Synthesizing Point Clouds for Training A Grasp Planner

R & D Proposal

December 15, 2017

Author:

Minh H. Nguyen

Advisors:

Prof. Dr. Paul G. Plöger

Alex Mitrevski

Maximilian Schöbel

1   Introduction

1.1   Motivation

- Finding a suitable grasp strategy for a given object is a challenging problem, especially in the area of service robotics where the query objects are usually unknown or not fully known beforehand, the viewing point of objects changes as the robot moves, and objects oftentimes need to be grasped in a certain way to perform specific tasks. Grasp synthesis is generally divided into analytical and empirical approaches [2, 9].

    - Analytic methods often optimize some metrics which measure one or multiple properties of a robotic hand, namely dexterity, object equilibrium and stability, or the ability to exhibit a certain dynamic behavior [2]. However, these methods often make assumptions about the object (i.e. shapes or poses) and therefore are prone to errors and do not generalize well to novel objects, while taking more time to match point clouds to known models [4].

    - Empirical or data-driven approaches generally sample from several grasp candidates and rank them according to a specific metric, which can be found heuristically or generated in simulation and on real robots [2]. These approaches do not rely on objects' physical model and therefore can generalize better to familiar or unknown objects compared to analytical approaches. Furthermore, learning to grasp via mimicking human grasping eliminates the need for task modeling. However, data collection for these methods is often tedious and costly [7].

- Mahler et al. [7] proposed a pipeline to generate synthetic data (Dex-net 2.0) for training Grasp Quality Convolutional Neural Network (GQ-CNN), a model which can rapidly predict probability of successful grasp from depth image. Despite achieving high precision in successful grasp prediction, the model is limited in how grasps are represented: the grasp's approach vector and camera view point is presumed to be perpendicular to the table, while the gripper configuration and wrist orientation used as training input are simplified to the angle of the parallel-jaw grasp axis with respect to the table. The model is therefore insufficient for use with mobile robots, where the camera pose often changes as the robot moves, and the grasp approach vector is not known beforehand.

- This research aims to extend the work by Mahler et al. [7] for use in mobile robotics, specifically on the Care-O-bot 3 platform [1], through synthesizing a new dataset using a more general grasp representation and training a new CNN model on this dataset.

---

[1] https://www.care-o-bot.de/en/care-o-bot-3.html

1.2   Prior Work

Figure 1 maps the different aspects that may affect generation of grasp hypotheses. As applied to a specific use case to be described in section 1.4, the proposed approach will focus on 3D object features and synthesizing grasp candidates for familiar objects. Section 1.3 will describe the robot platform, including arm and gripper configuration. This section will focus on surveying literature about object-grasp representation, 3D grasp synthesis, 3D grasp quality metrics, and task compatibility. In the final part, a description of the Dex-net data generation process will be included.

Figure 1: Aspects which may influence generation of grasp hypotheses [2].

1.2.1   Object-Grasp representation

Bohg et al. [2] parameterized grasps as

- the *grasping point* of the object where the gripper should be aligned,

- an *approach vector* from which the gripper shall approach the *grasping point,*

- the *wrist orientation* of the robotic hand,

- and an *initial gripper configuration.*

The article also categorized object-grasp representation approaches into ones which extract local (i.e. curvature, contact area with the hand) or global features (center of mass, bounding box). In generating Dex-net 2.0, Mahler et al. represent grasps in 2D images by aligning the image center to the gripper central point and the image's middle row to the grasp axis [7]. In general, however, object-grasp representation is highly dependent on the specific hardware setup (i.e. the robot arm and gripper).

1.2.2   3D grasp synthesis

One of the key element of grasping strategies is stability. As defined in [8], a grasp is stable if position error of object caused by a disturbance disappears after the disturbance vanishes. Two important subsets of stable grasps are *force closure* and *form closure.* For a force closure grasp, all fingers can apply force on object to produce wrenches in all directions. A grasp that can achieve force closure with frictionless point contacts achieves form closure [9].

Grasp synthesis is generally categorized into analytical and empirical approaches: analytical methods (figure 2) rely on geometric, kinematic and/or dynamic formulations, while empirical ones (figure 3) generally build a knowledge base of grasp experience for sampling grasp candidates and rank them based on some quality metric [9, 2].

1. Analytical approaches are categorized in [9] into methods which compute force-closure grasps and ones which consider task compatibility. Force-closure methods are further divided into methods which synthesize force-closure grasps and methods which attempt to find an optimal force-closure grasp using a grasp quality metric. Suárez reviewed grasp quality metrics extensively in [10], and a more recent review was done by Roa and Suárez in [8].

2. While Sahbani et al. categorized empirical approaches into techniques which focus on human observation and ones focusing on object observation [9], Bohg et al. argued that sorting these techniques based on the assumption of what is known about the query objects would better capture the diversity of the techniques [2]. Empirical approaches are thus divided in the latter article into ones which deal with known, familiar or unknown objects.
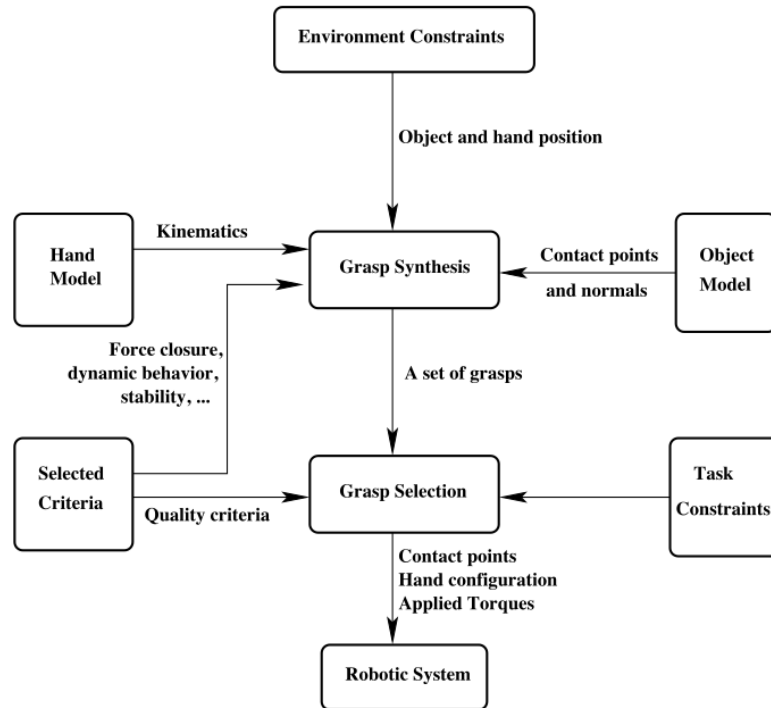
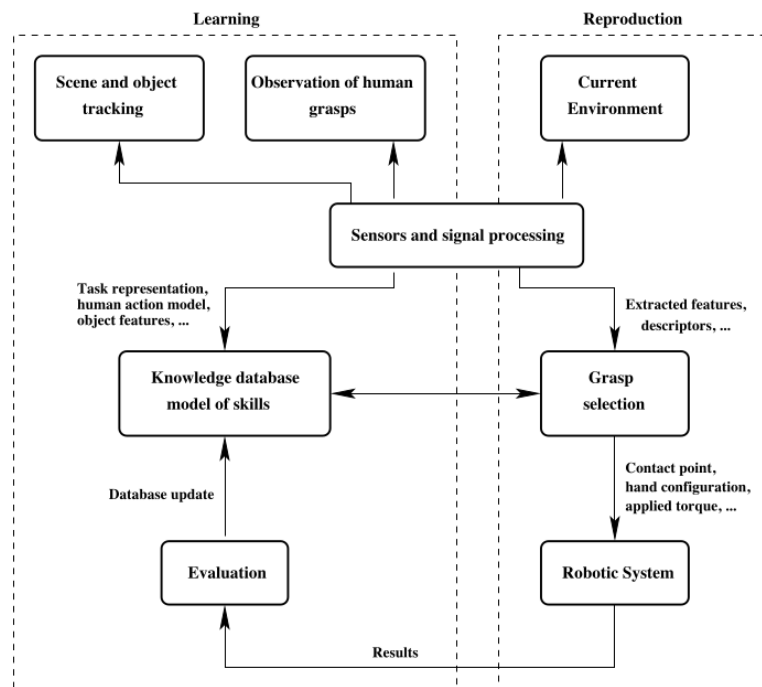Figure 2: Grasp synthesis using analytical approaches [9].

Figure 3: Grasp synthesis using empirical approaches [9].

1.2.3   3D grasp quality metrics

Group	Subgroup	Quality index	Criterion
Measures related to the position of the contact points on the object	Based on algebraic properties of G	Minimum singular value of G	Maximize
		Volume of the ellipsoid in the wrench space	Maximize
		Grasp isotropy index	Maximize
	Based on geometric relations	Shape of the grasp polygon[a]	Minimize
		Area of the grasp polygon	Maximize
		Distance between the centroid C and the center of mass CM	Minimize
		Orthogonality	Minimize
		Margin of uncertainty in finger positions[b]	Maximize
		Based on independent contact regions	Maximize
	Considering limitations on the finger forces	Largest-minimum resisted wrench	Maximize
		Volume of the Grasp Wrench Space	Maximize
		Decoupled forces and torques	Maximize
		Normal components of the contact forces	Minimize
		Coplanarity of the normals[a]	Minimize
		Task oriented measures	Maximize
Measures related to hand configuration		Distance to singular configurations	Maximize
		Volume of the manipulability ellipsoid	Maximize
		Uniformity of transformation	Minimize
		Finger joint positions	Minimize
		Similar flexion values	Minimize
		Task compatibility index	Maximize
		Safety margin	Maximize
Other measures		Biomechanical fatigue	Minimize
		Deviation in object pose	Minimize

[a] Applicable only to 2D and 3D planar grasps
[b] Applicable only to 2D grasps

Table 1: Grasp quality measures [8].

Roa and Suárez group grasp quality measures into approaches focusing on contact point position, hand configuration, and ones which combine both metric types [8].

- Contact point grasp quality measures focus on the object's properties, friction constraints and form/force closure conditions.

- Hand configuration measures focus on studying the properties of the hand-object Jacobian H.

- Grasp measures can be combined serially or in parallel to capture different aspects of a grasp's quality. In the serial approach, one quality measure is applied to find several grasp candidates, then another metric is used to choose the optimal candidate. The parallel approach combines the different measures in a single global index.

Table 1 illustrates how contact point and hand configuration related approaches can be further categorized into subgroups, and include metrics based on bio-mechanical fatigue and deviation of object pose which are not encapsulated by the aforementioned groups.

1.2.4   Task Compatibility

In addition to stability, a successful grasping strategy also has to take into account the task to be performed with the grasped object. However, Sahbani et al. [9] points out that incorporating task compatibility in to grasp synthesis, is still an open challenge because of the difficulty in modeling a task and the computational cost of finding a suitable grasp once a task is defined. These challenges are particularly relevant to analytic grasp synthesis, but is also not solved by empirical approaches. Methods which mimic human grasping alleviate the need for modeling tasks but is not fully automatic when facing new objects, while methods based on object observations have to generate many grasp candidates and face the same difficulty of task modeling as analytical approaches. One possible approach suggested by the article is to directly identify object features that are relevant to the requested tasks. Approaches to task modeling include methods based on manipulability ellipsoids [11] and task space polytopes [5], or building a task compatibility index measuring the similarity between the optimal directions of the manipulator and the movements required by a specific task [3].

1.2.5   Dex-Net 2.0 Data Generation

- Compute stable object poses from 3D mesh models from Dex-Net 1.0 [6], object meshes are rescaled to fit within the gripper physical constraints, and store poses with probability of occurrence above a threshold.

- Grasps are sampled using rejection sampling to ensure coverage of the object surface, ranked and selected using a grasp quality metric.

- 2D images are generated from grasp candidates as training data with the antipodal axis aligned to the middle row of the image.

1.3   Hardware Specification

The robot platform to be used for experimentation is the Care-o-bot 3 from Fraunhofer Institute for Manufacturing Engineering and Automation [2], mounted with a Schunk SDH [3] gripper and an ASUS Xtion PRO RGB-D camera [4].

Specification name	Value
Number of fingers	3
Maximum admissible finger length (mm)	155
Finger distance (mm)	66
Degrees of freedom	7
Degrees of freedom per finger	2
Degrees of freedom 2-finger-rotation	1

Table 2: Schunk SDH gripper specifications.

Specification name	Value
Depth image size	VGA (640x480) : 30fps QVGA (320x240): 60fps
Maximum admissible finger length (mm)	155
Field of view (Horizontal, Vertical, Diagonal)	58°H, 45°V, 70°D
Distance of use	Between 0.8m and 3.5m

Table 3: ASUS Xtion PRO depth camera specifications.

1.4   Use Case

This project aims at the manipulation tasks relevant to the Robocup@Home competition [5], i.e. picking up and transporting household objects directly or with containers, opening doors, etc. Figure 4 shows typical objects that the robot need to grasp at the competition. There are five main object groups [1]:

- Known objects: artificial and regular-shaped objects which are provided before competition for training.

- Alike objects: slightly different from known objects (i.e. in color or shapes), also provided before competition.

---

[2]https://www.ipa.fraunhofer.de/

[3]https://schunk.com/de_en/gripping-systems/series/sdh/

[4]https://www.asus.com/3D-Sensor/Xtion_PRO/

[5]http://www.robocupathome.org/

- Containers: objects which can be used for storage, pouring or transport such as trays or bowls.

- Special objects: requires special handling to accomplish a specific task. Examples are door handles, chairs, or walking sticks. Samples of these objects will not be provided before competition.

- Unknown objects: any objects that are unknown beforehand but can be grasped or handled.

Figure 4: Typical objects at the Robocup@Home competition [1].

1.5   Approach

- Define grasp plan representation in 3D.

- Choose a (or several) grasp synthesis strategy to generate grasp candidates.

- Choose a grasp quality metric for rating synthesized grasp candidates.

- Generate synthesis point cloud data with grasp based on defined representation following the Dex-Net 2.0 data generation process [7].

- Train a neural networks model on the synthesized data to calculate a quality measure for possible grasp candidates given an object.

1.6   Expected Results

- At minimum the project should result in a pipeline to generate training data for a model which can rapidly evaluate grasp candidates synthesized for detected objects.

- An extension to the project is implementing a full grasp planner for testing on the Care-o-bot platform, including object detection, grasp synthesis, grasp candidate ranking, and grasp execution.

- A stretch goal of the project is to incorporate task compatibility in the data generation pipeline.

2   Project Plan

2.1   Work Packages

- Literature research

- Choose an approach to represent object-grasp in point clouds and implement generation of this representation using point cloud data.

- Choose a grasp synthesis algorithm and generate synthesized grasps with the implemented object-grasp representation.

- Choose and implement a grasp quality metric for ranking synthesized grasp candidates.

- Synthesize training data for grasp quality model:

  - from a data set of 3D object models, generate point clouds from various view points of objects in various poses,
  - generate grasp candidates using the chosen object-grasp representation method,
  - and calculate grasp quality using the chosen metric.

- Train a model which can calculate grasp quality from grasp candidates synthesized for a detected object.

- Implement a grasp planner on the robot platform to test the trained model.

- Writing report.

2.2   Work Schedule

Name	Work																			
		2018, H1						2018, H2						2019, H1		
		Jan 2018	Feb 2018	Mär 2018	Apr 2018	Mai 2018	Jun 2018	Jul 2018	Aug 2018	Sep 2018	Okt 2018	Nov 2018	Dez 2018	Jan 2019	Feb 20..
Literature research	**54d**														
Research object-grasp representation	14d														
Research grasp synthesis	20d														
Research grasp quality metrics	20d														
Implementation	**108d**														
Generate 3D Object-grasp representation	20d														
Synthesize grasp from object models	20d														
Implement grasp quality calculation	20d														
Synthesize training data	20d														
Train grasp quality model	14d														
(optional) Implement grasp planner	14d														
Writing Report	45d														

References

[1] Loy van Beek et al. *RoboCup@Home 2017: Rules and Regulations.* June 11, 2017. URL: http://www.robocupathome.org/rules/2017_rulebook.pdf.

[2] J. Bohg et al. "Data-Driven Grasp Synthesis: A Survey". In: *IEEE Transactions on Robotics* 30.2 (Apr. 2, 2014), pp. 289–309. ISSN: 1552-3098. DOI: 10.1109/TRO.2013.2289018.

[3] Stephen L. Chiu. "Task Compatibility of Manipulator Postures". In: *The International Journal of Robotics Research* 7.5 (1988), pp. 13–21. DOI: 10.1177/027836498800700502. eprint: https://doi.org/10.1177/027836498800700502. URL: https://doi.org/10.1177/027836498800700502.

[4] Corey Goldfeder and Peter K. Allen. "Data-driven grasping". In: *Autonomous Robots* 31.1 (July 1, 2011), pp. 1–20. ISSN: 1573-7527. DOI: 10.1007/s10514-011-9228-1. URL: https://doi.org/10.1007/s10514-011-9228-1.

[5] Jihong Lee. "A study on the manipulability measures for robot manipulators". In: *Intelligent Robots and Systems, 1997. IROS '97., Proceedings of the 1997 IEEE/RSJ International Conference on.* Vol. 3. Sept. 1997, 1458–1465 vol.3. DOI: 10.1109/IROS.1997.656551.

[6] Jeffrey Mahler et al. "Dex-net 1.0: A cloud-based network of 3d objects for robust grasp planning using a multi-armed bandit model with correlated rewards". In: *IEEE International Conference on Robotics and Automation (ICRA).* IEEE. 2016, pp. 1957–1964.

[7] Jeffrey Mahler et al. "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics". In: *CoRR* abs/1703.09312 (2017). arXiv: 1703.09312. URL: http://arxiv.org/abs/1703.09312.

[8] Máximo A. Roa and Raúl Suárez. "Grasp Quality Measures: Review and Performance". In: *Autonomous Robots* 38.1 (Jan. 1, 2015), pp. 65–88. ISSN: 0929-5593. DOI: 10.1007/s10514-014-9402-3. URL: http://dx.doi.org/10.1007/s10514-014-9402-3.

[9] A. Sahbani, S. El-Khoury, and P. Bidaud. "An Overview of 3D Object Grasp Synthesis Algorithms". In: *Robotics and Autonomous Systems* 60.3 (Mar. 1, 2012), pp. 326–336. ISSN: 0921-8890. DOI: 10.1016/j.robot.2011.07.016. URL: http://dx.doi.org/10.1016/j.robot.2011.07.016.

[10] Raúl Suárez, Jordi Cornella, and Máximo Roa Garzón. *Grasp quality measures.* Institut d'Organització i Control de Sistemes Industrials, 2006.

[11] Tsuneo Yoshikawa. *Foundations of Robotics: Analysis and Control.* Cambridge, MA, USA: MIT Press, 1990. ISBN: 0-262-24028-9.