

# Guest Editorial

## Open Discussion of Robot Grasping Benchmarks, Protocols and Metrics

Automated grasping has a long history of research that is increasing due to interest from industry. One Grand Challenge for robotics is Universal Picking: the ability to robustly grasp a broad variety of objects in diverse environments for applications from warehouses to assembly lines to homes. Although many researchers now openly share code and data, it is challenging to compare and/or reproduce experimental results to identify which aspects of which approaches work best due to variations in assumptions and experimental protocols, e.g., sensors, lighting, robot arms, grippers, and objects.

In computer vision, the emergence of specific reproducible benchmarks advanced the field considerably and provide “gradients” that expose gaps in the state of the art. With physical experiments, however, performance depends crucially on the hardware and environment and it is not possible for every lab to experiment with the exact same conditions. Nor is exact uniformity desirable as methods must ultimately work across various sensors, robot arms, grippers, object sets, and environments.

Industrial practitioners characterize picking in terms of the three R’s: rate, reliability, and range (class of objects). One metric for comparison is Mean Picks Per Hour (MPPH), which is common in the logistics industry where it is recognized that human workers can operate in the range of 400-600 MPPH for warehousing operations [1]\*. This can be formalized as:

$\mathbb{E}[\rho]$  : Mean Picks Per Hour

$\mathbb{E}(\rho) = \nu * q$  (computed as mean over  $N$  grasp attempts)

where  $\nu$  is the mean grasp rate, or average number of attempts per hour, and  $q$  is the mean grasp reliability, or probability that each grasp attempt is successful (also known as the “success rate”). The mean grasp rate can be formalized as:

$$\nu = 1/(t_s + t_c + t_r)$$

where  $t_s$ ,  $t_c$ , and  $t_r$  are the average times for sensing, computation, and robot motion, respectively, in fractions of an hour. In some cases sensing, computation, and motion may occur in parallel.

Performance may vary based on the experimental conditions such as the robot arm used (cobot vs. industrial arm) or application considered (e.g., picking any object vs. picking a specific object). To aid in comparison, we recommend reporting the variables above with the number of trials and the experimental conditions below, including tolerances and standard errors where applicable:

**Procedure.** Results should differentiate between training and test procedures and describe several factors. First is the

environment configuration, which includes the workspace area for objects and the set of objects that will not be grasped such as bins, shelves, and tables. Another is the calibration sub-procedure, such as registering sensors to the robot coordinate frame. A third important factor is the sub-procedure for sampling initial object configurations, which may be random (e.g., placing objects in a container, shaking, and emptying into a workspace), or structured (e.g. packed boxes, items on shelves). An additional consideration is the sub-procedure for evaluating success and taking auxiliary measurements (e.g. which object was targeted), which may involve human responses, sensor readings, or machine learning. It is also important to describe any additional training procedures used for machine learning or to set environment parameters (e.g., location and distance between bins, shelves, and tables).

**Object Range.** Object characteristics such as shape, size, materials (e.g. rigid versus deformable), friction, mass distribution, and reflectance properties can reflect performance. The unique sets of objects used for training and testing should be reported with an image displaying each object without occlusions and either a reference CAD model for 3D printed objects or purchase information for other objects. We recommend using objects from published object datasets such as YCB [2], and when objects are left out or new objects are introduced, authors should provide an explanation for the object choice.

**Success Metrics.** The success metric used to evaluate performance should be described in detail. Example success criteria might include whether any object is dropped into a target bin or whether a specific object is placed in a desired pose. To evaluate MPPH, it is also important to report the runtime in seconds for each grasp attempt in terms of sensing, computation, and movement.

In addition, data on MPPH can include raw data collected from experiments such as videos, sensor readings, and sequences of robot arm joint angles to aid in replicability. Additional metrics that are useful to report are: the identity of targeted objects, confidence values from the grasp planner, and precision-recall metrics based on confidence values.

**Computer System.** The type and number of computer processors used at runtime, including both CPUs and GPUs, should be reported. It is also helpful to report the operating system and other specifications of the computer (e.g., network interface card bandwidth, USB bandwidth).

**Robot Arm.** Include details on the kinematics, workspace, payload, speed, accuracy/repeatability, software, price, and whether the manipulator is mobile or fixed. The robot arm manufacturer, model, and version should also be reported.

**Robot Grippers / Hands.** Include details on the number of finger contacts, actuation method (e.g., electric, pneumatic, etc.), control scheme (e.g., position, force, hybrid), kinematics, workspace, contact surfaces (e.g., area, friction, stiffness, materials, etc.), payload, speed, and repeatability.

**Sensors.** Include details on the modality (e.g., color vs. RGB-D), lenses, resolution, field of view, light sensitivity, and noise levels. The camera manufacturer, model, and version should be reported along with the camera configuration categorized by:

- *Single Fixed Viewpoint:* The system receives a color and depth images from one externally mounted fixed RGB-D sensor (overhead or over-the-shoulder).
- *Multiple Fused Viewpoints:* The system takes input from several fixed RGB-D sensors and combines them into one 3D estimate.
- *Moving Viewpoint:* The system takes input from one or more moving sensors on board a mobile manipulator or eye-in-hand sensor.

Also, some grippers have force, tactile, and haptic sensors. The manufacturer, model, version, coordinate frame of all such sensors relative to the robot, and their method of feedback to the control scheme should be reported.

**Controller.** Include details on control methods for the robot arm and gripper. Example controllers include cartesian position, impedance, hybrid position/force, or custom methods.

**Lighting.** Due to the effect on sensing, it is useful to report the following lighting conditions: lighting levels (lux), structured vs. ambient, and indoor vs. outdoor, including combinations.

**Application.** Performance may be categorized based on the target application, such as warehouse picking, pick and place, industrial kitting, stowing, assembly, and assistive care tasks such as instrumental activities of daily living (IADL).

**End.** Other factors may also be important. We welcome others in the community to join the online discussion, which began in Nov 2017, via email or the collaborative document:

<http://goo.gl/6M5rfw>.

We thank our colleagues who also provided input on the discussion: Sergey Levine, Sidd Srinivasa, Matt Mason, Howie Choset, Russ Tedrake, Vincent Vanhoucke, Raia Hadsell, Kurt Konolige, Tom Fuhlbrigge, Tye Brady, Juan Aparicio Ojea, and Peter Puchwein.

\* “A human is capable of performing .. at a rate of approx 400 [picks] per hour with minimal errors, while the best robot in the first APC achieved a rate of approx 30 [picks] per hour with [an 84% success rate].” [3]

#### REFERENCES

- [1] W. Lu, D. McFarlane, V. Giannikas, and Q. Zhang, “An algorithm for dynamic order-picking in warehouse operations,” *European Journal of Operational Research*, vol. 248, no. 1, pp. 107–122, 2016.
- [2] B. Calli, A. Singh, J. Bruce, A. Walsman, K. Konolige, S. Srinivasa, P. Abbeel, and A. M. Dollar, “Yale-CMU-berkeley dataset for robotic manipulation research,” *Int. J. Robotics Research (IJRR)*, vol. 36, no. 3, pp. 261–268, 2017.
- [3] N. Correll, K. E. Bekris, D. Berenson, O. Brock, A. Causo, K. Hauser, K. Okada, A. Rodriguez, J. M. Romano, and P. R. Wurman, “Analysis and observations from the first Amazon picking challenge,” *IEEE Trans. Automation Science and Engineering (T-ASE)*, vol. 15, no. 1, pp. 172–188, 2018.

JEFFREY MAHLER  
University of California, Berkeley  
Department of Electrical Engineering and Computer Sciences  
jmahler@berkeley.edu

ROB PLATT  
Northeastern University  
College of Computer Science  
platt@ccs.neu.edu

ALBERTO RODRIGUEZ  
Massachusetts Institute of Technology  
Department of Mechanical Engineering  
albertor@mit.edu

MATEI CIOCARLIE  
Columbia University  
Department of Mechanical Engineering  
matei.ciocarlie@columbia.edu

AARON DOLLAR  
Yale University  
Department of Mechanical Engineering and Materials Science  
aaron.dollar@yale.edu

RENAUD DETRY  
NASA Jet Propulsion Laboratory  
Mobility and Robotics Section  
Renaud.J.Detry@jpl.nasa.gov

MAXIMO A. ROA  
German Aerospace Center (DLR)  
Institute of Robotics and Mechatronics  
Maximo.Roa@dlr.de

HOLLY YANCO AND ADAM NORTON  
University of Massachusetts, Lowell  
Department of Computer Science  
holly@cs.uml.edu, Adam\_Norton@uml.edu

JOE FALCO, KARL VAN WYCK, AND ELENA MESSINA  
National Institute of Standards and Technology  
{joseph.falco, karl.vanwyck, elena.messina}@nist.gov

JURGEN ‘JUXI’ LEITNER  
Queensland University of Technology  
Australian Centre for Robotic Vision (ACRV)  
j.leitner@qut.edu.au

OLIVER BROCK  
Technical University of Berlin  
Department of Computer Engineering and Microelectronics  
oliver.brock@tu-berlin.de

LAEL ODHNER  
RightHand Robotics  
lael@righthandrobotics.com

ANDREY KURENKOV  
Stanford University and University of California, Berkeley  
Department of Computer Science  
andreyk@stanford.edu

MATTHEW MATL  
University of California, Berkeley  
Department of Electrical Engineering and Computer Sciences  
mmatl@berkeley.edu

KEN GOLDBERG  
University of California, Berkeley  
Department of Industrial Engineering and Operations Research  
goldberg@berkeley.edu