The Best of Two Worlds:
Merging Virtual and Real for Face to Face Collaboration

Desney S Tan[*], School of Computer Science, Carnegie Mellon University
Ivan Poupyrev, MIC Research Labs, ATR
Mark Billinghurst, HITlab, University of Washington
Hirokazu Kato, Faculty of Information Sciences, Hiroshima City University
Holger Regenbrecht, DaimlerChrysler Research and Technology
Nobuji Tetsutani, MIC Research Labs, ATR

ABSTRACT
In its simplest form, reality is merely information that is presented or acquired. Mixed Reality (MR) is built around the integration of real world physical and computer generated virtual information. We do not use the term *augmented reality (AR)* because we view the merging of both worlds as a symbiosis, with desirable properties from each accentuated and complementing each other, rather than the enhancement of one with the other. Collaborative MR allows multiple participants to simultaneously share a physical space while being surrounded by a virtual space that is registered with the physical. Because the MR world inherits the properties of real and virtual worlds, it is rich with social context, spatial cues, and tangible objects from the real world as well as flexible digital information from the virtual. We believe that Mixed Reality is a medium, largely unexplored, but very well suited for face-to-face collaboration.

Introduction
We first examine the advantages that real and virtual worlds afford us for face-to-face collaboration. In this way, we are equipped to find a region in the design space where interaction between the two worlds harnesses the full potential of each, creating a world that is greater than the sum of its parts. We discuss some of the issues that arise in this medium as well as some of our experiences building novel interfaces and applications.

Real World Collaboration:
Social Context, Spatial Cues, Tangible Objects
In real world face-to-face collaboration, we extensively use a wide spectrum of verbal and non-verbal cues including gestures, body posture, facial expressions, and the use of personal space. This information augments conscious communication by giving us social awareness of other users, making collaboration a rich interaction.

Unlike traditional digital systems, which typically have a small number of output channels, the real world presents a single continuous large display with practically infinite resolution. In fact, objects as well as the space that contains them convey useful information about the world. Just as we convey and derive a large amount of information from social actions, so too is this the case with the environment.

Another advantage of spatial cues is that they are strongly tied to memory. People are very proficient at using the spatial locality of objects to reference items in memory. Since spatial cues are the same for different users, they allow several people to reference information using the same cue. If everyone who had to read these proposals in the same room with each of them pinned on the wall, you could point to the first paper on the left and would never have to refer to the title "The Best of Two Worlds." In fact, at some point, the paper may be removed and you could continue to point to the empty space and talk about the paper as if it were still there.

The physical world is intrinsically a shared one. Although people each have personalized points of view, there exists a common space and means for discourse. Synchronous sharing and referencing of objects and viewpoints is as simple as handling the same object, passing it back and forth, or pointing and looking at an object. This leads to parallelism for multiple users performing tasks.

Manipulation of physical objects is so well integrated into our daily lives that most people do not think of the real world as possessing an interface. The tangible manipulation of physical objects provides simultaneous use by more than one user, as well as affordances for intuitive use. Consequently, they are able to spend their efforts on the task at hand rather than coordinating activities. Another advantage of physical objects is that they are rapidly re-configurable. One could, for example, tear a corner of a newspaper page and write a phone number and stick it on a board for others to see, use it as a bookmark, or fold it to describe a particular shape.

[*] Participating Student Author contact: Desney S Tan. ATR Media Integration & Communications Research Lab, 2-2-2, Hikaridai, Seika-cho, Souraku-gun, Kyoto 619-0288, Japan. 0774-95-1429 or 0774-73-9675. desney@cs.cmu.edu.

Virtual World Collaboration: Unbounded, Flexible Information

Virtual worlds were designed to augment humans and provide them with the capability to manipulate information in ways that are not normally possible in the real world. This includes efficient means to store, retrieve, process, and communicate artifacts as well as symbolic information and methodologies. Unlike the somewhat passive physical objects in the real world, information stored in this world can potentially be active, accessing other information and updating themselves.

Since virtual worlds are completely conceived and created in the mind of the human, it is under her complete control and is not bounded by physical laws. For example, it is not uncommon to scale time, quickly moving forward into the future or undoing actions that happened in the past. In addition, designers may scale space, creating objects that are of different sizes to convey spatial information or to make them easier to manipulate. Also, they can create objects that exist in multiple places at multiple scales or give users different views of the same object. The designer is constrained only by her own imagination in augmenting the collaborators with tools for interaction.

Previous work

Researchers working on Computer Supported Collaborative Work have recognized the advantages that real and virtual worlds afford. However, most work has been largely focused on building tools that leverage on a subset of one or the other for effective communication and collaboration. On a separate front, there has been a significant amount of work done in MR technology [1,9]. This includes the visual tracking and registration techniques, displays, input mechanisms, and interface design. In our work, we attempt to combine principles behind real world face-to-face collaboration with novel technology to create effective MR interfaces for face-to-face collaboration.

A large amount of work has been done in human-computer interaction research to explore different media and interfaces to present information for collaboration. Some common media used are the desktop (or some variation thereof), Virtual Reality (VR), and Tele/Video conferencing [2,4]. Desktop interfaces are a poor medium for complex collaboration because of their small display portals as well as their limited capabilities for interesting spatial input mechanisms. VR interfaces construct synthetic spaces where users can meet and interact. Because users are taken out of the real world to be completely immersed within the virtual environment, they often lose important social cues. Tele/video conferencing, while providing some social cues, is normally projected in 2D and does not give participants the full advantage of spatial context. Also, most systems are not easily scalable to large numbers of participants [3, 5].

Mixing Reality for Collaboration

In Mixed Reality, we combine the social, spatial, and tangible properties of the real world with the flexible, accessible, replicable information created in the virtual. The usual "window into the world" and "immersive" metaphors imply a seam between two separate realities and are, in our opinion, confining. In MR collaboration, we give the user a multi-sensory experience, permitting them to use the real world as their interface. With this interface, we provide intuitive control to have information when, where, and however users want it. Users can thus concentrate on the tasks at hand without attending to technology's needs.

In creating collaborative MR environments, we encounter new and interesting research questions that do not exist in each of the independent worlds. The main issue in such worlds is the seamless integration of real and virtual, along with the relation of one to the other (semantic, spatial, or otherwise). There are many different ways of linking and manipulating co-located real and virtual objects and it is not immediately apparent how to best create a single coherent interface.

Even though collaborators have different viewpoints in the real world, they see and can refer to the same objects. In the virtual, it is sometimes desirable for different people to see different interpretations of the same information, or even different information in the same space. For example the electrical engineers might like to see the electrical wiring diagrams in a building, while the civil engineers would prefer the structural diagrams, and so on. An important issue to consider is the fact that because we are grounded in the real world, collaborators must be able to use spatial references, such as pointing to a particular object, just as they would in the real world. In MR worlds, the issue of consistency within and among users is an important one. Users must have some awareness of what other users are seeing as well as the abilities that they have. They should also be able to easily control what other users see and know of them.

Building on the philosophy that we can and should combine different media to exploit the advantages of each, we extend our work to integrate different interfaces to expand the design space and facilitate heterogeneous collaboration across media. Researchers have designed and implemented a myriad of different interfaces, each supporting special functionality and being useful for different tasks. In our increasingly complex world, specialization will continue, and effective work will require new combinations of skills drawn from collaborative teams of specialists. Each individual has diverse requirements and preferences for the presentation and manipulation of information. This manifests itself in the evolution of different devices, interfaces, and media. Each individual must be allowed to employ existing tools and media in the collaboration process, not only to present their own ideas, but also to view and edit those of others. We explore combining MR with such media as traditional desktops and VR, as well as with users not augmented with any technology.

Our Mixed Reality Research

The fundamental elements in augmented and mixed reality systems are techniques for tracking user position and/or viewpoint direction within physical environment, methods for registering virtual objects in physical environments, and rendering and presenting them to the user via augmented reality display devices, e.g. head-mounted displays and large-screen projection screens.

Our work makes heavy use of computer vision techniques for tracking and registering virtual objects [6,7]. We mark physical objects with simple square fiduciary patterns with unique identifying symbols. The video output from either static or mobile video camera is captured and image processing techniques are applied to identify square markers and symbols in the middle of markers. Because we are tracking square markers of known size, the relative camera position and orientation can be found in real time. Once this is known, the virtual camera can be placed at the same position so that 3D virtual objects appear to be exactly attached to markers (inset in Figure 1). The technique, optimized for speed, allows for fast and reliable tracking (~30 fps latest Intel Pentium-based computers).

At least two display configurations can be used in our mixed reality environments to support face-to-face collaboration, each with its own advantages and disadvantages: head-mounted displays (HMDs) and wall-sized projection screens. The choice between these approaches leads to very different interaction styles and interface designs. In the case of HMDs, users are presented with a traditional mixed reality display. Their view of the real world is augmented with virtual objects so that they can see other participants and virtual images (Figure 1). The major shortcoming of HMDs is that they are obtrusive: the users have to wear devices. In the large-screen configuration, the users can interact with each other without need to wear any display devices (Figure 2). However, in this case only a portion of the workspace can be augmented with virtual objects and users cannot maintain natural communication with other participants without interrupting their interaction with virtual and physical objects. An interesting approach would be combination of two, which would allow both instrumented and non-instrumented users to interact in the same mixed reality environment. The choice of the approach is defined by application requirements and one such particular application is described in the following section.

Figure 1: Users interacting with HMDs. The projection screen in the background allows other users to watch and interact.

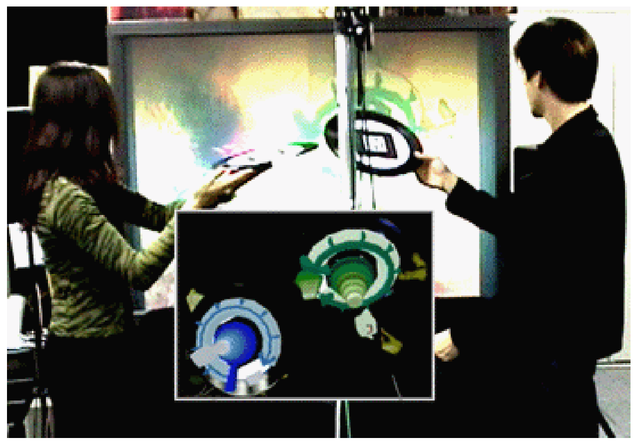

Figure 2: Users collaboratively playing music in Augmented Groove [8].

Application: Collaborative Design of Aircraft Instrument Panels

We explicitly consider implications of collaboration in real and virtual worlds, but also think about the applications in which such augmented collaboration is required and beneficial. We have explored such areas as collaborative engineering design, music, education, and leisure. We present one of our ongoing projects in the area of engineering design. In this project, we explore how MR interfaces can be used to support rapid collaborative prototyping in an industrial design application. Although our techniques are broadly applicable, our specific application area is the design of aircraft instrument panels, a joint research initiative carried out with support from DASA/EADS Airbus.

In our interface, we allow designers, engineers, human factors specialists, and aircraft pilots to quickly layout and rearrange a set of virtual aircraft instruments on a board simulating an airplane cockpit. The participants are able to easily and quickly evaluate the layout, rearrange instruments as necessary, add new instruments or remove those that are not needed. The design activity is inherently collaborative and involves team-based problem solving. It requires that the interface facilitates discussion and joint evaluation, allows use both of existing physical plans, schemes, documents and tools, but also of digital data.

The ability to simultaneously access and manipulate physical objects and digital data by several participants is a key requirement for the application. Furthermore, the interface should provide intuitive and easy techniques for traditional ‘house-keeping’ activities, e.g. removing virtual panels from the board, copying, deleting them, saving modifications, etc. The system that we have designed meets these goals by providing a tangible collaborative mixed reality interface, in which users arrange virtual instruments on a whiteboard by manipulating small marked cards that have images of virtual instruments attached to them (Figures 3,4).

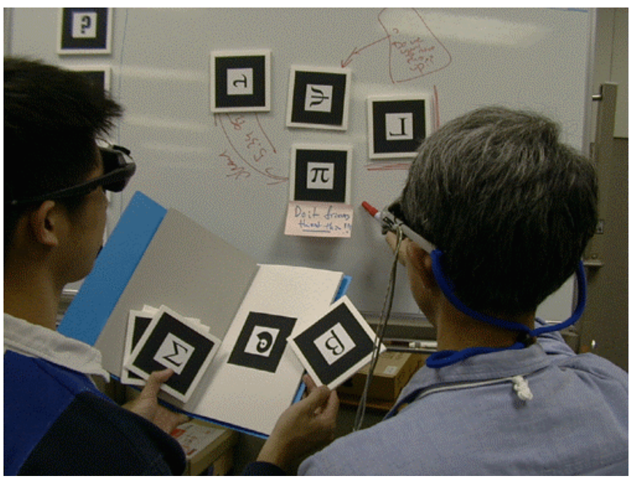

Figure 3: Users arranging instruments on the dashboard using physical markers.

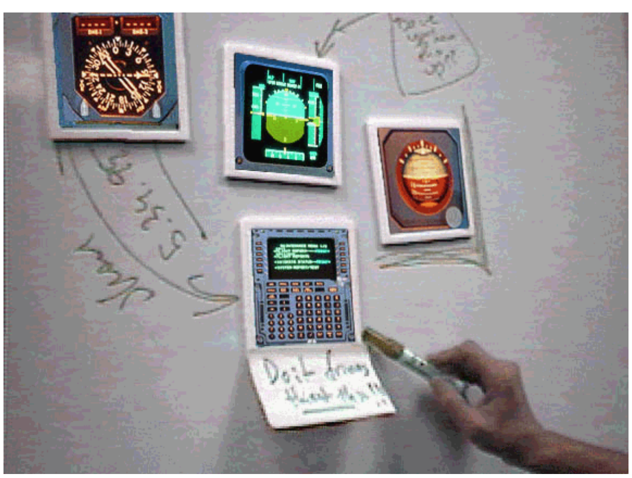

Figure 4: One user’s view of the real world augmented with virtual instruments.

The cards work as physical placeholders for virtual objects. The designer can assign any virtual instrument by copying it from the ‘mixed reality catalog’ of instruments: a real book with virtual models of the instruments overlaid on each page, an approach that was first introduced in the Magic Book system [1]. Users can copy the desired instrument onto the placeholder cards by placing them beside the book, i.e. docking the card to the book. The object moves from the page onto the card and becomes associated with that card. Each card is equipped with a magnet and can be placed on a whiteboard, built to resemble and represent an actual aircraft dashboard. To layout the instruments, several participants can rearrange cards on the board, physically manipulating virtual instruments, passing them to each other, adding new instruments and removing them from the board if needed.

Besides physical placeholders we have designed command widgets that allow the user to perform simple operations on cards, such as copying and deleting objects. To invoke the operation the user docks the appropriate widget and card to each other: for example docking a copy card with any instrument card would copy the instrument to the copy card or vice versa depending on whether the instrument card were empty or not.

Our early experiences have indicated a number of advantages to this approach. First, manipulation of physical objects to control virtual objects is quick and intuitive both for individuals and for groups of users. Second, participants are provided with immediate and rich feedback including visual, tactile, proprioceptive, as well as variety of social cues about other participants, leading to intuitive, effective and enjoyable interaction. Designers can also get a ‘real feel’ for the instrument panel, which is important in industrial design. Finally, the provision for select Windows, Icons, Menus, Pointing (WIMP) commands also allows designers to apply existing knowledge from familiar desktop interfaces to the mixed reality workspace. We are working to extend this system to allow users working within other media (such as on a desktop or in VR) to collaborate with the MR designers.

Conclusion

Examining the affordances of the real and virtual worlds in face-to-face collaboration, we can begin to construct a mixed space that seamlessly merges advantages of both. We described one of several ongoing projects in this area as an example of the utility of the MR medium. We believe that work in this area will stimulate further thought on issues of collaboration and provide us with the insight necessary to construct effective collaborative environments.

REFERENCES

1. Billighurst, M., Poupyrev, I., Kato, H., May, R., Mixing Realities in Shared Space: An Augmented Reality Interface for Collaborative Computing. Proceedings of ICME 2000. 2000. IEEE. pp. 1641-1644.
2. Billinghurst, M., Kato, H., Real World Teleconferencing. Proceedings of CHI'99, Extended Abstracts. 1999. ACM. pp. 194-195.
3. Grudin, J., and Poltrock, S. CSCW and Groupware: Experiences, State of Art, and Future Trends. Tutorial notes from CHI 2000.
4. Ishii, H., Miyake, N., Toward an Open WorkSpace: Computer and Video Fusion Approach of Team-Workstation. Communications of the ACM, 1991. 34(12): pp. 37-50.
5. Ishii, H., Ullmer , B., Tangible bits towards seamless interfaces between people, bits and atoms. Proceedings of CHI97. 1997. ACM. pp. 234-241.
6. Kato, H., Billinghurst, M., Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System, Proc. of 2nd Int. Workshop on Augmented Reality, pp.85-94 (1999). Proceedings of 2nd Int. Workshop on Augmented Reality. 1999. pp. 85-94.
7. Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., Tachibana, K., Virtual Object Manipulation on a Table-Top AR Environment. Proceedings of International Symposium on Augmented Reality. 2000.
8. Poupyrev, I., Berry, R., Kurumisawa, J., Nakao, K., Billinghurst, M., et al., Augmented Groove: Collaborative Jamming in Augmented Reality. Proceedings of SIGGRAPH'2000 Conference Abstracts and Applications. 2000. ACM. pp. 77.
9. Rekimoto, J., Nagao, K., The World through the Computer: Computer Augmented Interaction with Real World Environments. Proceedings of UIST'95. 1995. ACM. pp. 29-36.

PARTICIPATING AUTHOR BIOGRAPHY

Desney S Tan is a PhD student working with Dr Randy Pausch and the Stage 3 Research Lab at Carnegie Mellon University. His current research thrust is designing systems to surround users with creative influences in order to effectively augment the collaboration process. He works closely with Dennis Proffitt's Perceptual Psychology Lab at the University of Virginia in exploring the effects of technology in providing structure for organization of information and hence augmenting the human memory. His current focus is on building multiple media interfaces that allow participants to collaborate across various media.

Desney is currently an intern at the Media Integration and Communications Research Lab at the Advanced Telecommunications Research (ATR) Institute International in Kyoto, Japan. Here, he is working not only on developing new vision-based interaction techniques for mixed reality (MR), but also on integrating a system utilizing VR, MR and standard desktop interfaces for collaborative design of airplane instrument panels (a joint effort with the HIT-lab at the University of Washington and DaimlerChrysler). He has consulted for Walt Disney Imagineering, designing and user testing several VR attractions now at the DisneyQuest Interactive Theme Parks in Orlando and Chicago. It is this work that first got him interested in the collaboration across multiple media. He has also recently interned with the Adaptive Systems and Interaction group at Microsoft Research, building 3D navigation techniques for use with novel projection displays.

See also http://www.cs.cmu.edu/~desney/resume.htm