

# 4

## Integration of Shared Workspace and Interpersonal Space for Remote Collaboration\*

HIROSHI ISHII

*MIT Media Laboratory*

### ABSTRACT

Computer-based groupware and video telephony are the major technological components of remote collaboration support. However, integration of these two components has been a big design challenge. This chapter introduces the research effort to integrate the shared workspace created by groupware technology and the interpersonal space supported by video communication technology. TeamWorkStation and ClearBoard will be introduced as example systems which were designed to support focused real-time collaboration by distributed group members.

### 4.1 INTRODUCTION

“Groupware” is a label for computer-based systems explicitly designed to support groups of people working together. It is growing rapidly as a new application category in the computer industry [Eil91, Col93].

Most of the current groupware such as workflow systems and collaborative authoring tools are devoted to computational support and are designed under the constraint of limited communication bandwidth. However, the deployment of broadband digital networks opens a new future for multimedia collaboration environments that integrate real-time audio and video

---

\* This chapter is based on an article “Iterative design of seamless collaboration media” originally published in *Communications of the ACM*, Special Issue on Internet Technology, Vol. 37, No. 8, August 1994, pp. 83–97, with the permission of ACM.

communication links with computer-based shared workspaces [Bri92, Ly193]. Especially, the integration of two functional spaces, shared workspace (e.g. electronic shared whiteboard) and interpersonal space (e.g. videophone), is a critical interface design issue.

#### 4.1.1 Shared Workspace and Interpersonal Space

One major focus of groupware development has been the creation of virtual “shared workspaces” in distributed computer environments. Some groupware definitions take this workspace-oriented view, such as:

“Groupware . . . the computer-based systems that support groups of people engaged in a common task (or goal) and that provide an interface to a shared environment.” [Ell91]

Whiteboards and overhead projections of transparencies are examples of shared workspaces in face-to-face meetings. Participants can see, point to, or draw on a whiteboard simultaneously. An overhead projector makes handwritten or computer-generated documents visible to all participants in a room while permitting the speaker to point or draw. Shared workspace activities include sharing information, pointing to specific items, marking, annotating, and editing.

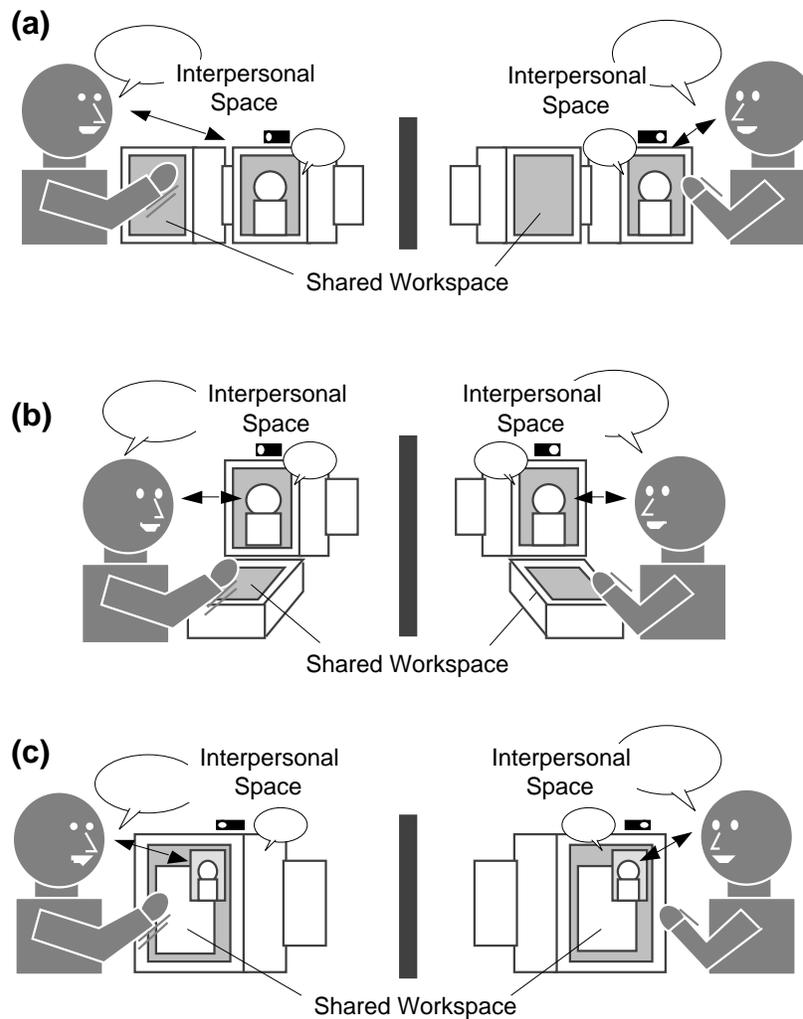
In a distributed, real-time collaboration, these activities can be supported by computer-based groupware, including

- shared screen systems such as Timbuktu [Far91]
- shared window systems such as VConf and Dialogo [Lau90], and
- multi-user editors such as Cognoter [Fos86], GROVE [Ell91], Commune [Bly90], Cave-Draw [Lu91], Aspects [Gro90], GroupSketch [Gre92], GroupDraw [Gre92], We-Met [Wol92], and TeamPaint (described later). Use of hand gestures in a shared workspace can be supported by shared video drawing media such as VideoDraw [Tang91] and Team-WorkStation [Ish90].

In face-to-face meetings, we speak, make eye contact, and observe each other’s facial expressions and gestures. These verbal and non-verbal channels are important in building confidence and establishing trust [Arg75, Bux92, Man91]. The focus of telecommunication technologies such as the video-phone and videoconferencing has been the creation of “interpersonal spaces” that maintain a sense of “telepresence” or “being there” [Hol92] through the visibility of gestures and facial expressions of distributed group members. “Media Space” is an example of such technologies. Originated by Xerox PARC [Bly93], it is an environment that integrates video, audio, and computer technologies, allowing individual and groups to work together despite being distributed geographically and temporally (see also Chapter 3 in this book [Mac99]). Recent developments include Cruiser (Bellcore) [Fis93], VideoWindow [Fis90], RAVE (Rank Xerox EuroPARC) [Gav92], and CAVECAT/Telepresence (University of Toronto) [Man91]. (“Media space”, originally the name of a specific system [Bly93], is used here in the sense of Mantei et al [Man91] as a general term to represent computer-controlled video environments.)

#### 4.1.2 Limitations of Existing Support Technologies

Both shared workspace and interpersonal space are present in ordinary face-to-face meetings and may be essential for remote real-time collaboration. Several media space technologies support both spaces.



**Figure 4.1** Typical screen arrangements in media space

Figure 4.1 illustrates three typical display arrangements of media spaces. In (a), a display providing a live video image of the partner's face adjoins a display showing the shared work. The ARKola simulation [Gav91] in the IIIF environment [Bux90a] and some nodes of CAVECAT [Man91] adopt this arrangement. SEPIA-IPSI media space [Str92] locates small custom-built desktop video devices (small monitors and cameras) on top of the computer screen. In (b), the displays are repositioned to resemble the situation of interacting across a table. VideoDraw [Tang91] and Commune [Bly90, Min91] adopt this arrangement. In (c), the live video images and the shared workspaces are incorporated into different windows of a single screen. TeamWorkStation, PMTC [Tan91], MERMAID [Wat90] and most of the recent PC-based desktop conferencing products employ this desktop-video technology.

### 4.1.3 Iterative Design of Seamless Collaboration Media

We have been exploring the future of collaboration media that make good use of real-time video through the iterative design of various groupware systems. Our research started in 1988 and was motivated by the study on shared drawing space [Tang91] in the Media Space environment [Bly93]. While most of the current video telephony has been designed to see “talking heads”, our goal is to go beyond this model and demonstrate new usage of video-mediated communication technologies. Video is powerful media for not only seeing talking heads, but also for creating shared workspaces and shared visual context for remote collaboration.

The following sections introduce the progression of iterative media design from TeamWorkStation to ClearBoard. These systems were designed to support focused real-time collaboration by distributed group members. The key concept behind our iterative design is “seamlessness”. Seamless design pursues the following two goals:

- *Seamlessness (continuity) with existing work practices*: People develop their own work practices after using a wide variety of tools and interacting with a large number of people. We believe the continuity with existing work practices and everyday skills is essential. Groupware that asks users to abandon their acquired skills and to learn a new protocol is likely to encounter strong resistance [Gru88].
- *Seamlessness (smooth transition) between functional spaces*: Collaboration requires us to shift among a variety of functional spaces or modes. Seamless design undertakes to decrease the cognitive load of users as they move dynamically across different spaces. For example, TeamWorkStation was designed to enable smooth transition between individual workspaces and shared workspaces by allowing users to keep using both familiar desktop tools and computer tools. ClearBoard realizes seamless integration of interpersonal space and shared workspace allowing people to use various non-verbal cues such as a partner’s gaze direction for smooth focus switching between these two spaces.

## 4.2 TEAMWORKSTATION-1 AND SEAMLESS SHARED WORKSPACES

People do a lot of their work without computers, or using different tools on different computer systems, and have developed their own work practices for these situations. Even in a heavily computerized individual workplace, users often work both with computers and on the physical desktop. Neither one can replace the other. For example, printed materials such as books and magazines are still an indispensable source of information. Therefore, when designing real-time shared workspaces, depending on the task and the media of the information to be shared (paper or computer file), co-workers should be able to choose either computers or desktops, and to switch between them freely. One person’s choice should be independent of the other members’ choices. Group members should be able to use a variety of heterogeneous tools (computer-based and manual tools) in the shared workspace simultaneously. To realize such a seamless shared workspace, we designed TeamWorkStation-1 (TWS-1) [Ish90, Ish91].

The key design idea of TWS-1 is a “translucent overlay” of individual workspace images. TWS-1 combines two or more translucent live-video images of computer screens or physical desktop surfaces using a video synthesis technique. Translucent overlay allows users to combine individual workspaces and to point to and draw on the overlaid images simultaneously. We chose video as the basic medium of TWS because it is the most powerful for fusing presentations of traditionally incompatible visual media such as papers and computer documents.



**Figure 4.2** Overview of TeamWorkStation-1 prototype

#### 4.2.1 System Architecture of TWS-1

Figure 4.2 shows an overview of the first prototype, TWS-1. Two CCD video cameras are provided at each workstation: one for capturing live face images of the group member, and the other for capturing the desktop surface images and hand gestures. TWS-1 provides two screens. The individual screen (private workspace) is on the left and the shared screen is on the right. These two screens are contiguous in video memory, and this multi-screen architecture allows users to move any application program window between the individual and shared screens by merely mouse dragging. Therefore, it is easy to bring your own data and tools from each personal computer into the shared workspace to use in remote collaboration. Hardcopy information can also be shared easily by placing it under the CCD camera (i.e. on the physical desktop). Figure 4.3 shows an image of a shared screen where two users are discussing the system configuration by annotating and pointing electronic diagrams in a drawing editor by hand.

The first prototype TWS-1 was implemented on Macintosh computers to provide small work groups (2–4 members) with a shared workspace. The system architecture of TWS-1 is illustrated in Figure 4.4 [Ish91]. The video network is controlled by a video server that is based on a computer-controllable video switcher and video effector. The video server gathers, processes and distributes the shared computer screen images, desktop images, and face images. Overlay of video images is done by the video server. The results of overlaying are redistributed to the shared screens via the video network.

#### 4.2.2 Experience of TWS-1

Through experimental use of TWS-1, we found that users liked the feature which allowed them to keep using their favorite individual tools, especially papers and pen, while collaborating in a desktop shared workspace. That is, there was no need to master the usage of new sophisticated graphware. The drawback of this overlay approach is that the results of collab-

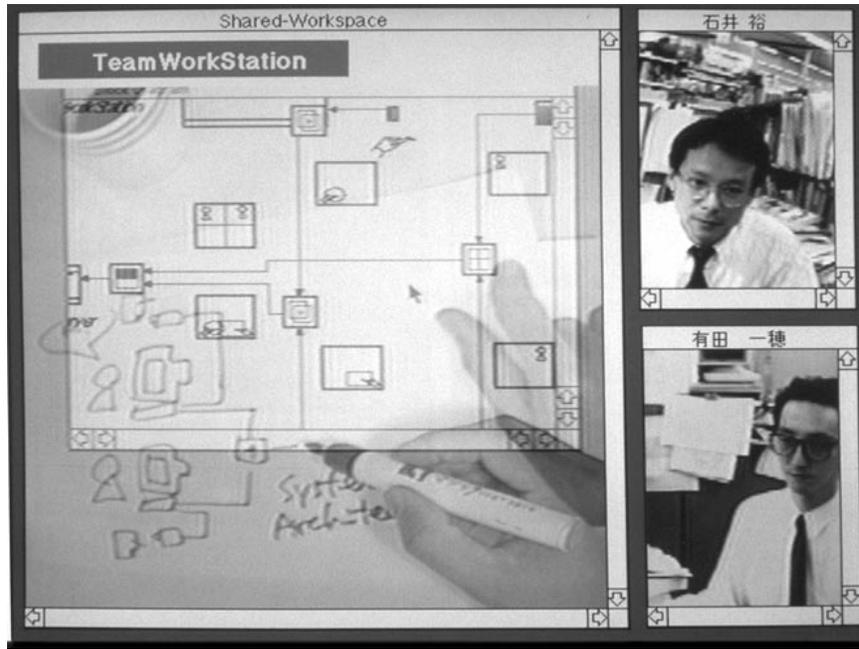


Figure 4.3 A shared screen of TeamWorkStation-1

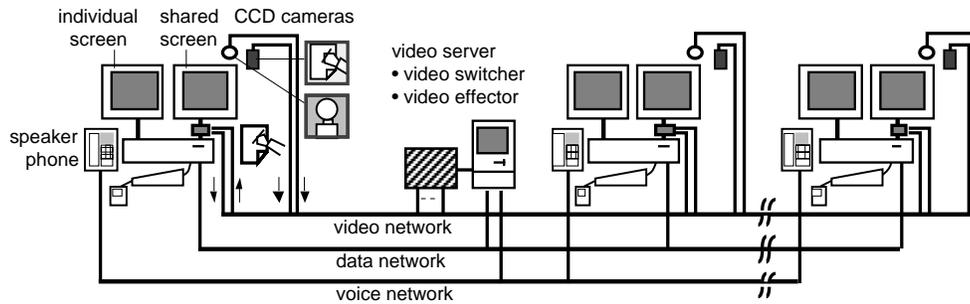
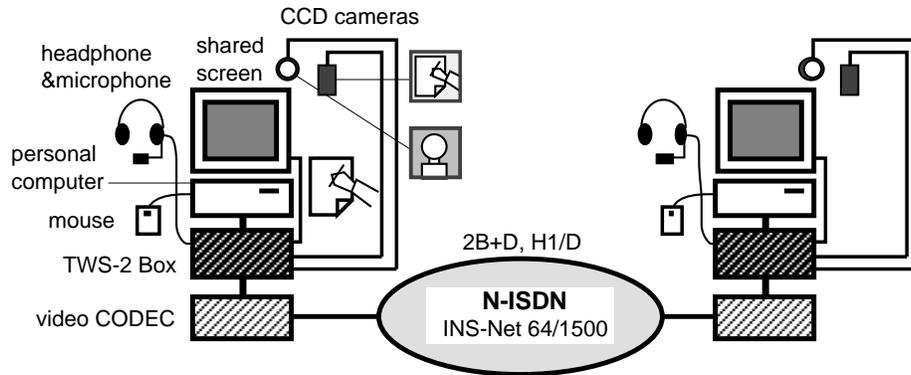


Figure 4.4 System architecture of TeamWorkStation-1 prototype

oration cannot be shared directly. Since individual workspaces are overlaid as video images, the marks and the marked documents occupy different “layers” in the shared screens. They are actually stored separately in different places in different media (in computer files or on paper). We mainly used a video printer or video tape recorder to record the results and the collaboration process.

“Shared workspace” is taken by many computer scientists to mean “data sharing”. However, we think it is not required that all the outcomes of the work-in-progress be directly “manipulable” by all the participants. We seldom felt the necessity to edit the other’s diagrams directly. If a diagram was to be changed, usually the originator would change it according to the comments made by the other. One reason appears to stem from the respect paid to the ownership of the outcomes. This seems to be a very natural feeling, even in a close collaborative session.



**Figure 4.5** System architecture of TeamWorkStation-2 prototype

The overlay solution provides us with a comfortable work environment, because the overlaid layers keep the person's own layer of work intact.

Since TWS-1 was designed for laboratory experiments to verify the concept of seamless shared workspaces, we did not pay much attention to the number of cables or the communication bandwidth. As a result, the system configuration became complex and difficult to maintain. This complexity prevented us from conducting the field tests using publicly available digital networks, and motivated us to start designing a completely new system, TeamWorkStation-2.

### 4.3 TEAMWORKSTATION-2 FOR N-ISDN

TeamWorkStation-2 (TWS-2) was designed to provide a shared workspace over narrowband ISDN (N-ISDN) Basic Rate Interface (2B+D) and the Primary Rate Interface (H1/D) using the CCITT H.261 standard of moving picture transmission [Ish93a]. We chose N-ISDN Basic Rate Interface as the target network because of its widespread availability in Japan.

We devised a new multi-user interface called ClearFace for TWS-2. ClearFace superimposes translucent, movable, and resizable face windows over a workspace image to enable more effective use of the normally limited screen space. We found users had little difficulty in selectively viewing either the facial image or the workspace image.

#### 4.3.1 System Architecture of TWS-2

We radically simplified the system architecture. Figure 4.5 shows the system architecture of TWS-2. We targeted dyadic communication to make the centralized video server unnecessary and to eliminate complexities that would arise from multipoint connection. The two TWS-2 terminals are connected by one ISDN link. Each terminal is composed of three major components: a TWS-2 box, a video codec, and a PC-9801™ personal computer. All video processing functions (e.g. translucent overlay, picture-in-picture) are supported at each terminal. All the hardware for video processing, camera control units, audio amplifiers, and power units were encapsulated into a single TWS-2 box.

The PC-9801™ computer is mainly used to control the video processing hardware in the TWS-2 box and the video codec. If direct sharing of information stored in the computer is



**Figure 4.6** Appearance of TWS-2 terminal

required, we can use screen-sharing software while overlaying desktop video images with the shared computer screen.

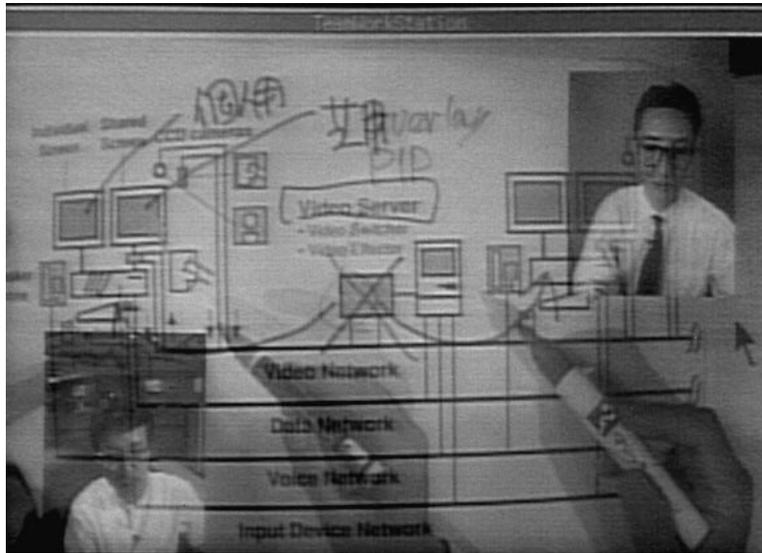
Figure 4.6 shows the appearance of a TWS-2 terminal in use. A headphone with a small microphone is provided for voice communication. Like TWS-1, TWS-2 provides two CCD cameras, one to capture the user's face image and another to capture the physical desktop image. The TWS-2 box also provides an external video input port that can be used to show recorded video clips by connecting a video player.

TWS-2 provides only one screen instead of the two screens (individual and shared screens) of TWS-1. The experimental use of the previous system, TWS-1, led us to observe that most work was done in the "desktop-overlay" mode in which only the video images of physical desktop surfaces are overlaid. We decided to make "desktop overlay" the basic service of TWS-2, and to make "computer screen overlay" an option. This decision led to the one-screen architecture of TWS-2.

Figure 4.7 shows a typical screen image of TWS-2 in use. In Figure 4.7, users are discussing the system architecture using hand drawing and gestures. This example demonstrates the important TWS feature that all of the collaborators share not only the results of drawing, but also the dynamic process of drawing and gesturing.

### **4.3.2 Experimental Use of TWS-2**

We have tested TWS-2 since 1992 by connecting our offices in Tokyo, Yokosuka, and Osaka by INS-Net 64. (The latest version of TWS-2 is available via Ethernet as well as N-ISDN.) We conducted several controlled laboratory experiments as well as tests of real work outside of laboratories [Ish93a]. Before we started the TWS-2 experiments, many people felt



**Figure 4.7** Design session via TWS-2

dubious about the ability of INS-Net 64 to support real-time activities because of their previous experience with the jerky displays of video phones. However, the subjects generally commented that they could smoothly interact with their partner and that they were absorbed in the task. Although the subjects noticed some delay and jerkiness in the remote desktop video image, these did not hinder subjects from concentrating on their work. However, all the subjects noted that they could not clearly see their partner's desktop image. This confirmed that the CIF (Common Intermediate Format) (352 pixels/line  $\times$  288 lines/picture) standard is definitely insufficient to see small characters or fine drawings in the remote documents.

### 4.3.3 Beyond the Video Phone

Video phones and videoconferencing are the most typical video applications that use N-ISDN, and they represent the effort at imitating "being there" which has long been the goal of telecommunication technologies [Hol92]. Real-time video is used only to see the remote partners' facial expressions, postures and gestures in these applications.

In contrast to these "talking head" applications, TWS-2 demonstrates a new direction for the usage of real-time video: the creation of a virtual shared workspace. The main focus of TWS-2 is not the imitation of face-to-face communication but rather the sharing of overlaid desktop images for collaboration.

The experiments to date confirm that TWS-2 has one large advantage over ordinary video phones as the pre-eminent N-ISDN service. The advantage is due to the bandwidth limitation and human perception. People are especially perceptive to changes in facial expressions. If facial expression is the main means of communication, even slight asynchronism between the voice and the movement of eyes and lips is immediately noticed, and makes smooth conversation difficult. Since the facial expression is always changing and the face and body are always moving, delay in transmitting the partner's image increases perceived discontinuities and hence increases the negative impression of users.

The main difference between the desktop and face images is that the desktop images are relatively static. Images of papers and the marks drawn on the papers do not change quickly. Only the hands move on the desktop when users gesture or draw. Thus the total amount of motion is far less than that experienced with video phone displays. This more static nature of the desktop surface increases the effective video frame rate. Although quick hand motions look jerky, TWS-2 users can be more productive than their video phone counterparts since they can visually share objects and work on them.

#### **4.4 SEAMLESS INTEGRATION OF INTERPERSONAL SPACE AND SHARED WORKSPACE**

One major focus of groupware development has been the creation of virtual “shared workspaces” in distributed computer environments. Shared workspace activities include sharing information, pointing to specific items, marking, annotating, and editing. These activities can be supported by computer-based groupware, including shared screen systems, shared window systems, and multi-user editors [El191] (see also Chapter 5 in this book [Pra99]).

In face-to-face meetings, we speak, make eye contact, and observe each other’s facial expressions and gestures. These verbal and non-verbal channels are important in building confidence and establishing trust [Bux92]. The focus of telecommunication technologies such as the video phone and videoconferencing has been the creation of “interpersonal spaces” that maintain a sense of “telepresence” or “being there” [Hol92] through the visibility of gestures and facial expressions of distributed group members.

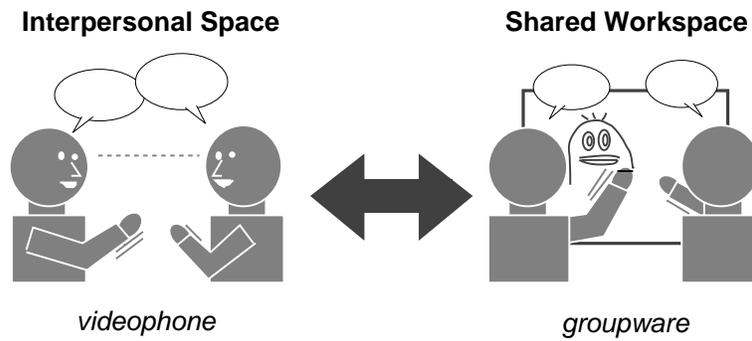
Both shared workspace and interpersonal space are essential for remote, real-time collaboration. Many desktop multimedia conferencing systems such as TeamWorkStation, PMTC [Tan91], and MERMAID [Wat90] support both spaces, but they have a major limitation: an arbitrary seam exists between the shared workspace and the face images. We realized that this problem was not just the superficial physical discontinuity of spatially separated windows. Absent are the non-verbal cues that would enable a smooth shift in attention between the shared workspace and the partner’s face image. Current groupware and videoconferencing technologies do not support these cues.

Lack of eye contact is another problem of TWS. Camera positioning prevents one person from knowing the direction of the other’s gaze; it could be directed toward the face image, toward objects in the shared workspace window, or elsewhere. A shift in focus is not apparent until accompanied by a visible gesture or an audible remark. Awareness of gaze direction and mutual eye contact are impossible.

ClearBoard is designed to overcome these limitations by seamlessly integrating interpersonal space and shared workspace (Figure 4.8). A design goal of ClearBoard is to allow a pair of users to shift easily between interpersonal space and shared workspace using familiar everyday cues such as the partner’s gestures, head movements, eye contact, and gaze direction.

##### **4.4.1 ClearBoard Metaphor**

The key metaphor of ClearBoard design is “talking through and drawing on a big transparent glass board”. Figure 4.9 shows “ClearBoard-0” which is the simple mock-up of this ClearBoard concept for co-located pairs of users. ClearBoard-0 consists of a glass board positioned between the partners on which they draw or post objects. ClearBoard requires less eye and



**Figure 4.8** Seamless integration of interpersonal space and shared workspace



**Figure 4.9** A simple mockup of the ClearBoard metaphor: ClearBoard-0

head movement to switch focus between the drawing surface and the partner’s face than is needed in either the whiteboard or the desktop environment. However, a real glass board has the problem that written text appears reversed to one’s partner; we were able to solve this problem by mirror-reversing video images in ClearBoard-1 and 2 as described below.

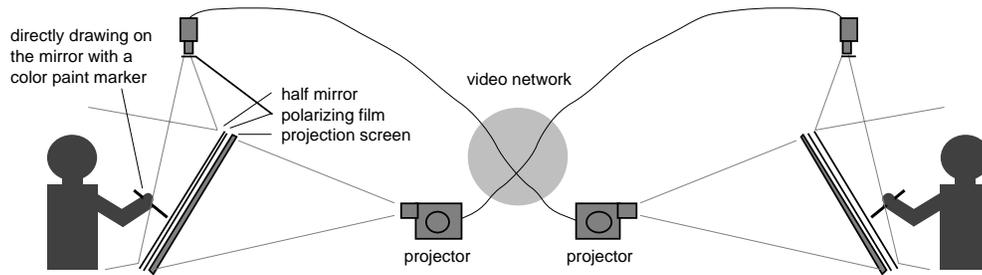
#### 4.5 DESIGN OF CLEARBOARD-1

Figure 4.10 shows ClearBoard-1, our first prototype to support remote collaboration [Ish92a]. Two users are discussing a route by drawing a map directly on the screen surface. Both users can share a common map orientation. The partner can read all the text and graphics in their correct orientation.

In order to implement the remote version of ClearBoard, we devised the system architecture called “drafter-mirror” architecture illustrated in Figure 4.11. Each terminal is equipped with a tilted screen, a video projector and a video camera. Users can write and draw directly on the surface of the screen using color paint markers. The video camera located above the screen captures the drawings and the user’s image as reflected by the half-mirror as a continuous



**Figure 4.10** ClearBoard-1 in use



**Figure 4.11** System architecture of the ClearBoard-1 prototype

video image. This image is sent to the other terminal through a video network and projected onto the partner's screen from the rear so that both users can share a common orientation of the drawing space. The partner can draw directly over this transmitted video image. This shared video drawing technique, which allows remote partners to draw directly over the video image of their co-workers' drawing surface, was originally demonstrated in VideoDraw [Tang91].

#### 4.5.1 Experimental Use of ClearBoard-1

Since 1990 this prototype has been used in experimental sessions. We observed effortless focus switching between the task and the partner's face. Users could read their partner's facial expression, achieve eye contact, and utilize their awareness of the direction of their partner's gaze. Easy eye contact even during drawing-intensive activities increased the feeling of intimacy and co-presence. No subjects reported difficulty with the mirror-reversal of the partner. This may be because our faces are quite symmetric, or our own images are reversed in mirrors.

We found that ClearBoard provides the capability we call "gaze awareness": the ability to monitor the direction of a partner's gaze and thus his or her focus of attention. A ClearBoard user can tell which screen objects the partner is gazing at during a conversation more easily and precisely than is possible in an ordinary meeting environment with a whiteboard.

To understand the implication of gaze awareness, we conducted a collaborative problem solving experiment on ClearBoard using the "river crossing problem." [Ish93b]. This experi-

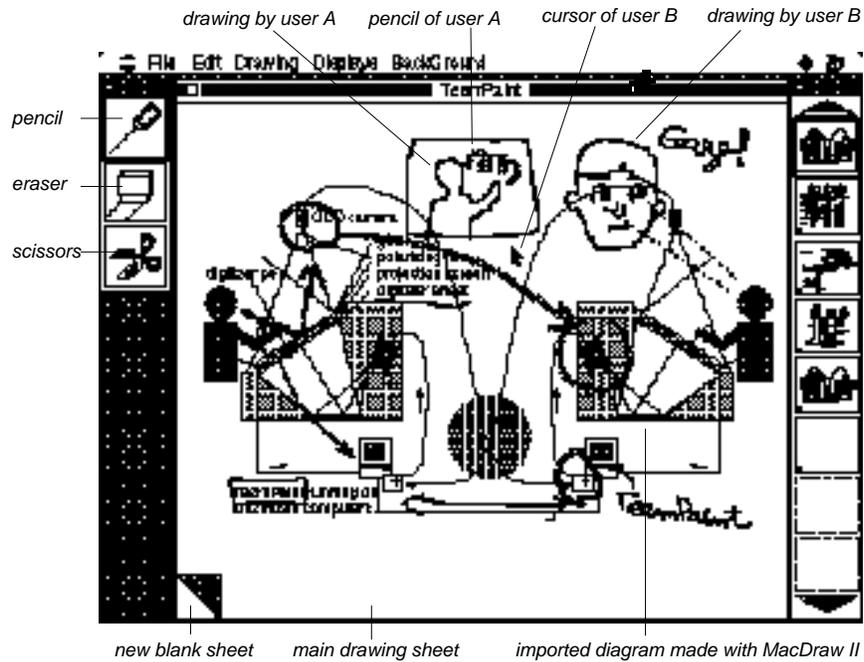


Figure 4.12 A screen of TeamPaint

ment confirmed that it is easy for the instructor to say which side of the river the student was gazing at. This information was quite useful in understanding the student’s thinking process and in providing advice. The importance of eye contact in the design of face-to-face communication tools is often discussed. However, we believe the concept of gaze awareness is more general and more important. Eye contact can be seen as a special case of gaze awareness.

An interesting and less critical confusion manifested itself when users directly drew over their partner’s image, playfully adding eye glasses or a mustache, for example. Clearly they had a “WYSIWIS” (what you see is what I see) expectation, not realizing that although the drawing is shared, the facial images are not, with each person seeing only the other’s image. Thus, the metaphor of the ClearBoard is not always entirely assimilated.

#### 4.6 DESIGN OF CLEARBOARD-2

In using this ClearBoard-1 prototype, we found several problems. The projected video image of a drawing was not sufficiently clear. Lack of recording capabilities was an obstacle to re-using the work results. To overcome these problems in ClearBoard-1, we decided to design a new computer-based prototype, “ClearBoard-2” [Ish93b]. Instead of using color paint markers, ClearBoard-2 provides users with “TeamPaint”, a multi-user computer-based paint editor and digitizer pen.



**Figure 4.13** ClearBoard-2 in use

#### 4.6.1 TeamPaint

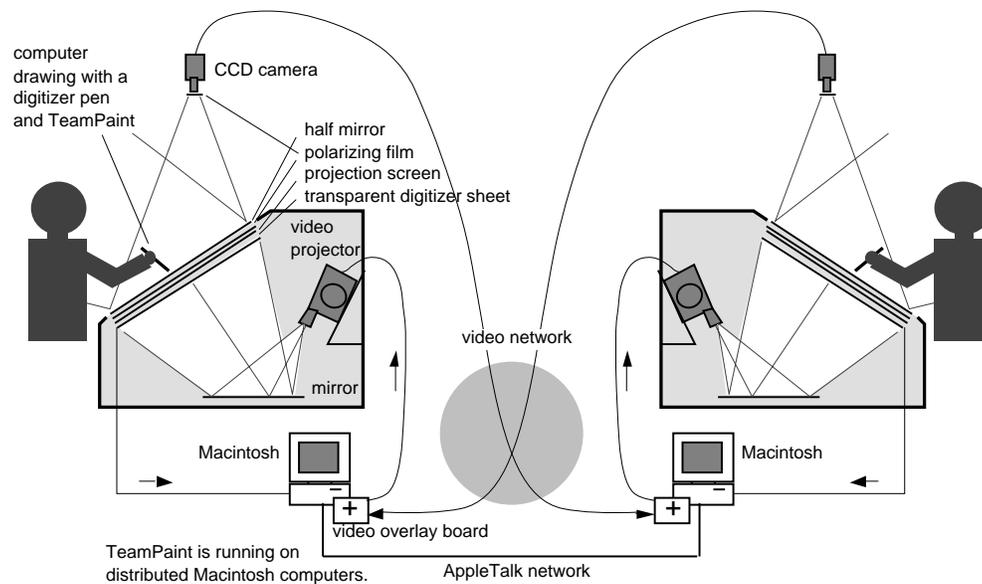
TeamPaint is a groupware application for shared drawing. It runs on networked Macintosh™ computers, and it is based on a replicated architecture. TeamPaint offers several functions: recording of working results, easy manipulation of marks, and the use of data held in computer files. TeamPaint provides an intuitive interface based on the metaphor of drawing on a sketch pad with a color pencil as shown in Figure 4.12.

Each user is provided with individual layers and can only modify his or her own layers by default. All members see the composite image of all the layers. Because each layer is isolated from the others, no access control is necessary. TeamPaint has no floor control mechanisms but enables simultaneous gesturing and drawing by multiple users. Gestures, in the form of cursor movements, and through them the drawing process, are visually shared by all members.

#### 4.6.2 The ClearBoard-2 System and Its Use

Using TeamPaint, transparent digitizer sheets, and electronic pens, we implemented a computer-based prototype, ClearBoard-2. Figure 4.13 shows the ClearBoard-2 prototype in use, and Figure 4.14 shows the system architecture of the prototype. The composite drawing image of TeamPaint is made to overlay the face images with a special video overlay board. The mixed RGB video image is projected onto the screen's rear surface. TeamPaint makes it easy to get a new blank sheet and the drawing marks are easier to see. The lower screen angle decreases arm fatigue, but gives the impression that the partner is under the screen, rather than behind it as in ClearBoard-1.

The use of RGB video and the chroma-keying overlay technique does increase image clarity. Furthermore, the capability of recording results and re-using the data produced in previous sessions or from any other application program promises to add tremendous value to an already practical tool. Through the use of ClearBoard-2, it was often observed that the user's gaze follows the partner's pen movements. We confirmed that "gaze awareness" is as well supported in ClearBoard-2 as it was in ClearBoard-1. One can easily tell which object on the TeamPaint screen the partner is looking at.



**Figure 4.14** System architecture of ClearBoard-2 prototype

We see the evolution from ClearBoard-1 to ClearBoard-2 as being very important. Computer and video-communication technologies have, until now, evolved independently. Although they have been loosely coupled using arbitrary multi-window interfaces in many desktop multimedia conferencing systems, they have never been integrated seamlessly from the users' cognitive point of view. ClearBoard-2 succeeds in naturally integrating the user interfaces of computer-based groupware with that of videoconferencing. We expect that the seamless integration of computer and video communication technologies will be an integral part of the next generation of collaboration media.

Moreover, ClearBoard-2 can be seen as an instance of the paradigm shift from traditional HCI (Human-Computer Interaction) to HHI (Human-Human Interaction) mediated by computers. We are interacting *not with* computers, but *through* computers.

We believe that the ClearBoard design is not only beyond the traditional desktop metaphor based on a multi-window interface, but also suggests a direction of "beyond being there" [Ho192]. We expect ClearBoard to be useful both as a collaboration medium and as a vehicle to investigate the nature of dynamic human interaction.

## 4.7 SUMMARY AND FUTURE WORK

This chapter has discussed the integration of shared workspace and interpersonal space for real-time remote collaboration, and has introduced an evolution of our seamless collaboration media design. TeamWorkStation (TWS) demonstrates a new usage of real-time video for collaboration, by providing distributed users with a seamless shared workspace. Using a translucent video overlay technique, real-time information such as hand gestures and handwritten comments can be shared, as can information contained in printed materials and computer files. Users can continue to use their favorite application programs or desktop tools, so

there is only a minor cognitive seam between individual workspaces and shared workspaces. TWS-2, a redesigned version which uses N-ISDN Basic Rate Interface, demonstrated the advantage of this application over ordinary videophones given the same bandwidth limitation.

In order to integrate the shared workspace and the interpersonal space seamlessly, we designed ClearBoard. ClearBoard-1 permits co-workers in two different locations to draw with color markers while maintaining direct eye contact and the use of natural gestures. Through experiments, we discovered that ClearBoard also supported the important feature of “gaze awareness”. In order to offer new functions, such as recording of working results, easy manipulation of marks, and the use of data held in computer files, we designed a computer-drawing version, ClearBoard-2. ClearBoard-2 supports shared drawing with TeamPaint and electronic pens.

Through the iterative design of these collaboration media, we believe it is most important to respect the skills that people use in everyday life [Bux90b]. The design focuses on basic skills such as drawing, gesturing, talking, gaze reading, and using computers. We believe skill-based design will lead to cognitive seamlessness.

We are now very much interested in how the next generation of collaboration media may augment the process of collaborative creation by people such as artists, musicians, designers and children. NTT’s vision video, “Seamless Media Design” [NTT93], illustrates our expectation of future collaboration media based on the ClearBoard concept.

#### 4.7.1 Broadband Digital Network

Although all these prototype systems except for TWS-2 were implemented using hybrid (analog video + digital data) networks, it is obvious that hybrid networks have serious limitations in extending their scale.

We expect that the new international telecommunication standard B-ISDN (Broadband Integrated Services Digital Network) and ATM (Asynchronous Transfer Mode) [Ly193] will provide a universal and scalable infrastructure for various collaborative applications including TeamWorkStation and ClearBoard. ATM is expected to be a common technology for both LAN (Local Area Networks) and WAN (Wide Area Networks). ATM also provides “bandwidth-on-demand” to meet the requirements of many applications.

Although N-ISDN provides users with fixed communication bandwidth, we expect that ATM technology will provide users with the flexibility to dynamically change the appropriate bandwidth and the balance between the frame rate and resolution of motion pictures on demand (based on the contents and the usage of video). For example, a TWS session using a detailed blueprint of a new building may require more bandwidth for higher resolution of shared documents compared with a TWS meeting with shared sheets of blank paper for freehand drawing. Competitive negotiation tasks may require both higher frame rate and resolution to read your colleague’s subtle facial expression rather than documents. ClearBoard requires much more communication bandwidth (higher resolution, higher frame rate, and less delay) than TWS since ClearBoard presents a life-size partner’s image and users want to read subtle and quick changes of a partner’s gaze.

Since required bandwidth changes dynamically both within a single application depending on the contents and usage of video, and among various applications, rapid reassignment of bandwidth on demand will be a critical feature to support seamless transitions among various collaboration modes.

### 4.7.2 From Multimedia to Seamless Media

“Multimedia” is now becoming a big buzz word in the computer and communication industries. As a result, the number of cables behind a computer, the number of features users need to understand, and the number of incompatible data formats are increasing beyond the limits of human cognitive capability. A variety of media (such as text, video, audio, graphics) and services (on-demand video, videoconferencing, electronic newspaper) are becoming available through a single powerful computer on the desktop and a broadband communication network named the “information super highway”. However, each medium and service are still separated from each other and they are not seamlessly integrated from a user’s cognitive point of view.

The communication channels of human beings are inherently multi-modal and seamless. It does not make much sense to decompose the representation of information into primitive data types such as text, video, audio, and graphics, and stress the “multi-ness” of the media. For example, we are speaking, gesturing, and drawing simultaneously in a design meeting. We have great skills to express ideas and understand each other in everyday contexts using all these media as a whole. We believe the multi-ness of media is not the main issue; how to integrate them into a seamless media, hiding the various low-level representations, discontinuities among primitive media, and complexity of underlying technologies is the core issue in designing new applications. “Multi-media” sounds like a premature label that represents a stage of media evolution from the mono-media to the seamless media.

### 4.7.3 Toward Ubiquitous Media and Augmented Reality

We hope that ClearBoard will change our concept of a wall from being a passive partition to being a dynamic collaboration medium that integrates distributed real and virtual spaces. We are now exploring a vision of new architectural spaces where all the surfaces including walls, ceilings, windows, doors and desktops become active surfaces through which people can interact with other spaces, both real and virtual. In these spaces, both computers and video must be inherently ubiquitous media [Bux94, Wei91]. Many challenges exist to achieve a seamless extension of spaces and their interconnections. Nevertheless, our design will be based on the natural skills and social protocols people are using in everyday life to manipulate and interact with information, artifacts and each other.

## ACKNOWLEDGEMENTS

I would like to thank Masaaki Ohkubo, Kazuho Arita, and Takashi Yagi at NTT for their contribution to building the TWS prototype, and Minoru Kobayashi at NTT for his contribution to implementing ClearBoard. I appreciate the contribution of Prof. Naomi Miyake at Chukyo University and Prof. Jonathan Grudin at University of California Irvine to the observational and experimental phase of this research. I thank George Fitzmaurice and Beverly Harrison at the University of Toronto for their careful comments on an early version of this chapter. Finally, the stimulating discussions with Prof. William Buxton at the University of Toronto on the principles of skill-based design and the ubiquitous media were greatly appreciated.

## REFERENCES

- [Arg75] Argyle, M., *Bodily Communication*. Methuen & Co. Ltd., London, 1975.
- [Bae93] Baecker, R. (Ed.), *Readings in Groupware and Computer-Supported Cooperative Work*. Morgan Kaufmann, San Mateo, 1993.
- [Bly90] Bly, S.A. and Minneman, S.L., Commune: A shared drawing surface. In *Proceedings of COIS '90*, pages 184–192. ACM, New York, 1990.
- [Bly93] Bly, S.A., Harrison, S.R. and Irwin, S., Media spaces: Bringing people together in a video, audio and computing environment. *Communications of the ACM* 36(1):28–47, January 1993.
- [Bri92] Brittan, D., Being there: The promise of multimedia communications. *MIT Technology Review*, pages 42–50, May/June 1992.
- [Bux90a] 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*, pages 11–34. North-Holland, Amsterdam, 1990.
- [Bux90b] Buxton, W., Smoke and mirrors. *Byte*, pages 205–210, July 1990.
- [Bux92] Buxton, W., Telepresence: Integrating shared task and person spaces. In *Proceedings of Graphics Interface '92*, pages 123–129. Morgan Kaufmann, Los Altos, 1992.
- [Bux94] Buxton, W., Living in augmented reality: Ubiquitous media and reactive environment. (unpublished paper).
- [Col93] Coleman, D. (Ed.), *Proceedings of Groupware '93*. Morgan Kaufmann, San Mateo, 1993.
- [Ell91] Ellis, C.A., Gibbs, S.J. and Rein, G.L., Groupware: Some issues and experiences. *Communications of the ACM* 34(1):38–58, January 1991.
- [Far91] Farallon Computing Inc., *Timbuktu 4.0 User's Guide*. Farallon Computing Inc., Emeryville, CA, 1991.
- [Fis90] Fish, R.S., Kraut, R.E. and Chalfonte, B.L., The VideoWindow system in informal communications. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '90*, pages 1–11. ACM, New York, 1990.
- [Fis93] Fish, R.S., Kraut, R.E., Root, R.W. and Rice, R.E., Video as a technology for informal communication. *Communications of the ACM* 36(1):48–61, January 1993.
- [Fos86] Foster, G. and Stefik, M., Cognoter, theory and practice of a collaborative tool. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '86*, pages 7–15. ACM, New York, 1986.
- [Gal90] Galegher, J., Kraut, R. and Egido, C., *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1990.
- [Gav91] Gaver, W., Smith, R. and O'Shea, T., Effective sounds in complex systems: The ARKola simulation. In *Proceedings of Human Factors in Computing Systems, CHI '91*, pages 85–90. ACM Press, New York, 1991.
- [Gav92] Gaver, W., The affordance of media spaces for collaboration. *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '92*, pages 17–24. ACM, New York, 1992.
- [Gre91] Greenberg, S. (Ed.), *Computer-Supported Cooperative Work and Groupware*. Academic Press, London, 1991.
- [Gre92] Greenberg, S., Roseman, M., Webster, D. and Bohnet, R., Issues and experiences designing and implementing two group drawing tools. In *Proceedings of HICSS '92*, pages 139–150. IEEE Computer Society, Los Alamitos, CA, 1992.
- [Gre88] Greif, I. (Ed.), *Computer-Supported Cooperative Work: A Book of Readings*. Morgan Kaufmann, San Mateo, CA, 1988.
- [Gro90] Group Technologies, Inc., *Aspects: The first simultaneous conference software for the Macintosh. Aspects User's Manual*, Group Technologies, Inc., Arlington, VA, 1990.
- [Gru88] Grudin, J., Why CSCW applications fail: Problems in the design and evaluation of organizational interfaces. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '88*, pages 85–93. ACM, New York, 1988.
- [Gru91] Grudin, J., CSCW introduction, *Communications of the ACM*, 34(12):30–34, December 1991.
- [Hol92] Hollan, J. and Stornetta, S., Beyond being there. In *Proceedings of Human Factors in Computing Systems, CHI '92*, pages 119–125. ACM, New York, 1992.

- [Ish90] Ishii, H., TeamWorkStation: Towards a seamless shared workspace. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '90*, pages 13–26. ACM, October 1990.
- [Ish91] Ishii, H. and Miyake, N., Toward an open shared workspace: Computer and video fusion approach of TeamWorkStation. *Communications of the ACM*, 34(12):37–50, December 1991.
- [Ish92a] Ishii, H. and Kobayashi, M., ClearBoard: A seamless medium for shared drawing and conversation with eye-contact. In *Proceedings of Human Factors in Computing Systems, CHI '92*, pages 525–532. ACM Press, May 1992.
- [Ish92b] Ishii, H., Arita, K. and Kobayashi, M., Toward seamless collaboration media: From TeamWorkStation to ClearBoard. *SIGGRAPH Video Review, CSCW '92 Technical Video Program*, Issue 87, Item 6. ACM, New York, 1992.
- [Ish93a] Ishii, H., Arita, K. and Yagi, T., Beyond videophones: TeamWorkStation-2 for narrowband ISDN. In *Proceedings of European Conference on Computer-Supported Cooperative Work, ECSCW '93*, pages 325–340. Kluwer Academic Publishers, Dordrecht, the Netherlands, September 1993.
- [Ish93b] Ishii, H., Kobayashi, M. and Grudin, J., Integration of interpersonal space and shared workspace: ClearBoard design and experiments. *ACM Transactions on Information Systems (TOIS)*, 11(4):349–375, October 1993. (a previous version of this paper was published in the *Proceedings of CSCW '92*, pages 33–42, ACM, November 1992).
- [Lau90] Lauwers, J.C., Joseph, T.A., Lantz, K.A. and Romanow, A.L., Replicated architectures for shared window systems: A critique. In *Proceedings of COIS '90*, pages 249–260. ACM, New York, 1990.
- [Lu91] Lu, I. and Mantei, M., Idea management in a shared drawing tool. In *Proceedings of European Conference on Computer-Supported Cooperative Work, ECSCW '91*, pages 97–112. Kluwer Academic, Dordrecht, the Netherlands, 1991.
- [Lyl93] Lyles, B., Media spaces and broadband ISDN. *Communications of the ACM*, 36(1):46–47, January 1993.
- [Mac99] Mackay, W., Media spaces: Environments for informal multimedia interaction In Beaudouin-Lafon, M. (Ed.), *Computer Supported Cooperative Work*, Trends in Software Series 7:55–82. John Wiley & Sons, Chichester, 1999.
- [Man91] Mantei, M., Baecker, R., Sellen, A., Buxton, W. and Milligan, T., Experiences in the use of a media space. In *Proceedings of Human Factors in Computing Systems, CHI '91*, pages 203–208. ACM Press, New York, 1991.
- [Min91] Minneman, S.L. and Bly, S.A., Managing á trois: A study of a multi-user drawing tool in distributed design work. In *Proceedings of Human Factors in Computing Systems, CHI '91*, pages 217–224. ACM Press, New York, 1991.
- [NTT93] NTT, Seamless Media Design (video). Presented at TED4 KOBE, May 1993. Also presented at CSCW '94 formal video session.
- [Pra99] Atul Prakash, Group editors. In Beaudouin-Lafon, M. (Ed.), *Computer Supported Cooperative Work*, Trends in Software Series 7:103–133. John Wiley & Sons, Chichester, 1999.
- [Str92] Streitz, N., Haake, J., Hannemann, J., Lemke, A., Schuler, W., Schuett, H. and Thuring, M., SEPIA: A cooperative hypermedia environment. In *Proceedings of Conference on Hypertext, ECHT '92*, pages 11–22. ACM Press, New York, 1992.
- [Tang91] Tang, J.C. and Minneman, S.L., VideoDraw: A video interface for collaborative drawing. *ACM Transactions on Information Systems (TOIS)*, 9(2):170–184, April 1991.
- [Tan91] Tanigawa, H., Arikawa, T., Masaki, S. and Shimamura, K., Personal multimedia-multipoint teleconference system. In *Proceedings of INFOCOM '91*, pages 1127–1134. IEEE Communications Society, 1991.
- [Wat90] Watabe, K., Sakata, S., Maeno, K., Fukuoka, H. and Ohmori, T., Distributed multiparty desktop conferencing system: MERMAID. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW'90*, pages 27–38. ACM, New York, 1990.
- [Wei91] Weiser, M., The computer for the twenty-first century. *Scientific American*, pages 94–104, September 1991.
- [Wol92] Wolf, C. and Rhyne, J., Communication and information retrieval with a pen-based meeting support tool. In *Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW '92*, pages 322–329. ACM, New York, 1992.

