Acquisition of Hand Information Using an Omni-directional Device

Kyuhei Honda* and Masayuki Hamauchi

Oita National College of Technology, 1666 Maki, Oita 870-0152, Japan

*Corresponding Author: honda@oita-ct.ac.jp

Abstract

In this paper, we propose a user interface for object manipulation with a multi-fingered robot hand. In order to achieve object manipulation, the user needs to tell the robot about task information such as desired contact points between the manipulated object and robotic fingers. Our teaching-by-showing system can measure the user's hand and its surroundings using an omni-directional camera and laser modules, while the camera which the user holds with several fingertips is regarded as a virtual object for object manipulation. The human hand is in contact with the cylindrical surface of the omni-directional camera so that our system can obtain the desired contact points for object manipulation by processing the camera image and using a support vector machine. The experimental results show the system can recognize which fingertip contacts the camera surface and can measure the contact positions of the fingertips.

Keywords: user interface, multi-fingered robot hand, omni-directional camera, laser module, support vector machine

1. Introduction

In recent years, robot hands have been expected to achieve complicated tasks not only for full automation in factories but also for working in human society such as offices, homes and so on. A multi-fingered hand⁽¹⁻⁴⁾ has the dexterity to achieve complicated tasks with many fingers and joints. However, the planning of its movement is problematic due to many degrees of freedom. Besides, few inexpensive human-machine interfaces for a multi-fingered hand have been so far developed for users without technical knowledge.

Several researches have been reported as

human-machine interfaces for controlling a robot hand. Some master-slave systems have been developed for a multi-fingered hand. Ueyama et al.⁽²⁾ utilize a data glove with multiple linear regression models to measure each joint angle of the human hand. Tsujiuchi et al.⁽³⁾ utilize a data glove to control a pneumatic hand. Wojtaraa et al.⁽⁴⁾ utilize a data glove to measure each joint angle of the human hand and a device to convey the feedback of exerted forces. In some systems, the posture of the human hand is estimated by processing images taken by cameras. Hoshino et al.⁽⁵⁾ apply the Self-Organizing Map method for clustering in order to retrieve posture and joint angles of the human hand based on the database of images. M.H. Jeong et al.⁽⁶⁾ estimate the hand posture with an active contour model with the power of several PCs in order to process a lot of images in real time. Master-slave systems have some problems such as the cost of special apparatus, the delay of signal transmission and the portability. Vision based systems have some problems such as various light conditions and the occlusion.

In this research, a human-machine interactive system has been developed for object manipulation which is achieved by using several fingertips of a multi-fingered robotic hand. In order to achieve object manipulation by a multi-fingered robotic hand, it is necessary to obtain the task information on contact positions of fingertips which are in contact with the manipulated object, the pose of the object, and so on⁽¹⁾. The proposed system can obtain the information on user's intention that includes the desired contact positions of fingertips and the desired pose of the object. The goal of this system is to enable a robot to manipulate an object while a user teaches. In other words, a robot learns user's intention by observing the human hand for object manipulation. In this system, a user holds and manipulates an omni-directional camera as an input device, which means that the user changes the pose of the camera so that the system can provide the desired pose of an object

to be manipulated by a robot hand. On the assumption that the fingertips of a robotic hand should be in contact only with the manipulated object during manipulation and that several fingertips of the human hand should be in contact with the cylindrical surface of the camera, an omni-directional image is taken including user's fingers and the environment. We also have attached laser modules to the omni-directional camera in order to obtain contact information of the human hand. In this paper, we mainly present a basic concept of this system, image processing for obtaining contact information of the human hand, and experimental results.

2. System overview

The overall structure of our system is shown in Fig.1. The system consists of the acquisition of information on the object pose and contact points between the human hand and the virtual object. This information is calculated by processing the image taken by an omni-directional camera which the user holds the omni-directional camera as a virtual object by using the user's fingertips. The obtained information is transmitted to the robot control system to control the actual robot hand.

In order to avoid complicated image processing, the omni-directional camera is surrounded by an artificial environment. For object pose estimation, we use an environment which has some markers on it whose positions are already known. By matching those known positions and measured positions of the markers in the image, the system can estimate the pose of the camera that the user holds. We already reported the principle of the pose estimation⁽⁷⁾. In this paper, we focus on the acquisition of contact information of the human hand. To further simplify the image processing, we simply use a white background around the camera.

Fig. 1. Overall structure of the system

Fig. 2. Omni-directional camera and laser module

Fig. 3. Contact state between finger and camera

The omni-directional device we use here is the combination of an omni-directional camera and laser modules which are shown in Fig.2. The omni-directional camera consists of a conventional camera and a mirror whose surface is hyperbolic. This has an advantage that it can obtain views of all the directions around the center of the camera. The image contains the information on fingertips as long as the fingertips touch the transparent plastic surface which supports the upper mirror of the camera. The laser module emits the laser beam which reflects on a finger surface as a spot. We have placed 15 laser modules in a circle around the camera in order to detect whether fingertips contact the plastic surface of the omni-directional camera.

3. Calculation method

3.1 Camera model

The camera coordinate is shown in Fig. 4. The origin of the camera is located at the focal point and the origin of mirror coordinate is located at the focal point of the hyperbolic surface. A point $\tilde{x}$ in the image plane corresponds to a point $\mathbf{X}_c$ in the camera coordinate by using $\tilde{x} = sKX_c$,

(1)

Fig. 4. Camera and mirror coordinate frames

where s is a scale factor, and $\mathbf{K}$ is a matrix which contains the information on the focal length, the image center and the transformation coefficients along the axes of the image plane. The gaze direction $\mathbf{v}$ is calculated based on the position in the image. The intersection point $\mathbf{X}_{mp}$ among the gaze line and the hyperbolic surface is calculated by substituting the line equation of gaze direction into the hyperbolic equation. The reflected direction $\mathbf{p}$ is expressed as

$$\mathbf{p} = \mathbf{v} - 2(\mathbf{v} \cdot \mathbf{n})\mathbf{n} \quad (2)$$

where $\mathbf{n}$ is a unit normal vector of at the point $\mathbf{X}_{mp}$ on the hyperbolic surface.

In this setting, a fingertip looks smaller in the image when it is located closer to the laser module and apart from the mirror.

3.2 Image processing

First, the system obtains the image which contains fingers and the environment (Fig.5a). In the image, each fingertip goes toward the center of the image. This is because the user holds the camera only with the fingertips in order to teach the robot to manipulate an object in robot task space. Next, the image which contains the regions of the fingers is obtained by excluding the environment (Fig.5b).

After that, the system obtains two types of images. One image is obtained by skin color extraction (Fig.5c). The threshold of extraction is determined in advance after measuring YUV values that the skin color of the human hand has under various light conditions. The region of the laser reflection is expressed in black. Those areas are filtered and labeled based on the threshold of area size. Five areas which correspond to the fingers of the user are extracted. Let G_{1i} ($i=[1..5]$) be a gravitational center of a

Fig. 5. Image processing

Fig. 6. Angle of a region from the x-axis

labeled area in this image. Let S_{1i} ($i=[1..5]$) be a size of a labeled area.

Another image is obtained by using the threshold of brightness (Fig.5d). Let G_{2i} ($i=[1..5]$) be a gravitational center of a labeled area in this image. Let S_{2i} ($i=[1..5]$) be a size of a labeled area. Let A_i be an angle of G_{2i} from the x-axis of the image (Fig.6). Let C be the image center. Let T_i ($i=[1..5]$) be a point in a labeled area which is the nearest to the image center C .

3.3 SVM

A support vector machine is an algorithm which can solve a classification problem. Our system uses two types of SVMs for obtaining contact information.

One SVM is used to detect whether each fingertip contact the surface of the camera. In other words, it is used to detect whether a laser spot is reflected on the fingertip or not. The image features which are used for this SVM are obtained by using two images shown in Fig.5c and Fig.5d with the following equations using symbols mentioned in section 3.2.

(Ft 1) The area size of the laser reflection is obtained by

$$S_i = |S_{2i} - S_{1i}| \quad (3)$$

If the laser reflection occurs, this value becomes larger because the area S_{1i} which expresses laser reflection is smaller than the area S_{2i} .

(Ft 2) The difference among two gravitational centers of one area against the image center is obtained by

$$Diff_i = | |G_{2i} - C| - |G_{1i} - C| | \quad (4)$$

If the laser reflection occurs, this value becomes larger because the gravitational centers of two images become different from each other.

(Ft 3) The value is obtained by dividing (Ft 2) by the distance between T_i and the image center.

$$Diff_norm_i = \frac{Diff_i}{|T_i - C|} \quad (5)$$

This is merely the normalized value of (Ft 2).

Another SVM is used to detect which labeled area belongs to the thumb of the user's hand. The image features which are used for this SVM are obtained with the following equations.

(Ft 4) The relative angle A_{ij} between G_{2i} and G_{2j} is obtained by

$$A_{ij} = |A_i - A_j|, \quad (6)$$

where the value is less than 180 degrees. The sum of A_{ij} is also obtained as follows:

$$SA_i = \sum A_{ij} \quad (7)$$

This value SA_{ij} tends to be larger if the area belongs to the thumb of the user's hand. Therefore, the ranking is used for recognizing the thumb via SVM as follows.

(Ft 5) The ranking R_i of the sum of A_{ij} is obtained by

$$R_i = \text{ranking}(SA_i, \{SA_1, SA_2, SA_3, SA_4, SA_5\}) \quad (8)$$

4. Experiments

The user held the surface of the omni-directional device in 59 different ways and gave the classifiers 295 data of fingers as training data. Several samples of training data are shown in Fig. 7.

After training, the classifiers were evaluated by using 50 data. Some of the data are shown in Fig. 8. As for

Fig. 7. Samples of training data

Fig. 8. Samples of evaluation data

Fig. 9. Misrecognized data (Far from mirror)

Fig. 10. Misrecognized data (Close to mirror)

obtaining the contact information on whether each fingertip contacts the surface, the classifier uses the features of equations (3), (4) and (5). With this classifier, the recognition rate is more than 95 %. Some of the data were misrecognized. The way of holding the surface and the image for one failure example are shown in Fig. 9. This failure of recognition occurred because the detected area of the laser spot was too small to judge that the fingertip detached the surface due to the position which was near the image center. In another experimental setup, the failure of recognition was observed (Fig. 10). This is because the detected area of the laser spot was too large since the fingertip was located near the mirror even though the

Fig.11. Only one fingertip is detached from the surface.

Fig.12. Two fingertips are detached from the surface.

Fig.13. Distance from base circle to fingertip.

fingertip contacted the surface.

The other classifier of SVM uses the features of equations (6) and (8). This classifier recognized which labeled area in the image belonged to the thumb perfectly. It is assumed that this is because the feature of the thumb is easily distinguished in the image. Based on the position of the thumb, the system recognizes the rest of fingers including the index finger, the middle finger, the ring finger, and the little finger. If the user holds the camera with the right hand, the system finds the location of the index finger in the image by searching it counterclockwise after starting from the position of the thumb. After that, the system finds the middle finger, the ring finger, and the little finger in the same way. On the other hand, if the user holds the camera with the left hand, the system finds the location of the index finger by searching it clockwise after starting from the position of the thumb. Two examples of the results are

shown in Fig. 11 and Fig.12. The type of each fingertip is expressed as a symbol of {t, i, m, r, l} which means {thumb, index, middle, ring, little}. Contact state is expressed as a symbol of {C, N} which means {Contact, Non-contact}. In fig.11, the user holds the camera with the right hand. The index finger is not in contact with the surface of the camera. In fig.12, the user holds the camera with the left hand. The middle finger and the little finger are not in contact with the surface of the camera.

The kernel type of SVMs we use here is Gaussian kernel and its sigma is 0.00042. The parameter is 1.0 to control a trade-off between the margin size and the penalty which is expressed by using a slack variable.

The relationship between the fingertip position from the bottom edge of the plastic cover of the omni-directional camera and its corresponding point in the image is shown in Fig.13. The value along the horizontal axis in Fig.13 shows the distance in pixel between a fingertip and the image center. By using this relationship, the system can calculate the contact points on the device surface once the fingertip positions in the image are obtained by image processing.

5. Conclusions

We present a human-machine interface for object manipulation by a multi-fingered robotic hand using an omni-directional camera. The experimental results show that this system provides with the information on the user's hand, which is needed for object manipulation by a multi-fingered robotic hand.

The system can recognize whether each fingertip contacts the omni-directional device or not. The recognition rate is more than 95 %. The system can also completely recognize the types of fingertips based on searching for the thumb. The contact positions of the fingertips on the device surface are also calculated by the geometric information of the omni-directional camera.

Our system works feasibly with inexpensive hardware structure. The system also has the advantage to be immune to the occlusion which occurs in other vision sensor systems. As for the future work, we will improve the accuracy and speed of the measurement and enable this system to deal with various kinds of objects for practical use.

References

- (1) T. Matsuoka, T. Hasegawa, and K. Honda: "A dexterous

manipulation system with error detection and recovery by a multifingered robotic hand", Proc. IROS Vol. 1, pp. 418–423, 1999

- (2) Y. Ueyama and K.Hoshino: "Application of Multiple Linear Regression Models to Robotic Hand Control with Data Glove," IEICE PRMU, Vol.104, pp.19-24, 2004
- (3) N. Tsujiuchi, T. Koizumi, S. Nishino, H. Komatsubara, T. Kudawara, and M. Hirano: "Development of Pneumatic Master-Slave Hand and Joint Control," Transactions of the Japan Society of Mechanical Engineers, Series C, Vol.74, No.741, pp.1267-1272, 2008
- (4) T. Wojtaraa, K. Nonami, et al.: "Hydraulic master–slave land mine clearance robot hand controlled by pulse modulation, Mechatronics", Vol. 15- 5, pp. 589-609, 2005
- (5) K. Hoshino et al:"Realtime Hand posture estimation with self-organizing map for stable robot control", IEICE Trans. on INF. & SYST., VOL. E89-D, No.6, pp.1813-1819, 2006
- (6) M. H. Jeong, Y. Kuno, N. Shimada and Y. Shirai: "Recognition of shape-changing hand gestures, IEICE Trans. on INF. & SYST., VOL. E85-D, No.10, pp.1678-1687, 2002
- (7) K. Honda and Y. Shigeyoshi : "Human-machine Interface for Object Manipulation by Multi-fingered Robotic Hand", Robomec2009, 1A1-D15, Pages, 2009