

Collaborative Augmented Reality

Mark Billinghurst, Hirokazu Kato

Science fiction has foreshadowed many of the great technical advances in communication. In Kubrick's 2001: A Space Odyssey, Dr. Floyd calls home using a videophone, one of the first film appearances of videoconferencing. Little more than a decade later, in Star Wars, collaboration is accomplished using live-sized virtual images superimposed over the real world. Twenty years later, desktop videoconferencing is freely available, while the virtual holo-phone is still far off fiction.

Ironically, in 1965, just three years before 2001 was released, Ivan Sutherland developed a technology that made it possible to overlay virtual images on the real world. Attaching two head worn miniature Cathode Ray Tubes to a mechanical tracker he created the first head mounted display (HMD). With this display users could see a simple virtual wire-frame cube superimposed over the real world, creating the first Augmented Reality (AR) interface. The term Augmented Reality is often used to refer to interfaces in which two- and three-dimensional computer graphics are superimposed over real objects, typically viewed through head-mounted or handheld displays.

Since that time, single user Augmented Reality interfaces have been developed that enable a person to interact with the real world in ways never before possible. For example, a doctor can see a virtual ultrasound image overlaid on a patient's body, allowing her to have "X-Ray" vision in a needle biopsy task, while a soldier has his own personal head up display that overlays targeting information on the battlefield. Azuma provides an exhaustive review of current and past AR technology and applications [Azuma 2001].

Although single user AR applications show great promise, perhaps the greatest potential use for Augmented Reality is for developing new types of collaborative interfaces. Augmented Reality techniques can be used to enhance face-to-face and remote collaboration in ways that is difficult with traditional technology. In this article we discuss some of the limitations of current collaborative interfaces and then describe collaborative Augmented Reality interfaces and the lessons that can be learnt from them.

Current Collaborative Technology

There are shortcomings with most current collaborative technology, especially when used to interact with spatial content. In face-to-face collaboration, people use speech, gesture, gaze and non-verbal cues to attempt to communicate in the clearest possible fashion. However, in many cases the surrounding real world or real objects play a vital role, particularly in design and spatial collaboration tasks. Physical objects support collaboration both by their appearance, the physical affordances they have, their use as semantic representations, their spatial relationships, and their ability to help focus attention. Real objects are also more than just a source of information, they are also the constituents of the collaborative activity, create reference frames for communication and alter the dynamics of interaction, especially in multi-participant settings [Minneman 96].

In contrast, most computer interfaces for co-located collaboration create an artificial seam between the real world and the shared digital task space. Ishii defines a seam as a spatial, temporal or functional constraint that forces the user to shift among a variety of spaces or modes of operation [Ishii 94]. People looking at a projection screen or crowded around a desktop monitor are less able to refer to real objects or use natural communication behaviors. For example, observations of the use of large shared displays have found that simultaneous interaction rarely occurs due to the lack of software support and input devices for co-present collaboration [Pedersen 93].

Recently, researchers have been investigating computer interfaces based on real objects. Ishii's Tangible Media group and others have explored how Tangible User Interfaces (TUI) can be used to support collaboration [Ishii 97]. TUI projects develop the theme of how real world objects can be used as computer input and output devices. Tangible interfaces are extremely intuitive to use because physical object manipulations are mapped one-to-one to virtual object operations. However, information display can be a challenge. It is difficult to dynamically change an object's physical properties, so most information display is confined to image projection on objects or augmented surfaces. In those Tangible interfaces that use three-dimensional graphics there is also often a disconnection between the task space and display space.

Augmented Reality technology has the potential to overcome many of these shortcomings and enable more natural communication. For co-located collaboration Augmented Reality can blend the physical and virtual worlds so that real objects can be used to interact with three-dimensional digital content and increase shared understanding. Tangible interaction techniques can be combined with AR display techniques to develop interfaces in which physical objects and interactions are equally as important as the virtual imagery and provide a very intuitive interaction metaphor. We call this combination of Tangible interaction and AR display *Tangible Augmented Reality* (Tangible AR). Later in the paper we show how application of the Tangible AR metaphor can produce AR interfaces that significantly enhance face-to-face collaboration.

Most technology for remote collaboration also affects communication. Audio-only interfaces remove the visual cues vital for conversational turn taking, leading to increased interruptions and overlap, difficulty in disambiguating between speakers, and in determining other's willingness to interact [Hindus 96]. With conventional video conferencing subtle user movements or gestures cannot be captured, there are no spatial cues between participants, the number of participants is limited by monitor resolution and participants cannot easily make eye contact. Speakers cannot know when people are paying attention to them or are able to hold side conversations. Sellen summarizes several decades of telecommunications research by reporting that the main effect on communication is the presence of mediating technology rather than the type of technology used [Sellen 95]. It is difficult for technology to provide remote participants with the same experience they would have if they were in a co-located meeting.

Recently, researchers have explored how desktop and immersive collaborative virtual environments (CVEs) can provide spatial cues to support group interactions. These

interfaces restore some of the spatial cues common in face-to-face conversation, but they require the user to enter a virtual world separate from their physical environment. In addition it is very difficult to transmit the same fidelity of non-verbal communication cues present in a face-to-face meeting and so create the same level of presence. Perhaps closest to the goal of remote perfect tele-presence are the Office of the Future work of Raskar et. al. [Raskar 98]. In this case multiple cameras are used to capture and reconstruct a virtual geometric model and live video avatar of a remote user. The output is displayed on a stereoscopic projection screen. While this work is most impressive, one drawback is that the remote user is represented as a 2.5 D virtual model. It is impossible to move all the way around the virtual avatar and occlusion problems prevent a complete model of the body being transmitted. The work also shares the common limitations of projection-based interfaces in that the display is not portable, and that virtual objects can only appear relative to the projection surface.

In contrast, for remote collaboration AR technologies can provide spatial audio and visual cues that overlay a person's real environment. The remote participants are added to the users real world rather than from it. So once again AR technology provides a seamless blending of reality and virtuality. In the next sections we describe examples that illustrate the benefits of AR technology for face-to-face and remote collaboration.

Augmented Reality for Face-to-Face Collaboration

Augmented Reality can be used to enhance a shared physical workspace and create an interface for three-dimensional CSCW. One of the first interfaces to show the potential of AR for face-to-face collaboration was the StudierStube project of Schmalsteig et. al. [Schmalsteig 96]. In this on-going project they use see-through head mounted displays to allow users to collaboratively view 3D virtual models superimposed on the real world (figure 1). They report users finding the interface very intuitive and conducive to real world collaboration, because unlike other interfaces, the groupware support can be kept simple and mostly left to social protocols.

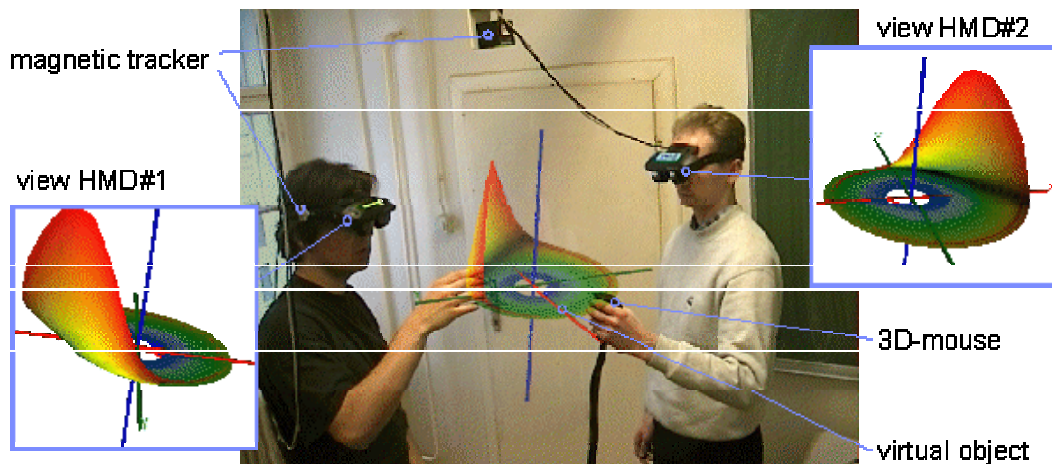


Figure 1: The StudierStube interface

The StudierStube researchers identify five key features of collaborative AR environments:

Virtuality: Objects that don't exist in the real world can be viewed and examined.

Augmentation: Real objects can be augmented by virtual annotations.

Cooperation: Multiple users can see each other and cooperate in a natural way.

Independence: Each user controls his own independent viewpoint.

Individuality: Displayed data can be different for each viewer.

The value of these characteristics is shown by several user studies that compare collaborative AR interfaces to other technologies. Kiyokawa et. al. have conducted an experiment to compare gaze and gesture awareness when the same task is performed in an AR interface and an immersive virtual environment [Kiyokawa 2000]. In their SeamlessDesign interface users are seated across a table from one another and engaged in a simple collaborative pointing task to compare gaze and gesture awareness between AR and VR conditions. Subjects performed significantly faster in the AR interface than in the immersive VR condition and also felt that this was the easiest condition to work together in. The performance improvement is largely the result of the improved perception of non-verbal cues that is supported by the collaborative AR interface.

Similarly, collaborative AR interfaces can produce communication behaviors that are more similar to unmediated face-to-face collaboration than to screen based collaboration. This is because when people collaborate at a table they can see objects on the table at the same time as each other, thus the task-space (the space containing the objects) is a subset of the communication space. However when users are collaborating in front of a screen the task space is part of the screen space, and may be separate from the interpersonal communication space. Thus while unmediated face-to-face collaboration and AR interfaces support seamless interaction, the screen-based interface may introduce a discontinuity that causes collaborators to exhibit different communication behaviors. In a recent experiment we explored this by comparing communication behaviors used to complete logic puzzle tasks in three conditions [Billingshurst 2002]:

- face-to-face collaboration with real objects
- co-located AR collaboration with virtual objects
- co-located projection screen based collaboration with virtual objects

The virtual objects were exact copies of the real objects and in the AR case they were attached to real objects so that Tangible AR manipulation techniques could be used.

Although the users did not feel that the face-to-face and AR conditions were very similar, they exhibited speech and gesture behaviors that were more alike in those conditions than in the projection condition. For example, in their deictic speech patterns there was no difference between the face-to-face and AR conditions, and their pointing behavior in the AR condition was more similar to the face-to-face condition than the projection condition. Deictic phrases are those that include the words "this" and "that" combined with pointing gestures. The tangible interface metaphor was particularly successful. Users felt that they could pick and move objects as easily in the AR condition as in the face-to-face condition and far more easily than the projection condition.

Although these results are encouraging they are just the beginnings of a deeper exploration. The first user studies have focused on object-centered collaboration where users are actively engaged in working together. However past work on teleconferencing has found that negotiation and conversational tasks are even more sensitive to differences between communication media. Future experiments need to be conducted where object manipulation is a small part of the task at hand.

Augmented Reality for Remote Collaboration

Our early AR Conferencing interfaces were developed to explore how AR could be used to support remote collaboration, and to provide gaze and non-verbal communication cues. In our first interface a user wore a lightweight headmounted display with a camera attached and was able to see a single remote user appear attached to a real card as a life-sized, live virtual video window [Billinghurst 2000]. The overall effect was that the conference collaborator appeared projected into the local user's real workspace (figure 2).

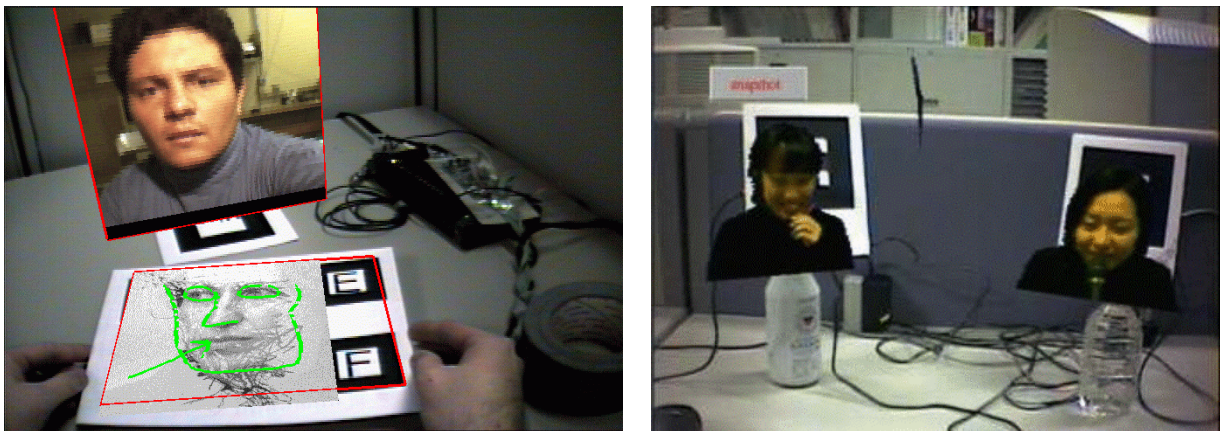


Figure 2: Live virtual video avatars for remote collaborative AR interfaces

Since the cards are physical representations of remote participants, our collaborative interface can be viewed as a variant of the Ishii's tangible interface metaphor [Ishii 97]. Users can arrange the cards about them in space to create a virtual spatial conferencing space and the cards are also small enough to be easily carried, ensuring portability. The user is no longer tied to the desktop and can potentially conference from any location, so the remote collaborators become part of any real world surroundings, potentially increasing the sense of social presence.

It is obvious that there are a number of other significant differences between the remote AR conferencing interface and traditional desktop video conferencing. The remote user can appear as a life-sized image and a potentially arbitrary number of remote users can be viewed at once. Since the virtual video windows can be placed about the user in space, spatial cues can be restored to the collaboration. Finally, the remote users image is entirely virtual so a real camera could be placed at the users eye point allowing support for natural gaze cues.

In a user study that compared AR conferencing to traditional audio and video conferencing subjects reported a significantly higher sense of presence for the remote user in the AR conferencing condition and that it was easier to perceive non-verbal communication cues [Billinghurst 2000]. One subject leaned in close to the monitor during the video conferencing condition, and moved back during the AR condition to give the virtual collaborator the same body space as in a face-to-face conversation!

In more recent work we have developed an AR conferencing interface that supports multiple remote users and applies alpha mapping techniques to extract video of the remote user from the background and create a more natural image (figure 2). In a user study with this interface users felt that the AR condition provided significantly more co-presence and improved the understanding of the conversational relationships between participants.

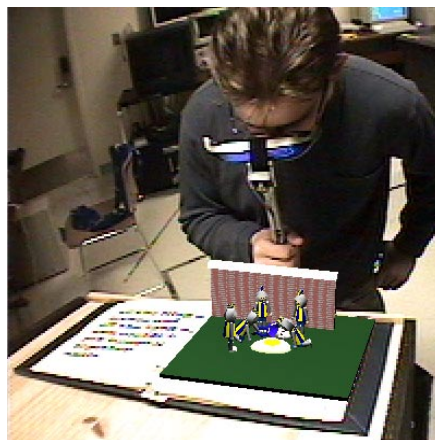
These results represents a powerful vision of remote conferencing; with remote participants not in separate windows on a screen but spread around the user's environment, positioned in space where the user prefers. Another intriguing characteristic of AR technology is its ability to support collaborative interfaces that smoothly move the users from real to immersive virtual environments, as discussed in the next section.

Transitional Interfaces

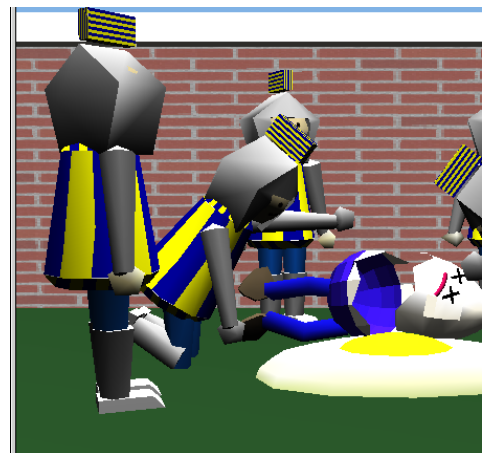
Physicality, Augmented Reality and immersive Virtual Reality are traditionally separate realms that people cannot seamlessly move between. However, human activity often cannot be broken into discrete components and for many tasks users may prefer to be able to move seamlessly between the real and virtual worlds. This is particularly true when interacting with three-dimensional graphical content. For example if collaborators want to experience a virtual environment from different viewpoints or scale then immersive VR may be the best choice. However if the collaborators want to have a face-to-face discussion while viewing the virtual image an AR interface may be best.



3a: Reality



3b: Augmented Reality



3c: Immersive Virtual Reality

Figure 3: Using the MagicBook to move between Reality and Virtual Reality

AR techniques can be used to develop interfaces that smoothly move between the real and virtual worlds. This is shown by the MagicBook work [Billinghurst 2001] in which a real book is used as a Tangible AR interface object. Several users can read the book like normal. However, when they look at the book pages through a handheld display they see three-dimensional virtual scenes overlaid on the real page. When users see a scene they like they can “fly” into it and experience it immersively (figure 3). In this way the book is a transport mechanism that can move people between reality and virtuality.

The MagicBook also supports collaboration. Several readers can read the same book. If these people then pick up their AR displays they will see the virtual models superimposed over the book pages from their own independent viewpoints. Multiple users can also be immersed in the virtual scene where they will see each other represented as virtual characters. More interestingly, one or more users can be immersed in the virtual world, while others are viewing the content as an AR scene. In this case the AR user will see exocentric views of miniature figures of the immersed users, while in the immersive world, users viewing the AR scene appear as large virtual heads looking down from the sky. Thus, unlike traditional collaborative interfaces, the MagicBook supports multi-scale collaboration on several different levels; as a physical object, as an AR objects, and as an immersive virtual space.

Conclusions

Augmented Reality techniques can be used to develop fundamentally different interfaces for face-to-face and remote collaboration. This is because AR provides:

- Seamless interaction between real and virtual environments
- The ability to enhance reality
- The presence of spatial cues for face-to-face and remote collaboration
- Support of a tangible interface metaphor
- The ability to transition smoothly between reality and virtuality

In this paper we have provided several examples of the types of interfaces that can be produced from taking advantage of these characteristics. Despite early promising results there is a lot of research work that needs to be done before collaborative AR interfaces are understood as well as traditional telecommunications technology. Rigorous user studies need to be conducted on a variety of tasks and interface types. Hybrid interfaces that integrate AR technology with other collaborative technologies need to be further explored. More socially acceptable display and input devices need to be developed. When these areas have been addressed then George Lucas’ vision of teleconferencing will become an everyday reality.

Acknowledgements

The authors would like to thank the many people who have contributed to the projects mentioned in the paper. In particular, Ivan Poupyrev, Kiyoshi Kiyokawa, Tom Furness and the staff and students at the University of Washington’s Human Interface Technology Laboratory. Without them none of this would have been possible.

References

- [Azuma 97] Azuma, R. A Survey of Augmented Reality, *Presence*, 6, 4, pp.355-385, 1997
- [Billinghamurst 2000] Billinghamurst, M., Kato, H. (2000) Out and About: Real World Teleconferencing. *British Telecom Technical Journal (BTTJ)*, Millennium Edition, Jan 2000.
- [Billinghamurst 2001] Billinghamurst, M., Kato, H., Poupirev, I. (2001) The MagicBook: A Transitional AR Interface. *Computers and Graphics*, November 2001, pp. 745-753.
- [Billinghamurst 2002] Billinghamurst, M., Kato, H., Kiyokawa, K., Belcher, D., Poupirev I. (2002) Experiments with Face to Face Collaborative AR Interfaces. *Virtual Reality Journal*, Vol 4, No. 2, 2002.
- [Ishii 94] Ishii, H., Kobayashi, M., Arita, K., Iterative Design of Seamless Collaboration Media. *Communications of the ACM*, Vol 37, No. 8, August 1994, pp. 83-97.
- [Ishii 97] Ishii, H., Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In proceedings of CHI 97, Atlanta, Georgia, USA, ACM Press, 1997, pp. 234-241.
- [Kiyokawa 2000] Kiyokawa, K., Takemura, H., Yokoya, N. "SeamlessDesign for 3D Object Creation," *IEEE MultiMedia, ICMCS '99 Special Issue*, Vol.7, No.1, pp.22-33, 2000.
- [Minneman 96] Minneman, S., Harrison, S (1996) A Bike in Hand: a Study of 3-D Objects in Design. In *Analyzing Design Activity*, N. Cross, H., Christiaans, K. Dorst (Eds), J. Wiley, Chichester, 1996.
- [Pedersen 93] Pedersen, E.R., McCall, K., Moran, T.P., Halasz, F.G. (1993). Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings. In *Proceedings of Human Factors in Computing Systems (InterCHI 93)* ACM Press, pp. 391-398.
- [Raskar 98] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin and H. Fuchs. The Office of the Future: A unified approach to image based modeling and spatially immersive displays. In *Proceedings of SIGGRAPH 98*, pp 179-188, 1998.
- [Schmalsteig 96] Schmalsteig, D., Fuhrmann, A., Szalavari, Z., Gervautz, M., Studierstube - An Environment for Collaboration in Augmented Reality. In *CVE '96 Workshop Proceedings*, 19-20th September 1996, Nottingham, Great Britain.
- [Sellen 95] Sellen, A. Remote Conversations: The effects of mediating talk with technology. *Human Computer Interaction*, 1995, Vol. 10, No. 4, pp. 401-444.