User-recognition system for an autonomous robot in human living environments

Yoshitaka Hane

Graduate School of Computer Science and System Engineering
Kyushu Institute of Technology
680-4, Iizuka City, Fukuoka, Japan, 820-8502
hane@mmcs.mse.kyutech.ac.jp

Eiji Hayashi

Department of Mechanical Information Science and Technology
Kyushu Institute of Technology
680-4, Iizuka City, Fukuoka, Japan, 820-8502
haya@mse.kyutech.ac.jp

Abstract

In this report, we describe a user-recognition system for an autonomous robot. The proposed method enables the robot to recognize the person who instructs it without placing any markers in human living environments. The system is based on differences information between some camera's images, matches the template, and recognizes color based on HSV information. The threshold can be lowered by combining features of these in formations. This system enables recognition under various environments.

A moving object is initially extracted by using the difference between frames. Next, to reduce the load, the image is narrowed to the area on which the difference is concentrated. The system matches the template and forecasts the user's position. In addition, it determines a flesh-colored area by using HSV information. When two or more people are present in the field of view, the system determines that the person with the largest-width silhouette is closest to the robot. After confirming that the person is standing still in its presence, the robot waits for the potential user to identify his or her intention to use the robot; this is signaled by the user raising his or her right hand. This process is repeated every 100 milliseconds.

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 the 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 of the drive mechanism. One CCD camera is installed on the head of the robot. It can be rotated to some sides (90 degrees to top direction, 65 degrees to lower degrees, 90 degrees to right direction, and 90 degrees to 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 user. The robot can be easily sent instructions from devices such as remote controls, personal computers and so on. However, this step is inconvenient. We consider that the robot should recognize human activity via its visual and auditory sensors, and thereby understand a command so that it may allow human users to act more naturally in their living environments. In light of this, we have developed a system based on images taken by the CCD camera installed in the robot aimed at determining the face of the human who has made a command. The information is passed on to the human instruction recognition system and the human tracing system.

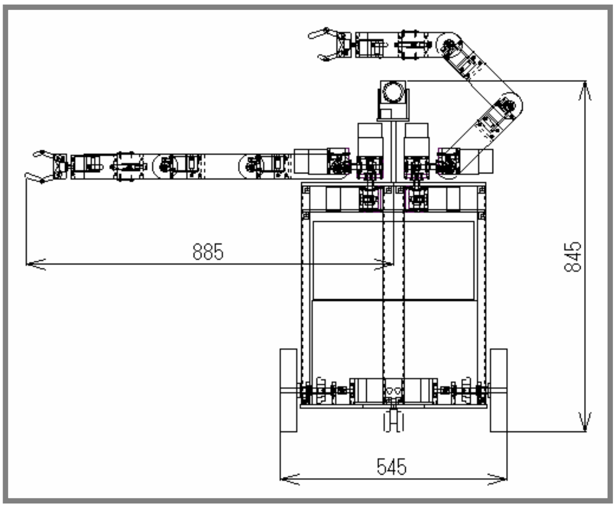


Fig.1 our developed robot

2 User-position recognition system

2.1 Outline of the system

To communicate with people, the robot must confirm the position of people within its field of view. We have thus developed a system that can achieve position recognition. The system is informed by the fact that people move slightly even after they stop, unlike a static object. It certain moving objects from the camera image, and selects one prospective area from them. It determines whether the image is that of a human being by comparing the image with certain human characteristics. When it is judged that the image is that of a human being, this system outputs the position information.

2.2 Method for estimating distance error

By this system, we consider that this system is able to operate in real time and that it is not affected by changes in background

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. It is hard to be affected by a change of a background. This allows this system to resist being affected by changes in background. First, our robot acquires an image via the camera and saves it. Next, it acquires a succession of images. They are compared with the first image until differences of more than a total of 2% appear. When differences of RGB color model exceed the threshold, the system determines that a difference appears. When a moving object is not detected, even if the system performs ten comparisons, it photographs the first image again. In this way, our robot acquires a binary image including a human outline (Fig.2).

Ⅲ. Determining one object for range processing

The purpose of this system is to determine the position of a human face. The view does not have to be obtained unless a human being is present in the image. Therefore this system can treat an image in less view in other processes. At first, our system creates a histogram of the binary image in the lengthwise direction, the image having been acquired in the previous process, and calculates the average value of the histogram. Some smaller-than-average values are taken as noise, and they are removed. The value of the histogram at a near position is considered one block. (Fig. 3) The widest of the blocks is judged the range of the human being. By doing so, the robot recognizes the nearest human being as a user. (Fig. 4) From this point on, only the range of this image is determined.

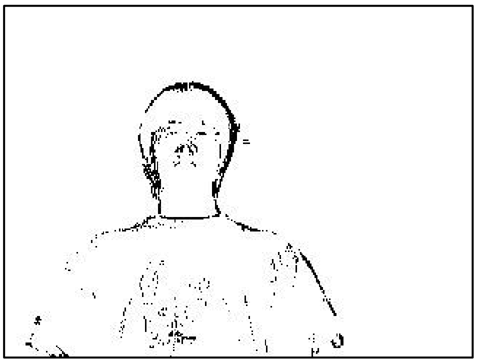
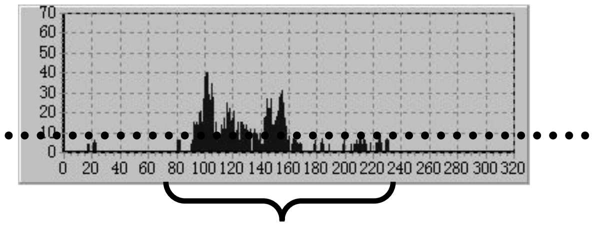


Fig.2 Difference between flames



One block

Fig.3 Lengthwise histogram of Fig2

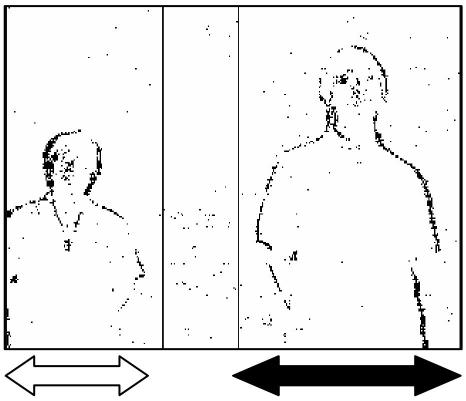


Fig.4 two human beings in view

Ⅳ. Template matching processing

When this system begins its operation, it reads the template imitating the head of a person. The outline acquired by 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. Thus, 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, using the template matching process, the identity and position of the human image. In this case, color information processing uses the HSV color model.(Fig.5)

When our system performs these processes, and it is determined that the head of a person is in view, it outputs the position information. The flow of this system is shown in Figure 6.

Fig.5 Check with HSV information

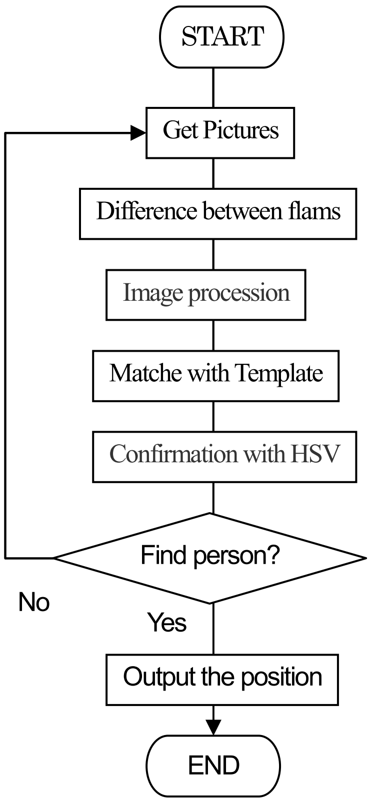


Fig.6 System flow

2.3 Experiment

The system's performance was evaluated under three conditions.

A: One person under normal lighting
B: Two person s under normal lighting
C: One person under strong lighting

The strong lighting condition in case C was a fluorescent light placed above the robot's head, coming into view when the robot turns its head upward. 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 able to identify a person and follow him or her throughout the field of view, but could not continue following the particular person and irregularly shifts its focus from object to object. 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. (Fig.7)

We consider that these results demonstrate that the system can detect a person and follow him or her in a human living environment.

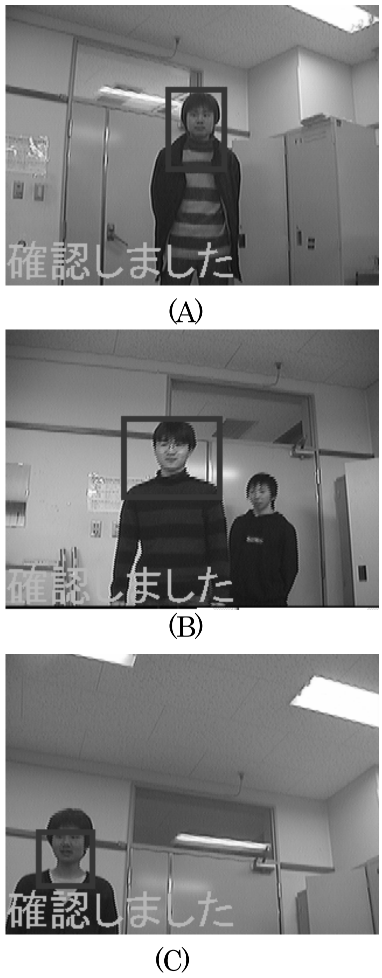


Fig.7 Experiment

3. Sign recognition system

3.1 Outline of the system

The user-position recognition system outputs the position of a person within the robot's field of view. However, the robot has to determine whether the person requests it. We created a rule in which the human operator, after the robot's face (camera) turns in a person's direction, must raise his or her right arm to the side of the face. After confirming the signal, the system recognizes the human being as the user.

3.2 Method of sign recognition

The robot recognizes a humans' position by repeatedly using the user-position recognition system. When these outputs show values close to one another's, it is determined that the person is a generally fixed position. The robot then initiates the sign recognition system. First, the system establishes a recognition range beside a human face and saves the image. The system photographs the range in succession again. The image is compared with an image obtained at the beginning of the process. In the range, when differences exist within the range and the user's skin color is detected, the system identifies the user's signal.

3.3 Method for the first position of the arm recognition

As soon as this robot recognizes the sign from the user, this system gets the first position of the hand. The difference image between before and after raising the user's right hand in prior process is cleared away the small noise by the Closing and Opening methods. Next, this system puts labels on left pixel blocks, judges the biggest blocks as the palm of hand, and erases the other small blocks. The slope of the palm is calculated with the next model. (Fig.8, 9, 10, Formula (1), (2))

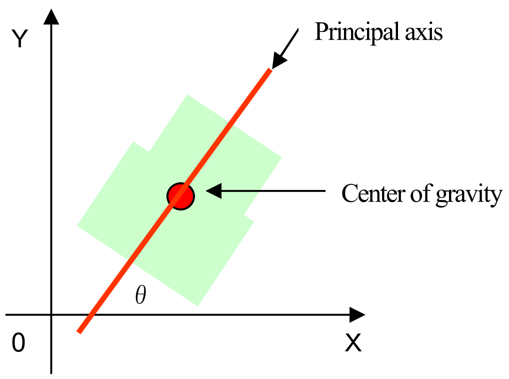


Fig.8 Principal axis

$$M(p,q) = \sum_{i,j} i^p j^q \qquad (1)$$

Formula (1) is for calculating "moment feature". The "i" and "j" are each position on the x-axis and the y-axis.

$$\tan^2\theta + \frac{M(2,0) - M(0,2)}{M(1,1)} \tan\theta - 1 = 0 \qquad (2)$$

Formula (2) is for calculating the slope of the principal axis. When this formula is used, the center of gravity has to move to the origin of the coordinates.

Fig.9 Recognition a signal from user

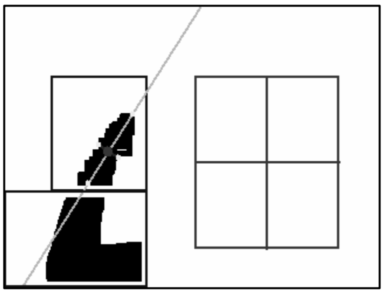


Fig.10 Different image of the right hand

4. Conclusions

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

Our next subject of study is the development of a system for tracking the movement of a human arm using an initial value as the position of the right arm which the sign recognition system outputted. The robot will communicate with people effectively by identifying the direction of the human arm and acting upon it.

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