Video Memo: Asynchronous Communication through a Video Message on a Table

Shun Nakaigawa
Department of Computer Science, University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki, Japan
nakaigawa@gmail.com

Shin Takahashi
Faculty of Engineering, Information and Systems
1-1-1 Tennodai, Tsukuba, Ibaraki, Japan
shin@cs.tsukuba.ac.jp

Jiro Tanaka
Faculty of Engineering, Information and Systems
1-1-1 Tennodai, Tsukuba, Ibaraki, Japan
jiro@cs.tsukuba.ac.jp

ABSTRACT
This paper presents a hand gesture-based tabletop system that enables asynchronous video communication through a "video memo." A video memo is a video message recorded on a table, such as a note about a book on the table. Using our prototype system, it is possible to record a video on a table with only hand gestures. The recorded messages can be put as an icon on a table in the current room or in a remote room. Then, the other user can view them with only hand gestures. This paper describes the interaction techniques and the implementation of our prototype system.

Author Keywords
Interactive tabletops; video communication; asynchronous communication; ubiquitous environment; interaction design.

ACM Classification Keywords
H5.2 (Information interfaces and presentation): User Interfaces – Interaction styles.

General terms: Design, Human Factors.

INTRODUCTION
Video is a powerful medium that can transmit a large quantity of real-world information, such as object forms, gestures, and sounds, to a person who is not currently available or to a person at a distant location. There are many systems that use the power of video. For example, videophone systems, such as Skype, support face-to-face real-time communication where the users communicate to each other through the screen of a PC or mobile/smart phones. ClearBoard and other systems [1,2,3] can transmit not only faces but also real-world information, such as writing on the desk, to each other. However, they basically handle real-time communications.

Our ultimate goal is to design an asynchronous video messaging service that can be used anywhere in the world. The service should enable the user to make a video anywhere easily, without wearing any device. The user should also be able to specify the receiver of the video message readily, which can then be viewed later by the specified user.

APCHI '12, August 28–31, 2012, Matsue-city, Shimane, Japan.

As a step towards this goal, this paper proposes the use of 'memo' as a metaphor for asynchronous video messages in the real world. That is, we represent a recorded video message as a memo that can be attached to a physical object, such as a book on a table. By attaching a memo to a physical object, people can readily accept that it is related to the attached object. For example, simply, by putting a memo on someone's table, it is clear that the owner of the table is the intended recipient.

In this paper, we describe our prototype system, designed and implemented according to the above metaphor. We call a recorded video message a "video memo." In our prototype system, the user can readily make a video memo on the table with hand gestures. He/she can record his/her voice and actions on the table. After the recording, the memo is 'put' somewhere on the table; that is, a video memo icon is displayed on the table by the projector. The user can also put a video memo on a remote table. The displayed video memo icon can be played back easily by simply pointing at it with a finger.

VIDEO MEMO SYSTEM

Video Memo
A video memo is a video message that is placed on a table like a paper note. We implemented the Video Memo System as a tabletop interface to record and replay video memos. A camera installed above the desk is used to record the scene and sounds around the desk. The angle of the camera is similar to the view of video maker (see Fig. 1) so that the user can record and share with the video viewer much real-world information, such as the forms of real objects, figures, text on paper, and pointing gestures. The user can start recording instantly just by slapping the table, enabling the user to quickly memorialize his idea.

The video memos created on the desk are shown as icons (Fig. 2). The user can run and watch the video memo by pointing at it with his/her finger. Video memos that have

not yet been played will be shown as bigger icons so that the user will notice them.

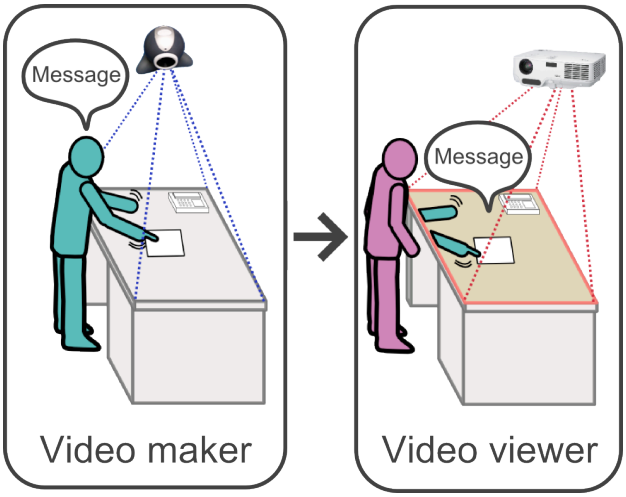

Figure 1. Concept image of the Video Memo System. The camera recognizes the video maker's hand gestures and records a video message. Additionally, the projector shows a video and an icon recorded on the desk for a video viewer.

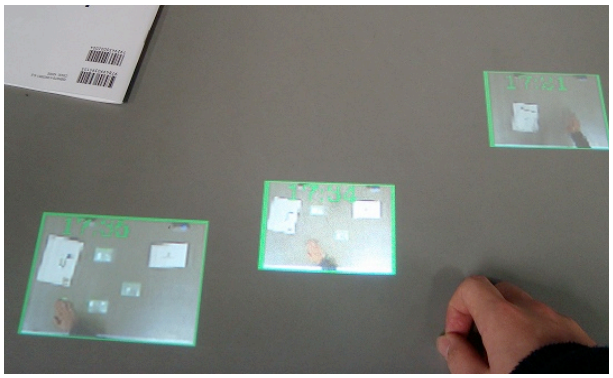

Figure 2. The Video Memo icon projected on the desk.

Pointing and selecting icons

A basic operation in this system is the selection operation, by pointing at an icon for one second. The pointing position is displayed as a round cursor. When the user selects an icon, the system provides visual feedback by changing the color and auditory feedback with sound effects.

Movement and deletion of a video memo

When the user points at a video memo icon, Delete and Move icons appear at its side. As the user selects the Delete icon, the video memo is deleted. As the user selects the Move icon, the video memo starts to move, following the pointing position. When the user stays still for one second, the video memo is put in the new position.

Creating video memo on the local desk

To create a video memo, the user first records a video using the installed camera. Then, he/she puts its icon on the desk. Both of these operations are achieved using hand gestures. We assigned the slapping gesture to these operations. That is, at first, the user starts recording the video by slapping the table. Then, he/she stops and puts its icon on the table also by slapping the table. The icon is placed wherever the user slaps. We used the slapping gesture because it resembles the action of putting notes on the table.

Besides putting a video memo on the table, the user can also scribble on the table with hand gestures (see Fig. 3). The video maker starts scribbling by making a circle with his/her thumb and forefinger (like holding a pen). Then, the trace of the user's hand is drawn on the table. Scribbling is stopped by breaking the circle. Scribbling can be used for drawing a figure or writing a short text note around the video memos on the table.

The steps for making a video memo follow:

1. The user slaps a hand on the desk to inform the system to start recording. The system will then start to record all of the actions and sounds that occur over the whole area of the desk.
2. The user makes a video message with sounds and actions using objects on the desk. For example, the user may point to a book with his/her finger and ask the viewer to bring it back to the library, or show a paper and recommend that it is worth reading.
3. The user finishes recording the video memo by slapping the desk again. The video memo is placed at the position where he/she slaps his/her hand.

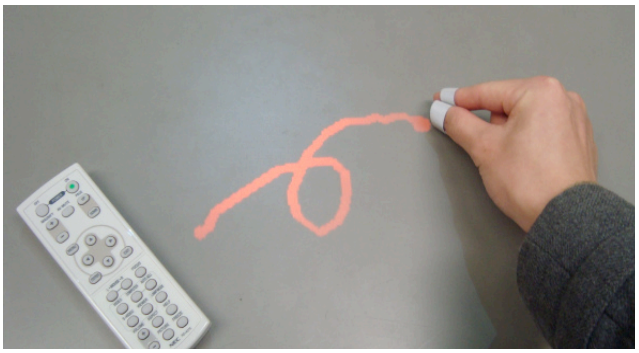

Figure 3. Scribbling gesture.

Putting a video memo on a desk in A remote room

In our system, it is also possible to place a video memo remotely, which is achieved in two steps: selecting a remote desk and making a video on the desk.

First, the video maker selects a remote desk through the image of the remote room. It is captured by the live camera

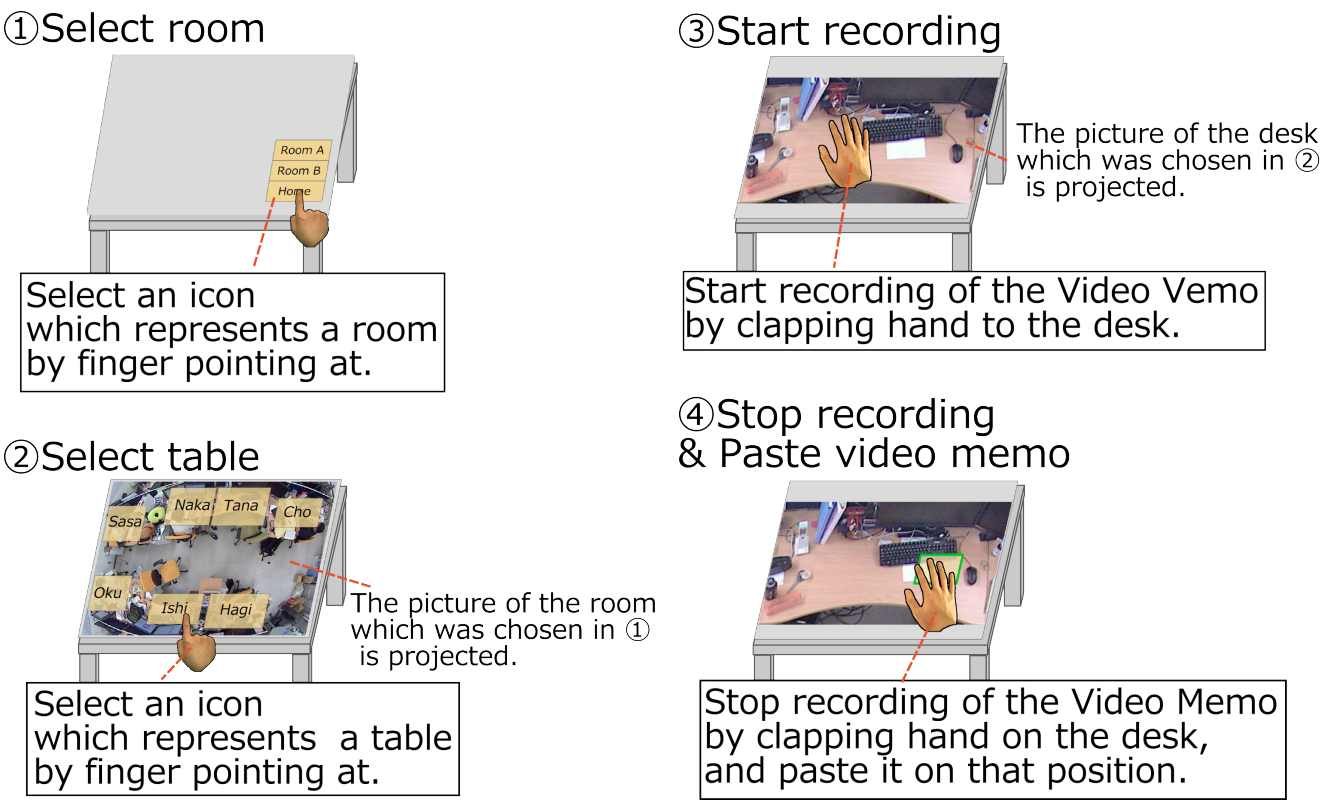

Figure 4. Recording procedure of a video memo to the desk in a remote living room.

installed at the ceiling of the room so that the whole room can be viewed from the camera. In our system, the destination of a video memo is not a person, but a location on the desk. The video maker selects a translucent rectangle label displayed over each desk in the room. We intend to partially cover the view of the desk with the label but the user can still see whether a colleague is sitting there.

When the user has selected the destination desk, the live camera image of the desk is projected on his/her local desk. Then he/she makes a video memo in the same way as when he/she makes and places it locally. The user can refer to objects on the remote desk in the message through their images.

Figure 4 shows the five steps for making a remote video memo:

0. The video maker requests showing the menu of the selectable remote rooms by pointing at the right edge of the local desk.
1. The video maker points to one of the menu items for one second to select the destination room and the live camera image of the room is projected onto the local desk.
2. The video maker then selects the destination desk by pointing a label displayed on each desk. Then, the system projects the live image of the selected remote desk onto the local desk.
3. The system starts recording when the video maker slaps the local desk.
4. Slapping the desk again stops the recording of the video memo, which is then put on the destination desk.

Viewing a Video Memo

A video memo is replayed by pointing at its icon displayed on the desk. When selected, it expands to the size of the desk and is replayed. Figure 5 shows a video memo replayed on the desk. Because the size of the replayed image is the same as that of the recorded area, the position and size of the objects are maintained in the projected replayed video image.

IMPLEMENTATION

Figure 6 shows the hardware setup of our system. For each desk, we installed a camera and a microphone for recording, and a speaker and a projector for replaying. The computers at each desk communicate with each other through the network. A wide-view live camera is also installed at the center of the ceiling to provide an image of each room. The software is implemented using C++ and OpenCV library.

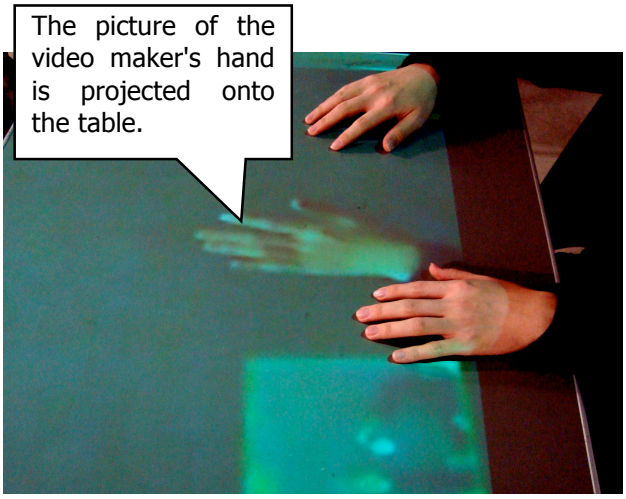

Figure 5. Video memo that is replayed.

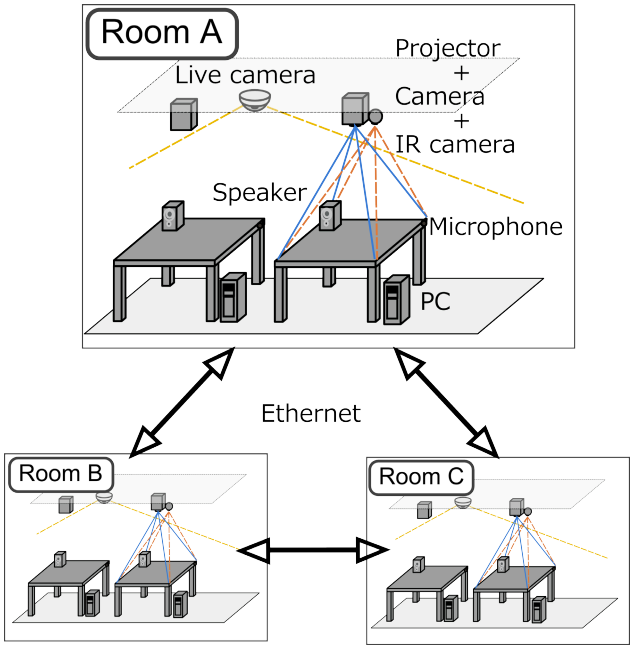

Figure 6. Hardware setup.

Finger recognition

To recognize the position of fingertips on the desk where the video image is projected, we use a near infrared (IR) camera and retro-reflective material tape, which is a material that strongly reflects light back to its source. The tape put on the fingertip strongly reflects the IR light from the IR LED around the IR camera, which enables robust recognition of fingertips on the projected desk. The system binarizes IR camera images with a threshold, depending on the environment, and regards the center of detected blobs as the positions of the fingertips.

Slapping recognition

We use a contact-type microphone to acquire solid vibrations. A microphone is installed on the desk to detect the slapping sounds of video maker. Sounds louder than the threshold and shorter than the threshold are recognized as candidate slapping gestures. Fingertips must also be recognized at the moment of a slapping gesture.

Scribbling recognition

To recognize a scribbling gesture, this system must determine whether the blobs of the fingertips contact each other. The video maker must wear tape on the thumb and forefinger to scribble. Figure 7(a) shows the camera image when two fingers are apart. When the user starts a scribbling gesture, the blobs will merge, as in Figure 7(b). The system recognizes this change as a scribbling gesture and starts drawing a line of the trace of the center of the gravity of the merged blobs. The system detects when the scribbling gesture ends, and it stops drawing.

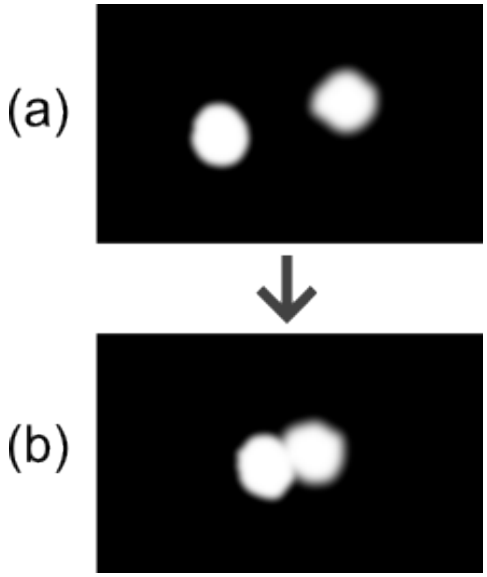

Figure 7. IR camera image when the user brought two fingertips close together. The white blobs are retro-reflective material tape worn on the fingertips.

Camera calibration

Our system uses homography transform to calibrate the camera. It is necessary to clip a remote or replayed camera image projected onto the desk to record and replay a life-size video with a normal camera. It is also important for

correct visual feedback for pointing gestures. Thus, we used homography transform. First, the user points to the four corners of the projected area with a mouse from the camera image when setting up the system. The system generates a homography matrix from these four coordinates and transforms each frame from the camera with the matrix.

RELATED WORK

Notification Collage is a multimedia asynchronous communication system [4]. Live camera images can be put on a shared visual bulletin board. In our system, a video is put on a real-world desk, which makes accessing it easier. Our system also can use video with real-world information.

t-Room is a video collaboration system that uses shared video screens [5]. This system encloses a user space with surrounding large LCD displays showing life-sized images. This system can communicate using video displayed on these displays to a remote system synchronously and asynchronously. In our system, we use a desk-type interface and gesture interaction using the metaphor of the memo. Hence, our system is an intuitive technique for exchanging messages in a ubiquitous environment.

CONCLUSIONS AND FUTURE WORK

Our video memo system records sounds and actions on the desk, and puts the memo on the local desk and also on a remote desk. It enables asynchronous communication between people in distant rooms. Recording and replaying the video memo can be done with only hand gestures, making it possible for the user to access them quickly.

In future work, we plan to implement the system in a smart room, so that a video memo can be placed anywhere in the room.

REFERENCES

1. Ishii, H., Kobayashi, M. ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact. Proc. of CHI'92, pp.525-532, 1992.
2. Ichikawa, Y., Okada, K., et al. MAJIC Videoconferencing System: Experiments, Proc. ECSCW'95, pp.279-292, 1995.
3. Kuzuoka, H., Yamashita, J., Yamazaki, K., and Yamazaki, A. Agora: a remote collaboration system that enables mutual monitoring. In CHI '99 extended abstracts on Human factors in computing systems (CHI EA '99), pp.190-191, 1999.
4. Rounding, M., Greenberg, S.: The Notification Collage: Posting Information to Public and Personal Displays, Proc. SIGCHI'01, pp.514-521, 2001.
5. Hirata, K., Kaji, K., Harada, Y., Yamashita, N., Aoyagi, S. t-Room: Remote Collaboration Apparatus Enhancing Spatio-Temporal Experiences, Adjunct Electronic Proc. CSCW'08, 2008.