Iterative Design of Seamless Collaboration Media: From TeamWorkStation To ClearBoard

Hiroshi Ishii, Minoru Kobayashi, and Kazuho Arita

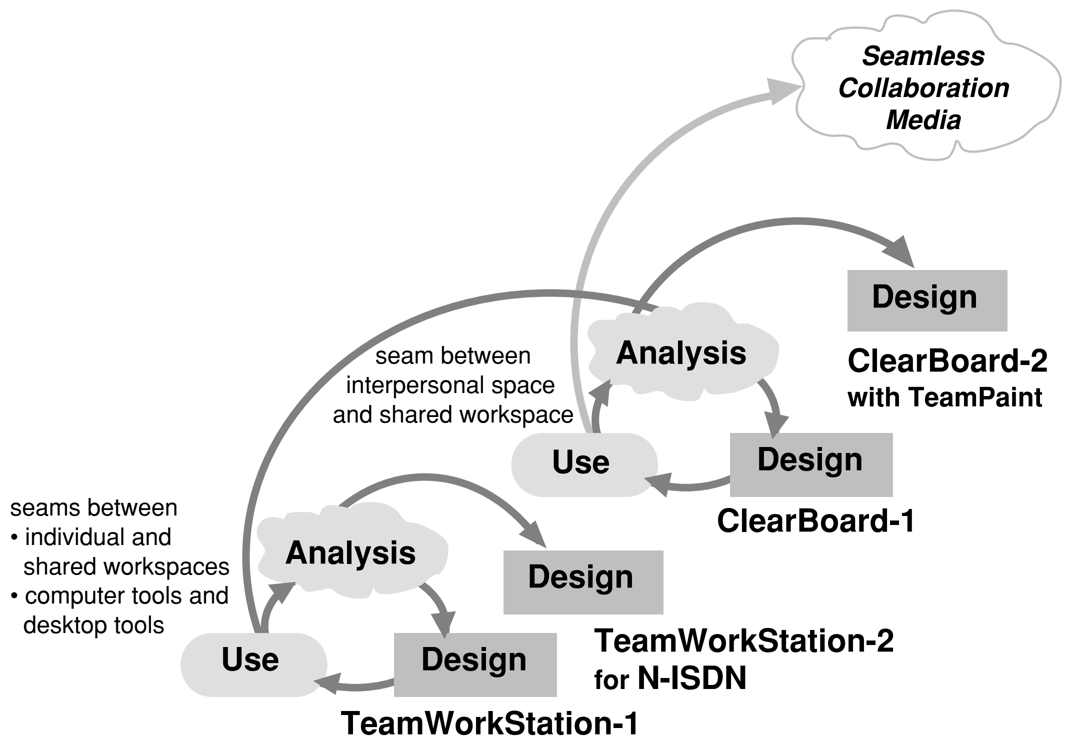


Fig. 1 Progression of our collaboration media design

Computer-Supported Cooperative Work (CSCW) emerged in the middle of the 1980s as an identifiable interdisciplinary research field focused on the role of computer and communication technology to support group work [Baecker 1993; Grudin 1991; Greenberg 1991, Galegher 1990; Greif 1988]. CSCW examines how people work together in groups and how *groupware* technologies can support and change them.

"Groupware" is a label for computer-based systems explicitly designed to support groups of people working together, and growing rapidly as a new application category in the computer industry [Ellis 1991; Coleman 1993]. Familiar groupware examples include electronic mail, bulletin boards, group schedulers, group decision support systems, workflow systems, collaborative authoring tools, screen sharing software, and so on.

Although most of the current groupware systems are devoted to the computational support and designed under the constraint of limited communication bandwidth, the deployment of broadband digital networks open a new future of multimedia collaboration environments that integrate realtime audio and video communication links with computer-based shared workspaces [Brittan 1992; Lyles 1993]. We have been exploring the future of collaboration media that make good use of broadband networks through the iterative design of various CSCW systems. Progression of our collaboration media design is illustrated in Fig. 1[1].

Our CSCW research started in 1988 and was motivated by the study on shared drawing space [Tang 1991] in the Media space[2]. The focus of our research is

1 A video tape that introduces the whole history of this collaboration media design is available as ACM SIGGRAPH video review, Issue 87 [Ishii 1992b].

2 "Media space" originated by the Xerox PARC [Bly 1993] is the environment that integrates video, audio, and computer technologies, allowing individual and groups to work together despite being distributed spatially and temporally. Recent developments include Cruiser (Bellcore) [Fish 1993], RAVE (Rank Xerox EuroPARC)

on the creation of *new application concepts* and the *new usage* of technologies, rather than the development of technologies themselves. Our research has taken a user-centered approach instead of a technology-driven approach, following Buxton's design principle: "*Let's do smart things with stupid technology today, rather than wait and do stupid things with smart technology tomorrow*."[3] For quick prototyping and verification of new application concepts, we have mainly used hybrid networks consisting of analog video networks and digital data networks.

This article introduces the progression of iterative media design from TeamWorkStation-1 and TeamWorkStation-2 to ClearBoard-1 and ClearBoard-2. These systems were designed to support focused realtime collaboration by distributed group members. The key concept behind our iterative design is "seamlessness." Seamless design pursues the following two goals:

1) 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, and we believe the continuity with existing work practices is essential. Groupware that asks users to abandon their acquired skills and to learn a new protocol is likely to encounter strong resistance [Grudin 1988].

2) 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 loads of users as they move dynamically across different spaces. For example, TeamWorkStation was designed to enable smooth transition between individual 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 so that people can use various non-verbal cues such as a partner's gaze direction for smooth focus switching between these two spaces.

TeamWorkStation-1 and Seamless Shared Workspace

People do a lot of their work alone, 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 of them can replace the other. For example, printed materials such as books and magazines are still an indispensable source of information. Therefore, when designing realtime shared workspaces, depending on the task and the media of the information to be shared (paper or computer file), coworkers should be able to choose either computers or desktops, and to switch between them freely. This 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) [Ishii 1990, 1991b].

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 by 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 a variety of traditionally incompatible visual media such as papers and computer documents.

System Architecture of TWS-1

Fig. 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 member, and the other for capturing the desktop surface images and hand gestures. TWS-1 provides two screens. The individual screen is on the left and the shared screen is on the right. These two screens are contiguous in video memory. This multi-screen architecture allows users to move any application program window between the individual and shared screens by just 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. Fig. 3 shows an image of a shared screen. Two users are discussing the system configuration by annotating diagrams in a draw editor by hand.

The first prototype TWS-1 is implemented on Macintosh™ computers to provide small work groups (2-4 members) with a shared workspace. The system architecture of the TWS-1 is illustrated in Fig. 4. In order to connect distributed workstations, an analog video network (NTSC and RGB) was developed and integrated with existing data network (LocalTalk™ network) and voice (telephone) network [Ishii 1991b].

The video network is controlled by a video server that is based on a computer-controllable video switcher and video effecter. 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

---

[Gaver 1992], CAVECAT/Telepresence (University of Toronto) [Mantei 1991; Buxton 1992].

[3] Prof. William Buxton at University of Toronto talked about this design principle in the invited lecture of IFIP WG8.4 conference at Crete in September 1990.

overlaying are redistributed to the shared screens via the video network.

Experience of TWS-1

Through the experimental use of TWS-1, we found that users liked the feature that allowed them to keep using their favorite individual tools, especially papers and pen, while collaborating in a desktop shared workspace. That is, there is no need to master the usage of new sophisticated groupware. The drawback of this overlay approach is that the results of collaboration can not be shared directly. 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 people to mean "data sharing." However, we think it is not required that all the outcomes of the work-in-progress to be directly "manipulable" by all the participants. We seldom felt the necessity of editing 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 come from the respect paid to the ownership of the outcomes. This is a very natural feeling, even in a close collaborative session. The overlay solution provides us with a comfortable 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 workspace, 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, as shown in Fig. 5. 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 (TWS-2).

Fig. 2 Overview of TeamWorkStation-1 prototype

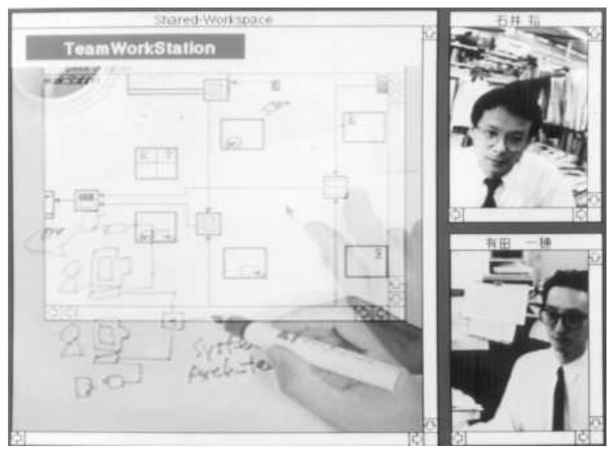


Fig. 3 A shared screen of TeamWorkStation-1

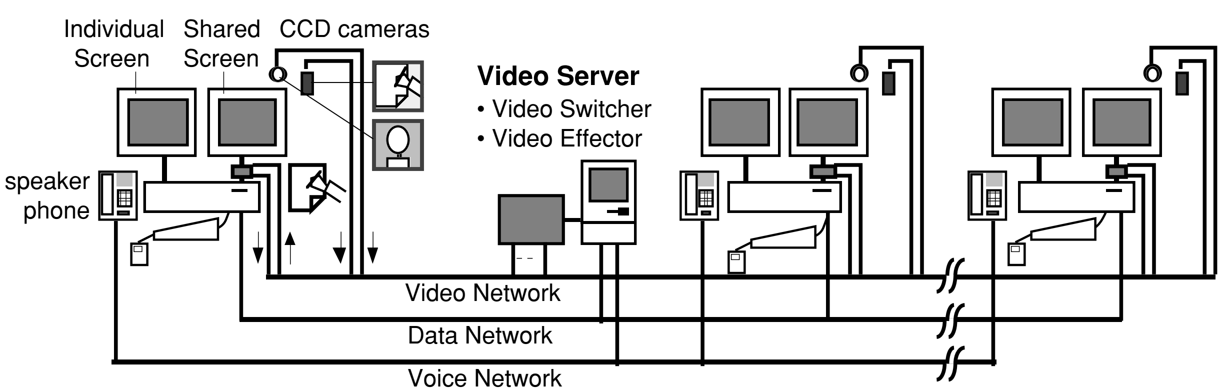


Fig. 4 System architecture of TeamWorkStation-1 prototype

Fig. 5 "Spaghetti" behind TWS-1

TeamWorkStation-2 for N-ISDN

TeamWorkStation-2 (TWS-2) was designed to provide the shared workspace over narrowband ISDN (N-ISDN) Basic Rate Interface (2B+D) and the Primary Rate Interface (H1/D) using CCITT H.261 standard of moving picture transmission [Ishii 1993a]. We chose N-ISDN, especially Basic Rate Interface as the target network because of its wide spread in Japan[4].

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 the more effective use of the normally limited screen space. We found users have little difficulty in selectively viewing either the facial image or the workspace image [Ishii 1991a].

System Architecture of TWS-2

We radically simplified the system architecture. Fig. 6 shows the system architecture of TWS-2. We targeted *dyadic communication* to make the centralized video server unnecessary and 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 unit 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 PC-9801™ computer is required, we can use *screen sharing software* while overlaying desktop video images with the shared computer screen.

Fig. 7 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 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 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. It is quicker to point to the part of remote documents with a finger than with a mouse pointer, and marking and annotating with a pen is also quicker and easier than using computer programs. Even though the documents were stored as Macintosh™ computer files, users preferred to print and share the hard copies (papers) as desktop-overlay rather than using direct computer screen overlay. Based on these observations of TWS-1 use, we concluded that the support of desktop images is more important than the support of computer screens. We decided to make "desktop overlay" the basic service of TWS-2, and to make "computer screen overlay" an option. This decision lead to the one screen architecture of TWS-2.

Fig. 8 and 9 show typical screen images of TWS-2 in use. In Fig. 8, user A and B are discussing the system architecture using hand drawing and gestures. In Fig. 9, teacher A is teaching calligraphy to student B using red ink to correct B's strokes made in black. This calligraphy example demonstrates the important TWS feature that all the collaborators share not only the results of drawing, but also the dynamic process of drawing and gesturing.

Experimental Use of TWS-2

We have tested TWS-2 by connecting our offices in Tokyo, Yokosuka, and Osaka by INS-Net 64 since 1992. We conducted several controlled laboratory experiments as well as the tests for real work outside of laboratories [Ishii 1993a]. Before we started the TWS-2 experiments, many people felt dubious about the ability of INS-Net 64 to support realtime activities because of their previous experience with the jerky displays of

---

4 In April 1988, NTT started an ISDN Basic Rate Interface service (called INS-Net 64) in Japan using the existing metallic cables providing two 64kbps B channels and 16kbps D channel. In June 1989, ISDN Primary Rate Interface service (called INS-Net 1500) was started using optical fiber cables. INS-Net 1500 provides 1.5 Mbps channel at the maximum. Since the introduction of these ISDN services, the number of subscriber lines has grown steadily, and there are more than 200,000 INS-Net 64 subscriber lines in Japan at the beginning of 1994.

videophones. However, the subjects generally commented that they could smoothly interact with their partner and they were absorbed in the task. Although the subjects noticed some delay and jerkiness in the remote desktop video image, these did not hinder users from concentrating on their work. However, all the subjects noted that they could not see the partner's desktop image clearly. This confirmed that the CIF (Common Intermediate Format) (352 pixels/line x 288 lines/picture) standard is definitely insufficient to see the small characters or fine drawings in the remote documents[5].

Beyond Videophone

Videophone and video conferencing are the most typical video applications that use N-ISDN, and they represent the effort toward imitating "being there" that has long been the goal of telecommunication technologies [Hollan 1992]. Realtime video is used just 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 realtime video: the creation of a virtual shared workspace. The main focus of TWS-2 is not on the imitation of face-to-face communication but on the sharing of overlaid desktop images for collaboration.

The experiments to date confirm that TWS-2 has one large advantage over ordinary videophones as the preeminent N-ISDN service. The advantage is due to the bandwidth limitation and human perception. People are especially perceptive to facial expressions. If facial expression is the main means of communication, slight *asynchronism* between the voice and the movement of eye 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 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 experienced with videophone displays. This more static nature of the desktop surface increases the *effective video frame rate*. Although quick hand motions looks jerky, TWS-2 users are happier and more productive than their videophone counterparts since they can visually share objects and work on them.

---

5 If both members need to see fine details, we expect facsimile to be the partner that offsets the weakness of TWS-2 in sharing detailed documents.

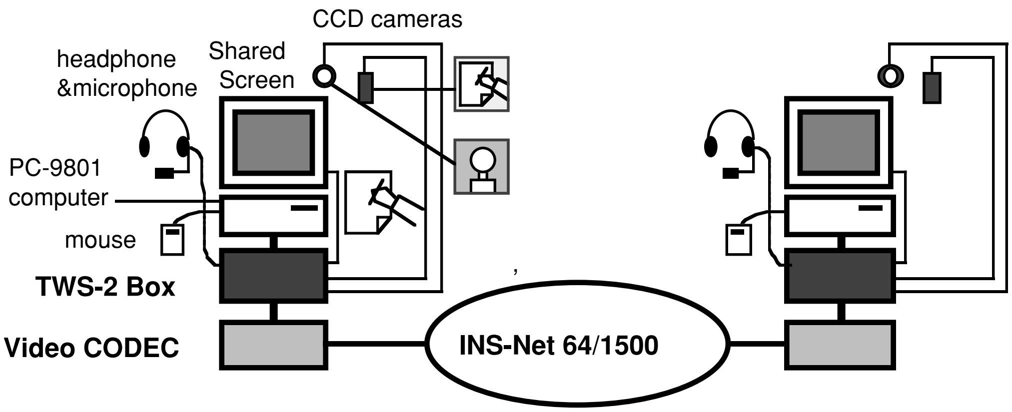
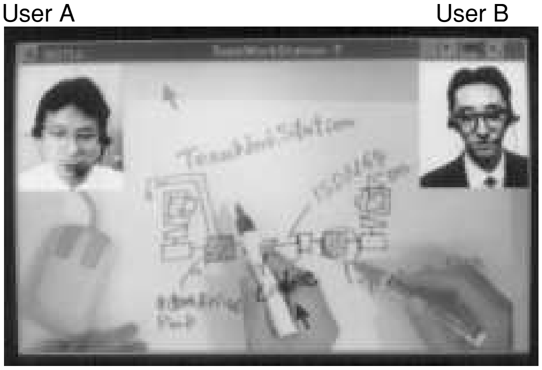


Fig. 6 System Architecture of TeamWorkStation-2 Prototype

CODEC TWS-2 Box PC-9801 Display
CCD cameras

Fig. 7 Appearance of TWS-2 Terminal

User A User B



A's hand B's hand

Fig. 8 Design Session via TWS-2

teacher A

A's hand B's hand student B

Fig. 9 Calligraphy Lesson via TWS-2

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 [Ellis 1991].

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 [Buxton 1992]. The focus of telecommunication technologies such as the videophone and video conferencing has been the creation of *"interpersonal spaces"* that maintain a sense of "telepresence" or "being there" [Hollan 1992] 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 [Tanigawa 1991], MERMAID [Watabe 1990] support both spaces, but they have a major limitation: an arbitrary seam between the shared workspace and the face images. We realized that the problem is not just the superficial physical discontinuity of spatially separated windows. Absent are the *non-verbal cues* that would enable a smooth shift of focus between the shared workspace and the partner's face image. Current groupware and video conferencing technologies do not support these cues.

Lack of eye contact is another problem of TWS (Fig. 3, 8, 9). 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 (Fig. 10). 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.

ClearBoard Metaphor

The key metaphor of ClearBoard design is "talking through and drawing on a big transparent glass board." Fig. 11 shows "ClearBoard-0" which is the simple mockup of this ClearBoard concept for co-located pair 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 head movement to switch focus between the drawing surface and the partner's face than is needed in either the whiteboard or desktop environment. 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.

Design of ClearBoard-1

Fig. 12 shows ClearBoard-1; our first prototype to support remote collaboration [Ishii 1992a]. 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.

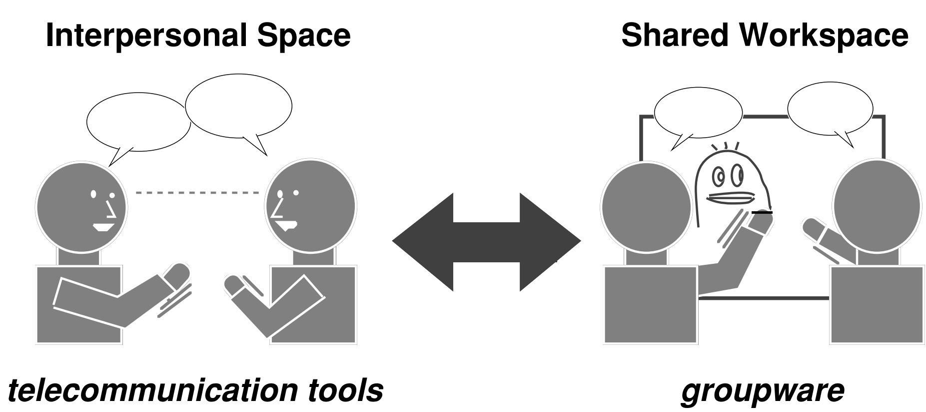


Fig. 10 Design goal: seamless integration of interpersonal space and shared workspace

Fig. 11 A simple mockup of ClearBoard metaphor: ClearBoard-0

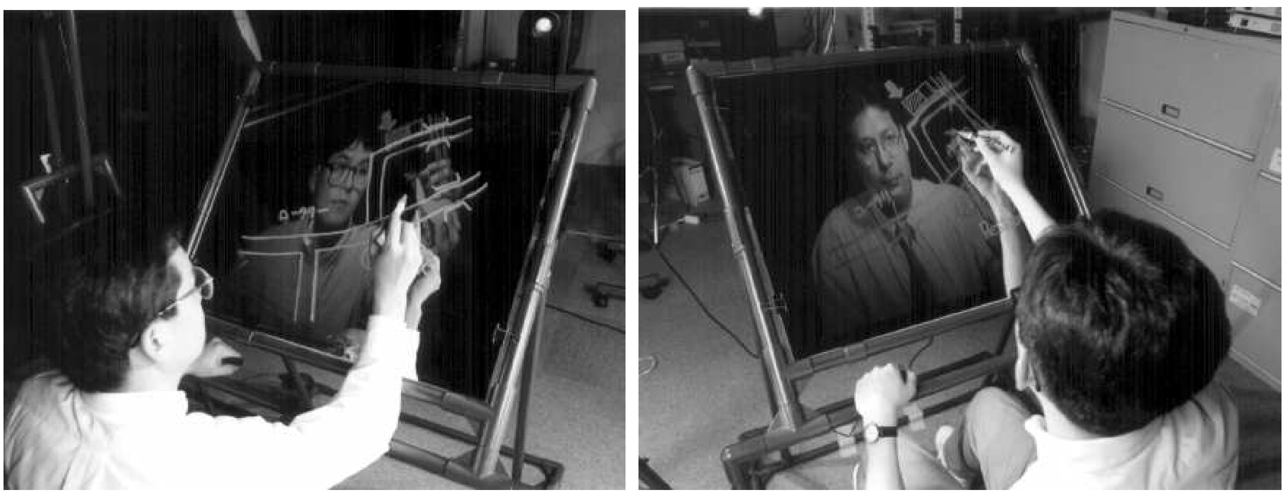


Fig. 12 ClearBoard-1 in use

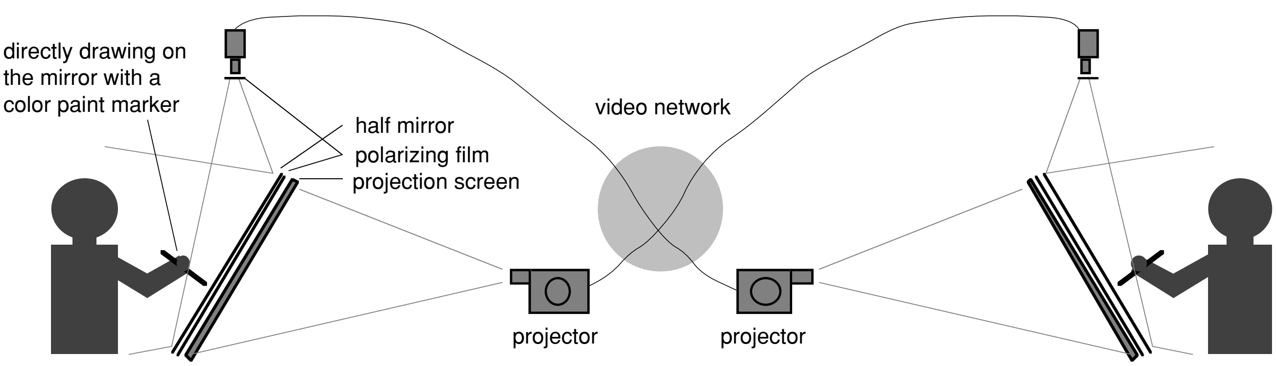


Fig. 13 Architecture of ClearBoard-1 prototype

In order to implement the remote version of ClearBoard, we devised the system architecture called "drafter-mirror" architecture illustrated in Fig. 13. 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 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 common orientation of the drawing space. The partner can draw directly over this transmitted video image[6].

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 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 what 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." [Ishii 1993b] This experiment confirmed that it is easy for the instructor to say which side of the river the student was gazing at, and this information was quite useful in understanding the student's thinking process and passing 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 generalized and is a more important notion. 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.

Design of ClearBoard-2

In using this ClearBoard-1 prototype, we found several problems. The projected video image of a drawing is not sufficiently clear. Lack of recording capabilities is an obstacle to reusing the work results. To overcome these problems in ClearBoard-1, we decided to design a new computer-based prototype, "ClearBoard-2" [Ishii 1993b]. Instead of using color paint markers, ClearBoard-2 provides users with "TeamPaint", a multiuser computer-based paint editor and digitizer pen.

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 Fig. 14.

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[7].

ClearBoard-2 System and Its Use

Using TeamPaint, transparent digitizer sheets, and electronic pens, we implemented the computer-based prototype, ClearBoard-2. Fig. 15 shows the ClearBoard-2 prototype in use, and Fig. 16 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 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.

---

[6] This shared video drawing technique, which allows remote partners to draw directly over the video image of their coworkers' drawing surface, was originally demonstrated in VideoDraw [Tang 1991].

[7] Awareness based on such a small tele-pointer may have limitations. Actual hand gestures have much more power of expression, and with ClearBoard-2, the real hand and pen gesture images that lie behind the tele-pointer augment the awareness provided by TeamPaint.

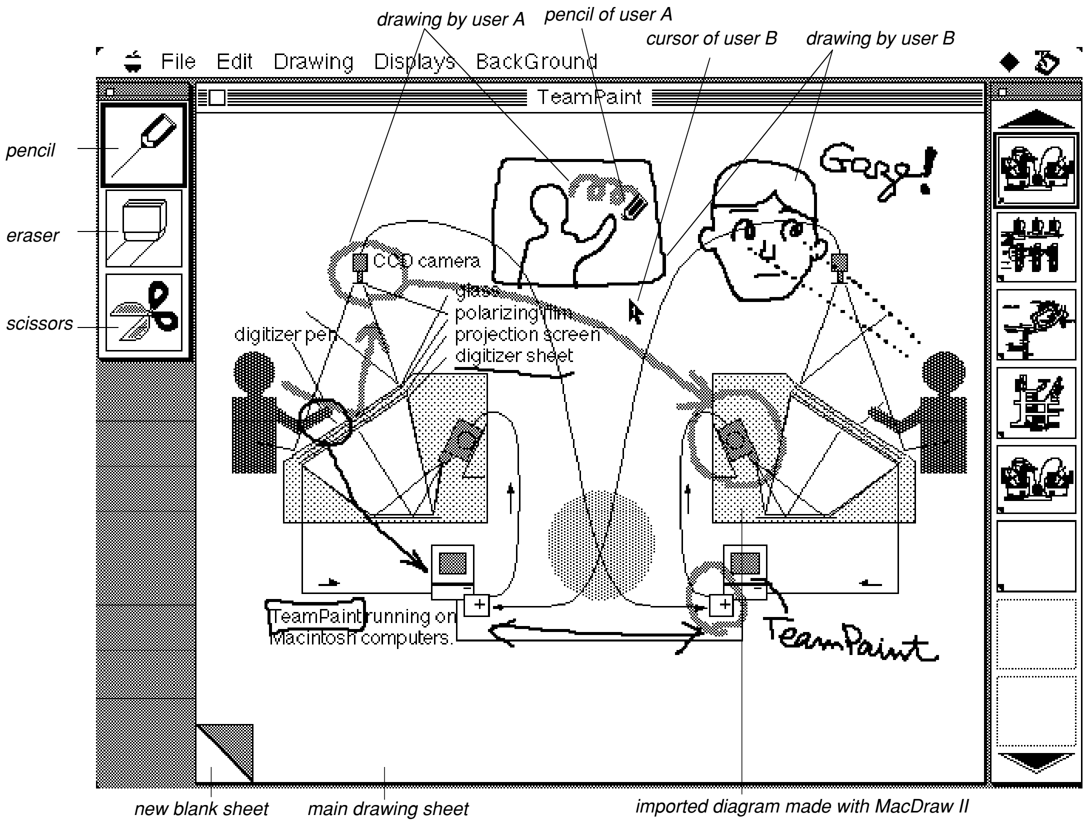


Fig. 14 A Screen of TeamPaint

Fig. 15 ClearBoard-2 in use

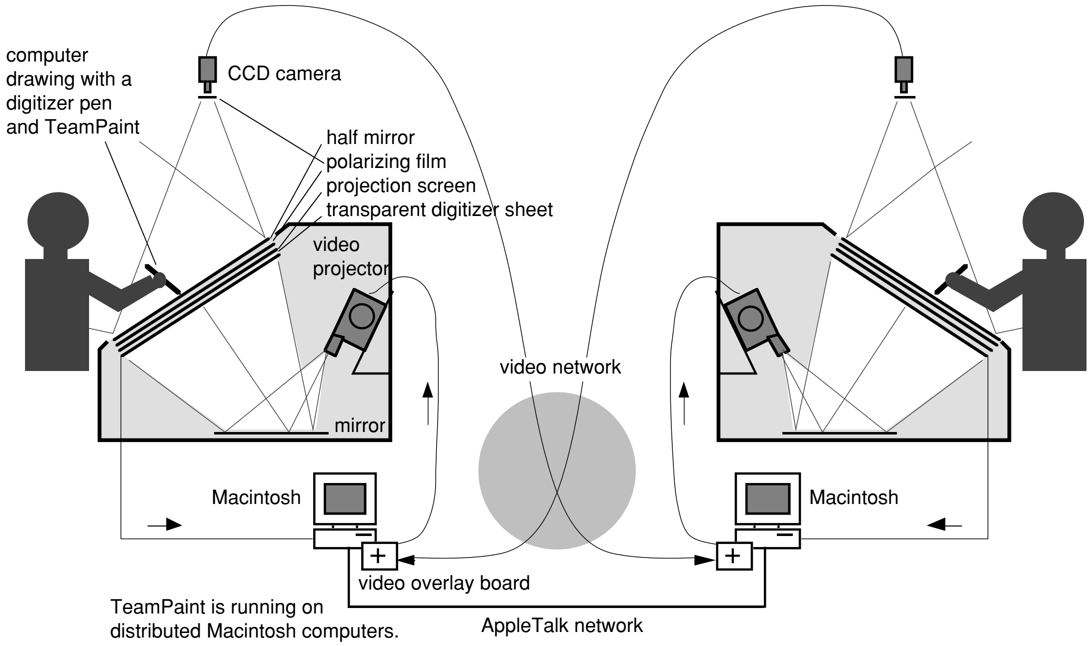


Fig. 16 System architecture of ClearBoard-2

The use of RGB video and the chroma-keying overlay technique does increase image clarity. Furthermore, the capability of recording results and reusing 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 as in ClearBoard-1. One can easily tell which object on the TeamPaint screen the partner is looking at.

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. We feel ClearBoard-2 is the first system that fully succeeds in naturally integrating the technology of computer-based groupware with that of video conferencing. 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 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" [Hollan 1992]. We expect ClearBoard to be useful both as a collaboration medium and as a vehicle to investigate the nature of dynamic human interaction.

Summary and Future Work

We have presented an evolution of our collaboration media design from TWS-1 to TWS-2 to ClearBoard-1 to ClearBoard-2. TeamWorkStation (TWS) demonstrates a new usage of realtime video for collaboration, by providing distributed users with a seamless shared workspace. Using the translucent video overlay technique, real-time information such as hand gestures and hand-written 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 and shared workspaces. TWS-2, redesigned version which uses N-ISDN Basic Rate Interface, demonstrated the advantage of this application over ordinary videophone 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 found the new 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 [8]. Our 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 kids. NTT's vision video, "Seamless Media Design" [NTT 1993] illustrates our expectation of future collaboration media based on the ClearBoard concept. Fig. 16 shows a snapshot of a session by kids from the vision video.

[a picture from the video is to be inserted.]

Fig. 16 Kids drawing robots via future ClearBoard

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 the 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) [Lyles 1993] 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 user 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 blue print for a new building may require more bandwidth for higher resolution of shared documents compared to a TWS meeting with shared sheets of blank paper for free hand drawing. Competitive negotiation tasks may require both higher frame rate and resolution to read your colleagues 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.

From Multimedia to Seamless Media

"Multimedia" is now becoming a big buzz word in 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 limitation of human cognitive capability. A variety of media (such as text, video, audio, graphics) and services (on demand video, video conferencing, electronic newspaper) are becoming available through a single powerful computer on the desktop and a broadband communication network called 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, but 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.

Toward Ubiquitous Media

We expect 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 [Buxton 1994; Weiser 1991]. Many challenges exist to achieve a seamless

8 The importance of skill-based design was originally proposed by Prof. William Buxton at the University of Toronto in the opening panel of ACM CHI '92.

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.

Acknowledgments

We would like to thank Masaaki Ohkubo and Takashi Yagi at NTT for the contribution to building the TWS-1 and TWS-2 prototype respectively. We 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. We thank George Fitzmaurice at the University of Toronto for his careful comments to the draft of this paper. Finally, the stimulating discussions with Prof. William Buxton at the University of Toronto on the principles of skill-based design and the new concept of ubiquitous media were greatly appreciated.

References

[Baecker 1993] Baecker, R. (Ed.) Readings in Groupware and Computer-Supported Cooperative Work. Morgan Kaufmann, San Mateo, 1993.

[Bly 1993] Bly, S.A., Harrison, S.R., and Irwin, S. Media Spaces: Video, Audio, and Computing. *Communications of the ACM Vol. 36, No. 1* January 1993, 28-47.

[Brittan 1992] Brittan, D. Being There: The Promise of Multimedia Communications. *MIT Technology Review*, May/June 1992, pp. 42-50.

[Buxton 1990] Buxton, W. Smoke and Mirrors. *Byte,* July 1990, pp. 205-210.

[Buxton 1992] Buxton, W. Telepresence: Integrating Shared Task and Person Spaces. *Proceedings of Graphics Interface '92,* Morgan Kaufmann Publishing, Los Altos, 1992, pp. 123-129.

[Buxton 1994] Buxton, W. Ubiquitous Media and the Active Office. *Communications of the ACM*, ACM, , 1994 (to appear)

[Coleman 1993] Coleman, D. (Ed.) Proceedings of Groupware '93, Morgan Kaufmann Publishers, San Mateo, 1993.

[Ellis 1991] Ellis, C.A., Gibbs, S.J. and Rein, G.L. Groupware: Some Issues and Experiences. Communications of the ACM, Vol. 34, No. 1, January 1991, pp. 38-58.

[Galegher 1990] Galegher, J., Kraut, R. and Egido, C. Intellectual teamwork: Social and technological foundations of cooperative work. Lawrence Erlbaum Associates, Hillsdale, NJ, 1990.

[Greenberg 1991] Greenberg, S. (Ed.) Computer-Supported Cooperative Work and Groupware. Academic Press, London, 1991.

[Greif 1988] Greif, I. (Ed.) Computer-Supported Cooperative Work: A Book of Readings. Morgan Kaufmann, San Mateo, Calif., 1988

[Grudin 1988] Grudin, J. Why CSCW Applications fail: Problems in the design and Evaluation of organizational Interfaces. Proceedings of CSCW '88, ACM, New York, 1988, pp. 85-93.

[Grudin 1991] Grudin, J. CSCW Introduction, *Communications of the ACM,* Vol. 34, No. 12, December 1991, pp. 30-34.

[Hollan 1992] Hollan, J. and Stornetta, S. Beyond Being There. *Proceedings of CHI '92*, ACM, New York, 1992, pp. 119-125.

[Ishii 1990] Ishii, H., TeamWorkStation: Towards a Seamless Shared Workspace, Proceedings of *CSCW '90*, ACM, October 1990, pp. 13-26.

[Ishii 1991a] Ishii, H. and Arita, K., ClearFace: Translucent multiuser interface for TeamWorkStation, Proceedings of *ECSCW '91*, September 1991, pp. 163-174.

[Ishii 1991b] Ishii, H. and Miyake, N., Toward an Open Shared Workspace: Computer and Video Fusion Approach of TeamWorkStation, *Communications of the ACM* , ACM, December 1991, pp. 37-50.

[Ishii 1992a] Ishii, H. and Kobayashi, M., ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye-Contact, Proceedings of *CHI '92* , ACM, May 1992, pp 525-532.

[Ishii 1992b] 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.

[Ishii 1993a] Ishii, H., Arita, K., and Yagi, T. Beyond Videophones: TeamWorkStation-2 for Narrowband ISDN. Proceedings of *ECSCW '93,* Kluwer Academic Publishers, Dordrecht, Netherlands, September 1993, pp. 325-340.

[Ishii 1993b] Ishii, H., Kobayashi, M. and Grudin, J., Integration of Interpersonal Space and Shared Workspace: ClearBoard Design and Experiments, *ACM Transactions on Information Systems (TOIS),* Vol. 11, No. 4, October 1993, ACM, New York, pp. 349-375 (previous version of this paper was published in the *Proceedings of CSCW '92*, ACM, November 1992, pp. 33-42).

[Lyles 1993] Lyles, B., Media Spaces and Broadband ISDN, CACM, Vol. 36, No. 1, January 1993, ACM, New York, pp. 46-47.

[NTT 1993] NTT. Seamless Media Design (video). presented at TED4 KOBE, May 1993 (to be submitted to CSCW '94 formal video session).

[Tang 1991] Tang, J.C., and Minneman, S.L. VideoDraw: A Video Interface for Collaborative Drawing. *ACM Transactions on Information Systems (TOIS),* Vol 9, No. 2, April 1991, New York, ACM Press, pp. 170-184.

[Tanigawa 1991] Tanigawa, H., Arikawa, T., Masaki, S., and Shimamura, K. Personal Multimedia-Multipoint Teleconference System. *Proceedings of INFOCOM '91,* IEEE Communications Society, 1991, pp. 1127-1134.

[Watabe 1990] Watabe, K., Sakata, S., Maeno, K., Fukuoka, H., and Ohmori, T. Distributed Multiparty Desktop Conferencing System: MERMAID. *Proceedings of CSCW '90,* ACM, New York, 1990, pp. 27-38.

[Weiser 1991] Weiser, M. The Computer for the Twenty-First Century. Sci. Am. September 1991, pp. 94-104.

CR Categories and Subject Descriptors:

H.4.3 [**Information Systems Application**]: Communications Applications----- *Computer conferencing and teleconferencing*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems-----*Video*; H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces---- - *Synchronous interaction*

General Terms: Design, Human Factors

Additional Key words:

CSCW, realtime collaboration, multimedia groupware, shared workspace, interpersonal space

About the Author:

HIROSHI ISHII is a visiting assistant professor at Computer Systems Research Institute of The University of Toronto, and a senior research engineer, supervisor of CSCW research project in NTT Human Interface Laboratories in Japan. He is currently an Associate Editor of ACM TOIS (Transactions on Information Systems). His research interests focus on human - computer interaction (HCI), CSCW, and the role of CSCW technologies in cross-cultural communication.
Author's Present Address: Until the end of May 1994: Computer Systems Research Institute, The University of Toronto, 6 King's College Road, Room 398, Toronto, Ontario, Canada M5S 1A1; email: ishii.chi@xerox.com
After June 1994: NTT Human Interface Laboratories, 1-2356 Take, Yokosuka-Shi, Kanagawa, 238-03 Japan; email: ishii.chi@xerox.com

MINORU KOBAYASHI Minoru Kobayashi is a researcher of NTT Human Interface Laboratories. He developed ClearBoard-1 and 2. Current research interests include HCI, CSCW, video-based groupware design and new input/output devices for visual communication tools.

KAZUHO ARITA is a research engineer of NTT Human Interface Laboratories in Japan. He is engaged in CSCW research and development of TeamWorkStation-2 for N-ISDN.

Authors' Present Address: NTT Human Interface Laboratories, 1-2356 Take, Yokosuka-Shi, Kanagawa, 238-03 Japan; email: ishii.chi@xerox.com, {minoru, arita}@aether.ntt.jp

Glossary

ATM
Asynchronous Transfer Mode

B-ISDN
Broadband Integrated Services Digital Network

CSCW
"Computer-Supported Cooperative Work (CSCW)" emerged in the middle of the 1980s as an identifiable interdisciplinary research field focused on the role of computer and communication technology in group work. CSCW examines how people work together in groups and how *groupware* technologies can support and change them.

Groupware
"Groupware" is a label for computer-based systems explicitly designed to support groups of people working together. Familiar groupware examples include electronic mail, bulletin boards, group schedulers, group decision support systems, workflow systems, collaborative authoring tools, screen sharing software, and so on.

CODEC
Equipment to code and decode the video and audio data for the transmission through digital network.

H.261:
H.261 is the video compression standard recommended by CCITT (International Telegraph and Telephone Consultative Committee) as the compression algorithm for transmitting moving pictures. TWS-2 uses CIF (Common Intermediate Format) whose resolution is 352 pixels/line x 288 lines/picture.

ISDN:
Integrated Services Digital Network

INS-Net 64:
INS-Net 64 service provides two 64kbps information channels (B channels) and one 16kbps signaling channel (D channel). Using the existing metallic cable, INS-Net 64 service requires only the addition of a compact DSU (Digital Service Unit).

INS-Net 1500:
INS-Net 1500 service offers digital communications at 64kbps, 384kbps, and 1.536Mbps. This service is accessed by means of optical fiber cable and a DSU (Digital Service Unit).