A Technique for Displaying Presence Information on a Live Camera Image Using 3-D Mask Objects

Jun Higuchi, Shin Takahashi, and Jiro Tanaka

Department of Computer Science
Graduate School of Systems and Information Engineering
University of Tsukuba
1-1-1 Ten-noudai, Tsukuba, Ibaraki 305-8573, Japan
{higuchi,shin,jiro}@iplab.cs.tsukuba.ac.jp

Abstract. Installing a live camera in an office or laboratory can provide reliable high-quality communication among collaborators who are physically separated, but may also lead to privacy concerns. We propose a technique for displaying presence information using a three-dimensional object as a mask that covers the image of people in the live camera system. Our technique enables information to be shared partially, enabling sufficient communication while alleviating the discomfort of feeling monitored by the camera. The system also clarifies less-visible activities such as typing text using a keyboard or moving a mouse of people in the camera image. We implemented the *Koka-ComeCam* system using our proposed technique.

Keywords: Live camera, Communication support, Privacy, 3-D, Visualization.

1 Introduction

Using a live camera to support communication among physically separated collaborators has become common [6, 8-10]. Live video provides more reliable and higher-quality communication to group members compared to text-based tools such as Instant Messenger [8]. For example, a live camera in an office or laboratory indicates at a glance who is there and what they are doing. Unlike video chat and video conferencing, which use Web cameras, a live camera provides continuous coverage of the daily routine in the room. Moreover, the image can be easily monitored from anywhere via the Internet. In addition, using a camera with a pan/tilt/zoom (PTZ) function enables video coverage of a wide area such as an entire room.

However, some people may feel uncomfortable about being recorded by a live camera. One common technique to alleviate such discomfort is to mask the video image (e.g., using a mosaic or blur), but this reduces the information provided by the camera and thus questions the merit of installing a live camera.

We propose using three-dimensional (3-D) virtual objects in the video image to partially mask images and therefore support live camera communication as well as privacy. The technique basically overlays an avatar on the images of people to mask

their direct appearance but identify them individually. The level of detail of the mask can be changed according to the level of privacy desired, and this setting is controlled by the person represented by the avatar. This enables one to balance what information is shared with privacy. Moreover, the mask represents each person's current and past activities, thereby presenting information that would otherwise be hidden in traditional masking systems.

The rest of this paper is organized as follows. In Section 2, we summarize our previous work and introduce the basic 3-D mask object. In Section 3, we describe the *Koka-ComeCam* system as it relates to our proposed technique. In Section 4, we explain the architecture and implementation details of the system. Related work is provided in Section 5, and the conclusion is given in Section 6.

2   Previous Work: *ComeCam*

We previously developed the *ComeCam* system [5], which supports communication between two ends of a live camera connection. The user[1] can display a mask in the video output that covers his or her image. The mask is a 3-D virtual object, which we call a *3-D mask object*. Unlike a mosaic or blur, a 3-D mask object functions as an avatar chosen by the user as shown in Figure 1. A 3-D mask object covers the user, but enables others to identify him or her based on its unique character and displayed position. Because the system manages each user's position as a 3-D coordinate relative to the size of the room, the mask can cover the user's area even if the view of the camera is changed using PTZ operations. We propose a new technique for displaying users' presence information based on this 3-D mask object, and describe the *Koka-ComeCam*[2] system implemented with the new technique.

Panned, tilted, and zoomed images
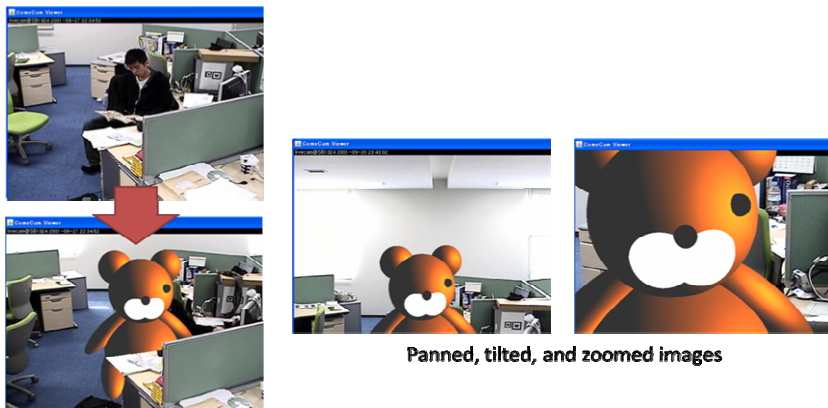
Fig. 1. An example of a 3-D mask object displayed on a live camera image

---

[1] This indicates the observed person in this paper.
[2] "Koka" means "effect" in Japanese.

3   Displaying Presence Information with a 3-D Mask Object

A 3-D mask object in *ComeCam* works as an avatar and is displayed according to the camera movement via a PTZ operation, but it is basically a mask for the privacy problem. Compared to a mosaic or blur, overlaying a virtual 3-D object on the real camera image enables user presence information to be represented more naturally and more effectively. In this section, we introduce our new technique and use it to implement the *Koka-ComeCam* system.

### 3.1   Changing the Mask According to User Privacy Level

Masking the video image blocks some information from remote users. Because it is preferable to give users control over masking, our system links masking to privacy settings. The user selects a privacy level of [high], [middle], or [low], as shown in Figure 2(left), or [none] for no mask. As the privacy level is lowered, the mask's appearance becomes simpler, and finally becomes a simple cube. The size of the mask area also becomes smaller. At the same time, the transparency of the mask increases, so that the user's video image can be gradually seen through the mask (Figure 2(right), a snapshot of *Koka-ComeCam*). Initially, the mask at the [middle] level of privacy is automatically chosen by the system (just after starting the application on the user's own PC).

Because the mask plays the role of avatar, even if the privacy level is [high], the minimum presence information, such as whether the user is available, is always maintained in the output video image. In contrast, the [low] and [middle] levels of privacy lessen the masking of the object so that it does not perfectly cover the user, making the user's actions and state more visible. However, it still provides a sense of security to the observed person because it obscures the user's direct appearance.

Users can change the level of privacy to suit their situation. For example, if a user wants to sleep for a few hours, and determines that it would not be a good idea to broadcast this to colleagues or the boss, the user could temporarily raise the privacy level. If, however, a user wants to be seen working, he or she could lower the privacy level. However, in some cases, changing the privacy level manually might be difficult. For example, if a person decides to take control of the camera and zoom in on a specific user, the system will automatically increase (or decrease) the privacy level; each user sets the conditions of the zoom function and other specific operations.

Fig. 2. Changing a mask according to user privacy level
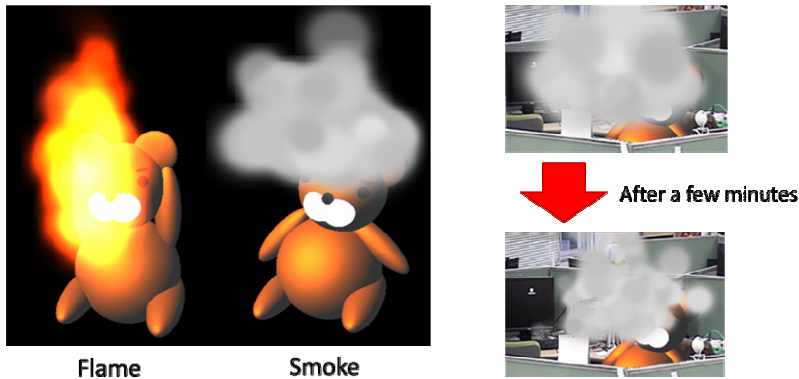
Fig. 3. Special effects represent user activities

From the standpoint of sharing information between research groups or laboratories, it is preferable to set the system at low privacy/high sharing and only occasionally switch to high privacy/low sharing. However, because users can control their own privacy settings, the potential exists for users to always use the high privacy/low sharing setting, thereby undermining the advantage of installing a live camera. To address this potential problem, the system assumes that live cameras are installed in two or more rooms and that users monitor the appearance of each others' rooms interactively. Thus, other room members are covered by the high-privacy mask as long as the viewing user does not decrease the privacy level. This maintains equality between groups.

3.2   Representing User Activities with a 3-D Mask Object

The technique of showing the mask described in Section 3.1 allows observers to access user information without compromising privacy, by masking and not completely blocking images of people. A live camera by itself can only display the current situation or activity. In addition, it remains unclear whether users are working or sleeping if they are facing a computer and not moving. To address this issue, we introduced graphical effects used in cartoon animations and computer games to provide additional information about users. Other systems have applied such effects to represent abstract information such as a person's feelings or atmosphere. We built on this to depict current and past activities of users graphically by overlaying additional information on the 3-D mask object. Figure 3(left) shows two examples. First, when the user starts working at his or her own computer, such as typing text using a keyboard or moving the mouse, a flame graphic is displayed around the mask. The flame grows or shrinks according to the level of keyboard or mouse operation. Second, when the user stops operating the computer and becomes inactive, a smoke graphic is displayed around the mask. The amount of smoke changes according to the amount of time spent inactive. In this way, the user's past activities can be inferred (Figure 3(right)). For example, smoke is displayed for some time after a user leaves the room (after the user logs out from the system). Thus, even if the user is no longer

Fig. 4. An example of a live camera image of Koka-ComeCam

displayed on the camera image, it can be inferred that the person was there working relatively recently. Because these effects are also represented as 3-D virtual objects, they can robustly follow any changes in the camera view made by a PTZ operation.

In addition, users can leave messages such as "I'll be back soon" near their mask. The message is displayed around the mask as a balloon. It is also possible to automatically show or hide messages according to the focus and the zoom level of the camera. For example, if a remote user zooms in on someone's image in the camera, the imaged user can set the system to display the message "Don't bother me! I'm busy now!"

A 3-D mask object masks users, but by overlaying additional information, our system enables the sharing of additional important information, such as activity status. Because these effects are displayed around each user's mask, user activities and status can be easily understood at a glance. In fact, when an effect is displayed, the mask moves as shown in Figure 3(left) to make it even easier to map between the mask and effect. Figure 4 shows an example of a live camera image using *Koka-ComeCam*. By simply viewing the overview of the room, observers can instantly know what people in the room are currently doing, thereby facilitating communication among group members.

Some users may be uncomfortable revealing this extra information to others, but these effects represent abstract information only, such as whether the keyboard and mouse are currently being operated, and thus our system strikes a better balance between information sharing and privacy than other existing systems.
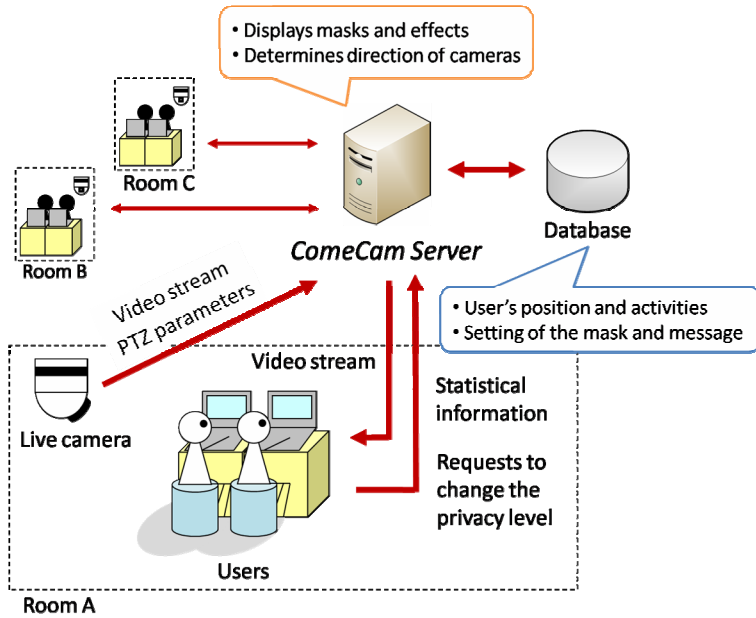
Fig. 5. System Architecture

4   Implementation

4.1   System Architecture

Figure 5 shows the architecture of *Koka-ComeCam*. The system consists of networked PTZ cameras, client PCs in each room, a database server that maintains information such as user positions and activities, and the central *ComeCam Server*, which connects all of the PCs. Each client PC runs the application with the interface to allow control over the mask and messages. Each client PC sends user statistics (e.g., keyboard activity) and requests (e.g., privacy level settings) to *ComeCam Server*. The server receives the video stream from each live camera, overlays masks and effects at each user's position according to the user's privacy level and activities, and then transmits the resulting animated video image to each user's PC. In this way, no raw images are transmitted directly to users.

Our system uses AXIS 214 PTZ[3] as a live camera. The parameters for this camera can be queried or set simply by referring to specific URLs. Each time the PTZ values of the camera change, *ComeCam Server* acquires the pan and tilt angle values (-180 - +180) to determine if the camera is pointing at the user. If the system decides that the camera is pointing at the user, it refers to his or her settings of the mask and messages from the database and overlays the mask according to this information. The paper [5] supplies more details about the algorithm that determines whether the camera is pointing at the user.

[3] AXIS 214 PTZ http://www.axiscom.co.jp/prod/214/

4.2 Displaying 3-D Mask Objects on a Live Camera Image

When the camera is moved using the PTZ operations, the two-dimensional area of the image to be masked changes considerably, and so it must be calculated in three dimensions. *Koka-ComeCam* uses 3-D graphical objects as masks, and places them virtually at the position of the user to cover them even if the camera is moved. For zooming, we match the horizontal angle of view in the perspective projection of the virtual camera and the real camera. The current angle of view of the real camera can be calculated by the zoom level (0 - 9999), which is acquired from the real camera. For panning and tilting, the current referenced coordinates of the real camera can be acquired by the pan and tilt angle value, so we match them to the referenced coordinates of the virtual camera. *Koka-ComeCam* acquires the PTZ parameters of the camera and redraws objects whenever the PTZ operation is executed. As a result, it synchronizes the display of objects with the movements in the real camera. We use OpenGL library to draw masks and effects.

To render 3-D mask objects and effects as if they exist in the real world, we display them after clipping according to the shape of real objects. To input the shape of real objects in the room, we put ARToolKit[4] markers on them (e.g., at the four corners of a desk) as shown in Figure 6(a) and capture their coordinates with the camera. ARToolKit is a library that determines the relative position of the marker to the camera in 3-D space. We create the 3-D virtual object of the real object (Figure 6(b)), clip the overlapping part of the 3-D virtual object (Figure 6(c)), and overlay it on the camera image (Figure 6(d)). In this way, the mask can be naturally integrated to the live camera image.

4.3 Setting up User Positions Using Markers

Those with the authority to use *Koka-ComeCam* are members of the workspace (e.g., office or laboratory) in which the system is being operated. Currently, users' positions are
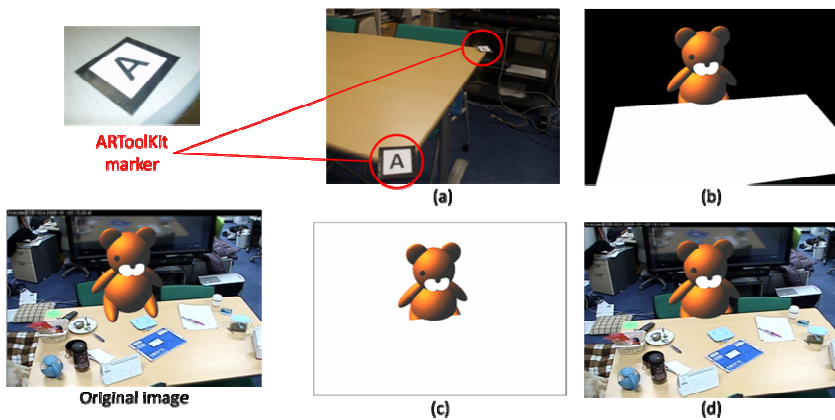
Fig. 6. Displaying the mask after clipping according to the shape of real objects

---

[4] ARToolKit http://www.hitl.washington.edu/artoolkit/

fixed based on their desk position, where most of their work is centered. The position of the user and the camera are handled as 3-D coordinates relative to the size of the room. Therefore, it is necessary to measure the absolute positions of the user and the camera at least once, but using ARToolKit and its markers can reduce the time to input these values. We capture the marker assigned to each user's position, convert it into coordinates relative to the size of the room, and store it in the database. In this way, we can update the user's position easily even if it changes by moving the position of the desk.

5   Related Work

Many systems use video to help provide the sense that distantly located collaborators are in the same room [2, 7, 4]. In particular, GestureCam [3] uses a moving camera mounted on an actuator to achieve high-quality results. These techniques aim to synchronize and improve collaborative work by sharing images for one-on-one communication, and thus require special devices. Our technique requires only a live camera to provide continuous coverage of the daily routine in the entire room.

Live video communication systems must balance information sharing and privacy. Hudson et al. [8] proposed the Shadow-View technique to address privacy, while Boyle et al. [6] used a mask-like mosaic and blur. Tansuriyavong et al. [9] used silhouettes to represent people, and other research [10] made images of people transparent to a certain degree depending on distance to the camera. Many techniques, however, do not support dynamic changes of the camera view via PTZ operations. Our technique uses a 3-D virtual object as a mask, and shares other information such as activity status using special effects.

A previous study [1] extended Instant Messenger so that users can recognize presence information visually from a 3-D display. The advantage of our technique is that because our display is based on a real video image, it can represent not only awareness information of remote users but also environmental information of remote places, and all the input unites naturally on the display.

6   Conclusion

We proposed a technique for displaying presence information using a 3-D object as a mask that covers the people in a live camera system. We applied the method to *ComeCam* and implemented *Koka-ComeCam* using the system. Our system provides much more awareness information of remote users while also protecting privacy. In future work, we will examine a new technique for representing user information, and discuss the balance between it and privacy.

References

1. Neustaedter, C., Greenberg, S.: Supporting Coherence with a 3D Instant Messenger Visualization. In: CHI 2002 Workshop on Discourse Architectures (2002)
2. Ishii, H., Kobayashi, M., Grudin, J.: Integration of Interpersonal Space and Shared Workspace: ClearBoard Design and Experiments. ACM Transactions on Information Systems 11(4), 349–375 (1993)

3. Kuzuoka, H., Kosuge, T., Tanaka, M.: GestureCam: a video communication system for sympathetic remote collaboration. In: 1994 ACM conference on Computer-Supported Cooperative Work, pp. 35–43 (1994)
4. Kuzuoka, H., Yamashita, J., Yamazaki, K., Yamazaki, A.: Agora: A Remote Collaboration System that Enables Mutual Monitoring. In: CHI 1999 extended abstracts on Human factors in computing systems, pp. 190–191 (1999)
5. Higuchi, J., Takahashi, S., Tanaka, J.: ComeCam: A Communication Support System Between Both Ends of The Live Camera Connection. In: Third International Conference on Collaboration Technologies, pp. 33–38 (2007)
6. Boyle, M., Edwards, C., Greenberg, S.: The Effects of Filtered Video on Awareness and Privacy. In: 2000 ACM Conference on Computer-Supported Cooperative Work, pp. 1–10 (2000)
7. Morikawa, O., Maesako, T.: HyperMirror: Toward Pleasant-to-use Video-Mediated Communication System. In: 1998 ACM Conference on Computer-Supported Cooperative Work, pp. 149–158 (1998)
8. Hudson, S.E., Smith, I.: Techniques for addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support System. In: 1996 ACM Conference on Computer-Supported Cooperative Work, pp. 248–257 (1996)
9. Tansuriyavong, S., Hanaki, S.: Privacy protection by concealing persons in circumstantial video image. In: Workshops on Perceptual/Perceptive User Interfaces 2001, pp. 1–4 (2001)
10. Tansuriyavong, S., Susuki, T., Chokchaitam, S., Iwahashi, M.: Privacy Conscious Video Communication System Based on JPEG 2000. In: International Workshop on Antenna Technology 2007, pp. 50–55 (2007)