USING SPATIAL CUES TO IMPROVE VIDEOCONFERENCING

Abigail Sellen and Bill Buxton*
Computer Systems Research Institute
University of Toronto
Toronto, Ontario
Canada M5S 1A1

John Arnott,
Arnott Design Group
33 Davies Ave.
Toronto, Ontario
Canada M4M 2A9

Figure 1. A user is seated in front of three Hydra units. Each Hydra unit contains a video monitor, camera, and loudspeaker.

INTRODUCTION

In this video we describe and demonstrate Hydra, a prototype system for supporting four-way videoconferencing. The design is intended to build as much as possible upon existing skills used in face-to-face discussions.

A conventional approach to multiparty videoconferencing is to support a four way meeting using a Picture-in-a-Picture (PIP) device. In this approach, each remote participant's image is placed in one quadrant of the screen of a single monitor. This common view is then distributed to each person. In addition, the audio from each participant is combined, and all voices emanate from a single loudspeaker.

*Abigail Sellen is now at Rank Xerox Cambridge EuroPARC, and the MRC Applied Psychology Unit, Cambridge, UK.

Because each participant has a single monitor, camera, and loudspeaker, PIP videoconferences are limited in their support of participants' ability to:

- establish eye contact with other participants;
- be aware of who, if anyone, is visually attending to them;
- selectively listen to different, parallel conversations;
- make side comments to other participants;
- hold parallel conversations.

Hydra, on the other hand, is intended to preserve the unique personal space that participants occupy in face-to-face meetings. In simulating a 4-way round table meeting, the place that would otherwise be occupied by a remote participant is held by a Hydra unit as shown in Figure 1. Each Hydra unit consists of a camera, monitor, and speaker. Hydra units are, in effect, "video surrogates" for the participants, occupying the physical space that would be held by people, if they were physically present. The technique used is similar to that of Fields (1983), although it was developed independently.

The result of this technique is that each participant is presented with a unique view of each remote participant, and that view and its accompanying voice emanates from a distinct location in space. The net effect is that conversational acts such as gaze and head turning are preserved because each participant occupies a distinct place on the desktop.

The fact that each participant is represented by a separate camera/monitor pair means that gazing toward someone is effectively conveyed. In other words, when person A turns to look at person B, B is able to see A turn to look towards B's camera. The spatial separation between camera and monitor is small enough to maintain the illusion of mutual gaze or eye contact. Looking away and gazing at someone else is also conveyed, and the direction of head turning indicates who is being looked at. Furthermore, because the voices come from distinct locations, one is able to selectively attend to different speakers who may be speaking simultaneously.

The ways in which the design of Hydra affects behaviour is currently being investigated experimentally. The first of these analyses appears in these proceedings (see the paper by Sellen). Preliminary analysis of the data indicates that Hydra is successful in supporting selective attention both visually and auditorily. In addition, the data show that Hydra does make aside and parallel conversations possible.

A key aspect of the success of the design of Hydra is the contribution of industrial design. We describe and illustrate this process. We also show one office with three prototypes designed by the Arnott Design Group, and contrast that with a room equipped with standard video equipment.

ACKNOWLEDGMENTS

This work was undertaken as part of the *Ontario Telepresence Project.* It has been sponsored by the Arnott Design Group, the Information Technology Research Centre of Ontario, Xerox PARC, IBM Canada's Laboratory Centre for Advanced Studies (Toronto), Apple Computer's Human Interface Group, and the Natural and Engineering Science Research Council of Canada. This support is gratefully acknowledged.

REFERENCES

Buxton, W. and Sellen, A. (1991). *Interfaces for multiparty videoconferencing.* Unpublished paper. Dynamic Graphics Project, Dept. of Computer Science, University of Toronto: Toronto, Canada.

Fields, C.I. (1983). Virtual space teleconference system. *United States Patent 4,400,724,* August 23, 1983.

Sellen, A.J. (1992). *Speech patterns in video-mediated conversations.* In these proceedings.