# G3DB: A Database of Successful and Failed Grasps with RGB-D Images, Point Clouds, Mesh Models and Gripper Parameters

Ashley Kleinhans Benjamin Rosman Michael Michalik Bryan Tripp Renaud Detry

## I. INTRODUCTION

Autonomous grasping is a problem that receives continuous attention from our community because it is both key to many applications, and difficult to solve. The complexity of robot grasping is counter-intuitive. For us humans, planning a grasp is a trivial task that requires neither considerable effort nor time to solve, it is difficult to imagine the tremendous challenges that our brain has to overcome in order to allow us to interact with objects with such ease. Firstly, much information is missing. The weight, mass distribution, or friction of an object are impossible to measure prior to manipulating an object, and these properties have a dramatic impact on the behavior of a grasp. We thus have to infer from past experience or common sense the most likely values for them, and adapt to actual object properties during the grasp. Another key piece of information that is missing is the object's 3D shape. Humans and mobile robots perceive the world from a single viewpoint, making the back and often the sides of an object inaccessible through our senses. We are thus forced to consider grasps for which at least one finger comes in contact with a part of the object that we cannot perceive. Further, the faces of an object that we do perceive are sampled through noisy sensors yielding unreliable depth or color. To complete the list of challenges associated with grasping, we note that grasps are parametrized in a high-dimensional space: six parameters to fix the position and orientation of the gripper, plus the parameters that define the shape of the hand – 25 for a human hand, 4 in the case of the Barrett hand considered in this paper. The space to explore to plan a grasp is high dimensional, and, given the mechanical complexity of grasping, good solutions are sparsely distributed.

In traditional grasp planning, we engineer a mapping from vision data to action parameters. Given the difficulties listed above, designing such a program by hand is a tedious task. Instead, research on grasping is now turning towards learning grasp planners from grasp examples, yielding a direct mapping from vision to grasp parameters. A learning-based grasp planner does not necessitate the operator to input hundreds of rules defining how to plan a grasp based on the shape of an object; it instead extracts those rules automatically from grasp examples. A common requirement of grasp learning algorithms is thus a database of grasp examples, representative of a variety of objects and relative hand-object configurations.

A database of grasp examples can be created either with a real robot, or in simulation. When using a real robot, grasp examples are generated either via teleoperation, kinesthetic teaching, human imitation, or autonomous exploration. In either case, each grasp takes at least a dozen seconds to execute. The process can be sped up by running the grasps in simulation, although the process of collecting or creating objects and hand models expensive. In either case, creating a grasp database is an expensive task. Fortunately, a set of grasp examples can often serve to train or test different grasp planners, allowing research groups to share grasp data. The Columbia Grasp Database is one example of such an effort [5]. Other groups have published object databases [1], [2], [9], [8], [11] that can be used in combination with a simulator to produce grasp data.

We are currently constructing a database of robot grasps, with a focus on providing rich perceptual information characterizing each grasp. Our aim is to create a database that can be readily used with limited resources, to train/test a variety of grasping algorithms.

Grasp algorithms come in all sorts of colors. Some learn a mapping from photometric images to grasp parameters, others rely on depth data or a 3D mesh. Certain algorithms make very few assumptions on the dynamics of grasping or object shape, and require a vast number of examples to build a workable model. Others focus on a smaller family of objects or object parts and are designed to work with a small training set. Some algorithms learn the position of the gripper and let the robot compute orientation and preshape parameters online from perceptual data, while other work attempts to model both the robot's wrist pose and the finger preshape.

In order to reach a large group of researchers, we are constructing a setup that can provide, for each grasp, the following data, both before and after a grasp attempt: (1) the pose of the hand, (2) the kinematic configuration of the fingers, as well as the 3D poses of their links, (3) a mesh model of the object, and its pose, (4) one RGB-D image and four point clouds; three are placed at $\frac{2\pi}{3}$ radians facing the object and one taken from the viewpoint of the gripper. We

*This work was supported by the Belgian National Fund for Scientific Research. The Council for Scientific and Industrial Research, South Africa. The DAAD-NRF Scholarship, South Africa. The University of Johannesburg, South Africa.

Ashley Kleinhans and Benjamin Rosman are with the Council of Scientific and Industrial Research, South Africa akleinhans@csir.co.za, brosman@csir.co.za

Bryan Tripp and Michael Michalik are with the University of Waterloo, Canada bptripp@uwaterloo.ca, mjmichal@uwaterloo.ca

Renaud Detry is with the University of Liège, Belgium renaud.detry@ulg.ac.be

aim to collect grasps of at least a hundred household objects, with hundreds of grasps per object.

Executing thousands of grasps on a real robot is a prohibitive task. Yet, building a database exclusively in a simulator is not ideal either, as the correlation between the outcome of simulated grasps and grasps executed on a real robot is limited [3], [7], [4]. Our strategy is to test a large number of grasps in a simulator, and to validate a fraction of those on a real robot platform, equipped with the same hand as the simulator. Validation will be performed by making the robot execute grasps that are as close as possible to a subset of the grasps tested in simulation. Evaluation of grasp success will be similar to that of the simulator. Validation data will, for instance, serve to compute the correlation between simulated and real-world outcomes, allowing users to associate credibility to the simulated data.

Our database will open several research avenues. Highly generalizable grasp models will be trained on the real data exclusively. Models that require a very large training set will have the choice of including simulated grasps in the training set, using the validation data to weight the applicability of the model. Another avenue of interest is to evaluate whether the success of the simulator in predicting real-world outcomes depends on object features, and to learn what sorts of objects (and in what pose) lead to good or bad simulations.

To simulate grasps, we opted for the VREP simulator. VREP is free for research and education, and many of its components are open-source. The advantage of VREP over other simulators such as Gazebo or GraspIt! are its robustness, ease of use, and stability. While we aim to provide a database that is accessible to users who do not wish or do not have the means of running a simulator, we will provide interested users with the ability to easily reproduce our results, and possibly complement them with, for instance, additional camera views, or different grasp success tests, with minimal effort. VREP is an ideal code and data sharing platform. It runs on Linux, Mac, and Windows, and it offers bindings to a number of different languages including Matlab.

We will make the database freely available to the community as soon as it is ready, and provide means of extending the database to interested parties.

## II. NEUROPHYSIOLOGY OF GRASPING

We plan to use this database to support applications of machine learning and models of neurophysiological grasp control. Machine learning methods (e.g. decision trees, convolutional networks, etc.) are typically more effective with greater numbers of training examples. In particular, the difference between performance on test vs. training data (overfitting) becomes smaller with more training data [10]. A larger training dataset therefore allows one to train a more sophisticated system without overfitting. Using simulations, we hope that it will be practical to obtain $10^5$ or more labelled grasps, and that this will support learning of sophisticated grasp controllers that generalize well to new objects.

We are interested in modelling the neural systems that control hand shape in primate grasping, particularly the anterior intraparietal area (AIP), frontal area F5, and related areas. Such a model would draw primarily from recordings of neural spikes in these areas, in different conditions of viewing and grasping objects. A large database of object shapes, hand postures, and grasp outcomes will provide further constraints that complement the neuron recordings. We hope that this will lead to more realistic and informative neurophysiological models.

We note that Kappler et al. [6] are currently working on a similar database.


## III. ACKNOWLEDGEMENT

We thank Marc Freese and the V-REP team for their help setting up the simulation environment.


***Index Terms*— Grasping, Database, Manipulation, Neurophysiology**

## REFERENCES


[1] The kit object database: `http://i61p109.ira.uka.de/ObjectModelsWebUI/`.

[2] The princeton shape benchmark: `http://shape.cs.princeton.edu/benchmark/`.

[3] Y. Bekiroglu, K. Huebner, and D. Kragic. Integrating grasp planning with online stability assessment using tactile sensing. In *IEEE International Conference on Robotics and Automation*, 2011.

[4] A. K. Goins, R. Carpenter, W.-K. Wong, and R. Balasubramanian. Evaluating the efficacy of grasp metrics for utilization in a gaussian process-based grasp predictor. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014.

[5] C. Goldfeder, M. Ciocarlie, H. Dang, and P. Allen. The Columbia grasp database. In *IEEE International Conference on Robotics and Automation*, 2009.

[6] D. Kappler, J. Bohg, and S. Schaal. Leveraging big data for grasp planning. In *IEEE International Conference on Robotics and Automation*, 2015.

[7] J. Kim, K. Iwamoto, J. J. Kuffner, Y. Ota, and N. S. Pollard. Physically based grasp quality evaluation under pose uncertainty. *IEEE Transactions on Robotics*, 2013.

[8] G. W. Kootstra, M. Popovic, J. A. Jørgensen, K. Kuklinski, K. Miatliuk, D. Kragic, N. Krüger, et al. Enabling grasping of unknown objects through a synergistic use of edge and surface information. *The international journal of robotics research*, 2012.

[9] A. Singh, J. Sha, K. S. Narayan, T. Achim, and P. Abbeel. Bigbird: A large-scale 3d database of object instances. In *IEEE International Conference on Robotics and Automation*, 2014.

[10] V. N. Vapnik. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5):988–999, 1999.

[11] W. Wohlkinger, A. Aldoma, R. B. Rusu, and M. Vincze. 3dnet: Large-scale object class recognition from cad models. In *IEEE International Conference on Robotics and Automation*, 2012.