2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)
Beijing, China, November 6-9, 2018

A Benchmarking Framework for Systematic Evaluation of Compliant Under-actuated Soft End Effectors in an Industrial Context

Panagiotis Sotiropoulos[1]*, Máximo A. Roa[2], Murilo F. Martins[1], Werner Friedl[2], Hussein Mnyusiwalla[1], Pavlos Triantafyllou[1] and Graham Deacon[1]
[1] Robotics Research Team, Ocado Technology, Hatfield, UK
[2] Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany

Abstract— **This paper presents an approach for systematic evaluation of robotic end effectors for an industrial use case that handles delicate, deformable, non-regular objects such as fruits and vegetables. To handle these objects, soft under-actuated hands are the most promising technology so far. However, the approach directions suitable for grasping objects are not usually easy to establish due to the under-actuation effect; therefore, we propose an experimental protocol to serve as framework for data collection, which aims to assess the best directions for grasping using a particular end effector. This protocol, focused on reproducibility and comparability, allows for a better understanding of how a particular hand embodiment influences grasping success for individual products. The protocol can also be used as an effective tool for redesign, and it was applied to evaluate two well-known soft hands (RBO Hand and Pisa/IIT SoftHand) for handling groceries.**

I. Introduction

Benchmarking has received increased attention in the robotics community due to the need of obtaining comparable and reproducible results in different subfields. A benchmark, considered as a standard to compare or assess the performance of a robotic (sub)system, requires well-defined standardized tasks with some quantitative evaluation of the results. Current efforts are focused on experimentally repeatable evaluations, often requiring restricted settings, or on competitions, mostly held once and focused on evaluating general system abilities. Factors that adversely affect the road towards a more scientific evaluation in robotics mainly include the lack of commonly available hardware, common protocols, and common datasets for the evaluations.

Robotic manipulation is one of the major fields in robotics with a very active community working towards standardization and comparability of results. This paper is centred on the evaluation of grasp capabilities, as they are the basis of any manipulation system. From our point of view, benchmarking of robotic grasping can be performed at four different levels:

1) Component level (hand evaluations): includes tests aimed to characterize and evaluate the intrinsic capabilities of the end effector, including its features (e.g. number of degrees of freedom and actuators, number of fingers, weight), and basic capabilities (e.g. finger and grasp strength, finger force tracking). General classifications of end effectors according to these criteria have been previously presented [1], [2], and more recently, a systematic view on evaluation of robotic hands at component level has been proposed by NIST (National Institute of Standards and Technology) [3].
2) Object level (hand-object evaluations): considers tests aiming to verify the capabilities of the hand to grasp a defined set of objects. A typical example of this category is the qualitative verification of the different grasp types that the hand can perform, typically referenced to a taxonomy of human prehension, for instance, Cutkosky's [4] and Feix's [5]. A more detailed characterization of grasping abilities requires the definition of a set of objects. A well-known example is the YCB object and model dataset [6], which has been published and distributed to multiple research groups worldwide; current efforts are focused on defining test protocols to effectively use such datasets.
3) Functional level (hand-arm evaluations): focuses on the evaluation of the joint hand and arm capabilities with limited visual input for providing the object pose, or simply considering a fixed pose of the object with respect to the robotic manipulator. Tests at this level mainly evaluate the planning strategy (either with model-based or model-free approaches) to derive a feasible grasping action for a given object. A functional proof of concept is then provided, which verifies if a given hand mounted on a specific arm is capable of performing certain grasping actions. A typical example was demonstrated during the IROS 2016 and 2017 grasping and manipulation competitions [7], where a set of predefined tasks were to be performed, with a fixed predefined initial and final pose of the objects and with almost no visual input to the system.
4) System level (task evaluations): considers the robotic system as a whole, including the full pipeline of perception, planning, control and possibly even a supervisory system for error correction. Tests at this level are centred on fulfilling a predefined task. Typical examples include the Amazon Picking Challenge [8], the DARPA robotics challenge [9], or even human-centred evaluation protocols applied to robotic grasping, like SHAP (Southampton Hand Assessment Protocol) [10].

This project has received funding from the European Union's H2020 programme under grant agreement No. 645599, project SoMa.
*Corresponding author email: p.sotiropoulos@ocado.com

The robotics community has been actively developing benchmarking approaches at the different levels described above, and more application-driven use cases show up, pushing forward the research in the field. This paper is centred on creating a benchmark protocol for robotic solutions capable of handling products in warehouses. One example of such application occurs at Ocado, the world's largest on-line only supermarket, where over 2 million items are packed every day, comprising over 270,000 customer orders of 50,000 distinct products. The variability of the product range imposes major challenges to robotic manipulation. This problem space calls for a systematic way to evaluate different grippers or hands that could work with the variety of products present in the warehouse.

Object-level evaluations are defined as the target tests, as the most important component for a picking system inside the warehouse is the end effector. Different object-level evaluation approaches have been proposed. The comparison of hand performance can be carried out for specific types of objects using graspability maps [11], which show for a specific object and a given hand the potential locations for the hand that lead to a force-closure grasp on the object. However, the approach is based on simulations, and a more realistic evaluation is desired for the use case. An empirical evaluation of gripper performance can be obtained through hand-in-hand (or hand-in-stick) tests, where a human holds the end effector and has some simple system to command opening and closing actions. Such a test for a picking scenario was used during the IROS 2016 grasping and manipulation competition [7]. One main disadvantage of such tests is that the human copes with limitations of the hand design by using his/her sophisticated sensorial, planning and control system to fulfil the task. In fact, it is well-known that amputees working with simple two-finger prosthesis are capable of doing an amazing number of activities of daily living.

A systematic approach to evaluate grippers at object level is then required, not only for selecting an end effector by comparing performances on a standardized test, but also as a tool for identifying the performance boundaries for each gripper (what kind of objects can the different hands grasp, what are the best approach directions for performing a grasp, and with what degree of robustness and flexibility), and for improving the design and iteratively achieving better behaviour of the grippers for accomplishing the intended task while considering the attributes relevant to the use case.

Although grasp performance intrinsically depends on several additional factors (arm, speed, environment), the focus of this paper is to evaluate the performance capabilities of the under-actuated hand itself. A test protocol and scoring method for systematic evaluation of grippers at object level is presented (Sections III and IV), covering the test procedures, the evaluation criteria, and the aspects to be reported for creating a meaningful benchmark for the manipulation community. As an application example, the protocol is used for evaluating and redesigning two state-of-the-art soft hands (Section V): the RBO Hand [12] and the PISA/IIT hand [13].

II. SOFT-HANDS IN THE USE CASE

Ocado is looking to increase the efficiency of its operation by employing automation for the picking and packing of shopping orders. Even though a large portion of the object range is amenable to handling by standard off-the-shelf robots and gripper technology, typically rigid bodies with regular geometry, there is still a multitude of objects that prove problematic. Fresh fruit and vegetables are good examples of this class of objects that typically do not exhibit a regular shape and/or are variable in shape and/or are deformable. The lack of a CAD model per product and the fact that the individual objects of the same type have variation in shape introduces further challenges on estimating the exact position and shape of the object to be manipulated. In addition, much of Ocado's range of fruit and vegetables comes as groups of items packaged together; often this packaging is made from flexible material (mesh), which constrains objects to be in close proximity to each other, but without exhibiting any easily pre-definable shape. This requires from the robotic hand itself and the grasp planning approach to be able to cope with some level of uncertainty. Soft manipulators are in this sense interesting, as the forgiving nature of a soft grasp can absorb the uncertainty involved. Moreover, even though a firm grasp could possibly increase the robustness during the manipulation phase, the pressure applied on the contact points between the hand and the product should not exceed the limit where the product would be damaged. Again, soft grippers can provide a better pressure distribution that favours a safe handling of the objects.

III. BENCHMARKING FRAMEWORK

This section presents the object-level protocol to test (soft) end effectors within the Ocado use case. The purpose of the proposed protocol is to assess the ability of a robotic end effector to robustly grasp an object with varying end effector approach directions and orientations but relying on minimum sensing. With a systematic evaluation, the protocol can provide insights into the optimum modes of operation of a given end effector.

Identifying the modes of operation for an end effector becomes particularly challenging when dealing with soft hands, where modelling their grasping behaviour can be a difficult problem. A method of computing a preferred grasp direction for grasping with under-actuated hands, the so-called closure signature, has been presented [14]. However, it is limited to top grasps; further investigation of the whole range of approach directions must be performed. Given that under-actuated soft-hands like the RBO Hand and the Pisa/IIT SoftHand cannot control the position of the individual joints, one of the aims of this work is to understand the ways that the environmental constraints (EC), i.e. interactions of the hand, the object and the environment, affect performance while grasping from different spatial directions. The importance of ECs to achieve robust grasps in human and robotic grasping has been discussed in [15].

The task to be executed can be described as *approach, grasp and lift* an object, which correspond to three distinctive phases: Non-contact/Pre-grasp phase, Grasp phase, and Manipulation/Post-grasp phase. A baseline strategy that implements this sequence is:

1) Move on a straight line from the initial pose towards the object,
2) Stop upon contact,
3) Close end effector,
4) Move upwards at a given speed during a certain time,
5) Stop the arm movement and hold the object for a predefined time,
6) Release the object

Note that in principle this strategy aims for power grasps, which are more relevant for soft under-actuated grippers. The application of the protocol to precision grasps would require the hand to stop at a certain distance from the object, as provided by a suitable grasp planner.

The protocol focuses on the grasping and lifting phase. We are interested in investigating the robustness of the grasp in different configurations of the hand during the lifting motion, which provides insights on the optimum hand orientation for the transport phase. The experiment is conducted with a fixed move up speed of 0.01 m/s and a holding time of 10 s. Variable speed and acceleration profiles would need to be considered to further investigate the dynamic nature of the post-grasp phase, where the trajectory of the wrist would be equally important for maintaining the grasp.

A. Setup Description

The object to be tested is located on top of a table, freely accessible from the top and from the sides. This setup allows evaluating the entire range of approach directions towards the object. Additional constraints are expected to be present on a real world scene, including constraints imposed by the storage crate and constraints imposed by neighbour objects in the same crate, yet these cases could be considered as a subset of the freely accessible case.

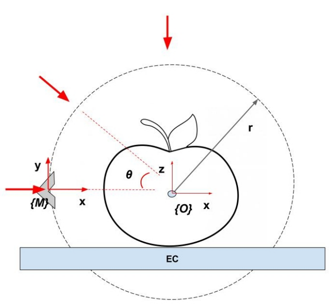

Fig. 1. Pregrasp poses and approach directions with varying angles for approach elevation (AE).

The focus of this protocol is to evaluate the manipulator performance at object-level, therefore the pose $^{R}T_{O}$ of the object coordinate frame $\{O\}$ with respect to the robot base coordinate frame $\{R\}$ can be predefined by the user so that no vision system is required. The initial pre-grasp hand poses are taken at a predefined distance r from the object, as illustrated on Fig. 1, where $\{M\}$ denotes the coordinate frame of the manipulator and θ denotes the elevation angle for the specific approach direction. $\{M\}$ is a virtual frame of reference located at the effective centre of grasp, as indicated by the particular hand manufacturer; it encodes the preferred orientation for executing top grasps. For non-symmetric objects, different azimuth angles have to be considered as well. Therefore the object has to be rotated with respect to the z axis of $\{O\}$, as illustrated in Fig. 2.

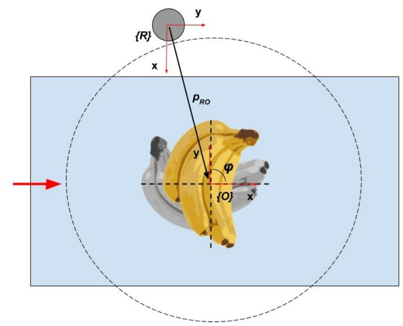

Fig. 2. Definition of object orientation with varying azimuth angle.

Different end effector orientations starting from 0^o to 180^o (Fig. 3) have to be employed to define different pre-grasp poses. However, symmetric orientations can be skipped for a symmetric end effector. An exemplary initial pose for the end-effector is illustrated in Fig. 4, considering a particular approach direction and hand orientation. Note that a reachability analysis has to be performed to ensure that all the intended poses can be reached by the robotic arm for the selected object pose, thus eliminating false negatives due to limited kinematics of the manipulator.

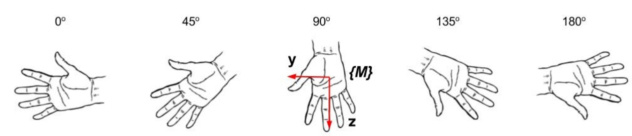

Fig. 3. Different hand orientations (HO) for defining a pregrasp pose.

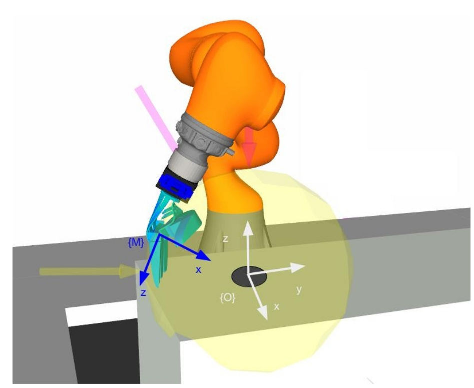

Fig. 4. Exemplary initial pose for the end effector with respect to the object.

B. Procedure and Robot/Hardware Description

The procedure can be executed by any robotic arm and any object. The object should be specified using the key attributes of average weight, size and shape. The system has to be tested on all hand orientations from all possible approach directions (azimuth and elevation). For practical purposes, a discrete set of orientations and directions is utilized, and must be specified as part of the protocol. On the approach phase, the initial contact of the hand with the object can be detected using a F/T (Force/Torque) sensor, which can also be used to generate tracing signals for the grasp evaluation. Grasping from each initial hand pose must be tested on a minimum of 10 attempts, and the average score along with the standard deviation must be reported, as defined in Section IV.

IV. Evaluation

This section is focused on the quantitative evaluation of the experiments with different robotic end effectors using the protocol defined in Section III. The results provide insights on the mode of operation of the end effector and identify its potential weaknesses, thus leading to an informed redesign.

An incremental scoring system is proposed to evaluate the ability of the end effector to grasp the object from different approach directions and hand orientations, and its ability to hold the object while performing a predefined movement, thus evaluating the robustness of the grasp. The maximum score is awarded for a successful attempt, defined when the hand grasps the object and holds it for the whole predefined period of time. Given that the ability of the end-effector to grasp the object is a prerequisite for a robust grasp, half of the total points are awarded upon a successful grasp that keeps the object in grip after it loses contact with the table. The robustness of the grasp is evaluated after this stage, and additional 0.5 points are awarded if the robot retains the object for the remaining of the transport phase. The score s_g for each attempt is then given by:

$$s_g = \begin{cases} 0 & \text{object not grasped} \\ 0.5 & \text{object lost in transit} \\ 1 & \text{grasp cycle completed} \end{cases} \tag{1}$$

The mean and standard deviation of a minimum of 10 attempts should be reported as the final score.

The scoring system is based on a final evaluation of whether the object stays within the hand at the end of the lift-up motion (using the F/T values). A method to measure grasp robustness based for instance on area of contact with the object or simply the number of contacts would require additional sensors on the hands, which are not always available (especially for soft hands) and would restrict the applicability of the protocol. Other scoring alternatives include calculating the maximum force that the hand applies on the object, which requires again a sensorized hand (or alternatively a sensorized object), or the maximum perturbation that the grasp can resist, which could be computed by applying rapid acceleration profiles on the robot while grasping the object. However, the proposed benchmark is centred on evaluating the grasping ability for different end effectors with a protocol easy to replicate, measure and report. As one of the applications of this benchmark is improving the design of end effectors, a detailed description of the common failure cases (and possibly suggestions for modifications) should be reported.

V. Case Study: RBO and Pisa/IIT Soft Hands

This section presents the practical application of the protocol and scoring system defined in Sections IV and III as a framework to evaluate two soft manipulators with very distinct characteristics: the RBO Hand [12] and the Pisa/IIT SoftHand [13]. Due to space constraints, experiments are presented for a single object, to exemplify the evaluation of the protocol and benchmark and to show its influence on the redesign of the end effectors. We deliberately chose to start with an apple, whose spherical shape is representative of a large portion of commonly available fruits and vegetables (e.g. limes, mangoes, tomatoes, oranges, potatoes, etc). Naturally, more objects of interest can be tested to cover a large variety of conditions of the particular use case of interest.

A. Details of experimental setup

Hardware and end effector – The experiments were performed using a 7-DoF KUKA LBR iiwa 14 R820 robot arm. A F/T sensor from Optoforce, model HEX-70-XE-1000N was attached to the robot flange, and each soft hand was fixed to the sensor using a 3D-printed attachment (Fig. 5). Both soft hands were actuated using one degree of actuation, with *fully closed* and *fully open* states. The speed at which the hands were closed was the maximum possible for each hand.

Fig. 5. Soft hands mounted on the KUKA LBR iiwa. Left: RBO hand. Right: Pisa/IIT SoftHand

Object attributes – We chose a single, near-spherical wooden apple as the object for the experiments. The wooden material guarantees no deviation in shape after repeated trials. The object diameter is 7 cm, and weighs 0.150 Kg. Additionally, the friction with the surface closely resembles that of a real apple.

Object location – We defined the position of the object on an IFCO[1] with respect to the robot base via empirical reachability analysis to allow testing of all the desired approach directions. The object was always upright (rotation is irrelevant due to object symmetry).

[1] IFCO is the leading global provider of industry-standard reusable packaging solutions for fresh foods

Force threshold – Following a preliminary set of experiments, the combined force threshold for contact detection in these experiments was set to 2.5 N.

B. Initial evaluation results

Trials – We performed trials with both soft hands for every combination of approach elevation (AE) and hand orientation (HO). For these experiments, 4 AE and 5 HO were used. Each trial was repeated 10 times, totalling 400 trials. The whole data collection procedure took over 26 hours.

Scoring – Applying the benchmark defined in Section IV to evaluate the trials, we obtained the scores (mean and standard deviation of 10 trials) presented in Tables I (RBO Hand) and II (Pisa/IIT SoftHand). The scoring is also shown in Figures 6 and 7 as grey-scale heatmaps that provide a quick visual representation of the hand-object performance; white regions represent the best poses for grasping the objects.

TABLE I

OCADO BENCHMARKING SCORES – RBO HAND

HO\AE	0°	30°	60°	90°
0°	0.0 ± 0.0	0.55 ± 0.14	0.9 ± 0.3	1.0 ± 0.0
45°	0.0 ± 0.0	0.95 ± 0.12	1.0 ± 0.0	1.0 ± 0.0
90°	0.0 ± 0.0	0.0 ± 0.0	1.0 ± 0.0	0.85 ± 0.1
135°	0.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0
180°	0.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0

TABLE II

OCADO BENCHMARKING SCORES – PISA/IIT SOFTHAND

HO\AE	0°	30°	60°	90°
0°	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	1.0 ± 0.0
45°	0.0 ± 0.0	0.0 ± 0.0	0.65 ± 0.45	1.0 ± 0.0
90°	0.0 ± 0.0	0.1 ± 0.3	1.0 ± 0.0	1.0 ± 0.0
135°	0.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0
180°	0.0 ± 0.0	0.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0

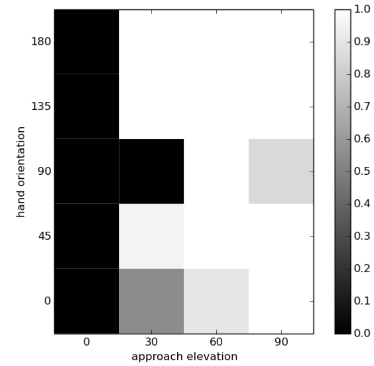

Fig. 6. Scoring heatmap – RBO Hand

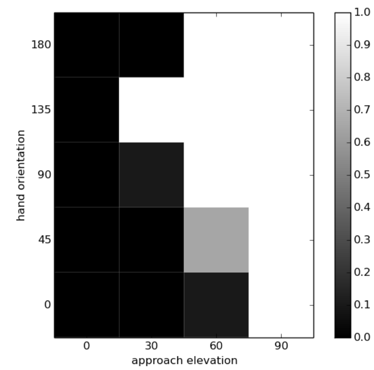

Fig. 7. Scoring heatmap – Pisa/IIT SoftHand

C. Design iteration and retesting

The analysis of the most common modes of failures and the report of the first benchmark tests were the basis for a redesign process for both hands. The analysis following the first evaluation cycle is presented in this section. Redesign criteria are outlined, and the evaluation of the new hand versions is presented.

In the first evaluation cycle we have noticed that even though the RBO Hand was rather successful in grasping the object when approaching from different directions and hand orientations, there were a number of failures on the post-grasping/lift-up phase (9.29% overall). A typical failure occurred on the 90° AE and HO on the heatmap in Fig. 6. On this configuration, vibrations coming from the arm during the lifting phase caused the fingers to lose contact with the object, thus leading to a failure. This failure was mainly due to the compliance of the fingers and their overall length, which would allow only a partial enclosure of the object, limiting the contact area between hand and object. To address these issues, longer and less compliant fingers were selected for the second version of the hand, while maintaining the same palm and thumb configuration. The added length aimed to increase the contact area during grasping, while the added stiffness aimed to increase the robustness of the grasp during the post-grasping phase.

Even though there were no post-grasp failures with the Pisa/IIT SoftHand, the length of the fingers and the limited contact area with the object on certain directions were the main reason behind most of the grasp failures observed. The hand kinematics, optimised for pinch grasps, would often result in the object being pushed away before the thumb could oppose its movement, a behaviour reflected at 45° HO in Fig. 7. A new version of the hand that addressed these issues incorporated longer phalanges to increase the contact area, and included also a compliant palm and updated kinematics to favour power grasps.

Following the same protocol and benchmark, the new versions of the hands were evaluated, and the generated heatmaps are illustrated in Figures 8 and 9 for the RBO Hand and the Pisa/IIT SoftHand, respectively.

D. Analysis of soft hands' performance

While the heatmaps presented in the previous section can provide a qualitative and visual comparison of the hands' performance and provide insights into the preferred modes

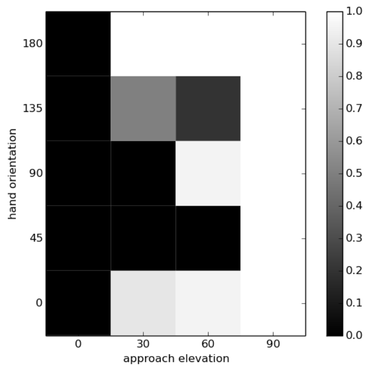

Fig. 8. Scoring heatmap – RBO Hand v2

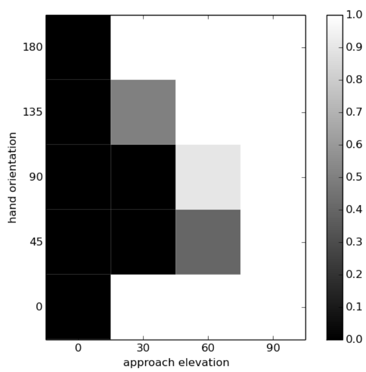

Fig. 9. Scoring heatmap – Pisa/IIT SoftHand v2

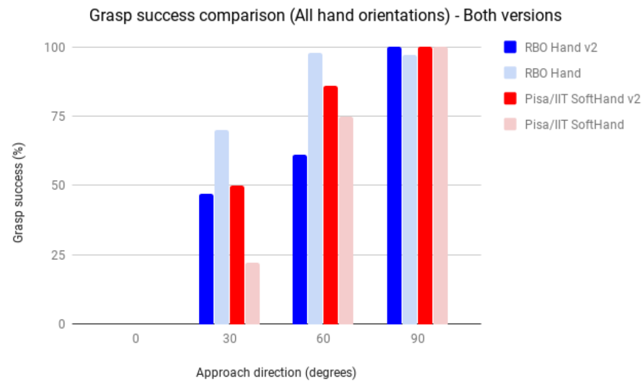

Fig. 10. Results of the proposed benchmark for all the tested hands

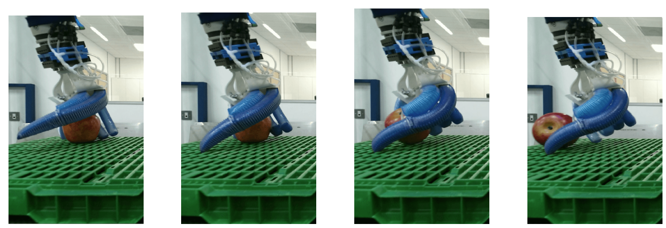

Fig. 11. Sequence of a typical grasp failure in the RBO hand v2

of operation for the hands, a more consolidated score is required. The proposed score is the overall success index across all the AE, as illustrated in Fig. 10.

Comparison across hand versions – Following the evaluation process for the second version of the RBO Hand, an improvement in grasp robustness was observed, with the post-grasp failure rate falling from 9.29% to 3.77%. However, in terms of overall results (Fig. 10), even though the success rate for 90° AE has increased with the new version, the overall success rate, particularly for 30° and 60° AE, has decreased. Given that less compliant actuators were introduced, the inflation sequence has been affected as the thumb and the palm were actuated later, thus causing failures on certain cases (45° and 135° HO and 30° and 60° AE in Fig. 8). An example of such a behaviour is illustrated in Fig. 11, where the fingers actuated earlier push the object away from the hand, before the thumb and palm could actually oppose the movement.

In the case of the Pisa/IIT SoftHand, the overall success rate for 30° and 60° AE has improved drastically from 22% to 50% and from 75% to 86%, respectively, while retaining the 100% success rate on the 90° case (Fig. 10). The grasp robustness has not been affected by the changes, as still no post-grasping failures were detected.

Using the environmental constraint – The results shown in Fig. 10 clearly indicate that top grasps, i.e. AE of 90°, with the approach direction perpendicular to the plane of the environmental constraint provided by the surface of the IFCO, are the most successful across all versions of the hands. However, side grasps, i.e. AE of 0°, with approach direction parallel to the EC plane, failed at every single trial. The fact that the success rate increases with higher elevation angles clearly supports the observation that ECs are of extreme importance for obtaining successful grasps with under-actuated soft hands [15], especially in hands that provide no sensorial information.

Using under-actuated soft hands, it is in general very difficult to create the necessary constraints for a successful grasp by commanding individual fingers to certain positions around the object. In the case of the Pisa/IIT SoftHand, a desired motion for individual fingers is only possible through elaborate motion compensation from the wrist, since all fingers are actuated by a single motor. In the case of the RBO Hand, fingers can be actuated individually, but the nature of the actuation makes it very hard to achieve sufficient accuracy. The role of the EC when grasping with under-actuated hands becomes evident in this case, where the environment can be used to constraint the movement of the object, thus allowing time for the softhand to close around it and securely grasp it. The behaviour is clearly demonstrated in the case of the second version of the RBO Hand, where the synchronization of the fingers and the thumb-palm actuators, which led to a series of failures as discussed before, does not affect grasping success for the 90° AE as the object is constrained between the hand and the environment, thus achieving a 100% success rate for every HO.

All failures at 0° can be explained by the force threshold of 2.5 N. Given the low weight of the object studied and the relatively low friction of the surface, the grasp was never triggered, and both hands ended up sliding the object off the surface. Initial tests with lower values of the threshold for comntact detection and hand actuation still resulted in 0% success, since the only constraint on the movement of the object is the friction with the surface, and the fingers push the object away from the palm before the thumb can come into contact with the object. Moreover, using lower values of the threshold led to a series of false positive triggers, as the threshold was close to the sensor's noise limits and thus created failures in grasps at 30° and 60° AE.

Asymmetric-anthropomorphic hands – Given that both hands are anthropomorphic and thus asymmetric, variations on the performance with respect to the hand orientation were expected. Indeed, all heatmaps show a common pattern of higher success rates towards the 135° HO and lower ones towards the 45° HO. In fact, at 135° both end effectors performed a near pinching grasping, which may have well been the reason for higher scores in such cases. The heatmap representation enables the potential end effector user to make these observations in a straightforward manner, while the systematic evaluation of the hand can identify modes of operation that would otherwise seem counter-intuitive, like the case of 135° HO for an anthropomorphic hand.

Note that even though the protocol was applied here to evaluate anthropomorphic hands, it is flexible enough to be applied to non-anthropomorphic hands and grippers.

VI. CONCLUSIONS

In this paper we have proposed a novel benchmark, including its protocol and scoring system (summarised in the Appendix) to systematically evaluate robotic end effectors, especially under-actuated hands. The protocol is general enough to be applied to any type of end effector. The application of the protocol was exemplified for the study of performance and ulterior redesign of two state-of-the-art soft hands, the RBO Hand and the Pisa/IIT SoftHand. The most effective regions with respect to the approach direction and hand orientation are identified for each hand. Unsuccessful attempts were discussed, and redesign parameters emerged and were used to improve the initial hand designs. The identification of the optimum way to use a particular end effector for a certain class of objects could significantly improve grasp success and robustness for autonomous grasping and manipulation strategies.

The presentation of the results using a heatmap provides an intuitive and immediate way of identifying the effective regions of approach for every hand, as well as facilitating comparison of different end effectors. The heatmap can also be used as input for a grasp planner, as it provides a rough estimate of the best regions (defined by the AE and HO) for approaching a desired object.

As a future work, the proposed protocol will be extended towards a system-level evaluation based on a pick and place task. This is tied with further investigation of the boundaries between success and failure, in order to identify which component (e.g., perception, end effector design, grasping strategy) has the largest influence on the overall system performance.

APPENDIX

The protocol and scoring system for the Ocado Benchmark using the YCB templates are summarized in the attached tables.

ACKNOWLEDGMENT

The authors would like to thank the H2020 SoMa consortium for the continued support, as well as the Robotics Research Team at Ocado Technology for insightful discussions.

REFERENCES

[1] D. Alba, M. Armada, and R. Ponticelli, "An introductory revision to humanoid robot hands," in *Proc. Int. Conf. on Climbing and Walking Robots - CLAWAR*, 2005, pp. 701–712.

[2] J. Bellter and A. Dollar, "Performance characteristics of anthropomorphic prosthetic hands," in *Proc. IEEE Int. Conf. on Rehabilitation Robotics*, 2011, pp. 921–927.

[3] J. Falco, K. V. Wyk, S. Liu, and S. Carpin, "Grasping the performance: Facilitating replicable performance measures via benchmarking and standardized methodologies," *IEEE Robotics and Automation Magazine*, vol. 22, no. 4, pp. 125–136, 2015.

[4] M. Cutkosky and R. Howe, "Human grasp choice and robotic grasp analysis," in *Dextrous Robot Hands*, S. Venkataraman and T. Iberall, Eds. Springer-Verlag, 1990, pp. 5–31.

[5] T. Feix, R. Pawlik, H. B. Schmiedmayer, J. Romero, and D. Kragic, "A comprehensive grasp taxonomy," in *Robotics Science and Systems - RSS, Workshop on Understanding the Humand Hand for Advancing Robotic Manipulation*, 2009.

[6] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, "Benchmarking in manipulation research: Using the Yale-CMU-Berkeley object and model set," *IEEE Robotics and Automation Magazine*, vol. 22, no. 3, pp. 36–52, 2015.

[7] Y. Sun, J. Falco, N. Cheng, H. Chosi, E. Engeberg, N. Pollard, M. Roa, and Z. Xia, "Robotic Grasping and Manipulation competition: task pool," in *Robotic grasping and manipulation*, Y. Sun and J. Falco, Eds. Springer-Verlag, 2018, pp. 1–18.

[8] N. Correll, K. Bekris, D. Berenson, O. Brock, A. Causo, K. Hauser, K. Okada, A. Rodriguez, J. Romano, and P. Wurman, "Analysis and observations from the first Amazon Picking Challenge," *IEEE Trans. Automation Science and Engineering*, vol. 15, no. 1, pp. 172–188, 2018.

[9] "DARPA Robotics Challenge 2015." [Online]. Available: http://archive.darpa.mil/roboticschallenge/

[10] "SHAP: Southhampton Hand Assessment Procedure." [Online]. Available: http://www.shap.ecs.soton.ac.uk/index.php

[11] M. A. Roa, K. Hertkorn, F. Zacharias, C. Borst, and G. Hirzinger, "Graspability map: A tool for evaluating grasp capabilities," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems - IROS*, 2011, pp. 1768–1774.

[12] R. Deimel and O. Brock, "A novel type of compliant and underactuated robotic hand for dexterous grasping," *Int. J. Robotics Research*, vol. 35, no. 1-3, pp. 161–185, 2016.

[13] M. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, and A. Bicchi, "Adaptive synergies for the design and control of the Pisa/IIT softhand," *Int. J. Robotics Research*, vol. 33, no. 5, pp. 768–782, 2014.

[14] M. Pozzi, G. Salvietti, J. Bimbo, M. Malvezzi, and D. Prattichizzo, "The closure signature: A functional approach to model underactuated compliant robotic hands," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2206–2213, 2018.

[15] C. Eppner, R. Deimel, J. Alvarez-Ruiz, M. Maertens, and O. Brock, "Exploitation of environmental constraints in human and robotic grasping," *Int. J. Robotics Research*, vol. 34, no. 7, pp. 1021–1038, 2015.

Ocado End effector Assessment Protocol

Reference No / Version Authors Institution Contact information	P-OMA-0.1 Panagiotis Sotiropoulos, Murilo F. Martins, Máximo A. Roa Ocado Technology / German Aerospace Center (DLR) p.sotiropoulos@ocado.com
Purpose	Assess robustness and reliability of manipulators for grasping deformable and delicate objects with specific characteristics, as defined by the Ocado use case.
Task Description	Systematically grasp objects varying approach elevation (AE) and hand orientation (HO).
Setup Description	**List of objects and their key attributes:**
	Description of the manipulation environment:
	Pose of the object:
Robot/Hardware Description	**Targeted robots/hardware:**
	Prior information provided to the robot: • ${}^{R}T_{O}$: the pose of object O relative to robot base coordinate frame $\{R\}$ • r: initial distance of end effector reference frame $\{M\}$ from object reference frame $\{O\}$
Procedure	**Generating the initial poses:** • Approach direction elevation angles to be tested: 0^o, 30^o, 60^o and 90^o; with 0^o being the direction where the manipulator is moving parallel to the table and 90^o the direction where the manipulator is moving perpendicular to the table towards the object. • The azimuth angle of approach is controlled by the object orientation. Object orientations to be tested with respect to the initial one: 0^o, 45^o, 90^o, 135^o and 180^o. The rotation is about the +z axis of the object coordinate frame $\{O\}$. For symmetric objects certain orientations may be skipped. • Manipulator orientations to be tested: 0^o, 45^o, 90^o, 135^o and 180^o. The rotation is about the +x axis of the manipulator coordinate frame $\{M\}$. For symmetric manipulators certain orientations may be skipped. **For every object to be tested:** 1) Generate the initial poses for the manipulator based on the approach direction elevation angle θ and the distance r. The distance may vary depending on the object size and the hardware setup. 2) A reachability analysis must be performed to define the pose of the object with respect to the robot base so that all initial poses are feasible for the robot-manipulator system. For non-symmetric objects, make sure that one of the principal axes of the object is aligned with the projection of the vector p_{RO} to the object coordinate frame x-y plane. 3) Place the object on top of your table on the selected pose, and drive the manipulator to the initial pose for a given hand and object orientation. 4) Perform a minimum of ten attempts for each configuration using the baseline strategy. 5) Place the object back to its initial pose and repeat the same procedure for all the remaining combinations of approach direction elevation angle, object orientation azimuth angle and hand orientations. **Baseline strategy:** • On the pre-grasp stage, the approach is performed with respect to the +x axis of the manipulator coordinate frame $\{M\}$, commanding the robot on Cartesian velocity mode with a speed of 0.01m/s. • The sensor threshold has to be set and reported. • For any hand orientation, if the object is displaced by a distance equal to its size along its principal axis from the initial object pose, the attempt is considered a failure. Stop the robot and move to the next trial. • On the post-grasping phase the move up motion is executed with respect to the +z axis of the Object coordinate frame $\{O\}$ at a speed of 0.01m/s. • The duration of the move up motion is 10 sec.
Execution Constraints	When placing the object make sure that the selected pose is a stable one and that the object does not move or roll by itself.

Ocado End effector Assessment Benchmark

Reference No / Version Authors Institution Contact information	B-OMA-0.1 Panagiotis Sotiropoulos, Murilo F. Martins, Máximo A. Roa Ocado Technology / German Aerospace Center (DLR) p.sotiropoulos@ocado.com
Adopted Protocol	Ocado End effector Assessment Protocol (P-OMA-0.1)
Scoring	• Please report the results using as a template Table I or II • Award points for each attempt according to the description on Section IV.
To submit	• Scoring tables per hand and object orientation • One heatmap per scoring table • Detailed description of the common failure cases • Suggestions for end-effector modifications should be reported if appropriate • Report the force threshold used to detect contact • Affordances that could not be described using the key attributes defined and that affect grasp success should be explicitly mentioned.