Rope Caging and Grasping

Tsz-Ho Kwok¹, Weiwei Wan², Jia Pan³, Charlie C.L. Wang⁴†, Jianjun Yuan⁵,
Kensuke Harada² and Yong Chen¹

Abstract—We present a novel method for caging grasps in
this paper by stretching ropes on the surface of a 3D object. Both topology and shape of a model to be grasped has been
considered in our approach. Our algorithm can guarantee
generating local minimal rings on every topological branches
of a given model with the help of a Reeb graph. Cages and
grasps can then be computed from these rings, and physical
experimental tests have been conducted to verify the robustness
of our approach.

I. INTRODUCTION

Caging, as compared to grasping, is non-dynamic, topo-
logical, and therefore robust to uncertainty. Given the noisy
point clouds or the noisy results of model-based estimation,
caging can find finger positions or configurations that con-
strain the target. The target may move inside the cage formed
by the fingers, but it will never escape from the cage. To cage
an object, we do not need to know its physical properties like
mass and inertia matrix.

Despite the advantages, caging has a major disadvantage
that the caged target is not immobilized and its exact position
and orientation are unknown. Caging can be used to catch
and constrain something, but it is difficult to be used to
assemble components or operate tools.

A. Motivation

The solution to overcome these disadvantages is grasping
by caging. For example, the workers in Fig.1 are picking and
placing wine port using caging. In the first step, they drop a
circular rope and wrap it around the thinner part of the port.
The port is caged by the rope after the actions of dropping
and wrapping. In the second step, they stretch the circular
rope by lifting it up and make the rope firmly in contact
with the port’s surface. While the first step only cages the
port and the port is not immobilized, the second step grasps
the object and enables the worker to place the port down to
an expected position.

Motivated by this example, in this paper we propose a
grasping by caging approach based on ropes. The seminal
work of grasping by caging was done by Rodriguez [1]
and Wan [2]. Rodriguez extended the idea of caging and
proposed F-caging. Given a finger formation, Wan, on the
other hand, recovered its configuration space and computed
the optimized configuration that could cage an object and
at the same time could shrink to grasping. Both papers pro-
posed interesting ideas but neither demonstrated successful
applications. Rodriguez’s work was more theoretical than
practical, and is more about a concept than some rigid-
form solutions. Wan’s work is applicable to 2D objects and
was demonstrated with 2D pick-and-place and multi-robot
transportation.

Unlike these works, our approach concentrates on ropes
and studies using rope to cage and grasp 3D objects with
slim waists or with topological holes. We employ reeb
graph to extract the position of slim waists or the holes
of object models, initialized some caging rings using the
extracted mesh segments, and estimate the stable grasps by
stretching the initial rings like a tapeline. We ensures caging
by requiring a finger ring be larger than a stretched caging
ring, and estimates the object’s position and orientation by
assuming that the grasping from this caging will converge
to the stretched caging ring. Our work can be used to plan
the grasps of rope or rope-like end-effectors. We demonstrate
the performance of our method with both a rope and a soft
gripper in the experimental tests.

B. Related Work

Grasping research in literature can be classified into two
major groups: force-closure and caging. Earlier works are
mainly based on the force-closure methods (e.g., state-of-the-
art simulators such as GraspIT and OpenGRASP). Recently,
approaches to realize stable grasping by a caging grasp
become more and more popular (ref. [1], [2]).

An interesting work is to use caging grasp to manip-
ulate articulated objects with handles [3]. With the help of
computing a topological loop on a given surface [4],
Stork et al. [5] can plan motions for clasping on objects
with handles. Our approach presented in this paper is more
general – an object to be grasped by our rope caging is not

The authors thank Zhongfeng Qian and Zheng Wang from the University
of Hong Kong, and Hao Zhang from Dorabot Inc. for assistance with the
design and fabrication of the soft gripper used in the experiment.

¹T. Kwok and Y. Chen are with the Epstein Department of Industrial
and Systems Engineering, University of Southern California. ²W. Wan and
K. Harada are with National Inst. of AIST. ³J. Pan is with the Department
of Computer Science, University of Hong Kong. ⁴C. Wang is with the
Department of Design Engineering, Delft University of Technology. ⁵J.
Yuan is with the Robotics Institute, Shanghai Jiaotong University
†Corresponding Author; Email: c.c.wang@tudelft.nl.

Fig. 1. With the help of ropes, objects can be stably caged in real world.
Our work is motivated by this caging method with a long history.
necessary to have a handle (e.g., the Venus model shown in Fig.6 and the spray container model shown in Fig.9(a)). Recently, more and more researches in robotics start to employ the geometric information of a 3D object. Curvature-based feature has been used by Calli et al. [6] to identify most proper place for grasping. They used concave points of a 2D elliptic Fourier descriptor of an object. Gaussian curvature is employed in [7] to conduct the segmentation of object for the grasping analysis. Tsuji et al. [8] modeled objects with multiple quadric surfaces and generated grasping posture by selecting the constricted parts. Zarubin et al. [9] proposed an idea of using geodesic balls on the object’s surface in order to approximate the maximal contact surface between a grasp and an object. Two types of caging grasps are developed: circle caging and sphere caging, where circle caging is similar to our rope caging. However, unlike our approach, their method does not provide the capability of capturing either the topological extremity such as handle or the geometric extremity like bottle-neck shape.

Reeb graph is a topological analysis tool that was previously employed to conduct a variety of applications in computer vision and graphics, including model decomposition [10], [11], parameterization [12] and shape matching [13]. In our approach, its concept is borrowed to analyze whether a topological branch on the given object is missed when constructing rings by uniformly sampling a scalar field. An algorithm is developed to ensure every branch has one ring. Note that, we do not need to construct the graph representation but only identify the critical points on a Reeb graph in our approach. Details can be found in Section III-A and III-B.

C. Contribution

The technical contribution of our work presented in this paper can be summarized as follows.

- We introduce the idea of stretching ropes on the surface of a 3D object to analyze the possibility of caging the model using a ring. Both topology and shape of a model to be grasped have been considered.
- A novel algorithm has been developed to generate minimal rings on every branches of the given model, which is guaranteed by the topology analysis with the help of a Reeb graph.
- Our approach has been evaluated in both simulation and physical experimental tests, where a rope and a soft gripper are employed in experimental tests.

Although the topology analysis based on a topological loop has been conducted in prior work (e.g., [5]) for grasping, to the best of our knowledge, this is the first approach that considers the minimal rings on all handles and branches.

Not only the topology but also the geometry information have been considered in our approach. As a result, the rope caging technique developed in this paper can be applied to models with very simple topology (e.g., genus zero models) if a local shape extremity can be touched by the elastic rope, which is similar to the real situation as shown in Fig.1. Cages and grasps can be computed from the rings generated by our algorithm (Section IV). Physical experimental tests have been taken to verify the validity of this approach on models with different shape and topology (Section V).

II. BACKGROUND

We assume each object to be caged has a corresponding 3D model represented by a polygonal mesh in the computer system. Following the notation in [14], a polygonal mesh \( \mathcal{M} \) is defined as a pair \((\mathcal{K}, \mathcal{V})\) where \( \mathcal{K} \) is a complex specifying the connectivity of vertices, edges and faces (i.e., the topological graph of \( \mathcal{M} \)) and \( \mathcal{V} = \{v_1, \ldots, v_n\} \) is the set of vertices defining the shape of polyhedral surface in \( \mathbb{R}^3 \). To simplify the implementation, every polygonal face in \( \mathcal{M} \) is subdivided into triangles. From \( \mathcal{K} \), it is very easy for our algorithm to get the adjacent vertices, edges and faces of a vertex in constant time; the same, the left/right faces of an edge and the three nodes/edges of a triangle can also be obtained by constant time cost. The geometric coverage of our method is confined to the domain of two-manifold polyhedral surface. Specifically, for every point on the surface of a two-manifold, there exists a sufficiently small neighborhood that is topologically the same as an open disk in \( \mathbb{R}^2 \). Such representation of a real object can be obtained via 3D reconstruction (e.g., KinectFusion [15]). In our approach, elastic ropes will be constructed and then actively moved (i.e., shrunk) on a mesh model \( \mathcal{M} \). Each rope is stored as an ordered list of surface points as \( \mathcal{L} = \{a_1, \ldots, a_m\} \), where a surface point could be located on an edge or inside a face of \( \mathcal{M} \). They are stored as attribute curves attached on \( \mathcal{M} \). Details can be found in [16]. This type of data structure can prevent pulling/pushing an elastic rope during the process of ‘stretching’ (i.e., \( \mathcal{L} \subset \mathcal{M} \) is ensured). The time-consuming step of collision detection can be avoided.

A Reeb graph [17] is a fundamental structure that encodes
the topology of a shape, which can be obtained by tracking the evolution of the level sets of a real-valued function $f$ on a manifold. Given a 3D model $M$, its topology can be represented by a Reeb graph with nodes denoting the change of topology and all surface regions on the same branch sharing the same topology (see Fig. 2 for an illustration). Every node on the Reeb graph corresponds to the critical level sets $f^{-1}(c)$. The pattern in which a branch meets a critical point reflects the change in topology of the level set $f^{-1}(t)$ as $t$ passes through the critical value $c$. By tracing all the critical points on a Reeb graph, all branches of a given model can be trimmed out by the level sets according to the critical values. The basic idea of our rope caging/grasping approach is to generate a few elastic ropes on each branch (see the blue cross-sectional rings in Fig. 2) and stretch them into the minimal length. The resultant rings on the surface of a given model will be employed to supervise the caging and thereafter grasping. The analysis based on Reeb graph can help us ensure no topological branch is missed when constructing the initial rings.

III. METHODOLOGY

This section presents the methodology of computing local minimal rings for rope caging and grasping. The topology analysis that guarantees not missing any branch of a given model $M$ is taken with the help of a geodesic distance field $G_M$. Critical points are found by a flooding algorithm. With the help of critical points, initial rings are constructed for each branch and thereafter stretched into local minimal rings. Figure 3 shows an overview of steps in our algorithm.

A. Scalar Field Based Topology Analysis

Topology analysis is taken on the surface of a given model $M$ to avoid missing any topological branch when generating initial rings. The concept of Reeb graph is employed in our analysis; however, we do not really construct a Reeb graph but only find out its critical points on $M$.

First of all, a scalar field $G_M$ is constructed on $M$. One simple option as used in [13] is the height field of a model. However, it can easily miss the handles that are nearly orthogonal to the z-axis. Instead, Approximate Geodesic Distance (AGD) field is employed in our approach. Assuming $M$ has been properly oriented, all mesh vertices that are on the ground are located and assigned as sources of AGD. $G_M(p)$ ($\forall p \in M$) is then computed according to these sources. In our implementation, a Dijkstra’s algorithm with virtual paths is employed (ref. [18]). More precise methods such as [19], [20] can also be used here; however, $G_M$ in our approach is not necessary to be precise as it is only for the purpose of topology analysis.

After computing $G_M$, we conduct a flooding algorithm akin to [10] to find out the critical points on $M$. The $G_M$ score of a face is defined as the maximal value of $G_M$ on its three vertices. A flooded region $F$ is first initialized by faces touched the ground. Other faces are then sorted by their scores. Starting from the face with lowest score, faces are added into $F$ one by one. Every time when a new face $f$ is inserted, we check the topology of $F$’s boundary, $\partial F$, by counting the number of loops in $\partial F$ – denoted by $|\partial F|$.

- $|\partial (F \cup \{f\})| > |\partial F|$: $f$ is a splitting face;
- $|\partial (F \cup \{f\})| < |\partial F|$: $f$ is a merging face;
- $|\partial (F \cup \{f\})| = |\partial F|$: $f$ is not critical.

Both splitting and merging faces are considered as critical (see the illustration in Fig. 4). In a critical face, the vertex with largest field value is recorded as a critical point. When three edges of a critical face are all on the boundary of $F$ before inserting it, the face is located as the tail of $M$. A critical point on this type of face is called $T$-critical, and other type of critical points are called $C$-critical as illustrated in Fig. 2.

B. Computing Initial Rings

In this step, a finite set of candidate rings are generated on the model. Each ring is constructed according to an isovalue. Specifically, for generating rings for $G_M(p) = c$, the isocurve consists of line segments extracted on the
surface of $\mathcal{M}$ by first searching faces with vertex scores both smaller and greater than $c$ and then producing a line segment intersect two edges in each of these faces. The endpoints of a line segment must be located on the edges with its two endpoints greater and smaller than $c$. The position of intersection is determined by the linear interpolation. Using the local connectivity of $\mathcal{M}$ stored in $\mathcal{K}$, the line segments can be linked up into one or a few loops.

Given a maximal field value $g_{\text{max}}$ on $\mathcal{G}$, the simplest method to generate rings is by isocurves with uniformly sampled $k$ isovales. Specifically, isocurves

$$\mathcal{G}(\mathbf{p}) = \frac{i}{k+1} g_{\text{max}} \quad (i = 1, \ldots, k)$$

(1)

are extracted to form initial rings uniformly (as shown in Fig.3(b)). Rings on an isocurve are generated by the aforementioned method. However, this simple method may miss some topological branches on a given model with the complex topology (e.g., when a handle happens in the region between $\mathcal{G}(\mathbf{p}) = \frac{i+1}{k+1} g_{\text{max}}$ and $\mathcal{G}(\mathbf{p}) = \frac{i}{k+1} g_{\text{max}}$). A more sophisticated algorithm is developed below to avoid such scenarios.

After constructed rings by uniformly sampled isovales (e.g., $k = 50$ is employed in all our tests), rings are inserted on those missed branches. We build a list $\mathcal{C}$ by the field values on all critical points in an ascending order. For example, when a model has $n$ critical points found in the topology analysis step, the list is like

$$\mathcal{C} = \{0, c_1, c_2, \cdots, c_n \} \quad (\forall i, c_i < c_{i+1}).$$

Note that both $C$-type and $T$-type critical points are included in this list. The last critical point is a $T$-critical point, $c_n = g_{\text{max}}$. Defining $\tau = \frac{k+1}{g_{\text{max}}}$, we then check if $|\tau c_i| = |\tau c_{i+1}|$. If this happens, no ring has been constructed on the branches between $c_i$ and $c_{i+1}$ by uniform sampling. We insert a ring at the middle of these two critical points. Specifically, initial rings are constructed on the following isocurves

$$\mathcal{G}(\mathbf{p}) = \frac{1}{2}(c_i + c_{i+1}) \quad (\forall i, |\tau c_i| = |\tau c_{i+1}|)$$

(2)

These initial rings will be further stretched on the surface of $\mathcal{M}$ to catch the local minimal of its shape.

C. Stretching for Local Minimal Rings

In this step of our algorithm, each initial ring is iteratively shortened on $\mathcal{M}$ until reaching a local minimum. As a result, rings generated by this method can capture the features such as bottle necks and bumpy regions and use them to cage an object to be grasped. These analogies learned from realistic lead to a smart method for caging/grasping 3D objects.

For a rope $\mathcal{L}$ stored as a list of surface points $\{a_1, \cdots, a_n\}$, the method how it is constructed ensures that

1) all points are located on the edges (as the rings/ropes are generated from isocurves);
2) the segment between two neighboring points lies on the same triangle (i.e., there is no segment across an edge).

These nice properties of a rope’s representation lead to a simple algorithm for stretching $\mathcal{L}$ on $\mathcal{M}$.

Points in $\mathcal{L}$ are moved one by one to meet an equal angle condition [16]. As shown in Fig.5(a), the optimal position of point $a_i$ on the edge that results in a minimal length of $||a_{i-1}a_i|| + ||a_ia_{i+1}||$ is a point that makes $\phi_i = \phi_0$. This can be easily proved by flattening two triangles adjacent to the edge onto a plane together with the points. An operator pushing points of a rope to move on edges of $\mathcal{M}$ is called edge operator.

More complicated scenario may happen after applying edge operators for a few iterations – quite a few very short segments crowded around a mesh vertex (see Fig.5(b)). Then, the rope is pushed to either the left (when the left surface angle $\alpha_L$ is smaller than the right surface angle $\alpha_R$) or the right (if $\alpha_L > \alpha_R$). Such an operator is called node operator. Again, the purpose of this operator is to make $\mathcal{L}$ satisfy the equal angle condition. Detail analysis and the implementation of these two operators can be found in [16].

The edge operator and the node operator are repeatedly applied to each rope initially constructed from isocurves. The iteration is stopped when no point on the rope can be moved any more. When all points on a rope have been located in the same triangle, this rope is removed from candidate rings as it indicates that the given model cannot be caged by this rope. Lastly, overlapped ropes stretched from different initial positions are merged to remove redundancy. Examples can be found in Fig.6. The resultant ropes are the candidates to be used in caging and grasping. Each rope is sampled into a set of points together with their surface normals, which are passed to the grasper controller to take further action.
Fig. 7. An illustration of computing the positions of finger-tips, $f_0$ and $f_1$, for caging.

IV. COMPUTING THE GRASPS

Algorithm 1: Computing the caging grasp using the stretched ring.

Data: a mesh model $\mathcal{M}$ and a local minimal ring $\mathcal{L}_s$

Result: A list of triples $\mathcal{G}_{\mathcal{L}_s}$ where each element indicates the position and configuration of the caging gripper: $(p_r, R_g)$

$$(\mathcal{P}, \mathcal{N}) \leftarrow \text{SampleRing}(\mathcal{L}_s);$$

for each $(p_r, n_r) \in (\mathcal{P}, \mathcal{N})$ do

\begin{align*}
    x_p &\leftarrow n_r; \\
    z_p &\leftarrow \text{CrossProduct}(n_{next}, n_r); \\
    y_p &\leftarrow \text{CrossProduct}(z_p, x_p); \\
    R_p &\leftarrow \{x_p; y_p; z_p\}; \\
    \{p_g, R_g, f\} &\leftarrow \text{ComputeHandConf}(p_r, R_p, \mathcal{P}, \mathcal{N}); \\
    \text{if NOT IsColliding}(\{p_g, R_g, f\}, \mathcal{M}) \text{ then} \\
    \quad \mathcal{G}_{\mathcal{L}_s} &\leftarrow \{p_g, R_g\};
\end{align*}

To determine the grasps, each local minimal ring is sampled evenly to compute the approaching vectors and orientations of a caging hand. Let $\mathcal{P}$ and $\mathcal{N}$ denote the set of sampled points and their corresponding normals. Their elements are $p_r$ and $n_r$. A local frame $[x_p; y_p; z_p]$ is constructed at each sample point to let $x_p$ and $y_p$ align with the normal and the tangent of the ring respectively. We use them together with $p_r$ and all other points on the sampled ring to compute the position, the orientation, and the finger-tip positions $f$ of the hand. This is implemented in the ComputeHandConf function as follows.

- First, $R_p$ is directly used as $R_g$ — that is the orientation of the hand.
- Second, the position of the hand is assigned as

\begin{align*}
    p_g &= p_r - \tau_g x_p, \\
    \text{(3)}
\end{align*}

where $\tau_g$ is a threshold gap between the hand and the sample point. Note that we assume all normal vectors sampled from a mesh model pointing inwards.

- Lastly, the positions of finger tips, $f$, are computed. The vector consists of $f_0$ and $f_1$, which presents positions of the first and the second finger-tips respectively. They can be computed by

\begin{align*}
    f_i &= p_r + \tau_f x_p + \tau_i y_p \quad (i = 0, 1), \\
    \text{(4)}
\end{align*}

where the threshold $\tau_f$ is assigned with a value smaller than the length of the fingers to ensure caging.

\begin{align*}
    \tau_0 &= \max \limits_{p \in \mathcal{P}} ((p - p_r) \cdot y_r) - \tau_{\text{tip}} \\
    \tau_1 &= \min \limits_{p \in \mathcal{P}} ((p - p_r) \cdot y_r) + \tau_{\text{tip}} \\
    \text{(5)}
\end{align*}

$\tau_{\text{tip}}$ is a fixed value to ensure the closure. $\tau_0$ and $\tau_1$ make the distance between two finger-tips smaller than the diameter of $\mathcal{L}_s$ along the direction of $y_p$ and therefore guarantee a successful cage.

An illustration about the positions of hand and finger-tips with reference a local minimal ring is given in Fig.7. After computing the configuration of a hand for grasping, we further pass it to a collision checking function, IsColliding, to ensure caging by the hand at this ring is collision-free. Only the collision-free caging grasps are saved as results.

Fig.8 shows the results of three examples. The left image of each sub-figure shows the local minimal rings, their
samples, and the local frames $x_p$, $y_p$, and $z_p$ at each sample using red, green, and blue colors. The middle image shows the computed caging grasps, where the green ones are the results returned by Algorithm 1 and the red ones are collided and deprecated candidates. The right image shows a clear view of the caging grasps, where the collided candidates are not plotted. During physical experiments, we use the "top" one or two candidates (the stretched rings that has largest difference from its initial states) of the rings shown in the left image to test rope caging and use the grasp candidate shown in the right figure to test soft robotic grippers. A robot will first approach the object with fully opened hand. Then, it closes the hand to form the planned cage. Finally, it fully wraps the object and immobilize it at a stable grasp.

V. PHYSICAL EXPERIMENT

In this section, we demonstrate the effectiveness of caging grasps based on minimal rings by using both a rope and a soft gripper. The grasps are tested on two containers as shown in Fig.9, which are served as the representatives for the objects with bottle-necks and for the objects with handles, respectively. The candidate caging grasps for these two objects computed by our approach are marked by the green curves in Fig.9.

A. Experiments with a Rope

In the first experiment, we use a rope to cage and grasp these two objects. As shown in Figs.10 and 11, the rope will start from a loose cage or at an unstable grasp. But after being stretched, it will eventually converge to a stable grasp as computed by our algorithm. For objects with more than one stable caging grasps, for instance the detergent container in Fig.9(b), the stretching direction will determine the final configuration of rope grasping.

B. Experiments with a Soft Gripper

In the second experiment, we use a soft gripper as shown in Fig.12. The soft gripper consists of two fingers made of silicone rubber and a driving mechanism. When the soft finger is dragged by a cable fixed at the fingertip and goes through the inside of the finger, it will bend the finger inwards. When releasing the cable, the supporting board of the finger (the back side of the finger) will serve as a spring and resume the finger into its original position. The cable dragging the finger first goes through a spring tube, which is designed to guide the direction of the dragging force along the tube. The cable is connected to a linear motor, which can generate very large force for dragging the cable. With the help of position feedback mechanism, this mechanism can control the positions of fingers precisely.

The soft gripper is actually similar to a rope, though its shape cannot be arbitrarily changed but is non-linearly controlled by driving forces. However, we find the caging grasps computed by our approach still provides a good prediction about the final stable grasp for the test objects. As shown in Figs.13 and 14, the soft gripper starts from unstable grasps. After suitable shaking perturbation, the gripper will end up at grasping object with a stable pose. Since the soft gripper is not a perfect rope, the final stable grasp pose is
not exactly same as the prediction provided by our algorithm, but it is close. In addition, similar to the rope case, we can find that for objects with more than one caging grasps, which grasp is achieved by the gripper is determined by the shaking direction and the pose of the object, as shown in Fig.14.

VI. CONCLUSION AND DISCUSSION

A new strategy called rope caging and grasping has been developed in this paper. Our method starts from generating an AGD field on the surface of a 3D model to be grasped. Critical points of the model’s Reeb graph according to the AGD field are extracted to segment the model into topological branches. Initial rings are then constructed on all branches by the isocurves of the AGD field. To capture the local shape extremities on the given model, the initial rings are further stretched to become shorter and shorter in an iterative way. The resultant local minimal rings are employed as candidates of grasping to be further verified in the simulation. Physical experiments have been taken at the end of this paper to verify the performance of our approach.

The current implementation using C++ takes about 62 sec. to compute all the minimal rings for a model with 5k faces on a PC with i7-4700MQ 2.4GHz CPU. The long computation time is mainly caused by the unoptimized code for extracting initial rings. One of our future works is to speedup the computation in this step of initial ring generation. Another planned work is to test this rope caging and grasping work on an integrated platform with robotic arm and grippers so that both motion planning and grasping are considered and tested.

REFERENCES