

# A Topology-based Object Representation for Clasping, Latching and Hooking

Johannes A. Stork

Florian T. Pokorny

Danica Kragic

**Abstract**—We present a loop-based topological object representation for objects with holes. The representation is used to model object parts suitable for grasping, e.g. handles, and it incorporates local volume information about these. Furthermore, we present a grasp synthesis framework that utilizes this representation for synthesizing caging grasps that are robust under measurement noise. The approach is complementary to a local contact-based force-closure analysis as it depends on *global topological features* of the object.

We perform an extensive evaluation with four robotic hands on synthetic data. Additionally, we provide real world experiments using a Kinect sensor on two robotic platforms: a Schunk dexterous hand attached to a Kuka robot arm as well as a Nao humanoid robot. In the case of the Nao platform, we provide initial experiments showing that our approach can be used to plan whole arm hooking as well as caging grasps involving only one hand.

## I. INTRODUCTION

Local contact-based methods for grasp stability assessment are a well-established area in robotics. Classical, physics-based frameworks such as the grasp quality measures of [1] determine if a grasp can withstand external forces of a known magnitude. Such force-closure based grasp synthesis approaches commonly assume that the manipulator and the grasped object are rigid and have known geometry and friction properties.

In addition to solely fixing an object in the hand, humans employ a vastly richer repertoire of grasps affording both in-hand manipulation and additional skills such as dragging, toting, pulling and hauling of soft and rigid objects. These activities are important in many situations and may require large forces which the robot might not be able to exert using point-contact based grasps. A mobile robot may also want to latch onto objects and fixed structures in the same way as humans do: consider for example clasping holding-bars in buses or subways, grabbing handrails, pushing a baby stroller, or using a pair of scissors.

While actions such as holding and lifting *do* depend on stability in terms of constraining forces and contacts, the classical methods of assessment and planning of this form of stability require detailed surface models. For the generation of stable grasps on complex objects with many contact points, no suitable analytical methods are currently available. Most common methods employ some form of sampling [2]–[4], various heuristics [5]–[7] or are based on previously learned probabilistic models of grasping poses for a specific

The authors are with the Centre for Autonomous Systems, Computer Vision and Active Perception Lab, School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden. {jastork, fpokorny, dani}@kth.se

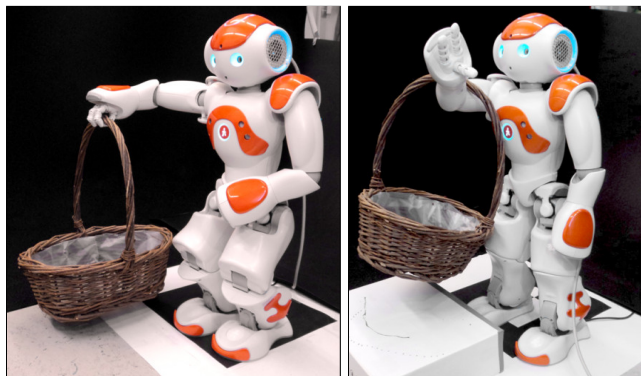


Fig. 1. Two example grasps are displayed: a Nao grasping a basket using a chain forming grasp and a whole arm hooking grasp.

class of objects [8], [9]. Furthermore, sparse or noisy data often prohibit a sufficiently good object reconstruction such as volumetric representations [10], [11] derived from a watertight surface reconstruction [12], or a sufficiently accurate normal direction estimate for force-closure analysis.

While force and contact stability could additionally be incorporated, it is not a priori required within our framework. In fact, a complete fixation of the object may even be counterproductive, e.g. in the case of using a pair of scissors. We believe that, if robots are to eventually reach the dexterity of an acrobat catching a rope in mid-flight or swinging with its hands hooked securely around a ring, methods based on *global rather than local* object properties can provide an alternative approach to grasp synthesis. The chain forming grasps presented in this paper exploit a global object property – namely a basis of the first homology group formed by *shortest closed loops* on the object [13], [14]. As in [14], we consider these as a basis for interlinking an object and robot hand or arm for the purpose of grasp synthesis.

The main contributions of this work are:

- We propose a new object representation for generating hypotheses of chain forming grasps from point cloud data. The method is capable of dealing with unknown and potentially deformable objects using global topological features of the object’s surface geometry.
- We evaluate our method on the Schunk, Armar III, DLR 1 and iCub hands with synthetic data, showing that our approach can be used for a wide range of hand kinematics.
- We provide real-world experiments with a Nao humanoid robot and a KUKA/Schunk setup, and with noisy partial view Kinect data for several rigid and soft objects.

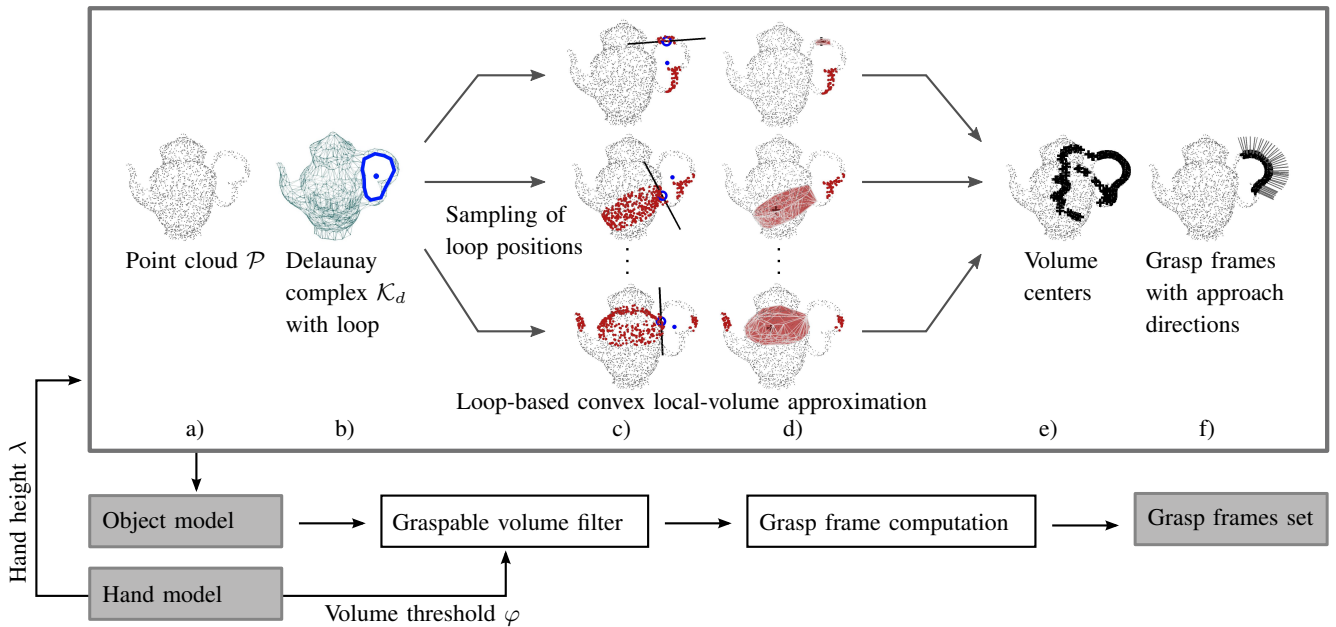


Fig. 2. System overview: Point cloud data  $\mathcal{P}$  and its refined Delaunay triangulation  $\mathcal{K}_d$  are shown in *a*) and *b*). A shortest loop and its center is depicted in blue. Selected loop points and loop tangents are marked, and associated slices of  $\mathcal{P}$  are indicated in red in *c*). Convex volumes and centers of the slice’s adjacent connected component are shown in *d*). Volume centers and approach vectors of selected grasp frames after filtering are indicated in *e*) and *f*).

### System Outline

Our method starts by building a geometric mesh description directly from 3D point cloud data and applies a topology-based algorithm [13] to identify shortest non-contractible closed loops in the model. Fig. 2 depicts an example point cloud, a mesh description and a loop on the left. Subsequently, a hand model and loop tangents are used to construct approximations of the local geometry for potentially graspable handle-like object parts. For each considered loop point, a slice of the point cloud, orthogonal to the tangent is extracted. The convex hull of the slice’s connected component that is adjacent to the loop point is determined as a local volume approximation. Slices and convex hulls can be seen in Fig. 2c,d).

Finally, these local volume approximations are filtered by an end effector-determined threshold and grasp parameters are computed for the remaining volume centers. The last step is visualized on the right side of Fig. 2. There, the volume centers are depicted as black plus-signs, while the grasp frames are described by approach vectors.

Although our approach, based on the detection of handle-like object parts, is limited to objects with *holes* there are many natural objects that feature such parts – several examples will be shown in the experimental evaluation.

In this work, we consider grasping in the broader sense of holding an object with or without the support of external forces such as gravity. Our system plans general *chain forming grasps* that maximize the interlinking between suitable parts of the object with an end effector. The resulting configurations can informally be described as clasping, latching and hooking grasps. For illustration, Fig. 1 shows clasping or latching on left and hooking on the right side. The grasps synthesized in this work establish caging grasps in a generalized practical sense and not necessarily in the

analytical sense. This means that the escaping probability of the object is low. For clasping or latching of rigid bodies, we used a random motion rigid body simulation to test if the object was caged in this sense. Hooking grasps such as in the left part of Fig. 1 cage the object only with respect to motions under task constraints. The basket in Fig. 1 will remain hooked only under reasonable expected motions and forces applied e.g. during walking. Note that not all styles of such caging grasps afford the same activities or provide similar flexibility. Hooking grasps could be useful in particular when two humanoid robots carry a basket together, since they impose fewer constraints on the two robots’ motion than a tight clasp of the object.

## II. MOTIVATION AND RELATED WORK

Point-contact, physics-based methods depending on the Coulomb friction model such as force-closure are a popular approach to grasp synthesis as reviewed in [15]. In [2], grasp hypotheses are generated by approaching an object from randomized starting directions and closing the robot’s hand until contact occurs. Optimal or viable grasps are then chosen based on a grasp quality function [1].

Recent work of [16] examines empirical or data-driven methods for grasping. As stated in the survey, the grasping process depends on many factors such as the physical properties of the object, the embodiment of the robot, available sensory information and task constraints. Thus, many approaches apply some form of heuristic directly in the state space to simplify aspects of the problem [5]–[7], [10], [16], [17]. Furthermore, a suitable simplification of the robot’s hand state space representation, such as the notion of postural synergies discussed in [18], can be used to simplify the control of the robot’s hand. Recently, several extensions of these ideas have been actively pursued [19]–[21].

There are also many approaches that perform a decomposition of objects and generate grasps on individual, decomposed parts. Examples include unions of parametrized shape primitives from manual matching [2] and automatic hierarchical fitting methods [22], [23]. While a superquadric decomposition tree may provide some flexibility [22], the high number of parameters may limit the viable complexity of considered objects. Hence, simple hierarchical box shape-based models have been used to represent objects from range data [23]. These methods relate to our approach since they approximate object volumes. However, they do not realize nor exploit *global structural information* such as the connectivity between parts or topological properties of the object.

Recently, Morse theory has been applied to determine topological skeletons such as Reeb graphs for shape decomposition [24]. This approach results in a decomposition into parts based on surface properties and relies on surface curvature to define part boundaries. The concept has been applied to semantic grasp planning [25], where the topological segmentation is annotated by automatic part recognition and shape retrieval. Based on the same idea, programming by demonstration for grasping and manipulation has been proposed [26], [27]. In these approaches, the Reeb graph is however only used for segmentation and transfer of grasp affordances by graph matching. The extracted information about the global object structure and the relation of the parts is not utilized to plan the grasps. Our method, in contrast, focuses on global topological information directly.

The medial axis representation, a volume-based object representation describing maximal inscribing spheres was introduced for the purpose of grasping in [10]. The representation consists of a union of sphere centers which captures the object’s geometric connectivity and allows to exploit local symmetry properties. Based on a tree structure derived from the medial axis, a set of heuristics is applied to generate viable grasps. The approach has been extended to a grid-based volumetric representation that examines symmetry properties of the grid-local neighborhood [11]. This method is similar to ours since it is volumetric and it directly considers structural information. However, the approach suffers from inherent instability under noise [28]. The underlying computational method for extracting the approximated medial axis [12] is based on the reconstruction of a watertight surface of the whole object. This depends heavily on a dense, full-view point cloud and is thus sensitive to noise. In contrast, we will show that our method extracts useful global structural object information even when the data is sparse, noisy or obtained from a partial view.

In contrast to force-closed grasps, caging grasps only bound an object’s mobility to prevent it from *escaping arbitrarily far* from the manipulator. A caging grasp provides a way to control an object without immobilizing it completely. Early research on caging has focused on a theoretical analysis and on efficient algorithms for simple hand mechanisms with few degrees of freedoms [29], [30] and considered mostly planar objects. Object concavity has been exploited

for caging with finger position maximally far away from the object [31]. Later, the study of three-dimensional caging by multi-fingered hands led to the definition of sufficient conditions and resulted in a cage planning system [32]. Caging configurations have been considered in relation to grasping as a waypoint or initial point to immobilizing grasps [33] and as a method to deal with uncertainty in grasping [34]. Applications in manipulation planning have been found, where rigidity constraints between an object and the robot can be neglected in planning when the object is caged [35].

To the best of our knowledge, caging research so far has only considered rigid objects. In this work, we particularly focus on chain forming grasps which attempt to interlink the object with the manipulator and result in a caging configuration if the object is rigid. However, in the experimental evaluation, we will also show that our approach can be used for caging soft or flexible objects with holes.

In summary, our approach prevents objects from escaping without necessarily creating immobilizing grasps. In particular, we exploit an object representation based on global structure information which allows us to apply simple enveloping grasp policies to grasp rigid and flexible objects.

### III. METHODOLOGY

In this section, we provide details on the detection of closed loops, the employed end effector models, the object representation and elaborate on our grasping strategy.

#### A. Detection of Loops

A key concept used in this paper are *closed loops* on a closed surface  $S \subset \mathbb{R}^3$  which are non-trivial as elements of the first homology group  $H_1(S)$ . When computed over a finite field,  $H_1(S)$  is a vector space whose elements correspond to equivalence classes of closed curves and where two loops are considered equivalent if they form a boundary of a part (a 2-cycle) of  $S$ . The first homology group is used in Algebraic Topology [36] and stays invariant under continuous deformations of  $S$  called homotopies. We are in particular interested in a basis of  $H_1(S)$  consisting of curves which are of *shortest length* and use them as features for grasp synthesis as in [14]. Since we reconstruct our surface representation from point-cloud data, we will work with a simplicial complex approximation  $\mathcal{K}_d$  of  $S$  for which the notion of  $H_1(\mathcal{K}_d)$  is also well-defined. We employ the software *ShortLoop* [37] to determine a basis of approximately shortest curves of  $H_1(\mathcal{K}_d)$ . An example of the aforementioned shortest loop basis is displayed in Fig. 2b.

For a watertight perfect mesh model, loops in  $H_1(S)$  can be further classified as *tunnel-loops* and *handle-loops* [38], with tunnel loops running around holes in an object. Tunnel loops traverse handle-like parts of an object and are hence most relevant for grasping. We however work with a rough simplicial complex reconstruction obtained from point-cloud data where this distinction is often impossible to make. Non-graspable elements of  $H_1(\mathcal{K}_d)$  will hence at first be included in our analysis and are only eliminated during a later collision detection phase.

End effector	height [m]	radius [m]	$\frac{\lambda}{2}$ [m]	$\varphi$ [m]
Schunk SDH	0.120	0.050	0.060	0.1562
Armar III	0.090	0.035	0.045	0.1140
DLR I	0.130	0.060	0.065	0.1697
iCub	0.085	0.025	0.425	0.0986

Fig. 3. Pre-shapes with GCP frames (top) and grasps on maximal volume cylinder (middle), see Section III-B. The arrows show  $\varphi$  (white) and  $\lambda$  (yellow). Table below: End effector description parameters.

### B. End Effector Related Parameters

We define two parameters for a robot hand which define the shape of the maximal cylinder that the end effector can secure by closing the fingers from a parallel pre-shape as seen in Fig. 3. The height of the cylinder,  $\lambda$ , describes the grasp width,  $\varphi$  describes the maximal extent of the cylinder. Parameter  $\lambda$  is used to build an object model that consists of hand-sized volumes while  $\varphi$  serves as a bound in grasp parameter generation to reject overly large volumes.

For an end effector, we define a grasp center point frame (GCP frame) that defines the center of the cylinder relative to the robot hand. All end effectors used in this work can be described by a set of individual fingers and a static palm. Our approach requires that at least one of these fingers can oppose the remaining fingers to form a parallel pre-shape.

The DLR Hand (12 DOF) is a four-fingered articulated robotic hand that is 50% larger than the average human hand. The hand of the child-sized iCub humanoid robot has 20 DOF. The five-fingered IAI-HAND-12 (10 DOF) for the ARMAR III platform is a symmetric three-jaw grasper. The Schunk SDH (7 DOF) is a 3-fingered fully actuated industrial robot gripper. We disabled the coupled rotation of two of the Schunk hand’s fingers, resulting in a 6 DOF manipulator.

### C. Object Representation

A loop, as defined in Sec. III-A, represents a topological and geometric attribute of an object. Therefore, we model objects with holes as a set of convex local-volume approximations of their approximated shortest loops. Considering chain forming grasps – created by closing the end effector around a segment of an extracted loop – we are interested in the local shape of the enveloped loop section. We infer this information from unordered full-view 3D point cloud data. The individual steps of the model building process are exemplified in Fig. 2a–e).

Assuming that the point cloud  $\mathcal{P} \subset \mathbb{R}^3$  is sampled from a closed surface, we start by refining the Delaunay complex of  $\mathcal{P}$  and denote the result  $\mathcal{K}_d$ . The refinement consists of removing all edges with length  $d$  or larger and eliminating all tetrahedra and faces involving those edges. Hence, the value of  $d$  limits the distance of two points up to which they

are still joined in  $\mathcal{K}_d$  by some edge, face or tetrahedron. We assume that  $\mathcal{P}$  is dense enough so that the triangulation results in a descriptive model of the observed object. The exact value of  $d$  can be chosen in an informed way by considering additional constraints, e.g. using a persistent homology analysis of  $\mathcal{P}$  in the form of barcodes [39]. In the experiments however, we select the value manually based on point cloud density and hand extent.

We apply a polynomial time algorithm [37] on the finite simplicial complex  $\mathcal{K}_d$  that results in a set,  $\mathcal{L}$ , of cycles of linear line segments, approximating a basis of shortest loops for  $H_1(\mathcal{K}_d)$ . In [14], we give further details of the process and its parameters. Considering a loop  $l \in \mathcal{L}$ , we approximate the loop-tangent,  $T_P$ , of a point  $P$  on  $l$ .

Next, we construct an orthogonal plane,  $O_P$ , containing  $P$  with normal direction  $T_P$ . Given a point  $P$  on a loop, we describe the object’s local shape by a convex volume,  $V_P$ . To this end, we consider an undirected graph,  $G_P$ , derived from the simplicial complex  $\mathcal{K}_d$ . The graph arises from the edges of  $\mathcal{K}_d$ , defined on the points of  $\mathcal{P}$  that are within distance  $\frac{\lambda}{2}$  to the plane  $O_P$  (Fig. 2c). Finally,  $V_P$  is defined as the convex hull of the connected component closest to  $P$  (Fig. 2d).

In the process described above, every point on an approximated shortest loop can be annotated with a hand-size related convex approximation of the loop-local volume. The extracted set of loops, loop points, and volumes forms our object representation.

### D. Grasping Process

We parametrize a grasp by prescribing the end effector’s GCP frame and consider the  $Z$ -axis as the approach direction. Given an end effector’s GCP and a convex local-volume approximation  $V_P$  at a loop  $l \in \mathcal{L}$ , as defined above, we move the GCP frame into the volume’s center. Since the center of  $V_P$  already defines the position, only the frames orientation is left to choose. The loop center point,  $M_l$ , is given as the center of gravity of the points in  $\mathcal{P}$ , supporting the linear line segments of  $l$ .

Beginning with the approximated tangent of the loop point,  $T_P$ , and using the vector  $Z = M_l - P$ , we construct an orthonormal coordinate frame located at the center of  $V_P$  that is identified with the GCP frame. We reject all sampled points,  $P$  on  $l$  that result in a maximal extension of  $V_P$  larger than the end effector’s volume measure threshold  $\varphi$ . If the robotic hand is not symmetric, we additionally include grasp parameters rotated around the frame’s  $Z$ -axis by 180 degrees for each accepted volume.

## IV. EXPERIMENTAL EVALUATION

We first evaluate the performance of the proposed method on synthetic data using four robotic hands. After that, we present results from a real setup using Kinect data, a 6 DOF Kuka arm with a Schunk SDH and a Nao humanoid robot.

### A. Experiments on Synthetic Data

We first demonstrate our approach on five real-sized objects by executing a simple grasp policy. From 3D polygon

mesh models [40], [41] we sample 4000 points on the visible outside surfaces by moving a virtual camera around the object. Thus, the average distance to the closest neighbor point is about 0.03 m. This point density is acceptable since we are interested in particularly exposed object features and since the rough object geometry is sufficient for our method. When refining the simplicial complex  $\mathcal{K}_d$  as described in Sec. III-C, we choose the maximal edge length,  $d$ , as 0.06 m. Noise, randomness in sampling and object concavity can lead to detection of small and negligible loops in  $\mathcal{K}_d$ . We reject small loops where the maximal distance between any two loop vertices is below 0.09 m.

When evaluating the method in simulation, we apply a straightforward grasping policy as proposed in [3], [42]. As in other recent work such as [11], we initialize the end effector in parallel pre-shape and align its GCP frame with the grasp frame (GF) derived from the object model. Next, the end effector is moved backwards along the frame’s  $Z$ -axis until no collision occurs. The resulting end effector pose and joint angles define the grasp. If the backwards movement does not remove all collisions within 0.08 m, we consider the grasp to be obstructed.

If a grasp is not obstructed, its reliability is tested by simulating pushing, turning and pulling interactions. While keeping the object fixed, we perform 10 separate sequences of 100 interleaved random rotations and translations of the end effector. Each random motion is executed until it leads to separation of the end effector and object or until contact is achieved. A grasp is considered failed if the end effector and object are separated at least once. We have found that the results do not change significantly for sequences with more than 60 motions.

To describe the quantitative results, we employ the terminology of a classification experiment in the following. In our approach, we denote non-separated i.e. secure grasps as *true positives*, and obstructed or failed grasps as *false positives*. Fig. 4 shows for each pair of end effector and object, the true positive rate (red), and the rates of failed (black), and obstructed (blue) grasps each. The rejection rate of the filtering step (yellow) describes how much of the handle-parts cannot be grasped by the particular end effector. The true positive rate ranges from 72.93% in the case of the DLR 1 hand with the *Bag* object (Fig. 5.2b) to 100% in the cases of Armar III and DLR 1 hands with the *Purse*, Schunk SDH with the *Travel Bag* and all cases of the iCub hand. The overall precision is 94.91%. The largest false positive rates are caused by the two large hands, DLR 1 and Schunk SDH. On the *Bag* object, the separability is predominant (Fig. 5.2b and Fig. 5.3b), while on the *Chair* object, most false positives are obstructions and are therefore caused by the simple grasp policy.

A qualitative analysis is conducted by inspecting Fig. 5, where we show the locations of successful, failed and obstructed grasps on the objects, next to robot poses of successfully executed grasps. The marker colors correspond to the colors in Fig. 4. It can be observed that the *accepted* grasp frames describe parts that are well suited for enveloping

grasps. Frames inside a large volume are rejected. For the *Purse* object, it is clearly visible how our approach respects the robot hand size by only selecting grasp frames on the top part of the handle for large hands, while the smaller hands still grasp from the side. This property can also be observed on the *Lawn Mower*, where the small hands cannot envelope the control lever box and therefore grasp frames are only generated for the large hands. For the *Bag*, the failed grasps are located on a narrow part of the belt which prevents the large fingers from closing properly.

However, most of the time at least one of the two rotated copies of a grasp frame does not lead to a failed grasp. The structure of the *Chair* results in obstruction of the large hands at the backrest since our approach does not consider the rest of the object when judging a loop-local volume approximation. Obstruction also occurs in corners since, there, the loop tangent used to define the grasp frame orientation is not completely aligned with the object shape and leads to a non-perpendicular placement by the grasp policy. Both the *Purse* and the *Travel Bag* object show that our approach can place large hands on relatively small handle-parts. One loop was detected on the *Purse*, *Bag* and *Wheel Bag*, two on the *Lawn Mower* and seven on the *Chair* object.

## B. Experiments on Real Robot Platforms

In the following experiments, we relax the assumption of the full-view point cloud and rely on subsampled noisy Kinect data to show the applicability of our approach in real-world scenarios. In the first experiment with a Kuka arm and a Schunk SDH hand, the sensor is calibrated relative to the robot frame while in the experiment with a Nao humanoid, the robot is placed on a visual marker in the scene after the scene has been analyzed, inducing additional uncertainty in the robot’s pose.

1) **Nao Humanoid Robot:** The experiments using a Nao robot consider two different ways of picking up a small basket: grasping the basket at the handle with chain forming grasps and hooking the handle employing one whole arm. For robots with limited workspace – such as a Nao – reachability and kinematic feasibility become important issues in a non-mobile setting. Therefore, the computed grasping frames need to be carefully selected. Here, we simplify the problem by applying a simple distance-based scoring function that selects the frame closest to the torso. Fig 6 shows the course of one grasping and one hooking experiment at different time instants from left to right.

In the grasping experiment, an iterative inverse kinematic solver is used and initialized at a starting position suitable for the selected target frame. As a sample application of chain forming grasps, the robot later rises up, turns its arm and lets the handle slide through the hand to establish the proper object configuration for carrying. In case of the hooking experiment, the robot executes a five-waypoint-based motion template parametrized by the selected target frame to pick up and lift the basket in one action. For these experiments the Kinect sensor was placed at a distance of about 2 m.

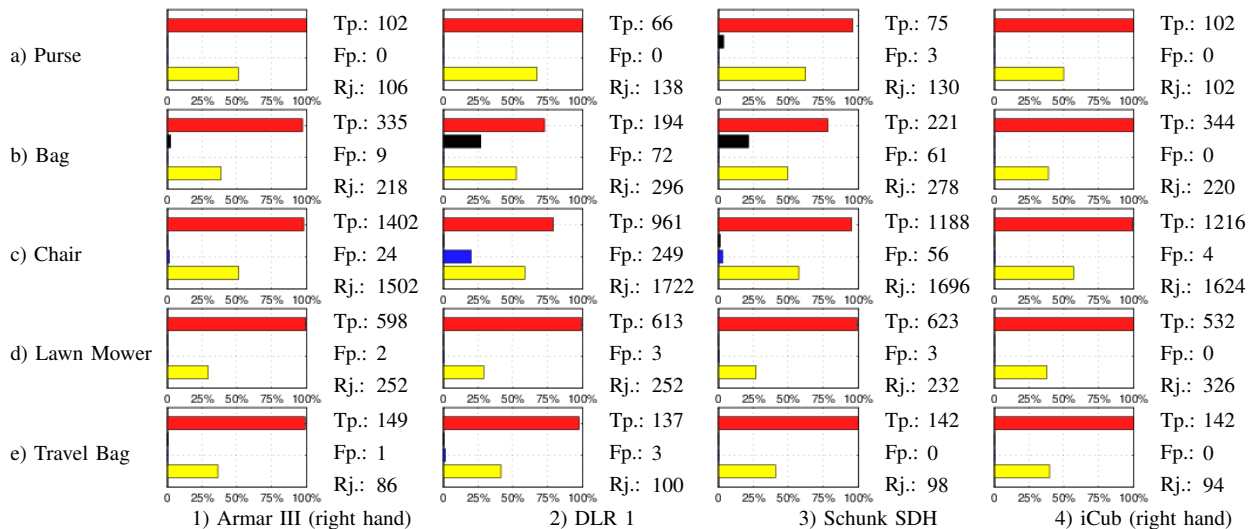


Fig. 4. Quantitative results of the synthetic experiment: The bar plots show the *true positive rate* of the applied heuristic, i.e. secure grasps, as the top bar (red). The *false positives* are divided into failed grasps (black), and obstructed grasps (blue). The bottom bar shows the *rejection rate* of the heuristic function (yellow). The colors are chosen according to the markers in Fig. 5. The absolute numbers for *true positives*, *false positives*, and *rejected* grasp frames (Tp., Fp. and Rj.) are given next to each plot.

2) **Kuka Arm & Schunk SDH:** In contrast to the synthetic data experiments, we employ an approach vector based grasp policy. We first move the Schunk SDH in pre-shape to a pose 0.2 m in front of the grasp frame. Then we move the end effector forward and finally close the fingers until a closed chain is formed. The experiment consists of executing this grasp policy on several grasp frames at representative locations for three different objects with subsequent lifting. Fig. 7 shows the robot grasping a children’s chair, a basket and a real-sized backpack. The images indicate that our approach results in secure chain forming grasps and the supplementary video confirms that each object, including the soft deformable backpack, can successfully be grasped.

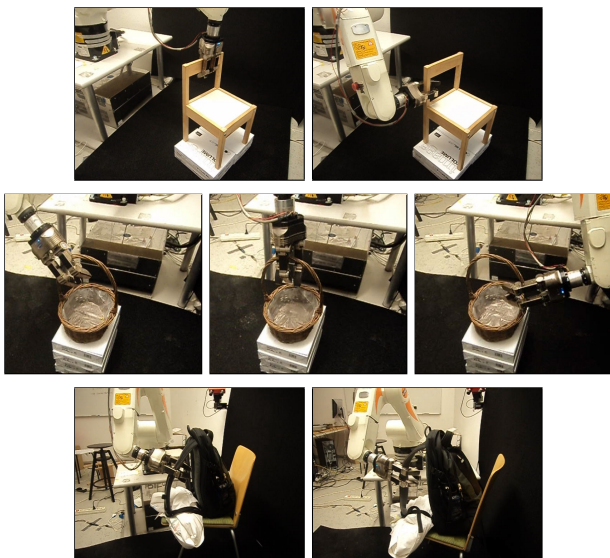


Fig. 7. Kinect data is used for the execution of grasp policies for several generated grasp hypotheses. Details of the execution can be seen in the supplementary video.

## V. CONCLUSION

We have presented work on generating grasps that enable clasping, latching and hooking of everyday objects that exhibit handle-like parts. The proposed method is topology-based and results in an object representation suitable for objects with holes. The representation is used to identify object parts suitable for grasping and incorporates local volume information about these. A whole grasp synthesis framework that utilizes such loops has been presented and evaluated on two real robot platforms.

The approach is complementary to a local contact-based force-closure analysis as it depends on global topological features of an object. We have shown that our grasp synthesis framework leads to caging grasps that are robust under real world measurement noise. As demonstrated, the approach is also applicable to soft and flexible objects. We have performed an extensive evaluation with four robotic hands in simulation as well as using a real setup with a Kinect sensor, a Schunk dexterous hand and a Nao humanoid robot.

Our current work is exploring the use of different hand pre-shapes for grasping and the use of an improved grasp policy featuring eigengrasps and motion planning. We are also interested in coupling this work with semantic information that constrains the set of viable grasps based on the robot’s intended task, as demonstrated in our previous work in [44], [45].

## ACKNOWLEDGMENT

This work was supported by FLEXBOT (FP7-ERC-279933), the Swedish Research Council and the Swedish Foundation for Strategic Research. We thank Yasemin Bekiroglu, Alejandro Marzinotto, and Francisco Vina for their help with the robot experiments.

## REFERENCES

- [1] C. Ferrari and J. Canny, “Planning optimal grasps,” in *IEEE ICRA*, vol. 3, 1992, pp. 2290–2295.

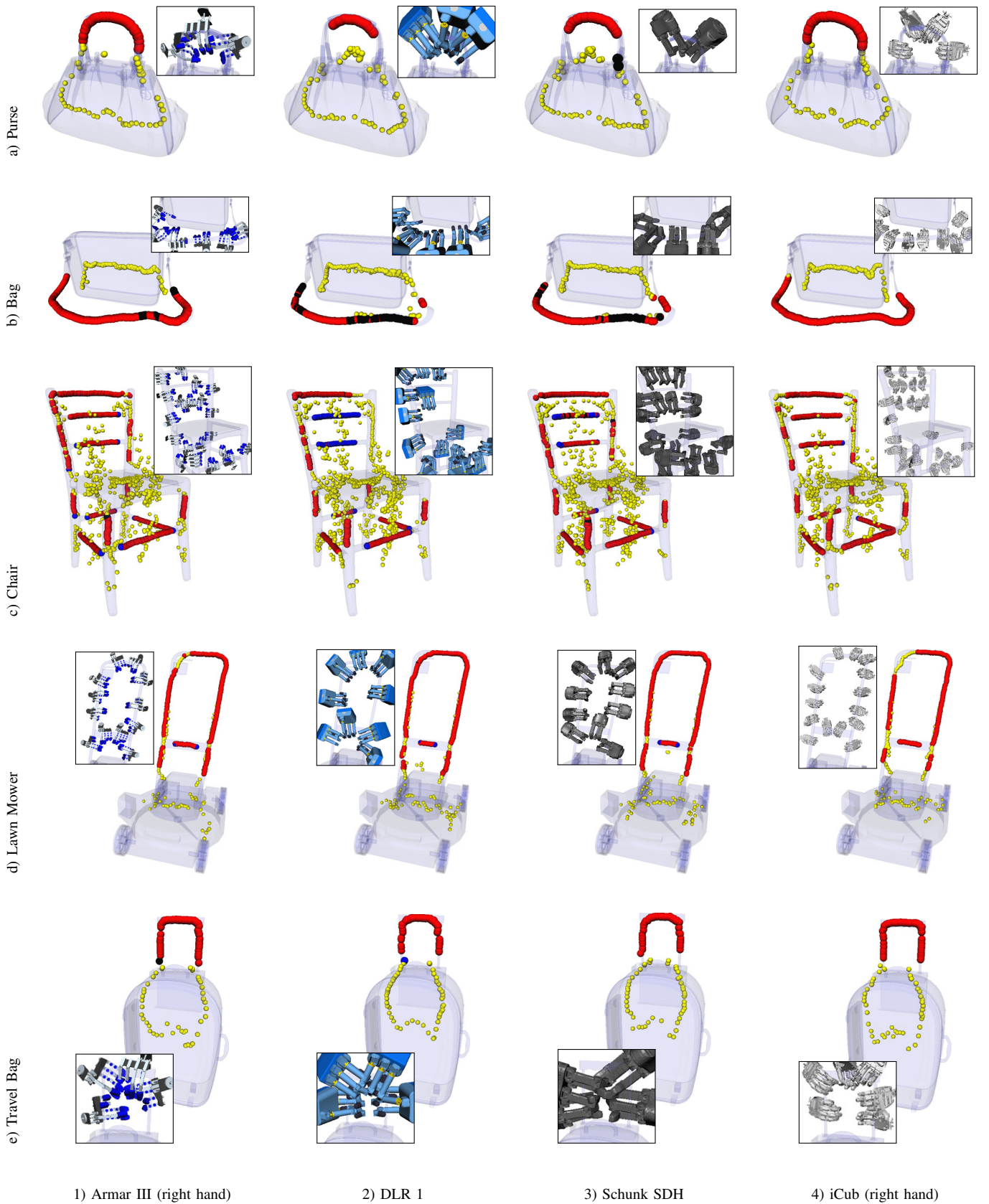


Fig. 5. Visualization of the experiments with synthetic data using a modified version of Simox [43]: For each pair consisting of a robot hand and an object, we display the generated grasp center points of a grasp frame (GF, see Sec. III-D and Sec. IV-A) and selected examples of successful grasps. The markers indicate positions of *rejected* GFs (yellow), *successful* GFs leading to clasp (red), *failed* GFs leading to separation (black), as well as GFs leading to positioning failures due to obstruction (blue). The colors here correspond to the bar plots in Fig. 4.



Fig. 6. Examples of the experimental evaluation performed with a Nao humanoid robot. *Top row*: The robot moves its right hand into approach position and subsequently grasps the handle. It stands up, lifts its arm and lets the basket slide into carrying position. *Bottom row*: The robot hooks and lifts the basket in one action by performing a grasp frame-parametrized motion template.

- [2] A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, "Automatic Grasp Planning Using Shape Primitives," in *IEEE ICRA*, 2003, pp. 1824–1829.
- [3] R. Pelossof, A. Miller, P. Allen, and T. Jebera, "An SVM Learning Approach to Robotic Grasping," in *IEEE ICRA*, 2004, pp. 3512–3518.
- [4] C. Borst, M. Fischer, and G. Hirzinger, "Grasping the Dice by Dicing the Grasp," in *IEEE/RSJ IROS*, 2003, pp. 3692–3697.
- [5] J. Seo, S. Kim, and V. Kumar, "Planar, Bimanual, Whole-Arm Grasping," in *IEEE ICRA*, 2012, pp. 3271–3277.
- [6] K. Hsiao, S. Chitta, M. Ciocarlie, and E. Jones, "Contact-Reactive Grasping of Objects with Partial Shape Information," in *IEEE/RSJ IROS*, 2010, pp. 1228–1235.
- [7] S. El-Khoury and A. Sahbani, "Handling Objects By Their Handles," in *IEEE/RSJ IROS 2008 Workshop on Grasp and Task Learning by Imitation*, 2008.
- [8] R. Detry, N. Pugeault, and J. Piater, "A probabilistic framework for 3D visual object representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1790–1803, 2009.
- [9] J. Glover, D. Rus, and N. Roy, "Probabilistic Models of Object Geometry for Grasp Planning," in *RSS*, 2008, pp. 278–285.
- [10] M. Przybylski and T. Asfour, "Unions of Balls for Shape Approximation in Robot Grasping," in *IEEE/RSJ IROS*, 2010, pp. 1592–1599.
- [11] M. Przybylski, T. Asfour, and R. Dillmann, "Planning grasps for robotic hands using a novel object representation based on the medial axis transform," in *IEEE/RSJ IROS*, 2011, pp. 1781–1788.
- [12] T. Dey and S. Goswami, "Tight cocone: a water-tight surface reconstructor," in *ACM symposium on Solid modeling and applications*, 2003, pp. 127–134.
- [13] T. Dey, J. Sun, and Y. Wang, "Approximating loops in a shortest homology basis from point data," in *ACM SoCG*, 2010, pp. 166–175.
- [14] F. T. Pokorny, J. A. Stork, and D. Kragic, "Grasping Objects with Holes: A Topological Approach," in *IEEE ICRA*, 2013.
- [15] A. Bicchi and V. Kumar, "Robotic Grasping and Contact: A Review," in *IEEE ICRA*, 2000, pp. 348–353.
- [16] A. Sahbani, S. El-Khoury, and P. Bidaud, "An overview of 3D object grasp synthesis algorithms," *Robotics and Autonomous Systems*, pp. 326–336, 2011.
- [17] M. Ciocarlie, "Low-Dimensional Robotic Grasping: Eigengrasp Subspaces and Optimized Underactuation," Ph.D. dissertation, Columbia University, School of Arts and Sciences, 2010.
- [18] M. Santello, M. Flanders, and J. Soechting, "Postural hand synergies for tool use," *The Journ. of Neurosc.*, pp. 10 105–10 115, 1998.
- [19] A. Bicchi, M. Gabbicini, and M. Santello, "Modelling natural and artificial hands with synergies," *Phil. Trans. R. Soc.*, pp. 3153–3161, 2011.
- [20] M. Gabbicini, A. Bicchi, D. Prattichizzo, and M. Malvezzi, "On the Role of Hand Synergies in the Optimal Choice of Grasping Forces," *Autonomous Robots*, pp. 235–252, 2011.
- [21] J. Romero, T. Feix, C. Ek, H. Kjellstrom, and D. Kragic, "Extracting postural synergies for robotic grasping," *Robotics, IEEE Transactions on*, vol. PP, no. 99, pp. 1–11, 2013.
- [22] C. Goldfeder, P. K. Allen, C. Lackner, and R. Pelossof, "Grasp Planning Via Decomposition Trees," in *IEEE ICRA*, 2007, pp. 4679–4684.
- [23] K. Huebner, S. Ruthotto, and D. Kragic, "Minimum volume bounding box decomposition for shape approximation in robot grasping," in *IEEE ICRA*, 2008, pp. 1628–1633.
- [24] S. Berretti, A. Del Bimbo, and P. Pala, "3D Mesh decomposition using Reeb graphs," *Image and Vision Computing*, pp. 1540–1554, 2009.
- [25] J. Aleotti and S. Caselli, "A 3D shape segmentation approach for robot grasping by parts," *RAS*, pp. 358–366, 2012.
- [26] —, "Part-based robot grasp planning from human demonstration," in *IEEE ICRA*, 2011, pp. 4554–4560.
- [27] —, "Manipulation planning of similar objects by part correspondence," in *IEEE ICAR*, 2011, pp. 247–252.
- [28] D. Attali, J. Boissonnat, and H. Edelsbrunner, "Stability and computation of medial axes – a state-of-the-art report," *Mathematical foundations of scientific visualization, computer graphics, and massive data exploration*, pp. 109–125, 2009.
- [29] E. Rimon and A. Blake, "Caging 2D bodies by 1-parameter two-fingered gripping systems," in *IEEE ICRA*, 1996, pp. 1458–1464.
- [30] C. Davidson and A. Blake, "Caging planar objects with a three-finger one-parameter gripper," in *IEEE ICRA*, 1998, pp. 2722–2727.
- [31] P. Pipattanasomporn and A. Sudsang, "Two-finger caging of concave polygon," in *IEEE ICRA*, 2006, pp. 2137–2142.
- [32] S. Makita and Y. Maeda, "3D multifingered caging: Basic formulation and planning," in *IEEE/RSJ IROS*, 2008, pp. 2697–2702.
- [33] A. Rodriguez, M. Mason, and S. Ferry, "From caging to grasping," *The International Journal of Robotics Research*, 2012.
- [34] W. Wan, R. Fukui, M. Shimosaka, T. Sato, and Y. Kuniyoshi, "Grasping by caging: A promising tool to deal with uncertainty," in *IEEE ICRA*, 2012, pp. 5142–5149.
- [35] R. Diankov, S. Srinivasa, D. Ferguson, and J. Kuffner, "Manipulation planning with caging grasps," in *IEEE-RAS Humanoids*, 2008, pp. 285–292.
- [36] A. Hatcher, *Algebraic topology*. Cambridge University Press, 2002.
- [37] O. Busaryev, T. Dey, J. Sun, and Y. Wang, "ShortLoop Software for Computing Loops in a Shortest Homology Basis," Software, 2010.
- [38] T. K. Dey, K. Li, and J. Sun, "On computing handle and tunnel loops," in *Cyberworlds, 2007. CW'07. International Conference on*. IEEE, 2007, pp. 357–366.
- [39] R. Ghrist, "Barcodes: the persistent topology of data," *Bull. Amer. Math. Soc.*, vol. 45, no. 1, p. 61, 2008.
- [40] Archive3d.net, <http://www.archive3d.net/>.
- [41] Trimble 3D Warehouse, <http://sketchup.google.com/3dwarehouse/>.
- [42] D. Berenson, R. Diankov, K. Nishiwaki, S. Kagami, and J. Kuffner, "Grasp Planning in Complex Scenes," in *IEEE-RAS Humanoids*, December 2007, pp. 42–48.
- [43] N. Vahrenkamp, T. Asfour, and R. Dillmann, "Simox: A Simulation and Motion Planning Toolbox for C++," 2010.
- [44] M. Madry, D. Song, and D. Kragic, "From Object Categories to Grasp Transfer Using Probabilistic Reasoning," in *IEEE ICRA*, 2012, pp. 1716–1723.
- [45] Y. Bekiroglu, D. Song, L. Wng, and D. Kragic, "A Probabilistic Framework for Task-Oriented Grasp Stability Assessment," in *IEEE ICRA*, 2013.