EDGE EXTRACTION USING IMAGE AND THREE-AXIS TACTILE DATA

Sukarnur Che Abdullah¹,², Takuya Ikai¹, Yusuke Dosho¹, Hanafiah Bin Yussof³, and Masahiro Ohka¹

¹Graduate School of Information Science, Nagoya University
²Faculty of Mechanical Engineering, Universiti Teknologi MARA
³Email: sukarnur.ca@gmail.com

Submitted: June 12, 2011      Accepted: August 22, 2011    Published: September 1, 2011

Abstract - This paper describes a hand-arm system equipped with optical three-axis tactile sensors and a binocular vision sensor. The vision compensates for the limitations of tactile information and tactile sensing, and vice versa. The tactile sensor can obtain geometrical data as real scale, while image data requires calibration to obtain length as a metric unit. Even if stereovision is used, we cannot obtain sufficient precision. In the evaluation test, the robotic hand equipped with tactile sensors traces an object including convex and concave portions to evaluate edge trace precision. Error of distance obtained by the binocular vision is around ±10 mm when distance between the camera and object is around 600 mm. When the hand-arm robot touches the convex portion of the object, size data obtained by the vision is modified within ±0.5 mm accuracy. Since the robotic finger is too thick to touch the bottom of the concave, size data of the concave portion obtained by tactile sensing includes relatively large error of around 4 mm. However, the robot finger can follow the contour with ±0.5 mm accuracy.
except for the bottom portion. Therefore, vision sensing is not sufficient for precise edge exploration and modification based on tactile sensing is required.

Index terms: Tactile Exploration, Edge Extraction, Image Data, Hand-arm Robot, Calibration.

I. INTRODUCTION

For a robot, it is difficult to recognize geometrical data of an object buried in a complicated background using vision. Even for human beings, it is sometimes difficult to find a specified object on a messy desk, for example. Since persons know that the relationship between the specified object and the background can be explored with trial handling such as pulling or twisting the object, it is very effective for robots to use tactile data to obtain the geometrical data of object.

Our research group studies the Artificial Tactile Affordance System (ATAS) to control a robot based on interaction between a robotic hand and the environment through the above trial. In a previous study [1], our robotic hand was able to obtain the ideal trajectory of fingers to twist a bottle cap based on tactile data acquired through handling the cap. Although the tactile sensor can obtain geometrical data as real scale, it cannot survey a wide area; its scanning is restricted to around the contact area between the finger and sensor. If visual image data are used in addition to the tactile data, they compensate for the above limitation of tactile information to enhance ATAS.

On the other hand, edge detection is one of the important components in both robot vision and tactile sensing. To evaluate our study, edge detection using robot vision and tactile sensing is surveyed in the following. First, research on robot vision has spanned more than four decades, resulting in development of several types of edge detectors [2]-[7]. Some of these implemented segmentation techniques use only the gray level histogram and spatial details, while others use fuzzy set theoretic approaches [8]. Edges are found by applying a derivative operator or a high-pass filter to each pixel of the smoothed image. Although edge detection using tactile sensing has a rather short history compared to robot vision, several researchers [9]-[15] have presented their work on edge detection with tactile exploration. Their studies are considered important milestones because they showed the effectiveness of tactile information for exploratory procedures [9], surface tracking [10], object-edge tracing [11]-[13], reconstructing shapes [14] and feature modeling [15].

When edge detection is performed by tactile sensing, the robotic finger has two inconsistent
specifications: one is as a thin-needle probe of a coordinate-measuring machine (CMM); the other is to lift up and handle an object, and a thick finger is required for this specification. To solve the above contradiction requiring simultaneously thin and thick shapes for the finger, many sensing points are required to be distributed on the finger surface.

Our ultimate goal is to incorporate visual sensation into ATAS. Since edge extraction involves basic data processing to perform several tasks such as object recognition and handling, edge extraction using image and tactile data is treated as the first step toward that ultimate goal.

In this paper, before establishment of the whole system, the vision system and the hand-arm system are individually checked. From results of stereovision experiments with a binocular eye system, the size of the obtained edge includes error of around ±20 mm when distance between the camera and object is around 600 mm. For the edge detection test, we adopt the Canny algorithm, which is included in OpenCV. In the edge detection test using a monocular eye, we adopt a piece of wood with convex and concave portions as a specimen. Although a relatively sharp object edge is obtained from vision data with the Canny algorithm, the edge data should be calibrated because they do not have metric units, while the distance obtained by the binocular eye system includes large error. Because of this, we calibrate the data using tactile sensing, in which the hand-arm robot equipped with three-axis tactile sensors is capable of sensing distributed contact data. In the evaluation test, the hand-arm robot traces the object to evaluate effectiveness of calibration using tactile sensing.

II. DATA INTEGRATION OF TACTILE AND VISUAL SENSING

Tactile sensing and visual sensing have advantages and disadvantages as shown in Table 1. If the disadvantages of one can be compensated by the advantages of the other, the two can be applied to allow a robot to perform several tasks.

Since the tactile sensor measures an object with direct touch, it can obtain geometrical data as real scale. This characteristic is convenient for measuring the object edge, while image data requires calibration to obtain length as a metric unit because the length unit obtained by image data is the pixel. Even if a stereovision system is used, we cannot obtain sufficient precision because it includes ±10-mm error when distance between the camera and object is around 600 mm as discussed below.
However, the tactile sensor cannot survey a wide area but is restricted to the contact area between the finger and object surfaces. If visual image data is used in addition to the tactile data, they can compensate for the above limitations of tactile information.

Table 1: Advantages and disadvantages of tactile sensor and vision sensor in recognizing object shape

<table>
<thead>
<tr>
<th>Aspects of sensor</th>
<th>Tactile sensor</th>
<th>Vision sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>Depends on contact condition</td>
<td>Depends on lighting condition and calibration is required</td>
</tr>
<tr>
<td>Detection area</td>
<td>Restricted to contact area</td>
<td>Wide area</td>
</tr>
<tr>
<td>Environmental/background subtraction</td>
<td>Not related</td>
<td>Image segmentation required</td>
</tr>
<tr>
<td>Room lighting</td>
<td>Not required</td>
<td>Fully dependent</td>
</tr>
<tr>
<td>Edge/contour location</td>
<td>Must be taught object position</td>
<td>Depends on image processing filter</td>
</tr>
<tr>
<td>Reliability/robustness</td>
<td>Moderate reliability and not too robust</td>
<td>Not related</td>
</tr>
</tbody>
</table>

In our project, we are attempting to develop a robotic system equipped with stereovision and tactile sensors as shown in Fig. 1. In the following chapters, we will explain the tactile data acquisition system and the visual data acquisition system.

Figure 1. Robotic system equipped with stereovision and tactile sensors
III. TACTILE DATA ACQUISITION SYSTEM

a. Three-axis tactile sensor

Since our three-axis tactile sensor has been explained in previous papers [15]-[18], the structure and functions of the tactile sensor are described briefly in the present paper. The tactile sensor is composed of a CCD camera, an acrylic dome, a light source, and a computer as shown in Fig. 2. The light emitted from the light source is directed into the acrylic dome. Contact phenomena are observed as image data, which are acquired by the CCD camera and transmitted to the computer to calculate the three-axis force distribution. The sensing element presented in this paper is comprised of a columnar feeler and eight conical feelers, which are arranged on the hemispherical dome in concentric configurations with 41 sub-regions. The sensing elements, which are made of silicone rubber, are designed to maintain contact with the conical feelers and the acrylic dome and to make the columnar feelers touch an object.

Figure 2. Three-axis tactile sensor system
When the three components of the force vector, $F_x$, $F_y$, and $F_z$, are applied to the tip of the columnar feeler, contact between the acrylic dome and the conical feelers is measured as a distribution of gray-scale values, which are transmitted to the computer. $F_x$, $F_y$, and $F_z$ values are calculated using integrated gray-scale value $G$ and the horizontal displacement of the centroid of gray-scale distribution.

Sensing elements are arranged on the acrylic dome in a concentric configuration. The acrylic dome is illuminated along its edge by optical fibers connected to a light source. Image data consisting of bright spots caused by the feelers’ collapse are retrieved by an optical fiber scope connected to the CCD camera.

Image data acquired by the CCD camera are divided into 41 sub-regions as shown in Fig. 3: the dividing procedure, digital filtering, integrated gray-scale values, and centroid. Image data obtained by CCD camera and the address of sensing-element displacement are processed on an image processing board. Then, force vector components $F_x$, $F_y$, and $F_z$ are calculated according to each local coordinate defined in Fig. 3.

Since the image warps due to projection from a hemispherical surface, software installed on the computer modifies the warped image data and calculates $G$, $u_x$, and $u_y$ to obtain the three-axis force applied to the tip of the sensing element.
b. Hand-arm robot

Figure 4 is a photograph of the present robot; the arm system has DOF of 5; each finger’s DOF is 3. To compensate for the lack of arm DOF, this robot uses its finger’s root joint as its wrist’s DOF. On each fingertip, it has the three-axis tactile sensor described in the previous chapter. We use the sensor information as an effective key to induce a specified behavior. A local coordinate system is embedded on sensing elements of the three-axis sensor attached to the fingertip. Figure 3 shows the relationship between the location of the sensing element and the local coordinate. Using coordinate transformation, component slippage vectors are calculated with respect to the global coordinate O-XYZ attached to the base.

Figure 4. Photograph of hand-arm robot

Position control of the fingertip is performed based on resolved motion rate control. In this control method, joint angles are assumed at the first step, and the current displacement vector is calculated with kinematics. Joint angles are adjusted through the current joint angle and the difference between the current displacement vector and the objective displacement vector to modify the joint angle in the next step. The modified joint angle is designated as the current angle in the next \( k + 1 \)-th step, and the above procedure is repeated until the displacement vector at \( k \)-th
step \( \mathbf{r}_k \) coincides with objective position vector \( \mathbf{r}_d \) within a specified error. That is, the following Eqs. (1) and (2) are calculated until \( |\mathbf{r}_d - \mathbf{r}_k| \) becomes small enough:

\[
\dot{\mathbf{r}}_k = J \dot{\theta}_k, \tag{1}
\]

\[
\theta_{k+1} = \theta_k - J^{-1}(\mathbf{r}_d - \mathbf{r}_k) \tag{2}
\]

where \( J \) is the Jacobian matrix obtained from kinematics of the hand-arm robot.

Figure 5. Algorithm of finger-speed estimator
c. **Control algorithm of arm-hand robot**

The hand is controlled according to velocity control. First, hand status enters “search mode” to make fingers approach an object with finger speed $v = v_0$. After the fingers touch the object, the hand status changes to “move mode” to manipulate the object with finger speed $v = v_m$. During both search and move modes, when the absolute time derivative of the shearing force of a sensing element exceeds a threshold $dr$, this system regards the sensing element as slippage. To prevent the hand from dropping the object, re-push velocity is defined as moving the fingertip along the counter direction of applied force.

However, if normal force of a sensing element exceeds a threshold $F_s$, the re-push velocity is canceled to prevent the sensing element from breaking. The hand is controlled by a control module applying total velocity obtained by adding the re-push velocity to the current velocity.

In our system, the sensor control program and hand control program are executed in different computers because CPU time is efficiently consumed using a multi-task program method. These programs are synchronized with the following five flags.

- **SEARCH**: Fingers search for an object with initial finger velocity $v_0$ until normal force of a sensing element exceeds a threshold $F_1$ or SLIP flag is raised.
- **MOVE**: This flag is raised whenever the robotic hand manipulates an object.
- **TOUCH**: This flag is raised whenever one of the fingers touches an object.
- **SLIP**: This flag is raised whenever the time derivative of shearing force exceeds a threshold $dr$.
- **OVER**: This flag is raised when normal force of a sensing element exceeds a threshold $F_s$ or a specific direction of the force component exceeds a threshold.

These flags are determined according to tri-axial tactile data and finger motions. Two modules, the flag analyzer and finger-speed estimator, mainly play the role of object handling. The algorithm of the finger-speed estimator is shown in Fig. 4.

In the flag analyzer, the TOUCH flag, SLIP flag and OVER flag are decided on. The flag analyzer regards finger status as touching an object when normal force of a sensing element exceeds $F_1$ or the absolute time derivative of the shearing force exceeds $dr$ (SLIP flag is raised). Whenever it regards finger status as touching an object, the TOUCH flag is raised. The OVER
flag is raised when normal force of a sensing element exceeds $F_2$ to prohibit the re-push motion. For this purpose, finger velocity is set as the opposite of the last step direction. In the finger-speed estimator, the velocity of the fingertip is determined based on the five flag values and conserved whenever contact status is not changed. Since in the edge-extraction experiment the robot hand should keep contact with the object during scanning, the MOVE and TOUCH flags continue to be raised after SEARCH flag is lowered because of first touch. Whenever the SLIP flag is raised, a sensing element of the largest normal force is determined and the re-push velocity of the finger is determined as an inward normal line of the sensing element. The re-push velocity is added to the current velocity, and the resultant velocity is applied to the control module. According to this re-push motion, the robot finger can apply contact force to normal direction of the object surface.

<table>
<thead>
<tr>
<th>Constants and threshold</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling interval for sensor</td>
<td>100 msec</td>
</tr>
<tr>
<td>Sampling interval for finger</td>
<td>25 msec</td>
</tr>
<tr>
<td>$F_1$</td>
<td>0.08 N</td>
</tr>
<tr>
<td>$F_2$</td>
<td>0.2 N</td>
</tr>
<tr>
<td>$dr$</td>
<td>0.04 N/sec</td>
</tr>
<tr>
<td>Initial velocity, $v_0$</td>
<td>1 mm/sec</td>
</tr>
<tr>
<td>Manipulate velocity, $v_m$</td>
<td>1 mm/sec</td>
</tr>
<tr>
<td>Re-push velocity, $v_p$</td>
<td>1 mm/sec</td>
</tr>
</tbody>
</table>

On the basis of the resultant velocity, the revolution velocity of each joint is calculated to control the hand-arm robot motion. Major constants for the control are shown in Table 2, which are defined from a primary calibration test that was conducted during development of the tactile sensor [13] For the edge-extraction experiment discussed in the next chapter, the difference between thresholds $F_1$ and $F_2$ should be small enough to have no influence on other experiments. After several trials, we selected 0.08 and 0.2 N for $F_1$ and $F_2$, respectively.
IV. VISUAL DATA ACQUISITION SYSTEM

a. Vision sensor
We have designed a binocular robot eye system to use for a test of stereovision. The binocular eye can view vision as a process of active data acquisition. Therefore, a central part of a typical active vision system is a camera-head as shown in Fig. 6, which can control the following parameters: orientation of the viewing-direction, which are neck-movements, pan and tilt; and further camera parameters, e.g. distance between two cameras.
In this paper, edge extraction is performed by monocular eye since we only study the edge for a static object and depth perception is not related. Furthermore, the process is performed off-line from tactile sensor edge exploration.

![Figure 6. Binocular robot eye system](image)

b. Edge extraction algorithm
Canny [2] defines certain desirable criteria for edge detection. Canny’s criteria achieve good localization, good detection, and a single response to a single edge, and are used in an optimization framework to derive optimal detectors for ridges, roof edges, and step edges. The optimal detector for step edges has an approximate implementation in which edge pixels are the
maxima in the gradient field of a Gaussian-smoothed image. The step edge detector performance is further enhanced by extending the operator point spread function along the edge.

In this research, we implement the Canny algorithm in OpenCV (cvCanny) image processing for edge detection transforms. The Canny Edge Detection algorithm takes the derivative of an image to find the gradients, then determines the direction of these gradients (vertical, horizontal, diagonal up, diagonal down). Then, if the amplitude of a given gradient is high enough (the high threshold), the algorithm will trace along that gradient in its direction until the amplitude falls below the low threshold, or the gradient changes direction sharply. The algorithm will also suppress the local non-maximum around the edges.

V. EXPERIMENTAL RESULTS AND DISCUSSION

a. Stereo vision test

In this paper, elements of the robot system shown in Fig. 1 are individually tested. For the binocular eye system, distance is measured between the origin of the binocular eye system and objects such as a cup on the table shown in the left photograph of Fig. 7. The distance calculation is performed by “cvFindStereoCorrespondenceGC” of OpenCV function to obtain a distance chart expressed by gray scale image data shown in the right image in Fig. 7. In this image, the gray scale decreases with increase of distance.

Figure 7. Distance measurement using binocular eyes
Figure 8 shows the relationship between graduation and distance. In this system, the distance cannot be obtained when it exceeds 730 mm. Since the distance between the binocular eye system and the hand is around 600 mm, the system can calculate the distance. However, since we obtain measurement error around ±10 mm from Fig. 8 because of 730 mm divided into 16, we cannot estimate the distance between the robot and an object using only the vision system with sufficient precision.

b. Tactile edge exploration

Figure 9. Specimen
We conducted an edge-exploring experiment using the two robotic fingers to explore real objects to supplement with vision. In this experiment, after the robot touches both sides of the object with two fingers, the hand moves linearly in the z-direction as shown in Figure 9. According to the algorithm shown in Figure 5, the hand-arm robot maintains contact with the sides of the object because it tries to make maximum normal force generated in a tactile element to maintain the range specified by two thresholds, $F_1$ and $F_2$.

Figure 10 shows the result of edge extraction with tactile data processing. We conducted a series of experiments using the two robotic fingers to explore real objects in order to obtain normal force distribution from our tactile sensor. The normal force distribution performs in the proposed control system by responding to the changes of object shape during the object-exploring tasks. In the experiment result as shown in Fig. 10, the black line is the fingertip position and the dots with several colors are the normal force distribution performed by 17 sensor elements of the optical three-axis tactile sensor.

Figure 10 a) shows the start position of fingertip 1 was at 30 mm, and moved down to 12-13 mm at around 60 sec. Then it goes up to 27 mm at 168 sec. Figure 10 b) shows the other way around: start position -30 mm and moving upward to -11 mm at 58 sec. Then, it goes down to reach -25 mm at 168 sec. These two figures show that the movement of exploration was started at the wide convex position, moved to a narrow concave position and then moved toward a convex position again.
During this scanning, sensing element accepting the highest normal force is changed as shown in Fig. 10. In the fingertip 1, at first element #17 accepts the highest normal force. After that, element #05, #00 and #01 show the highest normal force in turn. On the other hand, in fingertip 2 sensing element accepting the highest normal force is changed in the reverse order of fingertip 1 as #09, #01, #00, #05 and #17.

Furthermore, from the plotted graphs of x directional position (the local x coordinate is defined as hand open/close direction), the initial geometry of the object surface can be obtained because the black line traces the object contour. The fingers trace the contour from thick to thin width before around 60 sec in Fig. 8. During this period, since MOVE and SLIP flags are raised, decreased normal force is increased to keep constant normal force. On the other hand, since opening hand is performed by exceeding threshold $F_2$ after 60 sec. and this opening motion is intermittently performed, the finger trajectory is varied as stairs. Although this variation is invisible because of small variation, vibration of normal force becomes considerably large after 60 sec. Therefore, three-axis tactile sensing accomplishes smooth edge-tracing.

c. Visual edge detection

![Figure 11. Edge extraction with image data processing](image_url)

Since the binocular vision cannot precisely measure distance between the robot and the object, we adopt another strategy that modifies the image data with tactile data. First, we obtain contour data from the monocular data using Canny’s algorithm described in the previous section. The obtained contour is shown in Fig. 11. Although there are vague portions in the extract edge, it can extract the entire contour of the specimen. Since the hand-arm robot can measure the width of the specimen, the unit of image data obtained as pixels is changed to metric units.
d. Comparisons and evaluation of result with real image

The modified contour obtained from image data is shown in Fig. 12. In this figure, we superimposed two kinds of contours obtained from tactile sensing. One of them is obtained from distributed tactile sensing data; in this edge follower, the robot finger can follow the concave of the specimen because it can estimate the contact position even if the side of the fingertip touches the concave. The other is obtained from single tactile sensing data; this edge following test is performed to compare the data with the distributed tactile data; in this edge follower, the edge is measured by single sensing element #00 in Fig. 3.

In Fig. 12, since the image data is calibrated by tactile sensing data, at points A and B the contour obtained from image data almost completely coincides with tactile data. Although in the convex portion three curves coincide, there is spacing between contours obtained by visual and tactile sensing because the side of the fingertip touches the object surface before the center of the fingertip touches the object. In the single sensor case, the contact portion on the side of the fingertip cannot be estimated. Although the single sensor case does not completely reproduce the case of the force sensor with probe, a similar phenomenon will occur because the probe with a
larger radius of curvature cannot make its tip fit the bottom of the concave. The maximum error is 4.1 mm on the bottom of concave. The distributed tactile sensing case can follow the side of concaves as shown by C. Although the contour obtained from distributed tactile sensing coincides with the contour obtained from vision in portion C, single tactile sensing deviates from them and the maximum error is 2.7 mm. The above results suggest that even thick fingers can obtain the object contour if equipped with distributed tactile sensors.

VI. CONCLUSION

Geometric data are basic and necessary to manipulate an object. Although image-data processing is very effective for obtaining geometrical data, there are several problems such as excluding the background, adjusting lighting and coordinating transformation. Particularly, if we use a monocular eye system, we cannot obtain metric information. Even if we use a binocular eye system, obtained distance data include considerably large error.

On the other hand, object size is obtained as metric data via tactile sensing. Although the tactile sensing has attractive advantages, it also has disadvantages such as small survey area. In this study, we adopt a strategy that compensates for each disadvantage using both visual and tactile sensing.

In the evaluation test, the robotic hand equipped with tactile sensors traces an object including convex and concave portions to evaluate edge trace precision. Error of distance obtained by the binocular vision is around ±10 mm when distance between the camera and object is around 600 mm. When the hand-arm robot touches the convex portion of the object, size data obtained by the vision are modified within ±0.5 mm accuracy. Since the robotic finger is too thick to touch the bottom of the concave, size data of the bottom of the concave obtained by tactile sensing include relatively large error of around 4 mm. However, if contour data obtained by vision are applied to the concave portion, the hand-arm robot can extract precise contours of the object. Therefore, the combined sensing with vision and tactile sensing is very effective for precise edge exploration.
REFERENCES


