Research on Tactile-Gripping for Difficult-to-Grasp Objects

Yoshitaka Sakata
Kyushu Institute of Technology, 680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan

Gamolped Prem, Eiji Hayashi
Kyushu Institute of Technology, 680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan

Email: sakata.yoshitaka414@mail.kyutech.jp, gamolped.prem129@mail.kyutech.jp, haya@ics.kyutech.ac.jp

Abstract

This study focuses on the automation of food preparation and boxing in the food manufacturing industry. An important point of food grasping by robots is that the shape of the food should not be damaged. However, it is difficult for a conventional robot hand to perform this task perfectly. Therefore, an end-effector equipped with a camera-based tactile sensor has been developed to perform this task in previous studies. However, the performance of this end-effector depends on the reflectance of the target object, since it estimates contact based on the reflectance of light. We have developed a camera-based tactile sensor and contact estimation system to solve this problem. In addition, we have developed a pickup motion combined with object detection.

Keywords: Soft Robotics, Tactile Sensors, Deep Learning

1. Introduction

This study focuses on the automation of food preparation and boxing in the food manufacturing industry. In recent years, the need for robot automation has been increasing due to the expected worsening of labor shortages caused by the decline in the working-age population. Serving and boxing soft foods is a challenging task for robots. In a previous study, soft and deformable objects were defined as “difficult-to-grasp objects,” and an end-effector equipped with a vision-based tactile sensor was developed to perform these tasks. Although this end-effector has excellent ability to grasp difficult-to-grasp objects, it has a problem that accurate estimation of the grasping state becomes difficult when the light reflectance of the object is low. We developed a vision-based tactile sensor and grasp state estimation system to improve this problem. In addition, we have developed a pickup motion combined with object detection.

2. Robot Composition

The robot consists of a vertically articulated 6-axis arm, an RGB-D camera, and a hand tip equipped with a tactile sensor as shown in Figure 1.

The camera-based tactile sensor consists of a finger with Fin Ray structure, rubber-coated elastic skin, LED strip, and RGB camera as shown in Figure 2. Since it is necessary to grasp the object without crushing it, we designed it based on the structure proposed in [1]. For the placement of the elastic skin, LEDs, and cameras, we refer to [2] and [3]. The RGB camera takes pictures of the elastic skin and captures the deformation of the skin due to contact with an object. The contact is shown in Figure 3.

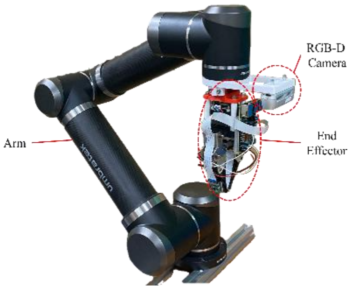

Fig.1 Robot Overview

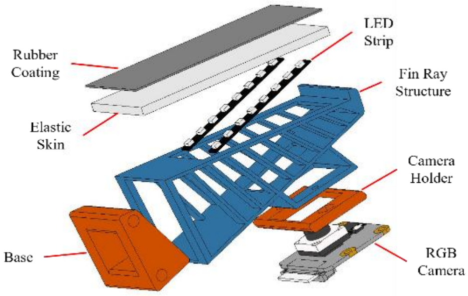

Fig.2 Composition of camera-based tactile sensor

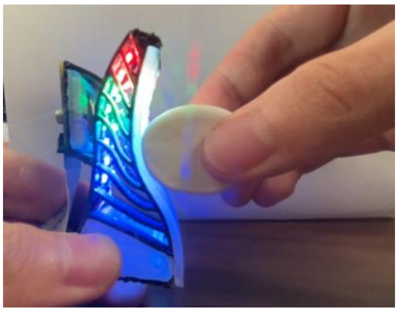

Fig.3 Deformation of tactile sensor due to object contact

3. System configuration

Figure 4 shows the system configuration. The system consists of object detection, grasp state estimation, motion planning, and paw control, and uses ROS (Robot Operating System) topic communication for inter-process communication.

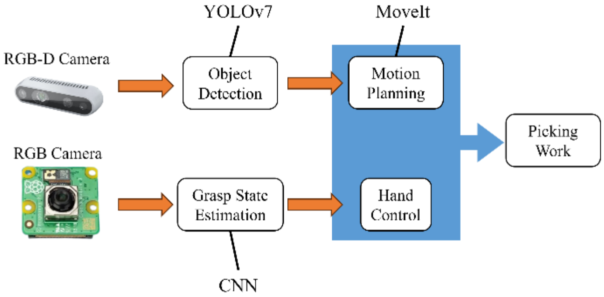

Fig.4 Pickup System Configuration

3.1. Object Detection

The algorithm uses YOLOv7, a real-time object detector. One hundred marshmallow images were used as the dataset, and 300 training runs were conducted. The algorithm was trained 300 times on a dataset of 100 marshmallow images. Object detection was performed using the trained model, and the 3D coordinates of the object were calculated by combining the center pixel of the detected object (Figure 5), depth information, and internal camera parameters.

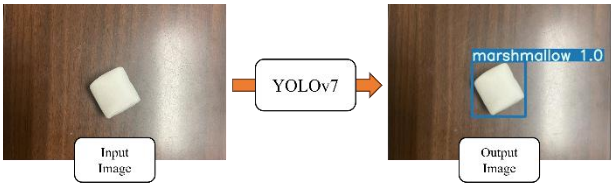

Fig.5 Object detection results

3.2. Motion Planning

The coordinates calculated by object detection are converted to a coordinate system based on the robot. MoveIt (motion path planning library) is used to plan the motion using this as the target position of the robot's hand.

3.3. Contact State Estimation Model

The grasping state estimation model is shown in Figure 6, where a convolutional neural network consisting of a convolutional layer, an all-joint layer, and an output layer was created using VGG-16 with RGB images as input. Here, the target of grasping is a marshmallow, and 214 images in the non-contacted state and 250 images in the contacted state were used for fine tuning.

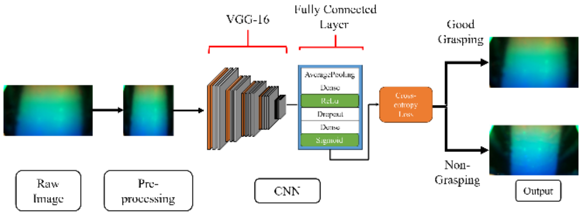

Fig.6 Contact State Estimation Model

3.4. End-Effector Control

The hand control using the estimation model is shown in Figure 7, where the video from the RGB camera is used as input for estimation. The model closes the fingers if they are not grasped and maintains the position of the fingers if they are in contact.

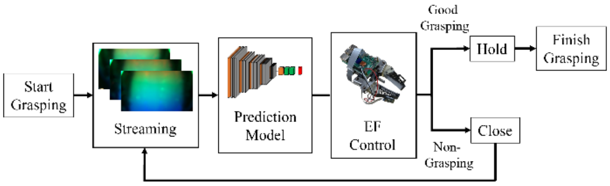

Fig.7 End-effector control using an estimation model

4. Experiment

Experiments were conducted to measure the performance of tactile sensors on cylindrical and ball-shaped marshmallows. Figure 8 shows an overview of the grasping experiment when the marshmallow was in contact with elastic skin areas A, B, and C, each about 11 mm in diameter. The size of the marshmallow before grasping and the distance between the fingers after grasping were measured 10 times each, and the amount of deformation of the marshmallow before and after grasping was calculated.

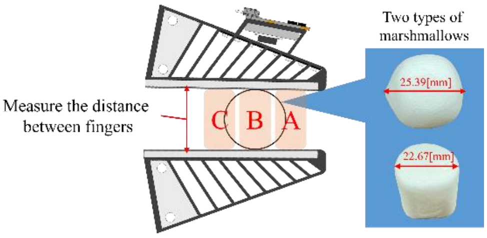

Fig.8 Experiment Overview

5. Results

Table 1 shows the average deformation and the ratio of deformation to the size of the marshmallow before grasping. It can be confirmed that the deformation of the ball-shaped marshmallow is larger than that of the cylinder-shaped marshmallow in the grasping of the ball-shaped marshmallow. The smallest deformation was observed in region C for both shapes. In the case of the cylindrical marshmallow, the deformations in regions A and B are almost the same, and in the case of the ball-shaped marshmallow, the difference between the deformations in regions A and B is large.

Table 1 Experimental Results

Contact Surface	Deformation (Cylinder)	Deformation Rate (Cylinder)	Deformation (Ball)	Deformation Rate (ball)
A	0.648	2.815	0.999	3.928
B	0.639	2.825	0.879	3.366
C	0.447	1.9997	0.827	3.359

6. Conclusion

We developed a vision-based tactile sensor and grasp state estimation system for grasping difficult-to-grasp objects. This system was then applied to the control of an end-effector to verify the performance of the tactile sensor on two different shapes of marshmallows.

The performance of the tactile sensor was verified for two shapes of marshmallows, and its effectiveness in grasping difficult-to-grasp objects was confirmed. Future work will include verification of the Fin Ray structure and the modification of the elastic skin and RGB camera positions. In addition, we aim to develop a deep learning system with multiple data inputs, such as a fingertip image from another camera in addition to the RGB camera image.

References

1. Emerson, J., Elgeneidy, K., (2020). Optimising Soft Fin Ray Robotic Fingers using Finite Element Analysis to Reduce Object Slippage. UKRAS20 Conference: “Robots into the real world” Proceedings, 43-45.
2. S. Q. Liu and E. H. Adelson, "GelSight Fin Ray: Incorporating Tactile Sensing into a Soft Compliant Robotic Gripper," *2022 IEEE 5th International Conference on Soft Robotics (RoboSoft)*, Edinburgh, United Kingdom, 2022, pp. 925-931.
3. S. Q. Liu, Y. Ma and E. H. Adelson, "GelSight Baby Fin Ray: A Compact, Compliant, Flexible Finger with High-Resolution Tactile Sensing," 2023 IEEE International Conference on Soft Robotics (RoboSoft), Singapore, Singapore, 2023, pp. 1-8.

Authors Introduction

Mr. Yoshitaka Sakata

He received bachelor’s degree in Engineering in 2024 from Intelligent and Control Systems, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Gamolped Prem

He received bachelor degree in Engineering in 2021 from Robotics and Automation Engineering, King Mongkut’s University of Technolgoy North Bangkok in Thailand and received masters in 2022 from Kyushu Institute of Technology at Hayashi Laboratory. He is currently a PhD student in Kyushu Institute of Technology at Hayashi Laboratory and conducts research about Robotics and AI.

Prof. Eiji Hayashi

Prof. Eiji Hayashi is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).