Reflexion: a responsive virtual mirror

Cian Cullinan and Stefan Agamanolis
Human Connectedness group
Media Lab Europe
Sugar House Lane, Bellevue, Dublin 8, Ireland
{cian, stefan}@mle.ie

ABSTRACT
Reflexion is an interpersonal video communication system that operates on the metaphor of a "magic mirror" in which one sees a reflection of oneself along with the reflections of other participants in remote locations, overlaid on a common backdrop image or movie. The system responds to visual and auditory cues to appropriately compose the scene and emphasize the center of attention. Participants can use brightly colored objects to control and interact with the system. In this paper, we describe the basic design features of Reflexion and some scenarios for how it might be used toward the aim of creating more intimate shared environments for communication, collaboration, and fun.

Keywords
Telepresence, media space, video conferencing, remote collaboration, video object tracking

INTRODUCTION
Traditional multi-party video conferencing systems often display participants in separate windows, in scenes reminiscent of the title sequence of the TV show *The Brady Bunch*. We believe the visual separation characteristic of these multi-window designs introduces a confrontational or divisional dynamic that can have a negative impact on a meeting or interaction even before it begins. In the absence of a multi-window design, some systems employ audio-based automated camera selection algorithms to switch between views of the active participants. This approach can result in an even greater sense of separation since individual users do not appear on the screen together, and it also limits awareness of the inactive participants.

Reflexion is a multi-point interpersonal communication system that broadly aims to address these issues and create a more intimate dynamic that fosters an enriched sense of togetherness between its users. The system employs audio and video analysis and an interface design based on the metaphor of a "magic mirror" with the goal of emphasizing the center of attention while preserving a sense of background awareness between all participants. Our overall aim is to create a system that responds and adapts to its participants in a way that truly enhances and extends natural human communicative abilities, beyond what is possible in a mere physical face-to-face interaction.

RELATED WORK
The idea of identifying the center of attention in a video conference can be traced back over 30 years to a "voice voting" system that automatically switched between multiple close-up camera views based on who is speaking [3]. In the *Personal Presence System*, users had personalized control over the size and positioning of the remote participants' video streams on the screen [2]. The *Clearboard* system explored the notion of seamlessly integrating shared work and interpersonal spaces into a common environment [4]. *HyperMirror* has experimented with blue screen technology to layer one participant into a scene consisting of another [5].

Figure 1. Screen shots of Reflexion. Transparency and layering are dynamically altered as participants become more or less active in the conversation. Two participants are shown, but 3 or more are possible.

SYSTEM DESIGN
Reflexion is the continuation of project named *Reflection of Presence*, which began at the MIT Media Lab over 5 years ago [1]. This new version consists of a reinvented infrastructure that employs higher resolution imaging and a lower-latency peer-to-peer strategy for audio and video transmission. Below is a brief description of some of the key features of our design:

Foreground segmentation

Each Reflexion station consists of a standard desktop PC, fitted with a video capture card, a video camera, and an echo canceling speaker/microphone. Using an in-house segmentation algorithm, the system extracts images of the participants from their backgrounds at full frame rate. This algorithm works on arbitrary backgrounds as long as they are static or only changing very slowly.

Mirror effect

The system horizontally mirrors the extracted participant images and combines them together into a single scene. The effect is one of a "magic mirror" in which the participant sees a reflection of him/herself as well as reflections of the remote participants. Every participant sees exactly the same thing, enhancing the sense of inhabiting a shared space.

Dynamic composition

The system tracks which participants are speaking and gesturing in order to judge who is the center of attention. Active participants are rendered opaque and in the foreground to emphasize their visual presence, while other less-active participants appear slightly faded in the background in a manner that maintains awareness of their state without drawing undue attention. In cases where multiple users are speaking, a policy of "first come first served" determines who retains the frontmost layer. The system smoothly transitions the layering and appearance of the participants as their interactions continue.

Object tracking

Participants may train the system to recognize brightly colored objects and later use these objects to interact with the system in various ways – as pointing devices for example. This enables more relaxed scenarios in which a traditional desk environment and standard interface devices like keyboards and mice may not be desirable.

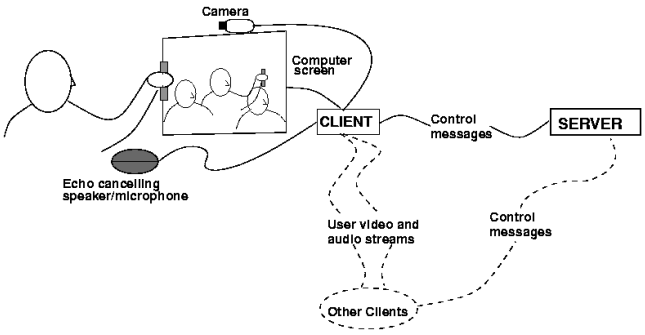

Figure 2. Diagram of a Reflexion station showing the peer-to-peer network topology for audio and video streams.

SCENARIOS

The basic Reflexion framework could be applied to a number of specific application scenarios, four of which are described below:

- Tele-meetings. The basic Reflexion framework could be used for formal telecollaboration as well as informal social interaction. The current system allows participants to manipulate documents or images in the background to create a sense of a shared workspace.
- Distributed classrooms. Reflexion could support synchronous distributed class meetings that focus attention on the lecturer while maintaining awareness of and between the students. The interface could be modified to handle a larger number of participants appearing at a smaller size.
- Multi-actor remote interactive theater: The Reflexion framework could be harnessed as an environment for remote theatrical productions in which the actors are physically distant from each other. The system could track the presence and gestures of the actors and cue sound effects or changes in backdrop.
- Watching movies or television: Reflexion can create an environment for watching movies or live television shows together with friends or family in far away places.

EXPERIENCE AND FUTURE DIRECTIONS

Informal testing of Reflexion during its development has been encouraging, particularly within the last scenario described above. Some first-time users report that seeing themselves on the screen is awkward, but this feeling appears to diminish with continued use. Many have reported liking this feature as it provides added awareness of how one is being portrayed to the other participants. Formal studies of these effects are planned. Future experiments will include scaling the system to work fluidly with very large numbers of participants and developing new techniques for more natural interaction.

ACKNOWLEDGMENTS

We wish to thank the Object-Based Media group at the MIT Media Lab, especially Parul Vora for her great ideas and help in testing. This research has been supported by sponsors and partners of Media Lab Europe.

REFERENCES

1. S. Agamanolis, A. Westner, and V. M. Bove, Jr., "Reflection of Presence: Toward more natural and responsive telecollaboration," *Proc. SPIE Multimedia Networks*, 3228A, 1997.
2. D. G. Boyer and M. E. Lukacs, "The Personal Presence System: a Wide-Area Network Resource for the Real-Time Composition of Multipoint Multimedia Communications," *Multimedia Systems, 4*, 1996, pp. 122-130.
3. R. C. Edson, et al., "Video Conference System using Voice-Switched Cameras," U.S. Patent 3,601,530, 1971.
4. H. Ishii, M. Kobayashi, and K. Arita, "Iterative Design of Seamless Collaboration Media," *Communications of the ACM, 37*, August 1994, pp. 83-97.
5. O. Morikawa and T. Maesako, "HyperMirror: Toward Pleasant-to-use Video Mediated Communication System," *Proc. CSCW '98*, ACM Press, pp. 149-158.