Face Detection System by Camera Array that Satisfies both Wide View and High Resolution

Kosuke Okabe* Hiroshi Mizoguch**
Takaomi Shigehara* Kazuyuki Hiraoka* Masaru Tanaka* Taketoshi Mishima* Shuji Yoshizawa*
*Department of Information & Computer Sciences
Saitama University
255 Shimo Okubo, Saitama 338-8570, Japan

**Department of Mechanical Engineering
Tokyo University of Science
2641 Yamazaki, Noda 278-8510, Japan

Abstract

This paper presents a vision system that is possible to detect human face in wide view. The system has an advantage that it is capable to simultaneously satisfy incompatible requirements of wide view and high resolution. The authors have studied human symbiotic and collaborative robot. As a part of our effort to realize such robot, the authors have been developing face recognition system. During the development, we have faced a trade-off between wide view and high resolution. When the resolution is increased, i.e. image is closed up, so that a face can be stably detected, field of view becomes narrower. On the contrary, when the filed of view is widened, then the resolution becomes lower and the face detection becomes fail. To resolve this trade-off, the authors introduce camera array of more than one cameras. The camera array enables to satisfy the above mentioned incompatible requirements simultaneously. The authors have already implemented a working prototype system of the camera array based upon the above idea. And we have conducted experiments using the prototype. The experimental results demonstrate feasibility and effectiveness of the idea behind the system. In our experiment, the field of view is widened at 1.7 times.

1 Introduction

The authors have studied human symbiotic and collaborative robot. As a part of our effort to realize such robot, the authors have been developing face recognition system [1]. During the development, we have faced a trade-off between wide view and high resolution. When the resolution is increased, i.e. image is closed up, so that a face can be stably detected, field of view becomes narrower. On the contrary, when the filed of view is widened, then the resolution becomes lower and the face detection becomes fail.

For example, this problem cannot be ignored when combining a speaker array [2] and the face recognition system. Since the view range of a single camera is relatively narrow, it is limited to the view range that the target person can hear sound. Thus, in this paper, the authors solve this problem by using the camera array which consists of two or more set cameras to extend the view range. The camera array satisfies both wide view and high resolution simultaneously.

2 Camera-Array System

2.1 Problem

Fig. 1 shows the relation between resolution and a view. Here, the method of carrying out face discovery is considered in this case using the face discovery system using template matching based on the correlation operation which authors developed. The size of a template is 16x16 pixels. As for the size of a picture, both Fig.1(a) and Fig.1(b) are 120x90 pixels. As shown in fig.1(a), since the resolution of the face for a screen falls, the problem that a face cannot be discovered occurs. On the contrary, if resolution is raised to the grade which can discover a face in the state of Fig.1, the view range of a camera becomes narrow. That is, the view range and resolution have the relation of a trade-off.

2.2 Solution

There is a method of pursuing a candidate person using two or more cameras [3]. In this paper, solution of the problem of the trade-off stated to the foregoing paragraph by using the camera array which consists of two or more fixed cameras is aimed at. Fig.2 shows the Configuration of Camera Array. As shown in a Fig.2, two or more cameras are aligned out so that a mutual view may overlap. A camera is connected with a computer, respectively and performs face detection independently. Each camera sends the information on a success or failure of face discovery in a charge view to the integrated server in a figure. In addition to this, when a face is discovered,

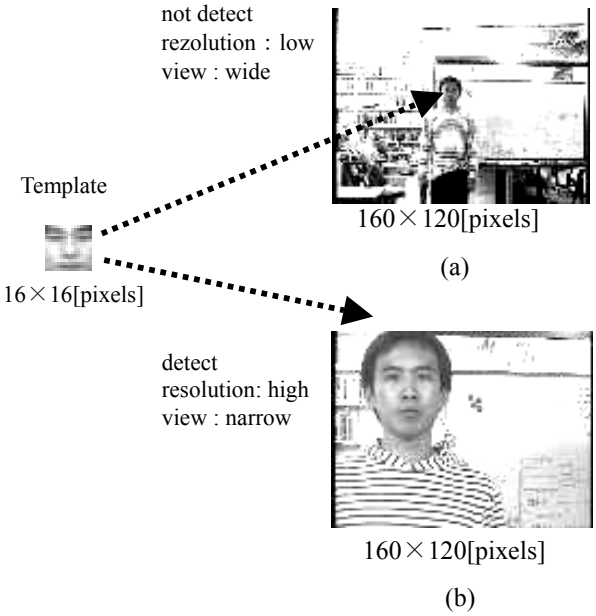


Fig. 1： Relation between resolution and a view

face position information is transmitted. Based on a face discovery success and failure, an integrated server judges the validity of the sent data. When people exist in the domain which the view of a camera overlaps, although a face will be found with two or more cameras, naturally the average value of the sent data in that case is determined as a face position.

Thus, the range of a view can be extended by using a camera array. Since two or more cameras share a view, it is not necessary to extend the view of each camera, therefore resolution does not fall. In consequence of this, the phenomenon in which a face cannot be discovered by the low resolution is avoidable.

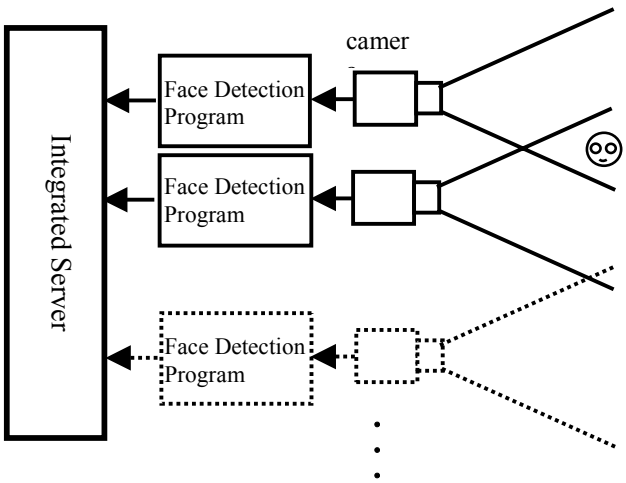


Fig. 2：Configuration of Camera Array

3 Experiment

3.1 Method

The authors have built a prototype camera array system in order to prove the validity of the above-mentioned idea. Fig. 3 shows the appearance of an experiment system.

Fig.4 illustrates the block diagram of the implemented camera array system. As shown in fig. 4, the method of determining the position of a face passed the information on the position of the face discovered by each face discovery program to the face position output program, and it realized by sending each position information to an integrated server using Socket.

The contents of an experiment are shown below. The field of view of one camera in camera array conform to the field of view of a single camera. In other words, as the number of the cameras in a camera array increases, a view becomes larger from a single camera. In this state, It investigates whether a face can be detected or not. The number of the cameras which constitute camera array is two. Face detection uses the face detection system by the vision which the author implemented. Furthermore, Table.1 shows the conditions for execution of face detection.

Table 1：Experimental Condition.

OS	Linux
CPU	Pentium4 2.8GHz
Multiple resolution	10 level
Skin Color Extraction	use
Size of Template	16×16[pixels]

Fig. 3： Figure of System

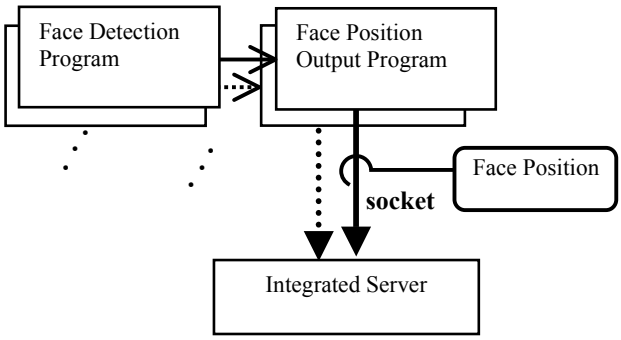


Fig. 4：Block Diagram of Camera Array

3.2 Experimental Result

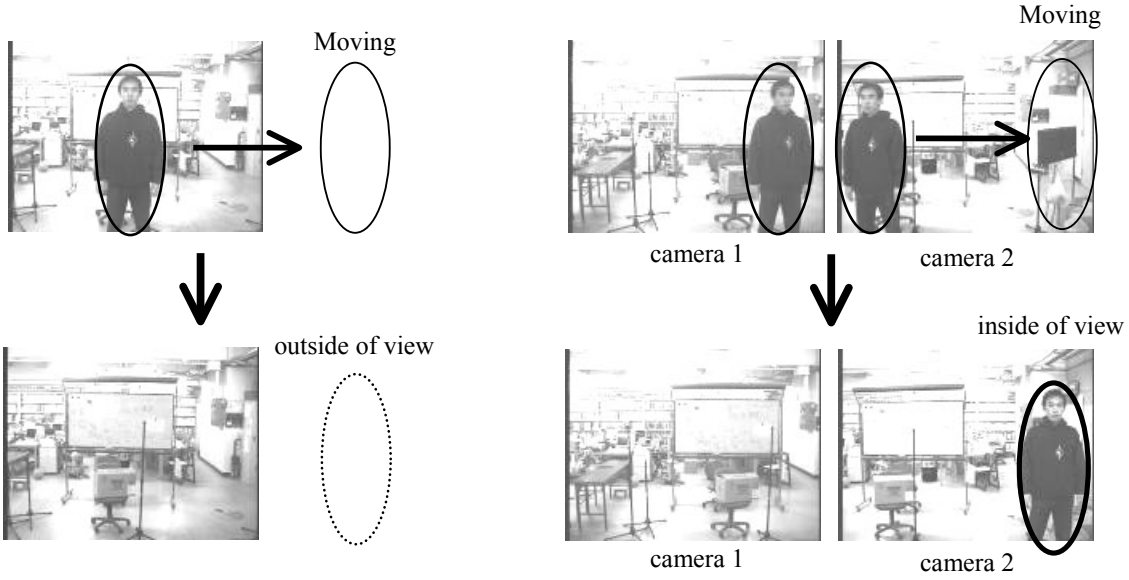
The result of the experiment is presented. Fig.5 shows comparison of the view of camera array and a single camera in actual operation environment. At this time, the size of the image of the single camera and camera array is 320x240 pixels. In the case of single camera of Fig. 5(a), if the person moves in the direction indicated by arrows , a single camera loses sight of person. To the contrary, in the case of Camera Array of Fig. 5(b), person is in the view of a camera 2. For this reason, when resolution is the same, it is the view that the camera array is larger than a single camera. There we measure actually width of a view. Fig.5 shows the actual measurement of the view range. Fig.5(a) is a view range in the case of one camera. and Fig.5(b) is a view range in the case of a camera array. In consequence of this, in the case of same resolution, the view of a camera array is 1.7 times as large as a single camera.

Next, the face discovery by the camera array is described in the following. Fig. 8 shows time transition of the face detection in each camera. In an experiment, person walks from the left end of a camera 1 to a camera 2. the position of detected face among Fig.8a is given a superimposition indication of the frame. Fig. In 8 (a), a candidate person moves to (6) from (1). In addition to this, (3) and (4) are each picture of the cameras 1 and 2 of same time. In consequence of this, Even if it disappears from the view of a camera 1, it turns out that it has detected in a camera 2 continuously.

Fig.8 (b) is the graph showing time transition of the position of the detected face. The vertical axis of Fig. 8(b) shows the position where a face is horizontal. namely, shows the value of the x-coordinate of Fig. 8(a). The value of the horizontal position of a face is a value which averaged the value acquired from cameras 1 and 2 in the integrated server. As shown in a figure, the center of a camera 1 is the points of origin. The horizontal axis of Fig.8b shows time. The unit of a horizontal axis is the number of frames. In consequence of this, A camera array can perform face detection in the range larger than a single camera.

4 Conclution

The single camera has trade-off that a large-scale field of view and big high resolution are incompatible. Then, the problem can be solved using the camera array. As a result, it succeeds in realizing face detection covering a 1.7 times



(a) Single Camera (b) Camera Array

Fig. 5. Comparison of a view

as many large area as this compared with a single camera. The future works are expansion of much more view range with three or more cameras and detection and simultaneous tracking of the face of two or more person.

Acknowledgements

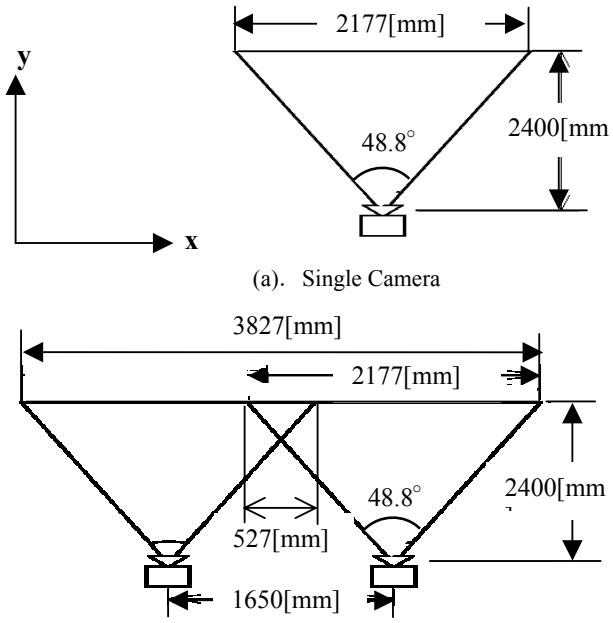
This work has been partly supported by CREST of JST (Japan Science and Technology Corporation). The workalso has been supported in part by Grant-in-Aid for Scientific Research by JSPS.

References

[1] HIDAI Ken-ichi, et al, "Robust Face Detection Against Brightness Fluctuation and Size Variation", *Proc. of IROS2000*, pp. 1379-1384 ,2000.

[2] MIZOGUCHI Hiroshi, et. al, "Implementation of Invisible Messenger System to Whisper in a Person's Ear Remotely by Integrating Visual Face Tracking and Speaker Array", *Proc. of SMC2002*, pp.WA2N4(1)-(5), 2002.

[3] Norimichi Ukita, et al , "Real-Time Cooperative multi-Target Tracking by Communicating Active Vision Agents", *Proc. of ICPR2002*, pp.14-19, 2002.



(a). Single Camera

(b). Camera Array (two cameras)

Fig. 5. View of Camera

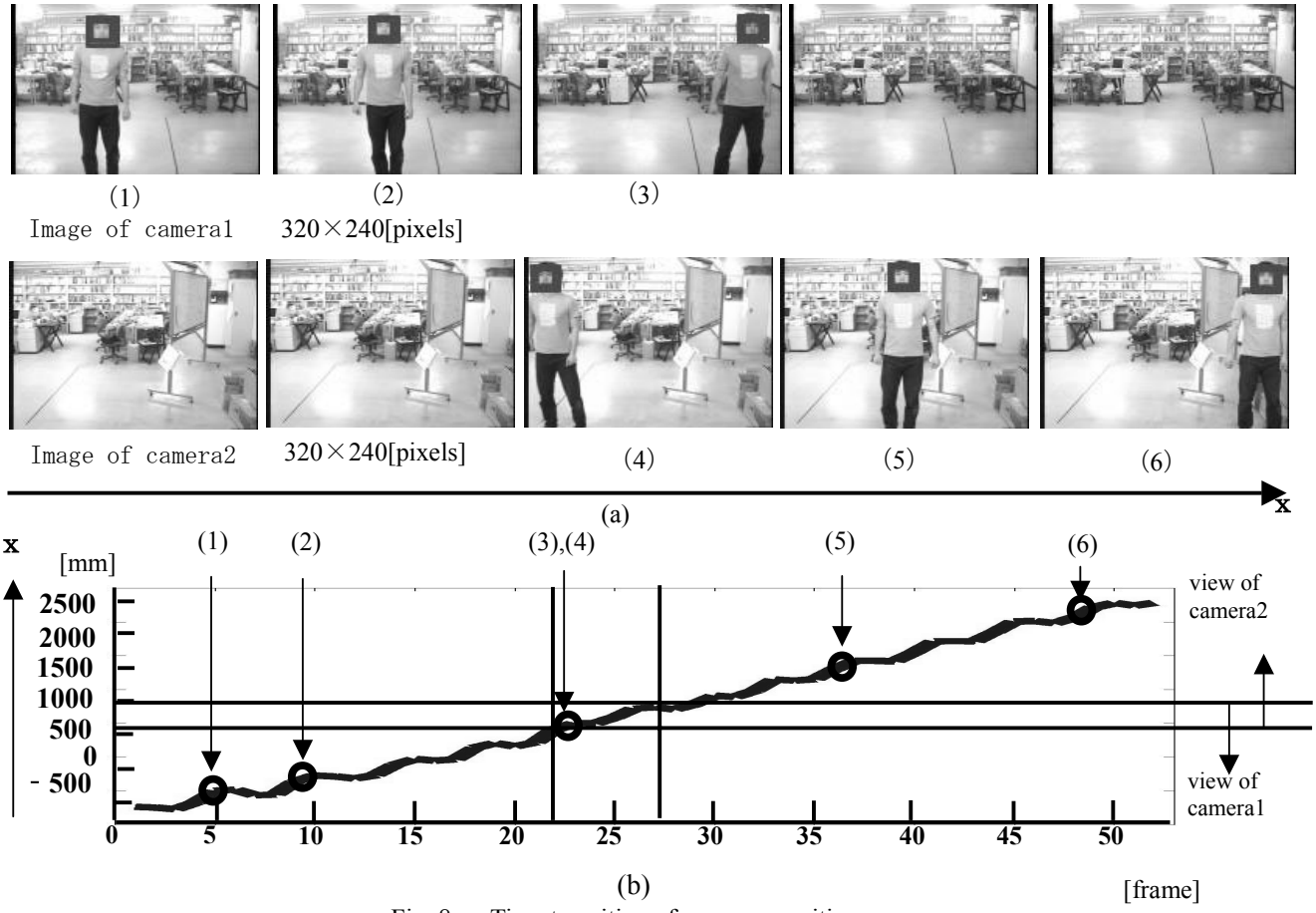


Fig. 8： Time transition of a person position