

# Automatic Grasp System with Sensor Feedback Based on Soft Gripper

PeiChen Wu<sup>1</sup>, Nan Lin<sup>1</sup>, Yifan Duan<sup>1</sup>, Ting Lei<sup>1</sup>, Lei Chai<sup>1</sup>, Xiaoping Chen<sup>1</sup>

**Abstract**—In this paper, we present a robotic automatic grasp system that is capable of grasping objects with a wide range of sizes and shapes firmly and delicately. The two main components of grasping object are grasp selection and grasp execution. The key new feature of our proposal is that we define three grasp primitives which are able to handle diverse objects. In grasp selection process, the algorithm needs to determine a better grasp primitive for each grasp position. This phase does not rely on precise object model but just the local length and width information of object. For different grasp primitive, there are different sensors triggered. We utilize decision-making tree method to build a rules set which guides grasp execution. The pressure value of suction cups, tactile information and the drive motor angle are needed in training process. We capture over 3000 data groups and labels in total to train the rules set with cross validation method. The true positive rate and true negative rate are 92.74% and 97.84% respectively of decision-making tree method. Our results also show that the gripper can grasp diverse diameter paper cups delicately without crushing them. Finally, we display the ability of our system by grasping diverse objects automatically.

## I. INTRODUCTION

The autonomous operation ability for robotic platform has been constantly evolving in recent years. This promotes the robotic applications in many scenes. However, manipulating objects in unstructured environments still has many unresolved items because of the variant shapes and materials. This paper mainly focuses on two components of object grasp, which are grasp selection and grasp execution.

The grasp selection is defined as: find the grasp configuration when the robotic hand grasping a object. The grasp configuration usually is a combination of hand posture and position. The more likely the grasps can be executed, the better the grasp configuration is. As for grasp execution, when the hand configuration is achieved, the hand should grasp object firmly and delicately. That means, there is no slip for hard object and the soft object will not be crushed. So, this process is object-dependent and is intrinsically related to onboard sensors for robot hand.

Grasp selection has been widely studied since the last century. Bendiksen and Hager [1] use 2-D image data to find the grasp positions which achieve the rigid body equilibrium conditions. Their approach relies on object information including knowledge of the object center of mass, object mass and coefficient of friction. Miller *et al.* [2] get grasp position by analyzing the model of shape primitives. Their

approach simulates in *GraspIt!*, which relies on the object models. There are also many work using 3-D object model for planning grasp [3], [4], the quality of model decides the grasp performance. However, it is hard to get high-quality model for a real-time robot platform.

In contrast to the approaches that relies on priori information of object, the methods only using the sensor data to determine grasp position are developed rapidly in recent years. Jain *et al.* [5] use a laser range finder to separate objects out of background and then pick them up. Pelossof *et al.* [6] plan grasp position by the *SVM* learning method. In [7], the work uses template-based learning method to select grasp position. In order to deal the novel objects, they made an assumption that similarly shaped objects can be grasped with similar grasp configurations. When dealing with image data, the convolutional neural network always works well. Levine *et al.* [8] and Jang *et al.* [9] train large convolutional neural network to predict the grasp configuration by monocular images. However, these approaches will break down when uncertainty in location of one grasp happens. Because there is no other feedback information for the grasp execution.

The information that the grasp selection relied on is always static information of target object, range and imaging sensors are suited for providing these information. However, these sensors have shortcomings in grasp execution process. Because there is some inevitable position error when using these sensors. Errors can arise from imperfect calibration, environmental occlusion, or even imperfect control of the robot [10]. During grasping, the interaction of target object and robotic hand generates valuable information. Those signals generated from the interaction often have tiny amplitudes and high frequencies. So, we need some extra sensors to capture those information to guide the grasp execution.

Tactile sensors work well for providing the information which is grasp execution needed. Petrovskaya *et al.* [11] combine tactile sensor and known object model to grasp target object. In [10], the authors realizes contact-reactive grasp by combining tactile sensor and partial shape information. However, both of them dose not avoid slipping and crushing. Shaw-Cortez *et al.* [12] achieve stable blind grasping which avoids slipping of target object based on tactile sensor. But this work dose not take soft objects into account, which easily deforms during grasping and even is crushed. In [13], the work is most similar to ours in grasp execution process, the authors use tactile sensor to achieve the firm and delicate grasp. The ointment is that the paper dose not tell how to select the grasp position. Beside the tactile sensors, there are also other sensors served as guiding grasp execution. For example, a Resistor Network Structure Proximity sensor

\*This work is supported by the National Natural Science Foundation of China under grant 61573333 and U1613216.

<sup>1</sup> PeiChen Wu, Nan Lin, Yifan Duan, Ting Lei, Lei Chai and Xiaoping Chen are with the School of Computer Science and Technology, University of Science and Technology of China, Hefei, 230026, China. iewupeichen@163.com

TABLE I  
DEFORMATION FORCE FOR SOME COMMON OBJECTS

Object Type	Deformation Force(N)
Paper Cup	>1.9
Breakable Puffed Packing	>3.5
Plastic Water Bottle	>5.0

is used to make pre-shaping for various objects in [14]. In this approach, the surface of each fingertip is brought into contact with the object surface very well. However, it can not guarantee the firm and delicate grasp.

In our previous work [15], we have designed a novel sucked-type soft hand. In this paper, we realize the firm and delicate grasp with this soft hand. Inspired by human, we will find the hand position and a better grasp primitive for the soft hand instead of the configuration of finger. For a person, when grasp a object, he will determine which grasping mode should be applied subconsciously. For example, pinching mode will be chosen when the target is a pen. Accordingly, when grasping a cup, the enveloping mode is employed. And by utilizing the tactile sensor information, we are able to grasp objects without dropping and crushing them.

## II. EXPERIMENTAL SYSTEM

Our hand is mounted on a 6 DOFs robot arm produced by KINOVA ROBOTICS company. Kinect depth camera is used to capture image data and position information. Owing to pressure sensors connected to the suction cups [15], we can get pressure information from the vacuum pressure. If the vacuum pressure changes, that means the suction cups is covered. There are four groups of suction cup in different position for one finger, we can know the relative location between target object and robot hand. Besides, we may infer whether the grasp is firm through the value of vacuum pressure. However, it works only when the suction cup is almost covered. Another shortcoming of capturing air pressure data is that the value of vacuum pressure changes slowly, i.e. the frequency of this respond is low. However, in order to prevent crushing object, the hand must quickly responds when the deformation is in the wind. Therefore, we need introduce some other sensors to get contact information even when the object and sensor contact marginally. And the sensor is able to respond quickly.

We roughly measure the minimal deformation force of some flexible objects in daily life, so as to provide more information for selecting tactile sensor. The value of deformation force is shown in Table I. The measurement tool we used is NK-30 push-pull dynamometer produced by HAND-PI company. Considering structural characteristics and the limit of space of our hand, we decide to use tactile push button switch as our tactile sensor. The sensor is produced by OMRON company and the number is B3F-4005. The operating force of the sensor is  $2.55 \pm 0.69\text{N}$ , and the pretravel is around 0.3mm. The size is 12mm\*12mm\*4.3mm(length, width and height). The sensor can quickly respond even press on the edge. The sensors' positions on our hand are shown

in Fig. 1. Our all experiments run on the open-source ROS(Robot Operating System) software. Information of the tactile sensors and air pressure is sampled at 1 KHz.

## III. GRASP PRIMITIVES

In this part, we define three grasp primitives to achieve robust grasp for common objects in life.

- **First primitive:** grasping objects with fingertips. There are such a kind of typical objects in daily life like chopsticks, pens, marbles, peanuts etc. They are so small or slender that they probably slip from the hand if you simply enclose it with your palm. So, when grasping these objects, human mainly depends on fingertips as show in Fig. 2(a). Inspired by this behavior of human, we claim the grasp primitive which mainly depends on fingertips. What's more, the algorithm makes suction cups of fingertip adjusted to the center of object, which result in holding object without dropping off even the object is a bit heavy.
- **Second primitive:** grasping objects with palm plying-up. This primitive is suit for thin and flat objects like mobile phone and books, which may not be able to be caught with small force area. As is known to us, the barycenter of most regular object is in their geometric center. If you only grasp one side of it, it is likely to fall off in the other side due to torque, not to mention that the object is a bit heavy. For this type objects we do have to find a solution to make gripper to contact more with there surface. As luck would have it, it is because the object is thin so that it could be put as deeply as possible in the gripper as shown in Fig. 2(b). Therefore, the object has more enough contact area with the hand.
- **Third primitive:** grasping objects with fingers wrapping. The algorithm adjusts the distance between finger root and the surface of object closed to the gripper according to the specific width of the object. In spite of a large part of the gripper is in a suspended state (not touching the object) by reason of object's size, our gripper can definitely grasp the object tightly as shown in Fig. 2(c). Because our gripper has both grasping force and suction force, which is a excellent property. The

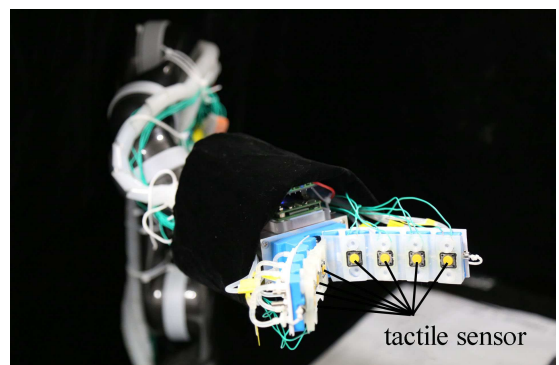


Fig. 1. Picture of our gripper with tactile sensors

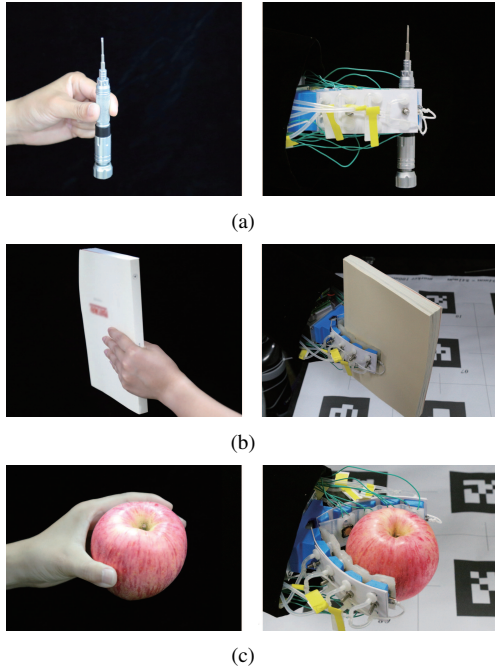


Fig. 2. Presentation photograph of human hand and our soft gripper with three grasp primitives. (a) Grasping screwdriver with first primitive. (b) Grasping book with second primitive. (c) Grasping apple with third primitive.

ingenious gripper structure endows the internal ability to grasp a large range of objects.

It's clear that, the size of object plays an important role in choosing grasp primitive. If the width and thickness are both small enough (width  $< 18\text{mm}$ , thickness  $< 36\text{mm}$ ), the gripper would choose the first grasp primitive. But when the object is wider and wider, only the second grasp primitive can solve the problem. In the same way, as long as the width of object is beyond threshold value, the gripper have to carry out the third grasp primitive to meet the demand regardless of how large the thickness is. Through three grasp primitives, our gripper can achieve the destination grasping most of the objects in daily life. And, different grasp primitive will definitely trigger different response combinations of sensors. This feature facilitates our feedback mechanism.

#### IV. GRASP SELECTION

##### A. Preprocessing

Using kinect camera and Point Cloud Library (PCL), we can obtain the Point Cloud of any objects with its Oriented BoundingBox. To simplify the description of our algorithm, we may as well create a new Cartesian coordinate system in the center of the BoundingBox of the object. Let the origin of the new coordinate system coincide be the midpoint of the BoundingBox and the XYZ axes be perpendicular to the three faces of the BoundingBox.

The core idea of our algorithm is to traverse the surface of the object from top to bottom in several directions to get the features demanded later. All the heights in all directions are possible grasp positions. For every possible grasp position,

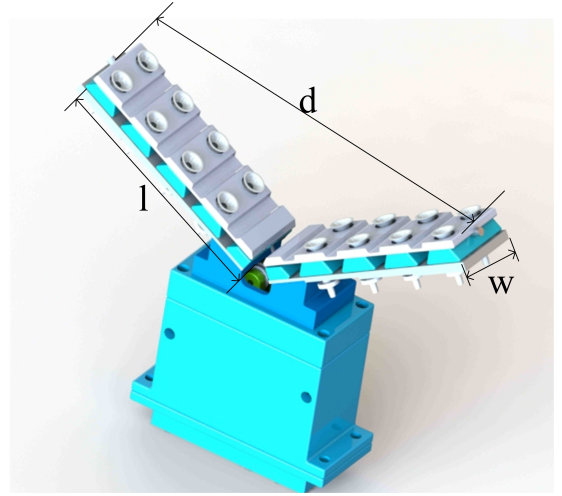


Fig. 3. CAD model of our soft gripper with some intrinsic parameters

we sort them to select the most suitable position for gripper to grasp and perform the most reliable grasp depending on the features.

Considering that humans prefer to grasp for which the wrist is oriented orthogonally to one of the object's principal axis, we assume that the grasping direction of the gripper is always orthogonal to an axis of the coordinate system. To simplify the description, we illustrate the algorithm with the x-axis as the grasp direction and the gripper plane perpendicular to the z-axis as an example.

Starting from the top of the BoundingBox, we slice the object from the outside to get the width of the object at equal interval. Having finished slicing at one height, the z coordinate value decrease at another equal interval, which is equal to half of  $w$ , where  $w$  denotes the thickness of the fingertip, which is showed in Fig.3. Then as the z coordinate decreases, we gradually acquire all the information we need in this direction.

Since we only use one kinect camera, we may not be able to know the exact width of the object. Here we can only calculate the distance between the center and the surface that can be observed by the kinect camera and then use it to take place of the width of the object. When the object is symmetrical, there is no deny that this method is effective. If not, a satisfactory success rate has also been achieved owing to the excellent performance of our gripper.

After enough samples of possible grasping position have been obtained, we process the data of the samples to get the features the algorithm demand.

##### B. Feature Acquisition

First of all, we explain some concepts which are used in this part as follow.

- **smoothness:** In order to make the suction cups on the gripper fully utilized, we define the smoothness as the concave and convex condition of the surface. The larger the value of smoothness is, the smoother the object is, the better the suction cup can be attached to the surface.

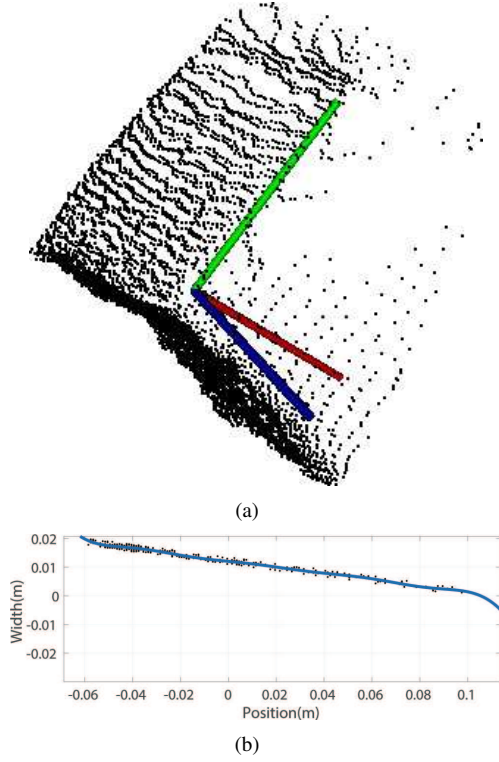


Fig. 4. Smooth example. (a) The PCL of a box and the coordinate system. (b) The width at different positions and fitted five-level Fourier series.  $s = 0.89$ , while the maximum value of the derivative is 0.11. Obviously, it is smooth. For labels in (b), the Position aligns the green axis and the Width aligns the red axis in (a).

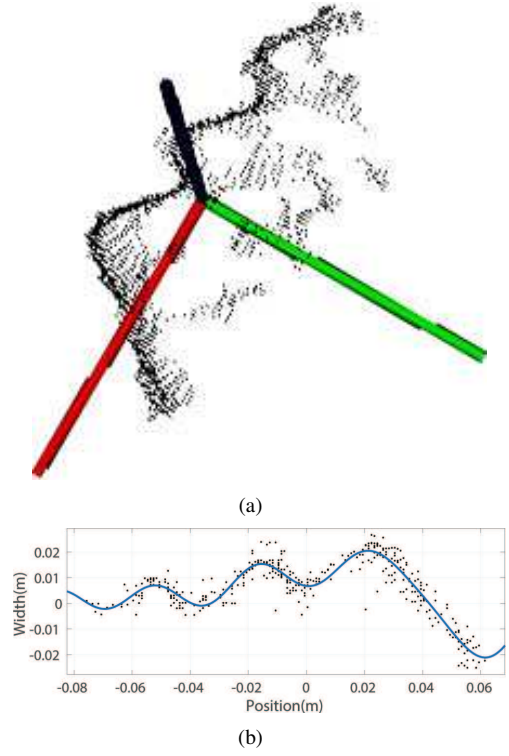


Fig. 5. Unsmooth example. (a) The PCL of a ladder and the coordinate system. (b) The width at different positions and fitted five-level Fourier series.  $s = -0.35$ , while the maximum value of the derivative is 1.35. Obviously, it is not smooth. For labels in (b), the Position aligns the red axis and the Width aligns the green axis in (a).

- **curvature:** If all goes well, our gripper can attach to the surface of the object because our gripper are soft and have excellent bending properties. So we figure out the value of curvature to describe the overall shape of the object. The difference from smoothness is that curvature describes the approximate shape of the object and smoothness focuses on the details of the surface of the object.

- **relative position:** The relative position to the center.

Then we will show the mathematics used to obtain these three parameters as follow.

First of all, we fit the width of a position obtained in the previous step to the form of a Fourier series:

$$f(x) = a_0 + a_1 \cos(xk) + b_1 \sin(xk) + a_2 \cos(2xk) + b_2 \sin(2xk) + a_3 \cos(3xk) + b_3 \sin(3xk) \dots \quad (1)$$

Study the absolute maximum of the second derivative of the function from the front to the end of the object. If the maximum is smaller than the threshold, it is considered smooth; else if not, it is considered rough. We define a constant  $s$  to describe the smoothness of the object:

$$s = 1 - (m/t) \quad (2)$$

Where  $m$  denotes the maximum and  $t$  denotes the threshold. Obviously, when the object is smooth following the definition,  $s > 0$ ; but if we find the  $s < 0$ , it means that the

object is not smooth. Here are two example. One is typical smooth shown in Fig. 4, the other is typical non-smooth shown Fig. 5.

Second, since we only need to know a rough shape, we fit the width of each position into a quadratic function which is relatively simple to see the variation tendency of maximum and minimum values in the specified interval ( $X_{max} - l, X_{max}$ ). Where  $l$  denotes the length of finger, which is shown in Fig. 3. There are five situations in total:

- maximum, minimum, maximum
- minimum, maximum, minimum
- maximum, minimum
- minimum, maximum
- No extremum

The latest case can be seen as not bending. For the other four cases, if there is a maximum value in the range, we can calculate a constant  $c$  to describe the curvature:

$$c = 1 - 4(m/d - 0.5)^2 \quad (3)$$

Where  $d$  denotes width between two fingertips after the gripper fully opens, which is shown in Fig. 3. Clearly, if the maximum is about half of the width, the value of  $c$  is largest. If the maximum is too large,  $c$  may be negative. If there is no maximum in the range, it can be seen as not bending.

Third, to make the grasp position as close as possible to the center of gravity of the object, we define a constant  $h$  to describe the relative position to the center. Since we can



TABLE II  
PARAMETERS AND THEIR DESCRIPTION

Parameter	Description
M	Grasping mode
T	Tactile information
P	Pressure value
A	Drive motor angle

not know where the object's center of gravity is, we use the center of the object to approximate the center of the object.

$$h = 1 - (Z - Z_{center}) / (Z_{max} - Z_{center}) \quad (4)$$

### C. Optimal Grasp Position Selection

So far, we have acquired all the information we need:  $s$ (smoothness),  $c$ (curvature),  $h$ (relative position). We use the formula  $w = \alpha s + \beta c + \gamma h$  to assign three parameters with different weights to sort the position. The first one is the optimal grasp position obtained by this algorithm.

Last but not least, depending on the width and thickness of the objection at this position, we can get a grasp primitive for each grasp position to grasp the object actually.

## V. GRASP EXECUTION

After the grasp selection, we get the target hand position and the grasp primitive. Then, the grasp execution starts. The core problem is how to stop closing the gripper duly, which can make the hand grips target objects firmly, without dropping, and delicately, without crushing them. Our approach for this utilizes information as more as we can to guide grasp execution. We can get the contact information from our tactile sensors. The value is binary, if a sensor is triggered, the value is 1. As for the air pressure, in our prior work [15], we find that the value of it will reach to  $-40\text{KPa}$  rapidly if there are something covering the suction cups. We binarize the pressure value: if the value of negative pressure is greater than  $-40\text{KPa}$ , we set the pressure value as 1. Otherwise, the value is set 0. Besides, we can also get the angle of the driving motor to make assistance in grasp execution.

The key point of the grasp execution can be described as determining whether we can grasp target object firmly and delicately according to the data we know. In other words, at each sampling moment, having a data group  $\{M, T, P, A\}$ , the autonomous algorithm need to make a correct decision with a set of rules. The description of the parameter in the data set is shown in Table II. We build the rules set using the decision-making tree method. Each property of decision tree is defined as below.  $T$  and  $P$  are expressed by 8-bit vector where every element is binary. As for continuous properties, we discretize them into finite sets. For example, the range of angles  $A$  of the driving motor varies from  $0^\circ$  to  $180^\circ$  and they are uniformly divided into 36 valid values.

## VI. EXPERIMENTAL RESULTS

In this section, we will display three experimental results. Firstly, the result of the decision-making tree mode is presented. Secondly, we use paper cups with diverse sizes to

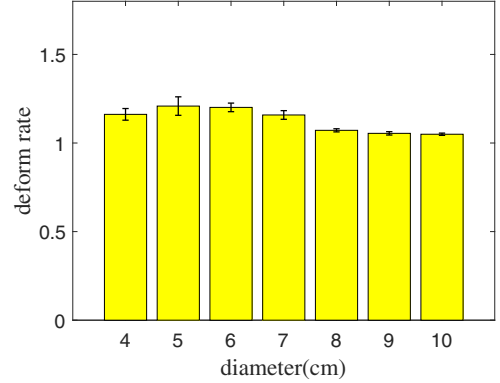


Fig. 6. Grasp experiments

test whether the hand can grasp delicately without crushing them. Finally, we perform automatic grasping with different objects.

### A. Result of Decision-Making Tree

During the grasp execution, if the gripper is able to hold target object, the label will be positive with the relevant data group. There is at most one positive label for once grasp execution. So, the negative labels are much more than the positive ones. We capture over 3000 data groups and labels in total. We using cross validation method ten times to get the rules set. For each train, the training set has 90% positive data and negative data. The average results are shown in Table III, where the values are rounding off. The true positive rate and true negative rate are 92.74% and 97.84% respectively. The false positives may caused by lacking sensors data because of target object's position error. And the false negatives more likely arise from the error labels. When capturing data, operators deem the gripper can hold target object in some moment, but it is not able to grasp firmly in fact.

### B. Deform Degree of Diverse Size Paper Cups

In this part, the gripper grasps diverse size paper cups whose diameter is from 4cm to 10cm. To quantize the deform degree, we introduce the deform rate for paper cups. In initial step, the paper cup is cycle looking from top. After grasping execution, if deformation occurs, the paper cup will look like ellipse from top. The deform rate is that the long axis of ellipse divides the short axis. The deform rate is always greater than 1. And the more severe deformation is, the larger deform rate is. For every paper cup, we test ten times to calculate the mean and standard variance of deform rate. The

TABLE III  
TABLE OF CONFUSION

Hypothesized Class	True Class		Total
	Positive	Negative	
	537	42	579
Positive	537	42	579
Negative	55	2498	2553
Total	592	2540	3132

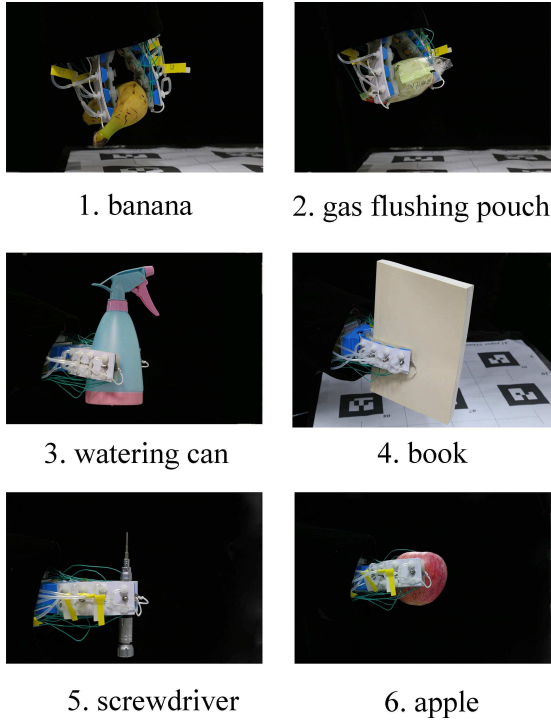


Fig. 7. Experimental results for grasping paper cups with diverse diameters

deform rate of all cups approximates 1 is shown in Fig. 6. That means the gripper can grasp soft objects delicately. And the value of standard variance shows the repeated accuracy is also well.

### C. Grasping Diverse Objects

We display the ability of our system by grasping diverse Objects automatically in this section. As shown in Fig. 7, it is evident that the gripper can grasp diverse objects which have wide range of sizes and shapes, from screwdriver to apple. Flat object such as book and irregular object like sprinkling are also caught. Moreover, the soft objects such as banana and gas flushing pouch can be grasped firmly and delicately.

## VII. CONCLUSION AND FUTURE WORK

This paper has presented a robotic automatic grasp system that is capable of grasping a wide range of sizes and shapes objects firmly and delicately. Typically, the algorithm firstly find the grasp position for a automatic grasp. Secondly, when the robotic hand has moved to the target position, the grasp execution starts. But for our proposal, there are some differences. For a human, when grasping a object, he will determine which grasp mode should be applied subconsciously. For example, pinching mode will be chosen when the target is a pen. Accordingly, when grasping a cup, the enveloping mode is employed. Inspired by this, we define three grasp primitives to handle diverse target object. So, in the grasp selection process, the algorithm needs to determine a better grasp primitive for each grasp position. To do this, the algorithm only need the local length and width information. Even when only partial object scans are

available, the algorithm still works. Once getting the grasp primitive, the grasp execution starts. The key point for this process is that when the hand should stop closing to grasp target object firmly and delicately according to the sensors' information. In this part, we use decision-making tree method to train a rules set. And the algorithm decides when to stop closing hand by the rules set. Our results show that our system has the ability to grasp diverse target objects without dropping or crushing them.

In our proposal, only when the target object is segmented, the algorithm starts work. In our future work, we would like to find the grasp position and primitive in a complex scene which may have redundancy objects, even some objects occlude. And in this work, there are also some false negatives and positives existing. To eliminate these effect, we will ascend the density of tactile sensors to provide more information during grasp execution. In other way, we will attempt to use other classifiers to guide the grasp execution.

## REFERENCES

- [1] A. Bendiksen and G. Hager, "A vision-based grasping system for unfamiliar planar objects," in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*. IEEE, 1994, pp. 2844–2849.
- [2] A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, "Automatic grasp planning using shape primitives," in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*, vol. 2. IEEE, 2003, pp. 1824–1829.
- [3] Z.-C. Marton, L. Goron, R. B. Rusu, and M. Beetz, "Reconstruction and verification of 3d object models for grasping," in *Robotics Research*. Springer, 2011, pp. 315–328.
- [4] Z. C. Marton, R. B. Rusu, D. Jain, U. Klank, and M. Beetz, "Probabilistic categorization of kitchen objects in table settings with a composite sensor," in *IROS*, 2009, pp. 4777–4784.
- [5] A. Jain and C. C. Kemp, "El-e: an assistive mobile manipulator that autonomously fetches objects from flat surfaces," *Autonomous Robots*, vol. 28, no. 1, p. 45, 2010.
- [6] R. Pelossof, A. Miller, P. Allen, and T. Jebara, "An svm learning approach to robotic grasping," in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, vol. 4. IEEE, 2004, pp. 3512–3518.
- [7] A. Herzog, P. Pastor, M. Kalakrishnan, L. Righetti, T. Asfour, and S. Schaal, "Template-based learning of grasp selection," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 2379–2384.
- [8] S. Levine, P. Pastor, A. Krizhevsky, and D. Quillen, "Learning hand-eye coordination for robotic grasping with large-scale data collection," in *International Symposium on Experimental Robotics*. Springer, 2016, pp. 173–184.
- [9] E. Jang, S. Vijaynarasimhan, P. Pastor, J. Ibarz, and S. Levine, "End-to-end learning of semantic grasping," *arXiv preprint arXiv:1707.01932*, 2017.
- [10] K. Hsiao, S. Chitta, M. Ciocarlie, and E. G. Jones, "Contact-reactive grasping of objects with partial shape information," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 1228–1235.
- [11] A. Petrovskaya, O. Khatib, S. Thrun, and A. Y. Ng, "Bayesian estimation for autonomous object manipulation based on tactile sensors," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE, 2006, pp. 707–714.
- [12] W. Shaw-Cortez, D. Oetomo, C. Manzie, and P. Choong, "Tactile-based blind grasping: A discrete-time object manipulation controller for robotic hands," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1064–1071, 2018.
- [13] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, "Human-inspired robotic grasp control with tactile sensing," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1067–1079, 2011.

- [14] K. Koyama, H. Hasegawa, Y. Suzuki, A. Ming, and M. Shimojo, "Pre-shaping for various objects by the robot hand equipped with resistor network structure proximity sensors," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 4027–4033.
- [15] N. Lin, P. Wu, X. Tan, J. Zhu, Z. Guo, X. Qi, and X. Chen, "Design and analysis of a novel sucked-type underactuated hand with multiple grasping modes," *unpublished*.