Sim-to-real Transferable Object Classification through Touch-based Continuum Manipulation

Huitan Mao*, Junius Santoso*, Cagdas Onal, and Jing Xiao

Abstract It is important to investigate object perception for classification or recognition based on touch sensing, especially when robots are operating in darkness or the objects are difficult to capture by vision sensors. In this work, we present a new form of continuum manipulator equipped with sparse touch sensing, validate the effectiveness of automatic generation of the touch-based continuum wraps, and the effectiveness of object classification based on the continuum wraps. Using the indirect object shape information encoded in the robot shape, we demonstrate that a classifier trained from the simulated continuum wraps is transferable to identify the real world objects with real continuum wraps.

Key words: Continuum manipulation, tactile sensing, object perception.

1 Introduction

In order to broaden the real-world application of robotics, the importance of advancing the capability of robotic manipulation is unquestionable. Most of the existing research in robotic manipulation assumes that the target object is known or visible, or can be made visible by adding a light source. However, in many real-world scenarios, such as underwater or in a smoke-filled room (in a building on fire), a target object may not be made visible. Therefore, it is necessary to investigate robotic manipulation based on other sensing modalities.

Tactile or force sensing can be very useful in providing more information about surrounding objects. Indeed, there is recent research on using tactile sensing for object recognition [1], exploration and manipulation [2, 3, 4], and shape estimation [5, 6]. Touch sensors are usually put on the finger tips of the grippers, which are

H. Mao is with Dept. of Computer Science, UNC Charlotte, e-mail: hmao4@uncc.edu; J. Santoso, C. Onal, and J.Xiao are with Robotics Engineering, WPI, e-mail: jsantoso@wpi.edu.
*The authors equally contributed to this paper.

used to touch a target object to collect contact points. Object recognition is achieved solely based on the collected contact points in [1] or in combination with visual perception [7]. Such touch-based perception is especially useful for perceiving transparent objects, which can be missed by visual sensing.

Continuum manipulators are more flexible for manipulating objects of a wide range of shapes and sizes in very cluttered environments due to their compliant and soft nature. In particular, a continuum manipulator can get into a small hole to fetch an object that a conventional, articulated manipulator cannot [8]. Since it is often dark in such a hole and difficult to bring in an extra light source, touch-based continuum manipulation is most desirable. If the target object is not visible, then object perception (i.e., classification and/or recognition) is also needed from touch-based continuum manipulation.

In our recent work [9], we introduced a shape-based approach for object classification and recognition through continuum manipulation. The main idea is to use the continuum robot as a tool to indirectly "measure" the object shape. That is, the shape of a continuum arm during whole-arm wraps of a target object, which can be transparent (and thus not visible), is used to indirectly characterize the shape of the object and this information is used for classification and recognition. For an object of any shape, a continuum wrap is generated automatically by a touch-based approach. However, the work is tested only in simulation with the assumption that the continuum arm is covered by tactile sensors. In reality, existing continuum manipulators are often not equipped with tactile sensors. One difficulty is that the body of the robot is deformable. A kinematics-based contact detection and localization approach for continuum robots is presented in [10], for which an external tracking system is required but may not always be available in real-world scenarios.

In this paper, we study touch-based identification of object categories using a new form of continuum manipulator consisting of origami-based modules [11] and tactile sensors attached at each section. Using these new manipulators with sparse tactile sensing, we aim to achieve touch-based wrapping of objects and experimentally validate the following conjecture: the shape-based classifier introduced in [9] and trained in simulation can be readily transferred to classifying real objects with real touch-based continuum manipulation. We envision the robot to be used in a search and rescue scenario where it could be exposed to low-light environment hence could benefit from the touch-based manipulation.

2 Technical Approach

We next explain the manipulators we built for this study, sensors, and the touch-based motion planning strategy for generating continuum wraps, and the classification method of unknown objects.

2.1 Manipulators and Sensors

The continuum manipulators we built for this study consist of multiple origami continuum modules connected in series. Each continuum module contains a foldable origami body, three brushed DC micro-motors with pulley systems, and a controller board that offers on-board sensory measurements, feedback control, and module-to-module communication. The foldable body is made out of Polyethylene terephthalate (PET) films and constructed based on the Yoshimura crease pattern. This unique tubular structure with a diameter of 7 cm is capable of bending in two directions and extending/retracting, while maintaining its structure and resisting torsion. The foldable structure is connected to an acrylic plate on the top and to the PCB on the bottom where the motors are secured. Three nylon cables secured to the motor shafts and spanning the length of the structure along the edges are used to drive the segment. Each motor is equipped with a magnetic encoder for position control.

We built two continuum manipulators for this study. Fig. 1(a) shows the 3-section manipulator used in the experiments with planar wraps, and each section has an active white module with motors on it and a passive green foldable body to expand the robot workspace. Improved from the 3-section manipulator, Fig. 1(b) shows the 4-section manipulator used for generating spatial wraps benefiting to its more compact and lightweight modules and larger touch sensor contact areas.

Each section is characterized using three parameters (s, κ, ϕ), where s is the section length, κ is the curvature, and ϕ is the orientation angle [8, 11]. Using the inverse kinematics of a continuum section developed in [11], we can then find the required cable lengths (l_1, l_2, l_3), that will shape the module into the desired configuration (s_d, θ_d, ϕ_d). The cable lengths are converted into encoder positions, which are then sent to the low-level controller as reference signals.

We constructed the touch sensors acting as bumper switches using two copper sheets adhered onto a parallel plate structure made out of the same material used for the origami collapsible body. One copper sheet is connected to the control board's digital I/O pin while the other is connected to ground reference. When the copper sheets touch each other due to depression of the structure, an electrical circuit is completed hence signaling a touch on the continuum section. For each continuum section we place the touch sensor at the module and sandwiched between consecutive modules as shown in Fig. 1.

2.2 Touch-based Continuum Wrapping

A general touch-based motion planning strategy is introduced in [9] for a continuum manipulator to progressively generate wraps around an object under the guidance of the contacts made along the way without knowing the object model. Starting from the initial configuration, the robot motion alternates between the *enclosing motion step* to make contacts with the object, and the *advancing motion step* to move forward towards wrapping around the object, until a wrap is formed or no further

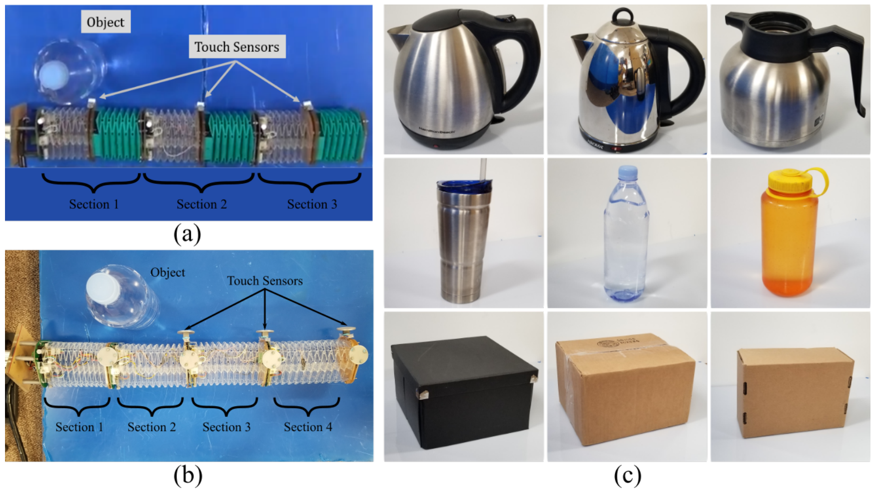


Fig. 1 An object (a transparent water bottle) was placed near the base of arms for manipulation: (a) a touch sensor was mounted in the middle of each origami module in a 3-section arm, (b) a touch sensor was mounted at the distal end of each module in a 4-section arm. The diameter of the continuum section is 7 cm. (c) A list of objects used in the experiments.

motion is feasible due to the physical limits of the manipulator. Such continuum wraps can be efficiently generated within hundreds of milliseconds (time of planning and collision checking combined) in simulation [9].

To achieve an *advancing motion step*, as described in [9], contact localization and estimation of the tangential and normal directions of the local contact patches are required; whereas, we relax this requirement in this paper through extrapolating the robot section endpoints based on their local frames. Therefore, our planner only needs to know whether each manipulator section is in contact or not from the tactile sensing to plan the next move, which makes it more effective to guide the manipulator hardware to achieve touch-based continuum wraps.

2.3 Classification of Unknown Objects

As introduced in Section 1, we aim to experimentally validate a shape-based classifier using continuum wraps [9]. The shape of the continuum robot wrapping around an object is described by a chord histogram descriptor, which approximates the robot shape using many 3D chords and statistically captures its shape based on the chord parameters. We first trained a linear SVM classifier in simulation using wraps generated around the simulated objects and next applied the trained model to classify the real-world objects using real-world wraps. The objects in simulation were scaled to roughly match the dimensions of the real-world objects.

3 Experimental Results with Planar Wraps

In our experiments with planar wraps, we used the 3-section manipulator as shown in Fig. 1(a). The robot manipulator is initially set at a straight-line configuration with full contraction, and the testing object is placed near the arm base. For each section, $s \in [0.085, 0.145](m)$, $\kappa \in [-10.69, 10.69](1/m)$, and $\phi \in [-\pi, \pi]$.

A linear SVM classifier is trained in simulation using wraps from 10 water bottles, 10 boxes, and 6 teapots. For each object one wrap was generated. Fig. 2 shows a few planar continuum wraps around the objects in simulation. We next conducted classification of three different real-world objects: a teapot, a water bottle, and a card box, through real-world continuum wraps.

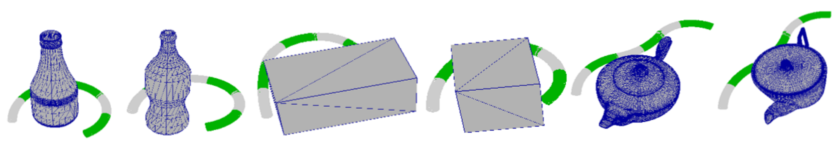


Fig. 2 A few examples of the planar continuum wraps generated on different objects in simulation. The simulated arm has 3 sections and each section is colored in white for the first half and in green for the second half.

The attached video shows the wrapping process of the real-world objects, and Fig. 3 shows a few motion snapshots. Note that the goal of such wraps is to encode the shape of an unknown object into the shape of the manipulator, as opposed to achieving tight grasps of objects with known models [8]. Therefore, the wraps do not need to be enclosing.

The final robot configurations commanded by the planner were used for classification. Table 1 summarizes object and wrap information and the classification results. It can be seen that the classifier learned solely from simulation is already effective in classifying the real-world objects, as the probabilities of correct classification are more than two times higher than that of a random guess (about 0.33). However, the box was mis-classified to be a bottle because the contact on section 1, which was closest to the base, was missed by the sensor, and therefore the planner kept commanding section 1 to bend more while it was actually stopped by the contact. Since the curvature of section 1 is a distinctive feature for classification as shown in Fig. 3, the classifier with the inaccurate data of a larger curvature resulted in the misclassification.

4 Experimental Results with Spatial Wraps

We use the 4-section robot arm (Fig. 1(b)) to conduct spatial continuum wraps around the objects to collect spatial shape information. The arm sections of this

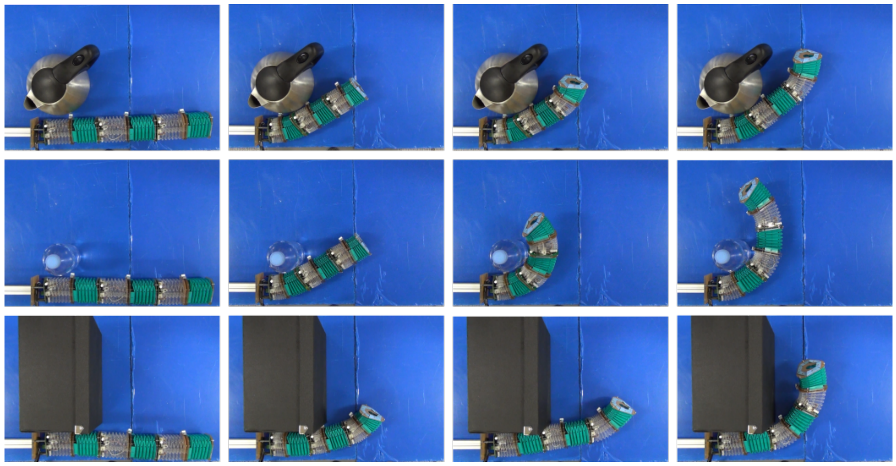


Fig. 3 The motion snapshots of the continuum wraps generated on the real objects. The rightmost sub-figure in each row shows the final wrapping configuration.

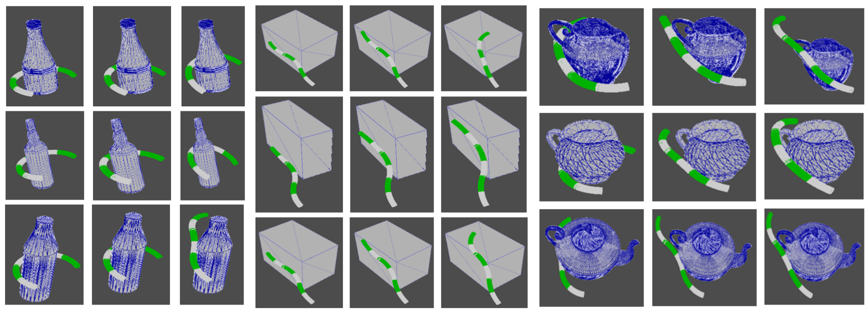
Table 1 Object dimension, number of intermediate configurations to generate the wraps, SVM prediction and its probability using 1 planar wrap for each object.

Object	dimension(cm)	# of configurations	SVM Prediction	Probability
bottle	$8 \times 8 \times 24$	59	bottle	0.81
box	$27 \times 16 \times 26$	61	bottle	0.83
teapot	$24 \times 18 \times 21$	42	teapot	0.71

robot are more compact (rather than having two modules connected as one robot section) and lightweight, which makes it more suitable for spatial wraps. It also overcame the problems of missed contacts with increased contact areas of the touch sensors. Different wraps covering different areas on the objects were conducted to collect object spatial shape information. The 3D chords generated from different wraps are accumulated into one histogram as an overall representation of the object shape. We next explain how the robot arm is lifted to conduct spatial wraps and present the object classification results.

4.1 Robot Arm Lifting

The arm is lifted up by keeping the first module of the robot at a certain configuration during the experiment and then changed when switching to other wrapping planes. See Fig. 5(b) for such examples when the robot arm is lifted up. The benefit of having this additional module is that the other modules (section 2, 3 and 4) can be fully used for generating the wraps, and they only need to undergo less strain and stress when wrapping around the objects. Alternatively, the modules could be mounted vertically to minimize the effect of gravity; however this configuration cannot always be possible in a real life scenario, hence we decided to proceed with the former.



(a) bottles (b) boxes (c) teapots

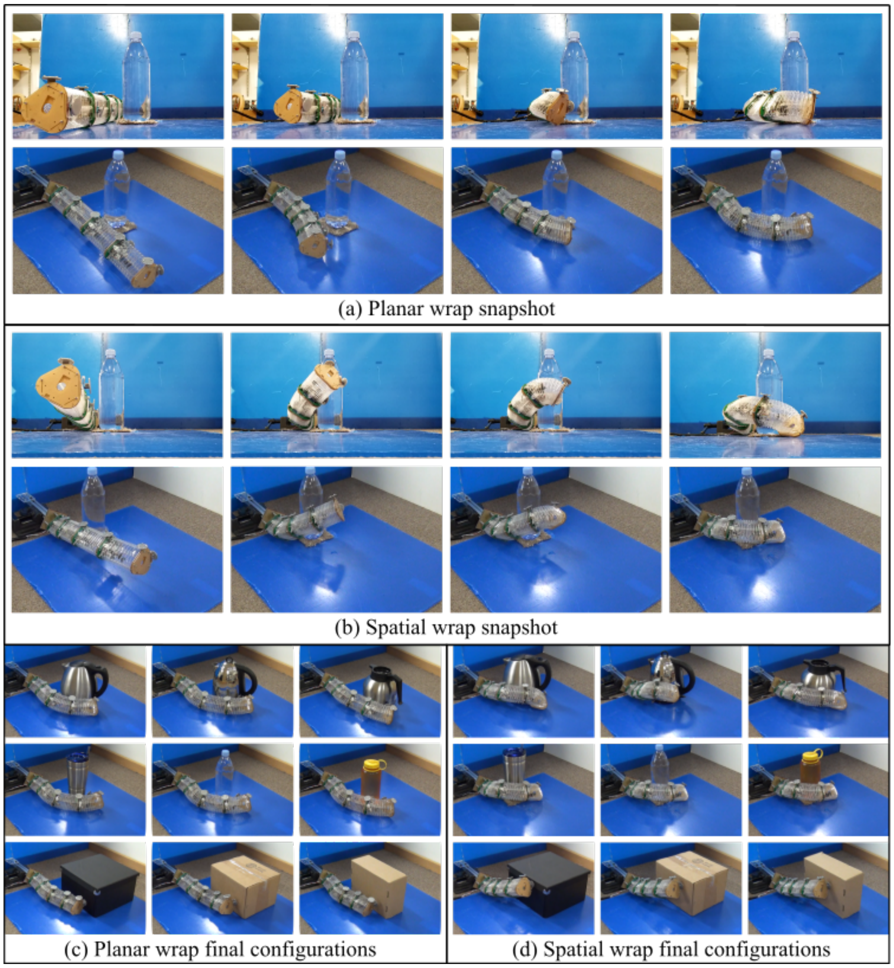
Fig. 4 The spatial wraps in simulation. Each row in each subfigure is the three wraps around the same object. The simulated arm has 3 sections and each section is colored in white for the first half and in green for the second half.

4.2 Classification Results

We trained a linear SVM classifier using the same set of training objects used in section 3, and the shape of each object was captured using 3 wraps. Fig. 4 shows the wraps around the training objects generated in simulation. Fig. 5 shows examples of such planar and spatial wraps around the real world objects. For testing dataset, we considered 3 object categories and 3 objects from each category (Fig. 1(c)).

The attached video shows the motion process of the continuum wraps. As mentioned earlier, such wraps are the result of local motion generation, and they are just used to encode the object shape onto the robot arm shape. Therefore, some (for instance the wraps around the boxes) may only locally conform the robot shape to the object shape and achieve partial wrapping of the object.

We noticed that the robot final wrapping configurations deviate from the robot motion commands sent by the motion planner (major cause for the mis-classification of the box in section 3). There are two main reasons. First, our robot arms currently do not have the proprioceptive sensors to achieve precise closed-loop control. Second, some contacts may be missed by the current sparse touch sensing on our robot arms. Therefore, in order to more precisely identify the final robot shape, we used an external vision tracking system to identify the robot final configuration when



(a) Planar wrap snapshot

(b) Spatial wrap snapshot

(c) Planar wrap final configurations

(d) Spatial wrap final configurations

Fig. 5 The motion snapshots of (a) planar and (b) spatial continuum wraps around the bottle, where the top and bottom rows show the motion of the same wrap from two view angles respectively, (c) final configurations for the remaining objects with planar wraps, (d) spatial wrap final configurations.

the wrapping process is terminated. This wrapping configuration is next used for generating the 3D chords and conducting the final object category classification.

Table 2 summarizes the objects used, the average number of robot configurations to generate the wraps, and the SVM prediction results. Overall, the classifier was able to correctly recognize 2 bottles, 3 boxes, and 2 teapots. The boxes are easier to be correctly identified since the robot arm conforms to the side surfaces of the boxes and therefore has distinctive straight sections in contact. The wraps of the bottles and the teapots typically have more curved robot sections but differ in lengths due to their dimensions. The confusion of classifying the bottles and the teapots is because sometimes the handles of the teapots (more distinctive features) are not captured. This can be improved by using longer robot sections and more dense touch sensing.

Table 2 Object dimension, average number of intermediate configurations to generate one wrap, SVM prediction and its probability using 3 spatial wraps for each object.

Object	dimension(cm)	avg. # of config.	SVM Prediction	Probability
bottle1	$8 \times 8 \times 24$	45	bottle	0.52
bottle2	$9 \times 9 \times 21$	43	teapot	0.56
bottle3	$8 \times 8 \times 21$	40	bottle	0.51
box1	$27 \times 16 \times 26$	12	box	0.74
box2	$25 \times 21 \times 17$	10	box	0.71
box3	$27 \times 10 \times 18$	15	box	0.78
teapot1	$24 \times 18 \times 21$	36	teapot	0.46
teapot2	$21 \times 15 \times 21$	40	teapot	0.51
teapot3	$22 \times 16 \times 16$	38	bottle	0.51

5 Experimental Insights

Our results have demonstrated that the shape-based classifier trained solely from simulation is able to generalize to real-world objects. This confirms our two key insights. First, because object classification is based on the shapes of the continuum arm wrapping around the objects and not the shapes of the objects directly, the classifier has the advantage of avoiding direct sensing and perception of the shape of an unknown target object as well as the associated limitations (such as low object visibility) and all the sensing uncertainties involved that can negatively affect classification accuracy. Second, the continuum wraps generated on objects in the same category have similar shapes, which are captured by the intrinsic parameters of the continuum arm, no matter if the objects wrapped are virtual or real.

Since conducting many real-world continuum wraps can be time-consuming, it is significant that the classifier trained purely in simulation showed considerable effectiveness in classifying real objects. This could make classifier training more efficient and feasible for classifying a large number of categories of many real objects from touch-based continuum wrapping.

6 Future Improvement

The overall system can be improved in multiple ways for better results and robustness. First, soft modules with longer length and more dense touch sensing can help better capture the unique features of the object shapes, for instance, the handles of the teapots. Moreover, more dense touch sensing can also help reduce the cases where the motion planner kept curving the robot sections while the actual contacts are blocking the robot motion. Second, more sophisticated gravity compensation should be considered in order to better lift up the robot arm and form spatial wraps around different areas on the objects. Third, since the final robot configuration is the result of both desired motion commands and the contacts with the object, the final shape of the robot can be better identified by equipping robots with proprioceptive sensors and feedback control rather than relying on an external tracking system.

The manipulators we built for this study are for object perception purposes and hence have limited payloads. Once the unknown objects are recognized via touch wraps, force-closure continuum graspings can be generated for further manipulation of the objects using the algorithms in [8]. We would also like to enable handling heavy objects by further improving the robot payloads.

Acknowledgements This work is supported by NSF grant IIP-1439695.

References

1. M. M. Zhang, N. Atanasov, and K. Daniilidis, "Active end-effector pose selection for tactile object recognition through monte carlo tree search," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.
2. N. Sommer, M. Li, and A. Billard, "Bimanual compliant tactile exploration for grasping unknown objects," in *International Conference on Robotics and Automation (ICRA)*, IEEE, 2014.
3. Q. Li, C. Schürmann, R. Haschke, and H. J. Ritter, "A control framework for tactile servoing.," in *Robotics: Science and systems*, 2013.
4. Z. Yi, R. Calandra, F. Veiga, H. van Hoof, T. Hermans, Y. Zhang, and J. Peters, "Active tactile object exploration using gaussian processes," in *IEEE/RSJ IROS*, 2016.
5. Y.-B. Jia and J. Tian, "Surface patch reconstruction from one-dimensional tactile data," *IEEE Transactions on Automation Science and Engineering*, vol. 7(2), 2010.
6. H. Mao and J. Xiao, "Object shape estimation through touch-based continuum manipulation," in *International Symposium of Robotics Research(ISRR)*, 2017.
7. P. Falco, S. Lu, A. Cirillo, C. Natale, S. Pirozzi, and D. Lee, "Cross-modal visuo-tactile object recognition using robotic active exploration," in *IEEE ICRA*, 2017.
8. J. Li and J. Xiao, "Progressive planning of continuum grasping in cluttered space," *IEEE Transactions on Robotics(TRO)*, vol. 32(3), 2016.
9. H. Mao, M. M. Zhang, J. Xiao, and K. Daniilidis, "Shape-based object classification and recognition through continuum manipulation," in *IEEE/RSJ IROS*, 2017.
10. A. Bajo and N. Simaan, "Kinematics-based detection and localization of contacts along multisegment continuum robots," *IEEE TRO*, vol. 28(2), 2012.
11. J. Santoso, E. H. Skorina, M. Luo, R. Yan, and C. D. Onal, "Design and analysis of an origami continuum manipulation module with torsional strength," in *IEEE/RSJ IROS*, 2017.