

Carpeno: Interfacing Remote Collaborative Virtual Environments with Table-Top Interaction

HOLGER REGENBRECHT

University of Otago, New Zealand

Information Science, P.O. Box 56, Dunedin

holger@infoscience.otago.ac.nz

Phone: ++64 3 479 8322

Fax: ++64 3 479 8311

MICHAEL HALLER

Upper Austria University of Applied Sciences, Austria

JOERG HAUBER & MARK BILLINGHURST

University of Canterbury, New Zealand

Abstract

Creativity is enhanced by communication and collaboration. Thus, the increasing number of distributed creative tasks requires better support from computer-mediated communication and collaborative tools. In this paper we introduce “Carpeno”, a new system for facilitating intuitive face-to-face and remote collaboration on creative tasks.

Normally the most popular and efficient way for people to collaborate is face-to-face, sitting around a table. Computer augmented surface environments, in particular interactive table-top environments, are increasingly used to support face-to-face meetings. They help co-located teams to develop new ideas by facilitating the presentation, manipulation, and exchange of shared digital documents displayed on the table-top surface. Users can see each other at the same time as the information they are talking about. In this way the task space and communication space can be brought together in a more natural and intuitive way. The discussion of digital content is redirected from a computer screen, back to a table that people can gather around.

In contrast, Collaborative Virtual Environments (CVE) are used to support remote collaboration. They frequently create familiar discussion scenarios for remote interlocutors by utilizing room metaphors. Here, virtual avatars and table metaphors are used, where the participants can get

together and communicate with each other in a way that allows behaviour that is as close to face-to-face collaboration as possible.

The Carpeno system described here combines table-top interaction with a CVE to support intuitive face-to-face and remote collaboration. This allows for simultaneous co-located and remote collaboration around a common, interactive table.

Introduction

In recent years computing and communication has become tightly connected so it is easier than ever before for remote teams to work together. Despite this, current remote collaborative tools do not support the easy interchange of ideas that occur in a face to face brainstorming session. In this case people are able to use speech, gesture, gaze, interaction with real objects and other non-verbal cues to rapidly explore different ideas. In addition, there is a need to provide technology that can capture and enhance face to face meetings, such as digital whiteboards and interactive tables.

The central question that we are interested in exploring is: how can we create a computer supported environment which enhances face-to-face collaboration while at the same time allowing remote team members to work as closely together as if they were all sitting around a single real table.

A tool dedicated to group processes has to support the inherent requirements of a creative environment (Kelly, 2001):

- The group members have to be able to communicate their ideas verbally and non-verbally, so they can build on top of each other's ideas.
- Group members need to be able to visualize ideas through use of sketching, image presentation and document sharing
- Group members need to be able to work with real world objects, including creating new or modify objects and showing examples to others.

The tool to be developed has to deal with three elements: creative people working in a creative space focusing on the creative task. Creative people are the target users, such as designers, and architects, who work in domains requiring original idea generation. The creative space is an environment which should be as close as possible to a face-to-face situation, which generally prove to be the most creative settings. Creative tasks are those where the goal is divergent rather than convergent thinking and where group result is supposed to be better than any individual outcome.

These requirements are challenging, however in this paper we present a prototype system that has many of the elements of an ideal interface for supporting face to face and remote collaboration. In the next section we review related work from earlier research in enhancing face to face collaboration and enabling remote collaboration. Then we describe two of our earlier prototype systems, cAR/PE! and Coeno, and our current integrated system, Carpeno, which uses elements from both of these prototypes. Finally we present an exploratory usability study which evaluates the Carpeno prototype and gives some directions for future research.

Related Work

Enhancing Face-To-Face Collaboration

Early attempts at computer enhanced face-to-face collaboration involved conference rooms in which each participant had their own networked desktop computer that allowed them to send text or data to each other. However, these computer conference rooms were largely unsuccessful partly because of the lack of a common workspace (Inkpen, 1997).

An early improvement was using a video projector to provide a public display space. For example the Colab room at Xerox PARC (Stefik et al., 1987) had an electronic whiteboard that any participant could use to display information to others. The importance of a central display for supporting face-to-face meetings has been recognized by the developers of large interactive commercial displays (such as the SMARTBoard DViT¹).

In normal face-to-face conversation, people are able to equally contribute and interact with each other and with objects in the real world. However with large shared displays it is difficult to have equal collaboration when only one of the users has the input device, or the software doesn't support parallel input. In recent years Stewart et al. coined the term Single Display Groupware (SDG) to describe groupware systems which support multiple input channels coupled to a single display (Stewart et al., 1999). They have found that SDG systems eliminate conflict among users for input devices, enabling more work to be done in parallel by reducing turn-taking, and strengthening communication and collaboration.

¹ <http://www.smarttech.com/>

In general, traditional desktop interface metaphors are less usable on large displays. For example, pull down menus may no longer be accessible, keyboard input may be difficult, and the mouse requires movement over large distances (Cao & Balakrishnan, 2004). A greater problem is that traditional desktop input devices do not allow people to use free-hand gesture or object-based interaction as they normally would in face-to-face collaboration. Researchers such as Ishii and Ullmer (Ishii, 1997) have explored the use of tangible object interfaces for tabletop collaboration while Streitz et al. (Streitz, 2003) use natural gesture and object based interaction in their i-Land smart space. In both cases people find the interfaces easy to use and a natural extension of how they normally interact with the real world.

In many interfaces there is a shared projected display visible by all participants; however, collaborative spaces can also support private data viewing. In Rekimoto's Augmented Surface interface (Rekimoto, 1999), users are able to bring their own laptop computers to a face-to-face meeting and drag data from their private desktops onto a table or wall display area. They use an interaction technique called *hyper-dragging* which allows the projected display to become an extension of their own personal desktop. Hyper-dragging allows users to see the information their partner is manipulating in the shared space, so it becomes an extension of the normal non-verbal gestures used in face-to-face collaboration. In this way the task space becomes a part of the personal space.

Enabling Remote Collaboration

Although being in one place and talking to another person face to face can be considered the gold standard for collaboration, it is not always possible, economical, or otherwise desirable for people to come together in the same location. In that case they alternatively rely on teleconferencing systems that support effective collaboration at a distance.

Many researchers from the fields of CSCW (Computer Supported Cooperative Work), HCI (Human Computer Interaction) (Gutwin & Greenberg, 1995), (Sellen, 1995) and Social Psychology (Short et al., 1976) have explored the complex issues around distant communication and remote collaboration. They have tried to understand how systems for remote collaboration should be designed to mediate human activities in a way that allows people at a distance to accomplish tasks with

the same efficiency and satisfaction as if being co-located - ideally even going beyond that (Hollan & Stornetta, 1992).

In that context, videoconferencing (VC) technology has always played and still plays an increasingly important role as it provides a rich communication environment that allows the real-time exchange of visual information including facial expression and hand gestures. A growing number of organisations nowadays use advanced video based collaboration-networks like for example the AccessGrid², or Halo³ system developed by HP for group-to-group meetings on a daily basis. Although the installation and operation costs for these systems seem high, they still prove effective at supporting tasks over a distance, thus making travel redundant. However, although systems like these are capable of producing videos with high grade audio and image quality, a remote encounter for people in front of the cameras often feels rather formal and artificial. The spontaneity and natural interaction that we take for granted in face to face meetings is inhibited by the absence of spatial cues (such as eye-contact), by the lack of a shared social and physical context, and by a limited possibility for informal communication. In fact, as various studies have proven, people's communication behaviour while being connected through a standard audio-video link more closely resembles that of people talking over a phone than of people talking from face to face. [2] [1]. While this might not greatly affect tasks that involve the exchange and the presentation of existing information and documents, it does have a negative impact on tasks of a more creative nature.

In an attempt to simulate traditional face-to-face meetings more closely and eventually overcome the formal and mediated character of standard videoconferencing interfaces, various three-dimensional metaphors have been developed in videoconferencing applications. Early work introduced spatially positioned video and audio streams into the conferencing space (FreeWalk (Nakanishi et al., 1998), Gaze (Vertegaal, 1999), VIRTUE (Kauff & Schreer, 2002)), but without the addition of virtual content to be discussed in such a meeting. In contrast, SmartMeeting⁴ provides a highly realistic conference

² <http://www.accessgrid.org/>

³ http://www.hp.com/hpinfo/newsroom/feature_stories/2005/05halo.html

⁴ <http://www.smartmeeting.com/>

environment with virtual rooms with chairs, whiteboards, multi-media projectors, and even an interactive chessboard, but without spatially placed video representations of the participants. AliceStreet⁵ makes use of a similar concept, although with a more minimalist virtual room design, but the participants are represented here as rotating video planes sitting around a virtual table at fixed positions and watching each other or a shared presentation screen capable of displaying presentation slides.

The common goal of all of these approaches is to improve the usability of remote collaboration systems by decreasing the artificial character of a remote encounter.

Mixed Presence Groupware

Systems that support multiple simultaneous users interacting on a single shared display are categorized as Single Display Groupware (SDG) (Stewart, 1999). If a shared visual workspace also supports distributed participants in real-time, one can label such a system as Multiple Presence Groupware (MPG) (Tang et al., 2004; see also Ashdown & Robinson, 2005). If placed into a place/time groupware matrix (see figure 1) it spans over the two places segments while still being synchronous.

	same place	different place
same time	face-to-face collaboration e.g. Coeno	real-time remote collaboration e.g. cAR/PE!
different time	co-located on-going work e.g. message boards	asynchronous distrib. work e.g. email work

MPG

Figure 1: Mixed Presence Groupware in place/time matrix

Tang et. al. identified only few MPG systems to date, a CAVE-like environment by SICS (Touch Desktop), Microsoft's Halo, a split screen environment for the

⁵ <http://www.alicestreet.com/>

Xbox, and two video-overlying systems without spatial arrangements of the participants. They found two main problems in using MPG systems: (1) Display disparity: considering the appropriate arrangement of persons and artefacts when using a mix of horizontal and vertical displays and (2) Presence disparity: the perception of the presence of others depending on whether s\he is co-located or remote. In our research presented in this article we will address both problems and try to find (partial) solutions.

System Concepts Used

Our approach is novel in that it combines and integrates several vital features found in other earlier work:

- We make use of a horizontal, interactive workspace to support creative group processes in a natural way and allow remote group members to be part of that process avoiding presence disparities.
- We combine interfaces of the remote and co-located worlds in a natural and easy-to-use way.
- We provide a system seamlessly combining a vertical and horizontal display system in a way that minimizes display disparities.
- We integrate the task space (data) within the work space (table environment) providing both with a task to focus on and a creative atmosphere.
- We offer private and public workspaces at different levels for all group members regardless of their location.

In the following we present in brief our earlier existing systems and how we combined them to create a novel collaborative environment.

3D Teleconferencing System: cAR/PE!

cAR/PE! is a teleconferencing system used with commonly available equipment: a PC with a web camera and a headset. It is designed for small group collaboration between Internet networked computers and it integrates data distribution and presentation with communication capabilities. cAR/PE! simulates a face-to-face meeting in a room and therefore uses the metaphor of a three-dimensional conference room (Regenbrecht et al., 2004).

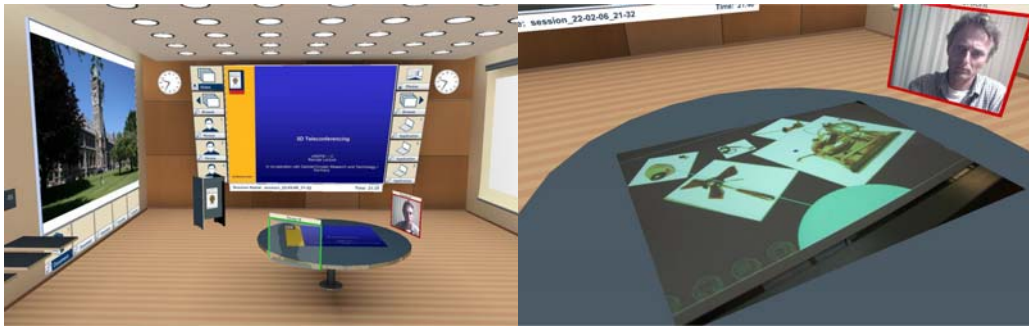


Figure 2: Views (screenshots) into the cAR/PE! room

All participants meet in this room and are represented by video avatars. The virtual room is “furnished” with a meeting table and several presentation screens to be used in a way as close as possible to a real world meeting (see figure 2). The participants can freely move around within this room, can place slides, movies, or pictures on the virtual screens or on the table, can share remote computer screens in an interactive way, and can put three dimensional virtual models onto the table to be discussed with others. The person’s movement within the room is visible to all other participants easing gaze and workspace awareness. This awareness is further supported by the provision of three-dimensional sound (in particular to hear others from the right direction even they are not in the current field of view).

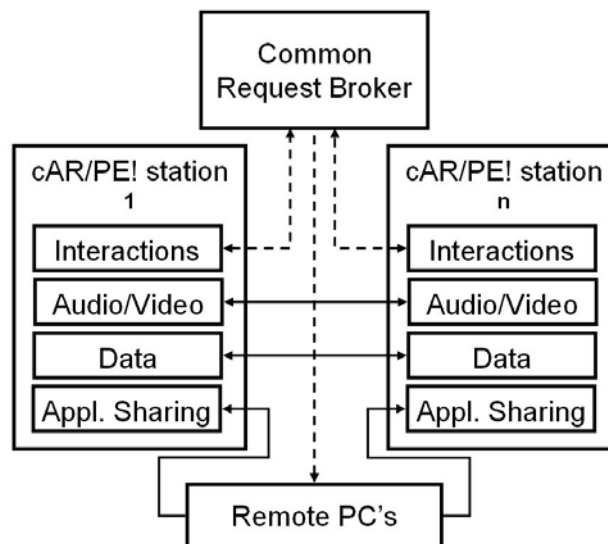


Figure 3: cAR/PE! connection scheme

From a technological point of view, cAR/PE! stations are connected via standard Internet as shown in figure 3. Up to six stations can be connected forming one virtual meeting space. The maximal number of stations depends on the bandwidth available and with standard ADSL connections three stations can be used with a good overall quality. All audio and video streams as well as the data distribution are implemented point-to-point, mainly for security reasons. All interactions

occurring in a session (e.g. the movement of the participants within the room or changing slides on the virtual projection screen) are sent to a common request broker, which delivers the results to all stations. Supplemental remote computers can be connected to this cAR/PE! network. The content of the displays of these computers is displayed within the virtual cAR/PE! environment and can be operated interactively from within the meeting room.

Given these capabilities, the cAR/PE! system allows for synchronous collaboration over a distance while trying to maintain the metaphor of a traditional face-to-face meeting. Remotely located participants are able to focus on their task and data (shared place) and to communicate in a natural way (shared space), because of the integration of both domains: data and communication. The system has been used in pilot installations in industry and academia and usability and social presence successfully evaluated with hundreds of subjects (Regenbrecht et al., 2004, Hauber et al., 2005, Hills et al., 2005).

Some desired interface functionality cannot be supported yet, because of the technology used, or the inherent limitations of this dedicated distant communication and collaboration tool. For instance, by its very nature tangibility input is not supported by any means. Users operate the system using a traditional mouse and therefore all interactions are virtual. To visualize ideas in a real world scenario one would probably use paper and pen or a whiteboard, in a mouse operated virtual room this is inconvenient and less natural. In addition, co-located collaboration and the transmission of most non-verbal cues is poorly supported, even when used in combination with a projection system.

Co-located Table-top System: Coeno

Collaborative table-top setups are becoming increasingly popular for creative tasks. Coeno, is a collaborative table-top environment that is designed for brainstorming and discussion meetings. In Coeno, we particularly focus on a novel ubiquitous environment for creative sketching, drawing, and brainstorming (cf. Figure 4).

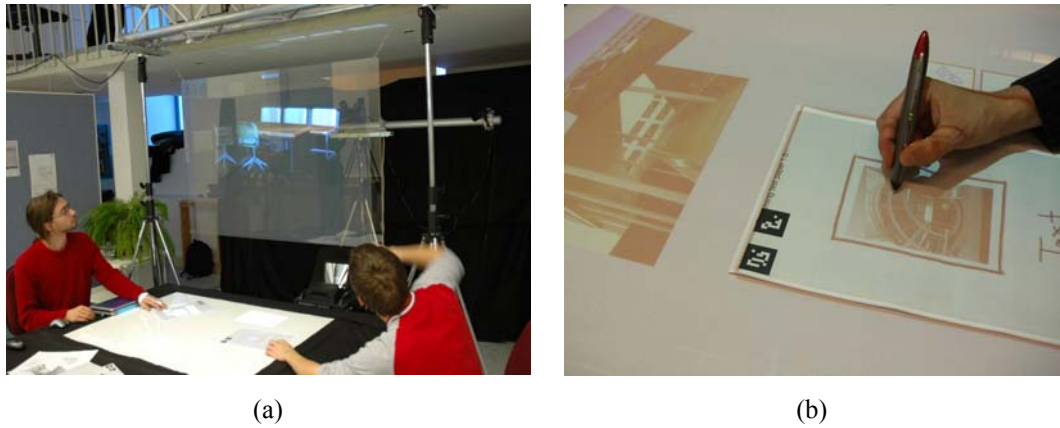


Figure 4: People can discuss and brainstorm by directly interacting with the table and presenting their results on a rear-projection screen (a). Moreover, we support natural input devices (e.g. digital pens) (b).

The application incorporates multiple devices and novel interaction metaphors supporting content creation in an easy-to-use environment. Our installation offers a cooperative and social experience by allowing multiple face-to-face participants to interact easily around the shared workspace, while also having access to their own private information space and a public presentation space.

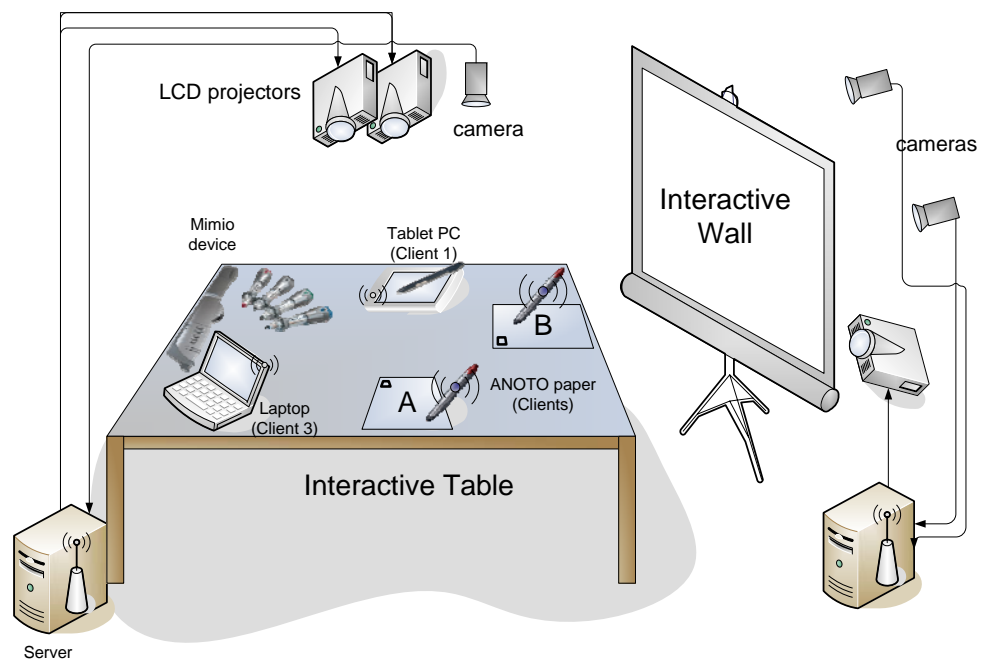


Figure 5: Coeno system configuration.

The installation itself consists of two main modules (cf. Figure 5):

1/ An Interactive Table, combining the benefits of a traditional table with all the functionalities of an interactive surface and display. The table allows people to easily access digital data and re-arrange both scribbles and virtual sketches in an intuitive way using different interaction tools.

2/ An Interactive Wall, consisting of an optically tracked rear-projection screen that displays digital content and captures gesture input. Combined with the Interactive Table, data can be seamlessly transformed from all presentation sources to the presentation wall.

The interface consists of two ceiling and one wall mounted projectors showing data on a table surface (Interactive Table) and on a rear-projection screen (Interactive Wall). All users can sit at the table and connect their own laptop and/or tablet PC computer to the display server. There is no limit as to how many clients can connect simultaneously to the system and the amount of co-located participants depends on the space around the table. In our case, typically 4-5 participants are involved in a meeting, where one of the participants usually leads the session.

Participants can interact with the table in several ways. They can either use their personal devices (e.g. tablet PC) wirelessly connected to the server, or a digital pen. Designers can create imagery on their own personal computers and “move” them to the interactive table for further discussion using hyper-dragging as proposed by Rekimoto et al (1999).

Unlike Rekimoto’s work, users can also use real paper in the interface. To digitally capture handwritten notes, participants use the Anoto⁶ digital pen system. These are ballpoint-pens with an embedded IR camera that tracks the pen movement on a specially printed paper covered with a pattern of tiny dots. We use the Maxell Pen-It device with Bluetooth wireless connectivity. In our tabletop interface, we also augment the real paper with projected virtual graphics. The paper itself is tracked by using ARTag⁷ markers, placed on top of each piece of paper. Thus, participants can make annotations on real content that is combined with digital content projected on top of the paper surface.

⁶ www.anoto.com

⁷ <http://www.cv.iit.nrc.ca/research/ar/artag/>

Participants are able to use the Interactive Table as a traditional whiteboard for brainstorming tasks. We integrated a MIMIO device⁸, with ultrasonic tracking, which enables participants to draw on the interactive table and create annotations in real-time. Finally, the Interactive Wall is a rear-projection system which allows an intuitive gesture based interaction on a wall screen. We use a transparent rear-projection screen and track the user's gestures with an infra-red (IR) camera setup. All of these devices can be used simultaneously and they combine input and output on one surface using several novel interaction metaphors. A closer description of the implemented interaction metaphors including a first pilot study is presented in Haller et al. (2005, 2005b).

In summary, the Coeno interface combines three different display spaces:

- *Private Space*: The users' own hardware device (e.g. laptop/tablet PC screen) and/or the area on the table around each participant. Other users cannot see the private information of the others.
- *Design Space*: The shared table surface (the interactive table), only visible to those sitting around the table. This space is mainly used during the brainstorming process.
- *Presentation Space*: The digital whiteboard which is visible to all people in the room and therefore part of the presentation space.

However, Coeno does not offer a remote, collaborative functionality. Therefore, we combined the advantages of cAR/PE! and Coeno into a first prototype, Carpeno, which is described in the next section.

A Combined Approach: Carpeno

Carpeno tries to overcome the barrier between co-located and remote collaboration while maintaining the interface advantages of table-top environments for creative group processes. Therefore a combination of the cAR/PE! and Coeno systems seems to be a promising approach. We will briefly introduce our conceptual idea and show a proof of concept with an initial, exploratory user study based on a first implementation of the concept. Our general concept is based around the obvious idea of combining the two approaches: (1) the table-top part of the Coeno environment and (2) the

⁸ www.mimio.com

teleconferencing elements of cAR/PE! in a wall projection mode. The goal is to link these systems as closely together as possible to allow for a borderless communication and interaction space. Figure 6 shows the setup in a simplified manner.

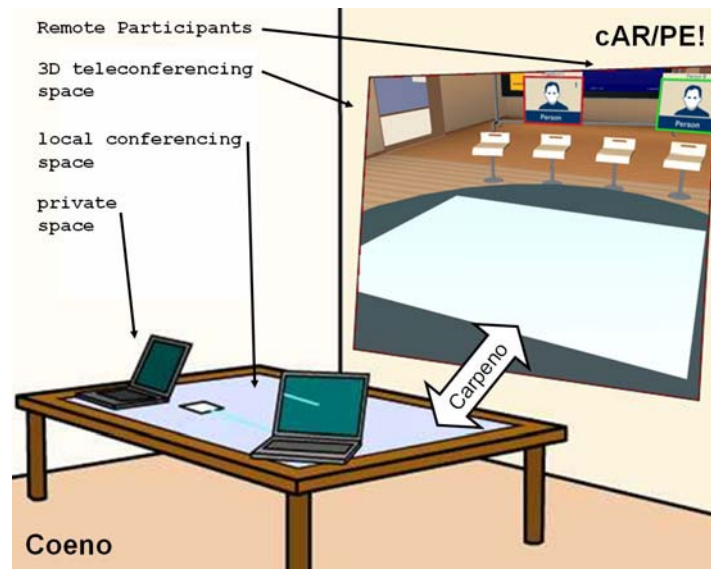


Figure 6: Carpeno Principle

Coeno's private space is preserved and the data and interface components are still used in the same or even enhanced way as the design space introduced earlier. The presentation space is replaced by a screen projection showing the remote cAR/PE! virtual meeting room environment. This should create the impression for the local participants of two tables placed next to each other: the physical local table and the remote virtual table, both interactive and suitable for information display. The remote cAR/PE! participants can still freely move around in the virtual space. With this they are able to form an own shared space out of reach and sight of the local participants (similar to their local shared space). Both sides of the setup are coupled via (1) the display of the video and audio streams, including their (changing) locations and (2) data transfer and interactions coupled between the systems. Figure 7 illustrates the new communication and interaction spaces with Carpeno.

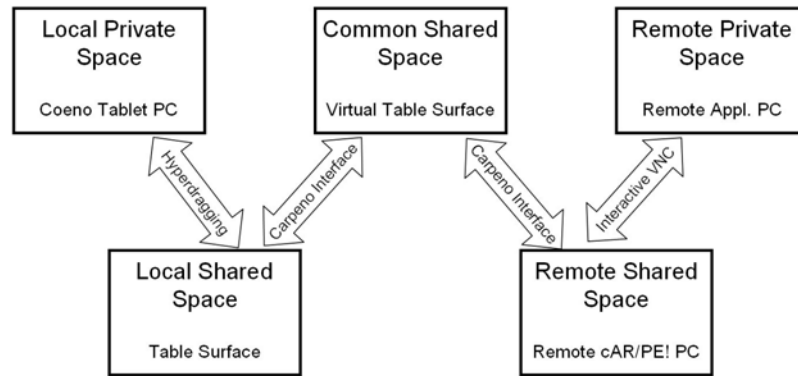


Figure 7: Carpeno Spaces

The central shared element between all participants (local and remote) is the virtual table within the (former) cAR/PE! environment, called the Common Shared Space. Local spaces are provided for each group: the local shared space on top of the physical table and the remote shared space everywhere within the cAR/PE! environment outside the reach of the local group. For example, the remote participants can choose a corner (and virtual table or presentation screen if needed) within the virtual environment and come back to the common shared space (virtual table) for discussions concerning the entire group.

The private spaces are on each side personal information systems (in most cases laptop computers or tablet PC's) connected to the Carpeno system, but only visible to the individuals. Digital content can be shared via hyper-dragging or screen sharing, visible to a sub-group (e.g. local only) or the whole group (e.g. on the virtual table). Furthermore the virtual presentation screen within the cAR/PE! environment can be made visible to all for group discussions.

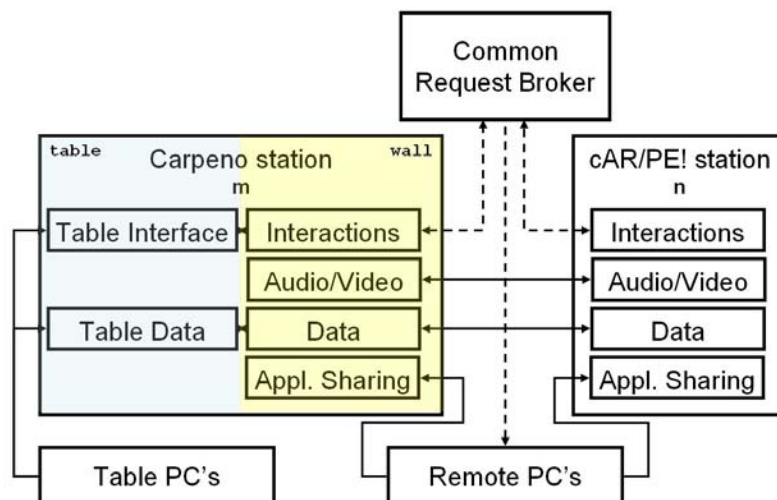


Figure 8: Carpeno Scheme

With this concept a new technological infrastructure and features have to be developed. Figure 8 illustrates how Coeno and cAR/PE! are linked together to form the seamless Carpeno system. As shown, the networked part of the cAR/PE! system remains almost entirely unchanged, while the data and interaction components are extended by the Coeno interface. We adopt a loosely coupled approach, where network messaging techniques are used as the main software technical method. With this we are able to control almost all of the aspects of the cAR/PE! part of the system with the Coeno part and vice versa.

A virtually infinite number of even mixed local and remote stations can be linked together without any system-inherent limitations. The main reasons not to do so are: (1) limited bandwidth and other networking issues, (2) the (virtual) placement of a certain number of persons and parties around one virtual table, and (3) interface issues that have to be solved beforehand (e.g. orientation of documents, pointers indicating interacting persons, etc). Currently two to six co-operating parties can be brought together in one Carpeno system without serious problems.

Prototype Implementation

The first implementation of our conceptual approach serves as a test bed for evaluating the feasibility of the Carpeno concept. Our focus therefore is set on building a functioning and tangible system to be used for testing rather than on providing the most comprehensive and complex solution first. We decided not to implement and integrate all features available in cAR/PE! and Coeno but rather to develop a system which can be initially tested in exploratory studies.

System

The initial version includes the following elements (see figure 9):

A vertical Plasma projection screen (WXGA resolution) displaying the remote shared space. The size of this screen was chosen to provide a wide field of view for the local party. The screen is accompanied with speakers to display the (spatially arranged) voices of the remote participants to the local group in a convenient way.



Figure 9: Carpeno v1.0

The local shared space is defined by a touch sensitive surface⁹ on which a projector (XGA resolution) shows the augmented surface content. With this setup one person at a time from the local group can directly interact with the digital content displayed simply by using his or her finger.

The augmented surface content is provided by the cAR/PE! system: An additional computer is rendering the same environment as shown on the vertical screen, but from a correct perspective from above the physical and virtual table. With this pre-configured setup we can ensure that both sides, local and remote, see the same content on the table.

To capture the live video stream of the local participant(s), we placed an Apple iSight camera on top of the Plasma display. While the image quality of the camera is superior for teleconferencing purposes, no real eye-to-eye contact can be achieved. In a standard situation, where the remote and local participants are sitting, this is still the best camera position, because it is close to the remote participant's eyes.

Within the shared cAR/PE! environment the virtual content on the table is provided via a VNC application sharing component. The Coeno system connected to the network is providing this screen stream and resides on an additional computer.

In summary, three components from the cAR/PE! system are involved in the Carpeno setup: (1) the remote participant working at a standard PC screen, (2) the

⁹ www.nextwindow.com

vertical screen (Plasma) of the local setup, and (3) the horizontal screen (touch screen) of the local setup. We have configured and calibrated these three components in a way that they form one, consistent spatial environment. The local private space is provided by a tablet PC standing beside the touch sensitive surface. It is used to prepare content to be discussed in the group and to drag and drop it to and from the local shared space using the hyper-dragging metaphor. While for the users this interaction is a transparent one, the actual technical process is implemented via VNC application sharing feeding the cAR/PE! applications. All three cAR/PE! components receive the same VNC stream and display it on top of the virtual table.

All computers involved in this initial Carpeno setup are linked via a dedicated network switch, ensuring the highest possible networking performance. While we could have chosen virtually any video and audio codecs in this network setup, eventually we opted for high quality videoconferencing standards (G.711 uLaw and H.261 CIF) to emulate an Internet connection.

In this version we have reduced the conceptual number of possible spaces to three to ease our exploratory studies. The virtual table (common shared space) and the projection onto the physical table (local shared space) are exactly overlaid to give the impression of one single table surface. Therefore, what the remote participants see on the virtual table is exactly the same what the local participants see. In addition, we abandoned the use of additional PC's on the remote side (remote private spaces) to avoid confusion about the interface in the first instance.

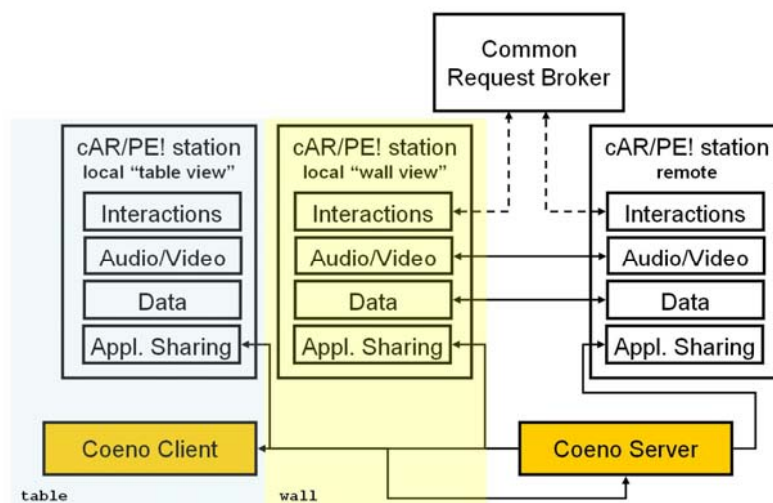


Figure 10: Carpeno v1.0 Implementation

Figure 10 illustrates our implementation. The Coeno system delivers all content via the application sharing functionality of cAR/PE! (sharing parts of the computer screen), while the interaction with the content of the common shared space is controlled by the touch sensitive surface. This system allows for actual communication and interaction within the Carpeno concept and serves as the basis for our exploratory user study described in the next section.

Exploratory Study

We conducted an informal exploratory study with our first prototype system. In total forty visitors at the ICAT2005 and Graphite2005 conferences participated in a hands-on evaluation during the exhibition of our system.



Figure 11: User Study at Conferences

Task

Two persons at a time took a seat at different parts of our booth. One part was configured as a Carpeno station as described in the Implementation section and the other part was set up as a cAR/PE! station using a standard PC and Monitor equipped with a headset and a web cam. If only one volunteer was available, one of the exhibitors took on the role of the second person at the cAR/PE! side.

Photographs of interesting looking devices that were invented during the last 200 years (taken from Collins (2004)) were then dragged onto the shared table by a moderator. The task for the participants was to collaboratively discuss what exactly the purpose of the displayed objects might be. If a device's function could be guessed correctly, that picture got removed from the table by the moderator.

All pairs had to discuss five to six different photographs in order to clear the table while playfully exploring the features of the Carpeno setup at the same time. To complete one round typically took between 5 and 8 minutes.

Questionnaire

After a team completed the task, both participants were asked to fill out a short questionnaire. Besides usability issues we were especially interested in finding potential research variables that would arise from the asymmetrical nature of our setup. Most results that are presented in the following section are therefore presented separately for cAR/PE! and Carpeno users.

Results

After each session users were asked to subjectively rate the experience by answering nine seven-point Likert-scale questions. The questions and their normalised scores are summarized in Figure 12.

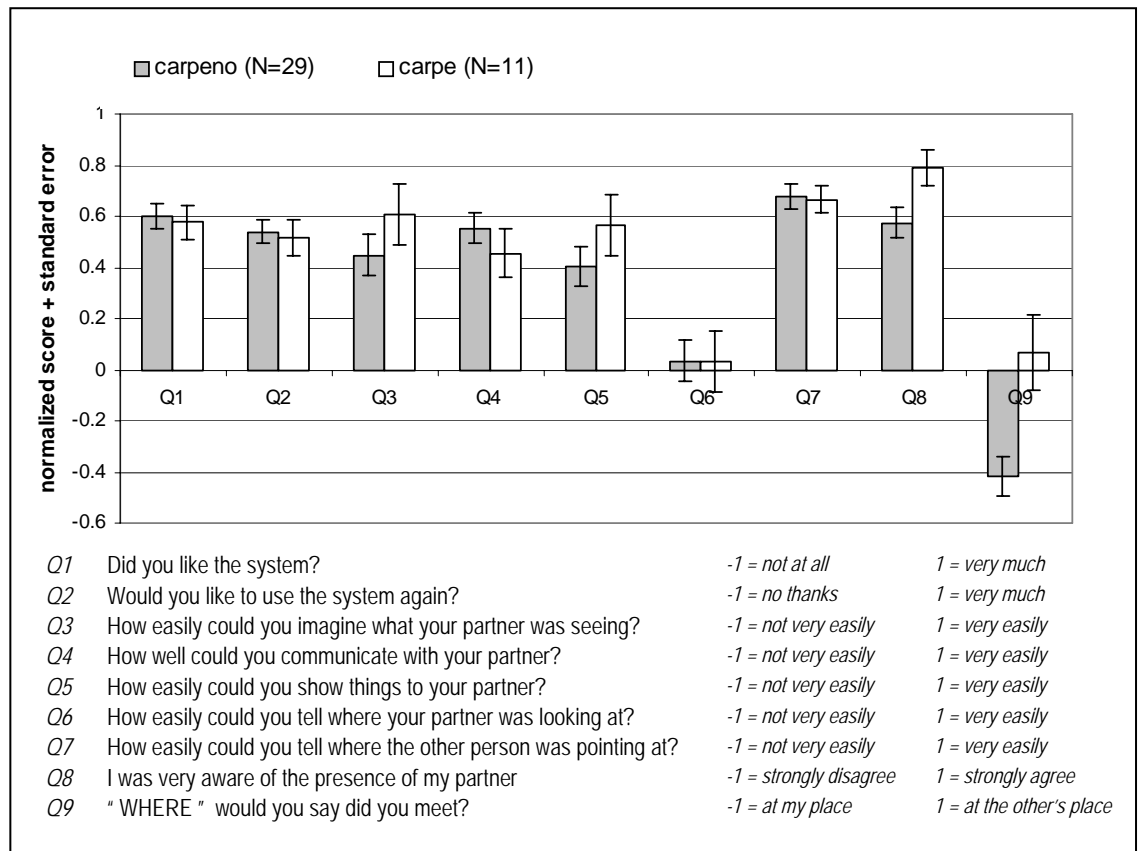


Figure 12. Questionnaire results by system

The scores in the satisfaction questions Q1 and Q2 show that both user groups liked the system. With the exception of question Q6, the answers on general

usability issues (Q3 to Q7) further show an overall positive response. The lower score of Q6 uncovers that users of both sides could not easily infer where the other person was looking at. This deserves further investigation but could be influenced by the fact that there was a very high task focus. No major differences in the usability scores emerged between the Carpeno and cAR/PE! side. However, cAR/PE! users were more aware of the other person's presence, as can be seen in the scores of question Q8, probably due to their undisturbed concentration on one screen surface (the monitor). The biggest difference between both user groups emerged in question Q9. Carpeno users felt much more that the meeting with the other person occurred "locally", i.e. around the physical table in front of them. On the other hand, cAR/PE! users thought the meeting took place more "remotely", situated somewhere in the middle between their and the other person's location.

Although we haven't carried out formal statistical tests in this exploratory study, we can derive some initial lessons:

- 1) The low gaze awareness that appeared in question Q6 suggests that this issue demands some more attention in our setup. Applying head tracking technology that allows users to control their video avatar simply by moving their heads could deliver some improvements and would get rid of the need for mouse-based navigation. In addition, other gaze awareness support could be integrated such as the "miner's helmet" metaphor (Vertegaal, 1999) that displays a lightspot at a person's centre.
- 2) The lower awareness of the partner's presence in the Carpeno setup might be a result of the carpe user "disappearing" from the Carpeno-user's screen when navigating to the other side of the table in the cAR/PE!room. This often led to confusion on the Carpeno side. Seeing the other person at all times therefore seems to be crucial for the awareness of the other's presence, even if the audio connection is maintained. In future experimental setups, we therefore have to limit the navigation space for the cAR/PE!-user to an area where s/he is always visible to the Carpeno user.

- 3) The clear result about the experienced location of the meeting (Q9) suggests that users are very much able to associate a remote encounter with a spatial reference frame somewhere between “here” and “there” as it is defined by the interface. To understand the effects on the user and how exactly we can move both interface types along this dimension will be part of our future research.

Discussion & Future Work

Our conceptual approach in bringing together co-located and remote collaboration into a single system as well as our first implementation suggests that the Carpeno interface has indeed great potential for enhancing remote face-to-face collaborative creative experiences. Our initial, exploratory user study with Carpeno and the numerous experiences with the single systems cAR/PE! and Coeno lead us to develop requirements a future Carpeno system should have and opens up new research areas to work on.

Our initial assumption was supported, that the combination of our two systems can compensate for the flaws in interfaces detected in the separated systems. In particular the incorporation of remote participants into the co-located collaboration is possible and the provision of a table-top environment for the remote participants is of great value, especially in creative tasks like brainstorming or general discussions involving some sort of media. Eventually we can provide a common shared space as well as local shared and private spaces at the same time.

Direct manipulation on the interactive table is intuitive and can be supported by different interfaces, depending on the particular task to be addressed. For our picture sharing application finger pointing was very appropriate. Participants have different preferences and different tasks require different input devices (e.g. digital pen, tablet PC, Mimio tracking device, etc.). Therefore, one of our goals is to test the different benefits of these devices.

The incorporation of a table as the central element of our interface (real and virtual) and the consequent integration into a meeting environment (also both real and virtual) leads to the reasonable approach of (“re-”) introducing spatial objects into the process and interface. On the physical side (real world) real objects can be

used as part of the creative group processes or as part of the interface (tangible user interface, see Billingham & Kato, 2002, Hauber et al., 2004, and Regenbrecht et al., 2002). On the virtual side (and within the virtual space) the use of 3D virtual objects representing the real world can be used also either as the object of discussion or as interface elements. Further research is needed here and should be based on existing findings and systems (in particular tangible and perceptual user interfaces, ubiquitous computing, 3D user interfaces).

For the sake of simplicity and to rapidly allow for an early exploratory study we've excluded some interfaces, which would be very relevant in non-experimental situations. We are going to amend the system with a shared digital whiteboard, better support for gesture communication, and pen-based interaction. Also, the (simultaneous) placement of documents in the shared spaces will be approached based on the experiences made with the single systems. For example, mechanisms already built-in into the Coeno system can be used for a "real estate" saving arrangement of documents onto the limited virtual and real table space.

While general gaze awareness could be provided with our Carpeno system, eye-to-eye contact is still not possible because of the different locations of the real camera and the virtual participant representation as a video stream. We are working on optical and/or IT solutions to allow for this essential aspect in certain task scenarios (like negotiations). The form of representation of the avatars itself (video stream on a moving virtual plane) was acceptable. This was already tested in earlier studies with the cAR/PE! system. However, to provide even better communication cues and channels, we are going to test, whether other forms of representations (e.g. with background eliminating methods) can even enhance the overall quality.

In addition, our first implementation was mainly limited to one remote and one local person. We are exploring how the system has to be modified to add more local or remote participants. Issues that must be addressed include concerns such as: Do all of the participants meet in the local (physical) or the remote (virtual) place? or How does informal communication between co-located participants affects the entire, creative process with the remote participants? These questions have to be answered in the future, involving more creative tasks besides brainstorming and/or picture sharing.

With our current, integrated approach the development of new interface metaphors and techniques considers the combined support for local and remote collaborative tasks at an early stage. It can be assumed that this consideration leads to more comprehensive and efficient interfaces suitable for both worlds, the local and the distant one. This could be a satisfactory contribution to tool and process development of a converging world of communication and information. Last but not least, communication quality can be improved in using the Carpeno approach. Especially support of non-verbal communication cues in relation to a high level of social presence seems to be essential and can be implemented on our current basis. For instance the introduction and evaluation of head-tracking, gaze and workspace awareness supporting techniques for natural gesture recognition, and eye-to-eye contact in remote settings are part of our future research.

Acknowledgements

We would like to thank Claudia Ott, Michael Wagner, Graham Copson and the Technical Support Group at Otago University, and all the participants in our experiments for their great support. In addition, we would like to thank DaimlerChrysler Research and Technology for supporting our work and the anonymous reviewers with their comments, which lead to some very relevant improvements. The Office of Tomorrow project is sponsored by the Austrian Science Fund FFG (FHplus, contract no. 811407) and VoestAlpine Informationstechnologie. Moreover, the authors would like to thank Daniel Leithinger, Jakob Leitner, and Thomas Seifried for the great work in the Coeno project.

References

- Ashdown, M. and Robinson, P. (2005). Remote Collaboration on desk-sized displays. Computer Animation and Virtual Worlds 16(1). Wiley InterScience, 41-51.
- Billinghurst, M. and Kato, H. (2002). Collaborative Augmented Reality. Communications of the ACM, 45(7), 64-70.
- Burgoon, J., Buller, D., & Woodall, W. (1996). Nonverbal Communication: The Unspoken Dialogue (2nd ed.): McGraw-Hill. Buxton, W. (1992). Telepresence: Integrating Shared Task and Person Spaces. Proceedings of Graphics Interface '92, 123-129.

Cao, X., Balakrishnan, R. (2004). VisionWand: Interaction techniques for large displays using a passive wand tracked in 3D. ACM Transactions on Graphics, 23(3). Proceedings of SIGGRAPH 2004. p. 729.

Collins, M. (2004). Eccentric Contraptions: An Amazing Gadgets, Gizmos and Thingamambobs. David & Charles.

Gutwin, C. and Greenberg, S. (1996). Workspace awareness for groupware. In Conference Companion on Human Factors in Computing Systems: Common Ground (Vancouver, British Columbia, Canada, April 13 - 18, 1996). M. J. Tauber, Ed. CHI '96.

Haller M., Billinghamurst M., Leithinger D., Leitner J., Seifried T. (2005). Coeno. Enhancing face-to-face collaboration. Proceedings of 15th International Conference on Artificial Reality and Telexistence, ICAT 2005, Dec. 5-8, 2005, Christchurch, New Zealand.

Haller M., Leithinger D., Leitner J., Seifried T. (2005). An augmented surface environment for storyboard presentations. ACM SIGGRAPH 2005, Poster Session, August, 2005, Los Angeles, USA.

Hauber, J., Billinghamurst, M., & Regenbrecht, H. (2004). Tangible Teleconferencing. In Proceedings of the Sixth Asia Pacific Conference on Human Computer Interaction (APCHI 2004), June 29th - July 2nd, Rotorua, New Zealand, 2004, Lecture Notes in Computer Science 3101, Springer-Verlag, Berlin, pp. 143-152.

Hauber, J., Regenbrecht, H., Hills, A., Cockburn, A. & Billinghamurst, M. (2005). Social Presence in Two- and Three-dimensional Videoconferencing. Proceedings of 8th Annual International Workshop on Presence, London / UK, September 21-23, 2005, 189-198.

Hills, A., Hauber, J., & Regenbrecht, H. (2005). Videos in Space: A study on Presence in Video Mediating Communication Systems. Short paper in Proceedings of ICAT 2005, December 5th-8th, 2005, University of Canterbury, Christchurch, New Zealand.

Hollan, J. & Stornetta, S. (1992). Beyond being there. In Proceedings of the SIGCHI conference on Human factors in computing systems Monterey, California, United States ACM Press, 1992, 119-125.

Inkpen, K. (1997) Adapting the Human Computer Interface to Support Collaborative Learning Environments for Children. PhD Dissertation, Dept. of Computer Science, University of British Columbia, 1997.

Ishii, H. and Ullmer, B., (1997). Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. Proceedings of Conference on Human Factors in Computing Systems (CHI '97), ACM, Atlanta, March 1997, pp. 234-241.

Kauff, P. & Schreer, O. (2002). An immersive 3D video-conferencing system using shared virtual team user environments. In Proceedings of the 4th international conference on Collaborative virtual environments Bonn, Germany ACM Press, 2002, 105-112.

Kelly, T. (2001). The Art of Innovation. Doubleday/Random House, New York.

- Nakanishi, H., Yoshida, C., Nishimura, T., & Ishida, T. (1998). FreeWalk: A Three-Dimensional Meeting-Place for Communities. In Toru Ishida (Ed.), Community Computing: Collaboration over Global Information Networks, John Wiley and Sons, pp. 55-89, 1998.
- Rekimoto, J, Saitoh, M. (1999), Augmented surfaces: a spatially continuous work space for hybrid computing environments. In CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems, 1999.
- Regenbrecht, H., Wagner, M., & Baratoff, G. (2002). MagicMeeting - a Collaborative Tangible Augmented Reality System. Virtual Reality - Systems, Development and Applications 6(3), Springer, 151-166.
- Regenbrecht, H., Lum, T., Kohler, P., Ott, C., Wagner, M., Wilke, W., Mueller, E. (2004). Using Augmented Virtuality for Remote Collaboration. Presence: Teleoperators and virtual environments, 13(3), 338-354.
- Scott, Stacey D. (2005). Territoriality in Collaborative Tabletop Workspaces. Ph.D. Dissertation, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada, Defended March 2005.
- Sellen, A. (1995). Remote Conversations: The effects of mediating talk with technology. Human Computer Interaction, 1995, Vol. 10, No. 4, pp. 401-444.
- Stefik, M., Foster, G., Bobrow, D., Kahn, K., Lanning, S., Suchman, L. (1987) Beyond the Chalkboard: Computer Support for Collaboration and Problem Solving in Meetings. Communications of the ACM 30(1), 32-47.
- Stewart, J., Bederson, B., Druin, A. (1999) Single Display Groupware: A Model for Co-Present Collaboration. In Proceedings of Human Factors in Computing Systems (CHI 99), Pittsburgh, PA, USA, ACM Press, pp. 286-293.
- Streitz, N., Prante, P., Röcker, C., van Alphen, D., Magerkurth, C., Stenzel, R., Plewe (2003). Ambient Displays and Mobile Devices for the Creation of Social Architectural Spaces: Supporting informal communication and social awareness in organizations. In: Public and Situated Displays: Social and Interactional Aspects of Shared Display Technologies, Kluwer Publishers, 2003. pp. 387-409.
- Tang, A., Boyle, M., & Greenberg, S. (2004). Display and Presence Disparity in Mixed Presence Groupware. In 5th Australasian User Interface Conference (AUIC2004), Dunedin, NZ. Conferences in Research and Practice in Information Technology, Vol. 28. A. Cockburn, Ed.
- Vertegaal, R. (1999). The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit Pittsburgh, Pennsylvania, United States ACM Press, 1999 pp. 294-301.
- Williams, E. (1977). Experimental Comparison of Face-to-Face and Mediated Communication: A Review. Psychological Bulletin, 84, (5), pp.963-976. 1977.