The Fourteenth International Symposium on Artificial Life and Robotics 2009 (AROB 14th '09),
B-Con Plaza, Beppu, Oita, Japan, February 5 - 7, 2009

Development of an autonomous-drive personal robot "object recognition system using a monocular camera"

Yuki Moritaka Eiji Hayashi
Department of Mechanical Information Science and Technology
Faculty of Computer Science and Systems Engineering , Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka Prefecture, Japan

Abstract: We are developing an autonomous personal robot able to perform practical tasks in a human environment based on information derived from camera images and a knowledge database. In the robot's adjustment to a human environment, it is very important that it be able to recognize the external world and identify specific objects. If an object cannot be identified, the robot cannot determine its next action plan. For this reason, we have developed an object-recognition system that allows an autonomous robot to identify objects within its environment. This object-recognition system is composed of both shape- and color-recognition systems that are based on a sensor using only a monocular camera. The shape-recognition system discerns objects' domains from their background by using region splitting processing. In the shape-recognition system, an object is identified based on its distinctive shape. The color-recognition system identifies an object based on its color information. Here we explain the algorithm used for the object search and the method for processing images. In addition, we describe the results of an experiment carried out to evaluate the system as it performed an object search.

Keywords: Personal robot, monocular camera, Image processing, Object recognition, Region splitting

Ⅰ.Introduction

In the near future, autonomous self-driving robots are expected to provide various services in human living environments. For this to occur, the robots must have a grasp of their driving environment. Therefore, systems to provide environmental recognition from the use of image information are being widely studied. However, they are very difficult to recognize only with the image information, because it is difficult to differentiate objects from their background. We, therefore, are developing an autonomous personal robot able to perform practical tasks in a human environment based on information derived from camera images and a knowledge database.

This object-recognition system is composed of both shape- and color-recognition systems that are based on a sensor using only a monocular camera. The object is identified by processing the image information which is provided by the camera. At first, the shape-recognition system checks to determine if the object is lying on a desk. Next, it discerns objects from their background by using region-splitting processing. This processing is performed based on HSV and RGB information in the camera image, and the system extracts only objects. The shape-recognition system identifies specified object shapes by comparing a shape template in the database to information about the extracted objects. The color-recognition system identifies an object based on its color information. In the processing flow, the system makes a histogram from the hue data based on the object's HSV information and in addition makes a histogram from the object data in the database. It then identifies the object by comparing these two histograms. An object-recognition system searches for the object using both of these recognition systems.

Ⅱ.System for robot

Our robot has a drive mechanism consisting of two front wheels and two back wheels. The front wheels are attached to a motor that operates the wheels on either side independently, while the back wheels function as castor wheels. This method has the advantage of allowing a small turning radius. In addition, in order to acquire image information, a single CCD camera with approximately 300,000 pixels is installed on the head of the robot and can be rotated to all sides by two motors. DC servo motors are used for the robot's drive mechanism, and position and speed control are achieved by the control system of the drive mechanism. The robot also has two arms and hands equipped with sensors, which enable it to respond to the various demands of humans. Finally, an installed wireless LAN can provide remote control for humans. All devices are controlled by a PC, and lead batteries supply the robot's electric power.

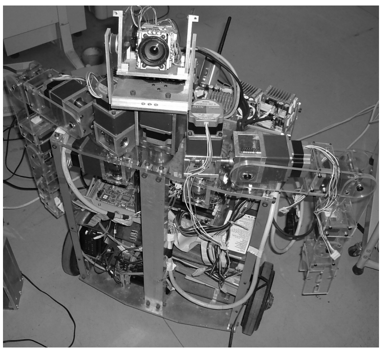

Fig. 1 Our developed robot

Ⅲ Object-recognition system.

3.1 Outline of the system

We developed an object-recognition system for robots that can search objects with image information captured by a monocular CCD camera. This system can search for an object on the assumption, for example, that the object is put on a desk. The system then searches for an object with the shape and color of the object registered by the database. The system notes the shape and color of the object and then step by step narrows down the objects.

3.2 shape-recognition system

In the shape-recognition system, an object is identified based on its shape. Objects that are carried in the hand in daily life are generally not complex in shape, so in this system especially notice the two-dimensional display of the object. The processing flow is as follows.

Ⅰ. Image Acquisition

The image obtained by the monocular CCD camera is read into a PC in the robot. This image is a 24-bit full-color image of the RGB data form.

Ⅱ. HSV conversion

The system converts 24-bit RGB image data into HSV data.

.

Ⅲ. Extraction processing of the desk surface

First, the robot extracts the desk surface which assumed that there is an object. In the processing flow, the system obtains image data from the desk surface by sampling. The robot then calculates the standard deviation of each HSV parameter from the acquired sample data. If the pixel of the image is within a value of threshold of the standard deviation, it is extracted as the surface of the desk.

Ⅳ. Presumption of an object domain

At first, the system performs label processing for all pixel groups except those recognized by the desk surface extraction process. The group of image pixels that lead to the search of an object is distinguished by this process. The system cannot judge accurate shape about the group of image pixels in contacting with the frame of the image. Therefore, this area around the image frame is excluded from the object region. The system determines the size of the object domain of the search object by using width data of the extracted object domain. The size of the object domain is calculated based on registered data for each object in the database. Figure 2 shows an example of the object domain established by this processing.

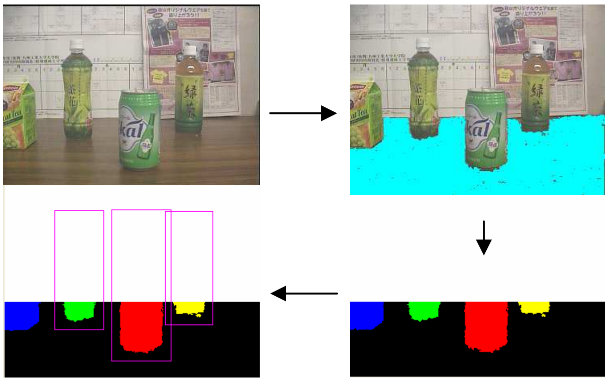

Fig. 2 Object extraction processing

Ⅴ. Region splitting processing

The system uses a plural region-splitting method to acquire only the object domain. The system distinguishes between the object domain and the background by this processing. The region-splitting processing uses two methods. One is a method that splits the domain by using integrated processing of the domain, and the other is a method that splits the domain by using the histogram of the image. The system has previously established knowledge of the desk domain, so it eliminates the desk domain from the processing area at this time. The system is able to extract only the object domain by combining these processes. Figure 3 shows the object domain extracted by this region-splitting processing.

Fig. 3 Result image of region splitting processing

Ⅵ. Template matching processing

This type of processing compares the shape of the extracted object domain with a template (shape data that was registered by the database). The size of the template changes according to size of the width of the object domain. The system partially computes the matching rate of the object shape. By this method, the system can obtain results that take into consideration differences in the shapes of the objects. The system calculates two matching rates, and determines the shape and position of the object as a shape of the template if it is both higher than the threshold and reaches its maximum.

3.3 Color-recognition system

The color-recognition system searches for an object based on color data for the object. The system acquires color data for the object which is searched for by the shape-recognition system. It compares the color data that was registered by the database with the color data of that object, which is then searched for by the shape-recognition system. If the feature of color data of two type accords, it judges as a search object by this processing. The processing flow is as follows.

Ⅰ. Color data Acquisition

The system acquires the color data for the object that was searched for with the shape-recognition system. The color data use information regarding Hue from the HSV information for the object.

Ⅱ. Making the histogram

The image is regarded as an assemblage of pixel values. The color data of Hue for the object is expressed as a value between 0 and 360. This processing divides the Hue data into 72 groups and counts the frequency of each pixel value. Figure 4 shows the histogram.

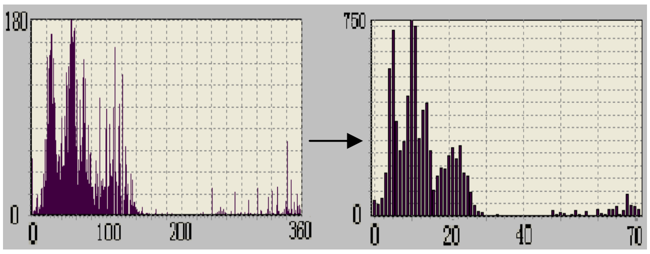

Fig. 4. Histogram

Ⅲ. Comparison of the histogram

Next, in order to identify color similarities, the system analyzes the correlation between the histogram of the object domain that was recognized by the shape-recognition system and the histogram of the search object that was registered by the database. First, the system calculates the regularization histogram. Next, two regularization histograms are overlapped, and the minimum value of the pair of values is taken. The total value from all these minimum values is set as a similarity. Color similarity becomes high when overlap of the histograms is large. In addition, this processing compares the average pixel value and the pixel value that appears most frequently. Figure 5 shows the overlapped regularization histograms obtained by this processing.

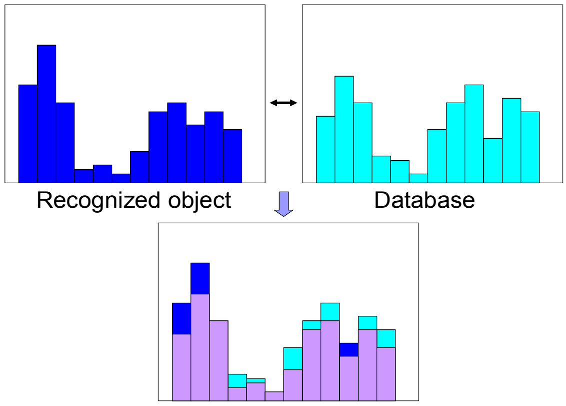

Fig. 5 Comparison of the histogram

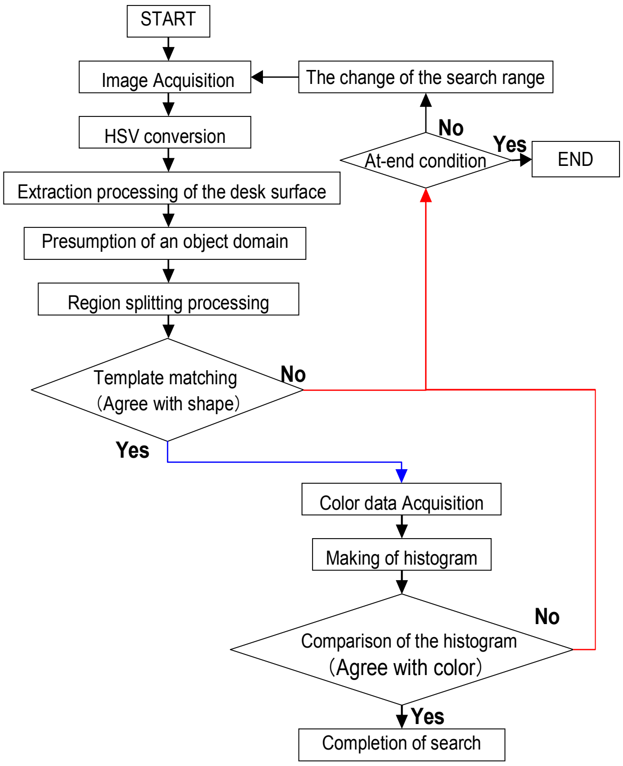

Fig. 6 System flow

Ⅳ. Experiment

We performed the following experiment to evaluate the performance of this system that searches for an object with the shape and color of the object registered by the database.

A: Experiments related to object shape

We conducted an experiment that involved searching for several different types of object forms.

B: Experiments related to object color.

We performed an experiment that involved searching for a specified object from an object that had been searched for by a shape-recognition system using color data that was registered by the database.

In the case of A, the experimental results show that an object of simple shape could be searched for with high probability. However, there was a case in which the system misidentified the object of similar form. There was also a case in which the system could not recognize depending on the results of region-splitting processing. In the case of B, the system was almost able to search for the object by using color data that was registered by the database because the histogram did not appear to be influenced by the visibility of the object. However, there was a case in which the system could not successfully carry out the search depending on the object. Figure 7 shows an example of results for the shape-recognition system, and Fig. 8 shows an example of results for the color-recognition system.

Fig. 7　Shape recognition system

Fig. 8　Color recognition system

Ⅴ.Conclusions

We have proposed a system that searches for objects using only a monocular camera. It is thought that objects held in the hand in daily life are generally not complex in shape. Experimental results show that objects of simple shape can be searched for with high probability. However, there was a case in which the system could not successfully search for the object based on various factors. We therefore must improve the accuracy of the object-recognition system by combination with other processing methods such as those that can display an object in three dimensions. For these reasons, we believe that the present system requires further improvement. In addition, the robot cannot yet acquire distance information of the object. We therefore want to mount on the robot this function. Our next subject of study is to develop a system that makes possible a holding function for the recognized object and a carrying task.

References

[1] Takuma Umeno, Eiji Hayashi, Development of an au tonomous personal robot "The visual processing system for autonomous driving", proceedings of the 10th intern ational symposium on artificial life and robotics、Vol.10th, pp.44, Beppu, Oita, Japan, 2005

[2]Yoshitaka Hane, Eiji Hayashi, User-recognition system for an autonomous robot in human living environments, proceedings of The twelfth international symposium on artificial life and robotics、Vol.12th,pp72-75, Beppu, Oita, Japan, 2007

[3]Toshifumi Uewaki, Eiji Hayashi, Development of an autonomous mobile system for an autonomous robot in an indoor environment, proceedings of The twelfth international symposium on artificial life and robotics、Vol.12th, pp52-55, Beppu, Oita, Japan, 2007

[4]K. Belloulata and J. Konrad, Fractal Image compression with region-based functionality, *IEEE Transactions on Image Processing*11 (2002) , p. 351