GRASPA 1.0:
GRASPA is a Robot Arm graSping Performance benchmArk

Fabrizio Bottarel[1,2,*], Giulia Vezzani[1,*], Ugo Pattacini[1], Lorenzo Natale[1]

ABSTRACT

As the state of the art of robotic manipulation and grasping keeps advancing [1], the need for protocols to benchmark proposed approaches grows just as steadily. The use of common object sets has emerged in latest years [2], however no dominant protocols and metrics to test grasp planning algorithms beyond simple success/failure rates have taken root yet. Moreover, reproducibility is often hindered by researchers having access to robot setups that feature different manipulators, grippers and perception tools with different limitations.

In an attempt to address these shortcomings, we propose version 1.0 of GRASPA, a benchmark to fairly test effectiveness of grasp planners on real test robot setups accounting for platform limitations that might affect the performance of the algorithm itself.

The main features of GRASPA are:

- Printable layouts of increasingly challenging grasping scenarios (e.g. Figure 1(b)
- A protocol to assess the robot reachability and the calibration of the vision system over the layouts.
- A grasp quality metric to evaluate grasping poses in simulation.
- A score to assess the grasp stability during the practical execution of the task.
- Possibility to benchmark the algorithm either in isolation or in clutter.
- A composite score to quantify the overall performance of the algorithm accounting for reachability and calibration limits of the test platform.
- Designed to be deployed on real robot setups.

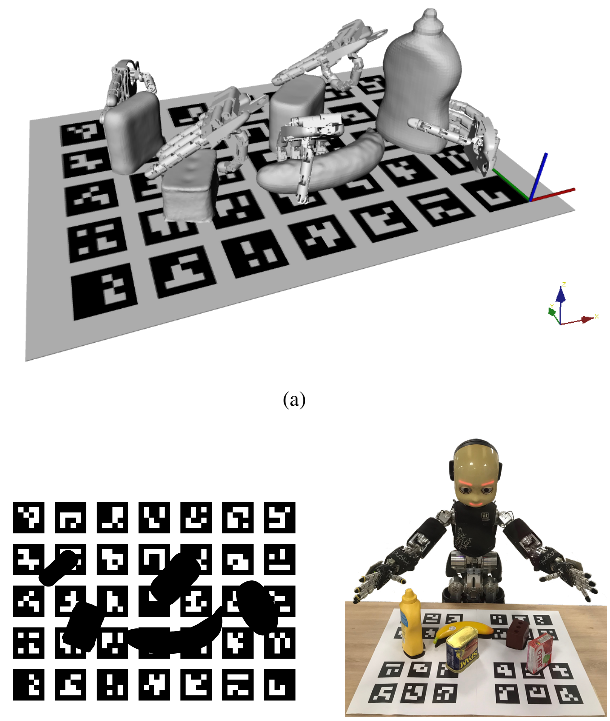
In order to allow for repeatability, the benchmarking protocol comes with predefined grasping scenarios of increasing complexity in terms of number and type of objects. Such scenarios feature objects from the YCB set [2] in fixed 6D poses with respect to the layout reference frame. They also include an ArUco [3] marker board to allow users to print them (Figure 1(b) shows an example) and locate such reference frame in robot camera images. GRASPA 1.0 proposes 3 layout scenarios.

*Equal contribution.

[1]Fabrizio Bottarel, Giulia Vezzani and Lorenzo Natale are with Istituto Italiano di Tecnologia, via San Quirico 19D, Genova, Italy name.surname@iit.it.

[2] Fabrizio Bottarel is also with Università di Genova, Genova, Italy.

This work was supported by the European H2020 project No. 730994 (TERRINet) and ERA-NET CHIST-ERA call 2017 project HEAP.



(a) (b) (c)

Fig. 1. (a) Simulated grasps rendered on one of the GRASPA layouts. (b) The printable layout relative to scenario in (a). (c) Deployment of the scenario on the real test platform.

Given a grasp pose planner, GRASPA is designed to output a score $\bar{S}_L$ for each scenario, $L \in \{1, 2, 3\}$. $\bar{S}_L$ is computed as a combination of scores, briefly outlined in Table I and detailed in the following. The outcome of scores $S0, S1, S2$ only depends upon the specific experimental setup being used, while $S3, S4, S5, S5, S6$ depend on the grasp synthesis and planning algorithm being benchmarked.

Reachability score. GRASPA accounts for grasps generated in poorly reachable regions of the workspace. We divide the layout area in 6 regions and define a set of poses uniformly distributed over the layout area, with different orientations. The user needs to have the robot reach (or attempt to) for these pre-defined poses and then acquire the ones actually reached by querying the forward kinematics. Each region is then scored according to the number of poses reached within a threshold, and each object inherits the score of the region it resides in.

Camera calibration score. This score is analogous to the reachability score, but is meant to evaluate camera calibration

Score	Score name	Meaning
$S0_k^L \in [0,1]$	Reachability score	Accounts for whether the object is located in a region characterized by a good reachability of the robot.
$S1_k^L \in [0,1]$	Camera-calibration score	Accounts for whether the object is located in a region characterized by a good calibration of the vision system.
$S2_k^L \in \{0,1\}$	Graspability score	Accounts for whether the object can be physically grasped and lifted by the robot, considering its shape and weight.
$\bar{S3}_k^L =\in [0,1]$	Grasp quality score	Accounts for how contacts are distributed on the object by simulating grasp closure in simulation and computing the grasp wrench space.
$\bar{S4}_k^L =\in [0,1]$	Binary success score	Accounts for whether the robot actually managed to grasp the object in real tests.
$\bar{S5}_k^L =\in [0,1]$	Grasp stability score	Evaluates the stability of the grasp during the execution of a fixed trajectory.
$\bar{S6}_k^L =\in [0,1]$	Obstacle avoidance score	(Only in cluttered mode) Accounts for how many objects the robot has hit while executing the grasp.
$\bar{S}_L$ where in isolation: $\bar{S}_m^L \in [0,2]$ in the clutter: $\bar{S}_m^L \in [0,3]$	Final per layout score	Combines all the scores in order to evaluate the grasp planner performance taking into account any limitation of the robotic platform used in real world tests.

TABLE I

SUMMARY OF THE BENCHMARK SCORES.

with respect to the layout area. GRASPA requires the user to perform the same reaching procedure, however the resulting pose must be read through the robot cameras (e.g. by affixing a marker to the robot end effector) instead of the forward kinematics.

Graspability score. Each object is assigned a binary score according to whether or not the object can be grasped and lifted by the manipulator and gripper setup. Grasps planned for un-graspable objects are not considered in the final score.

Grasp quality score. Given the end effector kinematics, planned poses are scored in simulation. For scenario L, object m and pose k we obtain sets of contact points $P_{m,k}^L$ resorting to the GraspStudio simulation suite [4]. Grasp quality is evaluated according to the ratio $\frac{\bar{r}(GWS_{k,m})}{r(OWS_m)}$ between the radii of the largest spheres [5] contained in the Grasp Wrench Space (GWS) and Object Wrench Space (OWS) [6].

Binary success and grasp stability scores. Planned grasps are then executed in the real robot setup. Binary success metric $S5$ evaluates whether the object is successfully grasped and lifted without falling for several seconds. $S6$ is evaluated by having the robot move the grasped object along 4 portions of a trajectory, assigning 0.25 for each segment in which the grasp prevents the object from falling. The trajectory consists of rotations around the object center of mass.

Obstacle avoidance score. GRASPA allows grasp planners to be benchmarked with objects in *isolation* (one at a time) or in *clutter* (all objects are present at the same time). In the latter case, if the grasp planner also takes care of collision avoidance during execution it can be accounted for by score $S6$. This is simply obtained by counting the number of objects hit during execution.

Final score. For scenario L, the k-th object is considered to be *reachable* if $S0_k^L > \tau_0$ and $S1_k^L > \tau_1$ and *graspable* if $S2_k^L = 1$. τ_0, τ_1 are user-defined thresholds. Final compound score $\bar{S}_L$ is computed considering the $S3, S4, S5, S6$ of each reachable and graspable object.

To summarize results, benchmark users must report the values of all scores for all objects. This way, readers can break down the results and gain insight on the performance of the tested algorithm and the shortcomings of the test platform. As an example application, we deployed GRASPA on the iCub humanoid robot and benchmarked the grasping pipeline proposed in [7]. For further details and results, please refer to [8].

The code and instructions for automatically computing the benchmark scores is publicly available[1], as well as a Docker container to ease installation and a cloud hosted environment to reproduce our results without requiring any installation.

REFERENCES

[1] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-Driven Grasp Synthesis - A Survey," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 289–309, 2015.

[2] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The YCB object and Model set: Towards common benchmarks for manipulation research," in *2015 International Conference on Advanced Robotics (ICAR)*, pp. 510–517, July 2015.

[3] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.

[4] N. Vahrenkamp, M. Kröhnert, S. Ulbrich, T. Asfour, G. Metta, R. Dillmann, and G. Sandini, "Simox: A Robotics Toolbox for Simulation, Motion and Grasp Planning," in *International Conference on Intelligent Autonomous Systems (IAS)*, pp. 585–594, 2012.

[5] C. Ferrari and J. Canny, "Planning optimal grasps," in *Proceedings 1992 IEEE International Conference on Robotics and Automation*, (Nice, France), May 1992.

[6] C. Borst, M. Fischer, and G. Hirzinger, "Grasp planning: how to choose a suitable task wrench space," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, (New Orleans, LA, USA), pp. 319–325 Vol.1, IEEE, 2004.

[7] P. Nguyen, F. Bottarel, U. Pattacini, M. Hoffmann, L. Natale, and G. Metta, "Merging physical and social interaction for effective human-robot collaboration," *Humanoid Robots (Humanoids)*, 2018.

[8] F. Bottarel, G. Vezzani, U. Pattacini, and L. Natale, "GRASPA 1.0: GRASPA is a robot arm grasping performance benchmark," *submitted to RA-L Special Issue on Benchmarking Protocols for Robotic Manipulation*, 2019. Under review.

[1] `https://github.com/robotology/GRASPA-benchmark`