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

The recognition of multiple people using an ocellus camera

Seiichi Tanaka 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
(0948-29-7793)
(tanaka@mmcs.mse.kyutech.ac.jp)

Abstract: Communication between a person and a robot is necessary for the robot to be able to function in a human living environment in such a way as to be helpful to the humans. A human being needs to send a command to a robot without having to use exclusive equipment, and this requires that the robot recognize the human being. For this purpose, I have developed in this study a system for recognizing multiple human beings whose images appear in a CCD camera. The purpose of this system is to enable a robot to recognize the position of a user with a stemma camera image.

From this image, the system detects an object which moves, and it then carries out a search in its data domain. It pays attention to the shape and color of the object and narrows it down to either a non-human object or a human being. In addition, it determines the position of a human face and decides if it is a human being from the color, size, and position of the face that it detected.

Keywords**:** Robotics, Image processing, Multiple people recognition

1 Introduction

Due to the insufficient number of workers in Japan's low-birthrate society, autonomous self-driving robots will be called upon to provide various services within human living environments. Robots are currently used in industry, where they simply perform a given motion previously made by humans. However, such robots are less useful for tasks in the home. We are developing an autonomous personal robot with the ability to perform practical tasks in a human living environment using information derived from sensors and a knowledge database.

Our robot has a drive mechanism composed of two front wheels and two back wheels. The two front wheels are attached to a motor, which operates them independently, while the back wheels are castor wheels. DC servo motors are used for the robot's drive mechanism, and position control and speed control are achieved by means of the control system for the drive mechanism. One CCD camera is installed on the head of the robot. It can be rotated to some sides (90 degrees in the top direction, 65 degrees to lower degrees, 90 degrees in the right direction, and 90 degrees in the left direction) by two DC motors. This camera contains approximately 300,000 pixels. All devices are controlled by a personal computer, and electric power is supplied by lead batteries.

To work, the robot needs to receive a command from the human. The robot can be easily sent instructions from devices such as remote controls, personal computers and so on. However, because this step is inconvenient, I developed a system which can recognize multiple human beings who appear in camera images.

With this system, I detect an object which moves from an image that is provided from a CCD stemma camera image, and then I search a data domain. I pay attention to the shape and color of an object which moves, and determine whether it is a human being and then estimate the position of the human face. Furthermore, I determine whether it is a human being by its size and color and the color under a thought, the position of the face which I detected when a human being is piled up.

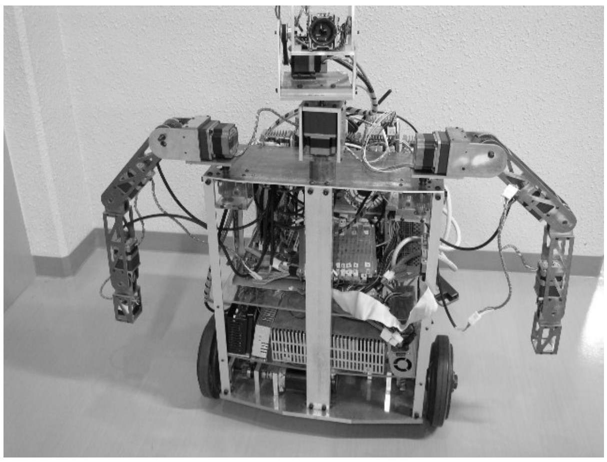


Fig. 1 Robot appearance

2 The recognition of multiple people system

2.1 Outline of the system

From an image provided by a CCD stemma camera, the system detects an object which moves and carries out a search in its data domain. It pays attention to the shape and color of an object and determines whether it is either a non-human object or a human being. In addition, it determines the position of a human face and whether it is a human being from the color, size, and position of the face that it detected.

2.2 Method for recognition of multiple people

Here in Section 2, we explain the method for obstacle detection. The flow for the recognition of multiple people is shown in Fig. 2.

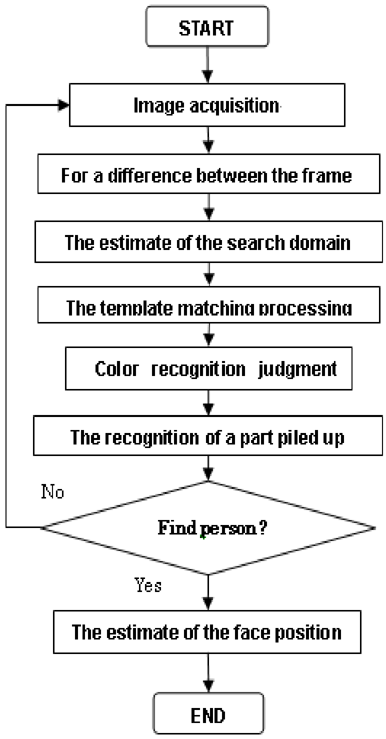


Fig. 2 System flow

Ⅰ. Image Acquisition

The image obtained by the CCD camera is read into a PC in the robot.

Ⅱ. Processing using the difference between frames

We use the difference between frames to extract certain objects moving within the camera's view. Because the system is not easily affected by changes in the background, it can resist the effects of such changes. First, our robot acquires an image from the camera and saves it. Next, it acquires a succession of images and compares them with the first image until differences greater than 2% appear. When differences in the RGB color model exceed this threshold, the system determines that a significant difference has appeared. If it does not detect a moving object, even if it has performed ten comparisons, it photographs the first image again. An image made from differences between frames is shown in Fig. 3.

Ⅲ. The estimate of the search domain

This process estimates the domain of a human being is from the information provided by the differences between the frames.

ⅰ. The detection of the maximum height

This system detects the maximum height Y of each X point in the image from the differential image which it acquired from the differences between the frames. The image from these detections is shown in Fig. 4.

ⅱ. Average

This system makes a smooth graph by creating an average by 40 pixels of values of the height Y which it detected at the maximum, because the position sensing of the domain is difficult only at the maximum height Y. The image of the average values is shown in Fig. 5.

ⅲ. The detection of the search domain

From the graph, the system detects the part which is at the top and estimates the position of a human being. Next, it scans from a person's position and detects the point below the top or the point that is lower than 1/3 from the height of the person's position. I assume that a part surrounded by points is a search domain.

Fig. 3 For a difference

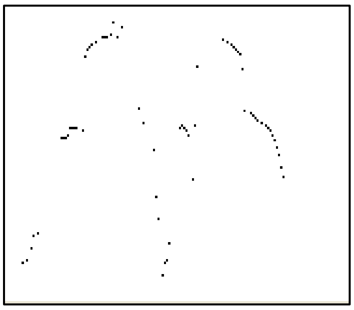


Fig. 4 Height at the maximum

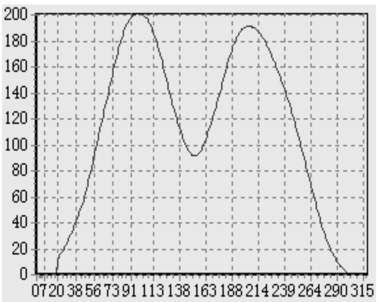


Fig. 5 The graph of the average

Ⅳ. Template matching processing

When this system begins its operation, it reads the template that imitates the head of a person. The outline acquired in the previous process is compared with the template. The size of the template changes according to size of the search range. Generally, this process requires a great deal of calculation time. Real time operation is achieved by reducing the number of comparisons. When the matching rate is higher than the threshold and reaches its maximum, the position is output as the position of a human face.

Ⅴ. Confirmation by using color information

An image is difficult to identify using only conventional processing. The system has to finally confirm that the image is that of a human being. Thus, skin color is used to ultimately determine this, using the template matching process to decide the identity and position of the human image. In this case, color information processing uses the HSV color model.

Ⅵ.The recognition of a part piled up

HSV converts the lower domain that this system took as a human face. Then it detects the part of the external color and performs 2 value. It makes a histogram of external color pixels for the X and Y coordinates and judges the position of the face from the histogram. The image which detected as an external color part is shown in Fig. 6. The image of the histogram is shown in Fig. 7.

When our system performs these processes and determines that the head of a person is in view, it outputs the position information.

Fig. 6 External color extraction

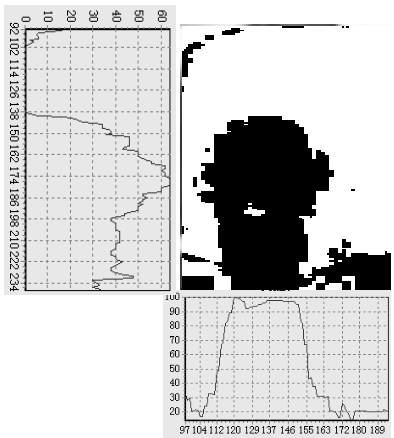


Fig.7 Histogram

2.3 Experiment

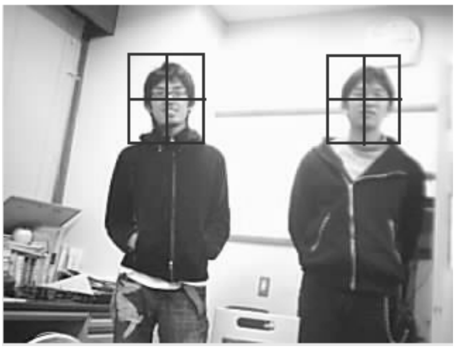
The system's performance was evaluated under two conditions.

A: Multiple people under normal lighting
B: A great number of people under normal lighting
C: Multiple people under strong lighting

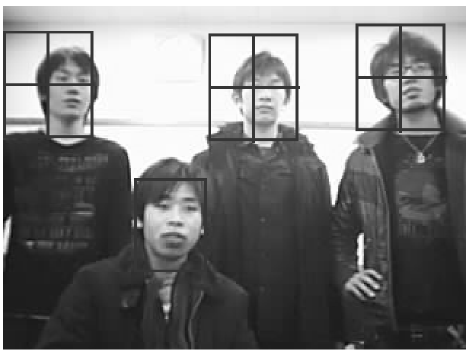
The strong lighting condition in case C was made by guessing a fluorescent lamp and sunlight right. The experiments were conducted in a conventional human living environment..

In case A, this system was seen to operate effectively. In case B, it was seen to detect a human, but when a head overlaps, detection is difficult. In case C, the system has a slight difficulty in identifying the person' s color, so the robot' s ability to follow the person under this condition deteriorates.

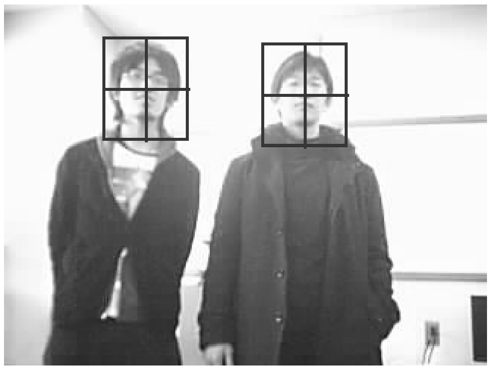
We consider that these results demonstrate that the system can detect multiple people.



(A)



(B)



(C)

Fig. 8 Experiment

3. Conclusions

We have proposed a system, composed of an ocellus camera and intended for use in a indoor environment, that extracts a human's position from a field of view. This system constitutes the beginning for communication between a robot and a human being.

Our next subject of study is the development of a program to recognize the position of the user.

References

[1] Toru Azuma, Yoshihiro Takenaga, Eiji Hayashi, Development of a self-driven personal robot、Artificial Life and Robotics(AROB 8th '03)、pp.1-3、Oita, Japan, 2003

[2] Yoshitaka Hane, Eiji Hayashi, User-recognition system for an autonomous robotin human living environments Artificial Life and Robotics(AROB 12th '07)、Oita, Japan, 2007