extension direction is to include multiple tactile exploration attempts. The study in this direction is performed by Vazquez et al. in [49]. Similarly to Publication I, they address the question of how good is a tactile sensor for tactile exploration. Apart from aforementioned extensions, the use of a more dexterous robot model could improve results due to the possibility to collect more contact information from a single manipulation attempt.

3. Grasping known objects under pose uncertainty

A grasp planning process for known objects consists of an offline planning phase and on-line operation. In the offline stage, a grasp experience database for every object model is generated. The major part of grasp candidate generation approaches use force-closure grasps and rank them according to the ϵ quality metric. The on-line phase includes object recognition and usually at least pose estimation. After that, the associated grasp hypotheses are retrieved from the database. In most cases when speaking about grasp planning of known objects it is assumed that object models are precise. However, uncertainty in object attributes, such as pose, can be crucial while performing manipulation tasks. More than that, even though geometrical models are good approximations of real objects, they are usually inaccurate especially for everyday household items.

3.1 Related work

If an object model is given together with already generated set of grasps candidates, the only problem is to determine an object pose and then find the best possible candidate grasp. When a 3D object model is known the challenge is to automatically generate a set of good grasp candidates. The quality of the grasp is usually based on physical properties and defined by quality measures such as the widely used ϵ-metric [50].

To simplify the process, many methods use different types of object decompositions to reduce total number of feasible grasps without trying them on objects. Some authors approximate objects with a collection of primitive shapes, likes boxes, cylinders or spheres. For example, Huebner et al. in [13] propose hierarchical minimum volume bounding box decomposition approach to approximate object shapes and, thereby, reduce the search space for good grasps. Instead of boxes superquadrics can be used for object shape approximation like in [51]. Przybylski et al. in [52] present a grasp planning approach, which operates on the grid of medial spheres object representation. To rank the grasps they calculate epsilon measure for force-closure. All above mentioned approaches have been implemented and tested in simulation.

One of the biggest criticism against using ϵ-metric and force-closure to rank grasp candidates is that a relatively fragile grasp can be classified as a force closure stable grasp [53]. To overcome this problem extra noise can be added to grasp attributes and only grasps for which some percentage of neighbouring grasps are also force-closure can be chosen. Such an approach is adapted by Weisz and Allen in [54]. They focus particularly on the ability of the ϵ-metric to predict grasp stability under object pose error.

Another way to generate grasp hypothesis is learning from human demonstration, which produces an experience database offline. De Granville et al. in [55] present a technique that learns mapping from object to grasp from human demonstration for reach-to-grasp actions. Detry et al. in [56] use human demonstrations to collect initial data, from which object specific grasp empirical density is built. The density is then used to sample grasp hypotheses. Romero et al. in [57] present a human-to-robot mapping system, where demonstrated human hand posture, including both grasp type and hand orientation, is classified and mapped to a specific robot hand. The method is first evaluated in simulation using 3D object models. The approach is also demonstrated on the humanoid platform [58]. A motion capture system is used to capture human grasp activities. Human upper body tracking, object tracking and hand pose estimation techniques are applied to analyze human grasps, which are then reproduced on the robot. Ekvall and Kragic in [59] present a method for generating grasps based on shape primitives and human demonstration. The system observes a human teacher, wearing a data-glove with magnetic trackers and recognizes the grasp type. The grasp type is then mapped on the set of robot hands. Finally, an approach vector is selected from an offline generated experience database. The authors conduct experiments with a simulated pose error in the simulation environment only.

Instead of grasp experience information a database of standard graspable objects can be constructed. Columbia grasp database [9] includes grasps for hundreds of objects and several robot hands. In this context grasp planning problem consists of recognizing an object, estimating its pose and retrieving a suitable grasp from the database. The authors also extend the work to deal with 3D partial range sensor data collected from a small number of viewpoints [60].

Another way for finding grasps is object affordance modeling. In cognitive robotics, the concept of affordances [61] describes the relationship between the agent and its environment through the effect of agent's interaction with the environment. A grasp affordance corresponds to the different ways a robot hand can be placed near an object to generate stable grasp after closing the hand. In the grasping related works, grasp affordances consider either the overall stability of the grasp [62, 63], or, for instance, accounting for a spesific task [64].

Learning is a powerful instrument in robotic systems as it allows to cope with uncertainties to some extent. In [63] learning based on experience is used on a real robot to acquire grasp affordances of an object. The learning process reduces a vision bootstrapped distribution of grasps to a smaller set containing only good grasps. Over the past years, reinforcement learning has been applied to various robotic manipulation problems including grasp planning. For example, in [65] reinforcement learning is applied to search for policies that optimize the chance of grasp success.

Multi-modal uncertainties are frequent in manipulation, especially based on tactile perception. While allowing to cope with model uncertainties reinforcement learning techniques do not usually consider uncertain beliefs, or use simple (e.g. Gaussian) uncertainty models. Tactile measurements are local in nature, and one tactile exploration attempt is not usually enough for precise object pose estimation.

One essential problem is how to represent such multi-modal errors. One way is to use particle filters (PFs). PF is a MCMC method, which represents probability distribution using a cloud of particles. Initially, particle filters were applied in robotic manipulation mainly for pose estimation [66, 67, 20]. Later, Platt et al. in [68] explore the idea of using Bayesian filtering to localize features embedded in flexible materials during robot manipulation. Zhang and Trinkle in [69] consider Grasp-SLAM problem as a filtering problem. They propose an approach that apply particle filter to improve knowledge of the system's physical parameters while simultaneously tracking the object during visual occlusion.

An alternative approach to solve grasping problem is probabilistic formulation. Already in nineties Goldberg et al. in [70] apply Bayesian framework to a grasping problem in the presence of an object's pose uncertainty. The method is evaluated on a parallel-jaw gripper grasping a 2D object. No sensor information is utilized by the approach. Later, Hsiao et al. in [71] propose a method for planning under uncertainty for robotic manipulation by partitioning the configuration space into a set of regions that can be treated as states in a partially observable Markov decision process (POMDP). The authors demonstrate the approach on simple planar problem of unknown object pose in a long time horizon. However, in contrast to the current study their work was mostly conceptual and addressed only small problems.

In [21] the authors present a decision-theoretic approach for task-driven manipulation of objects when there is uncertainty in relative pose between an object and a robot. Tactile sensors are used to detect contacts between the robot and the object. The approach operates with discretized observation and belief states. The resulting grasp is found as a maximum a posteriori belief. Petrovskaya et al. in [72] elaborate a new inference method, called Guaranteed Recursive Adapting Bounding (GRAB) for pose estimation problems. The method is tested in both simulation and on a real platform and it is further extended in [44]. Nevertheless, the approach is applied only for object localization. Moreover, in both [21] and [44] geometrical models of objects are given and stability information is not accounted during grasp planning. On the contrary, the presented work uses particle representation for the object attributes models and finds the most stable grasp by maximizing the expected stability.

Veiga and Bernardino in [73] propose an approach that uses Bayesian Optimization methods to search for the best grasp configuration by iteratively optimizing a suitable grasp criterion. They evaluate the methodology only in simulation on known objects.

A Bayesian framework for grasp planning under object pose or shape uncertainty as well as robot motion error is described in [74]. The framework combines the results from multiple object detectors and multiple grasp planning approaches and tries to find a consensus among them in order to result in a grasp robust to errors in both perception and execution. The authors claim that the use of a Bayesian formulation is essentially valuable, so that maximizing the expected success is clearly superior to executing the best action based on the maximum likelihood solution of the object attributes. The similar idea was expressed also by [54]. The approach is based on the grasp stability maximization over a uniform uncertainty in the object pose. This thesis is based on this line of work and extends it by considering the time series nature of sequential actions, looking at maximization of accumulated reward and updating the belief in the case of an unsuccessful attempt.

In some cases when, for example the predicted stability of a grasp is relatively small, it would be better to perform an exploration step to reduce the uncertainty in object attributes. Dragiev et al. in [75] present a framework for iterative grasping that utilizes two motion primitives: an explorative and exploitative grasp. However, the approach is focused on grasping previously unknown objects.

One way to find the most informative grasp is to use the measure of information content such as entropy. One big challenge is to estimate the entropy of the posterior distribution of object attributes represented by a set of particles.

Most studies in robotics which utilize information gathering approaches are related to SLAM applications and using Rao–Blackwellized particle filter (RBPF) [76, 77]. Stachniss et al. in [76] introduce an integrated approach for exploration, localization and mapping. They use RBPF to represent the posterior about the map and poses. The entropy is divided into 2 components: the entropy of the robot trajectory posterior and the map uncertainty weighted by the likelihood of the corresponding trajectory.

A similar approach for touch-based registration is proposed by Taguchi et al. [78]. The method performs 6DOF registration in a RBPF framework. Next robot motion is selected as a motion that provides the maximum information gain. An information gain from a proposed robot motion is estimated by the expected entropy that the RBPF distribution. The authors compare three methods for a particle-based density entropy estimation. An approximation based on kernel density estimation and estimation using particle weights only were shown to be superior for general distributions. In Publications III-IV a computationally efficient weight-based entropy estimation technique was elaborated. Comparing to Taguchi's work the method uses a basic particle filter instead of RBPF and solves a completely different task - stable grasp planning.

3.2 Planning for the maximally stable grasp

A probabilistic framework for grasp planning under uncertainty using online sensor information and simultaneously updating knowledge of object pose is described in Publications II-IV. When there exists a significant uncertainty in knowledge about object parameters a first grasp trial can easily fail. However, the failure can be detected using tactile sensors. The proposed approach utilizes the sensor measurements in order to refine information about object attributes. The most stable grasp is then found by maximizing the expected stability. Thus, the presented stability-maximizing approach allows planning and executing statistically optimally stable grasps, as well as collecting measurements and performing corrective motions, which reduce the uncertainty about the environment. The necessary probabilistic models are built using Gaussian process regression (GPR). A Markov Chain Monte Carlo approach is applied to estimate a goal object's pose and grasp stability while performing grasp attempts. A Bayesian approach is used allowing the marginalization over current knowledge to obtain estimates. The most stable grasp is, then, found by maximizing the posterior probability given models for object attributes and grasp stability.

The sequence of actions describing grasp stability maximization approach is shown in Figure 3.1. The process can be iterated until stability condi-

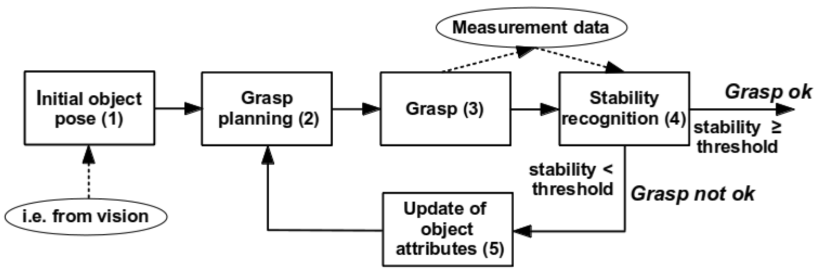


Figure 3.1. Stability maximizing framework model: *Step 1*: an initial uncertain estimate for oblect's pose is obtained (e.g. from vision); *Step 2*: planning for a grasp with uncertainty from the initial estimate is performed; *Step 3*: planned grasp is performed providing sensor measurements; *Step 4*: grasp stability is estimated using sensor data; *Step 5*: if the grasp is not stable, object attribute values are updated, new measurements are collected followed by re-planning for a new grasp. Adapted from Publication IV.

tions are satisfied.

The choice of probabilistic models for sensor-based manipulation and object knowledge refinement plays a key role in the performance of the approach. In a general form, a sensor-based manipulation model is depicted in Figure 3.2. From the Figure 3.2 it can be seen that the belief about object attributes O is updated over the time. Knowledge of O is, thus, refined using information from the performed action A and collected

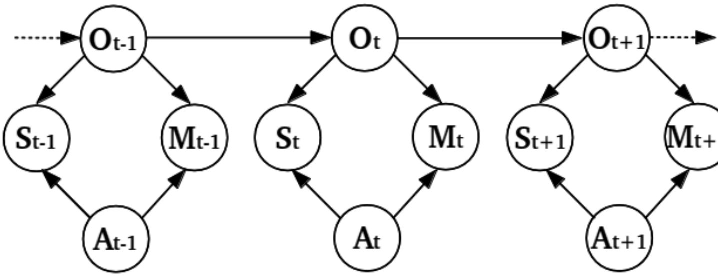


Figure 3.2. Probabilistic manipulation model (process of object knowledge refinement): object attributes (6DOF object's pose) O; success metric (grasp stability) S; action attributes (pose of the end-effector while grasping) A; online sensor measurements M. Adapted from Publication IV.

measurements M. The success S is determined by the action A (grasp) and object attributes O. More than that, measurement for the particular time stamp can be collected only after the action is taken. So, the prediction model of measurements should be defined.

To formulate planning for sensor-based grasping a probabilistic approach from [43] was applied. Two models are required to build a working system:

- $P(M|A,O)$ to describe relation between object attributes, grasp attributes and sensor measurements;

- $P(S|A,O)$ describing stability as a function of grasp action and object attributes.

Using these models the most stable grasp can be found as a maximum of the posterior distribution

$$\max_A \int P(S|A,O)P(O)dO. \tag{3.1}$$

To build the models GPR was applied. It allows, firstly, to build and update models quickly without using a simulator, which makes it possible to use them in a real environment. Moreover, GPR ensures generativity, so that models are able to account for the whole state space. More detailed description of using GPR for model building in the system is given in Publications II and IV. Readers interested in GPs can get a deeper understanding from [79]. Additionally, to reduce the computation time of the GP the Fully Independent Training Conditional (FITC) model [80] was utilized.

The theoretical framework was implemented using a particle based representation for object attributes, which is updated using Metropolis algorithm [81]. A benefit of particle based representation is that it is non-parametric and, thus, can represent any distribution. Each particle evolves independently and particles with low weights are reinitialized by resampling to preserve the representative power. The stability probability was maximized using Particle Swarm Optimization (PSO), an efficient global search algorithm. The reader can find more information about PSO in [82]. For the algorithm, which describes the proposed stability-maximizing grasp planning approach the reader is referred to Publications II and IV.

The framework was demonstrated both in simulation and on a real platform. GraspIt!-simulator [7] was chosen as a simulation environment for experiments. Barrett hand model was used in all experiments. Testing set consisted of two primitive-shaped objects, a cube and a cylinder, and 3 complex-shaped models, two different mugs and a pitcher. A table-top scenario and only top grasps were considered. Thus, the uncertainty was represented in three dimensions (x, z, α), where x, z are Cartesian coordinates and α is an orientation in degrees. Moreover, it was assumed that all objects are stationary during the process.

The main goal of the experiments was to demonstrate that using the proposed approach it is possible to refine the initially uncertain pose of the object during several grasp attempts while simultaneously improving grasp stability given estimates of object attributes and sparse measurement data collected during grasp attempts. Because of probabilistic nature of the approach the experimental results were represented using posterior distributions. Each run consisted of four grasps (the number of grasps was fixed instead of defining grasp stability threshold). For illustration, Figure 3.3 shows the initial and the final runs of the evolution of the second mug pose posterior. As can be seen the initial posterior distri-

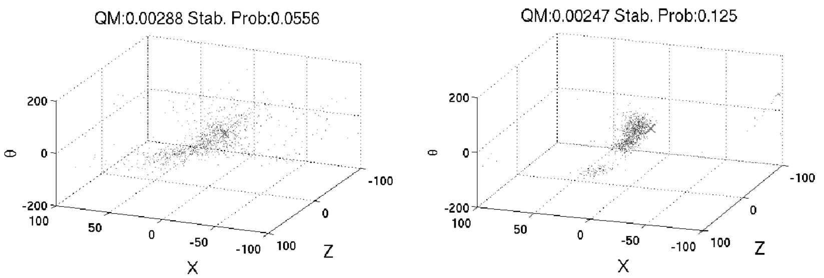


Figure 3.3. Grasping the mug at pose (-32,38,68): red cross denotes real object pose, QM is a stability quality metric computed in GraspIt! (ϵ-metric); Stab. Prob indicates predicted grasp stability probability. Adapted from Publication II.

bution is quite sparse, as it reflects the uncertainty in mug orientation. The final posterior is denser. Stability probability and ϵ-metric increase from run to run after each grasp attempt.

To show that the proposed probabilistic grasp planning framework actually improves the grasp quality, repeated experiments (100 runs) were performed for all 5 objects. The goal was to show that each successive grasp decreases the probability of unstable grasp and at the same time improves knowledge about the object pose. The results showed that the approach is able to refine object pose and, therefore, find stable grasp considerably more often after a few grasps. In general, the results can be compared to [21], where it was shown that 4 actions is required to achieve $80\% - 90\%$ probability of grasping an object in the similar tabletop scenario for a specific grasp. However, the authors of [21] use grid representation for object attributes instead of particle based one.

The method was also validated using a physical system consisting of a Melfa RV-3SB 6-DOF arm and a Robotiq 3-finger hand. The test object was a cardboard box. The stability S was a binary variable, defined experimentally by lifting the object. The positions from three finger actuators were recorded as the measurement data M. The goal was to show that the framework is able to find a stable grasp under uncertainty. After the first grasp, the posterior converged close to true position and the object aligned in the hand, but the orientation remained ambiguous (the stability probability only 26%). However, after the second grasp the probability increased up to 99% and the lifting of the object was successful. In the second experiment, the exploration capabilities of the approach were analyzed. The object was displaced such that the robot fails to grasp it during the first attempt. In this case posterior distribution consisted of two modes. After exploring the first mode wrong by chance the second grasp attempt also failed. However, the collected measurements supported the second mode and the third grasp was successful. This experiment demonstrated the benefit of the exploration stage in order to increase probability of achieving stable grasp. This lead to the idea of creating an explorative approach presented in Publications III-IV.

3.3 Exploration: planning for the most informative grasp

When uncertainty about an object pose is large, the predicted stability can be relatively small and such grasp can fail when executing. In these cases an exploration would be more beneficial than stability maximization. Publications III-IV present an extension of stability maximizing framework, which allows to balance between information gathering and grasp stability maximization. Whenever grasp stability probability is small, an exploration grasp is executed instead. This exploration grasp is chosen to minimize the expected entropy of the object attributes at the next time step given current knowledge of the object attributes. A particular challenge is to determine entropy for particle-based distributions. In Publications III-IV an efficient discrete estimate, which utilizes particle weights only, is proposed.

An extended by exploratory stage framework for probabilistic grasp planning can be described by Figure 3.4. As in the basic stability maximizing

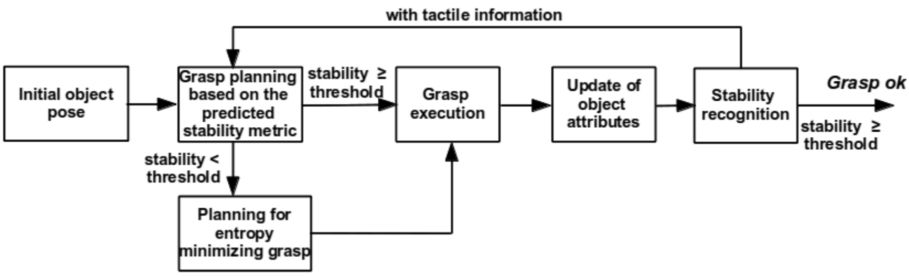


Figure 3.4. Exploratory framework model: An initial uncertain guess about the object's pose is known (e.g. from vision); Planning for maximally stable grasp using probabilistic stability model is performed; Grasp stability is predicted before execution; If the stability is less than a threshold re-planning for the most informative grasp is made (current grasp); If the stability is larger than a threshold maximally stable grasp is chosen as the current step; Chosen grasp is performed providing sensor measurements; Real grasp stability is estimated using sensor data; If the grasp is not stable re-planning for a new grasp with an updated belief is initiated. Adapted from Publication IV.

approach the process can be iterated until grasp stability criteria are satisfied.

The process of finding the most informative grasp is based on the entropy minimization for the posterior of object attributes $P(O_t|a_t, O_{t-1})$, given a grasp configuration a_t and previous object attributes estimate O_{t-1}. Objects are assumed to be stationary during the process, so that $O_t = O_{t-1}$. However, if required, the motion model $P(O_t|O_{t-1})$ can be included. To find the distribution the marginalization over unknown tactile measurements, which will be obtained after performing grasp using current object's pose estimate, can be done:

$$P(O_t|a_t, O_{t-1}) = \int_{M_t} P(O_t|a_t, O_{t-1}, M_t)P(M_t|a_t, O_{t-1})dM_t \quad (3.2)$$

The first factor describes the evolution of object attributes and can be derived using Bayes formula by applying stationarity assumption and marginalizing over uncertain object knowledge to obtain the normalization evidence term. As object distribution is represented by a particle set, an integral can be approximated by a weighted sum over particles. The second term describes the probability of measurements M_t after performing grasp a_t for current object belief O_t. The problem is to solve integral in Equation 3.2. A stochastic approach based on drawing random samples from $P(M_t|a_t, O_{t-1})$ is applied for this. This is possible, because GP is used for the measurements. So, one can draw particles from $P(M_t|a_t, t_{t,j})$, where the index j is selected proportional to the weights of the particle set O. For the detailed derivations and the algorithm, which illustrates this process the reader is referred to Publications III-IV.

One major challenge is to calculate the entropy for the distribution represented by the weighted set of particles. In Publication IV two techniques were introduced: weight-based and kernel-based approaches. The advantage of the weight-based method is its simplicity, only particle weights ignoring spatial locations are used for the entropy estimation. The set of particle weights is treated as a set of probability masses of a discrete probability distribution and the entropy for the discrete distribution is calculated. This approach does not produce absolute entropy values for different particle sets, because it does not consider the distance between particles. Nevertheless, the results can be compared for the same particle locations with different weights, which is enough to find the most informative action. Moreover, this approach is computationally efficient compared to more complex kernel-based estimation. Unlike the first technique, kernel-based entropy estimation is a non-parametric method, which utilizes kernel estimate composed of the collection of position samples and corresponding weights. Corresponding equations can be found in Publication IV.

As a basic stability-maximizing framework, the combined approach was evaluated in simulation using a similar setup. In all experiments a 2DOF case with uncertainty in x and z coordinates was considered. The first experiment followed the Figure 3.4. Examples of the initial and the final posterior distributions for the mug are shown in Figure 3.5. The posterior distribution for the object's pose converges faster when applying the combined approach, which unifies the stability evaluation and entropy minimization to obtain more information about an object.

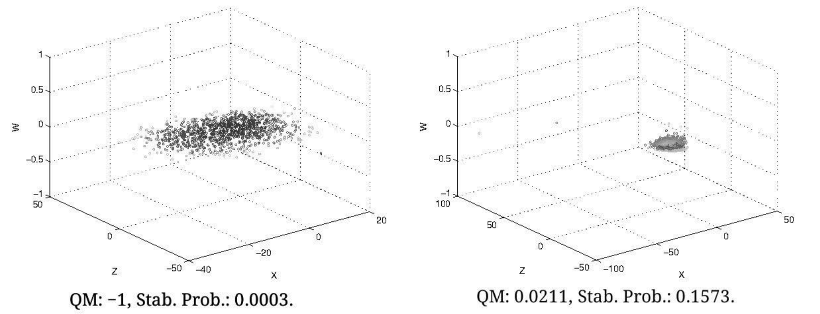
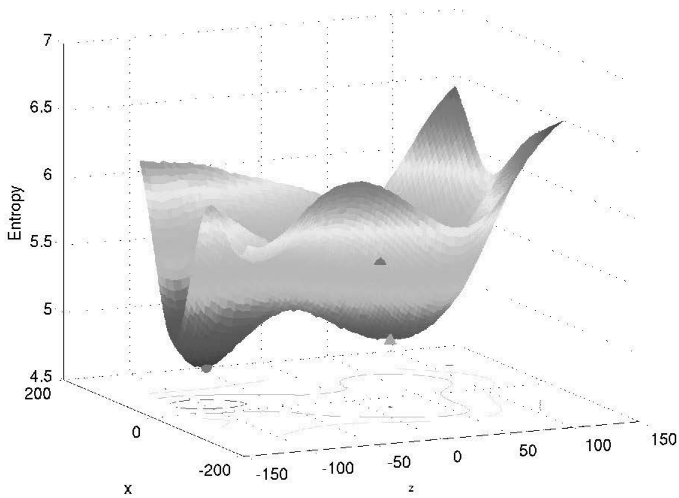


Figure 3.5. Particle distributions for grasping a mug: initial (left) and final (right). Adapted from Publication IV.

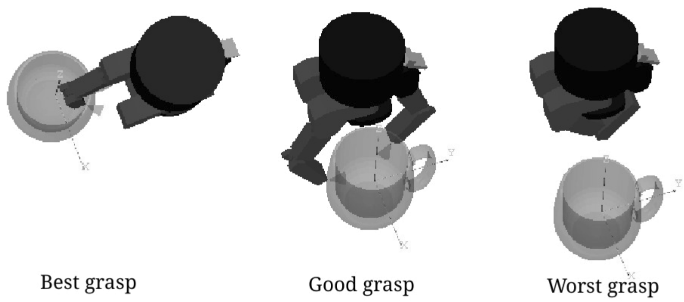
For both testing objects, a cube and a mug, the approach outperformed a method based purely on maximization of the stability metric. The quality metric increased from step to step, so that the uncertainty in the object's pose decreased and, as a result, both stable and informative grasp was found. To compare the approaches an experiment with the mug, where four maximally stable grasps were performed in a sequence each time updating the belief about object's pose according to the initial approach, was conducted. Resulting stability probability was slightly smaller than for the exploratory framework. However, ϵ-metric was not positive, which means that the found grasp was not force-closure.

Repeated experiments were performed to show that the grasp quality is actually improving also for the extended approach. The results after thirty tests on three test objects (cube, mug and cylinder) for randomly chosen object's poses reflected the general trend of increasing the ϵ-metric value after each exploratory step done for all testing objects.

To ensure that exploratory framework in practice allows effectively reduce the uncertainty, an additional experiment for two step optimization was conducted. Thus, the entropy was calculated for both initial particle set and the set updated after the grasp execution. Three different grasp configurations were chosen according to their predicted entropy values. These grasps as well as the predicted entropy landscape for different actions are depicted in Figure 3.6. After performing the grasps object attributes were updated and entropy values were recalculated. Although using weight-based entropy estimation the resulting absolute entropy values are not comparable over time, the prediction correctly preserved the order, that is, the case where the smallest entropy was predicted resulted in the smallest entropy and vice versa.



(a) The shape of the entropy function for 2DOF case: *red circle* - configuration with a smallest entropy(best grasp); *green triangle* - configuration with a small entropy(good grasp); *magenta triangle* - configuration with a large entropy(worst grasp).



(b) Grasps with different entropy values

Figure 3.6. Two step optimization experiment

To study that the stochastic approach is applicable for entropy estimation one more experiment was performed. The goal was to analyze behaviour of the entropy function dependent on the number of iterations in the estimation procedure. The entropy estimates were computed for the cube in case of 1DOF uncertainty in x coordinate for 1000 and 10000 iterations. The results showed that in both cases the entropy change behaviour is similar. Moreover, an error in best grasp location is not high for the case of smaller number of iterations. However, for the case of 1000 iterations the function is smooth in general, but contains some noise peaks because of randomness. Thus, to obtain smoother behaviour more iterations can be performed.

To justify the choice of a weight-based technique for entropy estimation a comparison with a kernel-based method was performed. To demonstrate that more computationally efficient weight-based estimation provides similar results to more complicated and time consuming kernel-based estimation the ”surface maps” of the entropy using both approaches were built. The hills and valleys in both cases are in the same locations, which confirms that the results are consistent. Thus, the simpler and computationally more efficient weight-based technique provides results similar to the kernel-based method and, therefore, it should be preferred.

3.4 Discussion

Publications II-IV propose a novel probabilistic framework for sensor-based grasp planning. The framework allows planning for stable grasps while simultaneously reducing uncertainty about the environment. All models used within the framework are purely data-driven, so perfect knowledge about object and manipulator attributes is not required. Two approaches utilizing general framework were presented. The first one is based on stability maximization. The second one extends the first one by employing an entropy-based exploration procedure, which allows the interplay between information gathering and maximizing grasp stability.

Both approaches are demonstrated using MCMC methods, particle representation for object attributes, and PSO for optimizing the grasps. A computationally efficient and simple weight-based technique is used for entropy estimation. Results of experiments conducted in simulation lead to the following conclusions. First of all, the proposed probabilistic grasp planning framework improves the grasp quality, as each successive grasp increases the grasp stability probability by reducing the uncertainty in the object's pose. Both approaches successfully prove the viability of the probabilistic grasp planning framework and provide reasonably good results for different test objects, which indicate that both of them can be applied for different grasp planning tasks. The resulting posterior distributions are adequately dense in both methods and stability probabilities are improved with each subsequent grasp attempt. The extended exploratory framework outperforms basic stability maximizing framework in both stability probabilities and quality metrics. Thus, an information gathering step reduces more efficiently the uncertainty about an environment when the predicted success probability is low. Additionally, real experiments confirm the viability of the framework and demonstrate the benefits of utilizing an exploration procedure.

Despite the promising results of the framework more experiments both in simulation for repeatability and especially on a real platform should be conducted. Only stability maximization approach is experimentally validated, which is enough to demonstrate the viability of the general probabilistic approach. However, the entropy minimizing approach could be also tested with a real robot and different test objects. The approaches are tested on a small set of mostly simple-shaped objects. Thus, the framework could be verified using more complex objects. The major simplification of the paper is the modeling of the grasping. Only top grasps are considered. So, the issue of higher-dimensional uncertainty is an important point to be addressed in the future work.

To improve the performance of the framework the probabilistic models can be modified. For example, grasp stability probability model $P(S|A,O)$ could be extended to include tactile measurements $P(S|A,O,M)$, which already was done in [40]. Adding sensor readings could benefit pose and stability estimation, as, for example, it would allow measuring contact surface types and shapes. Including a motion model into the framework would allow more accurate detection and tracking of the changes in the object pose during grasp attempts. One possible direction of extension is multi-step optimization, where the goal is to find a sequence of actions which maximize a success metric in a time horizon. Another direction is to modify the framework to deal with other types of uncertainties, like object shape and category as well as looking at the task-specific grasps. This direction is studied in Publications V-VI.

4. Category-based grasping

Object manipulation and especially grasping are important abilities for a robot acting in an unstructured environment. Many everyday objects share common features and characteristics, so that they can be grouped into categories. When a robot is required to grasp familiar objects, one can speak about category-based grasping. For manipulation and grasping tasks such natural criteria for categorization is functionality aspects, object's utility in performing a certain task [83]. Similar objects can be grasped in a similar way. Moreover, grasping can be constrained by a particular task, e.g. pouring or transporting. Thus, category-based grasping is closely related to task specification, because particular tasks cannot be performed with arbitrary objects, but only with objects belonging to a specific category. This chapter is focused on task-specific grasping of similar objects from a known category using visual information as well as utilizing tactile feedback.

4.1 Related work

Category-based grasping is most often performed by data-driven approaches. In case when a comprehensive database of object models and associated grasp configurations is available the task is easy to solve - there is always an object in the database which is a good fit for a goal object to be grasped. However, usually such databases are not available and their construction is a time-consuming and computationally intensive task. More than that, finding a similar object in the database is not trivial, because real measurements obtained from sensors are partial and noisy and fitting such data is difficult.

A broad review of data-driven approaches for grasp synthesis and methodoligies for grasp sampling and ranking is provided by Bohg et al. in [11].

Speaking about grasping familiar objects they distinguish the following categories of methods: discriminative approaches, which learn a discriminative function to separate bad and good grasp configurations, methods for grasp synthesis by comparison, in which grasp hypothesis for a specific object are found by searching for a similar object or its graspable part in the database containing associated good grasps, and approaches, which learn generative models of the whole grasp process.

Discriminative approaches mainly differ in object features which are used and also in a way how grasp candidates are parametrized. Some methods consider only graspable parts of objects, others learn multiple contact points or full grasp configurations.

Among approaches which are based on 3D data the study of El-Khoury and Sahbani [84] can be mentioned. They present a method which imitates human choice of the graspable component of the object, its handle. First, the object is decomposed into parts and each part is approximated by a superquadric. An artificial neural network is used then for classification. A grasp can be obtained by computing force-closure grasps on the handle. Pelossof et al. [85] use a single superquadric to approximate an object. They utilize Support Vector Machine to define what is a good grasp for a robot hand. Both of the aforementioned approaches were experimentally validated only in simulation with the assumption that precise 3D object models are known.

Boularias et al. in [86] propose a probabilistic approach for grasp learning based on Markov Random Fields. The goal of the method is to find the maximum a posteriori labeling of point clouds for new objects. Although the method also relies on 3D data for learning, it was tested not only in simulation, but on real 3D scans of different types of objects. However, it remains unclear how the approach would generalize to more object classes and sensor data.

There are number of techniques which mainly rely on more simple 2D data to distinguish bad and good grasp locations. For example, Saxena et al. in [87] propose a grasp learning algorithm that predicts points where to grasp an object as a function of images. Instead of using labeled training dataset as in [87], Montesano and Lopes in [88] present an algorithm that actively learns good grasping points by autonomously exploring different feature values on different objects. The approach combines beta-binomial distributions and non-parametric kernel technique to define full grasping probability distribution. The approach was tested using a real humanoid platform.

Even though algorithms that utilize 2D data are easier to implement, it is not always possible to infer a full grasp configuration using 2D alone as the problem is underconstrained. That is why the approaches that use both 2D and 3D visual data allow to learn functions that can take more parameters of a grasp into account. For instance, Saxena et al. [89] extend their earlier work [87] by incorporating 3D point cloud features in addition to 2D features, which enhances the prediction of grasp stability and allows to infer more grasp parameters like approach vector and finger spread. Rao et al. in [90] utilize segmentation on color and depth cues in order to achieve good classification rates. They employ a supervised learning method using both image and depth data to determine whether a given segment is graspable or not. Le et al. in [91] propose a method that learns the most stable fingertip placements. They apply Support Vector Machines to learn grasp hypotheses using relevant features extracted from both 2D and 3D data. Bohg and Kragic in [92] present a method that instead of local features apply the concept of shape context, which encodes global 2D object shape. For learning they use a supervised learning approach, in which the classifier is trained on labeled synthetic images.

A second group of approaches is based on finding similar objects or their parts in an experience database for which good grasp configurations are available. Thus, Curtis and Xiao in [93] build a comprehensive knowledge base for grasping consisting of 3D object types in simulation environment. The types are represented using Gaussian distributions over basic shapes. To infer a good grasp for a new object its low-level features are used to find the most similar type in the knowledge database.

Higher-level features are utilized by Hillebrand and Roa in [94]. They propose a method for transferring grasps between objects of the same category through warping the surface geometry along with the contact points of a grasp. The warped contacts are, then, locally replanned to ensure grasp stability. The approach is tested only on one mug category. Moreover, full 3D models for both source and target objects are required, which might not be always available. Failures on the most dissimilar mug show that the method generalizes poorly in case of large shape variability inside a category. Ben Amor et al. in [95] adapt the similar idea of contact warping onto new objects. They present an imitation learning approach for learning and grasp generalization based on human demonstrations. They demonstrated the approach on a real robot and objects from a mug category. However, their goal is not to find stable grasp configurations but to generate reach-and-grasp actions.

Recently, Stouraitis et al. [96] extend the work in [94, 95] to include functional grasps and tests on a wider set of objects. They modify the warping process to avoid mapping the complete geometry of the source and target objects. The approach exploits global and local shape similarities to wrap contact points. However, the method still requires full 3D object models and requires improvements in functionality prediction. Moreover, the authors validate the approach only in simulation environment.

Few studies are devoted to learning generative models for grasp synthesis that is based on identifying common structures from a number of examples. On example is given by Montesano et al. in [97], where they address the problem of learning affordances through robot-environment interaction. The general model utilises Bayesian networks to capture dependences between actions, objects and effects and to infer causality relationships. The method was validated in an imitation game, where a robot should repeat an effect demonstrated by a human with a a object. Thus, a robot should perform inference in the learned network to choose the action with the highest success probability. Song et al. in [64, 98] consider the problem of inferring full grasp configuration using object category together with task constraints as variables in the Bayesian network. The effectiveness of the approach was demonstrated using synthetic data and human hand model only.

In real-world scenarios a good grasp should not be only stable but it should be suitable for a particular task to be performed with an object. Thus, task constraints should be accounted during the process of grasp synthesis. Traditional approaches study task-specific grasping mainly in the force domain [3, 6]. Le et al. in [99] describe a data-driven approach to grasp synthesis. They begin with constructing a database of human grasps. To identify candidate grasp they introduce a shape-matching algorithm that utilizes shape features that contain contact normal information. Finally, they perform a task-based pruning using an anatomically-based grasp quality metric. The approach is dedicated for human-like hands only and was tested just in virtual environments.

Usually, both category and task should be accounted in order to find a useful grasp [64, 98]. Dang and Allen [100, 101] propose an example-based approach to generate semantic grasps, stable grasps that are functionally stable for a specific task. An affordance semantic map relates local object features to a set of predefined semantic grasps for different tasks. The resulting grasp is synthesized using the Eigengrasp planner [102]. The approach was demonstrated both in simulation and on a real platform. Nevertheless, the method requires full 3D models of objects.

In case when only partial sensor data is available a full object model can be estimated from multiple observations. Goldfeder and Allen in [103] use only synthetic data to construct a knowledge base and also utilize the Eigengrasp planner to generate grasps. Nevertheless, they utilize observations from real sensors to look up the most similar object and its pose in the database.

Some approaches combine 3D partial data and 2D images [104, 105]. They account for object category and task, but grasps are generated for objects that already exist in the database. Bohg et al. in [106] present an approach towards autonomous grasping of objects according to their category and a given task that also uses both 2D and 3D data. The grasp is predicted by a Bayesian network using only the most similar object model from the database.

The way how object and grasps are represented plays an important role in transferring grasps between objects. Recently, Pokorny et al. [107, 108] present a novel representation, the Grasp Moduli Space, in which objects are parametrized using smooth differentiable functions. This space can be utilized to continuously deform various surface/grasp configurations in order to generate grasps for a new object. However, the method is applicable only for objects without holes and full point clouds, as the smooth parametrization deteriorates when only partial data is available. Detry et al. in [109, 110] construct a low-dimensional space in which object parts with a similar shape are close to each other. The aim is to generalize grasps to novel objects by defining the object parts by which objects are often grasped. The overall shape similarities and object categories are not considered in the process. The method was tested using synthetic data as well as on a robot using real sensor measurements.

The approach proposed in Publications V belongs to the class of grasping familiar objects and more specific to the group of grasp synthesis by comparison. Unlike other grasp synthesis by comparison approaches discussed earlier the proposed framework does not require full 3D models for test objects. The method uses partial sensor data only and does not require the construction of a large database of models. What is distinguishes the proposed framework from all aforementioned methods is that the approach accounts for all training objects in the category during the optimization process, which allows to better generalize for new objects and handle larger shape variations.

The general methodology of maximizing grasp stability under uncertainty is presented in Publications II-IV. The vision-based approach for for task-specific grasping of novel objects from a known category presented in Publication V adopts the similar concept with different realization considering uncertainty in object's shape instead of pose. Publication VI extends the probabilistic framework from Publication V by incorporating tactile feedback in order to improve the estimate and sequentially replan increasingly stable grasps.

4.2 Task specific vision-based grasping

In many applications optimal grasps should not be only stable in physical aspect but also viable for a particular task. For instance, tools such as screwdrivers, knives or hammers, need to be grasped by their handles to use them. Task specific grasping is closely related to category-based grasping, because in real scenarios a specific task can be performed only with objects from a particular category.

A probabilistic approach for task-specific stable grasping of objects with shape variations inside the category is proposed in Publication V. Firstly, the approach is able to generalize from a sparse set of example objects and associated grasps to novel objects from the same category. Moreover, the method does not require full 3D object models, it operates with incomplete measurements from a single RGB-D image. The main contributions of the work are: (a) the idea of maximizing grasp stability is modified to cover the shape uncertainty; (b) the method accounts for all training objects during optimization step, which ensures better generalization for new objects and allows to cope with larger shape variations; (c) unlike most data-driven techniques the approach deals with a sparse training set; (d) the method exploits partial point clouds obtained from a single RGB-D snapshot. To concentrate on grasping, such problems as category recognition and detection of affordances were left out the scope of the work.

The general framework for task-specific category-based grasping can be informally described by the diagram in Figure 4.1. The procedure consists of offline and online parts. During the offline stage a training set together

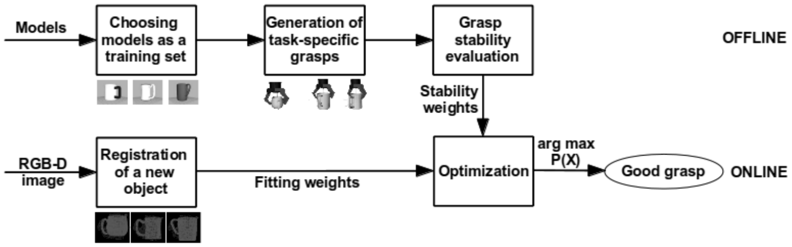


Figure 4.1. Category-based grasp generation framework: *OFFLINE stage*: 1) the training set of objects, that expresses shape variability, is chosen; 2) task specific grasps for each object are obtained (i.e. using user interaction), relative poses and corresponding stability metrics are stored; *ONLINE stage*: 1) partial point clouds for new objects are extracted from RGB-D image and registered against each training model to obtain fitting scores (metric of similarity); 2) fitting scores and stability weights are used during optimization process to determine the grasp with maximum expected stability that consistent with the training grasps. Adapted from Publication V.

with model grasps are generated and grasp stabilities are stored. It is done once per each category. The online phase is executed for each new object in the category and includes registration and optimization parts. The output from the online operation is a task-specific grasp with maximum expected stability.

It is assumed that category of objects is known a-priori and several 3D models from this category are given together with one or several corresponding task-specific grasps. The goal of the framework is to generalize from known examples to a novel object that is not included in the training dataset. The generalization is gained by applying a probabilistic approach to find the grasp which is maximally stable and at the same time consistent with a given task accounting for shape differences and possible variability in grasp location.

The general model for finding an optimal grasp as the maximum of the expected gasp stability and task compatibility over object shape variability can be described by the following equation

$$\arg\max_A E[P(S \wedge T|A,\delta)] = \arg\max_A \sum_{i,k} P(T_{i,k}|A,O_i;\theta_{i,k})P(S_{i,k}|O_i)P(O_i|\delta_i) =$$
$$\arg\max_A \sum_{i,k} P(A|T_{i,k},O_i;\theta_{i,k})\psi(q_{i,k})\phi(\delta_i), \tag{4.1}$$

where A is a 6DOF pose of the robot hand relative to the object, S denotes grasp stability and T denotes task constraint compatibility. $P(T_{i,k}|A,O_i;\theta_{i,k})$ is a probability that a grasp located at A is task compatible given model grasp location $\theta_{i,k}$. $P(S_{i,k}|O_i) \equiv \psi(q_{i,k})$ is a probability (stability weight) that the training grasp k for model i is stable as a function of stability metric $q_{i,k}$. $P(O_i|\delta_i) \equiv \phi(\delta_i)$ denotes a probability (fitting weight) that the model i can be used to generate grasps for the target object with fitting error δ_i obtained from the registration. The final equation can be obtained by applying the Bayes formula with uniform prior to the first term and omitting normalization terms because of maximization. For the density function $P(A|T_{i,k}, O_i; \theta_{i,k})$ three options were studied: Gaussian distribution, regularized Gaussian distribution (to avoid local optima) and Laplace distribution. For a more detailed description and formulation of model distributions the reader is referred to Publication V.

To find a stable task compatible grasp for a new object inside the category, a numerical optimization approach constrained by the environment geometry was applied. The process repeated the number of times equals to the number of training objects each time starting at the grasp configuration generated for the model object (local optimal solution). This process ensured that the final grasp would be in the neighbourhood of the grasps for the similar objects in the database. Finally, the result corresponding to the maximum of the objective function was selected. To obtain fitting weights the registration procedure was performed. A single RGB-D snapshot from a Kinect stereo camera was used to obtain a partial point cloud of an object after applying a planar supporting surface heuristic. Fast Point Feature Histrogram [111] and Iterative Closest Point algorithm [112] were used for alignment. More details about registration are given in Publication V.

The approach was tested both in simulation and using a real robot. The experiments in simulation were conducted in order to show that using several models in order to generate stable grasp is more beneficial compared to utilizing just the best match object. More than that, several task-specific grasps were generated to see how the method generalizes over several categories and tasks. Graspit! simulation environment with Barrett hand model were chosen for the experiments. Columbia Grasp Database [9] with object models from Princeton Shape Benchmark [113] was selected as a source for object categories. A leave-one-out cross validation was performed for two object categories: mugs (7 models) and bottles (11 models). Three task-specific grasp configurations (from the top, side and handle) for the mugs and one configuration (from the side) for the bottles were generated in simulation. Both object and robot hand poses together with automatically generated epsilon quality measures were stored for these grasps. The approach based on selecting the grasp corresponding to the best match object in the database with the highest fitting score (best single grasp) was used as a baseline approach for methods comparison. The registration part was performed using Point Cloud Library, and the algorithmic part was implemented in Matlab.

Experimental results with mug and bottle categories using both proposed framework and baseline best single grasp approach showed that the new approach using Laplace distribution outperformed the baseline technique for both mugs and bottles. The basic Gaussian approach performed the worst because of getting stuck in the local optima and was not used later in the real experiments. Regularized Gaussian approach just slightly outperformed the baseline. Thus, to show the advantages of the new approach, experiments on a real platform were conducted. Moreover, there could be a bias in the simulation results because of utilizing complete models of target objects. However, in failure cases baseline approach most often initially collided with the object or did not touch it at all. In case of regularized Gaussian and Laplace approaches the main problem was the lack of precision, so that small perturbations in robot hand locations could make such grasps stable.

An additional experiment with tools (4 hammers and 2 knives) was conducted in order to show that the proposed approach is able to generate stable grasps not only for objects inside the same category but also for other objects similar to the training models. Both hammer and knife have elongated shape and can be divided into handle and working parts. One grasp from the handle was manually generated for hammers as training data. As a result of the experiment, the task-specific category-based technique outperformed the baseline approach and found stable grasps for both knives. Thus, the proposed approach was able to generate stable grasps for objects from the other class, which are similar to the models in the database.

A KUKA LBR4+ robotic arm with a 3-fingered Barrett BH8-282 hand was used for real experiments. The testing set consisted of 5 mugs varying in shape and size. The training set was the same as in simulation. The stability of the grasp was evaluated from human observations after lifting and manually disturbing the object. The experimental results demonstrated that the proposed approach outperformed the baseline method with both distribution types, with the Laplace method performing the best. For the proposed approach the most frequent reason for failures was reachability problem, when the robot was not able to find an Inverse Kinematics solution. This problem can be solved by including more constraints in the optimization process. The baseline approach failed the most often because of shape dissimilarities. In this case the best match model in the database considerably differed from the target object in size or shape. The proposed approach avoided this problem because it accounted not only the most similar model but all training objects. Figure 4.2 shows an example for one mug when the baseline approach failed and the Laplace approach was able to find stable grasp. More experimental details can be found in Publication V.

Figure 4.2. Resulting grasps: (left) Best single match; (right) Laplace model. Adapted from Publication V.

To further analyze the proposed method it was studied how the use of several models changes the grasps. For this purpose real object point clouds were registered with the models, then the grasp was optimized and projected back to the models using the corresponding registration result. The goal was to see how the resulting grasp differs from the model grasps. Figure 4.3 illustrates grasps generated for one of the testing mugs projected back to model objects. The main observations from the experiments were: resulting grasps differ from model grasps (the approach interpolates over multiple models and provides a degree of generalization); the registration can fail for single object model as the fitting weight is not able to capture if the object was registered correctly (i.e. model 3 in Figure 4.2 was registered upside down), but the use of several models can solve this problem; small fitting weights often indicate registration failures (see model 4 in Figure 4.2), however, low probabilities decrease the effect of such models on the final grasp.

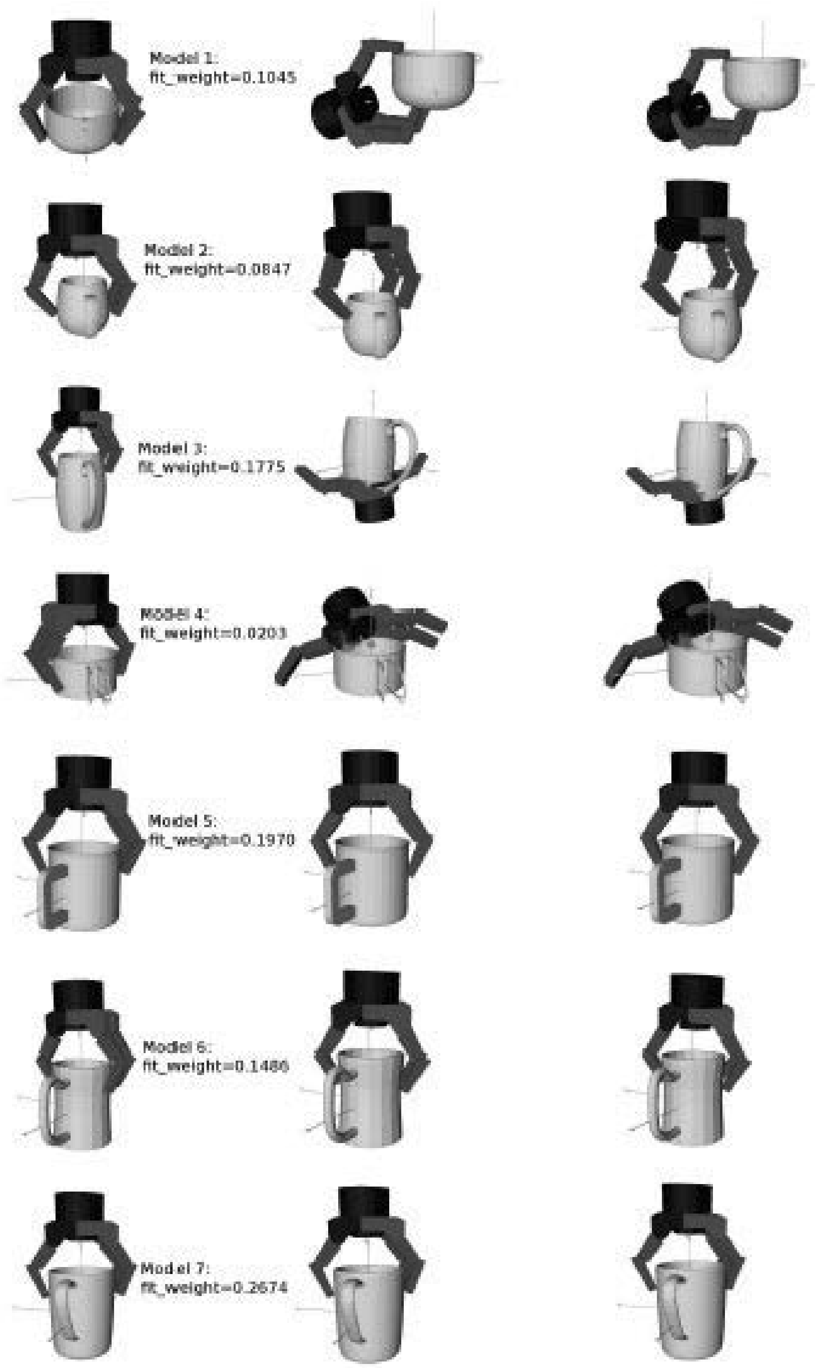


Figure 4.3. Resulting top grasps generated for the white mug: *(left)* model grasp; *(center)* regularized Gaussian model; *(right)* Laplace model; *fit_weight* - fitting probabilities (weights). Adapted from Publication V.

4.3 Using tactile feedback to improve the performance

The framework presented in Publication V was extended to incorporate tactile sensor feedback in order to improve the estimate and to sequentially replan increasingly stable grasps. Experimental results showed that the initial approach described in Publication V sometimes produces partially stable grasps. Vision is able to provide an initial estimate for the grasp, but because of sensor errors and self-occlusion it is not accurate enough to ensure the stability of the resulting grasp. The modified approach combines the ideas of planning for maximally stable and task compatible grasps, using vision for producing initial estimates, and tactile-based grasping. Knowledge about object parameters such as shape and pose, represented by probability distributions over the model objects, are updated based on collected tactile feedback. Thus, if the grasp resulting from the vision-based approach is not sufficiently stable, tactile information is used to replan the most stable task-specific grasp.

The overall process of finding an optimal grasp for the new object inside the particular category can be expressed by the sequential pipeline shown in Figure 4.4. The upper part of the diagram corresponds to the basic visual-based approach presented in Publication V. The lower part conforms to the updates based on the tactile feedback.

The extended approach updates object shape and pose estimates by maximizing a posteriori probability given tactile measurements, current estimates and executed grasp configuration. Mathematically the process of model updating from time $t-1$ to t can be expressed by:

$$\underset{w_i(t),\vec{p}_i(t)}{\arg\max} P(w_i(t),\vec{p}_i(t)|M(t),G(t-1),w_i(t-1),\vec{p}_i(t-1)) = \tag{4.2}$$

$$\underset{w_i(t),\vec{p}_i(t)}{\arg\max} P(M(t)|w_i(t),\vec{p}_i(t),G(t-1))P(w_i(t-1),\vec{p}_i(t-1)), \tag{4.3}$$

where $w_i(t)$ is a shape goodness-of fit weight (probability), $\vec{p}_i(t)$ is an object pose for each model i, $M(t)$ denotes tactile measurements and $G(t)$ indicates a grasp configuration. The resulting Equation (4.3) is obtained by applying the Bayes formula and omitting constants which do not affect the maximization.

The update of shape and pose attributes is performed independently. First, goodness-of-fit weights and tactile measurements $M = \hat{m}$ are fixed and object poses $\vec{p}_i$ are optimized. To find the pose a standard unconstrained nonlinear trust-region optimization approach is used to maxi-

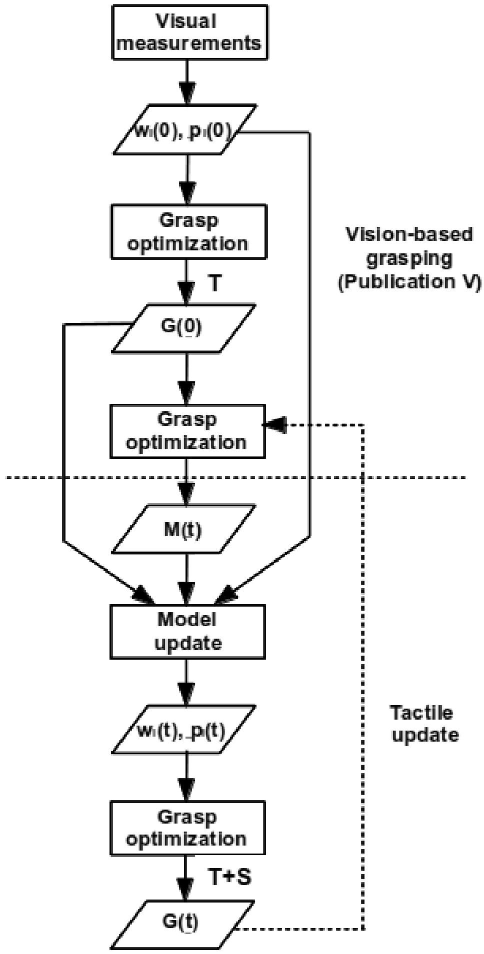


Figure 4.4. Combined framework for task-specific grasping of similar objects: Vision-based technique provides initial estimates of objects parameters expressed by shape goodness-of-fit weights (probabilities) $w_i(0)$ and object poses $\vec{p_i}(0)$ for each model i. An initial grasp configuration $G(0)$ is optimized by using visual estimates and accounting for task constraints (**T**) modeled as pose similarity to demonstrated task-specific grasps. If the grasp after execution is not stable enough tactile information $M(t)$ is collected. Model is, then, updated to find pose and shape estimates, $w_i(t)$ and $\vec{p_i}(t)$, with maximum a posteriori probability using current estimates of shape and pose, executed grasp configuration and tactile measurements. Next, grasp optimization is repeated by considering both task specificity and grasp stability (**T+S**). After that grasp is executed, new tactile data is collected and the decision about new round of optimization is made. Adapted from Publication VI.

mize the probability

$$\vec{p}_i(t) = \arg\max_{\vec{p}_i} P(M = \hat{m}|\vec{p}_i(t), G(t-1))P(\vec{p}_i(t-1)), \tag{4.4}$$

The probability of measurements is modeled using GPR and FITC (to reduce computations) with data from simulation. The prior about the pose is modeled using a Gaussian distribution. After all new poses $\vec{p}_i(t)$ are found, goodness-of-fit weights can be updated for each model i as a ratio of the likelihood of each model to the total likelihood

$$w_i(t) = \frac{P(M = \hat{m}|\vec{p}_i(t), G(t-1))}{\sum_i P(M = \hat{m}|\vec{p}_i(t), G(t-1))}. \tag{4.5}$$

The process of updating poses and fitness weights can, then, be iterated in order to improve the resulting grasp configuration.

The updated object pose and shape estimates are further utilized during grasp optimization. The general model for finding stable task-specific grasp can be formulated as:

$$P(S \wedge T|G, w, \vec{p}) = P(S|G, w, \vec{p})P(T|G, w, \vec{p}) \tag{4.6}$$

The main assumption is that stability is independent of task constraints given object poses and goodness-of-fit weights. The task compatibility is defined as similarity to manually generated task-specific grasps and is modeled by a sum of Laplace distributions, each centered at a demonstrated grasp. The stability probability is modeled using GPR and is evaluated using the negative exponential mapping function. All equations and detailed descriptions of the models used in the approach are given in Publication VI.

To test the hypothesis that tactile feedback can increase the quality of the grasps, experiments in simulation were conducted. Graspit! simulator in combination with Matlab and models from CGDB mug category were used as experimental setup. 100 top grasp configurations were randomly generated in the manually defined area to approximate the results from the vision-based approach. Cross-validation leave-one-out test was applied with the mugs. GPR was used to construct the probabilistic models of tactile measurements (finger joint angles) and stability of the grasp. The process of generating data for GP and defining hyperparameters is discussed in Publication VI. In general the experiment followed the flowchart shown in Figure 4.4. Out of 100 trials 45 resulting grasps become more stable than initial grasps, 34 less stable and 21 remained the same (both initial and resulting grasps were unstable). Overall the procedure improved the quality of the grasps, because the approach resulted in

25 stable grasps having initial unstable guesses and failed to find stable grasps for initial stable proposals only in 21 cases. For equal and worse groups of grasps tactile information collected from only one grasp was not able to decrease the uncertainty in the model poses. To see how the results can be further improved using the proposed approach another two rounds of optimization were performed. From totally 42 initially unstable grasps 15 stable grasps were obtained after the second round and 5 more after the third round. This demonstrated the ability of the method to collect more information about an object and, as a result, incrementally improve the quality of the grasps over the time horizon. The sequence of sequential improvements of grasps is illustrated in Figure 4.5.

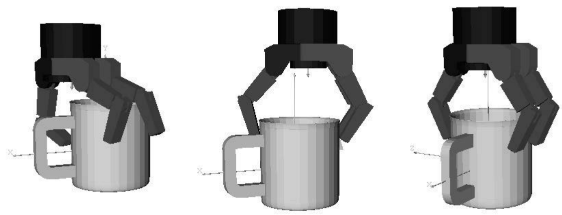


Figure 4.5. Sequence of resulting grasps: *(left)* initial unstable grasp, *(middle)* unstable grasp after first round of grasp optimization, *(right)* stable grasp after second round of grasp optimization. Adapted from Publication VI.

To study the superiority of the combined approach over the purely vision-based technique and to show how the grasp quality can be improved after collecting more tactile information experiments with a real robot were performed. A KUKA LBR4+ robotic arm with a 3-fingered Barrett Hand BH8-282 were utilized in experiments. Similarly to the previous study the table-top scenario and top grasp configurations were considered. The test set consisted of 8 mugs varying in shape and size, and also differing from the training models (the same CGDB mugs as in simulation). For real experiments the models using GPR were rebuilt with different mean, covariance and hyperparameters. This step is described in Publication VI experimental part. The combined approach was tested by following the general procedure described by equation 4.6. The search was performed in 4DOF space (top configuration fixed 2DOF). First, the modified vision-based approach was applied, the resulting grasp was performed and it was decided continue or not. If the grasp was unstable the tactile update was performed and the stability was checked again. If the grasp was still unstable, a second round of optimization was applied. If at any stage the grasp was considered stable, the mug was lifted and the stability was checked manually. For all objects 2 or 3 trials were done (totally 19 sequences). The baseline approach succeeded only in 3 of 19 cases, the modified approach improved 5 out of 16 grasps already after the first round and 6 out of 9 (2 grasps failed during the first round) after the second round. Figure 4.6 shows a sequence of resulting grasps. Thus,

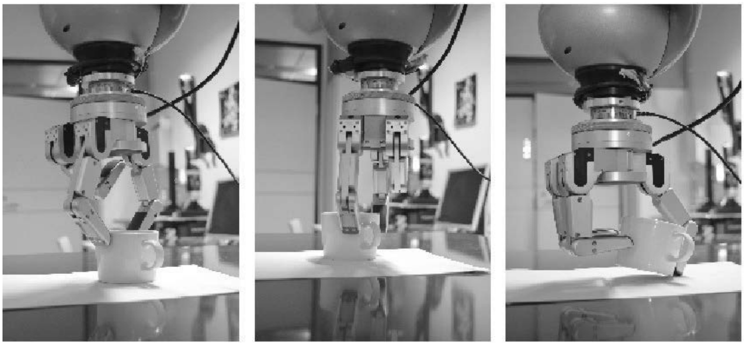


Figure 4.6. Resulting grasps (3 steps sequence): *left* - vision-based approach, *middle* - combined approach (first trial), *right* - combined approach (second trial). Adapted from Publication VI.

the modified approach incrementally improved the quality of the grasp by collecting more information about the graspable object. The most typical failure case was when the mug slipped out the hand after closing the fingers. This problem can be solved by modifying the closing hand procedure.

4.4 Discussion

In Publication V, a novel approach for task-specific stable grasping of objects with shape variations inside a category was presented. The method accounts for all training objects during the optimization according to their importance based on fitting and stability aspects, which ensures better generalization properties to handle larger shape variability compared to the traditional approaches based on most similar model's grasp. The proposed method is close to data-driven approaches during the model building stage, but it does not require a construction of a large experience database as it operates already with a single task-specific stable grasp per object. Because of its probabilistic nature the general model for finding stable grasps for familiar objects can cope with shape uncertainties and is able to find a grasp that is at the same time the most stable and task compatible. The experimental results with tools showed that the approach can generalize also for similar objects from different categories.

The results from the simulation and on a real platform demonstrated the superiority of the new approach over the baseline best match grasp technique. So, these findings demonstrate the ability of the statistical data-driven approach to generalize from individual examples. This can be treated complementary to approaches that perform deformable shape alignment or grasp adjustment to achieve generalization. However, the advantage of the statistical approach is that it can deal with partial data of the target object, which is a challenging task for deformable alignment.

Nevertheless, they are several open problems in applying the approach. First of all, the reachability problem that can be solved by modifying the set of constraints in the optimization part. Secondly, the registration procedure should be improved to provide better results in case when there is no perfect match between the target and model objects. Partially stable grasps (objects moved a bit inside the hand during lifting) can be improved by performing further adjustments. One possible way, which was implemented in Publication VI is to collect sensor measurements, e.g. from tactile sensors, and use this feedback to replan the grasp. In Publication V an object category is assumed to be known. The study can be extended to deal with uncategorized objects and as a first step perform category recognition. Moreover, the detection of affordances, i.e. if an object affords a particular action, is not considered. The task-based criterion of a grasp comes from human demonstration and represents a label for the training data. However, this criterion can be learned from human demonstrations, e.g. [64, 98].

In Publication VI the vision-based method for task-specific grasping of objects from a known category was modified to include tactile feedback in order to improve the quality of the grasp. Thus, the new approach combines RGB-D and tactile measurements in a probabilistic framework which allows to decrease the uncertainty about object pose and shape by incrementally collecting more information about the object and generating more stable grasps. The approach accounts for both task constraints learned from human annotated grasps and grasp stability modeled using GPR. Experimental results demonstrated that including tactile measurements in the optimization process can improve the grasp performance over vision even for objects significantly varying in shape. More than that, further iterations of collecting tactile experience can further improve the grasp quality in case when earlier grasps are not satisfactory.

To improve the performance of the method and avoid failure cases the closing fingers procedure can be modified, e.g. by slowing down the robot motion. The generalizability of the approach has not been experimentally studied. Thus, experiments with categories other than mugs should be done to verify the generalization boundaries of the framework.

5. Conclusion

This dissertation concentrates in developing approaches to address the challenges in grasp planning for known and familiar objects under different types of uncertainty such as object location and shape by perceiving the environment using sensors. Prior grasp planning approaches often assume perfect knowledge about target object attributes. However, in real applications geometric models of the objects are often incomplete and inaccurate. Firstly, the lack of exact geometric information can be compensated by sensory feedback. Moreover, probabilistic formulation of grasp planning can provide a robot with the capability to cope with uncertainties. Application of probabilistic models in robotic manipulation is a prominent direction because probabilistic models allow to represent uncertain beliefs and some of them can handle even multi-modal uncertainties, e.g. in tactile manipulation because of the local nature of tactile measurements. The publications in this thesis propose methods that in presence of uncertainty in object attributes allow to find stable and useful grasps.

The initial research towards grasp planning under uncertainty was presented in Publication I by looking at how much information a robot can optimally learn from a single tactile exploration attempt. Going further, Publication II proposed a probabilistic approach for grasp planning under pose uncertainty using on-line sensory information and simultaneously updating the knowledge about object attributes. MCMC methods were utilized to sample the evolving probability distributions and Bayesian approach was used to obtain the result by marginalizing over the current knowledge. The core of the approach was a Bayesian network for object knowledge refinement based on grasp stability maximization that modelled the relationships between object attributes, action (grasp) attributes, on-line sensor readings and success metric. This general framework allowed to accomplish statistically optimal grasp planning, while simultaneously reducing uncertainty about the environment. The model can be further extended to include more dependences between variables. In Publications III-IV an extension of the basic stability maximizing framework was developed. The modified approach unified the ideas of stability maximization, information gathering by minimizing the entropy and using sensor's feedback. An information gathering was performed by exploratory entropy-based procedure. An efficient discrete entropy estimate that uses only particle weights was proposed to measure the entropy of a distribution of the object pose attributes represented by a set of particles. Thus, the combined approach allowed to alternate between stability maximization and entropy minimization that allowed to improve the stability of the resulting grasp by decreasing the uncertainty in knowledge about object attributes.

All aforementioned techniques were focused on grasping known objects. The studies in Publication IV-V were concentrated on grasping familiar objects belonging to the same known category. Moreover, the task aspect was also taken into consideration because usually the target grasp should be not only stable but compatible for a particular task to be performed with an object. A task-specific category-based probabilistic method described in Publication IV allowed to generalize from a sparse set of training examples to novel objects. The idea of stability maximization from Publication II was taken in the new context to cover shape uncertainty. An approach used RGB-D vision data and dealt with partial point clouds. An approach in Publication VI extended the vision-based method by incorporating tactile sensor feedback in order to iteratively improve the beliefs about unknown object pose and shape and to generate better grasps.

The experimental results in simulation and on a real platform showed the viability of the proposed methods. More experiments with other test objects and categories as well as models modifications can be done in order to improve the performance. In the future it would be interesting to apply deformable object registration for a grasp planning to obtain fuzzy correspondences by modeling the deformation between the sparse point sets [114, 115]. In this dissertation a task for a grasp is assumed to be known, although, the framework could benefit by including a detection of object affordances. Moreover, grasping of unknown objects is a prominent research area and further studies could be conducted in this direction. The results of Publication VI showed that using both visual and tactile sensors is more beneficial for grasp performance comparing to a single sensor. So, a set of sensors could be used in order to reduce the uncertainty by collecting more information about a goal object and improve a quality of a grasp. One more possible direction of progress is to go further to multi-step optimization in order to find a series of actions which maximizes the success metric in a longer time horizon. For instance, partially observable Markov decision process [116, 117] can be used to model a process of planning under uncertainty with imperfect sensing . Nevertheless, in general, the work presented in this dissertation is a step towards creating autonomous robot which can be placed in a new unstructured environment and which then knows how to adapt its behaviour to a new environment in order to perform a particular task.

References

[1] R. Suárez, M. Roa, and J. Cornella. Grasp quality measures. Technical report, Technical University of Catalonia, 2006.

[2] J. Kerr. *An analysis of multifingered hand*. PhD thesis, Stanford University, Dapartment of Mechanical Engineering, Dec 1984.

[3] Z. Li and S. Sastry. Task oriented optimal grasping by multifingered robot hands. In *1987 IEEE International Conference on Robotics and Automation (ICRA)*, volume 4, pages 389–394, Mar 1987.

[4] B.S. Baker, S. Fortune, and E. Grosse. Stable prehension with three fingers. In *Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing*, STOC '85, pages 114–120, New York, NY, USA, 1985. ACM.

[5] M.R. Cutkosky. *Robotic Grasping and Fine Manipulation*. Kluwer Academic Publishers, Norwell, MA, USA, 1985.

[6] G.I. Boutselis, C.P. Bechlioulis, M.V. Liarokapis, and K.J. Kyriakopoulos. Task specific robust grasping for multifingered robot hands. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, pages 858–863, Sept 2014.

[7] A.T Miller and P.K. Allen. Graspit!: A versatile simulator for robotic grasping. *IEEE Robotics and Automation Magazine*, 11:110–122, 2004.

[8] R. Diankov and J. Kuffner. Openrave: A planning architecture for autonomous robotics. Technical report, Robotics Institute, Pittsburgh, PA, July 2008.

[9] C. Goldfeder, M. Ciocarlie, H. Dang, and P.K. Allen. The columbia grasp database. In *2009 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1710–1716, May 2009.

[10] D. Kappler, J. Bohg, and S. Schaal. Leveraging big data for grasp planning. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015.

[11] J. Bohg, A. Morales, T. Asfour, and D. Kragic. Data-driven grasp synthesis – a survey. *IEEE Transactions on Robotics*, 30(2):289–309, April 2014.

[12] D. Kragic, A.T. Miller, and P.K. Allen. Real-time tracking meets online grasp planning. In *2001 IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 2460–2465, 2001.

[13] K. Huebner, S. Ruthotto, and D. Kragic. Minimum volume bounding box decomposition for shape approximation in robot grasping. In *2008 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1628–1633, May 2008.

[14] A. Collet, D. Berenson, S.S. Srinivasa, and D. Ferguson. Object recognition and full pose registration from a single image for robotic manipulation. In *2009 IEEE International Conference on Robotics and Automation (ICRA)*, pages 48–55, May 2009.

[15] Z.C. Marton, D. Pangercic, N. Blodow, and M. Beetz. Combined 2D-3D Categorization and Classification for Multimodal Perception Systems. *The International Journal of Robotics Research*, 30(11):1378–1402, September 2011.

[16] J.R. Flanagan, M.C. Bowman, and R.S. Johansson. Control strategies in object manipulation tasks. *Current Opinion in Neurobiology*, 16(6), December 2006.

[17] T. Okada and S. Tsuchiya. Object recognition by grasping. *Pattern Recognition*, 9(3):111–119, 1977.

[18] W.D. Hillis. *Active Touch Sensing*. A.I. memo 629. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1981.

[19] N. Gorges, S.E. Navarro, D. Goger, and H. Worn. Haptic object recognition using passive joints and haptic key features. In *2010 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2349–2355, May 2010.

[20] A. Petrovskaya, O. Khatib, S. Thrun, and A.Y. Ng. Bayesian estimation for autonomous object manipulation based on tactile sensors. In *2006 IEEE International Conference on Robotics and Automation (ICRA)*, pages 707–714, May 2006.

[21] K. Hsiao, L. Kaelbling, and T. Lozano-Perez. Task-driven tactile exploration. In *Proceedings of Robotics: Science and Systems*, Zaragoza, Spain, June 2010.

[22] H. Dang, 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 (ICRA)*, pages 5917–5922, May 2011.

[23] H. Dang and P.K. Allen. Stable grasping under pose uncertainty using tactile feedback. *Autonomous Robots*, 36(4):309–330, 2014.

[24] E. Hyttinen, D. Kragic, and R. Detry. Learning the tactile signatures of prototypical object parts for robust part-based grasping of novel objects. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015.

[25] M.L. McLaughlin, G. Sukhatme, and J. Hespanha. *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001.

[26] R. S Johansson and J. R. Flanagan. Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews Neuroscience*, 10(5):345–359, 2009.

[27] K.C. Sahoo and C. Menq. Localization of 3-d objects having complex sculptured surfaces using tactile sensing and surface description. *ASME Journal of engineering for industry*, 113(1):85–92, 1991.

[28] L. Zhu, H. Luo, and H. Ding. Optimal design of measurement point layout for workpiece localization. *ASME Journal of engineering for industry*, 131(1), 2008.

[29] Leon D. Harmon. Automated touch sensing: A brief perspective and several new approaches. In *1984 IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 326–331, March 1984.

[30] R.D. Howe. Tactile sensing and control of robotic manipulation. *Advanced Robotics*, 8(3):245–261, 1993.

[31] A.M. Okamura and M.R. Cutkosky. Haptic exploration of fine surface features. In *1999 IEEE International Conference on Robotics and Automation (ICRA)*, volume 4, pages 2930–2936 vol.4, 1999.

[32] O. Kroemer, Christoph H. Lampert, and J. Peters. Learning dynamic tactile sensing with robust vision-based training. *IEEE Transactions on Robotics*, 27(3):545–557, June 2011.

[33] S. Chitta, J. Sturm, M. Piccoli, and W. Burgard. Tactile sensing for mobile manipulation. *IEEE Transactions on Robotics*, 27(3):558–568, June 2011.

[34] P.A. Schmidt, E. Maël, and R.P. Würtz. A sensor for dynamic tactile information with applications in human–robot interaction and object exploration. *Robotics and Autonomous Systems*, 54(12):1005–1014, 2006.

[35] S. Caselli, C. Magnanini, and F. Zanichelli. Haptic object recognition with a dextrous hand based on volumetric shape representations. In *IEEE Int. Conf. on Multisensor Fusion and Integration, Las Vegas, NV*, pages 2–5, 1994.

[36] P.K. Allen and K.S. Roberts. Haptic object recognition using a multi-fingered dextrous hand. In *1989 IEEE International Conference on Robotics and Automation (ICRA)*, pages 342–347 vol.1, May 1989.

[37] A. Schneider, J. Sturm, C. Stachniss, M. Reisert, H. Burkhardt, and W. Burgard. Object identification with tactile sensors using bag-of-features. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, pages 243–248, Oct 2009.

[38] Z. Pezzementi, E. Plaku, C. Reyda, and G.D. Hager. Tactile-object recognition from appearance information. *IEEE Transactions on Robotics*, 27(3):473–487, June 2011.

[39] A. Bierbaum and M. Rambow. Grasp affordances from multi-fingered tactile exploration using dynamic potential fields. In *2009 9th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 168–174, Dec 2009.

[40] Y. Bekiroglu, J. Laaksonen, J.A. Jorgensen, V. Kyrki, and D. Kragic. Assessing grasp stability based on learning and haptic data. *IEEE Transactions on Robotics*, 27(3):616–629, June 2011.

[41] K.T. Gunnarsson and F.B. Prinz. Cad model-based localization of parts in manufacturing. *Computer*, 20(8):66–74, Aug 1987.

[42] Y. Chu. *Workpiece Localization: Theory, Algorithms and Implementation*. Hong Kong University of Science and Technology, 1999.

[43] J. Laaksonen and V. Kyrki. Probabilistic approach to sensor-based grasping. In *ICRA 2011 Workshop: Manipulation Under Uncertainty*, 2011.

[44] A. Petrovskaya and O. Khatib. Global localization of objects via touch. *Robotics IEEE Transactions on Robotics*, 27(3):569–585, June 2011.

[45] Y. Chebotar, O. Kroemer, and J. Peters. Learning robot tactile sensing for object manipulation. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, pages 3368–3375, Sept 2014.

[46] J. Kennedy and R. Eberhart. Particle swarm optimization. In *IEEE International Conference on Neural Networks, 1995.*, volume 4, pages 1942–1948 vol.4, Nov 1995.

[47] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *SCIENCE*, 220(4598):671–680, 1983.

[48] X. Yang. *Nature-Inspired Metaheuristic Algorithms: Second Edition*. Luniver Press, 2010.

[49] A.S. Vazquez, R. Fernandez, A. Lopez, E. Valero, I. Payo, and A. Adan. In-hand object localization: Simple vs. complex tactile sensors. In *SENSORS, 2014 IEEE*, pages 1710–1713, Nov 2014.

[50] C. Ferrari and J. Canny. Planning optimal grasps. *1992 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2290–2295, 1992.

[51] C. Goldfeder, P.K. Allen, C. Lackner, and R. Pelossof. Grasp planning via decomposition trees. In *2007 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4679–4684, April 2007.

[52] M. Przybylski, T. Asfour, and R. Dillmann. Planning grasps for robotic hands using a novel object representation based on the medial axis transform. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011)*, pages 1781–1788, Sept 2011.

[53] R. Diankov. *Automated Construction of Robotic Manipulation Programs*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Sept 2010.

[54] J. Weisz and P.K. Allen. Pose error robust grasping from contact wrench space metrics. In *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pages 557–562, May 2012.

[55] C. de Granville, J. Southerland, and A.H. Fagg. Learning grasp affordances through human demonstration. In *2006 International Conference on Development and Learning*, 2006.

[56] R. Detry, E. Baseski, M. Popovic, Y. Touati, N. Kruger, O. Kroemer, J. Peters, and J. Piater. Learning object-specific grasp affordance densities. In *2009 IEEE 8th International Conference on Development and Learning (ICDL)*, pages 1–7, June 2009.

[57] J. Romero, H. Kjellstrom, and D. Kragic. Modeling and evaluation of human-to-robot mapping of grasps. In *2009 International Conference on Advanced Robotics (ICAR)*, pages 1–6, June 2009.

[58] M. Do, J. Romero, H. Kjellstrom, P. Azad, T. Asfour, D. Kragic, and R. Dillmann. Grasp recognition and mapping on humanoid robots. In *2009 9th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 465–471, Dec 2009.

[59] S. Ekvall and D. Kragic. Learning and evaluation of the approach vector for automatic grasp generation and planning. In *2007 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4715–4720, April 2007.

[60] C. Goldfeder, M. Ciocarlie, J. Peretzman, H. Dang, and P.K. Allen. Data-driven grasping with partial sensor data. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, pages 1278–1283, Oct 2009.

[61] J.J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, 1979.

[62] C. Barck-Holst, M. Ralph, F. Holmar, and D. Kragic. Learning grasping affordance using probabilistic and ontological approaches. In *2009 International Conference on Advanced Robotics (ICAR)*, pages 1–6, June 2009.

[63] R. Detry, D. Kraft, A.G. Buch, N. Kruger, and J. Piater. Refining grasp affordance models by experience. In *2010 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2287–2293, May 2010.

[64] D. Song, K. Huebner, V. Kyrki, and D. Kragic. Learning task constraints for robot grasping using graphical models. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, pages 1579–1585, Oct 2010.

[65] R. Platt. Learning grasp strategies composed of contact relative motions. In *2007 7th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 49–56, Nov 2007.

[66] K. Gadeyne and H. Bruyninckx. Markov techniques for object localization with force-controlled robots. In *2001 International Conference on Advanced Robotics (ICAR)*, Budapest, Hungary, 2001.

[67] S.R. Chhatpar and M.S. Branicky. Particle filtering for localization in robotic assemblies with position uncertainty. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, pages 3610–3617, Alberta, Canada, 2005.

[68] R. Platt, F. Permenter, and J. Pfeiffer. Using bayesian filtering to localize flexible materials during manipulation. *IEEE Transactions on Robotics*, 27(3):586–598, June 2011.

[69] L. Zhang and J.C. Trinkle. The application of particle filtering to grasping acquisition with visual occlusion and tactile sensing. In *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3805–3812, May 2012.

[70] K.Y. Goldberg and Matthew T. Mason. Bayesian grasping. In *1990 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1264–1269 vol.2, May 1990.

[71] K. Hsiao, L.P. Kaelbling, and T. Lozano-Perez. Grasping pomdps. In *2007 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4685–4692, April 2007.

[72] A. Petrovskaya, S. Thrun, D. Koller, and O. Khatib. Towards dependable perception: Guaranteed inference for global localization. In *IARP Workshop on Technical Challenges for Dependable Robots in Human Environments*, Toulouse, France, Jun 2010.

[73] F. Veiga and A. Bernardino. Active tactile exploration for grasping. In *Proceedings of the ICRA 2013 Workshop on Autonomous Learning*, 2013.

[74] K. Hsiao, M. Ciocarlie, and P. Brook. Bayesian grasp planning. In *2011 IEEE International Conference on Robotics and Automation (ICRA) Workshop on Mobile Manipulation: Integrating Perception and Manipulation*, Shanghai, China, 2011.

[75] S. Dragiev, M. Toussaint, and M. Gienger. Uncertainty aware grasping and tactile exploration. In *2013 IEEE International Conference on Robotics and Automation (ICRA)*, pages 113–119, May 2013.

[76] C. Stachniss, G. Grisetti, and W. Burgard. Information gain-based exploration using rao-blackwellized particle filters. In *Proc. of Robotics: Science and Systems (RSS)*, Cambridge, MA, USA, 2005.

[77] C. Stachniss, G. Grisetti, and W. Burgard. Recovering particle diversity in a rao-blackwellized particle filter for slam after actively closing loops. In *2005 IEEE International Conference on Robotics and Automation (ICRA)*, 2005.

[78] Y. Taguchi, T.K. Marks, and J.R. Hershey. Entropy-based motion selection for touch-based registration using rao-blackwellized particle filtering. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011)*, pages 4690–4697, Sept 2011.

[79] C.E. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[80] E. Snelson and Z. Ghahramani. Sparse gaussian processes using pseudo-inputs. In *Advances in Neural Information Processing Systems*, volume 18, pages 1259–1266, 2006.

[81] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.

[82] J. Kennedy and R. Eberhart. Particle swarm optimization. In *IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948, Perth, Western Australia, nov/dec 1995.

[83] M. Stark, P. Lies, M. Zillich, J.L. Wyatt, and B. Schiele. Functional object class detection based on learned affordance cues. In *2008 International Conference on Computer Vision Systems (ICVS)*, May 2008.

[84] S. El-Khoury and A. Sahbani. Handling objects by their handles. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008): Workshop on Grasp and Task Learning by Imitation*, 2008.

[85] R. Pelossof, A. Miller, P. Allen, and T. Jebara. An svm learning approach to robotic grasping. In *2004 IEEE International Conference on Robotics and Automation (ICRA)*, volume 4, pages 3512–3518 Vol.4, April 2004.

[86] A. Boularias, O. Kroemer, and J. Peters. Learning robot grasping from 3-d images with markov random fields. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1548–1553, Sept 2011.

[87] A. Saxena, J. Driemeyer, and A.Y. Ng. Robotic grasping of novel objects using vision. *Int. J. Rob. Res.*, 27(2):157–173, February 2008.

[88] L. Montesano and M. Lopes. Active learning of visual descriptors for grasping using non-parametric smoothed beta distributions. *Robot. Auton. Syst.*, 60(3):452–462, March 2012.

[89] A. Saxena, L.S. Wong, and A.Y. Ng. Learning grasp strategies with partial shape information. In *Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 3*, AAAI'08, 2008.

[90] D. Rao, Q.V. Le, T. Phoka, M. Quigley, A. Sudsang, and A.Y. Ng. Grasping novel objects with depth segmentation. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2578–2585, Oct 2010.

[91] Q.V. Le, D. Kamm, A.F. Kara, and A.Y. Ng. Learning to grasp objects with multiple contact points. In *2010 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5062–5069, May 2010.

[92] J. Bohg and D. Kragic. Learning grasping points with shape context. *Robot. Auton. Syst.*, 58(4):362–377, April 2010.

[93] N. Curtis and J. Xiao. Efficient and effective grasping of novel objects through learning and adapting a knowledge base. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, pages 2252–2257, Sept 2008.

[94] U. Hillenbrand and M.A. Roa. Transferring functional grasps through contact warping and local replanning. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012)*, pages 2963–2970, Oct 2012.

[95] H. Ben Amor, O. Kroemer, U. Hillenbrand, G. Neumann, and J. Peters. Generalization of human grasping for multi-fingered robot hands. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2043–2050, Oct 2012.

[96] T. Stouraitis, U. Hillenbrand, and M.A. Roa. Functional power grasps transferred through warping and replanning. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015.

[97] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor. Learning object affordances: From sensory–motor coordination to imitation. *IEEE Transactions on Robotics*, 24(1):15–26, Feb 2008.

[98] D. Song, C.H. Ek, K. Huebner, and D. Kragic. Multivariate discretization for bayesian network structure learning in robot grasping. In *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1944–1950, May 2011.

[99] Y. Li, J.L. Fu, and N.S. Pollard. Data-driven grasp synthesis using shape matching and task-based pruning. *IEEE Transactions on Visualization and Computer Graphics*, 13(4):732–747, July 2007.

[100] H. Dang and P.K. Allen. Semantic grasping: Planning robotic grasps functionally suitable for an object manipulation task. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012)*, pages 1311–1317, Oct 2012.

[101] H. Dang and P.K. Allen. Semantic grasping: planning task-specific stable robotic grasps. *Autonomous Robots*, 37(3):301–316, 2014.

[102] M.T. Ciocarlie and P.K. Allen. Hand posture subspaces for dexterous robotic grasping. *The International Journal of Robotics Research*, 28(7):851–867, 2009.

[103] C. Goldfeder and P.K. Allen. Data-driven grasping. *Auton. Robots*, 31(1):1–20, 2011.

[104] M. Madry, D. Song, C.H. Ek, and D. Kragic. "Robot bring me something to drink from": object representation for transferring task specific grasps. In *2012 IEEE International Conference on Robotics and Automation (ICRA): Workshop on Semantic Perception, Mapping and Exploration (SPME)*, 2012.

[105] M. Madry, D. Song, and D. Kragic. From object categories to grasp transfer using probabilistic reasoning. In *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1716 –1723, May 2012.

[106] J. Bohg, K. Welke, B. Leon, M. Do, D. Song, W. Wohlkinger, A. Aldoma, M. Madry, M. Przybylski, T. Asfour, H. Marti, D. Kragic, A. Morales, and M. Vincze. Task-based grasp adaptation on a humanoid robot. In *10th IFAC Symposium on Robot Control*, 2012.

[107] F.T. Pokorny, K. Hang, and D. Kragic. Grasp moduli spaces. In *Proceedings of Robotics: Science and Systems*, Berlin, Germany, Jun 2013.

[108] F.T. Pokorny, Y. Bekiroglu, and D. Kragic. Grasp moduli spaces and spherical harmonics. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014.

[109] R. Detry, C.H. Ek, M. Madry, J. Piater, and D. Kragic. Generalizing grasps across partly similar objects. In *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3791–3797, May 2012.

[110] R. Detry, C.H. Ek, M. Madry, and D. Kragic. Learning a dictionary of prototypical grasp-predicting parts from grasping experience. In *2013 IEEE International Conference on Robotics and Automation (ICRA)*, pages 601–608, May 2013.

[111] R.B. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (fpfh) for 3d registration. In *2009 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3212–3217, May 2009.

[112] P.J. Besl and N.D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb 1992.

[113] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. In *In Shape Modeling International*, pages 167–178, 2004.

[114] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, Dec 2010.

[115] B. Jian and B.C. Vemuri. Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645, Aug 2011.

[116] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101:99–134, 1998.

[117] J. Pajarinen and V. Kyrki. Robotic manipulation of multiple objects as a pomdp. *Artificial Intelligence*, 2015. (in press).

Aalto-DD 9/2016

Advanced robots such as mobile manipulators offer nowadays great opportunities for realistic manipulators. Physical interaction with its environment is an essential capability for service robots when acting in unstructured environments such as homes. Thus, manipulation and grasping under uncertainty has become a critical research area within robotics research.

This dissertation explores approaches to address the challenges in grasp planning for known and familiar objects under different types of uncertainty such as object location and shape by perceiving the environment using sensors.

ISBN 978-952-60-6619-6 (printed)
ISBN 978-952-60-6620-2 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Electrical Engineering and Automation
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS