Grasp It like a Pro: Grasp of Unknown Objects with Robotic Hands based on Skilled Human Expertise

Chiara Gabellieri[1,3], Franco Angelini[1,2,3], Visar Arapi[1], Alessandro Palleschi[1,3], Manuel G. Catalano[2], Giorgio Grioli[2], Lucia Pallottino[1,3], Antonio Bicchi[1,2,3], Matteo Bianchi[1,3] and Manolo Garabini[1,3]

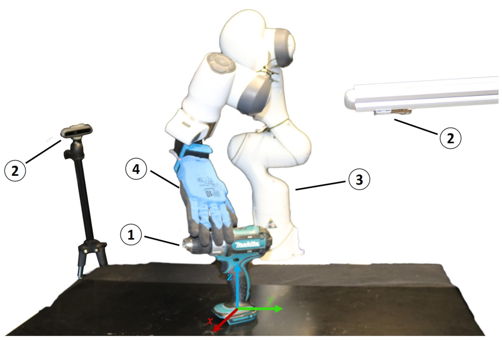


Fig. 1. The experimental setup is composed of: (1) the object to be grasped placed on a table; (2) two Intel RealSense Depth Cameras D415 to sense the object; (3) a Panda manipulator by Franka EMIKA; (4) the Pisa/IIT SoftHand.

***Abstract*— We observed that humans not only are obviously better at grasping with their own hands than robots, but are also when using the same hardware hand as the robot, provided they train for some time. We, therefore, propose that a skilled human user manually operates the robotic hand to grasp a number of elementary objects, consisting of different boxes. A Decision Tree Regressor is trained on the data acquired from the human operator so to generate hand poses able to grasp a general box. This is extended to grasp unknown objects leveraging upon the state of the art Minimum Volume Bounding Box (MVBB) decomposition algorithm that approximates with a number of boxes the shape of an unknown object, based on its point cloud. We tested the proposed approach on a Panda manipulator equipped with a Pisa/IIT SoftHand, achieving a success rate of 86.7% over 105 grasps of 21 different objects.**

I. INTRODUCTION

Grasping of previously unseen objects, an almost obvious task for humans, still represents an open and very challenging problem in robotics.

In the latest years, *data-driven* approaches have received the greatest attention [1]. These methods rely on the quality of the data sets of ground truth grasps, the creation of which is often a time consuming and not trivial work. Using simulated ground truth grasps may facilitate the data set generation, as done, e.g., in [2]. On the other hand, despite the good effort represented by [3], a reliable and computationally efficient grasp simulator is not yet available for a lot of new under-actuated and/or compliant robotic hands. Different approaches have been tried: in [4], the data set is generated from videos of humans grasping objects with their hands. A different idea is that associating each ground truth grasp not to an object but to its *local* shape may allow the creation of slim data sets [5].

To create a database of grasps in a time-efficient way, we propose that a skilled human operator performs experiments using the robotic hand to grasp only a set of cuboids instead of general objects, dramatically reducing the number of trials. This approach is then generalized to grasp unknown objects by relying on state of the art decomposition algorithms (see e.g. [6]) that allow approximating an object with cuboids.

II. GRASP PLANNING ALGORITHM

The procedure starts with the acquisition of the object point cloud (see Fig.2(a)). Once the point cloud is obtained, the MVBB algorithm approximates it with cuboid boxes (see Fig. 2(b)). The approximating boxes are ranked based on a criterion that considers both their distance from the centroid of the point cloud (as done in [7]) and their density in terms of points, to favor both accessibility of a box and the accuracy of the object local approximation it provides, respectively. The first box is the candidate box to be grasped (see Fig. 2(c)) in order to grasp the object. At this stage, the DTR, trained on the data registered from the skilled human operator as described in Fig. 4, predicts a set of possible grasping poses to grasp the candidate box (see Fig. 2(d)). Eventually, a candidate grasp pose is selected among the collision-free ones based on geometric considerations and it is sent to the robot (see Fig. 2(e)). If no collision-free grasp can be selected, the second candidate box is considered, and so on.

III. EXPERIMENTAL VALIDATION

The effectiveness of the method is verified grasping objects from a table using a robotic manipulator to reproduce the end-effector pose planned by the proposed method. The experimental setup is depicted in Fig. 1. In order to validate the method, we tested it on 21 previously unseen objects, shown in Fig. 3(a). Each object is tested 5 times with different position on the table and orientation along the vertical axis, randomly chosen by an operator. A total of

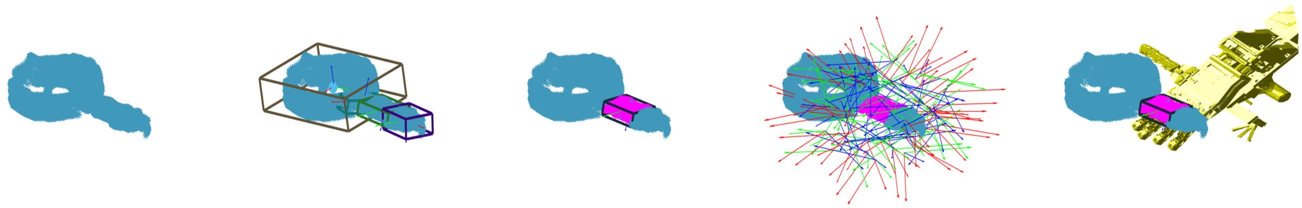
This work was supported in part by the European Unions Horizon 2020 research and innovation program as part of the projects ILIAD (Grant no. 732737), and in part by the Italian Ministry of Education and Research in the framework of the CrossLab project (Departments of Excellence).

[1]Centro di Ricerca "Enrico Piaggio", Università di Pisa, Largo Lucio Lazzarino 1, 56126 Pisa, Italy

[2]Soft Robotics for Human Cooperation and Rehabilitation, Fondazione Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy

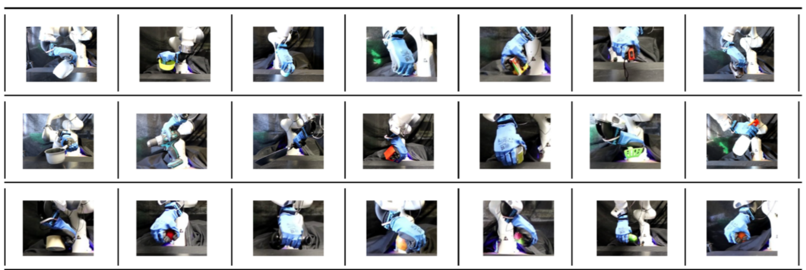
[3]Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Largo Lucio Lazzarino 1, 56126 Pisa, Italy

chiara.gabellieri1@gmail.com

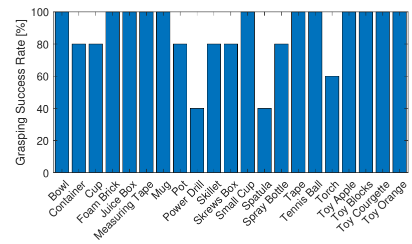


(a) Point Cloud (b) Boxes (N=4) (c) Box Candidate (d) Grasps (e) Grasp Candidate

Fig. 2. Main steps of the grasp planning procedure. The point cloud of the object is acquired (2(a)) and decomposed into bounding boxes (2(b)). A candidate box is selected (2(c)) and a set of poses suitable for grasping it is predicted (2(d)). Finally, a candidate pose is selected (2(e)).



(a) Grasp on the tested objects. From left to right: Bowl, Container, Cup, Foam Brick, Juice Box Measuring Tape, Mug, Pot, Power Drill, Skillet, Skrews Box, Small Cup, Spatula, Spray Bottle, Tape, Tennis Ball, Torch, Toy Apple, Toy Blocks, Toy Courgette, Toy Orange



(b) Grasping success rate for each object. Each item has been tested 5 times.

Fig. 3. Grasped objects on the left, and corresponding results on the right.

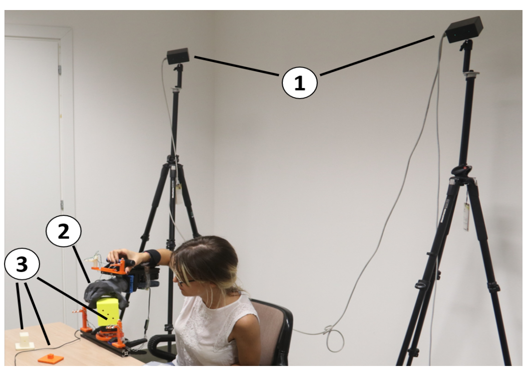


Fig. 4. Acquisition of the data from an expert human user for the database generation using the PhaseSpace cameras (1) to record the grasp poses. A set of 56 boxes (e.g., 3) are grasped with a Pisa/IIt SoftHand (2).

105 grasps have been tested. After closing the fingers, the end-effector of the robot is lifted of 150mm in 5.5s and if the object does not fall the grasp is considered successful.

The results are reported in Fig. 3(b). The average grasping success rate, among all the 105 trials is equal to 86.7%. It emerges that the objects more challenging to grasp are the Spatula and the Power Drill. Probably this is due to the flatness in the first case, which brings the hand fingers to remain stuck against the table while closing, and to the weight in the second case (1190 g). Indeed, heavy objects require more solid grasps to be lifted.

IV. Conclusions

In this work, we proposed and validated a planning algorithm for grasping with robotic hands that encodes the expertise of a skilled user, trained in the robotic hand use. We tested the proposed approach using the Pisa/IIT SoftHand achieving a percentage of grasp success of 86.7% over 21 previously unseen objects. We performed 5 grasps per object for a total of 105 grasps. Future work will focus on the integration of force measurements in the grasp planning as possible indicators of the grasp quality. Validation with different types of robotic hands will be carried out to show and explore the generality of the proposed method.

References

[1] A. Sahbani, S. El-Khoury, and P. Bidaud. An overview of 3d object grasp synthesis algorithms. *Robotics and Autonomous Systems*, 60(3):326–336, 2012.

[2] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg. Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics. *arXiv preprint arXiv:1703.09312*, 2017.

[3] A. Rocchi and K. Hauser. A generic simulator for underactuated compliant hands. In *SIMPAR*, pages 37–42. IEEE, 2016.

[4] C. Della Santina, V. Arapi, G. Averta, F. Damiani, G. Fiore, A. Settimi, M. G. Catalano, D. Bacciu, A. Bicchi, and M. Bianchi. Learning from humans how to grasp: a data-driven architecture for autonomous grasping with anthropomorphic soft hands. *RA-L*, 4(2):1533–1540, 2019.

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

[6] K. Huebner, S. Ruthotto, and D. Kragic. Minimum volume bounding box decomposition for shape approximation in robot grasping. In *ICRA*, pages 1628–1633. IEEE, 2008.

[7] M. Bonilla, D. Resasco, M. Gabiccini, and A. Bicchi. Grasp planning with soft hands using bounding box object decomposition. In *IROS*, pages 518–523. IEEE, 2015.