Privacy protection by concealing persons in circumstantial video image

Suriyon TANSURIYAVONG
Electrical Engineering, Nagaoka University of Technology,1603-1 Kamitomioka,Nagaoka,Niigata, 940-2188,Japan.
Telephone number: +81-258-47-9538
E-mail: suriyon@vos.nagaokaut.ac.jp

Shin-ichi HANAKI
Electrical Engineering, Nagaoka University of Technology,1603-1 Kamitomioka,Nagaoka,Niigata, 940-2188,Japan.
Telephone number: +81-258-47-9524
E-mail: hanaki@vos.nagaokaut.ac.jp

ABSTRACT

A circumstantial video image should convey sufficient situation information, while protecting specific person's privacy information in the scene. This paper proposes a system which automatically identifies a person by face recognition, tracks him or her, and displays the image of the person in modified form such as silhouette with or without name, or only name in characters (i.e. invisible person). A subjective evaluation experiment was carried out in order to know how people prefer each modified video image either from observer or subject viewpoint. It turned out that the silhouette display with name list seems to be most appropriate from the balance between protecting privacy and conveying situation information in circumstantial video image.

Keywords

circumstantial video image, privacy, face recognition, person identification

1. INTRODUCTION

In order to promote informal communications in office environment, several attempts have been made, which include sending circumstantial video images all the time between places relating to group activities. In the VideoWindow system[1], coffee rooms as public spaces were connected by circumstantial video images and in the PolyScope[2], coarse images of individual rooms or common areas with several control conditions were shown each other on the terminal display for privacy concern.

The circumstantial video image provides awareness information[3]-[5] that is considered to be useful to help starting conversation either in face-to-face talking or talking to a person through the screen in the telecommunication system.

The circumstantial video image is useful to provide awareness information, however, it cannot but give persons observed by camera some mental pressure caused by the fact that they are always monitored. In the VideoWindow system[1], it was observed that some people intentionally got rid of the camera field so that they should not be observed by the other party.

In the PolyScope[2], in order to reduce a mental resistance of users, only low resolution images were used, which still give some awareness information to others, but no detail of circumstantial objects irrelevant to the person's privacy are conveyed.

The circumstantial video images are sometimes useful for security monitoring purposes. In such a purpose, details are important with an image as fine as possible. Therefore, It is necessary to provide the means that protect personal privacy while conveying enough information about the circumstantial objects.

The present paper proposes and realizes a system that makes a circumstantial video image protect specific person's privacy information and convey sufficient situation information in the scene.

The realized system can automatically process in real time to identify a person by face recognition, tracks him or her, and conceal the person in the screen by displaying that person's image in the modified form such as silhouette with or without name, or only name in characters that is equivalent to an invisible person instead of the original video image, while it clearly displays the other portions.

2. USAGE OF CIRCUMSTANTIAL VIDEO IMAGE

The circumstantial video images provide awareness information that helps in making decision in the following situations;

(1) To start talking to a person.

Through the circumstantial video images, one can check if the person aimed at is at the place or not, and further estimate if the person is in an appropriate situation to visit and start talking or not.

(2) To find out the location of a person designated at a telephone call.

When one picks up telephone handset in response to ringing, there is no problem, if the person designated is within the scope. It is troublesome, however, if the person designated is not seen near and still strongly requested. Then, one has to start looking for the person. The circumstantial video images will provide opportunity to find out the person in seek, or provide information, at least, that the person is not in the room where the video shows.

(3) To know the availability of shared resources.

PUI 2001 Orlando, FL USA

In the laboratory, measuring apparatus, specific tools, or small free meeting spaces apart from office desks are often shared. In such a situation, it is convenient to know from the remote location through video images, whether the resource is occupied by somebody or not.

(4) To monitor for security purposes.

When something unusual is found in the circumstantial video images, for example, smoke comes out from the machine in the empty room, or a strange person is going to carry out a personal computer from the laboratory room, an appropriate action should immediately be taken. For this specific purpose, video image should be as clear as possible.

3. REALIZATION OF THE PERSON CONCEALMENT SYSTEM

3.1 System Configuration and Functions

By making use of video cameras, the developed system identifies a person entering into the room and tracks him/her, then it conceals the person on the display, if it is necessary. These processes are executed automatically in real time for the circumstantial video image obtained from the video camera. The face recognition technique was employed so that people can be identified without taking any explicit action; such as carrying something like badge, sensing device, etc.

In order to obtain a front face image appropriate for recognition process with enough resolution, a video camera was set near the entrance door. Another camera that takes circumstantial video images of the most areas of the room is set, which also covers the entrance door as a part of the taken picture.

Only the modified video image taken by camera C2 is displayed to others as the circumstantial video image, and the face image taken by camera C1 is not sent to others.

The system configuration is shown in Fig.1. The image from the camera C1 set near the entrance door is used for the person identification. It is input into the PC Linux, and the recognition results are transmitted to the host workstation through the LAN.

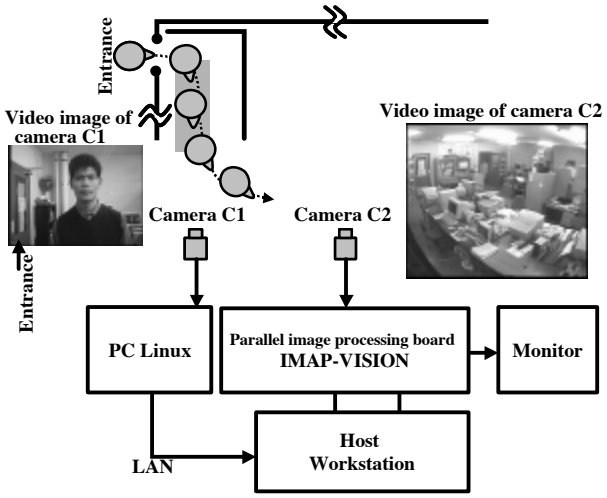
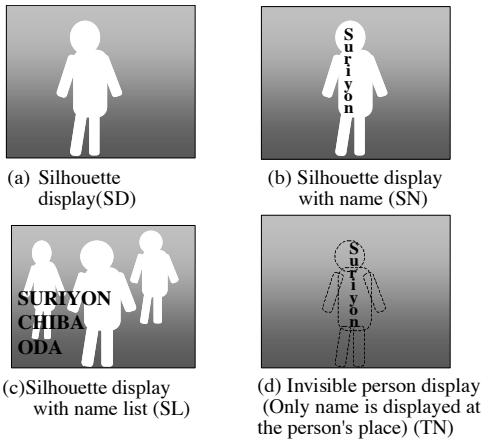


Fig. 1 System configuration.



(a) Silhouette display(SD)

(b) Silhouette display with name (SN)

(c)Silhouette display with name list (SL)

(d) Invisible person display (Only name is displayed at the person's place) (TN)

Fig.2 Sketch of person concealment.

If the recognized person is a member of the group who needs privacy protection, then the host workstation gives the order to a parallel image processor IMAP-VISION[6] to modify the image portion of the person by an appropriate concealing method in the circumstantial video images taken by the camera C2. A person is identified and given an ID-code by the face recognition when he/she enters the room, and the image portion of the person in the circumstantial video images obtained by the camera C2 is tracked by giving the same ID-code. Therefore, the image of a person is easily concealed by one of the following techniques (sketch of concealed pictures are shown in Fig.2);

(1) Silhouette display (SD),

(2) Silhouette display with name (SN),

(3) Silhouette display with name list (SL), and

(4) Invisible person display (TN), Only a name is displayed at that place without person's image.

3.2 The concealment display techniques

- Silhouette display (SD)

It offers information about the number of people in the scene and a rough idea about posture. It is difficult, however, to guess who the person is, or facial expression because no detail is displayed.

- Silhouette display with name (SN)

It provides information same as the silhouette display (SD), adding the name of the target person.

- Silhouette display with name list (SL)

It displays name list and silhouette separately. All names are displayed in the list at the corner of the screen. User cannot know completely which silhouette is whose, if there are more than two people in the scene.

- Invisible person display (TN)

It will first display the old background image, where no person is, and then the person's name is overlayed at the corresponding position of the present location of the person.

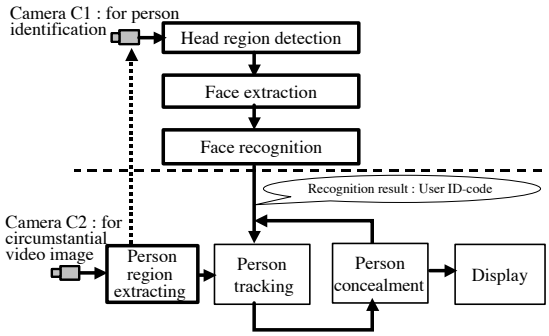


Fig. 3 Processing flowchart.

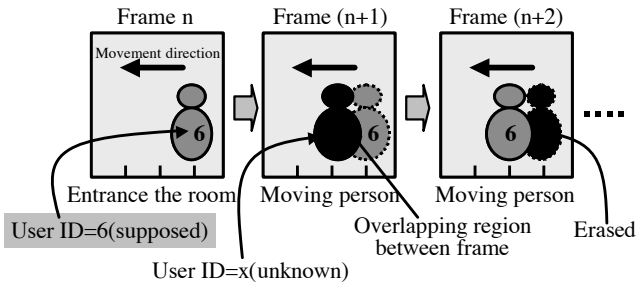


Fig. 4 Label propagation in person tracking process.

3.3 Processing flowchart

The process flowchart is shown in Fig.3. A person region is extracted from the image acquired from camera C1 for the person identification by image differential technique with the background image. Face region are extracted from the person head region at the top of the body. The result of face recognition is output as a user ID-code.

Then person tracking is carried out with holding a label corresponding to the user ID-code when a person moves around. Then the person in circumstantial video image taken by camera C2 is concealed and displayed.

3.4 Face Recognition

Several techniques were proposed for the person identification by face recognition[7]-[9]. The recognition technique using EigenFaces[7] requires much calculation time. In contrast, the face recognition technique using the mosaic image[8]-[9] requires shorter time, and moreover, the process is much simpler. For the real time processing requirement, the face recognition technique using the mosaic image is adopted.

As shown in Fig.1, the entrance route is modified to lead a person to proceed toward the camera C1, so that the person's front face image is acquired.

A face recognition program is implemented on the PC (CPU: Pentium3 600MHz, OS: Linux) in the C++ language. The resolution of captured image is 320?240 pixel. The system takes about 100msec from an image is taken to the face recognition process is completed or 10 frames/second.

Recognition rate was evaluated by 18 persons. Only half of them were registered into the face dictionary as a member of the group. The others were not registered, supposed to be strangers.

Each person entered the room five times. For one entry, during he/she is walking into the room through the path that was shown in Fig.1, it takes 2-3 seconds in average. It means the system can capture face image and carry out the face recognition process about 20-30 times for each entry. The recognition ID is given to the input face only when the sequence of the same person was consecutively observed more than 7 times in one entry, else it is decided that entering person is a stranger. The recognition rate for registered members was about 93%, and for a stranger the reject rate was about 97%.

3.5 Person tracking

In Fig.4, a person is assumed to walk from the right to the left in the scene. First, at the frame n, user ID-code is assigned to the person region as the recognition result. At the frame (n+1), by detecting the overlapping portion of the person region in the previous and the present frame, the label value to the person region is propagated into the present frame. For the following frames, similar propagation process will be applied, thus the system can track the person by assigning the same user ID-code to the person's image moving in the scene.

This process can be carried out in high speed by using the parallel processing capability of IMAP-VISION[6]. IMAP-VISION consists of 256 processors. The resolution of an image in IMAP-VISION is 256×240 pixel.

The average process time of the person tracking and the concealment display was 6.3 msec. For a typical practical case the distant from camera to the person is assumed to be 3.5 m, therefore 1 pixel is equivalent to the length of 0.02m at the object. As, the width of a person is about 20 pixels (=0.4m), if the movement of the person is less than 0.4m between frames(time interval =1/30 sec), the system can track that person. In the other words, this system can track the moving person slower than 12 m per second.

To conceal the person region, one of displaying method shown in Fig.2 is chosen in advance. The displayed images are shown in Fig.5.

At present, by this system it is easy to select the person concealment type of (b) - (e), and not to conceal a person unregistered beforehand such as a strange man.

4. EXPERIMENT AND EVALUATION

A subjective evaluation was carried out with 30 students(25 Japanese, 5 foreigners) in order to know how people prefer each modified video image either from observer or subject viewpoint.

At the observer viewpoint, the raw circumstantial video image is favorable, and the silhouette display(SD) is most unfavorable.

Conversely, from the subject viewpoint, the display of raw circumstantial video image is not preferred. It turned out that the silhouette display(SD) and the silhouette with name list display(SL) obtained the result with less mental resistance of subject while obtaining higher score from observers.

a) Original video image

b) Silhouette display (SD)

c) Silhouette display with name (SN)

d) Silhouette display with name list (SL)

e) Invisible person display (TN)

Fig. 5 Displays of various person concealment.

5. CONCLUSIONS

This paper proposed and realized a system, which convey sufficient situation information in circumstantial video image while protecting specific person's privacy information in the scene.

This system can automatically identifies a person by face recognition, tracks him/her, and displays the image of the person-part in modified form in order to conceal it for privacy protecting.

At present, this system can provide different images for different purposes. For example it can provide full detail to security guards watching for strangers, but only silhouettes, perhaps with names, to people who use the images for other purposes such as communications.

As a result of an experiment, in several applications considered, the silhouette with name list display seems to be most appropriate as the observer-subject compromise for the balance of protecting privacy and conveying situation information in circumstantial video image.

6. ACKNOWLEDGEMENTS

The authors thank Mr. Masahiro CHIBA who developed the face recognition software. They also thank all lab members and others students for their cooperation to the experiments.

7. REFERENCES

[1] Robert Fish, et al.: “The VideoWindow System in Informal Communication”, ACM Conf. Computer-Supported Cooperative Work(CSCW'90),ACM,pp.1-11(1990)

[2] A. Borning and M. Travers: “Two Approaches to Casual Interaction over Computer and Video Networks”, Proc. ACM Conf. Human Factors in Computing Systems CHI'91, New Orleans, pp.13-19(1991)

[3] Paul Dourish: “Exentding Aweaeness Beyound Sychronous Collaboration”, Workshop on Awareness in Collaborative Systems CHI’97 (1997)

[4] Michael Boyle, Christopher Edwards and Saul Greenberg: “The Effects of Filtered Video on Awareness and Privacy”, Proceedings of the CSCW'00, ACM (2000)

[5] Robert Root: “Design of a Multi-media Vehicle for Social Browsing”, CSCW-88, pp.25-38(1988)

[6] Sholin Kyo, et al. : “A Programming Environment for a Linear Processor Array Real-time Image Processing System”, IPSJ Trans., 39,6,pp.1790-1800(1998)

[7] M. Turk and A. Pentland: “Eigenfaces for recognition”, J. Cognitive Neurosci., 3, 1, pp.71-86(1991)

[8] Makoto KOSUGI, et al. : “Human-Face Recognition Using Mosaic Pattern and Neural Networks”, IEICE Trans., J76-D-II, 6, pp.1132-1139(1993)

[9] Makoto KOSUGI, et al. : “Human-Face Search and Location in a Scene by Multi-Pyramid Architecture for Personal Identification”, IEICE Trans., J77-D-II, 4, pp.672-681(1994)