Working through Walls:
Mediating Cooperation in Dynamic Spaces

Christine A. Halverson, Luc E. Julia, and Adam J. Cheyer*

Computer Human Interaction Center (CHIC!), SRI International, 333 Ravenswood Ave, Menlo Park, CA. 94025, USA, krys@speech.sri.com, luc.julia@sri.com

*VerticalNet Inc., 3408 Hillview Ave, Palo Alto, CA 94304, USA, acheyer@verticalnet.com

1. Abstract

An important part of supporting cooperative work is to *support* the work rather than becoming an obstacle *between* the people and their task. We are striving to achieve this kind of invisible support by developing systems that provide interaction via modalities other than keyboard and mouse. We present two prototype systems that we have developed that use speech and pen based interaction. One system—OfficeMATE—provides portable connections to embedded data to help a visitor interact more successfully in a new environment. The other system—EMCE—provides meeting support that includes both the more obvious collaborative support within a meeting as well as external support to link to the user's stationary desktop machine. OfficeMATE can also act within the EMCE environment as a more personalised access to the meeting. We feel that OfficeMATE and EMCE are both examples of computer supported cooperative work (CSCW) systems that encourage a seamless transition between desktop and meeting.

2. Introduction

As designers of computer systems, we want to provide computer *support* to users' work rather than getting *between* them and that work. To do this we first need to make interfaces more natural and easy to use by bringing real world methods of interaction, such as speech and pen, to computer interfaces. One advantage of this approach is that people already know how to navigate and interact with the world around them. People normally interact with each other and their environment, often with individual objects, by using a combination of expressive modalities, such as spoken and written words, tone of voice, pointing and gesturing, facial expressions, direct interaction, and body language. Observation shows that people often interact with computers and other devices as if they too were social actors [1]. People also interact by using a variety of modalities in order to express intent. They do so even when the only interaction these devices would accept are unimodal, that is with a single method of interaction such as a mouse click, keystrokes or a button push. Who hasn't seen someone gesture while talking on the phone? Or threaten to kick a recalcitrant machine?

A second intent of our work depends on two assumptions about the nature of computer supported cooperative work (CSCW). First, that the cooperative in CSCW refers to co-operation between the individual and the device, as well as cooperation between people. Furthermore, with different notions of cooperation, and because the kinds of cooperation to be supported will necessarily change with the context, then in some cases this means a change in the appropriate user interface. As Bannon and Schmidt have argued we need to understand what CSCW *is* in order to build applications for it [2, 3]. Bardram [4] argues that perhaps the problem is that "cooperative work" really is dynamic—different things at different times and different places.

We also recognise that much of the work practices that CSCW systems are meant to support involve mobility between locations, casual interactions in hallways, as well as organised meetings with and without all participants in the same location.

To this end, we developed two prototype systems designed to address the variety of needs, as we see them, of collaborative work today. OfficeMATE (Multimodal Augmented Tutorial Environment) is a portable notebook-sized slate that a user can carry around with them. EMCE (Enhanced Multimodal Conferencing Environment—pronounced like "MC" the abbreviation of master of ceremonies) is a room-based system for meetings. Both systems support interaction by pen and voice. Both prototypes can be seen as augmenting the space that users inhabit, hopefully in a way that is supportive and natural. Furthermore, the prototypes have an integrated, as well as stand-alone, use patterns.

In this paper we first present our rationale for the design of these systems, and discuss the related work. Next we discuss the implementation of both prototypes, and finally we talk about our preliminary user experiences and directions for the future.

3. Design Motivations

Our approach in this case has been a very pragmatic one. We observed meetings and the experiences of visitors—both those at our institution and in visits to other institutions. Visitors have some experiences that are common to residents in meetings, but they also have some different needs. Like travellers everywhere, visitors need help finding their way around, keeping track of where they are, as well as keeping up with the ordinary tasks of a productive work life. Our goal was to help visitors function in a novel environment while facilitating their work practice.

A particular challenge is that visitors are, almost by definition, mobile. They often do not have the option of being at a desktop, yet they are not that different from the workers they are visiting. There is a growing recognition in the CSCW community that much of the work in an office is done in informal interactions. Many of these informal interactions take place spontaneously, such as in a hallway, but just as often they are in a meeting room or an office, sometimes on the fringes of more formal activities. Many of these interactions end with the participants going off to look up information or ask others for information. This is partially because they may not have ready access to the information they need where they are. It may also be because the methods for interacting with computers where the information is available is not conducive to continuing the kind of discussion in which they are engaged. That is, asking for the information they need would get in the way of their current task.

We conceived a variety of ways in which to support visitors. Having done previous work with multimodal maps for travel settings [5-7], and believing that interfaces need to become more natural and invisible for devices to become ubiquitous, we chose multimodal interaction via speech and pen. While we take Weiser's admonition about the current intrusiveness of speech interaction to heart [8], we feel that ultimately interaction by voice and pen will ease and enhance the user experience rather than complicate it.

Both of these insights are key to our approach. We will discuss each one in additional detail before introducing our application solution.

3.1 User Interface Technologies

User interface (UI) design is about creating an interface between the user and a computer system. Ideally, this interface should be both intuitive and efficient. Until recently, interface design has been hampered by the restrictions of technology. A history of user interface design is a history first of the available hardware (keyboards and later mice) and software (command line, graphical user interface—GUI, and now windows, icons, manipulation, and pointers—WIMP). Given the limitations of the hardware and software an important focus has been to make it easy to learn and use. Over the history of computing, we have slowly taught people to consider typing more natural than speech [9]. In our training and adaptation to the limits of the technology, we forget that a more natural way to interact with computers may be the way we interact with other people. We have long been used to using multiple modalities to express everything from emotion (angry gesturing and sharp words) to direction (pointing while saying "turn left at that light"). Work by Reeves and Nass [1] implies that humans may by default treat computers a social partners, regardless of whether computers facilitate that kind of interaction.

While the technology may not yet be there to take a finely attuned combination of speech and gesture and translate it into computer system commands, it has come a long way. Automatic speech recognition (ASR) technology has been under development for over 25 years, with considerable resources devoted to developing systems which can translate speech input into character strings or commands. We are just beginning to see expansion in the application of speech technology to user interfaces. Though the technology may not have gained wide acceptance at this time, industry and research seem committed to improving the technology to the point that it becomes acceptable. While speech may not replace other input modalities, it may prove to be a very powerful means of human-computer communication.

Although we like to think that speech is a natural form of communication [10, 11], it is misleading to think that this means that it is easy to build interfaces that will provide a natural interaction with a machine [12]. Another consideration is that despite the naturalness of speech, it takes time and practice to develop a new form of interaction [13, 14]. Speech user interfaces (SUIs) are evolving as we learn about problems users face with current designs and work to remedy them. We feel that much of the potential power of speech is as an additional modality coupled with drawing and writing—pen based interactions.

Our first work with extended input peripherals and alternative interface metaphors focused on adapting a user's interaction with a pen and piece of paper to the electronic realm. In the TAPAGE/DERAPAGE applications [Figure 1, Left], a user can conceptualise a complex nested table or flowchart, draw a rough freehand sketch of the concept, then engage in an interactive dialog with the system until the desired product is realised [15]. Interactions consist of natural combinations of both pen and speech input – a user can cross out an undesirable line, draw in new additions, and reposition lines or objects using commands such as "put this over here." In these applications, we tried to capture the nature of a pen/paper experience, while enhancing the paper's role to become a partner in the process. These interfaces are capable of following high-level instruction and taking an active part in the construction of the document.

A second project focused on applying the metaphor of "smart paper" to the domain of maps, where the goal is to manipulate and reason about information of a geographic nature [Figure 1, Right]. Inspired by a simulation experiment described in [16], we developed MMap, a working prototype system of a travel planning application. Using MMap users could draw, write, and speak to the map to call up information about hotels, restaurants, and tourist sites [6]. A set of collaborative agents helps the user to find the right information through a reactive, multimedia, interface. A typical utterance might be: *"Find all French restaurants within a mile of this hotel" + <draw arrow towards a hotel>.*

The research challenges in constructing such systems are in how to develop a multimodal engine capable of blending incoming modalities in a synergistic fashion, and able to resolve the numerous ambiguities that arise at many levels of processing. One problem of particular interest was that of reference resolution (anaphora). For example, given the utterance "Show photo of *the* hotel", several distinct computational processes may compete to provide information. For example, a natural language agent may volunteer the last hotel talked about, the map process might indicate that the user is looking at only one hotel, and a few seconds later, a gesture recognition process might determine that a user has drawn an arrow, or circled a hotel.

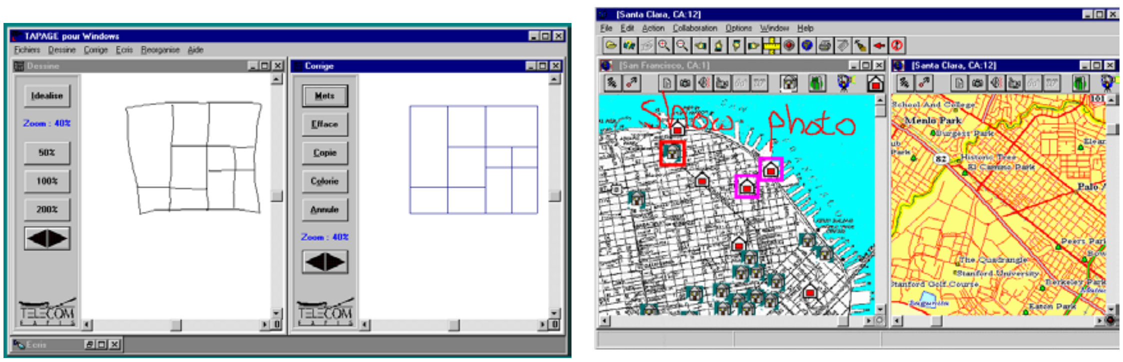


Figure 1. On the left, TAPAGE/DEREPAGE helps users draw complicated tables. On the right, MMap helps travellers navigate the city and provides them with information about hotels, restaurants, and sites to visit.

3.2 CSCW, Mediated Spaces, and Mobility

Mobility is becoming an increasing concern, both in the world at large, as well as within the framework of CSCW. Despite the proliferation of devices targeted to our more mobile lifestyle—such as pagers, cell phones, and PDAs—mobility issues have been largely ignored with respect to collaborative work. Notable exceptions include Whittaker et al. .[17], Bellotti and Bly [18], and Luff and Heath [19]. A more detailed look at their findings is worthwhile as it supports the design choices we have made in OfficeMATE and EMCE.

Whittaker et al. [17] recognise that much of workplace communication is informal, matching workers' peripatetic work habits. This informal communication supports a range of functions: executions of tasks; coordination of group activities; and social functions such as team building. They call on the research that indicates the importance of physical proximity for scientific collaboration [20], and reinforce it with their own findings of workplace interactions. Their findings indicate that over half the observed conversations involved documents, suggesting a shared workspace that supports annotation and simple interaction rather than a completely shared editor. Conversations were more frequent in offices than in common areas, but often took place in others' offices.

Bellotti and Bly [18] in their study of how a product design team actually works, confirms both Whittaker et al and what many of us know—to do their work people are often away from their desks and moving around. Some of this movement can be characterised as local (within a building) while some is more far ranging (from across the street to a different town). The commonality across all cases is that many of the typical support tools, such as email and or video conferencing, require the worker to be tied to a specific location, often their desk. Clearly, support that is more mobile is needed. However, to provide this support requires a more extensive infrastructure.

Luff and Heath [19] further underline the importance of taking mobility of working life seriously by presenting case studies from three apparently very different workplaces: a control room setting of the London Underground, medical consultations, and constructions sites. They point out how even in situations like a control room, where the workers are clearly tied to the location, there is a large reliance on the mobility of other workers and artefacts for the work to get done. Some of this mobility, like that noted by Bellotti and Bly, is very local, while some is more spread out. Their studies "reveal how the mobility of personnel and artefacts is critical to communication and collaboration" (p306) and confirm those from other settings such as air traffic control [21] and ship navigation [22]. Yet, as they point out, the CSCW field has been "principally concerned with enhancing the shared facilities for individuals on fixed workstations."

Heath, Luff and Sellen [23] in a discussion of the more classic CSCW media spaces, particularly video connectivity, point out that many of the assumptions about what work consists of tends to conflict with how people actually work. In particular, they argue that the focus on supporting face-to-face communication for video connections actually interferes with the being able to deploy a system that supports collaborative work.

In our own experience, we have observed that many of the "obvious" intuitions about workplaces are often wrong—perhaps because we are just too close to see the behaviour we engage in every day. Naturalistic studies of work, like those reported on by Heath et. al. [23], and our own experiences [21, 24, 25], confirm that many systems designed for collaborative work actually get in the way of the successful completion of that work in a variety of ways. It is difficult to transition between tasks on related projects, and perhaps more difficult to transition documents. Everyone has been at the meeting where they know they've just recently seen the perfect document to support a point—but they can't remember where it is or what it was in reference to. In addition, they can't get up and go back to their desktop computer to look.

Perhaps more important is how specialised systems demand our focus and adaptation to ways of interacting that derail us from the real task at hand. Such intense focus on the method and means of interaction tie up resources that we would otherwise use in interaction. As Heath et al. so aptly put it.

> "An individual's ability to contribute to the activities of others and fulfil their own responsibilities relies upon peripheral awareness and monitoring; in this way information can be gleaned from the concurrent activities of others within the "local milieu", and actions and activities can be implicitly coordinated with the emergent tasks of others."

Furthermore, we are used to interacting with each other by coordinating actions and activities around various shared artifacts and objects. Handing someone an object that they can annotate to their personal preference is preferable to limiting the scope of what they can do along a narrow range. As Stefik et al [26] point out, the promise of using computers for meeting support is that they can save state, and provide a way to retrieve it, in a manner that is not easy with a chalkboard. In their early work with Colab they point out the importance of understanding how meetings really work and the necessity of providing private as well as public space.

At the time when Xerox PARC began their investigations into computer support of collaborative work, we really were tied to the desktop. Since they began their work the notion of ubiquitous computing has gained popularity as well as much of the technical infrastructure to support it. In the following scenario, we present our notion of how we see OfficeMATE and EMCE could be used.

4. Scenario of Use

In designing and implementing these two applications, we were guided by our observations of the visitor's experience at SRI and the basic structure of meetings. We created a scenario of how we hoped these applications might be used. This scenario became increasingly fleshed out as development progressed. We offer it here as a guide to, and context for, the rest of the paper.

> Jane works for a large company on the East Coast that is looking at some technology created at SRI. Jane makes her first visit to SRI and has a whole day of visits scheduled all over the multi-acre campus. She arrives at the engineering building where there is no receptionist. An interactive kiosk welcomes her, with a very animated character known as InfoWiz. After a short interaction with InfoWiz, which confirms her schedule for the day, she is instructed where to get her badge, and to pick up a small computer at the badging office.
>
> As soon as she turns the computer on, OfficeMATE asks Jane where she wants to go. She indicates that she needs to find her first appointment of the day. OfficeMATE displays the day's schedule and tells her that the first appointment is with Christine Halverson. Jane asks for directions and OfficeMATE starts its guidance by automatically displaying the appropriate building map, showing where she is and where she needs to go. At one point, OfficeMATE tells Jane that she is in front of Doug Engelbart's office: "... you know, the guy who invented the mouse". When, Jane asks for more information about Doug, OfficeMATE answers her request by displaying some information on the slate and asking if she would like to bookmark it for later, read it herself, or have it read to her. She requests that it be bookmarked.
>
> When Jane arrives at Christine's office she notices that Christine has one of the small slates as well as the more typical desktop and laptop systems. They begin their meeting and as Christine gives her an overview of her work she picks up her own OfficeMATE slate to display a diagram to Jane. Jane uses the note-taker facility to take notes on the discussion. At one point, Jane has a question, and Christine brings up an Internet browser with the relevant Web page that amplifies her answer. At the end of the meeting Jane asks for some additional information to review. Christine asks to use Jane's OfficeMATE and using the stylus is able to copy the file from her desktop machine onto Jane's OfficeMATE.
>
> Later in the day, Jane writes on OfficeMATE's map "bathroom?" She doesn't use speech to ask her way for privacy reasons, and OfficeMATE doesn't use speech to show her the way to the bathroom. By using its speaker ID capabilities from earlier in the day, OfficeMATE guides Jane to the closest women's bathroom.
>
> Late in the afternoon, Jane has to attend a meeting that involves some of those she has been meeting with at SRI, as well as colleagues at her company on the East Coast. InfoWiz on her OfficeMATE slate reminds her of the meeting and directs her to the conference room. As she walks in the door she is automatically logged into the conference system and sees her display change to that of that of EMCE. As she takes a place at the table she sees Christine, her first appointment that day, to her

left, and a stranger she hasn't met to her right. She touches the icon that is to her right on the EMCE display and sees that the man is Luc Julia.

As the meeting starts she begins to take notes. At one point she explains a slide from one of her colleagues and draws directly on the slide to do so. Some notes she took during a conversation earlier in the day become important in the meeting. After discussion the other participants need a copy of the notes, so Jane drags the icon of her notes onto the printer icon in the side bar and the right number of copies is printed out by the conference room printer.

At the end of the meeting the session is archived, including the slides, the public annotations, and any documents that have been added. This archive will be accessible to all the participants later.

At the end of the day Jane decides to get her notes from the day in two separate ways. She asks for a printout, which is waiting for her when she turns in her badge and the slate on her way out the door. OfficeMATE also emails the notes to Jane so that she'll have access to them when she gets back home.

This entire scenario is not completely implemented. However, much of it is implemented in the two prototypes as well as some supporting infrastructure. In the next section, we discuss specifics of that implementation.

5. Design Implementation

As we discussed earlier, we approached the design of these two prototype applications from a mostly pragmatic approach. A critical part of the scenario presented is the integration of several technologies necessary to support the different modalities of interaction, as well as the access to different pieces of information. The integration is possible because of the Open Agent Architecture (OAA) developed at SRI [27]. OAA is a distributed infrastructure that provides the means for bringing together multiple component technologies in a flexible, plug-and-play manner. Components can be written in different programming languages[1] and be distributed over multiple computers.

[1] Current languages include Java, C, C++, Visual Basic, Prolog, Lisp, Delphi, and WebL.

5.1 Implementation of OfficeMATE

The OfficeMATE slate is intended to be a small tablet that consists of a touchscreen with a pen interface. In the current implementation, we are using a commercially available laptop sized pen computer (Figure 2).

Figure 2. OfficeMATE prototype

There were many separate components buried in the scenario presented above. The first one is InfoWiz, the animated character that helps direct Jane and answer her questions. Then there is the multimodal map Jane interacts with to find her way, and locate specific kinds of places (i.e. the bathroom). Finally, OfficeMATE is connected: to the local area network (LAN), as well as the World Wide Web, and the conferencing system of EMCE. We'll discuss each component in turn.

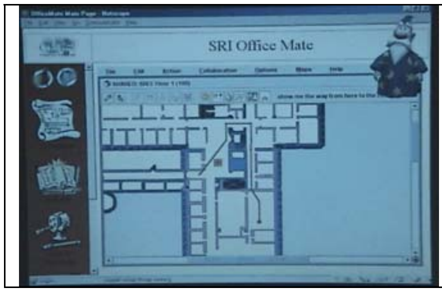
5.1.1 InfoWiz

The InfoWiz character in OfficeMATE is based on our previous work with the InfoWiz kiosk [28]. It integrates character animation graphics, speech recognition, natural language interpretation, simple dialog management, text-to-speech generation, and a repository of knowledge about InfoSpace. InfoSpace refers to a set of WebPages covering a subset of SRI, for which there is coded knowledge accessible to InfoWiz. The plug-and-play nature of OAA allows us to integrate off-the-shelf components with state-of-the-art research efforts. For example, all information is presented in a web browser, which can be either Netscape Communicator or Internet Explorer running under Windows. The speech recognition is an SRI-developed speaker-independent continuous speech recognition system, now commercialized in an SRI spin-off: Nuance Communications[2]. Natural language processing is handled by a mixture of SRI's DCG-NL parser [29] and Nuance's natural language (NL) API.

---

[2] http://www.nuance.com

An initial prototype of OfficeMATE is implemented in a browser-based interface. The initial screen uses a frames based approach to present the available applications and to explain the basics of interacting with OfficeMATE (Figure 3).



The OfficeMate computer will help you enjoy your day at SRI. You may choose any of the applications presented in the lefthand column by tapping its image or by saying the application name. All applications accept either handwritten or spoken commands. As you move through SRI's hallways, OfficeMate may update your display automatically, depending on where you are walking, to provide information in context.

Figure 3. On the left is a view of the building floor plan. On the right is the text of the initial screen.

5.1.2 Multimodal Map and Speaker ID

The multimodal map described in the scenario builds on our experience with multimodal maps in a variety of environments ranging from military applications to travel planning and most recently personalized map help in the car [5-7, 30]. Although the scenario focuses on voice input our multimodal map application takes input requests via gestures (e.g. arrows, lines, circles, cross-out or delete marks), handwriting, voice, or a combination of pen and voice.

The map currently covers two floors of one building at SRI encompassing the Speech and AI Labs. A hallway infrastructure of in-building trackers provides cues to OfficeMATE as to where it is with respect to the map. This information in turn helps InfoWiz access the appropriate data for where it is.

Speaker identification has not been implemented in either of these prototypes, but it is a part of other prototypes [31], and is targeted for inclusion. The foundation of OAA under all these prototypes makes integrating other functionality relatively easy, if not completely trivial.

5.2 Implementation of EMCE

EMCE is also implemented with the Open Agent Architecture (OAA) [27], mentioned above. Physically, an EMCE conference room has three main differences from traditional conference rooms: it has touchscreens embedded in the table to serve as personal display devices, it has an electronic whiteboard, and each location at the table has a microphone and earphone. The rest of the hardware is what one would find in a traditional conference room: a computer projector, VCR, printer, and so on.

Each participant must log in to the meeting upon arrival. In the case of the scenario this is mostly automatic. Information about each user is stored in a text database and is retrieved by the system as necessary. Each user can either sit down at an existing console or bring in a personal laptop that connects to the system through a wireless Ethernet card, using the DHCP protocol. OfficeMATE is designed to make this connection seamlessly.

EMCE is designed to work in two different types of meeting environments: collaborative and led. (We'll focus on collaborative here, but see Ionescu and Julia [32] for more information). In a collaborative meeting, each participant is considered to have equal permissions to place objects into the projected space and write public comments on objects.

The interface is designed to work with a pen on a touchscreen. Both the embedded screens and OfficeMATE provide this style of interaction. (For people using a laptop without touchscreens, a mouse and keyboard may be substituted as input devices.) The main goal in designing our interface was to make it as intuitive as possible to use within the context of a meeting. Whenever possible, if a function is performed electronically in EMCE, it is executed in a manner as similar as possible to it is performed in an unaugmented environment. For features that are new with the context of EMCE, and are not an augmented version of a previous meeting functionality, we tried to retain the idea of the virtual room and to make the interface work in a manner relative to the physical space. Our goal is that users should need to learn as little as possible when entering an EMCE conference room. Our intent is not to change their manner of meeting participation, but to augment and facilitate it.

We present some of the key components of EMCE as outlined in the scenario. This includes orientation within a virtual view, shared annotation of slides, printing within the meeting framework, note taking, and later access. Other functionality, such as passing private notes with *stickies*, is discussed in more detail in Ionescu and Julia.

5.2.1 A Virtual View

Our idea of creating a physical representation of the space was influenced by a project at IBM in which objects can be dragged between computers in a room according to their physical relationships in space [33]. When users log in to EMCE, they see a virtual representation of the meeting table in front of them. All meeting participants are represented as icons along the edge of the screen. As users log in and out, the icons appear and disappear. The icons always appear in the desktop relative to where a person is sitting (Figure 4).

Thus, for person A's display, person A's icon appears at the bottom of the screen and the icon of person B, who is physically sitting directly in front of person A, appears at the top of person A's screen. The icons are flipped for person B: person B is at the bottom of the screen, and person A is at the top of the screen.

The icons representing the participants are useful in several ways. The primary purpose is to obtain personal information. If a user passes the pen over an icon, the name of the person whom the icon represents appears. If the user then presses down with the pen on the icon, a dialog box with information about that person appears. The information provided includes e-mail address, position, office number, telephone number, and so on. Thus, a meeting participant always knows the identity of all the other people logged in to EMCE for a given meeting.

The concept of the virtual space is also used regarding some of the functionality directly linked to hardware. While a person does not need to know exactly where in a room the various pieces of hardware are located (such as the printer or projector), each of these objects is given a virtual space represented by an icon on the screen. A user can interact with the hardware by dragging and dropping both the icons themselves and files onto the icons.

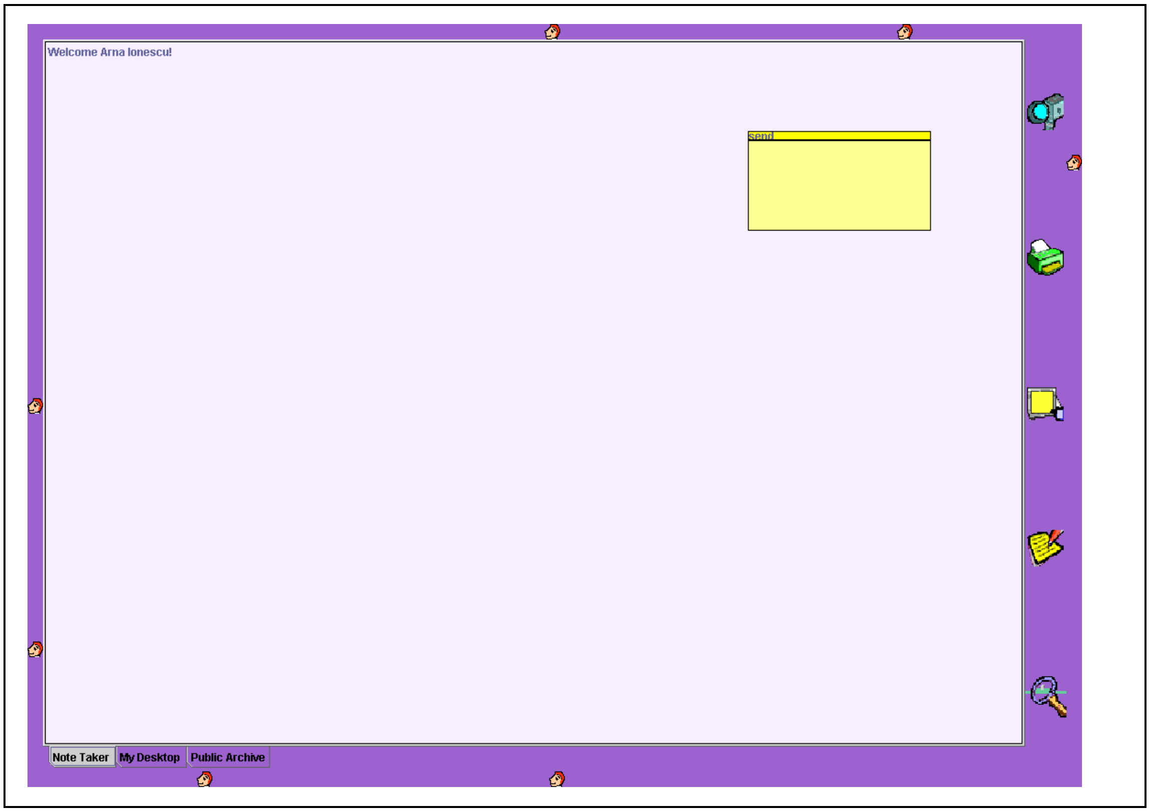


Figure 4. Main screen of EMCE showing seven people in the meeting. This persons place is centered in the bottom frame.

To create the virtual view, EMCE must know where each participant is located in the room. Eventually, we want to have a proactive system, with a database of voices, which can use speaker identification methods that will make user login unnecessary. For now, however, the system is reactive and each location in the room is assigned a letter that a user must enter upon logging in. A central agent checks for discrepancies (multiple people in one location), and also provides a list of users and locations that can then be processed by an algorithm to determine each user's personal view of the meeting table.

5.2.2 Public Archive

People frequently want to review an image that was previously projected for all participants to see. Thus, EMCE archives all the objects that have been placed on public view. Participants, who can browse the archive at any time during the meeting, see a stack of tabs, with the most recently projected view on top. A whiteboard image can also be included in the stack. In addition, two check boxes labeled "public" and "private" deal with annotation issues. A participant can use the public archive to write private notes as well notes for all participants to see. So that people do not accidentally write public notes they intended for private use, a private checkbox is given priority. If both public and private boxes are checked, notes being written are private. A user who wants to remove an annotation from view can scribble over it, the gesture will be recognized and the note removed. To avoid confusion about what is being viewed publicly, public annotations appear in a different color. Furthermore, the object on the top of the stack being viewed is slightly masked to let the user know that that object is being projected.

The user may still write on the projected object, but the constant visual feedback is a reminder of what is being viewed on the projected screen.

The public archive and projected view has the potential of leading to numerous blunders, and we designed this interface with the idea of minimizing these blunders. If a user decides to re-project an older view from the meeting, then that view must be dragged onto the projection icon, thus placing a second version on the top of the archive stack. The checkboxes can be used to state whether private and/or public annotations should be included in the new view. One aspect of the projected view is that any diagram or simple drawing is beautified. For instance, as a user draws a table, the lines of the table are straightened, and the written text inside the table is changed to typed text [15]. Thus, cleaner documents are produced, and after the meeting there is no need for someone to process the ideas and concepts discussed. The presentable form is created automatically as the participants work, thus saving time.

5.2.3 Minutes

The projected archive serves as a timeline for what occurred during a meeting. As an object is viewed, even if it was viewed earlier, it is placed on the top of the archival view. At the end of the meeting, all the images in the stack are processed, and minutes are created in HTML. The minutes, which can be viewed with any Web browser, build a timeline of the images from the projected view, with sound and video clips of the meeting available upon demand. Public annotations that were not removed are included, and private minutes files could be created to allow users to view their personal annotations as well. It might also be beneficial at some point to merge the projected items with a person's own notes to create a very personalized set of minutes. In some cases, users may want a document not placed on the public stack to be included in the meeting minutes. EMCE has a virtual "minutes keeper" that appears in the form of an icon, and anything dragged onto it is included in the minutes document. EMCE also archives the minutes for access during subsequent meetings. Discussion topics are tracked so that a list of past meeting minutes relevant to the topic at hand can be proactively created. At any time, a participant can bring up an old set of minutes either by selecting one from the list or by entering a keyword.

5.2.4 Note Taker

Meeting notes are often cumbersome to take and most frequently are lost. Although a laptop can be used for note taking, most people grew up taking notes with pen and paper, and they feel less comfortable doing so with a keyboard. (In addition, there is a social cost to the clacking of the keyboard). Furthermore, many people include drawings as well as text in their notes, so a keyboard is inappropriate. We solved the problem by creating an electronic note taker that is used much like pen and paper. Because the system is electronic, the notes can easily be archived for easy retrieval and the text and diagrams can be beautified for easier reading. An extra feature, not possible with traditional pen and paper note taking, is the ability to add pictures or objects that are presented during a meeting. A user can include an image in personal notes by simply dragging it into the note taker area. Eventually, the note archive should include the minutes created with a user's personal notes. We believe that both the personal note taker and the ability to write private annotations on the projected archive will be used because they serve slightly different purposes. One feature gives a user a blank slate on which to write and draw whatever is needed, and the other provides an explicit background on which to add personal reactions to a given object.

6. Discussion

With OfficeMATE and EMCE's basic features implemented we are ready to have them stand the real test—ongoing use by visitors and resident researchers at SRI. As we build the more permanent infrastructure necessary to support such an effort over time, we are designing studies that will help us understand if our interpretation of collaborative needs matches their actual use. The informal response to the prototypes is promising so far, and the features we have implemented are well supported by findings from previous workplace studies.

Many of those studies emphasize the mobility of office workers on both very small and local scales, as well as across larger distances. OfficeMATE supports mobility in two ways. One is the ubiquitous computing notion of picking up the right kind of support where you are going, and having it provide access to the things you need there. EMCE on the other hand provides mobility support by helping move the right data from your desktop to where you are, in the meeting. While the larger distance that EMCE supports is not the geographic distribution expected in most CSCW studies, it is an important difference in use that needs to be evaluated.

We have tried to get around problems of not carrying a computer into meetings in two ways. First, we have tried to make carrying more feasible by using a combination of a smaller form factor coupled with embedded infrastructure. In addition, providing content that is helpful for visitors adds incentive to the use of OfficeMATE. We hope that this in turn, will add incentive for its use as a collaborative tool. Second, by replacing keyboard and mouse interaction with speech and pen interaction we feel makes many interactions more natural.

The annotation and note taking features in EMCE support the findings from workplace studies that simple annotation is more appropriate than full collaborative writing applications. Using pen for both handwriting recognition and drawing, both on private space and public, provides the right kind of interaction tools for these uses. Although a keyboard is still available in the prototypes, we hope to ease the use of digital media by not requiring specialised skills, such as typing, or learning a new method of interaction like writing Graffiti. In this way we hope to fulfil the promise of early digital meeting environments by using computers for what they are great for: saving the state of your meetings and work and providing access to it at a later date.

Acknowledgements

We would like to acknowledge the work of the members of the Computer Human Interaction Center, CHIC! (http://www.chic.sri.com) at SRI International, its partners for the hardware and software they provided (http://www.chic.sri.com/projects/partners.html), and its visitors, without whom this work could not be done. We especially want to thank Arna Ionescu, David Hoze, and Manuel Godbert for their work on EMCE and OfficeMATE.

References

1. Reeves, B. and C. Nass, *The Media Equation: How people treat computers, television, and the new media like real people and places.* 1996, Stanford, CA.: CSLI Publications and Cambridge University Press.
2. Bannon, L. and K. Schmidt, *CSCW: Four characters in search of a context.*, in *Studines in Computer SUpported Cooperative Work*, J. Bowers and S. Benford, Editors. 1991, North-Holland: Amsterdam. p. 3-16.
3. Schmidt, K. and L. Bannon, *Taking CSCW seriosly: Supporting articulation work.* Computer Supported Cooperative Work (CSCW): An International Journal, 1992. **1**(1-2): p. 7-40.
4. Bardram, J. *Designing for the Dynamics of Cooperative Work Activities*. in *Computer Supported Cooperative Work (CSCW)*. 1998. Seattle Washington: ACM.
5. Cheyer, A. and L. Julia. *Multimodal Maps: An agent based approach.* in *Cooperative Multimodal Communication (CMC95)*. 1995. Eindhoven, The Netherlands.
6. Cheyer, A. and L. Julia, *Multimodal Maps: An agent-based approach.*, in *Multimodal Human-computer Communication, Lecture Notes in Aritifical Intelligence #1374*, Bunt, Beun, and Borghuis, Editors. 1998, Springer. p. 111-121.
7. Julia, L. and J. Bing. *TravelMATE: A demonstration of SRI's multimodal augmented tutoring environment*. in *IWAR 99*. 1999. San Francisco.
8. Weiser, M., *The world is not a desktop.* ACM interactions, 1993. **1**(1): p. 7-8.
9. Karat, C.-M.,C. Halverson, D. Horn, and J. Karat. *Patterns of entry and correction in large vocabulary continuous speech recognition systems.* in *CHI'99-Human Factors in Computing Systems*. 1999. Pittsburgh, PA.: ACM.
10. Clark, H.H. and S.E. Brennan, *Grounding in Communication*, in *Shared Cognition: Thinking as social practice.*, J. Levine, L.B. Resnick, and S.D. Behrand, Editors. 1991, APA Books: Washington, DC.
11. Oviatt, S., *Predicting spoken disfluencies during human-computer interaction.* COmputer Speech and Language, 1995. **9**: p. 19-35.
12. Yankelovich, N., G.A. Levow, and M. Marx. *Designing Speech Acts: Issues in speech user interfaces.* in *CHI 95*. 1995. Denver, CO.
13. Karat, J., *Scenario use in the design of a speech recognition system.*, in *Scenario-based design.*, J. Carroll, Editor. 1995, Wiley: New York.
14. Lai, J. and J. Vergo. *MedSpeak: Report creation with continuous speech recognition.* in *CHI 97*. 1997. Atlanta GA: ACM.
15. Julia, L. and C. Faure. *Pattern recognition and beautification for a pen based interface*. in *ICDAR*. 1995. Montreal, Canada.
16. Oviatt, S. *Multimodal interfaces of dynamic interactive maps*. in *CHI'96-Human Factors in Computing Systems*. 1996. Vancover, BC, CANADA: ACM.
17. Whittaker, S., D. Frohlich, and O. Daly-Jones. *Informal workplace communication: What is it like and how might we support it?* in *CHI'94 - Human Factors in Computing Systems*. 1994. Boston, MA.: ACM.
18. Bellotti, V. and S. Bly. *Walking away from the desktop computer: Distributed collaboration and mobility in a product design team*. in *Computer Supported Cooperative Work (CSCW)*. 1996. Cambridge, MA.: ACM.
19. Luff, P. and C. Heath. *Mobility in Collaboration*. in *Computer Supported Cooperative Work*. 1998. Seattle WA: ACM.
20. Kraut, R., C. Egido, and J. Galegher, *Patterns of contact and communication in scientific research collaboration.*, in *Intellectual Teamwork*, J. Galgeher, R. Kraut, and C. Egido, Editors. 1990, Lawrence Erlbaum Press: Hillsdale, NJ. p. 149-170.
21. Halverson, C., *Inside the cognitive workplace: new technology and air traffic control.*, in *Dept. of Cognitive Science*. 1995, University of California: San Diego. p. 371.
22. Hutchins, E., *Cognition in the Wild*. 1995, Cambridge, MA.: MIT Press.
23. Heath, C., P. Luff, and A. Sellen. *Reconsidering the Virtual Workplace: Flexible Support for Collaborative Activity*. in *European Conference on Computer Supported Cooperative Work*. 1995. Stockholm, Sweden.
24. Ackerman, M.S. and C.A. Halverson. *Organizational memory: Processes, Boundary Objects, and Trajectories.* in *32nd Hawaiin International Conference on Systems Science (HICSS99).* 1999.
25. Ackerman, M.S. and C.A. Halverson. *Considering an organization's memory.* in *Computer Supported Cooperative Work*. 1998. Seattle WA: ACM.

26. Stefik, M. G. Foster, D. Bobrow, K. Kahn, S. Lanning and L. Suchman, *Beyond the Chalkboard: Computer support from collaboration and problem solving in meetings.* Communications of the ACM, 1987. **30**(1): p. 32-47.
27. Martin, D., A. Cheyer, and D. Moran, *The open agent architecture: A frameword for building ditributed software systems.* Applied Artifical Intelligence: An International Journal, 1999. **13**(1-2): p. 91-128.
28. Cheyer, A. and L. Julia. *InfoWiz: An animated voice interactive information system.* in *Agents 99 (WS Communicative Agents)*. 1999. Seattle, WA.
29. Dowding, J., *et al. GEMINI: A natural language system of spoken-language understanding*. in *31st Annual Meeting of the Association for COmputational Linguistics*. 1993. Ohio State University, Columbus OH: ACL.
30. Julia, L.,A. Cheyer, J. Dowding, H. Bratt, J. M. Gaawron, E. Bratt, and R. Moore*. How natural inputs aid interaction in graphical simulations.* in *VSMM98*. 1998. Gifu, JAPAN.
31. Julia, L., L. Heck, and A. Cheyer. *A speaker identification Agent*. in *AVBPA 97*. 1997. Crans Montana, Switzerland.
32. Ionescu, A. and L. Julia. *EMCE: A multimodal environment augmenting conferencing experiences*. in *First Australian User Interface Conference (FAUIC'00)*. 1999, submitted. Canberra, Australia.
33. Rekimoto, J. and M. Saitoh. *Augmented Surfaces: A spatially continous work space for hybrid computer environments*. in *CHI 99 Conference Companion*. 1999. Pittsburgh PA: ACM.