

# Visual Detection of Opportunities to Exploit Contact in Grasping Using Contextual Multi-Armed Bandits

Clemens Eppner

Oliver Brock

**Abstract**—Environment-constrained grasping exploits beneficial interactions between hand, object, and environment to increase grasp success. Instead of focusing on the final static relationship between hand posture and object pose, this view of grasping emphasizes the need and the opportunity to select the most appropriate, contact-rich grasping motion, leading up to a final static grasp configuration. This view changes the nature of the underlying planning problem: Instead of planning for static contact points, we need to decide which environmental constraint (EC) to use during the grasping motion. We propose a method to make these decisions based on depth measurements so as to generate robust grasps for a large variety of objects. Our planner exploits the advantages of a soft robot hand and learns a hand-specific classifier for edge-, surface-, and wall-grasps, each exploiting a different EC. Additionally, we show how the model can continuously be improved in a contextual multi-armed bandit setting without an explicit training and test phase, enabling the continuous improvement of a robot’s grasping skills throughout life time.

## I. INTRODUCTION

Humans routinely exploit environmental contact during grasping and manipulation: picking up a coin by sliding it to the edge of a table, putting a key in a lock, or getting peanuts from a bowl. Contact with the environment is beneficial because it reduces the uncertainty present in free-space positioning and creates easily measurable feedback events [1]. The importance of this principle is reflected in the recent popularity of compliant, under-actuated robotic hands [2]–[4]. And even though these hands are designed to exploit contact with the environment, the majority of existing grasp planners do not incorporate these benefits. Instead, they regard the environment as an obstacle and explicitly avoid contact with it.

In this paper, we show how to plan grasps that leverage the benefits of exploiting environmental contact. We characterize the specific conditions under which contact-exploiting grasp strategies are successful. This characterization is based on feature descriptors extracted from depth data. Our approach to grasp planning outsources the fine details of a grasp to the soft hand, i.e. it exploits the shape adaptability of soft hands and the possibility of extensively making contact with the environment during grasping. As a result, the grasp planning problem becomes much simpler and low dimensional. Instead of searching the configuration space of the hand for

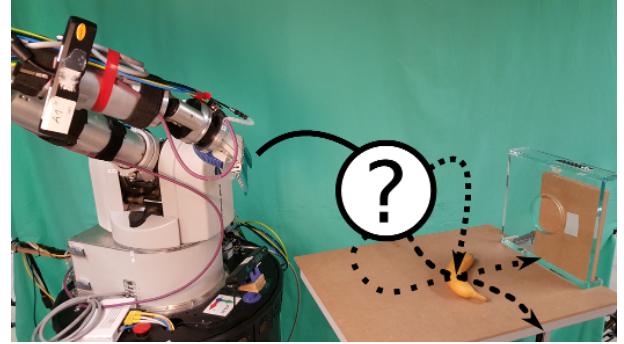


Fig. 1. The main focus of this paper is the visual detection of opportunities in the environment to perform powerful candidate grasps that exploit the environment and to decide based on a single depth image of an on-board camera which of these grasps is the most promising to perform. To solve this planning/decision problem in an data-efficient way, we use a multi-armed bandit formulation that trades off exploration and exploitation.

suitable grasps, we only have to select the most suitable of a small number of powerful contact-exploiting strategies: edge-, surface-, and wall-constrained grasps.

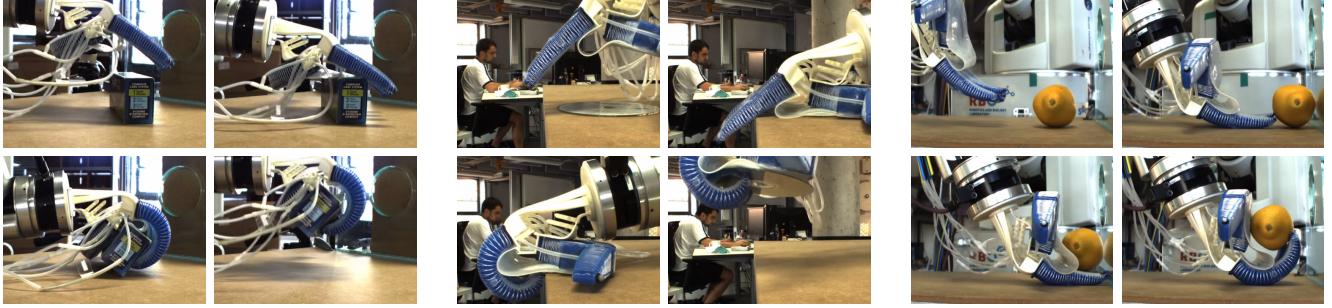
We formulate grasp strategy selection as a multi-armed bandit problem [5]. This allows us to learn outcome models for each strategy from scratch while also finding the most promising strategy for each novel problem scenario as quickly as possible based on prior experience. Among the compared bandit algorithms, we show that UCB [22] with Gaussian Processes performs best while modeling rewards via k-NN and acting  $\epsilon$ -greedy [6] is less susceptible to reward disturbances. In contrast to traditional grasp planners, our method does not require *a priori* knowledge in the form of geometric models of the environment or object. Instead, it plans directly on the sensor input of an RGB-D camera. We estimate models for each grasp type that relate point feature histograms [7] to grasp success. Although past grasping methods have done this [8], our description explicitly incorporates information about the environmental features that are relevant in a grasp. This paper contributes an experience-driven modeling of grasp outcomes, a selection procedure based on these strategy-specific models, and experiments that show our method generating grasp strategies for an anthropomorphic soft hand and a 7-DOF manipulator.

## II. RELATED WORK

According to our contributions, we decompose the relevant related work into three categories: publications that focus on grasps that exploit the environment, work which represents grasps based on sensor input to deal with novel scenarios, and approaches that look at how to improve grasp perfor-

All authors are with the Robotics and Biology Laboratory, Technische Universität Berlin, Germany.

We gratefully acknowledge the funding provided by the Alexander von Humboldt foundation and the Federal Ministry of Education and Research (BMBF), by the European Commission (EC, SOMA, H2020-ICT-645599) and the German Research Foundation (DFG, Exploration Challenge, BR 2248/3-1).



(a) Surface-constrained grasp for a box

(b) Edge-constrained grasp for a CD

(c) Wall-constrained grasp for a lemon

Fig. 2. The different grasping strategies: Each strategy consists of a series of controllers that are triggered based on contact or location events.

mance based on accumulated experience. For each category we emphasize similarities and differences concerning the problem at hand and the proposed solution.

#### A. Environmental Constrained Grasping

More traditional grasp planning techniques either ignore the environment or treat it as an obstacle with which contact must be avoided [9]. There are few grasp approaches which take pre-grasp manipulations with the environment into account, e.g. sliding flat objects across a table before picking them up [10]. Other methods categorize the grasp opportunities present in the environment, such as planes and edges, and represent their topology in a graph structure [11] but ignore the effect of object properties in grasp planning. One reason for exploiting the environment are motion errors. Compensating those errors via contact can be formalized as an optimal control problem [12] or as a sampling-based motion planning problem [13]. In contrast to this paper, all these methods do not focus on how to generalize these motions across a variety of objects.

#### B. Descriptors for Modeling Grasp Success

Since general models based on contact forces require a lot of prior knowledge, a variety of grasp representations have been proposed that describe the object's local geometry in a grasp-variant frame [8]. The advantage of the majority of these representations is that they can be used with raw sensor data, thus, allowing grasp planning for unknown objects. Additionally, they can be learned from different data sources including real-world trials, which circumvents the problems that occur when relying on realistic physics-based simulations [14]. Examples are matching depth templates [15], histogram-like features [16], or image-based descriptors derived with convolutional architectures [17], [18]. Our method is similar, but differs w.r.t. the fact that our features exploit information of the relationship between object and environment by aligning the depth data to the main axis of the exploited EC. Another difference is that our proposed solution must predict the outcome of an time-extended contact interaction.

#### C. Grasp Selection as Exploration vs. Exploitation

The data-driven techniques cited above usually consider a training phase, requiring a large set of grasp examples,

followed by a test phase in which the learned grasp model is applied. We focus on the problem of acquiring grasp experiences to make more informed decisions while concurrently trying to grasp as successful as possible. Active learning approaches to grasping favor grasps that have a high probability of success or large uncertainty. Slagamicoff et al. [19] learn approach directions based on object dimensions, while Montesano et al. [20] use expected improvement as exploration criteria to find the most promising grasping point based on image features. Krömer et al. [21] were the first to use a contextual bandit setting to find grasps. They estimate the reward and its uncertainty of grasps in  $SE(3)$  via Gaussian process regression and apply a variant of UCB [22] to choose a grasp. Given a grasp model that is expensive to evaluate, the bandit setting can also be used to find good grasps with as few samples as possible. Laskey et al. [23] present a 2D grasping model based on uncertainty in shape, pose, friction coefficient, and gripper approach and explore grasps using a Bayesian MAB algorithm. A similar model was extended to 3D [24]. Here, grasps are represented as local depth maps along the approach direction. In both of these models the gap between simulation and reality was not considered. Bandit formulations are also used for large-scale grasp acquisition on real robotic systems [17], [25].

### III. CANDIDATE GRASPING STRATEGIES

We consider three different grasping strategies: surface-constrained, edge-constrained and wall-constrained grasps (see Fig. 2). All of these grasps are composed by alternating between simple straight-line motions and sensor event triggers, such as contact. This motion representation is similar to finite state machines or hybrid automata [26]. Its main advantage is the ability to reduce uncertainty by introducing unambiguous discrete sensor events.

The **surface-constrained grasp strategy** consists of a top-down motion until the wrist-based force-torque measurements exceed a fixed threshold. Subsequently, fingers are closed while the hand is maintaining contact.

The **edge-constrained grasp strategy** starts with a similar top-down movement until contact but then slides the object towards an edge following a straight line. A location-based event triggers upon arrival at the edge and fingers close.

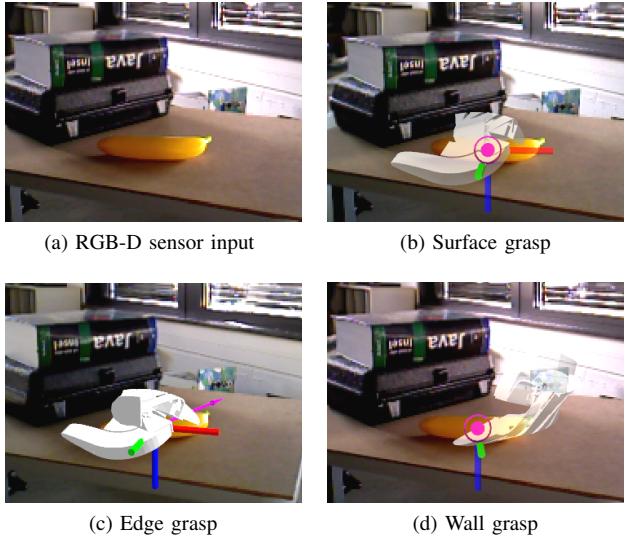


Fig. 3. Detected grasp strategies for a banana on a table: The camera image is overlaid with the grasp frames and a mesh depicting the hand orientation. The backs of the two books give rise to a possible wall grasp strategy.

This strategy allows the fingers to access the object from underneath, which enables grasping heavy or thin objects.

The **wall-constrained grasp strategy** first goes into contact with a surface and then pushes the object towards a wall-like surface. As soon as another contact event triggers the fingers close.

All strategies extensively exploit the environment during grasping, which is especially suited for soft manipulators. We use the RBO Hand 2 [4], a pneumatically actuated anthropomorphic hand made out of silicone that inherently adapts to the shape of a variety of objects when being inflated. In contrast to traditional, stiff hands, unplanned contacts do not necessarily lead to catastrophic outcomes.

#### *Instantiating Candidate Strategies From RGB-D Input*

We detect the opportunity to execute one of those grasp strategies based on geometric features in the environment, extracted from RGB-D sensor measurements [11]. A graph is constructed whose nodes represent regions in the workspace that are accessible within a single action: sliding, caging, free-space movement, plus the three aforementioned grasps. These regions are directly extracted from visual features, e.g. sliding is related to planar patches, an edge grasp to a convex edge, a wall grasp to a concave edge etc. Within this graph a directed edge connects two nodes if it is possible to transit from one action to another which depends on hand pose and contact state (for details see [11]). Finally, a graph search is used to find all possible paths that end in a grasp. These paths can be reformulated as hybrid automata and make up the set of candidate grasp strategies. While previously [11] the object pose was assumed to be known a priori, we now use a simple heuristic to determine what to grasp: we use the largest segment close to the largest surface. Fig. 3 shows an example scene with three detected candidate strategies.

#### IV. GRASP FEATURES TO INFORM GRASP SUCCESS

The candidate grasp strategies are planned based on the geometry of the environment and the rough location of the object to be grasped; object knowledge is ignored. Although in EC grasping the interactions with the environment dominate the object-specific interactions, information about the object is still crucial to predict grasp success.

To compute features for each of the three strategies, we use different local grasp frames. The grasp frame for the wall strategy is aligned with the normals of the wall and support surface. The surface grasp's frame is aligned with the normal of the support surface and the orientation of the hand. Finally, the frame for the edge strategy is aligned to the direction of the edge and the normal of the adjacent plane. These choices of grasp frames reflect the insight that the success of the strategies is mostly invariant in the direction of the contact normal of the environmental contact that is being exploited. We crop the 3D points in the local neighborhood of the candidate grasp frame and analyse three geometric descriptors to characterize the associated grasp strategy:

**Shape Distributions** [27] are signatures that are based on samples from a shape function such as the distance between two random surface points or the angle between three points. We use the distance measure and sample it 2000 times within the grasp region. Distances are discretized into a histogram of 128 bins, each representing a space of 3 mm.

**Shape Histograms** [28] decompose the 3D space into bins, which can have the shape of shells, sectors or combinations of them. We use 128 slices of 2 mm along the height axis with the intention of capturing geometric properties that influence grasp success. A wall grasp succeeds only if the fingers can slip underneath the object which depends on its flushness with the support surface. Since a depth sensor will not measure points in case of these cavities, our shape histogram should capture this property.

**Point Feature Histograms** [7] are a popular feature descriptor for 3D object detection and recognition. They calculate a signature based on the relation of 3D points and their normal information. We use these histograms to characterize the local neighborhood of the grasp frame.

We assume that similarity in these feature spaces also translates to similarity in the resulting physics of the grasps. This is an oversimplification, since effects of mass and friction, for example, cannot be estimated from those features. Still, when casting the problem of predicting grasp success as a binary classification problem based on the geometric descriptors, we can show that a substantial amount of grasp outcomes can be predicted correctly (Sec. VI-B). We train classifiers with an automated machine learning framework [29] and use grasp experiences in simulation and on a real robot.

#### V. CHOOSING AMONG DIFFERENT STRATEGIES

Instead of simply using the classifiers trained above to decide between the three grasping strategies, we would like to continuously update them and make decisions that maximize the expected long-term grasp success. This can be formulated



Fig. 4. Experimental setups with opportunities to do all three types of grasp strategies: *Left*: Barrett WAM with RGB-D sensor and RBO Hand 2 in front of a table and a vertical structure (transparent). *Right*: Simulated RBO Hand 2 in the SOFA simulation framework.

as a multi-armed bandit (MAB) problem [5]. Each arm corresponds to one grasping strategy and the reward is binary, indicating grasp success. Since the underlying reward structure is unknown a MAB algorithm needs to balance exploration (selecting a grasp it has not tried before) and exploitation (choosing the grasp that performed best so far). In our case, the contextual multi-armed bandit scenario is even more appropriate, since the reward distribution depends on the particular object that is grasped. In the contextual MAB setting, the agent receives a context before making the decision. We use the features calculated from RGB-D input as context and estimate the grasp success for each strategy. Popular algorithms for contextual MAB problems include LinUCB, LinTS, and GP-UCB [30], which differ in the way they model the estimated reward, how they are updated when facing new information, and how actions are chosen based on the estimated reward. We evaluate those algorithms including one which models reward via a simple k-nearest neighbor classifier and chooses actions in an  $\epsilon$ -greedy fashion [6].

## VI. EXPERIMENTAL RESULTS

We perform grasping experiments in simulation and on a real robot to show that (*a*) the apparently limited set of the three presented grasping strategies captures a wide variety of objects; (*b*) selecting the best strategy can be learned from data; (*c*) using our MAB formulation, the learning process can happen efficiently in an incremental fashion.

Our simulation experiments are based on SOFA<sup>1</sup> [31]. To simulate the RBO Hand 2 [4], we use a compliance-based constraint solver [32]. Each finger is modeled as a Cosserat beam with empirically identified stiffnesses. Skinning is used for determining the collision geometry. The simulation includes the hand, an object, and environmental constraints such as a table surface, wall, or edge (see Fig. 4). Object meshes are used from the KIT object models database [33] which contains mainly supermarket products. To find random initial object poses that are in static equilibrium we run a separate simulation in which the randomly oriented objects fall on a planar surface until they come to rest. A depth sensor is simulated using the intrinsic calibration parameters derived from an Asus XTion Pro with added Gaussian noise



Fig. 5. The 22 objects used for the real-world experiments

whose standard deviation scales quadratically with measured depth [34]. We run all three grasping strategies for ten poses of each object, resulting in 4020 simulated grasps.

The real-world experiments are conducted on a 7-DoF Barrett WAM platform, including an Asus XTion Pro RGB-D camera on the robot's forearm, a six-axis ATI Gamma force-torque sensor on the wrist, and the RBO Hand 2 as an end-effector. The experimental setup (see Fig. 4) contains a table with a vertical structure that can be used as a wall constraint. In each trial the robot first goes to a pre-defined viewing configuration and uses the RGB-D input to plan a grasp. The object set, shown in Fig. 5, contains 22 items that differ widely in shape, rigidity, surface friction and mass. For each of the three strategies every object was randomly placed on the table 10 times, totalling a number of 630 grasp attempts.

### A. Coverage of Detected Grasping Strategies

Before focusing on the high-level decision between different grasping strategies, we need to show that those options actually solve a significant amount of problem settings. This is done by looking at the performance of the most successful grasping strategy for each problem scenario. We would like to ensure that there are only very few objects/poses that cannot be grasped by any of the three strategies.

The real-world results (see Fig. 6) show that for each object there is at least one strategy that is able to grasp it. The most problematic objects are the gamepad and the marker. The geometry of the gamepad makes it most suitable for a surface-constrained grasp. But since this grasp type can only exert moderate forces that counteract gravity the relatively heavy gamepad is lost 6/10 times. The thin long-shaped marker is most suited for a wall grasp, but half the times the fingertips fail to slip underneath it.

The simulated data shows a different picture, here the strategies only cover 41 % of all problem scenarios (see Fig. 6). This is due to multiple reasons. The KIT object set contains a lot of cuboid-shaped supermarket objects which are relatively large w.r.t. the hand. Twenty large objects could not be grasped by any strategy. The contribution of the edge-constrained grasp strategy is extremely low (39/1340 successes) because of the lack of thin, easy to slide objects that do not topple over (the most successful one was a can of fish). Another problem was that the simulation framework did not allow to differentiate between friction

<sup>1</sup><http://www.sofa-framework.org>

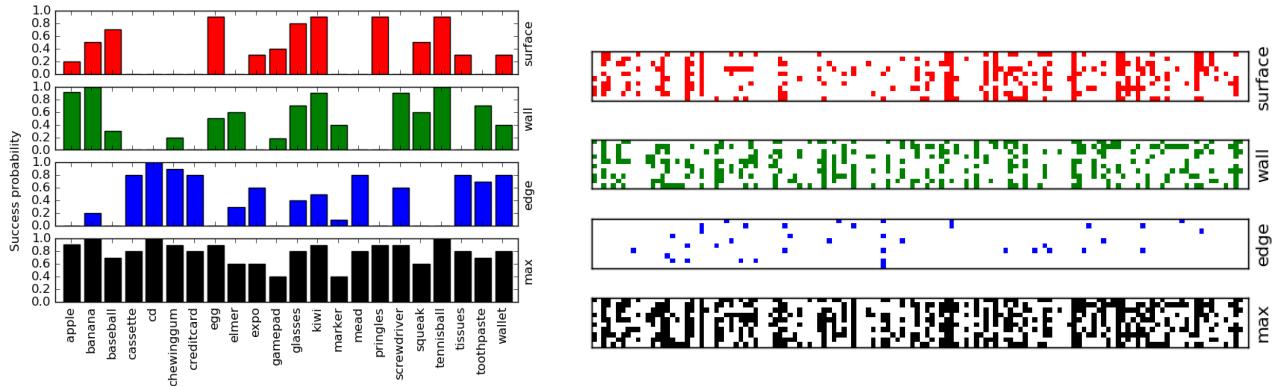


Fig. 6. *Left:* Success rate of the three different grasping strategies for 22 different objects: The success rate is based on ten grasp attempts per object and strategy. *Right:* Each dot represents a successful grasp in simulation. Along the x-axis are the 134 objects and the y-axis depicts ten different poses. The last row named 'max' is the maximum over all strategies per object (and pose). It is an upper performance bound given the robot would choose the optimal grasp in each situation. It also shows that the real-world data set is much better covered than the KIT object set used in simulation.

Feature \ Data Set	Shape Distributions	Shape Histograms	Point Feature Histograms
Simulation (surface)	0.78 (0.56)	0.85 (0.70)	<b>0.88 (0.76)</b>
Simulation (edge)	0.97 (0.94)	0.97 (0.94)	0.97 (0.94)
Simulation (wall)	0.77 (0.53)	0.79 (0.58)	<b>0.86 (0.64)</b>
Real-World (surface)	0.72 (0.44)	0.81 (0.62)	<b>0.86 (0.72)</b>
Real-World (edge)	0.71 (0.42)	<b>0.78 (0.57)</b>	<b>0.78 (0.57)</b>
Real-World (wall)	0.62 (0.25)	0.65 (0.30)	<b>0.79 (0.58)</b>

TABLE I

ACCURACY ( $F_{0.5}$ -SCORES) FOR DIFFERENT GRASP FEATURES

coefficients of table, hand and object, which affected mostly the sliding phase of the edge grasp.

We have shown that on a real-world object set all objects can be grasped by at least one strategy. But the data also shows that the best strategy differs from object to object. Being able to predict the match between strategy and object is the focus of the next experiment.

### B. Grasp Features to Predict Grasp Outcome

We compared the different grasp features presented in Sec. IV based on the accuracy and  $F_{0.5}$  measure of the trained classifiers. We include the  $F_{0.5}$  measure because it weighs recall lower than precision and in grasp detection it is favorable to find at least one robust grasp rather than finding all possible grasps. The data sets for simulation contain 1340 samples for each grasp strategy while the real-world sets contain 220 samples each. They were split into training and test set 4 : 1 using stratified sampling. The ensemble classifiers are trained via AUTO-SKLEARN [29] which searches a structured hypothesis space including multiple types of classifiers and preprocessing methods.

The results of the performance on the test set are listed in Table I. In general, the point feature histograms are the most suitable feature descriptor to predict grasp success for all strategies in simulation and in the real-world. The high scores for the edge strategy in simulation are misleading since it is a highly unbalanced data set with very few successful grasps.

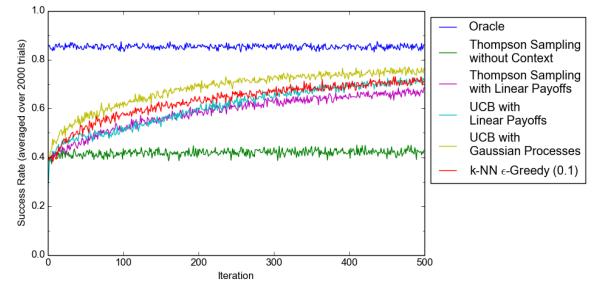


Fig. 7. Success rate of five different exploration strategies for grasping the 22 objects of the real-world data set, averaged over 2000 trials

### C. Exploration-Exploitation Tradeoff

Since the last experiment showed that point feature histograms are the most suited features to describe our grasp strategies, we will use them going forward. In order to efficiently and constantly learn from our grasp experience, we evaluate multiple contextual MAB algorithms on our problem. In each turn the robot faces a random object in a random pose, the three planned candidate grasps and their feature descriptors. The goal is to find a working grasp for any object and pose as quickly as possible. The results (see Fig. 7) show five different methods: A Thompson sampling scheme which ignores context, LinTS, LinUCB, GP-UCB, a k-NN ( $k = 3$ ,  $L_2$  norm on point feature histograms) reward model with an  $\epsilon$ -greedy scheme, and an oracle which represents the highest possible reward (the accumulated difference w.r.t. the oracle is usually called the regret). It can be seen that GP-UCB performs best, followed by k-NN, while the context-less alternative stays constant.

In a second experiment, we modified the reward that is observed by the robot. This resembles the scenario in which a grasp failure was erroneously recognized as success and vice-versa. The comparison between GP-UCB and the k-NN method (Fig. 8) shows that the latter is better at dealing with noise (only 10 % noise equalizes performance).

## VII. LIMITATIONS

The simulation data set has shown that, especially for larger objects, the proposed set of strategies might be too

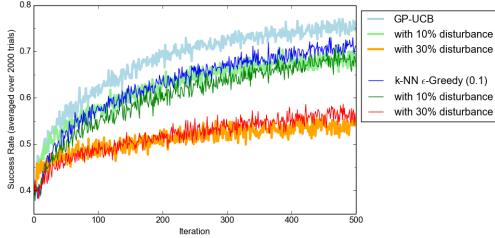


Fig. 8. Success rate of GP-UCB and  $\epsilon$ -greedy k-NN with varying degree of noise in the observed rewards, averaged over 2000 trials

limited. Another strong assumption is the presence of environmental opportunities such as edges and walls to detect the proposed strategies. Both drawbacks can be mitigated by expanding the set of grasp types and exploitable contact structures. Additionally, the algorithm only modifies a few grasp parameters such as the approach direction based on its extracted representation of the environment; the geometric description of the grasp itself has no influence (apart from rejecting it). Allowing modifications of grasp parameters would increase the strategies' applicability. Finally, the proposed geometric features only depend on depth measurements. They do not capture all relevant physical properties needed to predict grasp success, e.g. the baseball was rolling away during a wall grasp while the tennis ball succeeded (see Fig. 6). RGB features could correlate with such properties [18].

### VIII. CONCLUSION

We presented a grasp planner for unknown environments capable of exploiting contact with the environment. To develop our method, we relied on three contact-exploiting grasping strategies. We showed that it is possible to learn perceptual models that predict the success of these strategies based on a single depth image alone. We formulated the problem of deciding among this small number of grasping strategies while improving the perceptual models' predictive power from experience as a contextual multi-armed bandit problem. The quality of the decision improves by incorporating context in the form of perceptual information.

### REFERENCES

- [1] C. Eppner, R. Deimel, J. Alvarez-Ruiz, M. Maertens, and O. Brock, "Exploitation of environmental constraints in human and robotic grasping," *IJRR*, vol. 34, no. 7, pp. 1021–1038, 2015.
- [2] M. G. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, and A. Bicchi, "Adaptive synergies for the design and control of the Pisa/IIT SoftHand," *IJRR*, vol. 33, no. 5, pp. 768–782, 2014.
- [3] A. M. Dollar and R. D. Howe, "The highly adaptive SDM hand: Design and performance evaluation," *IJRR*, vol. 29, no. 5, pp. 585–597, 2010.
- [4] R. Deimel and O. Brock, "A novel type of compliant and underactuated robotic hand for dexterous grasping," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 161–185, 2016.
- [5] H. Robbins, "Some aspects of the sequential design of experiments," in *Herbert Robbins Selected Papers*. Springer, 1985, pp. 169–177.
- [6] Y. Yang, D. Zhu, et al., "Randomized allocation with nonparametric estimation for a multi-armed bandit problem with covariates," *The Annals of Statistics*, vol. 30, no. 1, pp. 100–121, 2002.
- [7] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms for 3D registration," in *Proc. of ICRA*. IEEE, 2009, pp. 3212–3217.
- [8] 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, 2014.
- [9] C. Goldfeder, M. Ciocarlie, H. Dang, and P. K. Allen, "The columbia grasp database," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 1710–1716.
- [10] D. Kappler, L. Chang, M. Przybylski, N. Pollard, T. Asfour, and R. Dillmann, "Representation of pre-grasp strategies for object manipulation," in *Int. Conf. on Humanoid Robots*, 2010.
- [11] C. Eppner and O. Brock, "Planning grasp strategies that exploit environmental constraints," in *Proc. of ICRA*, 2015, pp. 4947 – 4952.
- [12] M. Toussaint, N. Ratliff, J. Bohg, L. Righetti, P. Englert, and S. Schaal, "Dual execution of optimized contact interaction trajectories," in *Int. Conf. on Intelligent Robots and Systems (IROS)*, 2014.
- [13] C. Phillips-Grafflin and D. Berenson, "Planning and resilient execution of policies for manipulation in contact with actuation uncertainty," in *Workshop on the Algorithmic Foundations of Robotics (WAFR)*, 2016.
- [14] D. Kappler, B. Bohg, and S. Schaal, "Leveraging big data for grasp planning," in *Proceedings of the IEEE International Conference on Robotics and Automation*, May 2015.
- [15] A. Herzog, P. Pastor, M. Kalakrishnan, L. Righetti, J. Bohg, T. Asfour, and S. Schaal, "Learning of grasp selection based on shape-templates," *Autonomous Robots*, vol. 36, no. 1-2, pp. 51–65, Jan. 2014.
- [16] D. Fischinger, A. Weiss, and M. Vincze, "Learning grasps with topographic features," *The International Journal of Robotics Research*, vol. 34, no. 9, pp. 1167–1194, 2015.
- [17] L. Pinto and A. Gupta, "Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours," *arXiv preprint arXiv:1509.06825*, 2015.
- [18] S. Levine, P. Pastor, A. Krizhevsky, and D. Quillen, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *arXiv preprint arXiv:1603.02199*, 2016.
- [19] M. Salganicoff, L. H. Ungar, and R. Bajcsy, "Active learning for vision-based robot grasping," *Machine Learning*, vol. 23, no. 2-3, pp. 251–278, 1996.
- [20] L. Montesano and M. Lopes, "Active learning of visual descriptors for grasping using non-parametric smoothed beta distributions," *Robotics and Autonomous Systems*, vol. 60, no. 3, pp. 452–462, Mar. 2012.
- [21] O. B. Kroemer, R. Detry, J. Piater, and J. Peters, "Combining active learning and reactive control for robot grasping," *RAS*, vol. 58, pp. 1105–1116, Sept. 2010, ACM ID: 1838783.
- [22] P. Auer, "Using confidence bounds for exploitation-exploration trade-offs," *Journal of Machine Learning Research*, vol. 3, no. Nov, pp. 397–422, 2002.
- [23] M. Laskey, J. Mahler, Z. McCarthy, F. T. Pokorny, S. Patil, J. van den Berg, D. Kragic, P. Abbeel, and K. Goldberg, "Multi-armed bandit models for 2D grasp planning with uncertainty," in *Proc. of CASE*, 2015.
- [24] J. Mahler, F. T. Pokorny, B. Hou, M. Roderick, M. Laskey, M. Aubry, K. Kohlhoff, T. Kröger, J. J. Kuffner, and K. Goldberg, "Dex-net 1.0: A cloud-based network of 3D objects for robust grasp planning using a multi-armed bandit model with correlated rewards," in *2016 IEEE ICRA*, 2016, pp. 1957–1964.
- [25] J. Oberlin and S. Tellex, "Autonomously acquiring instance-based object models from experience," in *Proc. of ISRR*, 2015.
- [26] M. Egerstedt, "Behavior based robotics using hybrid automata," in *Hybrid Systems: computation and control*. Springer, 2000.
- [27] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Trans. Graph.*, vol. 21, no. 4, pp. 807–832, Oct. 2002.
- [28] M. Ankerst, G. Kastenmüller, H.-P. Kriegel, and T. Seidl, "3D shape histograms for similarity search and classification in spatial databases," in *ISSD*. Springer, 1999, pp. 207–226.
- [29] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," in *Proc. of NIPS*. MIT Press, 2015, pp. 2755–2763.
- [30] L. Zhou, "A survey on contextual multi-armed bandits," *CoRR*, abs/1508.03326, 2015.
- [31] J. Allard, S. Cotin, F. Faure, P.-J. Bensoussan, F. Poyer, C. Duriez, H. Delingette, and L. Grisoni, "SOFA - an Open Source Framework for Medical Simulation," in *MMVR 15*, vol. 125, Feb. 2007, pp. 13–18.
- [32] M. Tournier, M. Nesme, B. Gilles, and F. Faure, "Stable Constrained Dynamics," *ACM Trans. Graph.*, vol. 34, pp. 132:1–132:10, 2015.
- [33] A. Kasper, Z. Xue, and R. Dillmann, "The kit object models database: An object model database for object recognition, localization and manipulation in service robotics," *IJRR*, pp. 927–934, 2012.
- [34] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.