2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Visual Feedback Control of an Underactuated Hand for Grasping Brittle and Soft Foods

Ryogo Kai,[1] Yuzuka Isobe,[1] Sarthak Pathak,[2] and Kazunori Umeda[2]

Abstract— **This paper presents a novel method to control an underactuated hand by using only a monocular camera, not using any internal sensors. In food factories, robots are required to handle a wide variety of foods without damaging them. To accomplish this, the use of underactuated hands is effective because they can adapt to various food shapes. However, if internal sensors such as tactile sensors and force sensors are used in the underactuated hands, it may cause a problem with hygiene and require complicated calibration. Moreover, if external sensors such as cameras are used, it is necessary to grasp foods without damaging them by using external information such as images. In our method, to tackle these problems, a camera is used as an external sensor. First, contact between the hand and the object is detected by using the contours of both, obtained from a camera image. Then, to avoid damaging the object, the following information is extracted from camera images and observed: the centroid of both the hand and object, the deformation of the object, and the occlusion rate of the hand. Furthermore, to prevent the object from dropping while the robotic arm is in motion, the distance between the centroid of the hand and the object is calculated. The experiments were conducted using twelve different food items.**

I. INTRODUCTION

Research on the use of robots to handle deformable objects is advancing [1]. One example of a deformable object is food. Due to variations in shape and hardness based on different types of food, handling food with robots can be challenging. Therefore, in food factories, human labor is necessary. However, in aging societies like Japan, there is a concern about declining productivity due to labor shortages. Therefore, robotic automation such as robot manipulation in food factories is attracting increasing attention [2], [3].

In food manipulation, the following four factors are considered crucial: (i) handling various types of foods, such as brittle and soft foods; (ii) ensuring hygiene; (iii) enabling easy installation in factories; and (iv) avoiding damage to the food items.

To fulfill (i), an underactuated hand, which can adapt to the object's shape, is used. Since the underactuated hand directly touches the food items, it is important to ensure hygiene as mentioned in (ii). However, when using internal sensors like tactile or force sensors within the hand to manipulate objects, it can be difficult to keep the hand clean because cleaning it with heat or water may cause the sensors to fail. Moreover, if internal sensors are used on the hand's surface, repeated contact with the object may break the sensors, which can pose a risk of foreign-object contamination. The use of internal sensors affects not only (ii) but also (iii). When using any sensors, calibration is necessary. If a large number of hands equipped with internal sensors are introduced into multiple factories, calibration will be required for each hand, and the system installation may take a long time. To fulfill (ii) and (iii), it is better to use external sensors, such as cameras, instead of internal sensors. Humans use hands to grasp objects. However, they use information obtained through eyes to determine if an object is not too crushed or if it has been dropped. Thus, the robot can be controlled to satisfy (iv) by using images from the camera, just as humans use their eyes. When using a camera, markers are often used to observe the hand pose and the object's deformations. However, if markers are attached to the hand or the object, it may cause a problem with (ii) and (iii). Therefore, the object must be grasped without any markers.

An additional difficulty in food grasping is that foods have different amounts of hardness (i.e., there are brittle and soft foods). If the hand does not grasp with the force appropriate to the hardness, the object may be dropped or damaged. Therefore, it is necessary to detect whether an object is hard or soft and to grasp it without damaging or dropping it.

In this paper, we propose a method to control an underactuated hand by using only a monocular camera, without any internal sensors or markers. To grasp the object without damage, both the hand and the object are observed via an image. First, contact between the hand and the object is detected based on the contours of both. Next, to avoid grasping with excessive force, how the hand is moving and how the object is deforming are continuously observed via the image using four indexes: the hand's centroid, the object's centroid, object deformation detected from the optical flow, and the occlusion rate of the hand. Furthermore, to prevent the object from dropping during the robotic arm motion, the distance between the centroid of the hand and the object is observed via an image. The effectiveness of the proposed method is verified through experiments using real foods. The novelties and contribution of the proposed method are as follows:

- To achieve easy installation of the system and hygienic handling of various types of food, an underactuated hand is controlled using only a monocular camera as an external sensor.
- An object is grasped without damaging it, even without its hardness information, by observing the movement

[1]The Precision Engineering Course, Graduate School of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan. {kai, isobe}@sensor.mech.chuo-u.ac.jp

[2]The Department of Precision Mechanics, Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan. {pathak, umeda}@mech.chuo-u.ac.jp

of the underactuated hand and the deformation of the object.

II. RELATED WORKS

A. Robotic hand

To handle a wide variety of food items such as brittle and soft foods, robotic hands with high adaptability are being developed. One example is a soft hand [4]–[7]. This hand is made of soft material and has a high degree of freedom, allowing it to grasp objects gently while adapting to the object's shape. Another example is an underactuated hand [8]. This hand has fewer degrees of freedom compared to the soft hand, but it can also adapt to the object's shape. In this paper, an underactuated hand is used as a basic study of object grasping by a soft hand.

B. Sensor

Many studies of object grasping using soft hands or underactuated hands are using internal sensors. For example, tactile sensors or force sensors have been attached to the surface of the hand for force feedback control [9]–[12]. These studies may be able to prevent the hand from grasping objects with excessive force because contact force can be obtained from sensors mounted on the robot hand. However, the problem is that the force can only be measured in the contact area. When the contact area is small, only local forces and deformations of the object can be detected, and the deformation of the object may not be fully detected, which may result in damage to the object. In addition, several studies used cameras as internal sensors called vision-based tactile sensors [13]. She *et al.* [14] and Liu *et al.* [15] installed a camera inside the hand to perform hand-posture estimation and object classification. Other studies have used vision-based tactile sensors with parallel hands [16], [17]. These methods can acquire not only the hand's information, such as its posture and grasp position, but also details about the object itself, such as its type and shape. However, in the case of grasping food, food particles or dirt may adhere to the hand, making it impossible to acquire information from the camera.

Furthermore, as mentioned above, using internal sensors poses challenges for maintaining hygiene because it is difficult to clean the hand with heat or water. Additionally, repeated contact of the sensor with the food may cause sensor parts to be mixed in with the food. Another concern is that it takes time to introduce the hand into food factories. Since all sensors should be calibrated, if multiple sensors are used for a single hand, and a number of such hands are used in a factory, the time required for calibration becomes significant.

Cameras are often used as external sensors, the counterpart to internal sensors. Barrie *et al.* [18] used deep learning to estimate the grasping force based on the deformation of the soft gripper obtained from the image. Grady *et al.* [19] proposed a method that uses images and deep learning to estimate the pressure distribution when a soft hand contacts a horizontal plane, such as desk, which can be used to help grasp the object. However, these methods using deep

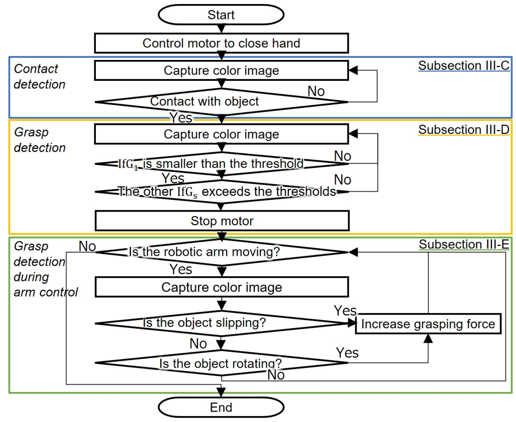


Fig. 1. Flow chart of the image feedback control process.

learning have the problem of needing a prepared dataset for the specific hand. Morgan *et al.* [20] recognized the interaction between the hand and the object by attaching markers to them. Nguyen *et al.* [21] attached markers to the object to detect its deformation. These methods are useful for recognizing the state of both the hand and the object, respectively, from an image. However, since it is impossible to put a marker on the food, it is important to control the hand without attaching any markers to the hand or the object.

As described above, the challenge is to handle food without using internal sensors or markers.

III. METHODS

A. Assumption

In this research, a two-finger underactuated hand is used. No internal sensor is attached to the hand, only a monocular camera is used as an external sensor. Brittle and soft foods are used, which are challenging for a robotic hand to grasp without causing any damage. Moreover, to consider cases where the same type of food has different hardnesses, no information about the object's hardness is known in advance. The hand is controlled with actuators, such as servo motors and pneumatic actuators. The hand closes at a constant speed from an open position and grasps the object in place. During this control process, visual feedback control is performed.

B. Overview

In our method, three types of detections for the image feedback control of the underactuated hand are introduced: *contact detection, grasp detection,* and *grasp detection during arm control.* Refer to the flow chart in Fig. 1 to understand the step-by-step process of the method. In the first step, *contact detection* is executed. Here, the contours of both the hand and the object are obtained from an image. Then, contact is detected based on the overlap of the contours. In the next step, *grasp detection* is performed. To grasp an object without excessive force to avoid damaging it, IfG$_s$ (Indexes for Grasping) are used. IfG$_s$ contain the following four indexes: centroid of the hand, centroid of the object, deformation of the object, and occlusion rate of the hand. Using $\mathrm{IfG_s}$, it is observed from the image how the hand moves and how the object is deformed. After *grasp detection*, the robotic arm is controlled to start moving. Finally, *grasp detection during arm control* is performed to detect the slipping and rotation of objects during the movement of the robotic arm. If the object is about to slip during the arm movement, the grasping force of the hand is increased to avoid dropping the object. With these three types of detection, the underactuated hand can grasp objects of various hardnesses and sizes without dropping or damaging them. In the following sections, we will provide details on the three types of detections.

C. Contact detection

The robotic hand starts in the open state. The contact between the hand and the object when the robot hand is closed is detected from an obtained image. First, the input image as shown in Fig. 2 is converted to a Hue Saturation Value (HSV) image. Next, as shown in Fig. 3, the regions of the hand and the object are obtained from the HSV image using the threshold process. Then the contours of the hand and the object are obtained from Fig. 3. In Fig. 4(a), the contours of the hand and the object are represented by blue and red lines, respectively. As the hand closes in Fig. 4(a), the contours of the hand and the object overlap. In Fig. 4(b), the pixels where the lines overlap are highlighted in magenta. If the number of overlapped pixels exceeds the threshold, contact between the hand and the object is detected. After contact is detected, the system proceeds to *grasp detection*.

D. Grasp detection

In this section, whether excessive grasping force is being applied to the object is detected so as not to damage the object, regardless of its hardness. When grasping a hard object that does not deform, the underactuated hand comes to a halt after contacting the object. On the other hand, when grasping a soft object that deforms after contact, the hand may continue to close its fingers. In this way, the hardness of the object affects how the hand can move and how the object can deform after contact. Therefore, for both brittle and soft objects, it is important to observe the hand and the object to avoid damaging the object.

Here, for observation, $\mathrm{IfG_s}$ (Indexes for Grasping) are used to detect excessive grasping force. $\mathrm{IfG_s}$ contain the following four indexes: the movements of both the hand and the object ($\mathrm{IfG_1}$ and $\mathrm{IfG_2}$), object deformation ($\mathrm{IfG_3}$), and the occlusion rate of the hand ($\mathrm{IfG_4}$). $\mathrm{IfG_1}$ and $\mathrm{IfG_2}$ are used to consider the grasping of brittle objects, $\mathrm{IfG_3}$ and $\mathrm{IfG_4}$ are used to consider the grasping of soft ones. Since the hardness of the object is unknown, all the $\mathrm{IfG_s}$, $\mathrm{IfG_1}$ through $\mathrm{IfG_4}$, are calculated for each object. If $\mathrm{IfG_1}$ is smaller than the threshold or the other $\mathrm{IfG_s}$ exceeds the given threshold, the hand is controlled to stop closing to prevent the crushing of a brittle object or collapsing of a soft object. After *grasp detection*, the system proceeds to *grasp detection during arm control*.

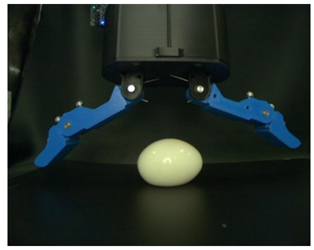


Fig. 2. Images captured by a monocular camera.

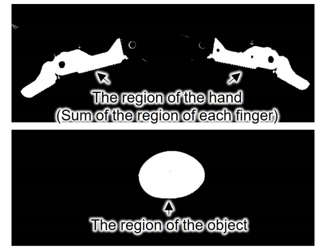
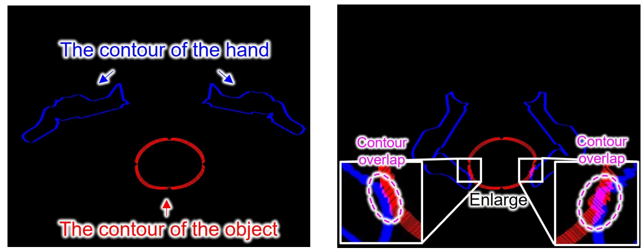


Fig. 3. The regions of the hand and the object extracted from Fig. 2.



(a) Before contact. (b) After contact.

Fig. 4. Contact detection. The magenta colored pixels represent the contact of the robotic hand and the object.

$\mathrm{IfG_1}$: centroid shifts in the robotic hand When grasping a brittle object that has a hard surface, the underactuated hand stops moving after contacting the object. After contact with the brittle object, the hand does not move but continues actuating to close. This causes an increase in the grasping force, and it may lead to crushing or damaging the object. To tackle this problem, whether excessive force is being applied is detected by observing the movement of the hand during actuation. For detection of the hand's movement, the centroid of each finger is used. The centroid is calculated for each region of the finger from Fig. 3. In Fig. 5, the centroid of each finger is indicated by a blue dot. Here, (u_l, v_l) and (u_r, v_r) are expressed as the position of the centroids of the left and right fingers, respectively, on the image coordinate system. l_{shift} and r_{shift} are introduced as the displacement of each finger and calculated using the following equations.

$$l_{shift} = (u_l(t-1) - u_l(t))^2 + (v_l(t-1) - v_l(t))^2 \quad (1)$$

$$r_{shift} = (u_r(t-1) - u_r(t))^2 + (v_r(t-1) - v_r(t))^2 \quad (2)$$

where t and $t-1$ represent the current and previous frames, respectively.

Then the movement $\mathrm{IfG_1}$ of the hand is calculated as follows.

$$\mathrm{IfG_1} = l_{shift} + r_{shift} \quad (3)$$

If $\mathrm{IfG_1}$ is smaller than a threshold during a certain time, it indicates that the hand is not moving. In this case, it is determined that the hand has grasped a brittle object. Then, to avoid applying additional grasping force and damaging the object, the actuator's drive is stopped.

$\mathrm{IfG_2}$: object's movement along the v axis on an image coordinate system When grasping an object, such as a small object, the hand comes to a halt after contact, but soon

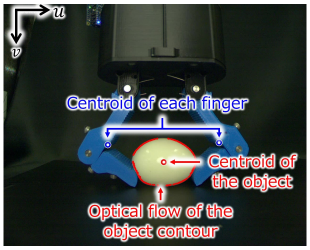


Fig. 5. IfG_1: the centroid of each finger.

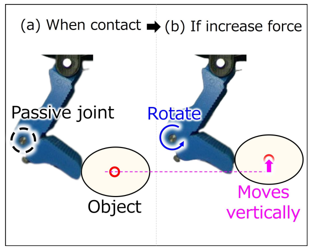


Fig. 6. IfG_2: lifting in the v direction.

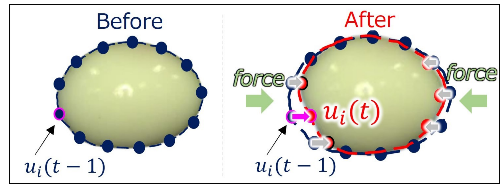


Fig. 7. IfG_3: deformation of the object.

the object moves vertically. Because of the movement of the object, IfG_1 exceeds the threshold. Therefore, whether excessive grasping force was applied to the object cannot be detected. To deal with such a case, another perspective is needed.

Because an underactuated hand is used, its passive joint, which is indicated by the blue arrow in Fig. 6(b), rotates after contact. This causes vertical movement of the fingers. Additionally, if excessive grasping force is applied to the object, the object moves vertically. Therefore, by detecting the vertical movement of the object, it is possible to grasp the object without applying excessive force.

The object's centroid, represented as red circle in Fig. 5, is used to detect its movement. The centroid is calculated from the object region extracted as shown in Fig. 3. Defining the object's movement as the displacement of it, IfG_2 is calculated by the following equation.

$$\mathrm{IfG}_2 = v(t-1) - v(t) \tag{4}$$

Here, v represents the value of the centroid along the v axis on the image coordinate system. t and $t-1$ represent the current and previous frames, respectively. After this detection, the actuator is controlled to stop driving so as not to crush the object.

IfG_3: object deformation Grasping a soft object with excessive force may cause large deformation of the object, resulting in damage to it. To prevent applying excessive force to the soft object, the deformation of the object is detected from an image.

The deformation is detected by calculating the movement of all points on the object's contour in each frame. By assuming that the force to crush the object is applied in the horizontal direction, only the movement of the points along the u axis in the image coordinates is used. In Fig. 5, the u axis is indicated in the upper left corner of the figure. To calculate the movement of points, the Lukas–Kanade optical flow [22] technique is used. The result of applying the optical flow method to the object contour is indicated by the red points in Fig. 5. Fig. 7 shows the calculation procedure. In the frame $t-1$ before deformation, the position along the u axis of each point on the object's contour is denoted by $u_i(t-1)$. Grasping forces, indicated by the green arrows, are applied horizontally to the object. Then a point at $u_i(t-1)$ is assumed to move along the magenta arrow shown in the right image. The new position of the point at $u_i(t-1)$ after the movement is defined as $u_i(t)$. The deformation of the object at the currently observed point is defined by the difference between $u_i(t-1)$ and $u_i(t)$. Considering all N points on the contour, the overall deformation IfG_3 of the object is calculated by the following equation.

$$\mathrm{IfG}_3 = \frac{1}{N}\sum_{i=1}^{N} |u_i(t) - u_i(t-1)| \tag{5}$$

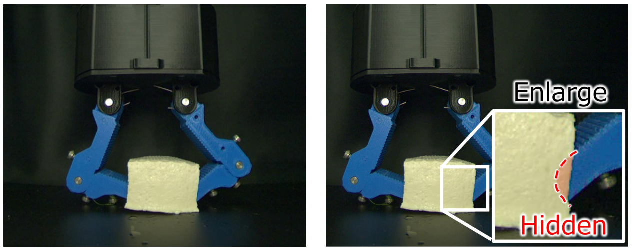
In (5), for normalization, the sum of the displacement along the u axis is divided by the number N of points on the contour. This normalization enables detection of the deformations of objects of different sizes.

IfG_4: occlusion rate of the robotic hand Fig. 8(a) shows an example of grasping a soft object. In this case, it may not appear that the object is deforming in the image. However, in reality, deformation occurs in areas that are not visible in the image. Therefore, IfG_3 is not sufficient to detect the deformation of the soft object; it may lead to damage to the object because of excessive force. On the other hand, in Fig. 8(b), the object appears undeformed, but occlusion is occurring where the hand is hidden by the object. Based on this fact, in our method, the degree of occlusion of the hand is observed to detect the deformation of the object. Therefore, it is possible to grasp a soft object without excessive grasping force.

To detect occlusion of the hand, the dimensions of both the hand and the object are used. The total dimensions of the hand before *grasping detection* are introduced as $H(0)$. The total dimensions of the hand and the object in *grasping detection* in frame t are denoted as $H(t)$ and $O(t)$, respectively. These dimensions of both the hand and the object can be obtained from Fig. 3. Using these dimensions, the occlusion rate IfG_4 of the hand is calculated by the following equation.

$$\mathrm{IfG}_4 = \frac{H(0) - H(t)}{O(t)} \tag{6}$$

In (6), IfG_4 is defined not only as the difference between $H(0)$ and $H(t)$, i.e., the dimensions of the occluded region, but also as the difference divided by $O(t)$. The reason is explained as follows. The dimension of the occluded region varies according to the grasping force. The grasping force required so as not to drop the object depends on the mass of the object. Meanwhile, even with the same hardness, the mass of the object changes with its size. Therefore, considering not only the hardness but also the size of the object is important to grasp it with the appropriate force. In



(a) Grasping a soft object. (b) An enlarged view of the contact area.

Fig. 8. Occlusion of the robotic hand.

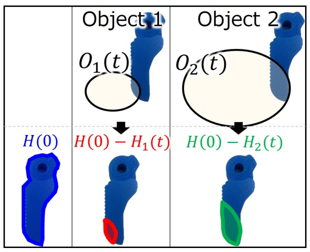


Fig. 9. IfG$_4$: relation between object size and occluded dimension.

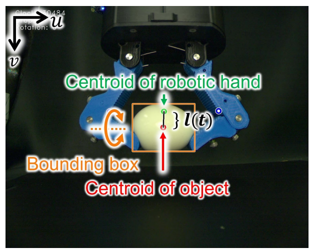


Fig. 10. Slip and rotation detection.

Fig. 9, objects 1 and 2 are assumed to have the same hardness but to be different in size and mass. The occluded regions when sufficient grasping forces are applied to objects 1 and 2 are shown in red and green, respectively. $H(0)$ indicates the dimension enclosed by the blue boundary in the lower left of Fig. 9. Similarly, in frame t, $H(t)$ is calculated for the same region as $H(0)$, where not occluded by the object, by measuring its dimension. Furthermore, $O_1(t)$ and $O_2(t)$ in Fig. 9 correspond to $O(t)$ in (6). Due to the difference in size, the dimensions of the occluded region $H(0) - H_1(t)$ for Object 1 are smaller than those for Object 2, denoted as $H(0) - H_2(t)$. Therefore, normalizing changes $H(0) - H(t)$ in the robotic hand's dimension by dividing it by the object's dimension $O(t)$ allows the grasping force to be adjusted appropriately for objects of different sizes.

E. Grasp detection during arm control

In the pick-and-place task using a robotic arm, it is crucial to maintain the quality of food by preventing the object from dropping. Moreover, moving the robotic arm in an unstable state, when the object is prone to rotation or slippage, may lead to dropping of the object. Therefore, using the images to detect slipping and rotating of the object makes it possible to grasp the object without dropping it.

Slip detection When an object slips, the relative position between the hand and the object changes. Therefore, using the centroid of the hand and the object enables the detection of situations where the object held by the hand is slipping. In Fig. 10, the centroids of the hand and object are represented by the green and red dots, respectively. Here, the centroid of the hand is defined as the average point of each finger centroid. The line connecting the green and red dots is shown in black. The length of the black line can be represented as $l(t)$. Especially before moving the robotic arm, the length is denoted as $l(0)$. When the object is slipping, the centroid of the object moves in the positive direction along the v axis in the image coordinate system. That is, $l(t)$ becomes greater than $l(0)$. Thus, the difference in the length of the black line, denoted as l_{diff}, is calculated as follows.

$$l_{diff} = l(t) - l(0) \tag{7}$$

When l_{diff} exceeds the threshold value, the object is detected to be slipping. After the detection, the robotic hand is closed to increase the gripping force, preventing the object from further slipping and dropping. Additionally, by substituting the current value of $l(t)$ with $l(0)$ after increasing the gripping force, slip detection is continuously performed during the arm movement.

Rotation detection It is also necessary to detect the rotation that causes the object to slip or drop. To detect rotation, a bounding box enclosing the contour of the object is used. In Fig. 10, the bounding box is indicated by the orange rectangle. By calculating the changes in height of the rectangle, the rotation of the object is detected.

The difference in height h_{diff} between the initial height $h(0)$ before moving the arm and the height $h(t)$ during arm movement is obtained using the following equation.

$$h_{diff} = |h(t) - h(0)| \tag{8}$$

When h_{diff} exceeds the threshold value, the object is detected to be rotating. After this detection, the robotic hand is closed as in *slip detection*, and the value of $h(t)$ will be updated as the same procedure in *slip detection*.

IV. EXPERIMENTS

A. Experimental setup

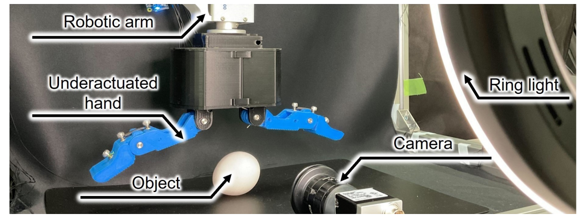


Fig. 11. Experimental condition.

To verify the effectiveness of the proposed method, we conducted grasping experiments. The experimental environment is shown in Fig. 11. In this experiment, we used the Yale OpenHand Project's Model T-42 [8] for the underactuated hand, the MG400 robotic arm from DOBOT, the acA1300-200uc camera from Basler, and the 4 mm fixed-focus lens from Edmund Optics' UC series. As objects to be grasped, food items were selected as shown in Fig. 12. To test the same foods in different sizes, some items were reproduced by changing the orientation of the objects, such as Pose A or Pose B. After *grasp detection*, the robotic arm was moved along the path defined in Fig. 13. For each object, we performed ten trials.

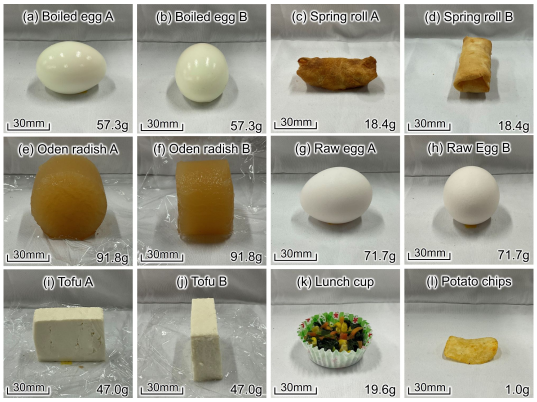


Fig. 12. Tested objects of real food items.

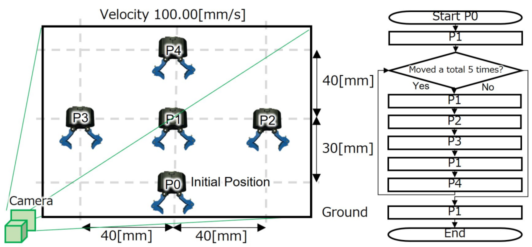
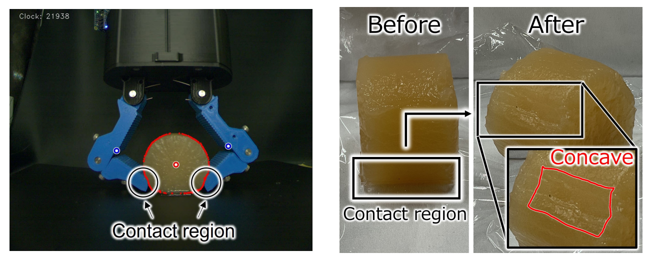


Fig. 13. Trajectory of the robotic arm.

TABLE I. Experimental results

Object	Pose	Evaluation value	
		$e_{success}$	e_{damage}
(a) Boiled egg	A	100	0
(b) Boiled egg	B	100	0
(c) Spring roll	A	100	0
(d) Spring roll	B	100	0
(e) Oden radish	A	100	1
(f) Oden radish	B	100	1
(g) Raw egg	A	90	0
(h) Raw egg	B	100	0
(i) Tofu	A	100	0
(j) Tofu	B	100	0
(k) Lunch cup		10	0
(l) Potato chips		100	0



(a) Small contact region.

(b) Damage.

Fig. 14. Damage on Oden radish A.

B. Experimental results

Evaluation of the experimental results was conducted based on the grasping success rate ($e_{success}$), and the number of observed damages was confirmed visually (e_{damage}). $e_{success}$ is calculated as the success rate of grasping without dropping the object until the end of the trajectory. The results are presented in Table 1. Except for (k) Lunch cup, a success rate of over 90% was achieved. This could be attributed to closing the hand during object slippage or rotation to prevent droppage. Additionally, except for (e) Oden radish A and (f) Oden radish B, the objects were grasped without causing any damage. This was because of the detection of object movement and deformation, which enabled the avoidance of excessive grasping force.

C. Discussion

In the case of (k) Lunch cup, although *grasp detection* was successful, there were frequent failures in lifting the object from position $P0$ to position $P1$, as shown in Fig. 13. (k) Lunch cup exhibited varying degrees of deformation, depending on the grasp location. In the experiments, we selected less deformable parts as the grasp position. However, due to the low friction coefficients of the hand and the small contact areas, the necessary friction force to support the weight of the lunch cup was not generated. Therefore, the object slipped off. To tackle this problem, we are considering increasing the number of fingers on the hand and enlarging the contact area to achieve a secure grip without slippage.

The state after *grasp detection* for (e) Oden radish A is shown in Fig. 14(a). At that time, IfG_3 was detected appropriately, but the object was grasped only by the fingertips of the hand. In addition, during the arm movements, there were false detections of the object dropping, leading to increased grasping force. Because of these two indexes, locally large forces were applied, which caused damage, as shown in Fig. 14(b). To prevent localized grasping forces, we aim to calculate the grasp position that increases the contact area.

(g) Raw egg A dropped because the surface of the object was hard and slippery. The object slipped in the depth direction of the image due to disturbances caused by the arm movements, leading the object to drop. To tackle this problem, we are considering calculating the grasp position to make it possible to geometrically constrain the object.

V. CONCLUSION

In this paper, we proposed a visual feedback control method with an underactuated hand for object grasping using a monocular camera. Without using any internal sensors or markers, grasping of brittle and soft food items was performed. First, contact between the hand and the object was detected using the overlap between the two contours. Then, to avoid damaging the food, we used the centroid of the hand and the object, the deformation of the object, and the occlusion rate of the hand. Moreover, the slippage and rotation of objects were detected during the movement of the robotic arm. In the experiments, the grasping success rate was above 90% for all cases except (k) Lunch cup. Further, all the food items were successfully grasped without causing any damage except in the cases of (e) Oden radish A and (f) Oden radish B.

In the future, we aim to calculate the grasp position where the hand can adapt to object shapes.

REFERENCES

[1] J. Zhu *et al.*, "Challenges and Outlook in Robotic Manipulation of Deformable Objects," in IEEE Robotics & Automation Magazine, Vol. 29, No. 3, pp. 67–77, 2022.

[2] The Japan Food Machinery Manufacturers' Association, "FOOMA Japan 2023 International Food Machinery & Technology Exhibition," foomajapan.jp. https://www.foomajapan.jp/int/ Accessed on July 16, 2023.

[3] Z. Wang *et al.*, "Challenges and Opportunities in Robotic Food Handling: A Review," Front. Robot. AI, Vol.8, 2022.

[4] Z. Wang *et al.*, "3D Printed Soft Gripper for Automatic Lunch Box Packing," in Proc. of IEEE Int. Conf. on Robotics and Biomimetics (ROBIO), pp. 503–508, 2016.

[5] Y. Yamanaka *et al.*, "Development of a Food Handling Soft Robot Hand Considering a High-Speed Pick-and-Place Task," in Proc. of IEEE/SICE Int. Symp. on Syst. Integration (SII), pp. 87–92, 2020.

[6] D. Holland *et al.*, "The Soft Robotics Toolkit: Shared Resources for Research and Design," Soft Robotics, Vol. 1, No. 3, pp. 224–230, 2014.

[7] Soft Gripping, "SoftGripper," soft-gripping.com. https://soft-gripping.com/softgripper/ Accessed on July 16, 2023.

[8] R. R. Ma *et al.*, "A Modular, Open-Source 3D Printed Underactuated Hand," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 2737–2743, 2013.

[9] Z. Lu *et al.*, "GTac-Gripper: A Reconfigurable Under-Actuated Four-Fingered Robotic Gripper with Tactile Sensing," in IEEE Robotics and Automation Letters, Vol. 7, No. 3, pp. 7232–7239, 2022.

[10] Z. Zhou *et al.*, "A Sensory Soft Robotic Gripper Capable of Learning-Based Object Recognition and Force-Controlled Grasping," in IEEE Trans. on Automation Science and Engineering, pp. 1–11, 2022.

[11] T. N. Le *et al.*, "Safe Grasping with a Force Controlled Soft Robotic Hand," in Proc. of IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC), pp. 342–349, 2020.

[12] B. S. Homberg *et al.*, "Robust Proprioceptive Grasping with a Soft Robot Hand," Auton Robot, Vol. 43, pp. 681–696, 2019.

[13] S. Zhang *et al.*, "Hardware Technology of Vision-Based Tactile Sensor: A Review," in IEEE Sensors Journal, Vol. 22, No. 22, pp. 21410–21427, 2022.

[14] Y. She *et al.*, "Exoskeleton-Covered Soft Finger with Vision-Based Proprioception and Tactile Sensing," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 10075–10081, 2020.

[15] S. Q. Liu *et al.*, "GelSight EndoFlex: A Soft Endoskeleton Hand with Continuous High-Resolution Tactile Sensing," in Proc. of IEEE Int. Conf. on Soft Robotics (RoboSoft), pp. 1–6, 2023.

[16] W. Li *et al.*, "L^3 F-TOUCH: A Wireless GelSight with Decoupled Tactile and Three-Axis Force Sensing," in IEEE Robotics and Automation Letters, Vol. 8, No. 8, pp. 5148–5155, 2023.

[17] A. Yamaguchi and C. G. Atkeson, "Implementing Tactile Behaviors Using FingerVision," in Proc. of Int. Conf. on Humanoid Robotics (Humanoids), pp. 241–248, 2017.

[18] D. D. Barrie *et al.*, "Deep Learning Method for Vision Based Force Prediction of a Soft Fin Ray Gripper Using Simulation Data," Front. Robot. AI, Vol. 8, 2021.

[19] P. Grady *et al.*, "Visual Pressure Estimation and Control for Soft Robotic Grippers," in Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), pp. 3628–3635, 2022.

[20] A. S. Morgan *et al.*, "Towards Generalized Manipulation Learning through Grasp Mechanics-Based Features and Self-Supervision," in IEEE Trans. on Robotics, Vol. 37, No. 5, pp. 1553–1569, 2021.

[21] P. V. Nguyen *et al.*, "Wet Adhesion of Micro-Patterned Interfaces for Stable Grasping of Deformable Objects," in Proc. of IEEE/RSJ International Conf. on Intelligent Robots and Systems (IROS), pp. 9213–9219, 2020.

[22] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision" in Proc. of the 7th International Joint Conference on Artificial Intelligence, Vol. 2, pp. 674–679, 1981.