Doree Duncan Seligmann
Avaya Labs

Proximity as an Interface for Video Communication

Nicolas Roussel, Helen Evans, and Heiko Hansen
LRI & INRIA Futurs

When you're far away, your own image reflects in it like in a distorting mirror, blurred and imprecise. As you move toward it, however, it becomes clearer and more accurate. By the time you reach it, the reflection is almost perfect, as in a conventional mirror. What you see is not a simple optical reflection but a video image captured, processed, and displayed in real time.

What appears to be a mirror is in fact an interactive video communication system—MirrorSpace—connected to similar devices in other places. As someone passes in front of one of the other devices, a vague silhouette appears on the one in front of you, mixed with your own image (see Figure 1). As he approaches, his image becomes sharper, allowing you to recognize him. When he reaches the device, his face gets fully merged with yours, letting you look into each other's eyes.

Evolution of the project

MirrorSpace was originally conceived as a prototype for the interLiving project (http://interliving.kth.se/) of the European Disappearing Computer initiative (2001–2003). This project focused on the design of technologies to support communication among family members located in different households. For three years, three Swedish and three French families collaborated with a multidisciplinary research team with expertise in computer science, social science, and design.

One of the technologies we investigated in this project was the exchange of still images or video streams between households. Previous work on video, including our own, has shown that it's well suited for coordination and informal communication.[1] However, traditional video technologies designed for work settings don't necessarily fit in home settings where space can be tight and serve multiple purposes, and where the relationships between family members can be complex. Our work on MirrorSpace started as we got interested in the use of space and distance in video-mediated communication.

(Un)use of space in video-mediated communication

One of the advantages of video over audio or text-based systems is the ability to transmit nonverbal information. However, while many studies have focused on eye gaze and gesture in video-mediated communication, little work has been carried out on proxemics,[2] one of the most fundamental elements of nonverbal communication. See the "Proxemics" sidebar for an overview.

Physical proximity to other people is a form of communication that we all use, although we're barely aware of it. Space and distance let us define and negotiate the interface between private and public, particularly during the moments leading up to contact. By altering our physical distance from other people in a space, we communicate subtle messages—such as our willingness to engage into dialogue with them, the desire for more intimacy, or a lack of interest. For each communication situation, we have a distance that we find appropriate. Certain feelings

Figure 1. Images from an early MirrorSpace concept video (August 2002).

1070-986X/04/$20.00 © 2004 IEEE

Published by the IEEE Computer Society

Proxemics

The term *proxemics* refers to the study of spatial distances between individuals in different cultures and situations. It was coined by E.T. Hall in 1963 when he investigated man's appreciation and use of personal space. Hall's model[1] lists four distances that North Americans use in the structuring of personal dynamic space: intimate (less than 18 inches), personal (between 18 inches and 4 feet), social (between 4 and 12 feet), and public (more than 12 feet). For each communication situation, we have a distance within these four categories that we find appropriate. If the perceived distance is inappropriate, we become uncomfortable and usually adjust it by physically moving closer or farther away, or even simply turning our head or looking in another direction.

Reference

1. E.T. Hall, *The Hidden Dimension: Man's Use of Space in Public and Private*, Doubleday, 1966.

Daniel Rozin's Mirrors

Daniel Rozin has created several installations that use image processing techniques to turn a set of motorized, nonreflective surfaces into a virtual mirror. His *Wooden Mirror*, for example, is made of 830 square pieces of wood lit from above that can be tilted up and down individually, appearing lighter or darker depending on the angle. The whole array can thus display a rough reflection of whatever is in front of it (Figure A). *Trash Mirror* is a similar Rozin installation, made of 500 irregular pieces of trash collected on the streets of New York. Yet another piece, *Shiny Balls Mirror*, consists of 900 hollow metal tubes with polished chrome balls placed in them. Here, the brightness of each "pixel" is controlled by moving the ball in (darker) or out (brighter) of the tube. The display thus serves as a mirror in two ways: as a whole, but also as it reflects the viewer 900 times on the shiny balls.

Figure A. Daniel Rozin's Wooden Mirror.

or emotions, for example, are difficult to share unless the two partners are close.

Existing systems for video-mediated communication fail to take proxemics into account. Although some of the people who design the systems understand the importance of proxemics, they fail to give it much consideration or to provide the support it requires. Systems are usually designed for a specific task, corresponding to a certain interpersonal distance. Physical constraints often make it impossible for people to come closer to the device than expected or to move away from it.

MirrorSpace: Design concept

Our work on MirrorSpace focuses on creating a video communication system that takes physical proximity into account. We're particularly interested in how people's interactions can trigger smooth transitions between situations as extreme as general awareness of remote activity (where anonymity is preserved) to close and intimate forms of communication. As the name suggests, MirrorSpace relies on a mirror metaphor. This system's key characteristics include the original placement of the camera combined with translucently overlaying the images and using a proximity sensor combined with a blur filter.

As a cultural artifact, the mirror has a prominent position in the creation and expression of aesthetics. Throughout Western culture, in narratives such as the Narcissus myth, *Snow White*, or *Through the Looking Glass*, the mirror has come to symbolize many things—including vanity, deception, identity, or a passage to another world. Unsurprisingly, numerous artists and designers have picked up on these meanings and taken advantage of the universal and irresistible fascination for self-image to explore the boundaries between the analog and digital worlds. Examples of these works include Christian Möller's *Electronic Mirror* (http://www.christian-moeller.com), Scott Snibbe's *Screen Series* (http://www.snibbe.com), Camille Utterback's *Liquid Time* (http://www.camilleutterback.com), and Daniel Rozin's various mirrors (http://fargo.itp.tsoa.nyu.edu/~danny; also see our sidebar, "Daniel Rozin's Mirrors").

Sometimes we can also perceive a mirror as a surface for mediating communication with its own rules and protocols. As many subway commuters know, making eye contact with a stranger through a reflecting surface is usually considered less intrusive than direct eye contact. Because humans already associate the mirror with this idea of reaching out to other people and other spaces,

Figure 2. Looking into each other's eyes.

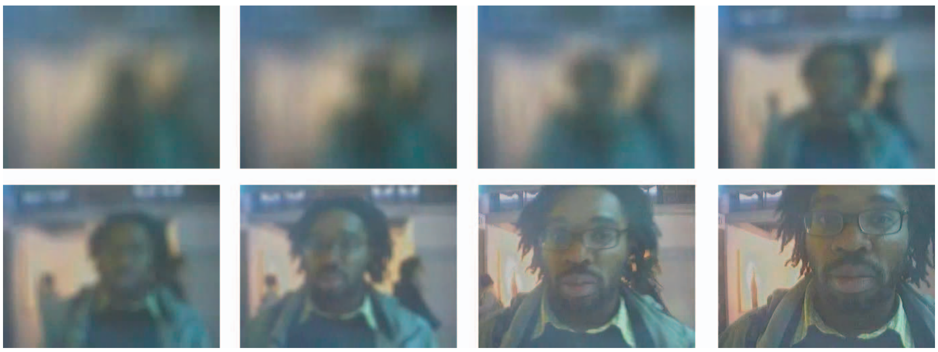

Figure 3. From peripheral awareness to close communication by moving toward the device.

we believe it's the ideal enabling metaphor for establishing a new communication experience.

Several video communication systems have recently used a mirror metaphor to provide seductive and pleasant-to-use interfaces. As demonstrated by Morikawa and Maesako,[3] this metaphor helps reduce the psychological distance between local and remote participants by displaying them side by side. However, because the camera is usually placed atop or beside the display, the remote people never seem fully engaged and appear to be looking slightly off, in another direction. To give the impression of looking into someone's eyes, the viewer has to look at the camera and thus can no longer see where the other person is looking.

MirrorSpace superimposes the live video streams from all the connected places on a single display on each site so that people see their own reflection combined with the ones of the remote persons. We felt it was important for people to actually look into each other's eyes and possibly merge their portraits into one, so we decided to place the camera on the screen, rather than beside it. This setup allows participants to come close to the camera while still being able to see the remote people and interact with them (see Figure 2).

Boyle et al.[4] showed that a blur filter is an effective way of masking out potentially sensitive information in an always-on video link. They also proposed to adapt the blur level to the distance between the user and the communication device, although their system only used three different levels. In contrast, instead of creating a series of shared spaces corresponding to particular interpersonal distances, MirrorSpace aims to create a continuum of space that will allow a variety of interpersonal relationships to be expressed.

MirrorSpace includes a proximity sensor that measures the distance to the closest object or person in front of it. A blur filter is applied on the images to visually express a distance computed from the local and remote sensor values. Blurring distant objects and people lets the up-close viewer perceive distant movement or passing with minimum involvement. It also offers a simple way of initiating or avoiding a change to a more engaged form of communication by simply moving closer (see Figure 3) or farther away.

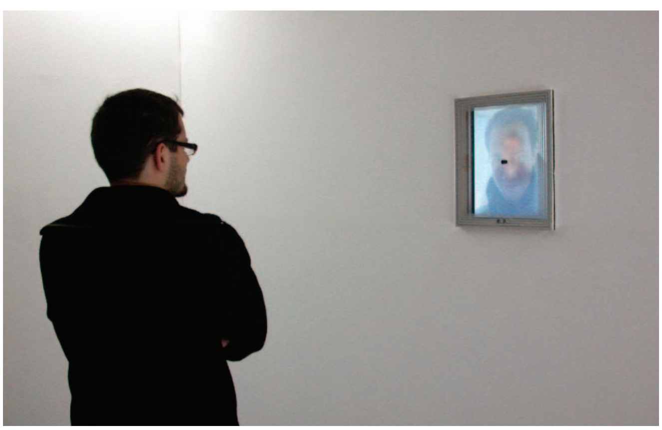

Figure 4. MirrorSpace installation at Mains d'Oeuvres (Paris, May 2003).

MirrorSpace installations

Several pairs of MirrorSpace prototypes have been built and presented to the public as an interactive video installation in four art exhibitions in February, May (Figure 4), July, and November 2003 (Figure 5). These exhibitions gave us the opportunity to observe a large number of people interacting with MirrorSpace in a controlled technical environment. Each of these exhibitions was also an occasion to refine the prototypes' design and explore new software or hardware possibilities.

Each prototype is made of a thin-film technology liquid-crystal display flat screen, a universal serial bus camera, an ultrasonic proximity sensor, and a computer that runs dedicated software. We designed the prototypes to minimize their technological appearance. The computer and all the wires are hidden from users. The screen and its attached sensors are placed into a wooden box, protected by transparent glass (Figure 6). The screen is oriented in portrait mode and part of the protective glass is covered with mirror film to further push the augmented mirror metaphor.

The image sensor and the camera lens are placed on the protective glass that covers the screen, and then connected to the camera's logic board using hair-thin isolated wires running over the glass. The proximity sensor is placed at the bottom of the screen and connected to the computer via a serial interface.

The software uses the videoSpace library[5] to capture images in real time and send them to the other prototypes, along with proximity sensor values. We're able to connect more than two prototypes, although we never did for the exhibitions. The software applies a two-pass incremental blur filter on each image. The resulting images are then flipped horizontally to produce the expected mirror effect and superimposed using OpenGL.

The system uses the distances measured by all the connected prototypes to compute the blur level to apply to each image. We've investigated three computation modes so far: the first mode uses the distance between people and their screen, the second one uses the sum of these distances, and the third one computes a virtual relative distance from them. Although the software lets you specify a different mode for each prototype, the configurations used for the exhibitions always imposed a strict "what you see is what I see" (WYSIWIS) condition.

Initial user reactions

Several hours of video were shot during the exhibitions, showing visitors interacting with the prototypes and what was displayed on the screens. Although the context isn't exactly representative of a remote video communication, a number of observations are worth reporting, as they're probably related to the nature of MirrorSpace itself rather than this particular context.

Although we tried our best to avoid it, a small

Figure 5. MirrorSpace installation at the Centre Pompidou (Paris, November 2003).

Figure 6. Close up showing the proximity (bottom) and image sensors (center).

delay (up to 500 milliseconds) was sometimes perceived between the capture and the display of images. Surprisingly, most people didn't pay attention to it and some liked it: they were running back and forth to play with their own image and see the blur effect in action. Some people even thought the delay had been introduced on purpose. This illustrates the important difference between the technical preoccupations (focusing mostly on function) usually associated with digital video and how users perceive a system like MirrorSpace that focuses on the use of the images and user experience. We discuss this further in the sidebar, "About Video and Time Delay."

Almost all visitors of the exhibitions agreed on one point: interacting with MirrorSpace is fun. Proximity sensing helps create an intimate relationship between users and the system. As we said, many of them played with their own image and the blur effect. People didn't hesitate to make a fool of themselves and many took pic-

About Video and Time Delay

Artists like Dan Graham already use time-delay mechanisms in mirror-based installations to let viewers see themselves as both subject and object. (A description of opposing mirrors and video monitors on time delay is available from http://www.sfmoma.org.) We believe that one of the reasons why people weren't bothered by the delay when interacting through MirrorSpace is that it affected both the remote person's image as well as their own simultaneously and was thus immediately perceived and understood. It isn't clear, however, whether the understanding would be the same in the case of a real remote communication.

tures or recorded video clips of themselves and others interacting with the system.

When they saw another person appearing next to them on the screen, many people turned around, looking for that person behind them. This clearly shows that MirrorSpace creates a sense of shared space and that it's perceived as a mirror more than a video communication system. In fact, the majority of the people didn't think about the camera at all. Only after playing with the system for some time did they suddenly ask with surprise, "Where is the camera?"

People who were visiting the exhibitions with friends or relatives tried to overlay their faces. Some went as far as kissing each other. At the same time, other persons were surprised and even disturbed to find strangers able to come so close to them. In that case, they backed away, which made their own image disappear smoothly with the blur effect. This shows that MirrorSpace supports at least part of our accustomed body language.

Directions for future work

One important step for future studies will be the building of other MirrorSpaces. We plan to deploy and demonstrate the system in various other contexts (for example, family households, different buildings of the same research group, and so on). This should help us collect more qualitative and quantitative data about the system's use. In particular, it should be easy to measure the actual time people spend at each distance according to Hall's classification.[2]

We're investigating several technologies that would let us embed the image sensor in the protective glass itself. We're also working on the design of an auditory equivalent to MirrorSpace that could be combined with it in future installations. The challenge here is to design an equivalent to the blur effect that would provide general audible awareness of people far away from the sensor and spoken communication with them as they move closer.

More information on MirrorSpace—including the source code, some images, and videos—is available at http://insitu.lri.fr/~roussel/. **MM**

References

1. W. Mackay, "Media Spaces: Environments for Informal Multimedia Interaction," *Computer-Supported Cooperative Work Trends in Software Series*, John Wiley & Sons, 1999.
2. E.T. Hall, *The Hidden Dimension: Man's Use of Space in Public and Private*, Doubleday, 1966.
3. O. Morikawa and T. Maesako, "HyperMirror: Toward Pleasant-to-Use Video Mediated Communication System," *Proc. ACM Conf. Computer-Supported Cooperative Work* (CSCW 98), ACM Press, 1998, pp. 149-158.
4. M. Boyle, C. Edwards, and S. Greenberg, "The Effects of Filtered Video on Awareness and Privacy," *Proc. Conf. Computer-Supported Cooperative Work* (CSCW 00), ACM Press, 2000, pp. 1-10.
5. N. Roussel, "Exploring New uses of Video with VideoSpace," *Proc. Int'l Federation for Information Processing's Working Conf. on Eng. For Human–Computer Interaction* (EHCI 01), LNCS 2254, Springer, 2001, pp. 73-90.

Contact author Nicolas Roussel at roussel@lri.fr; contact authors Helen Evans and Heiko Hansen at {helen, heiko}@hehe.org.

Contact Artful Media editor Doree Duncan Seligmann at Avaya Labs, 666 Fifth Ave., 11th floor, New York, NY 10103; doree@avaya.com.