The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

A user recognition system using a stemma 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 a user 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.

This system detects a moving object from an image provided by the stemma of a CCD camera and estimates a search domain. Using color and form information, it judges whether a human being exists in a camera image. In addition, using the zoom function of the camera and an obtained detailed image, it detects the position of the eyes and nose. It also detects features around obtained portions of an image. Finally, it identifies a user by comparing the features with information in a database.

Keywords**:** Robotics, Image processing

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 in 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 by 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 a 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 a user who appears in camera images.

This system detects a moving object from an image provided by the stemma of a CCD camera and estimates a search domain. Using color and form information, it judges whether a human being exists in a camera image. In addition, it obtains a detailed image of a face using the zoom function of the camera, detecting the position of the eyes and nose from an obtained detailed image. It also detects features around obtained portions of an image. Finally, it identifies a user by comparing the features with information in a database.

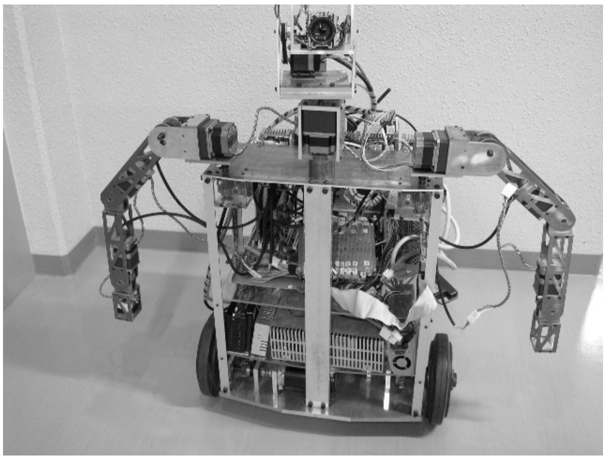


Fig. 1 Robot appearance

2 A user recognition system

2.1 Outline of the system

From an image provided by a CCD stemma camera, the system detects a moving object and carries out a search in its data domain. It pays attention to the shape and color of the object and determines whether it is 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. Finally, it takes a zoom picture of the face and determines its features from the obtained expanded image.

2.2 Method for a user recognition system

Here in Section 2, we explain the method for obstacle detection. The flow for a user recognition system is shown in Fig. 2.

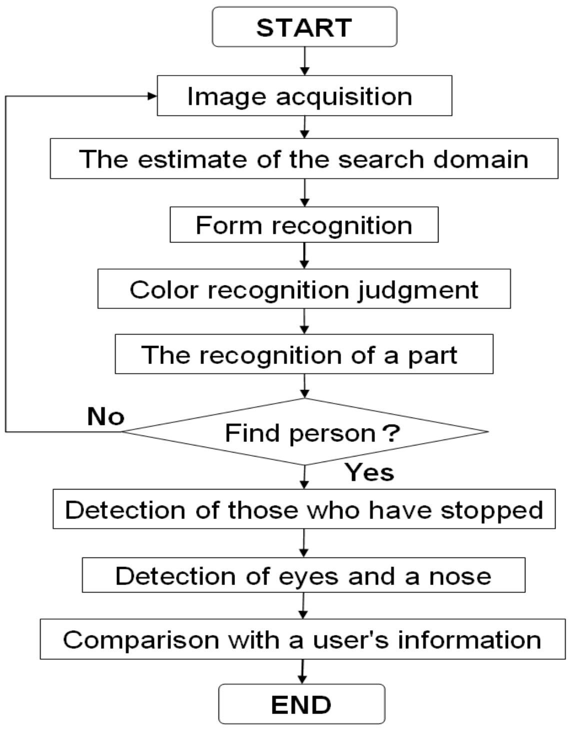


Fig. 2 System flow

Ⅰ. Image Acquisition

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

Ⅱ. The estimate of the search domain

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

ⅰ.The detection of the difference between frames

This system uses 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. When differences in the RGB color model exceed a threshold, the system determines whether a significant difference has appeared. An image made from differences between frames is shown in Fig. 3.

ⅱ. The detection of the maximum height

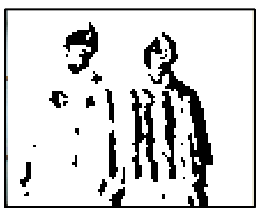
This system detects the maximum height Y of each X point in the image from the differential image which it acquired from the differences between the frames. The image from these detections is shown in Fig. 4.

ⅲ. Average

This system makes a smooth graph by creating an average by 40 pixels of the values of the height Y which it detected at the maximum. It does this 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.

ⅳ. 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.



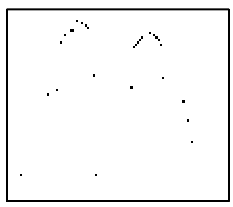


Fig. 3 For a difference   Fig. 4 Height at the maximum

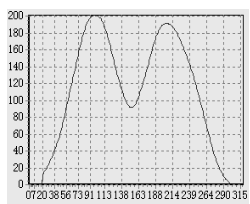


Fig. 5 The graph of the average

The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Ⅲ. Form recognition by template matching

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 the 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.

Ⅳ. Confirmation by using color information

An image is difficult to identify using only conventional processing. The system has to definitely 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.

Ⅴ.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. It makes a histogram of external color pixels for the X and Y coordinates and then judges the position of the face from the histogram. The image which is detected as an external color part is shown in Fig. 6, and 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

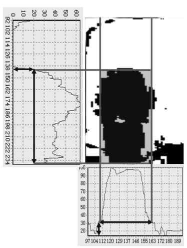


Fig.7 Histogram

Ⅵ. Detection of those who have stopped

In order to detect a human being who tries to take a robot and communication, the human being who is not moving within the picture is distinguished from the human being who is detected. At this time the system detects the human being who has previously appeared in the processing and whose face has not changed its position in the image.

The system zooms on that position in order to acquire information from a detailed image of the face. The image which is detected as that of the stop person is shown in Fig. 8, and the zoom image is shown in Fig. 9.

Fig. 8 Stop person

Fig. 9 The zoom image

Ⅶ. Detection of the part of a face

In order to judge a user, the eyes and nose in a face are recognized. Therefore, the frequency of a certain angle in the image is detected. The detected image at a frequency of 90 degrees is shown in Fig. 10.

The system creates a histogram from the features at the detected frequencies and specifies the eyes and a nose as the features that are detected as a result of the histogram. The image of the detection of eyes and a nose is shown in Fig. 11.

Fig.10 90 degrees

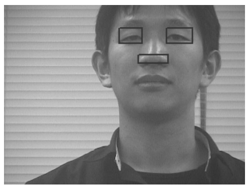


Fig.11 the detection of the part

Ⅷ. Comparison with a user's information

This system judges a user using two processings. First, the position of the detected eyes and nose is compared with the information in a database. Next, it averages the frequency component of each angle. Then it matches these averages with the template of the eyes in a database. The image of detection with the frequency component of each angle is shown in Fig. 12.

Fig. 12 The frequency component of each angle

2.3 Experiment

We performed the following experiment to evaluate the performance of this system for user recognition using a stemma camera.

We have recognized those who are not a user and a user using this system.

The system has recognized users registered by the database. However, when the zoom function of a camera was not appropriate, the system made incorrect detections. When a not user is recognized, It may have been rarely recognized as the user.

The image of user recognition is shown in Fig. 13. The image of not user recognition is shown in Fig. 14.

Fig. 13 User

Fig. 14 Not user

3. Conclusions

We have proposed a system, composed of a stemma camera and intended for use in a indoor environment, that extracts a user's position from a field of view. This system constitutes the beginning for communication between a robot and a human being. Experimental results show that a user can be detected with high probability. However, there was a case in which the system could not successfully search for the user because of various factors. Therefore we propose raising the recognition accuracy by making a nose and a mouth consistent. For these reasons, we believe that the present system requires further improvement. Our next subject of study will be to develop a system that makes possible the recognition of a user's gesture.

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,pp. 72-75, 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, pp. 52-55, 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