3

Media Spaces: Environments for Informal Multimedia Interaction

WENDY E. MACKAY
Aarhus University

ABSTRACT

Distributed organizations, with distributed cooperative work, are a fact of life. How can new technologies help? Distributed video is an appealing choice, carrying more contextual information than voice alone and, arguably, better at conveying subtle cues, such as the emotional states. Although new commercial systems are being introduced, they focus primarily on providing new technology. Most are based on relatively simple extensions of two existing models of communication: formal meetings become videoconferences and telephones become videophones. However, research in computer-supported cooperative work has tried to emphasize the user, with models based on Shared Workspaces (to support shared work on a common task), Coordinated Communication (to support structured communication to serve a specified purpose), and Informal Interaction (to support informal, unplanned and unstructured interactions). Although mediaspaces can incorporate all three, they emphasize informal communication, providing people working together at a distance with interactions that they take for granted when they are co-located. This chapter describes some of the pioneering work in media spaces, with more detailed descriptions of our own work at Rank Xerox EuroPARC (RAVE for our own use in the laboratory and WAVE, to support engineers working collaboratively between facilities in England and the Netherlands), concluding with a discussion of the technical, user interface and social issues involved in designing media spaces.

3.1 INTRODUCTION

Telephones, faxes, electronic mail and the World Wide Web have transformed work, enabling people to work together, even when they live in different countries and in different time zones. Yet long-distance projects are still difficult, even when cultural and organizational differences

Computer Supported Cooperative Work, Edited by Beaudouin-Lafon

are taken into account. Why? One important reason is the lack of informal social contact that people have when they work in the same physical location [Hea91]. People who are co-located benefit from chance encounters in hallways or chats before and after meetings, resolving problems before they become critical. Working in the same physical environment helps people discover shared interests and develop a sense of community. Implicit knowledge about the state of each other's work can prevent misunderstandings or resentment: If I see that my colleague's report is sitting in her "out" basket, ready to send, I can avoid asking her about it and thus avoid offending her. When people are separated geographically, much of their informal knowledge about each other disappears and communication becomes much more formal. Attempts to address this with additional meetings and reports often serves to exacerbate the situation and emphasizes the differences between groups.

Moran and Anderson [Mor90] identify three fundamental approaches to supporting co-operative work at a distance: Shared Workspaces [Tan90, Min91, Ish92, Ols91a, Ols91b], which emphasize people working cooperatively on a common task, Coordinated Communication [Win89, Ell99], in which people communicate in a structured fashion for some purpose (such as decision-making), and Informal Interaction, in which people engage in unplanned and unstructured interactions. Chapters 4 and 5 in this book [Ish99, Pra99] address shared workspaces and Chapters 1 and 2 [Ehr99, Ell99] address coordinated communication. This chapter is most concerned with informal interaction, providing people working at a distance with the kinds of informal interactions they enjoy when working in close proximity to each other.

The explosion of networked computing through the World Wide Web and the decreasing cost of video technology have made distributed video a popular choice for addressing the problem of providing distributed social context. However, whenever new technology is created, it usually begins as an imitation of something that already exists. Not surprisingly, then, most commercial distributed video systems are modeled after one of two familiar forms of communication: telephone calls and formal meetings. Videophones are based on the model of a telephone call in which a caller establishes a video and audio link to a second party. When the call is initiated, the phone rings in the other location; if the other person is available, he or she decides whether or not to accept the call and complete the connection. The call continues as long as both parties participate; when one hangs up, the connection is broken. Videoconferencing is the other common model, usually involving specially-designed conference rooms. A common arrangement uses one video camera to capture people sitting at a table and a second, overhead camera, to capture documents or slides. Live video images are sent to one or more remote video conference rooms, via telephone or satellite and projected onto wall screens. Often, a separate speaker phone is used to enhance the quality of the voice. Desktop videoconferencing is a low-cost alternative, designed to be used with computers in the office. A small video camera is usually placed on top of the monitor and digitized images are sent to a window on another participant's screen. Some of these systems can handle several video images at once, although the computer monitor quickly runs out of screen space. Audio can be a problem if it is delayed and people often use telephones or speaker phones in addition to the on-line video.

Telephone and conferencing models represent a limited subset of the ways in which video can support distributed cooperative work. The purpose of this chapter is to describe the concept of a media space, which attempts to extend distributed video to include a variety of forms of communication, ranging from informal encounters and peripheral awareness to focused, formal meetings. The difference between media spaces and most commercial distributed video systems has little to do with technology and everything to do with the way in which the technology is embedded into the social environment. Understanding these social issues is essential for understanding how to design and introduce effective media spaces. As Moran and Anderson [Mor90] explain in their description of EuroPARC's RAVE media space:

> "EuroPARC's concern is not simply with artifacts and their enabling technologies, but with understanding the processes and relationships which such artifacts support, including the processes by which they are designed. The discipline of design must involve a constant movement back and forth between the design and use of technologies and reflection upon those activities."

3.2 EARLY MEDIA SPACES

The idea of a videophone has been around for a long time. In the early 1960s, AT&T demonstrated a prototype "PicturePhone" at the Seattle World's Fair, which allowed callers to view each other on small video monitors, set up in expanded telephone booths. The set-up optimized lighting conditions and video camera position to simulate face-to-face contact. (Note that callers could not actually call someone they knew; they had to wait for a stranger at another PicturePhone booth to arrive before having a conversation.) Although touted to be the phone of the future, it never really caught on and was ultimately deemed a failure [Nol92]. Another interesting experiment was the Hole-In-Space by video artists Galloway and Rabinowitz [Gal80]. They created a real-time video/audio connection between two sites in Los Angeles and New York. Pedestrians walking by could see full-size images of people walking in the corresponding location 3000 miles away. People not only stopped and stared, but often would respond to the remote conversation and begin talking to the passersby at the other end.

The term "Media Space" was coined by Stults and his colleagues at Xerox PARC [Stu86], who developed what was probably the first real media space. The cost of video had begun to drop in the 1980s, making it possible to link a laboratory in Palo Alto, California with a related laboratory in Portland, Oregon. Stults defined a media space as:

> "An electronic setting in which groups of people can work together, even when they are not resident in the same place or present at the same time. In a media space, people can create real-time visual and acoustic environments that span physically separate areas. They can also control the recording, accessing and replaying of images and sounds from those environments."

The Xerox Media Space was originally designed to model the informal types of communication that occur in hallways and in common areas, re-establishing the possibility of informal communication for people located apart from each other. The goal was to create a technology-supported analog to the mailroom or cafeteria; places where people naturally congregate informally and chat, with one conversation leading into another. An important aspect of this media space was that the connections were always there: only the people came and went. Conversations or meetings did not have a formal start or stop; they simply represented ongoing interactions among people. Subsequent media space research has emphasized the role of informal interaction as its key goal, although many have been extended to include facilities for shared workspaces and coordinated communication.

The period of the late 1980s and early 1990s was an active period in media space research. Several laboratories embarked on major long-term projects in which members of the laboratories both developed and lived in their media spaces. Although they share a number of characteristics, each media space was clearly shaped by the particular social and physical environments in which they were established and reflect different research goals. Some of the groups collaborated closely, with researchers moving back and forth between laboratories. In particular, Xerox PARC, EuroPARC, University of Toronto and Université de Paris-Sud shared researchers, software and numerous design discussions.

Bly et al [Bly93] describe their experiences with the Xerox PARC Media Space, including the three-year experiment linking their laboratory with Portland and the on-going evolution of the Media Space even after the Portland laboratory was closed. In the beginning, they used physical buttons to establish or replace video connections. Lab members experimented with different configurations and handled privacy in a very mechanical way, by turning off the microphone or turning the video camera towards the wall. They examined how their own social and work relationships changed as they used the Media Space and highlighted the need for additional research in user tailorability of the interface and support for managing privacy issues [Ols91b]. Over time, the Media Space was expanded from four offices, several public areas and the link to Portland, to include multiple offices in both sites and a variety of video devices attached to the network. This pioneering work at Xerox PARC influenced the development of the next major media space, created at PARC's sister lab in England, EuroPARC.

3.3 RAVE: EUROPARC'S MEDIA SPACE

Rank Xerox EuroPARC was founded in 1987 as a laboratory of Xerox PARC, located in Cambridge, England. The building, Ravenscroft House, has 27 rooms with five open areas spread over four floors. Each floor has two "pods" separated by a central stairwell, which causes a surprising degree of isolation among lab members. The layout of the lab simulates some of the problems people face when they must work together, but are physically separated. The lab decided to encourage cooperative work and foster social interaction by offering lab members ubiquitous audio, video and data interconnectivity within the building [Bux90]. The small size of the laboratory (approximately 30 staff and researchers) made it possible for everyone to have a media space node: everyone lived and worked in both the physical space and the media space. This global participation enabled EuroPARC to explore a variety of social as well as user interface and technical issues and provided insights into how to provide similar levels of social contact for people working together but at a distance.

RAVE (the Ravenscroft Audio Visual Environment) was not designed to replace face-to-face communication but rather to support work and social interactions, ranging from informal casual encounters to formal planned cooperative tasks [Gav92b]. RAVE was built with off-the-shelf analog audio and video technology, using several kilometers of coaxial cable to connect all analog devices to a computer-controlled 64×64 analog switch. This approach provided very high- quality video images and stereo sound; but was limited to a single building (extending it further would have been prohibitively expensive) and required a major recabling effort whenever nodes were moved. Figure 3.1 shows the basic set-up of EuroPARC's RAVE media space.

Each office and many of the common areas were equipped with media space "nodes" with a PAL video camera, a monitor, a microphone, a mixer to handle multiple audio inputs, stereo speakers and an optional foot pedal for controlling audio (Figure 3.2). Audio and video connections were managed from client applications running on either LISP machines or, later,

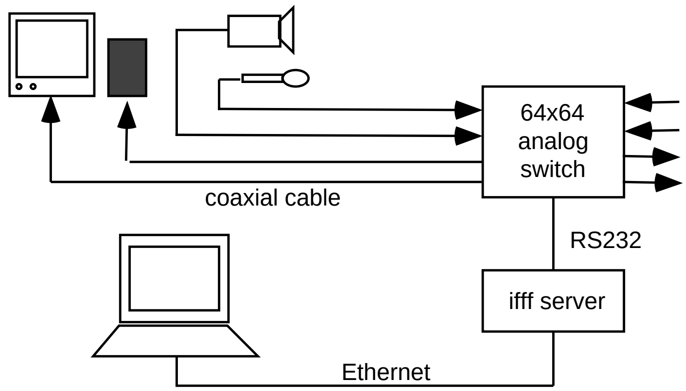

Figure 3.1 The RAVE media space consists of individual nodes connected via coaxial cable to a computer-controlled analog switch

Figure 3.2 A typical configuration for an analog media space node, with a video camera, microphone, video monitor and workstation

Unix workstations. Users had complete control over the position of the equipment, including location of cameras and microphones. They could turn equipment on or off, either electronically or physically (e.g. putting on a lens cap or unplugging it). Some nodes were equipped with additional video picture-in-picture (PIP) hardware, which permits simultaneous connections with up to four video sources on the same monitor. Connections to remote media spaces were created by connecting digital codecs to the analog switch via ISDN lines.

The *iiif* server [Bux90], running on Unix, controlled the analog switch and managed audio and video connections among media space nodes, as requested by client applications. In addition to allowing easy point-to-point connections within the building, *iiif* also guaranteed privacy and security. Different versions of the *iiif* server were used to set up media spaces at Xerox PARC, the University of Toronto and the Université de Paris-Sud, with individualized client applications designed according to the technical and social needs at each location.

Figure 3.3 Lab members can glance at each other or maintain long-term connections, such as this one between two administrative staff members

3.3.1 The RAVE User Interface

The RAVE user interface enabled users to display views from different video cameras located around the building and, if given permission, set up a two-way audio-video connection with any other node connected to the analog switch. Figure 3.3 shows one of the longest-running connections, between the reception desk on the first floor and the personal assistant to the director on the top floor. They established a permanent “Office Share”, with continuously-available video of the other person. Audio connections were made only when a foot pedal was pressed, to increase privacy. When the participants wanted to glance at others in the building or make other video connections, they did so directly and then returned to the default Office Share. Other members of the lab came to rely on their Office Share connection, using it to talk to each other and avoid running up and down the stairs. The effect was of a permanent “hole-in-space”, changing everyone’s psychological perception of the physical layout of the lab.

The user interface to RAVE evolved over the years, as users requested new functionality and when the entire lab shifted from a LISP-machine-based to a Unix-based environment. The original RAVE interface was based on user-tailorable on-screen buttons that accessed different functionality. Tailorability was particularly important, allowing users to explore different kinds of connectivity and express individual differences, ensuring everyone a choice in how they were represented within the media space.

The Buttons interface grew out of research at Xerox PARC [Hen86] and EuroPARC [Mac90]. Instead of typing commands or selecting from a menu, users could interact with an on-screen graphical object that ran relevant commands. They could look inside the button and tailor its functionality, as well as change its appearance, copy or even e-mail it to other users. Since buttons could be parameterized, users could change application-specific variables and edit the encapsulated code. This flexibility allowed lab members to explore the RAVE media space and develop the services that were most useful to them. The earliest buttons provided relatively low-level functionality, such as making or breaking a specific connection. Over

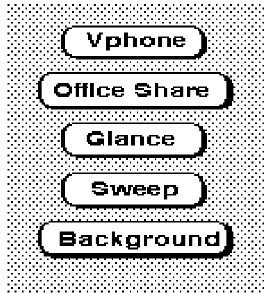

Figure 3.4 RAVE buttons provide users with services that reflect varying degrees of engagement

time, buttons evolved higher-level functions, providing encapsulated services, with built-in assumptions about handling issues such as privacy.

One of the most interesting features of RAVE is the ability to shift easily from peripheral to focused views. The five buttons shown in Figure 3.4 offer different levels of interaction based on the level of engagement required: from *Background* views operating at the periphery of attention, to the unobtrusive presence of a *Sweep* through the building or an informal *Glance*, to the shared awareness enjoyed by participants in an *Office Share*, to the full engagement of a *Vphone* conversation.

Vphone: A highly-focused form of interaction with two-way audio and video connections. Like telephone calls, one party must explicitly initiate the call and the other must explicitly accept the connection. The call ends when one party hangs up. Participants used this when they wanted to discuss a specific topic.

Office Share: The physical connection is technically identical to a Vphone call; the difference is that the participants decide on the connection in advance, after which they do not explicitly initiate or terminate individual calls. Participants choose whether or not to include permanent audio as well as video connections. Office Shares facilitated a range of communication, from passive awareness to highly-focused interaction. Long-term Office Share users, with connections lasting for months or even years, claimed it was like sharing a physical office without many of the annoyances.

Glance: A brief (three-second) one-way video connection to another node. The person being glanced at first hears an audio cue (or the name of the person), then the connection is established, after which another audio clue indicates that the Glance is complete. Lab members define in advance who has permission to glance at them. Glances provided a quick and unobtrusive method of determining whether or not someone was around or currently busy, similar to glancing into someone's physical office.

Sweep: A brief (one-second) one-way view of a series of pre-authorized nodes, local or distant. Users could customize their sweep patterns to include the most useful public (and, if authorized, private) nodes, in order to find out who was around and generally what was going on in the lab.

Background: A long-term view of a particular location that acts as the default view of the media space. Technically, Background is indistinguishable from an Office Share. However, most Background connections are of public areas that do not require specific permission, unlike the pre-arranged connections to a particular office. Although the view from the roof was popular with people in windowless offices; the most popular Background was the EuroPARC

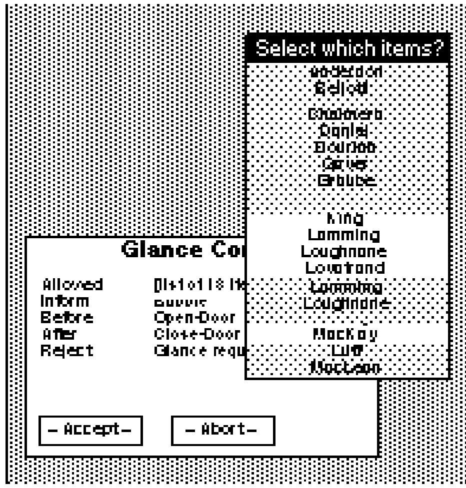

Figure 3.5 Users can select who has the right to glance at them through the RAVE media space

commons. Since everyone in the lab visited the commons regularly to check mail or get a cup of coffee, it provided everyone with a low-level consciousness of who was in the building and when informal gatherings (such as afternoon tea) had started.

Godard [Dou91a] defined and controlled these and other services, directing all connections made by the *iiif* server. Based on preferences expressed by users (such as who had permission to glance at them), Godard would determine whether or not to perform specific requests for services, requesting additional input when necessary. Before performing an authorized connection, Godard would record the previous connection (such as an Office Share or Vphone call), ensuring that the user would return to the correct state after the service was performed.

One of the primary benefits of this architecture is that participants can interact with RAVE in terms of high-level services rather than low-level physical connections. Each service includes a representation of the user's intentions and makes it possible to embed information used to protect privacy. Participants can decide in advance who has the right to perform each service, as opposed to making on-going decisions about low-level connections, which helps to balance the tradeoff between privacy and access. Figure 3.5 shows an example of a service control panel. In this case, the user looks at a list of lab members and indicates those who have the right to establish a Glance connection.

In practice, the default settings of such lists are very important. If the default is that everyone has the right to glance unless explicitly deleted from the list, a refused Glance request can be taken as an explicit, personal rejection. However, if the default is that no one has the right to glance unless specifically added to the list (as at EuroPARC), then refused Glance requests may simply indicate a non-updated Glance list, which is less likely to be viewed as an insult. This poses particular problems for new members of the lab. Since people rarely update their Glance lists, even during project and group changes, newcomers will find themselves unable to glance at many of their colleagues and feel excluded from the media space. The result is that long-term members may have much greater access to the media space than newer members, often without realizing it. Seemingly-innocuous decisions about default settings may have long-term effects on the ultimate acceptance of a media space within an organization.

3.3.2 Notification with Auditory Cues

Lab members not only wanted control over the types of connections they made, but also wanted feedback about when and what kinds of connections were in use. Because Godard had some understanding of the participants' intentions, it could not only distinguish among physically identical connections such as Vphone, Office Share and Background, but it could also determine and deliver the appropriate feedback. Users could request a variety of notification types, such as presenting a message on their workstations. However, the most popular notifications were more subtle. Gaver's work on auditory icons [Gav86] and the affordances of audio [Gav91b] was incorporated into RAVE, providing real-world auditory cues that indicated what was going on. For example, when someone glanced at another person, Godard triggered a sound (the default was that of a door opening) as the connection was being made. Three seconds later, when the Glance was terminated, another sound was triggered (usually a door closing). Other sounds were associated with other kinds of connections, indicating the corresponding intent. A knock or telephone ring signaled a Vphone, footsteps indicated a sweep and a camera whir flagged when a single-frame snapshot had been taken. Gaver [Gav92a] explains why real-world sounds are particularly effective:

- Sound indicates the connection state without requiring symmetry; providing information without being intrusive.
- Sounds do not require the kind of spatial attention that a written notification would.
- Non-speech audio cues often seem less distracting and more efficient than speech or music (although speech can provide different sorts of information, e.g., who is connecting).
- Sounds can be acoustically shaped to reduce annoyance [Pat89]. Most sounds involve a gradual increase in loudness to avoid startling listeners.
- Finally, caricatures of naturally-occurring sounds are an intuitive way to present information. The sound of an opening and closing door reflects and reinforces the metaphor of a glance and is thus easily learned and remembered.

A number of other researchers have explored the role of audio in distributed collaborative work and media space settings. Researchers at PARC [Mor97] and MIT [Kob97] explored the problem of browsing for audio data. Seligmann et al [Sel95] examined the cues that people use to understand ordinary telephone calls and then looked at the more complex information needs required in multi-party, multimedia conversations. They found that needs for "assurance" became more complex and that users needed information about connectivity, presence, focus and ways of distinguishing between real and virtual activities.

Godard, with auditory cues, provided control, feedback and intentionality, three prerequisites for privacy, at very little cost in terms of intrusiveness. Godard used a system called Khronika [Lov91] to handle auditory events. Khronika is an "event notification service" that supports selective awareness of planned and electronic events, announcing when a video connection has been made, reminding people about upcoming meetings, providing information about visitors and even gathering people to go to the pub.

Khronika was based on three fundamental entities: *events, daemons and notifications*, as shown in Figure 3.6. Events were organized within a class hierarchy, each with a class, start time and duration. Specific events included meetings, visitors, arriving e-mail and RAVE Glances. Events could also be manipulated as more abstract classes, such as "professional", "electronic" and "entertainment". Event daemons produced notification events when they detected specified event types. Users could constrain daemons, enabling them to select only the

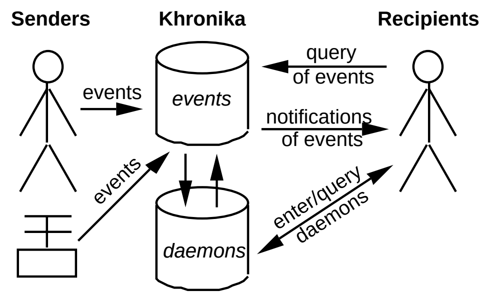

Figure 3.6 Khronika maintains an event database entered both by people and on-line systems. Daemons watch for specified events and post notifications when they are detected

events they wished to hear about. For example, a user could create a daemon to look for all EuroCODE project meetings in the conference room and generate a notification five minutes prior to the meeting. Users could also specify different notifications, such as an e-mail message a day prior to an important meeting, with a pop-up screen message an hour before and a synthesized speech message five minutes before the start time.

General-purpose non-speech audio cues were also popular. For example, the sound indicating the start of a meeting began with sound of people murmuring at low-volume, followed by a gavel sound to indicate the precise start time. Such sounds provided low-level peripheral awareness of events, enabling people to shift their attention to them when necessary and ignore them otherwise. Gaver et al [Gav91a] found that non-speech audio feedback changed both participants perception of the system and their tendency to collaborate while using it.

A button interface to Khronika let users browse the event database as well as create new events and daemons. Another interface, *xkhbrowser*, showed a calendar with events spanning different periods of time. Events could be displayed according to their level of specificity, enabling users to quickly view the kinds of events they were interested in. We compared the use of two EuroPARC calendar systems [Dou93]: Khronika's and the paper-based system managed by the administrative staff. We discovered that even though the information was ostensibly identical, users were influenced by their knowledge of the source of the information, with correspondingly different levels of trust in different kinds of information. If there was a conflict between the two systems, users would try to determine which calendar system was more likely to be accurate, given the particular piece of information. (For example, the person in charge of the brown bag lunch seminar series was known to input the events into Khronika, so users tended to believe the Khronika-based information. In contrast, the administrative staff would track people's travel schedules, thus the paper-based calendar was assumed to be more accurate for that kind of information.) Thus the social context played as important a role as technological functionality in determining how users interacted with each system.

3.3.3 Long-Distance Awareness

EuroPARC maintained a close association with her sister laboratory in the United States, Xerox PARC. Maintaining live video connections was too expensive, so members of the lab investigated different ways of linking the two media spaces.

Polyscope [Bor91] distributed low-resolution (200×150 bits) digitized images captured approximately every five minutes from each media space (assuming the owner had given

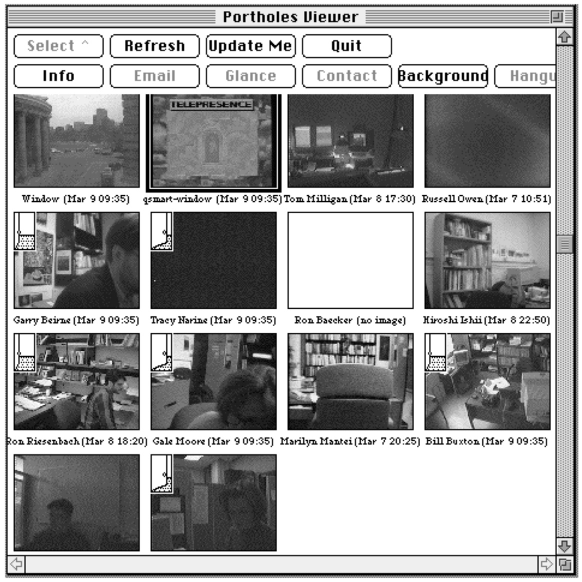

Figure 3.7 Portholes presents a collection of recently-captured still images from the media spaces at EuroPARC and Xerox PARC

permission). A simple animation facility looped several images in sequence, to provide a jerky, but usually effective, sense of motion and a way to disambiguate scenes. The button interface allowed users to select an image and immediately establish a Glance or Vphone connection.

Portholes [Dou91b] was developed to help lab members stay in touch by sharing more-frequently captured still images from the respective media spaces (Figure 3.7).

Both Polyscope and Portholes allowed multiple nodes from several remote locations to be presented simultaneously, providing passive awareness of distributed workgroups without making explicit video connections. They offered spatially-distributed but asynchronous functionality, which complemented the synchronous but single-channeled video services from each media space.

3.3.4 Observations of Use

Users define the social protocols surrounding the use of their electronic communications. For example, people try to avoid annoying each other. In the early days of electronic mail, outsiders were surprised that people working physically next to each other would still send each other electronic mail. This was not due to a preference for technology-based rather than human-based communication; it was simply a matter of courtesy. Calling over the wall or telephoning is an interruption, whereas sending e-mail allows the recipient to respond at their convenience.

Similar social protocols developed in response to the media space. Over time, media space users began to change it, creating uses not originally anticipated by the system's creators. Subtle characteristics in the technology suggested new uses or resulted in changes in how the system came to be understood [Suc91]. An interesting example was the use of the RAVE media space for "projecting" presence, taking advantage of the knowledge that members of the lab had a shared peripheral awareness of the EuroPARC commons. People usually waited until a critical mass of people had assembled in the commons before appearing in person. People thus showed up in two waves: the first few arrived. When three or four were there, suddenly everyone else appeared. People took advantage of their ability to work until the last minute (and avoided "wasting time") and were still assured of arriving when the meeting actually started.

Sitting in the commons, in view of the camera, is a way for a researcher to broadcast his or her availability, letting colleagues know that it is acceptable to come up and chat. For example, I had about an hour before I needed to catch a plane and five people I needed to talk to before I left. I decided to sit in the commons and found that, one-by-one, all five came to talk. Each person monitored my meetings on their monitors and came up when they could see I was ready to talk to someone else. Everyone coordinated their activities, managing to find appropriate times to meet, without wasting their time waiting for me or wasting my time waiting for them [Mac92].

Interestingly, when members of the administrative staff sat in the commons, their message was the opposite. They were on a break and it was not acceptable to ask them to do something (although it was fine to have a chat). The commons had an explicitly "video free" section, so that people who wanted to avoid being seen could do so easily and naturally. The overall effect was that of having the common area right outside one's door, but without the noise. Whenever the link to the commons broke down, such as when the equipment was being upgraded or there was work on the building, members of the lab reported the sense of being slightly disoriented and feeling out of touch.

Sharing the same physical office with someone can be annoying, especially if you have different tastes in music or are prone to talking all the time. Office Shares, particularly for people who worked late at night, proved to be extremely comforting. Without listening to each other, we could still sense each other's presence (and when the other person was ready for a break). To a somewhat lesser extent, the Portholes connection to PARC provided a similar sense of comfort, since someone was always there no matter how late the EuroPARC crowd worked (given the nine-hour time change from England to California).

Heath and Luff [Hea91] observed lab members using long-term RAVE connections and found that video sometimes undermined the effectiveness of subtle communicative gestures. For example, since the camera and monitor are offset, a person looking at the monitor will appear to be looking down slightly when displayed on the other monitor. Experienced media space users learned to shift their gaze back and forth between the "natural" view for them (i.e. looking at the person on the monitor) and the "effective" view (i.e. creating the appearance of eye-contact by looking directly into the camera). Visitors entering an office with an Office Share would sometimes be confused, thinking the person on the other end of the media space connection was looking at them when they were in fact looking elsewhere.

Gaver et al [Gav93] examined the effect of giving users a choice among four different views, rather than the usual single-camera, face-to-face view. The additional views included an "in context" view, showing people and objects in relationship to their workspace, a "desk view", using a high-resolution monochrome camera to view documents and either a "dollhouse" view specific to the experimental task, or a "birds-eye" view showing most of the room. They found that face-to-face views were rarely used when people were actively involved in a collaborative task; the exception was when participants engaged in negotiation about task strategy. This study suggests that the camera setup for media spaces should change when users want to engage in collaborative tasks.

3.4 OTHER MAJOR MEDIA SPACES

Several other media spaces were developed in the same time frame as RAVE. The US West Advanced Technologies Telecollaboration project [Bul89] was similar to the PARC/Portland link, supporting a small group of people sharing several projects who were located in Denver and Boulder, Colorado. The media space included several offices, a conference room and public areas at both sites. Users could “call” to get a private office-to-office audio/visual connection, “look around” to get a video-only connection, and “videoconference” to support multiple participants in the conference rooms. Public areas were continuously connected, as at PARC.

BellCORE was also very active in media space research at this time, creating Cruiser [Roo88] and the VideoWindow [Fis90], both controlled by a system called Rendezvous [Hil94]. In contrast to the Xerox PARC approach, which emphasized letting the users evolve the characteristics of the media space, BellCORE researchers followed a theoretical approach, focusing on the role of informal communication [Kra88, Fis93]. Cruiser was based on the model of walking down a hallway and glancing into open offices to see who was there. All connections were reciprocal, in that the person doing the glancing was always seen by the person being glanced at. Participants could control the access to their images and could also establish two-way connections in the course of a “cruise”. Cruiser was designed to encourage spontaneous, informal communication, but often resulted in longer-term Office Share. Cooperstock [Coo92] reports on the iterative design of Cruiser, describing how users and designers influenced the design of the system over four iterations.

VideoWindow was more similar to the PARC/Portland Media Space link, with two large-screen displays located in two public areas on different floors of the research building. The link was available continuously for three months and was designed to support informal communication among the 50 researchers and staff in the area. People would arrive to get their mail or have a cup of coffee and engage in conversation with the people physically present as well as the people located at a distance.

The Montage system from SunSoft [Isa93] explored how to use video to help members of distributed groups develop a sense of “teleproximity”, helping collaborators find opportune times to interact with each other by using reciprocal glances to “peek into someone’s office”. The system also provided access to an on-line calendar, e-mail and on-screen note facility. As with Cruiser, video connections appeared in a small window on the computer screen, rather than on a separate monitor. Researchers found that most glances did not result in interactive communication [Tan94]. Issacs [Isa95] reports on their experiences using video in a broadcast setting, as opposed to in a smaller forum with a local audience. They found that the audiences preferred watching the multimedia presentations by speakers in the broadcast setting, while the speakers themselves preferred the intimacy of local talks.

The University of Toronto engaged in two major media space projects, in collaboration with researchers at Xerox PARC and EuroPARC. CAVECAT [Man91, Gal91] was implemented with software from EuroPARC’s RAVE system and supported approximately ten offices within a single building. In addition to the media space features from RAVE, researchers explored the problem of integrating shared drawing facilities with shared presence [Pos92]. Sellen [Sel92a] studied the speech patterns in video-mediated conversations. She

Figure 3.8 Hydra incorporates a video camera, monitor and directional microphone into a small table-top unit that can act as a proxy for distant participants of a meeting

compared same-room and video-mediated conversations, using two interfaces, including a system called Hydra [Sel92b]. Figure 3.8 shows three Hydra units, each with a small video camera, monitor and directional microphone, which act as proxies for distant meeting participants. Real and video-mediated participants could react to each other as they would if all were co-present in the same room. Sellen found quantitative differences between face-to-face conversations and the two video-mediated interfaces. Although there were no significant quantitative differences between the two video-mediated interfaces, there were significant qualitative differences, with users preferring the Hydra interface. Olson and Olson [Ols95] studied the role of adding video to remote-collaborations. They found that users of audio-only connections had more difficulty communicating but that there were no basic differences between the quality of work performed in face-to-face settings and in settings with both high-quality audio and high-quality video.

The University of Toronto's follow-on project, called Telepresence, experimented with media spaces outside of a laboratory setting and included studies of the ways in which media spaces changed the social relationships among people working at a distance [Har94]. Both Toronto media spaces used an icon of a door, displayed in various states to indicate the user's level of accessibility (Figure 3.9). A fully-open door indicated that anyone could make a full two-way audio/video connection, whereas a door ajar enabled people to glance, but required a ring or further interaction from the user in order to make a full two-way audio/video connection. When the door was shut, glances were not authorized and further interaction was required to establish a two-way audio/video connection. When the door was locked, no video connections were possible. Yamaashi et al [Yam96] describe another extension in which users were given two views: a wide-angle view to show the context of the office and a more detailed shot linked seamlessly together. They also explored the use of sensors placed in the physical environment to provide contextual cues to remote users of the media space. For example, the state of the physical door to the office (i.e. open, ajar, closed or locked) controlled the state of the on-screen door icon.

A later Xerox PARC media space, called Kasmer [Bly93], was created to support a much

Figure 3.9 The original doors interface allowed users to select from four different door states: open, ajar, closed and locked

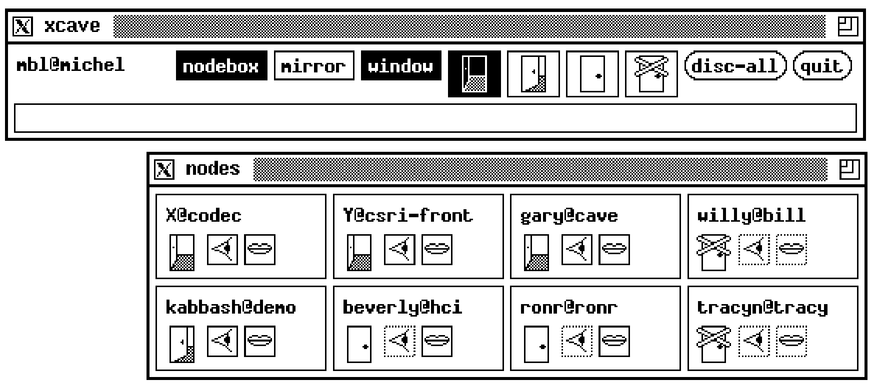

Figure 3.10 The interface to the Kasmer media space. The upper control panel allows users to select services. The lower display panel shows other users and their current level of accessibility

larger group of people in different groups within the laboratory, as well as offering codec links to external sites (Figure 3.10) . The underlying software was borrowed from CAVECAT, University of Toronto's system, and RAVE, from EuroPARC. The system was designed to balance frequent, easy communication within groups, while also providing less frequent communication to distant or external groups. Each working group included 10–25 participants and media space nodes; each with their own social conventions and models of use. Adler and Henderson [Adl94] describe their experiences with a 9-month Office Share connection within this environment. Mynatt et al [Myn97] explore the differences between on-line and physical space, arguing that media spaces reinterpret physical space through the positioning of audio and video elements and argue that activities derived from one space do not translate well into other spaces.

Another extension of the media space work is on-going at the Université de Paris-Sud, based on work the authors did at EuroPARC, PARC and University of Toronto. Mediascape

Figure 3.11 The Mediascape media space, with the standard user interface (left) and an electronic Post-It note (right)

[Rou98] was originally built with the *iiif* server, but has since been reimplemented with a set of custom HTTP servers, enabling users to embed access to the media space into any document on the World Wide Web. Unlike the numerous “webcams” that have appeared, which show a single view from a fixed camera, Mediascape is a full media space, with facilities for managing connections among multiple sites, notifying users and controlling access. The user interface is highly customizable since it is a plain HTML document. The standard interface is shown in Figure 3.11 (left). Images are updated every few minutes as in Portholes. Passing the cursor over an image initiates a Glance. A double-click establishes a Vphone connection, according to both users’ expressed availability. Mediascape uses the same door metaphor as CAVECAT. In the figure, one user has locked his door (bottom left) and another has left his door ajar (bottom right). Additional services include Post-It, to leave a message on someone’s computer screen, Grab, to grab a still image from the media space, and Dvideo to send pre-recorded or live digitized video. Electronic Post-It notes are implemented by remotely controlling the recipient’s Web browser (in this case Netscape Navigator). The image in the note has the same capabilities as in the interface: it updates every few minutes and can be used to establish other media space connections.

3.4.1 Building Upon the RAVE Media Space

The RAVE media space provided an infrastructure for other research projects as well. As part of a European ESPRIT project, called EuroCODE [Mac95a, Mac98], we were responsible for designing a multimedia communication system for engineers building a bridge across Denmark’s Great Belt (Storebaelt) Waterway. This project developed a radically-different interface to the media space: a paper engineering drawing. We developed Ariel (Figure 3.12) which detects individual engineering drawings via their barcodes. In the prototype shown, the drawing was placed over a large (A0 size) graphics tablet. A projector on the ceiling projected menus and other computer-generated information, including media space images, onto the paper. Here, the user is establishing a Glance connection with the author of the drawing, in order to discuss possible changes. The user could also associate any audio, video or text information with any part of the drawing and capture handwritten notes which could be sent to colleagues.

Figure 3.12 Ariel lets construction engineers access the media space and a hypermedia annotation system via paper engineering drawings. The user selects the media space option from the control section of the paper engineering drawing (upper left). Ariel projects a menu and the user selects Glance with the graphics tablet pen, which establishes a three-second connection

3.5 WAVE: A DETAILED CASE STUDY

WAVE [Pag93] was an attempt to test the media space concepts learned from RAVE in a real-world setting: a distributed product development organization within a large multinational corporation. WAVE differed in several important respects from the RAVE media space. RAVE existed within a single building and was designed to encourage communication among people who had other forms of informal and formal communication available to them. WAVE was more similar to the original PARC Media Space and Portholes, in that the participants were distributed in both space and time. (However, Portholes had to span eight, sometimes nine, time zones, whereas WAVE involved a one-hour difference between England and the Netherlands.)

RAVE was able to take advantage of point-to-point analog video connections, with an analog video switch and kilometers of coaxial cable, to provide high-quality images and sound with no delay. Portholes was restricted to still images displayed on a computer screen, with occasional dial-up links with a low-resolution video link and a rather annoying audio delay. The distances between the WAVE sites caused us to consider different technical solutions for distributing video, which had a corresponding impact on the user interface and social use of the media spaces. One research goal was to find out the acceptable thresholds for video quality under various media space conditions, given the bandwidth and cost constraints of international long-distance links. Another critical difference from RAVE and Portholes was that the participants were not researchers, but engineers creating a product. For them, the media space was like a telephone or fax, a technology to be used only if it supported the work at hand.

We studied an engineering design center in England, which took designs created in Japan or the United States and localized them for the European market. The organization was subject to the typical pressures of any high-tech company in a highly innovative, competitive and turbulent market: They had to increase customer satisfaction, maintain a technological edge and improve quality while decreasing costs. Product development had to reduce time-to-market, streamline processes, adapt to rapid technological change, while making efficient use of resources.

The organization had a sophisticated telecommunications infrastructure: a corporate telephone system based on leased lines (users need only dial an extension to reach other sites); voice conference calls; answering machines; beepers; electronic mail; fax; and sophisticated (and expensive) satellite videoconference facilities. All engineers and administrative staff had either a workstation or a computer terminal on their desks. Yet, in spite of this infrastructure, engineers spent a great deal of their time traveling; travel accounted for over 10% of the product development budget. At the time of our study, the travel budget was being cut and managers were interested in finding ways to reduce the need for face-to-face meetings. We met with the director of the division and interviewed all of his senior managers and many of their staff, who often took us on tours of their work areas after the interview. Interviews were generally open-ended, although each began with a set of standard questions, including the person's role in the organization, a description of his or her work (either a project or a function), as well as any communication breakdowns and strategies for addressing them. We also attended regularly-scheduled live and video-mediated meetings [Pag93].

We chose a major product development project in a critical stage within its two-year life cycle. The product was designed in England and was being assembled in a factory in the Netherlands, requiring complex communication and coordination between the two sites. The English engineers understood the product design, whereas the Dutch engineers understood the manufacturing problems and maintained the relations with the local suppliers. We identified two situations with serious coordination and communication needs: cooperation between design and manufacturing engineers and configuration management. After further analysis of their work patterns, we installed two media space connections: a dial-up video phone between the desktop of an engineer in England and the shop floor of the factory in the Netherlands, and an Office Share between the desktops of two people sharing administrative tasks across the two sites.

3.5.1 Analyzing the Existing Videoconference System

Early forecasts of the success of videoconferencing and video telephony were wildly optimistic. Egido [Egi88], in her analysis of why videoconferencing systems fail, cites an early 1970s prediction that a full 85% of meetings would be conducted by videoconferences by the end of the decade. Yet, videoconferencing has been slow to be accepted, despite major financial investments by corporations. Because people easily equate media spaces with videoconferencing, we were interested in how the people in the organization we studied felt about their existing satellite-based videoconferencing system.

Each site (in Europe and the United States) had a special meeting room set up to accommodate six people at a table, with two video cameras to capture each group of three. A ceiling-mounted camera was used to transmit images of objects or documents. Images of colleagues at remote sites were projected on two large video monitors opposite the table, with images of documents presented on a third monitor in the middle. One user likened it to being on a television quiz show, with opposing teams lined up, facing each other.

The videoconference room was designed for highly-stressful project checkpoint meetings, which had priority over everything else. Others could schedule meetings when the room was not already booked, usually to address critical problems that arose. Such meetings generally involved technical people who made use of the ceiling cameras to discuss design documents. (They had no facilities for bringing hardware prototypes into the meeting room, nor did they have facilities for sharing electronic documents.)

Like Egido, we found mixed, mostly negative reactions to the videoconferencing system. High-status managers were most likely to find it useful: they controlled the meetings and appreciated the reduction in travel costs. For example, during a period in which cross-Atlantic travel was eliminated, one manager said "It really came into its own during the Gulf war; [its] use has really increased since then." Another found it "good for sharing problems and project status; [although not] for general information exchange".

Interestingly, most others found it to be divisive, increasing the adversarial nature of the relationship among the participants. These users were individual contributors who used the system to negotiate issues and solve problems. Several people described their concerns as follows: "There is lots of friction. If people [already] have positions, being able to see them doesn't help to bridge the gap. You see a panel of people; it's a stand-off situation. It encourages antagonism" and it "is not good for problem solving . . . you react differently to body language on it versus face-to-face. The etiquette changes." One person described his weekly Friday meetings: "It would end the week horribly . . . it was pretty bloody. Emotions fly across the airwaves."

In summary, most people viewed face-to-face meetings as the optimal form of communication. Telephones were useful, but only for certain kinds of communication. The video conference system was viewed as useful by upper-level management, but created adversarial relationships among the participants. We were interested in whether or not a media space, with its emphasis on informal interaction, would provide better communication and reduce the adversarial quality of the interactions found with the videoconference system.

3.5.2 Design Center – Shop Floor Link

Since stopping the production line was very expensive, the engineers had a basic rule of thumb: if a problem arose that they did not think could be solved with telephone or fax in less than four hours, the design engineer got on a plane and flew to the Netherlands. The media space was seen as a way to reduce the latter.

One end of the link was on the desktop of a system integration engineer in England, responsible for ensuring that all sub-systems worked together. He knew most of the designers on the project and could quickly contact the appropriate person whenever a problem arose on the shop floor. The other end of the link was the shop floor itself. We installed equipment on a cart which could be moved to any part of the manufacturing line. We used two codecs, based on the H.272 standard (designed for desktop videoconferencing using public ISDN networks), connected by a 64 Kb/s data line (Figure 3.13). The codec in England was connected to an ISDN telephone via an X.21 interface; the ISDN telephone was used for dialing and for displaying line status messages. Unfortunately, the Netherlands did not have ISDN available at the time, so we used a switched 64 Kb/s IDN line. This made the set-up on the Dutch side a bit more complicated. We connected a 64 Kb/s modem to an X.21 controller, which was connected to the codec and to a VT100 terminal. In order to dial and disconnect the line, the Dutch users had to type several commands on the VT100 terminal.

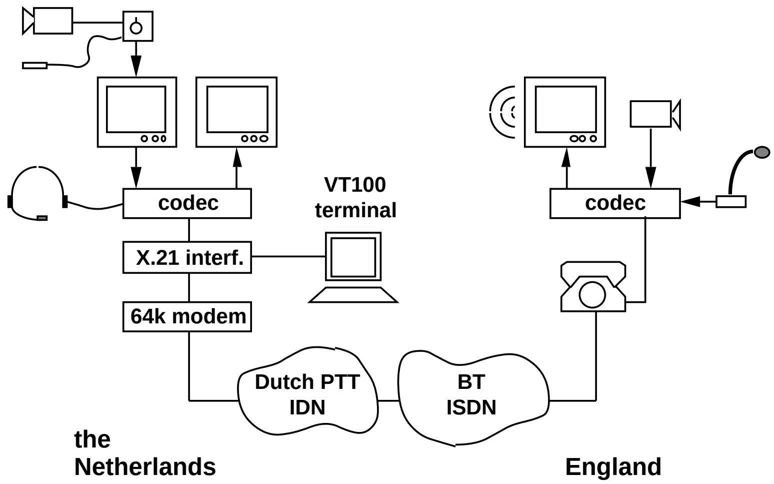

Figure 3.13 WAVE link between the design center in England and the factory in the Netherlands

Figure 3.14 Image captured from the WAVE link. Note the use of the small hand-held camera to show a close-up view of the problem

In England, we installed a single camera, a monitor with built-in speakers and a directional microphone. We also installed a videotape recorder to capture video from the Netherlands and audio in both directions. On the Dutch side, we installed two cameras: a standard-sized camera was clamped to the cart to provide an overview, and a miniature camera (1cm$\times$5cm) with a flexible cable to show small details (i.e. 2mm size type). We also installed two monitors, for incoming and outgoing video. Since there was a great deal of background noise on the shop floor, we provided headphones with a built-in microphone.

The link on the shop floor was up for two weeks; during this period, we spent two days at each site sitting next to the equipment and observing what users did. For the remaining time, we collected videotapes of the video and audio going through the link and later interviewed the people who used the link (Figure 3.14). Users complained about the poor audio quality, poor video resolution and the lack of reliability. Yet they were elated when they were able to solve problems without traveling. As one manager said, “This technology is a pain . . . However yesterday we used it and it saved us a trip, therefore I will let my people spend more time on it. I would rather have this lousy link than nothing. There are many things here which are not perfect that we have to cope with, this will be another one.” The following examples show how they used the system:

Show a problem: A manufacturing engineer shows something going wrong on the production line and asks designers for explanations, solutions, or changes. Video is very important, not only to improve communication, but also to overcome the initial skepticism and mistrust. Looking at the problem together and jointly working on the solution helps to overcome cultural differences (not only between Dutch and English, but also between manufacturing and design roles), fostering a cooperative attitude towards solving the problem and “getting things done”, rather than arguing an abstract problem over the phone to “pass the ball”.

For example, following a part change, the packaging also had to be changed, making a new cut into the cardboard and assembling the pieces so they would fit together. A packaging engineer was able to use the video link to show, step-by-step, how to make the cut and assemble the pieces, while the Dutch engineers repeated each step. In half an hour, a problem which would have required a trip was solved. The engineers were enthusiastic about the video link, saying that it had been particularly useful to do each step of the assembly at both sites, watching each other’s actions.

Show a solution: A design engineer demonstrates the correct way of doing something on the shop floor. The video link allows the engineers to go through each step of the process, performing it simultaneously at both sites to ensure that everyone understands the solution and its consequences. The video link is much faster and more direct; allowing participants at both sites to see the solution, which increases confidence in the solution and trust between the two sites.

For example, a manufacturing engineer showed a programmer in England a software bug by pointing the camera at the display and keyboard of the product so that the programmer could see what was going wrong. The programmer told the manufacturing engineer which keys to press and they found another, related bug. The problem might have been described over the telephone, but the manufacturing engineer felt that the programmer was skeptical about the bug and assumed the manufacturing engineer was doing something wrong. Using the video link allowed him to actually see the problem and try things out; it also allowed him to locate the bug precisely.

Cooperative problem-solving: A problem is shown and engineers at both ends of the link brainstorm solutions, discuss ideas, point at causes and try out experiments on the machine. Unlike a standard videoconference, the participants rarely look at each other’s faces and concentrate on the technical problem to be solved.

For example, the paper feed mechanism worked well on the prototypes, but did not work reliably on the units coming off the manufacturing line. Although it seemed to be a manufacturing problem, the manufacturing engineers were interested in suggestions from the designers. Six Dutch engineers showed the problem to three English engineers, who were able to brainstorm and test various solutions. Because there was only one set of headphones on the shop floor, the Dutch engineers passed around the miniature camera to show things and let one person handle the audio communication to England.

The management made a cost/benefit evaluation of effectiveness of the link and concluded that the system was useful for solving problems on the shop floor and that the link saved at least two trips during the course of the experiment. They were not happy with the unreliability of the link, but said, “when it works and is used for the right application, it is a very powerful tool”.

3.5.3 Configuration Management Link

The second media space link was installed between two planner analysts, located in England and the Netherlands. The planner analysts were responsable for configuration management, including tracking design changes, evaluating the cost of parts and changes and maintaining the inventory of the thousands of parts which make up a product acting as a bottleneck for all design changes. They registered each change, evaluated the cost and submitted them to the weekly Change and Control Board (CCB), held at the videoconference facility. Senior management at both sites would review and approve change requests, based on cost, timing, quality and technical issues. The planner analysts spent a great deal of time on the phone (10–30 calls per day) and managed 30–50 change requests per week.

We installed an Office Share type of link with a continuous video connection, active throughout the day. Because a continuously-available ISDN line would have been too expensive, we used the corporate TCP-IP network, using existing 128 Kb/s leased lines. We installed a Videopix video digitizing board on each of their Sun workstations, using two software packages: *vfctool*, which came with the Videopix board, and IVS, a public domain software developed at INRIA (Figure 3.15).

Vfctool grabbed frames from a video digitizing board shared over a LAN, without compression. With the network traffic between the Netherlands and England, it took up to six minutes to update an image with a size of 320×240 pixels and 8 gray levels. IVS [Tur93] was designed to support video and audio conferences over the Internet and achieved higher refresh rates by compressing the images according to the H.261 standard. IVS transmited the compressed data stream over an IP network using the User Datagram Protocol (UDP) and took about 20 to 30 Kb/s of bandwidth. We used QCIF images (176×144 pixels) with eight gray levels and obtained a refresh rate of one frame every two to four seconds, according to network traffic. IVS was very robust to packet loss and network overload; the only problem was that sometimes the video window was closed down. However the software never crashed and the user could restore the link with a couple of mouse clicks.

The link to support configuration management ran for a full six days, spread over a period of about one month. We spent several hours observing users while the link was up and interviewed them periodically. We abandoned Vfctool, because the refresh rate was too slow and it was unreliable as a source of information. Although they enjoyed putting up messages such as “Good morning”, “I’ll be back at 3pm” or “I am on holiday today”, the planner analysts were frustrated that the image was usually out-of-date and that the person in the image was often no longer there. We switched to IVS, with a smaller video window and lower image resolution, but a much higher refresh rate. The planner analysts regularly checked the IVS image before calling each other (at least ten times per day). They particularly enjoyed the Office Share on Monday nights, when they would often work until midnight preparing for the Change and Control Board meeting the next morning. They said that the link provided them with “remote solidarity”, encouraging them to drink coffee together and keep working until they were done. Another, more subtle aspect was how they communicated that they did not

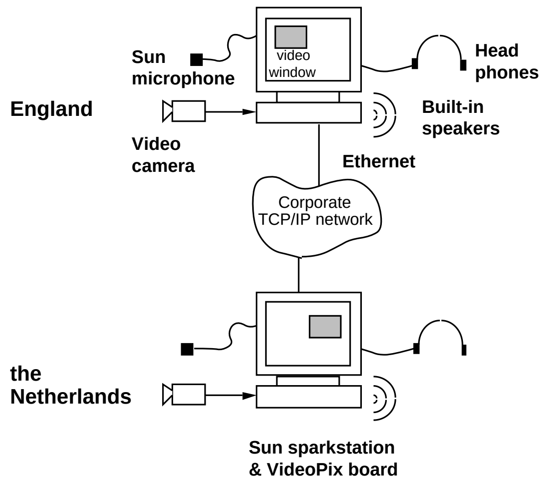

Figure 3.15 Office Share between planner analysts located in England and the Netherlands

want to be disturbed, by explicitly holding the phone, actively working through a large pile of papers or by moving out of the range of the camera. Their informal ways of communicating were encoded and decoded with no effort or attention; in most of the cases they were unaware that they were communicating. In one case, a planner analyst said: "Yesterday I saw you were talking with . . . ", but was later unable to remember that he had seen the person via the Office Share. Other people in the building also used the Office Share. Passersby would wave at the remote planner analyst and sometimes used it to find someone or talk to someone at the other site. People adapted easily to the link; after an initial period of self-consciousness they quickly forgot about the camera and responded to the other person.

3.6 ETHICAL ISSUES

Video is a very powerful medium, perhaps too powerful. One of the biggest issues is privacy: how do we balance the benefits of a relatively open media space with individuals' needs for privacy? Privacy issues are multi-dimensional and are greatly affected by the culture of the organization in which the media space is placed and the purposes for which it was created. Gaver [Gav92a] identifies four issues that must be disentangled when thinking about privacy:

- *Control:* Users want to control who can see or hear them at any time.
- *Knowledge:* Users want to know when somebody is in fact seeing or hearing them.
- *Intention:* Users want to know what the intention of the connection is.
- *Intrusions:* Users want to avoid connections that disturb their work.

Fish et al [Fis93] point out that the tradeoff between privacy and functionality involves a conflict between the desirability of control and knowledge and the intrusion implied by activities needed to maintain them. Explicitly acknowledging every connection provides control, but the requests themselves would be intrusive. Similarly, if every glance results in seeing someone's face on the monitor, it demands some sort of social response and may well disrupt previously-existing connections. If users must specify and be informed of the exact intention of every connection, the media space is no longer lightweight and is bogged down with continuous demands for the user's attention. The challenge is to provide users with control and notification, but in a lightweight and unobtrusive way.

Since privacy issues are affected by the social context in which the media space is embedded, it is not possible to simply create an "ideal technology" that is appropriate in every setting. The Xerox PARC Media Space, RAVE and WAVE media spaces worked within an atmosphere of trust because the participants knew each other and worked together. From a management perspective, it was also very important to enforce the idea that "turning off the media space" was acceptable behavior; allowing people at all levels of the organization to feel comfortable. (Contrast this to the experiences of users of the videoconferencing system in the WAVE study, in which high-level managers were very satisfied with the system, but everyone else found it to be disruptive.) Larger media spaces, such as Kasmer, have had some problems, when people suddenly found themselves being glanced at by people they don't know. Web-based media spaces such as Mediascape have world-wide reach, with correspondingly lower levels of trust among the users and require greater levels of privacy protection.

The organizations that created the media spaces in this chapter each developed their own safeguards to privacy, making judgments about how to balance privacy concerns while still making the media space worthwhile. In Cruiser [Roo88], all connections had to be symmetrical; such that hearing or seeing someone implied that that person could also see or hear you. The Media Space [Bly93] took the opposite extreme; all video connections were fully open with both audio and video links. This worked well when the media space involved close-knit members of a small work group, but eroded when others from other parts of the organization joined the media space and were seen to be "voyeurs". RAVE was based on the notion of services, such as Glance, in which the user's intention was incorporated into the service. Users decided in advance who had permission to glance at them; letting them avoid giving permission each time. RAVE specified access levels per person, which were rarely updated over time, whereas CAVECAT used the door metaphor to establish dynamic access levels, but did not distinguish who requested which type of access. Mediascape combines the two models: users can customize access rights according to the current state of the door as well as the origin of the call. This is especially important since Mediascape is accessible through the World Wide Web.

Media space designers need to explicitly consider a set of ethical issues when handling video [Mac95b]. People should be informed of the presence of live cameras. (Unfortunately, most of us are largely unaware of the myriad security cameras that capture and record video of us every time we use a bank teller, shop in a store or even walk down the street. Such uses of video increase our insecurity when contemplating media spaces.) At EuroPARC, a mannequin with a sign around his neck was a light-hearted way of letting people know they were in the range of the cameras. Displaying the camera's image on an adjacent monitor is also effective. People should be able to easily detect when a camera is left on all the time, such as in commons areas. People should be able to figure out when they are on camera and have the opportunity to avoid it by moving out of range. Recording video is especially problematic, since video taken out of context can be used in ways that may cause viewers to completely misinterpret what happened. People should know when video is being recorded and be given the opportunity to stop. Once recorded, people should have the ability to view the recorded material and consent to any further use of the material, by giving their informed consent.

3.7 CONCLUSION

Distributed video is not a single, unitary phenomenon that can understood simply at the level of the technology it incorporates. What is important is the way in which the video (and associated technologies) are set up and used within a social setting. Media spaces, with their emphasis on informal and open-ended as well as formal communication, are an important new approach for supporting distributed cooperative work groups. Media space designers must consider the context in which their technology will be used and ensure that uses can easily adapt them to meet the specific needs of their users. Media spaces are still in their infancy with much research to be done. However, as video costs continue to drop and as the Web becomes ubiquitous, media spaces promise to provide an effective means for supporting distributed, collaborative work.

ACKNOWLEDGEMENTS

The RAVE media space at EuroPARC and its extensions were a collaborative effort of many people, including Bob Anderson, Victoria Bellotti, Michel Beaudouin-Lafon, Bill Buxton, Kathy Carter, Ian Daniel, Paul Dourish, Bill Gaver, Christian Heath, Lennart Lvstrand, Allan MacLean, Tom Milligan, Mike Molloy, Tom Moran, Toby Morrill, Daniele Pagani, Gary Olson, Judy Olson and Randy Smith. Our collegues at Xerox PARC in Palo Alto also contributed research ideas, particularly Annette Adler, Sara Bly, Steve Harrison, Austin Henderson, Scott Minneman and John Tang.

REFERENCES

[Adl94] Adler, A. and Henderson, A., A room of our own: Experiences from a direct office share. In *Proceedings of Human Factors in Computing Systems, CHI '94* (Boston, MA), pages 138–144. ACM Press, New York, 1994.

[Bly93] Bly, S., Harrison, S. and Irwin, S., Media spaces: Bringing people together in a video audio and computing environment. *Communciations of the ACM* special issue, 36(1): 29–47, January, 1993.

[Bor91] Borning, A. and Travers, M., Two approaches to casual interaction over computer and video networks. In *Proceedings of Human Factors in Computing Systems, CHI '91* (New Orleans, LA), pages 13–19. ACM Press, New York, 1991.

[Bul89] Bulick, S., Abel, M., Corey, D., Schmidt, J. and Coffin, S., The US WEST Advanced Technologies Prototype Multimedia Communications System. GLOBECOM '89: In *Proceedings of the IEEE Global Telecommunications Conference* (Dallas, Texas), 1989.

[Bux90] Buxton, W. and Moran, T., EuroPARC's integrated interactive intermedia facility (iiif): Early experiences. In *Proceedings of the IFIP WG8.4 Conference on Multi-User Interfaces and Applications* (Herakleion, Crete), pages 11–34. North-Holland, 1990.

[Coo92] Cool, C. Fish, R., Kraut, R., Lowery, C., Iterative design of a video communication system. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '92* (Toronto, ON), pages 25–32. ACM Press, New York, 1992.

[Dou91a] Dourish, P., Godard: A flexible architecture for AV services in a media space. *EuroPARC working paper*, 1991.

[Dou91b] Dourish, P. and Bly, S., Portholes: Supporting awareness in a distributed work group. In *Proceedings of Human Factors in Computing Systems, CHI '92* (Monterey, CA), pages 541–547. ACM Press, New York, 1992.

[Dou93] Dourish, P., Bellotti, V., Mackay, W. and Ma, C., Information and context: Lessons from a study of two calendar systems. In *Proceedings of COCS '93* (San Francisco, CA). ACM Press, New York, 1993.

[Egi88] Egido, C., Videoconferencing as a technology to support group work: A review of its failure. In *Proceedings of the Conference on Computer Supported Cooperative Work, CSCW '88* (Portland, OR), pages13–24. ACM Press, New York, 1988.

[Ehr99] Ehrlich, K., Designing groupware applications: A work-centered approach. In Beaudouin-Lafon, M. (Ed.), *Computer Supported Cooperative Work*, Trends in Software Series 7:1–28. John Wiley & Sons, Chichester, 1999.

[Ell99] Ellis, C.A., Workflow technology. In Beaudouin-Lafon, M. (Ed.), *Computer Supported Cooperative Work*, Trends in Software Series 7:29–54. John Wiley & Sons, Chichester, 1999.

[Ens92] Ensor, J., Ahuja, S., Connaghan, R., Pack, M. and Seligmann, S., The Rapport multimedia communication system. Demo in *Conference Companion: Human Factors in Computing Systems, CHI '92* (Monterey, CA), pages 49–59. ACM Press, New York, 1992.

[Fis90] Fish, R., Kraut, R. and Chalfonte, B.L., The VideoWindows system in informal communications. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '90* (Los Angeles, CA), pages 1–13. ACM Press, New York, 1990.

[Fis93] Fish, R., Kraut, R., Root, R. and Rice, R.E., Video as a technology for informal communication. *Communciations of the ACM* special issue, 36(1): 48–61, January, 1993.

[Gal80] Galloway, J. and Rabinowitz, S., Hole-In-Space: *Mobile image videotape*. Santa Monica, CA.1980.

[Gal91] Gale, S., Adding audio and video to an office environment. In *Studies in Computer Supported Cooperative Work*. Bowers and Benford (Eds.), pages 49–62, Elsevier Science Publishers B.V., 1991.

[Gav86] Gaver, W.W., Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2:167–177. 1986.

[Gav91a] Gaver, W.W., Smith, R.B. and O'Shea, T., Effective sounds in complex systems: The ARKola simulation. In *Proceedings of Human Factors in Computing Systems, CHI '91* (New Orleans, LA), pages 85–90. ACM Press, New York, 1991.

[Gav91b] Gaver, W.W., Sound support for collaboration. In *Proceedings of the European Conference on Computer Supported Cooperative Work, ECSCW '91* (Amsterdam, the Netherlands), pages 293–308. ACM Press, New York, 1991.

[Gav92a] Gaver, W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K. and Buxton, W., Realizing a video environment: EuroPARC's RAVE system. In *Proceedings of Human Factors in Computing Systems, CHI '92* (Monterey, CA), pages 27–35. ACM Press, New York, 1992.

[Gav92b] Gaver, W.W., The affordances of media spaces for collaboration. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '92* (Toronto, ON), pages 17–24. ACM Press, New York, 1992.

[Gav93] Gaver, W., Sellen, A., Heath, C. and Luff, P., One is not enough: Multiple views in a media space. In *Proceedings of Human Factors in Computing Systems, InterCHI '93* (Amsterdam, the Netherlands), pages 335–341. ACM Press, New York, 1993.

[Har94] Harrison, B., Mantei, M., Beirne, G. and Narine, T., Communicating about communicating: Cross-disciplinary design of a media space interface. In *Proceedings of Human Factors in Computing Systems, CHI '94* (Boston, MA), pages 124–130. ACM Press, New York, 1994.

[Hea91] Heath, C. and Luff, P., Disembodied conduct: Communication through video in a multimedia office environment. In *Proceedings of Human Factors in Computing Systems, CHI '91* (New Orleans, LA), pages 99–103. ACM Press, New York, 1991.

[Hen86] Henderson, D.A. and Card, S., Rooms: The use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. *ACM Transactions on Graphics*, 5(3): 211–243, 1986.

[Hil94] Ralph Hill, Tom Brinck, Steven Rohall, John Patterson and Wayne Wilner. The RendezVous architecture and language for constructing multiuser applications. *ACM Transactions on Computer Human Interaction*, 1(2):81–125, June 1994.

[Isa93] Isaacs, E. and Tang, J., What video can and can't do for collaboration: A case study. In *Proceedings of ACM Multimedia '93* (Anaheim, CA), pages 199–206, 1993.

[Isa95] Isaacs, E., Morris, T., Rodriguez, T. and Tang, J., A comparison of face-to-face and distributed presentations. In *Proceedings of Human Factors in Computing Systems, CHI '95* (Denver, CO), pages 354–361. ACM Press, New York, 1995.

[Ish92] Ishii, H. and Kobayashi, M., Integration of interpersonal space and shared workspace: Clearboard design and experiments. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '92* (Toronto, ON), pages 33–42. ACM Press, New York, 1992.

[Ish99] Ishii, H., Integration of shared workspace and interpersonal space for remote collaboration. In Beaudouin-Lafon, M. (Ed.), *Computer Supported Cooperative Work*, Trends in Software Series 7:83–102. John Wiley & Sons, Chichester, 1999.

[Kob97] Kobayashi, M. and Schmandt, C., Dynamic soundscape: mapping time to space for audio browsing. In *Proceedings of Human Factors in Computing Systems, CHI'97* (Atlanta, GA), pages 194–201. ACM Press, New York, 1997.

[Kra88] Kraut, R. and Egido, C., Patterns of contact and communication in scientific research collaboration. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '88* (Portland, OR), pages 1–13. ACM Press, New York, 1988.

[Lov91] Löstrand, L., Being selectively aware with the Khronika system. In *Proceedings of the European Conference on Computer Supported Cooperative Work, ECSCW '91* (Amsterdam, the Netherlands), pages 265–278. ACM Press, New York, 1991.

[Mac90] MacLean, A., Carter, K., Moran, T. and Lövstrand, L., User-tailorable systems: Pressing the issues with Buttons. In *Proceedings of Human Factors in Computing Systems, CHI '90* (Seattle, WA), pages 175–182. ACM Press, New York, 1990.

[Mac92] Mackay, W.E., Spontaneous interaction in virtual multimedia space: EuroPARC's RAVE system. *Imagina '92*, Monte Carlo, Monaco. 1992.

[Mac95a] Mackay, W.E., Pagani D.S., Faber L., Inwood B., Launiainen P., Brenta L. and Pouzol V., Ariel: Augmenting paper engineering drawings. Video in *Conference Companion: Human Factors in Computing Systems, CHI '95* (Denver, CO), pages 421–422. ACM Press, New York, 1995.

[Mac95b] Mackay, W.E., Ethics, lies and videotapes. In *Proceedings of Human Factors in Computing Systems, CHI '95* (Denver, CO), pages 138–145. ACM Press, New York, 1995.

[Mac98] Mackay, W.E., Augmented Reality: linking real and virtual worlds. In *Proceedings of ACM Conference on Advanced Visual Interfaces, AVI '98* (L'Aquila, Italy), pages 13–21. ACM Press, New York. 1998.

[Man91] Mantei, M., Baecker, R., Sellen, A., Buxton, W., Milligan, T. and Wellman, B., Experiences in the use of a media space. In *Proceedings of Human Factors in Computing Systems, CHI '91* (New Orleans, LA), pages 203–208. ACM Press, New York, 1991.

[Min91] Minneman, S.L. and Bly, S.A., Managing a trois: A study of a multi-user drawing tool in distributed design work. In *Proceedings of Human Factors in Computing Systems, CHI '91* (New Orleans, LA), pages 217–224. ACM Press, New York, 1991.

[Mor90] Moran, T.P. and Anderson, R.J., The workaday world as a paradigm for CSCW design. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '90*. (Los Angeles, CA), pages 381–393. ACM Press, New York, 1990.

[Mor97] Moran, T., Palen, L., Harrison, S., Chiu, P., Kimber, D., Minneman, S., van Melle, W. and Zellweger, P., "I'll get that off the audio": A case study of salvaging multimedia meeting records. In *Proceedings of Human Factors in Computing Systems, CHI'97* (Atlanta, GA), pages 202–209. ACM Press, New York, 1997.

[Myn97] Mynatt, E., Adler, A., Ito, M. and O'Day, V., Design for network communities. In *Proceedings of Human Factors in Computing Systems, CHI'97* (Atlanta, GA), pages 210–217. ACM Press, New York, 1997.

[Nol92] Noll, A., Anatomy of a failure: Picturephone revisited. *Telecommunications Policy*, pages 307–316. May-June 1992.

[Ols91a] Olson, G. and Olson, J., User-centered design of collaboration technology. *Journal of Organizational Computing*, 1:61–83, 1991.

[Ols91b] Olson, M.H. and Bly, S., The Portland experience: A report on a distributed research group. *International Journal of Man–Machine Studies*, 34, 1991.

[Ols95] Olson, J and Olson, G., What mix of video and audio is useful for small groups doing remote real-time design work? In *Proceedings of Human Factors in Computing Systems, CHI'95*

(Denver, CO), pages 362–368. ACM Press, New York, 1995.

[Pag93] Pagani, D. and Mackay, W.E., Bringing media spaces into the real world. In *Proceedings of the European Conference on Computer-Supported Cooperative Work, ECSCW '93* (Milan, Italy). ACM Press, New York, 1993.

[Pat89] Patterson, R.D., Guidelines for the design of auditory warning sounds. In *Proceedings of the Institute of Acoustics 1989 Spring Conference*, 11(5):17–24, 1989.

[Pos92] Posner, I. and Baecker, R., How people write together. In *Proceedings of the Twenty-fifth Hawaii International Conference on Systems Sciences* (Kauai, Hawaii), volume IV, 1992.

[Pra99] Prakash, A., Group editors. In Beaudouin-Lafon, M. (Ed.), *Computer Supported Cooperative Work*, Trends in Software Series 7:103–133. John Wiley & Sons, Chichester, 1999.

[Roo88] Root, R.W., Design of a multi-media vehicle for social browsing. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '88* (Portland, OR), pages 25–38. ACM Press, New York, 1988.

[Rou98] Roussel, N., Towards a toolkit for building media spaces. *LRI Technical Report*, Université de Paris-Sud, Orsay, France, 1998. http://www-ihm.lri.fr/˜roussel/Mediascape.

[Sel92a] Sellen, A., Speech patterns in video-mediated conversations. In *Proceedings of Human Factors in Computing Systems, CHI '92* (Monterey, CA), pages 49–59. ACM Press, New York, 1992.

[Sel92b] Sellen, A., Buxton, W. and Arnott, J., Using spatial cues to improve videoconferencing. Video in *Conference Companion: Human Factors in Computing Systems, CHI '92* (Monterey, CA), pages 651–652. ACM Press, New York, 1992.

[Sel95] Seligmann, D., Mercuri, R. and Edmark, J., Providing assurances in a multimedia interactive environment. In *Proceedings of Human Factors in Computing Systems, CHI'95* (Denver, CO), pages 250–256. ACM Press, New York, 1995.

[Stu86] Stults, R., Media space. *Xerox PARC technical report*, 1986.

[Suc91] Suchman, L. and Trigg, R., Understanding practice: Video as a medium for reflection and design. In *Design at Work: Cooperative Design of Computer Systems*, Greenbaum and Kyng (Eds), Lawrence Erlbaum, Hillsdale, N.J., 1991.

[Tan90] Tang, J. and Minneman, S., VideoDraw: A video interface for collaborative drawing. In *Proceedings of Human Factors in Computing Systems, CHI '90* (Seattle, WA), pages 313–320. ACM Press, New York, 1990.

[Tan94] Tang, J. and Rua, M., Montage: Providing teleproximity for distributed groups. In *Proceedings of Human Factors in Computing Systems, CHI '94* (Boston, MA), pages 37–43. ACM Press, New York, 1994.

[Tur93] Turletti, T., H.262 Software codec for videoconferencing over the Internet. *INRIA Technical Report* No. 1834, Sophia Antipolis, France, 1993.

[Wel93] Wellner, P., Interacting with paper on the DigitalDesk. In *Communications of the ACM*, 36(7):86–96, July 1993.

[Win89] Winograd, T. and Flores, F., *Understanding Computers and Cognition: A New Foundation for Design*. NJ: Ablex, 1986.

[Yam96] Yamaashi, K., Cooperstock, J., Narine, T. and Buxton, B., Beating the limitations of camera-monitor mediated telepresence with extra eyes. In *Proceedings of Human Factors in Computing Systems, CHI '96* (Vancouver, BC), pages 50–57. ACM Press, New York, 1996.