

Buxton, W., Sellen, A. & Sheasby, M. (1997). Interfaces for multiparty videoconferencing. In K. Finn, A. Sellen & S. Wilber (Eds.). *Video Mediated Communication*. Hillsdale, N.J.: Erlbaum, 385-400.

Interfaces for Multiparty Videoconferences¹

William A. S. Buxton

Computer Systems Research Institute, University of Toronto

&

Alias | Wavefront Inc., Toronto

Abigail J. Sellen

Rank Xerox Research Centre (EuroPARC), Cambridge

Michael C. Sheasby

SOFTIMAGE/Microsoft, Montreal

Abstract

We describe how conventional approaches to multiparty video conferences are limited in their support of participants' ability to: establish eye contact with other participants; be aware of who is visually attending to them; selectively listen to different, parallel conversations; make side comments to other participants and hold parallel conversations; perceive the group as a whole; share documents and artifacts; and see co-participants in relation to work-related objects. We present some design alternatives to these conventional videoconferencing approaches, describe the prototypes we have developed, and discuss their experimental evaluation.

¹ This version of the document includes some typographical corrections and minor changes in wording to improve clarity. No changes have been made which would impact the meaning or content of the original.

Introduction

Designing to support multiparty videoconferences — conferences which involve more than two sites or more than two people — presents challenges beyond the design of simple point-to-point video systems. In a conventional videoconferencing set-up, everyone is seen on one or two monitors, cameras are fixed, and what control is available comes typically through a cumbersome interface. Even in the two-party situation, such technology presents a variety of problems including lack of eye contact, limited fields of view, asymmetrical control of cameras, and difficulty in knowing how one's voice is perceived by one's co-participant. Contrast this with meetings where all participants are in the same room — each person is their own "display" (i.e., their body), has full control over their own "camera" and "speaker" (i.e., their eyes and voice), and is fluent with the user interface to this "technology".

This chapter focuses on the particular problems of supporting multiparty meetings with video. In some respects, multiparty meetings exacerbate the problems inherent in two-party video meetings. In other respects, they present problems specific to the multiparty case. By experimentally evaluating conventional approaches to multiparty videoconferencing, we are able to explicate many of these problems. We then suggest design alternatives in the form of prototype systems which are themselves subjected to empirical evaluation. The primary intent of this chapter is to communicate the rationale behind our different design ideas, what we have learned from implementing and evaluating them, and the direction that we are heading in the future.

The PIP Approach

The most common method of supporting multiparty video conferences involving several sites is to use a picture-in-picture (PIP) approach². Using this technique, a single video screen is tiled into sub-screens, each containing the output of one video source. Figure 1 shows a schematic of a 4-way PIP, where the picture from each camera appears in a different quadrant of the screen. Figure 2 shows a user involved in a 4-way conference using such a PIP device.

This technique has the advantage that all parties can see each other. It is also technologically straightforward and reasonably economical, and therefore commonly used in commercial systems.

² The term "picture-in-a-picture" can also refer to the technique of embedding one small video picture into a larger full-screen background picture. We use the term here to refer to a 2 by 2 video tiling approach.

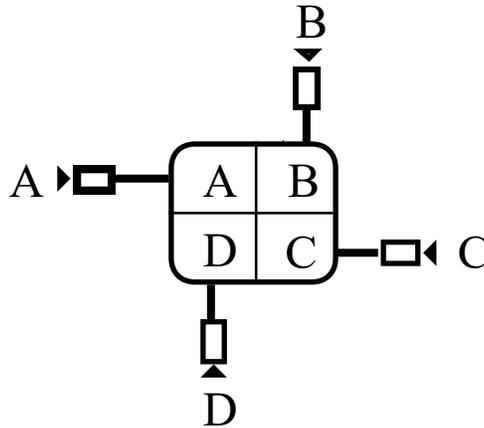


Figure 1. The output of multiple cameras A, B, C and D (each at different sites) shown tiled, in separate quadrants of the screen. Typically, the images are combined at a central location using the PIP device. The output is then broadcast to each participant.

One obvious problem with this approach is that it breaks down as the number of remote sites increases due to the decreasing size of the tiled images. But closer consideration reveals a number of other problems in supporting multiparty videoconferences this way.

First, participants using this approach are limited in their ability to establish eye contact with other participants, and to be aware of who, if anyone, is visually attending to them. Because there is a single camera and monitor, participants cannot tell who is looking at them as opposed to the other participants. Neither can they establish eye contact with any one of the participants to the exclusion of the others (mutual gaze). Further, because all participants occupy the same general area in the visual field (i.e., a single monitor), there is no need to turn the head to speak or listen to different participants. One can assume that supporting head-turning and gaze is an important consideration, as they have been shown to serve communicative functions as well as helping to manage turn-taking and floor control (Argyle et al., 1973; Exline, 1971).

Participants using this approach are also limited in their ability to listen to simultaneous conversations. One significant factor contributing to this problem is the way the audio is configured. Typically, the audio from all participants comes from a single speaker. In contrast, when people physically occupy the same room, separate speech streams emanate from different points in space. It is this in part which makes it possible to selectively attend to ongoing parallel conversations (the "Cocktail-Party Effect", (Cherry, 1953; Egan, Carterette & Thwing, 1954). This is made difficult when these spatial cues are eliminated.

These problems taken together represent serious design deficiencies which motivated us to try a different approach which would offer support for selective gaze and head-turning, and for selective listening.



Figure 2. A four-way videoconference using a PIP device. All participants see the same split screen, which includes an image of themselves.

Hydra

We call our first alternative design *Hydra*. The basis for the Hydra system is to preserve the notion of personal space to preserve the everyday skills of conversational acts afforded by such space.

The underlying concept behind Hydra is to replace each of the remote meeting participants with a video surrogate (Sellen, Buxton & Arnott, 1992).³ In simulating a 4-way round-table meeting, the place that would otherwise be occupied by a remote participant is held by a camera, monitor and speaker, as shown in Figure 3.

³ After designing and implementing this system, we became aware of earlier work that used the same approach (Fields, 1983). So, the basic approach is not original, just uncommon. But like conventional videoconferencing, just because it has been done before does not mean that there is not significant room for improvement.



Figure 3. A four-way videoconference using Hydra. Each Hydra unit contains a video monitor, camera, and loudspeaker. A single microphone conveys audio to the remote participants.

Using this technique, each participant is presented with a unique view of each remote participant, and the view and accompanying voice of each 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.⁴ Furthermore, because the voices come from distinct locations, one can selectively attend to different speakers who may be speaking simultaneously.

We carried out a series of empirical studies to examine and quantify the behavioural differences between Hydra and the PIP system (Sellen, 1992; 1995). These studies

⁴ Note that it is important to preserve the geometry of the virtual roundtable with Hydra. That is, if I appear across the table from you, you must similarly do so from me. Likewise, the person to my left must be to your right, etc. This is solved by assuming that whoever calls the meeting configures the seating of the table. Software can then ensure that the "seating" is consistent for all participants.

focused primarily on objective measures of speech such as turn length, amount of simultaneous speech, and floor control parameters.

We hypothesized that the lack of support for selective gaze and head-turning, and for selective listening in the PIP system would affect conversational interaction and make certain conversational acts difficult in comparison to the Hydra system. For example, we predicted that turn-taking might be adversely affected with the PIP system, and that holding parallel conversations and making side comments to others in a group would be difficult.

While there was no significant difference between the PIP and Hydra approach with respect to some measures of turn-taking behaviour, Hydra did, as expected, support parallel and side conversations. No such conversations were observed in the PIP approach. A majority of subjects expressed a preference for Hydra in their subjective evaluations, citing the ability to selectively attend both visually and auditorily as the major reasons for preferring it over the PIP system. Some subjects commented that Hydra has much more of an interactive "feel" about it than the PIP approach to multiparty meetings. Thus, the results are in line with the original intentions motivating the design of Hydra.

We are exploring ways to further exploit the properties of the preserved personal space. For example, by adding a proximity sensor to each Hydra unit, one will be able to establish a private audio link to another participant by leaning towards that person's unit. The gesture is the same as in everyday conversation, and conventional social mores are preserved, since the others can see not only that one person is making a side comment, but to whom. Once this mechanism is in place, and with the benefits of dedicated speakers for each participant, we hope to support parallel conversations, side comments, and breaking into conversational sub-groups even more effectively. All these important aspects of conversations and meetings are poorly supported by existing technology.

Since this system was developed, Ichikawa, Okada, and colleagues (Ichikawa et al., 1995; Okada et al., 1994) have developed a multiparty system which contains some of the same properties of Hydra. The MAJIC system projects life size images on a semi-transparent surface allowing cameras to be placed behind the screen. Speakers are also placed behind the screen image of each participant. Thus, the MAJIC system also provides support for selective gaze and head turning. The much larger images may be a much better approach for many multiparty situations. However, because it uses projection and large screens, one drawback of the system is that it does not sit unobtrusively on a desktop, but is an altogether more imposing type of configuration, with less flexibility to be moved around and combined with other systems, as will be described in the last section of this chapter.

LiveWire: Voice Activated Switching

While Hydra appears promising for small meetings of up to about four participants, like the PIP system, the approach does not scale up very well to larger groups.

Furthermore, despite design ideas that we have developed to minimize the effect, Hydra is equipment intensive.

As a result, we have been looking at alternative design approaches. One which we have implemented is often used in broadcasting and in some videoconferencing systems. The approach, our implementation of which is called *Livewire* (see Sheasby, 1995), involves changing who is visible on the monitor over time, depending on who is speaking. This is illustrated in Figure 4.

The following assumptions formed the basis of this approach:

- All non-speakers see the current speaker "full screen".
- The speaker sees the previous speaker.
- Only one person "owns" the screen at any one time.

The advantage is that the approach scales up well to large groups. It is also an interactive system responding to the dynamics of the conversation. However, it also has some serious drawbacks which were revealed in our empirical studies (Sellen, 1995):

- 1) Subjects commented that they quickly lost a sense of the larger group—people who were not speaking had virtually no presence in this system.
- 2) Speakers complained that they got no feedback or confirmation from the system that they were being seen by the others since any speaker continues to see the previous speaker. Subjects remarked that this was a serious problem.
- 3) The fact that LiveWire allows people only to monitor the speaker and not other people's reactions to what is being said was perceived to be unnaturally restrictive. Subjects disliked the fact that they had no control or choice over whom to monitor.
- 4) The ability to have side conversations, or to make side glances and other non-verbal communicative acts was not supported as it was in Hydra. This was largely due to subjects having difficulty, especially at first, in assessing who was looking at whom.
- 5) The automatic switching was often distracting and inappropriate, especially when people in the group coughed or laughed, causing the screen to quickly switch from one person to the next.

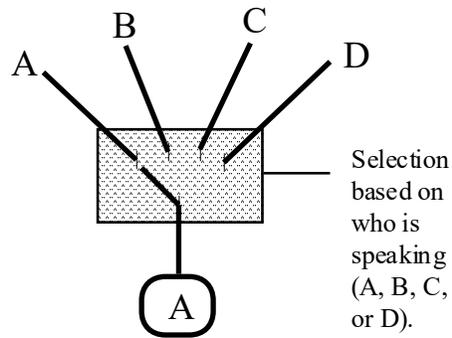


Figure 4. Voice-Activated switching. "Livewire" is an implementation of a voice-activated switching system. The voice of the speaker causes the speaker's image to be seen full frame on all other screens.

These design flaws represent considerable problems for systems like Livewire that depend on voice-switched full screen images. Not only have we found that this sort of "tunnel vision" is inappropriate in a multiparty situation, but that the lack of control over this selective view is also problematic. When Livewire was compared with the PIP system and an audio-only system (Sellen, 1995), the majority of the subjects said they liked the PIP system best, preferring the Livewire system only slightly more often than having no video at all.

While obviously not an ideal solution to supporting multiparty conferences, one advantage of developing Livewire was to allow us to assess a system similar to what is commercially available, and to use it to compare our alternative designs to current practice. In addition, evaluating the shortcomings of such systems can serve as a basis for further design innovations, as is described in the next section.

The Brady Bunch: LiveWire Meets Portholes

Taking some of the shortcomings of Livewire into account, we designed a successor which builds on the voice-switching approach. A prototype of this system, which we have named the *Brady Bunch*,⁵ was implemented and evaluated (Sheasby, 1995).

Building upon the Livewire technology, the Brady Bunch was partially inspired by two systems developed at Rank Xerox EuroPARC and Xerox PARC: *Portholes* (Dourish & Bly, 1992), and its predecessor, *Polyscope* (Borning and Travers, 1991). In brief, Portholes, (illustrated in Figure 5), is a system which repeatedly takes and distributes snapshots of the workgroup to the workgroup. The images are shot using one or

⁵ Named because of its similarity in appearance to the opening credits of that TV series of that name.

more frame-grabbers which have access to the group members' video cameras (without disrupting other uses of the cameras, such as conferencing). The individual snapshots are subsampled and distributed over the local (or wide) area network servicing the group, and then combined with the shots of others in the group. The net effect is that each group member receives relatively recent still pictures of the office or workspace of each group member, which are displayed on their workstation. Portholes also has embedded functionality that permits users to access one another over the accompanying A/V network. Hence, it has a control as well as an awareness function.

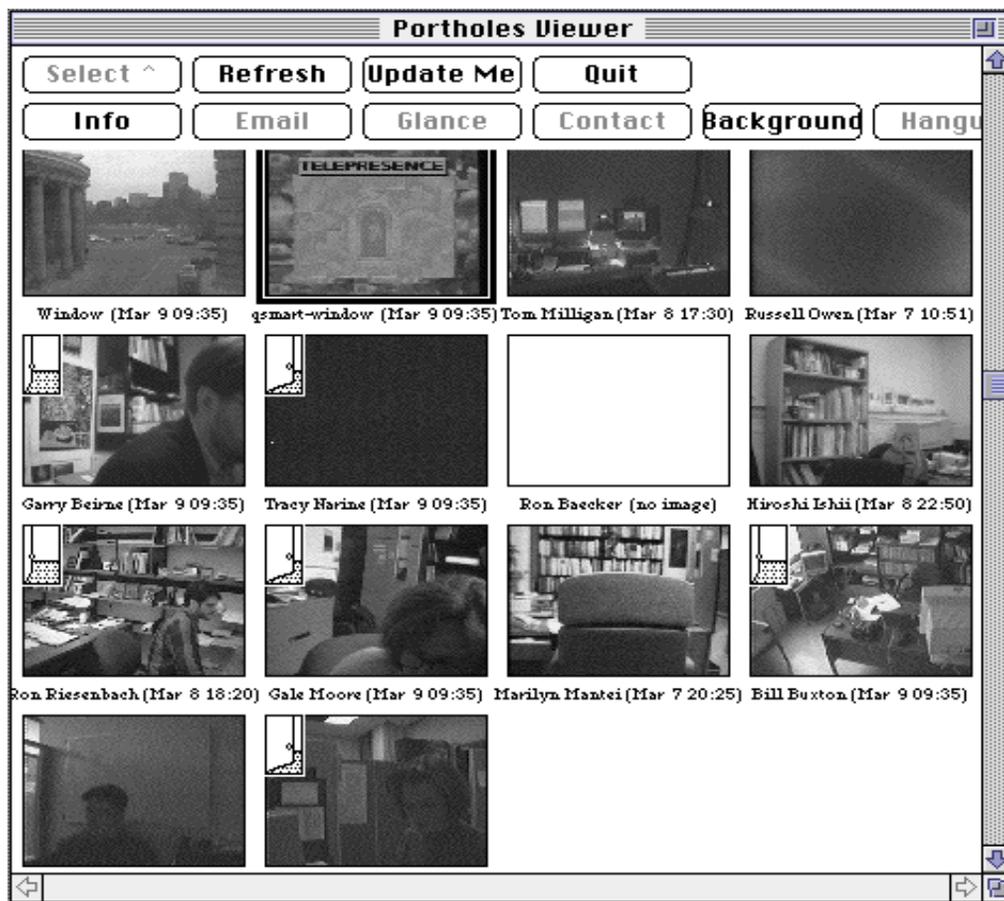


Figure 5. The Telepresence implementation of Portholes. Every 5 minutes, a snapshot of each member of the workgroup is distributed to all other members. In the Telepresence implementation, this is accompanied by an icon of that member's door icon, which indicates that person's degree of accessibility. The resulting tiled image of one's workgroup affords a strong sense of who is available when. It also can serve as a mechanism for making contact, finding phone numbers, and avoiding intruding on meetings.

The *Brady Bunch* design combines the Portholes/Polyscope approach with *Livewire*. A live voice-switched image is supported by a set of slow-scan video images. The static images are snapshots of the other meeting participants, grabbed using a technique similar to Portholes. While the initial design placed the slow-scan images in a ring around a larger live image directly on the workstation monitor, the first implementation of the Brady Bunch (Sheasby, 1995) placed the live image on a separate monitor, leaving the slow-scan images on the user's workstation desktop.⁶

The Brady Bunch was designed to be used in focused group interaction, where all group members play an active role in a discussion. In normal operation, the current speaker is displayed in the large Livewire image, while the other meeting participants are displayed in the slow-scan images. The slow-scan images provide a sense of the context of the larger group and give group members who are not talking some presence in the meeting. This addressed the first problem that we found with Livewire.

The second problem of lack of feedback was addressed by the addition of an "on camera" indicator to the Livewire system. This consisted of superimposing a red dot on the live image displayed in the current speaker's video monitor to confirm to them that they were being viewed by the others.

The third and fourth problems — the ability to glance at others, and to have side conversations with them— was addressed by the addition of two features. The first feature allows a user to "glance" at another user (view someone other than the speaker in the main window) by clicking on that person's slow-scan image. That person is then displayed as full motion video on the live monitor, replacing the speaker. This allows participants to override the voice-activated switching system to monitor non-speaking members of the meeting. The second feature allows two users to have "side conversations" by allowing them to drop out of the group meeting to communicate privately with each other. In this mode, pairs of users can communicate via a private and secure audio-video link. The method of connecting like this is similar to that for glancing at another user but involves acceptance by the remote user.

In face-to-face meetings, there are many inherent visual cues that convey the fact that one is being glanced at. To provide this kind of information in the Brady Bunch system, we used the slow scan images to present status cues. For example, if one was being glanced at, the name of the person glancing would alternate with the word "glancing" in the slow-scan window representing that person. Requests for side conversations were handled similarly.

⁶ One benefit of this was that it made it simple to determine a user's focus of interest for the purposes of evaluation (i.e., whether the user was looking at the current speaker on the live monitor, the rest of the group on the workstation monitor, or elsewhere).

The Brady Bunch was tested using the board game 'Diplomacy'. In this game of strategic negotiation, players attempt to dominate a stylized map of the world by invading one another's territory (see Figure 6). The rules are set up so that a player is unlikely to win alone; the players are intended to form alliances with one another to win specific battles. The point of the game is that players must negotiate with skill and persuasiveness, since treaties can be ignored and cheating one's allies is common behaviour.

The game was chosen because it depends heavily on the accurate assessment of the sincerity of a distant user. In this respect the game reflects actual negotiation, a common and important business practice. Thus, although difficult to measure, a player's success at the game is directly related to the translation of their face-to-face communication skills to the teleconferencing medium.

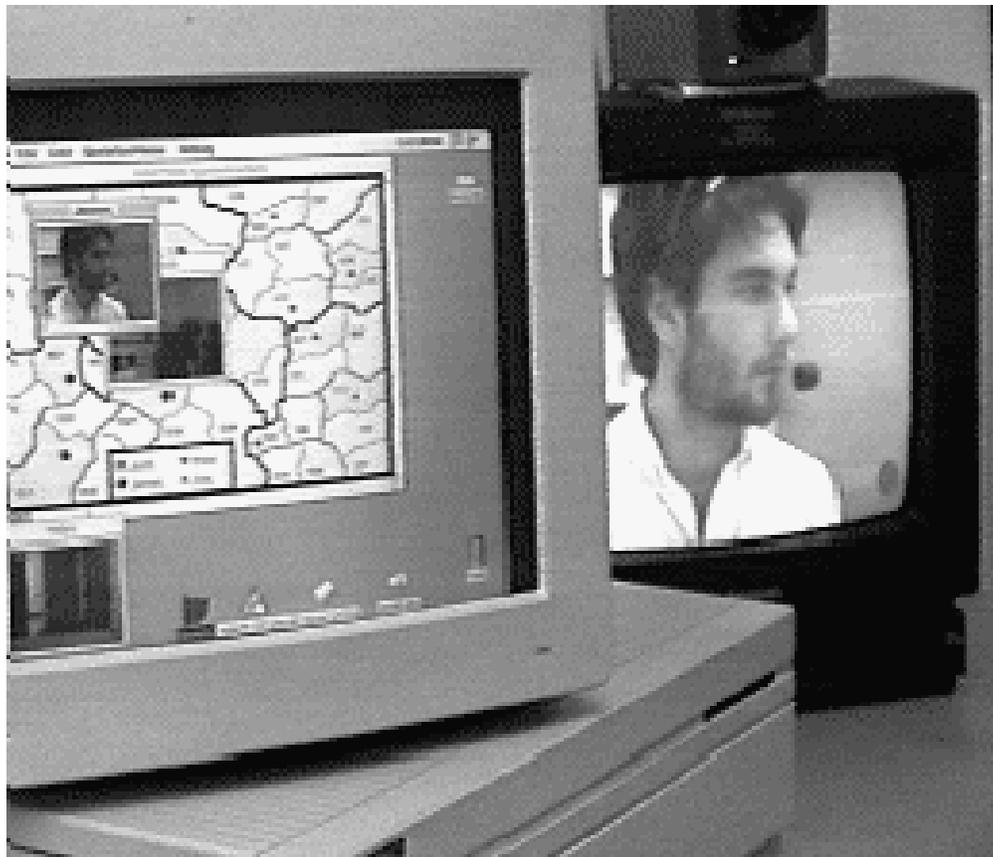


Figure 6. The Brady Bunch Approach used with the game "Diplomacy". A full-motion voice-switched video image of the current speaker on a separate monitor is supported by slow-scan images of all meeting participants in separate windows on the workstation display.

In the experiment, subjects made heavy use of the glance and side conversation features in the Brady Bunch system, although the difference between them appeared hazy to some subjects. During these side conversations, users could be seen to spend a great deal of time visually monitoring each other as if trying to assess the truth of what the other was saying. Thus, the ability to monitor someone other than the speaker, and to break into conversational sub-groups was shown to be important, at least in this kind of game situation.

The experimental evaluation also revealed that users wanted the system to enable them to engage in side-conversations of more than two people. They also wanted the system to provide them with information about when side conversations or glances were occurring between participants other than themselves.

In a subsequent version of the Brady Bunch, we intend to explore better ways of providing feedback. One potential solution is to highlight the borders of the slow-scan windows of users to tell each participant who is viewing them. For example, if I am talking, under normal circumstances, all participants' borders will be highlighted to indicate that everyone is viewing me. If I then lose the floor, the windows revert to their normal state. If I am *not* talking, I may still be glanced at by others, which would be indicated by those people's windows being highlighted. Notice that this solution removes the need for the red "on camera" dot in the live monitor.

What is missing in this approach, however, is the provision of feedback to users to tell them that *other* people are glancing at or are having side conversations with each other. Private conversations between distant users could be indicated with another form of highlighting, but other solutions need to be explored, such as altering the layout of the windows to indicate connections between distant users.

This method of providing information about who is attending to whom is intended to compensate for the lack of head turning and gaze cues people use in everyday conversation, and which we have sought to provide in Hydra. We hope to experiment to see whether this kind of compensation is effective.

In addition, like most existing practice, this approach does not have the spatial audio cues that formed the basis of Hydra. We may be able to effectively spatially distribute the individual voices using techniques such as those described by Ludwig et al. (1990), and Cohen & Ludwig (1991).

Integration of Shared Person and Task Spaces

To this point, our discussion has focused exclusively on the shared "person space" of the participants involved in a multiparty conference. The discussion is not complete, however, without considering the shared space of the documents, applications or other artifacts that are part of the reason for meeting. This we call the shared "task

space". What we have argued elsewhere (Buxton,1992), is that these two notions of shared space must be seamlessly integrated to achieve a true sense of *telepresence*. This is as important in the multiparty case as it is in the dyadic meetings. However, the multiparty case presents additional design and engineering complications. In this section, we shall outline two examples of how we have approached this problem.

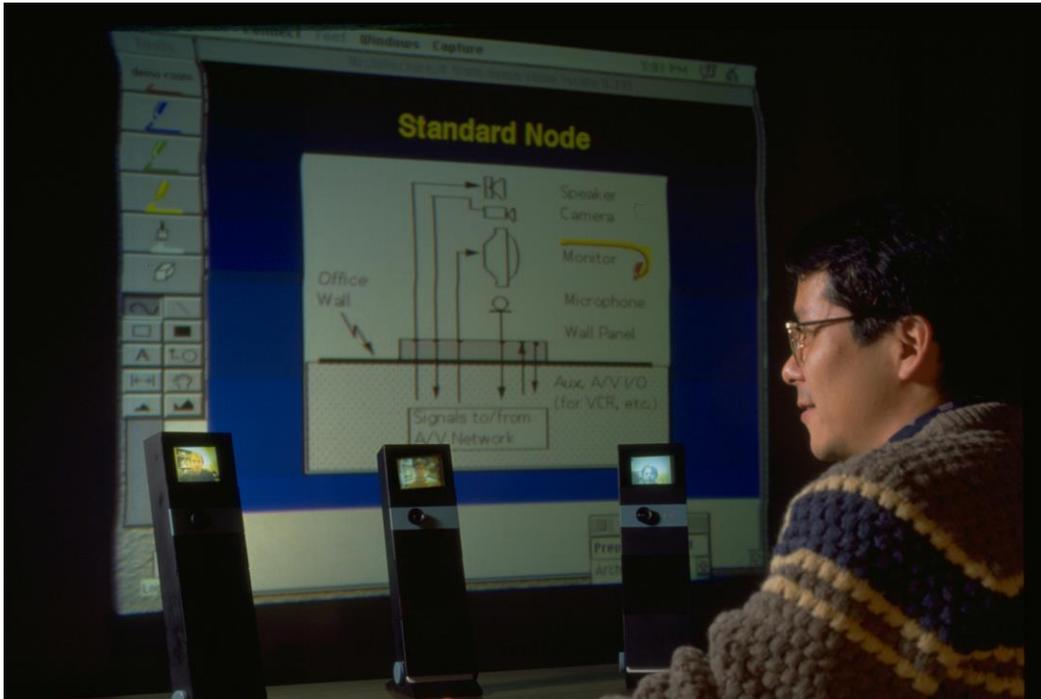


Figure 7. Shared task and person space. A multiparty meeting concerning a technical drawing is illustrated. The technical drawing is displayed on the large screen behind the Hydra units (which are used for the shared presence of the participants). Each participant can see and mark up the technical drawing. The configuration supports gaze awareness towards people and document.

The Electronic Whiteboard Case

In this case, illustrated in Figure 7, we combine the Hydra technique with a large data display, reminiscent of the Xerox *Liveboard* (Elrod, *et al.*, 1992). In this configuration, the Hydra devices function much as before. The major addition is that the sense of gaze awareness afforded by Hydra now extends to the shared document displayed on the large screen, as well as among the participants. For example, if the other participants are looking at the current speaker, and that speaker looks up at the large screen, the other participants will be aware of this, and follow the speaker's gaze. Thus, this configuration not only provides access to shared documents, but also gives some sense of one's co-participants' orientation to those shared documents.

It is worth briefly contrasting this configuration with the ClearBoard system of Ishii *et al.* (1993). Briefly, ClearBoard superimposes the image of the remote person on the work surface. In the dyadic case, this affords excellent and seamless fine grain gaze awareness. However, while elegant, the technique breaks down in the multiparty case. Hence, our need to pursue other design alternatives.

Finally, note that in at least one way, this electronic configuration improves upon the analogous “same place” configuration. Assuming that the configuration is replicated for all participants, each participant has the electronic whiteboard right in front of them. In contrast, in the same place, round-table situation, some participants would have to turn partially or completely around to see the physical whiteboard.

The Active Desk Case

In the case, illustrated in Figure 8, we have configured a Hydra unit on the top right corner of an electronic drafting table, or *Active Desk*.⁷ With the Active Desk, the user’s “desk-top” is an interactive 100x66 cm flat rear projection display. There is no desktop computer, nor desktop metaphor — the desktop *is* the computer. Electronic documents, shared or otherwise, appear on its surface, and one interacts with them with a stylus, keyboard, or some other input device.

Key to this configuration is the fact that the Hydra units can be placed around the periphery of the desk, thereby affording a seamless way of integrating conversation and collaborative interaction with a document. Overall, the approach has been to model the social and interaction skills seen in the everyday world: that is, people standing around a drafting table, discussing the document, and changing their gaze from document to person by simply raising/lowering, or turning their heads. Again, this approach tries to provide some support for conveying people’s orientation to shared, work-related documents.

Worth noting is how the previous two examples can be combined. Imagine that the person shown in Figure 7 is also working on an Active Desk. Furthermore, let us assume that information on an individual’s desk is their private space, and information on the electronic whiteboard is public. From the resulting relationship between space and function, the power of gaze awareness is extended. Now, for example, I can tell if you are looking at me, at the public space, or your private notes. Our assumption (one which we are exploring more formally) is that these additional cues — being based on everyday skills — facilitate the quality of the interaction and the naturalness of the ensuing dialogue.

⁷ The Active Desk was developed jointly by the Ontario Telepresence Project and the Arnott Design Group of Toronto.

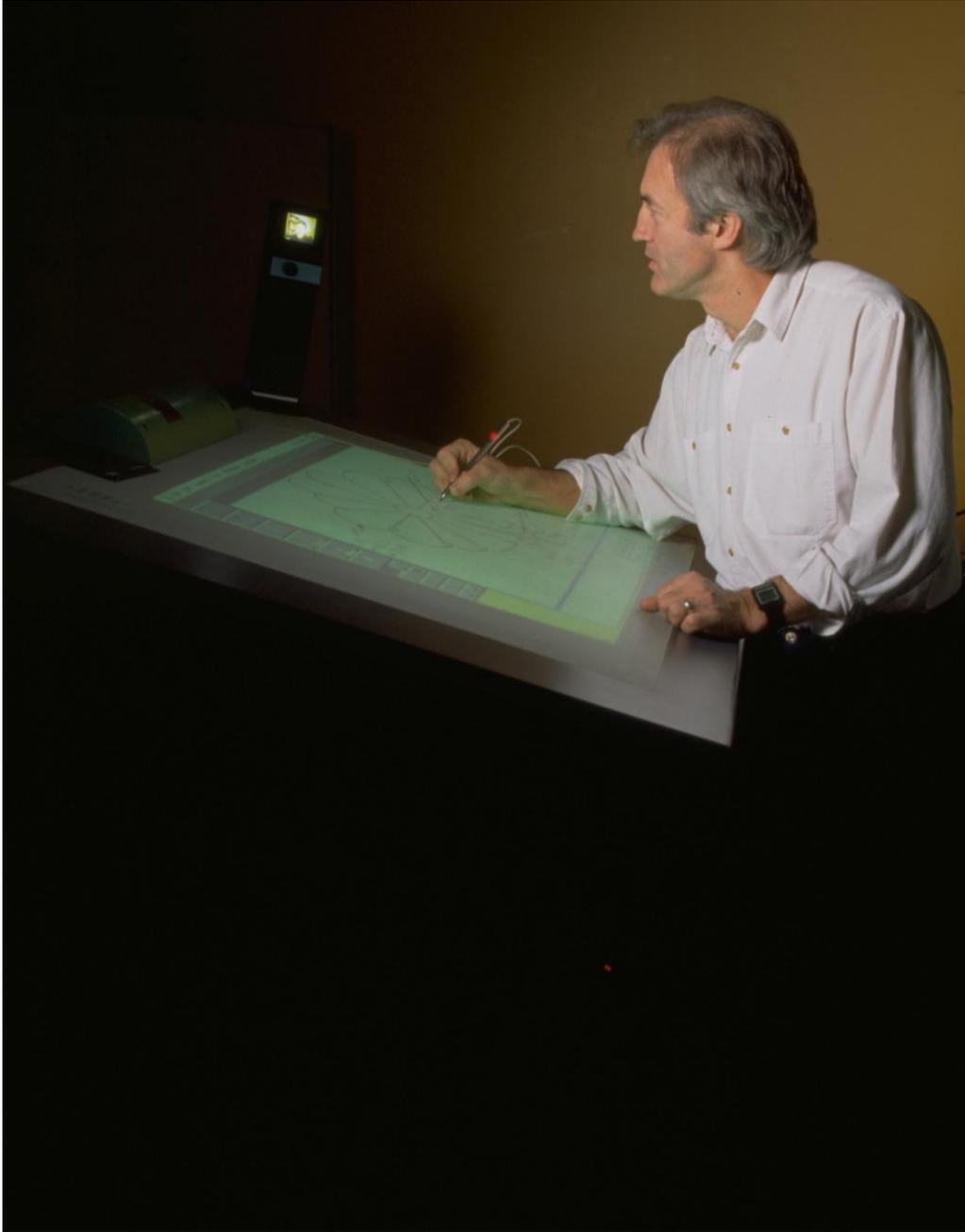


Figure 8: Like Figure 7, this image illustrates the use of distinct spaces for person and task space. While a Hydra device serves as a proxy for the remote participants(s) in both, the relative position of the person and task space are different, as is the posture of the task space. The configuration here affords interactions such as would occur when two or three colleagues were working on a design around a drafting table.

Conclusions

The effective support of multiparty meetings involving remote participants is a challenging design problem. Establishing a sense of co-presence, or *telepresence* of people, artifacts (such as documents), and spatial/gestural speech acts inherently implies a shared sense of place and space. Merely being able to see or hear remote participants or a shared document is just a prerequisite rather than the end goal of an effective system. The work described in this paper illustrates a number of everyday conversational and communicative acts which are poorly supported by existing technologies. For example, both the picture-in-a-picture and voice-switching techniques have been shown to be 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;
- perceive the group as a whole, in order to sense the "mood" of the group, for example;
- share documents and artifacts, and see co-participants in relation to those objects.

Some of the more unconventional approaches we have described provide much better support for these aspects of multiparty meetings, and as much as possible we have tried to evaluate and assess the extent to which they do so. We have also tried to document the particular design problems that still exist, and suggest how the designs might be improved. So far, we have found that the process of evaluation acts to inspire new design possibilities as much as it reveals design flaws.

The design space for multiparty video systems is rich and the issues are important. Our view is that in any such investigation, field trials and experiments with real subjects are critical. The dilemma is that to test, one needs a working system without making too much of an investment in a working system that has not been tested. Clearly, this is a case for iterative design and rapid prototyping, as we hope we have demonstrated in this chapter.

Acknowledgements

We would like to acknowledge the contribution of the Arnott Design Group of Toronto for the design and fabrication of the *Hydra models*. We also would like to thank Sara Bly, Paul Dourish, Bill Gaver, and Hiroshi Ishii for their helpful comments on

earlier drafts of this chapter. The work described has been supported by the Ontario Information Technology Research Centre (ITRC), the Natural Sciences and Engineering Research Council of Canada (NSERC), Xerox Palo Alto Research Center (PARC), Rank Xerox Research Centre, Cambridge, The Arnott Design Group, Toronto, Object Technology International, Ottawa, Digital Equipment Corp., Maynard, MA., and IBM Canada's Laboratory Centre for Advanced Studies, Toronto. This support is gratefully acknowledged.

References/Bibliography

- Argyle, M., Ingham, R., Alkena, F. & McCallin, M. (1973). The different functions of gaze. *Semiotica*, 7, 10-32.
- Borning, A. & Travers, M. (1991). Two approaches to casual interaction over computer and video networks. *Proceedings of CHI '91*, ACM Conference on Human Factors in Software, 13-19.
- Buxton, W. (1992). Telepresence: integrating shared task and person spaces. *Proceedings of Graphics Interface '92*, 123-129.
- Buxton, W. & Moran, T. (1990). EuroPARC's Integrated Interactive Intermedia Facility (iif): early experience, In S. Gibbs & A.A. Verrijn-Stuart (Eds.). *Multi-user interfaces and applications*, Proceedings of the IFIP WG 8.4 Conference on Multi-user Interfaces and Applications, Heraklion, Crete. Amsterdam: Elsevier Science Publishers B.V. (North-Holland), 11-34.
- Dourish, P. & Bly, S. (1992). Portholes: Supporting Awareness in a Distributed Work Group. *Proceedings of CHI '92*, 541- 547.
- Cherry, E.C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, 22,61-62.
- Cohen, M. & Ludwig, L. (1991). Multidimensional audio window management. *International Journal of Man-Machine Studies*, 34(3), 319-336.
- Dourish, P. (1991). Godard: a flexible architecture for A/V services in a Media Space. Unpublished manuscript, Rank Xerox EruroPARC, Cambridge.
- Elrod, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., Tang, J. & Welch, B. (1992). Liveboard: A Large interactive display supporting group meetings, presentations and remote collaboration, *Proceedings of CHI'92*, 599-607.
- Egan, J.P., Carterette, E.C., & Thwing, E.J. (1954). Some factors affecting multichannel listening, *Journal of the Acoustical Society of America*, 26, 774-782.
- Exline, R.V. (1971). Visual interaction: The glances of power and preference. In J. K. Cole (Ed.) *Nebraska Symposium on Motivation* Vol. 19, 163-206, University of Nebraska Press.

- Fields, C.I. (1983). Virtual space teleconference system. *United States Patent 4,400,724*, August 23, 1983.
- Gaver, W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K. & Buxton, W. (1991). Working Together in Media Space: CSCW Research at EuroPARC. *Proceedings of the Unicom Seminar on Computer Supported Cooperative Work: The Multimedia and Networking Paradigm*. London, England, 16-17 July.
- Ichikawa, Y., Okada, K., Jeong, G., Tanaka, S., & Matsushita, Y. (1995). MAJIC videoconferencing system: Experiments, evaluation, and improvement. In *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work (ECSCW '95)*, (Sept. 10-14, Stockholm, Sweden), H. Marmolin, Y. Sunblad, & K. Schmidt (Eds.). Dordrecht, Netherlands: Kluwer, 279-292.
- Ishii, H., Kobayashi, M., & Grudin, J. (1993). Integration of interpersonal space and shared workspace: ClearBoard design and experiments. *ACM Transactions on Information Systems (TOIS)*, Vol. 11, No. 4, pp. 349-375.
- Ludwig, L., Pincever, N. & Cohen, M. (1990). Extending the notion of a window system to audio. *IEEE Computer*, 23(8), 66-72.
- Mantei, M., Baecker, R., Sellen, A., Buxton, W., Milligan, T. & Welleman, B. (1991). Experiences in the use of a media space. *Proceedings of CHI '91, ACM Conference on Human Factors in Software*, 203-208.
- Okada, K., Maeda, F. Ichikawa, Y. & Matsushita, Y. (1994). Multiparty videoconferencing at virtual social distance: MAJIC design. In *Proceedings of CSCW '94*, (Oct. 22-26, Chapel Hill, NC), R. Furuta & C. Neuwirth (Eds.). New York: ACM Press, 385-394.
- Sheasby, M. C. (1995). *Brady Bunch and the LiveWire engine: Peripheral awareness in video teleconferencing*. M.Sc Thesis, Dept. of Computer Science, University of Toronto, June 1995.
- Sellen, A. (1992a). Speech patterns in video mediated conversations. *Proceedings of CHI '92*, Monterey, CA.
- Sellen, A. (1995). Remote conversations: The effects of mediating talk with technology. To appear in *Human-Computer Interaction*, Vol.10, No. 4.
- Sellen, A., Buxton, W. & Arnott, J. (1992). *Using spatial cues to improve desktop video conferencing*. 8-minute videotape. CHI '92.