ClearSpace: Mixed Reality Virtual Teamrooms

Alex Hill, Matthew Bonner, Blair MacIntyre

Georgia Institute of Technology, 85 Fifth Street NW
30308 Atlanta, United States
{ahill, matt.bonner, blair}@gatech.edu

Abstract. We describe ClearSpace, a tool for collaboration between distributed teamrooms that combines components of virtual worlds and mixed presence groupware. This prototype is a starting point for exploring solutions to display and presence disparity by leveraging model-based user representations. We describe our deployed system and a mirroring approach that solves several problems with scaling up ClearBoard style portals to a common virtual space. We also describe techniques for enforcing consistency between heterogeneous virtual and physical contexts through system-managed awareness.

Keywords: Mixed Reality, Distributed Groupware, Mixed Presence Groupware

1 Introduction

In our global, information-driven economy, distributed workforces have become a fact of life. Highly distributed groups face significant challenges to collaboration and to forming a shared group culture. While issues such as working asynchronously or the loss of physical cues like gestures are key, perhaps the thorniest problem is a distributed workforce's heterogeneity. An in-house user study by Sun Microsystems found that 70% of distributed meetings involved a combination of co-located teamroom users and remote users [1]. Workers may be in offices or on trains, conversing at a table or on a mobile, or in one of a thousand other work contexts. Crafting a shared work context from a potentially infinite number of individual contexts is no easily specified task. Both virtual worlds and Mixed Presence Groupware (MPG) researchers have confronted this problem. Virtual worlds address user heterogeneity by creating a new, virtual world to be shared. MPG systems in contrast carefully track physical cues to bring remote users into an existing workspace, creating a shared "mixed" reality. Our ClearSpace project explores the capacity of a system to blend these the extremes of physically and virtually focused systems, combining virtual worlds and MPG approaches and gracefully handling discrepancies in device configuration to support many heterogeneous users. We have designed a system based on the open-source Project Wonderland virtual world toolkit, which uses a magic mirror technique to mix the physical and virtual work contexts into a single persistent work environment. In our prototype system, we have developed a whiteboard-style application inspired by Ishii's Clearboard [2], with virtual analogues. Rather than using video, we use several tracking systems to map physical motion to articulated avatars in the virtual world.

Our system has served as a base for exploring solutions to presence and display disparity as described by Tang et al [3]. We are particularly concerned with consistency between physical and virtual contexts, especially user perspectives and the physical or virtual placement of work artifacts within the combined teamroom. We believe realizing consistency on many levels including appearance, affordance and physical to virtual mapping is critical to overcoming presence and display disparities. The next section covers related work and the following section describes our prototype system. Section 4 discusses how a requirement for multiple access points to teamroom documents lead to a mirror-based approach rather than a window approach. Section 5 discusses the management and awareness of workspace documents. Section 6 describes how system managed awareness facilitates scaling across heterogeneous physical to virtual configurations.

2 Related Work

The virtual worlds approach to remote collaboration, exemplified by cAR/PE! and GAZE, minimizes individual physical context by creating a new, shared virtual context [4, 5]. In both systems, video representations of users are placed in a virtual environment or teamroom. GAZE and cAR/PE! also incorporate some degree of head tracking to communicate gaze or control position in the virtual world. While simple tracking can help communicate physical cues, no support for co-located users is provided, and all users may as well be remote. Such enforced homogeneity may overcome certain challenges of creating a shared context, but it does so at the expense of natural co-located interaction. Mixed Presence Groupware (MPG) emphasizes the physical. These systems, such as VIRTUE, Carpeno and VideoArms center on an augmented room or artifact like a table or vertical display that becomes a shared context [6-8]. Sophisticated human tracking systems are used to communicate as many physical cues as truthfully as possible. Rather than bringing users into a new virtual environment, these systems seek to bring remote users into the same physical realm as co-located users. Such specialized environments can be costly, making these systems difficult to scale. Users in a different context, perhaps on a mobile device at a field site, are unable to participate.

Display and presence disparity, described by Tang *et al*, are particularly important to our work [3]. Display disparity refers to the difficulty of sharing contexts or viewpoints. Users may have different sized displays or displays may be at different orientations, making referring to work artifacts cumbersome. Presence disparity refers to the feeling of co-presence that exists between collaborators. Users collaborating through a remote system are limited to whatever tracking or communication signals are supported, which can lead to a distinct lack of presence. A user whose communicative repertoire has been impoverished by their means of accessing the system may feel noticeably 'apart' from other collaborators. Explicit concern with display and presence disparity is a key difference between ClearSpace and some closely related work. Both VIRTUE and Carpeno for example combine interactive table environments with virtual worlds [6, 7]. In either system, one can work with a remote collaborator on a horizontal interactive table, or look 'across the table' at a vertical screen for visual representation of their partners. Our system also combines elements of MPG and virtual worlds, but our focus on display and presence disparity has led to interesting differences in our design. Perhaps the most obvious of these is our use of "mirroring", through which remote users appear to be physically located alongside physical users. Through this approach, we hope to create a vertical display that doesn't form a cognitive 'wall' between collaborators, even without the aid of a conjoined collaborative table.

3 System Configuration

Ishii's ClearBoard project had the benefit of placing video cameras near the remote user location behind the interaction surface, which also allowed the system to capture interactions with the whiteboard surface. A more conventional placement of a camera at surface depth does not allow the capture of hand interactions with that surface. A further drawback of using video in general is that it highlights differences in the user context such as lighting, clothing, background and even camera resolution and connection bandwidth. In contrast, a model-based representation tends to mask these differences, which aligns well with our approach, which is to reduce display and presence disparity through consistent representations for all users. We understand that any model-based representation we develop will fail to capture the subtlety and range of video. However, we believe that the increasing use of human tracking technology (i.e. Microsoft Kinect) suggests that these approaches will eventually capture as much if not more fidelity than video alone. Furthermore, we feel that starting from a model-based approach allows us to consider a much greater range of scenarios than normally apply when using video alone.

Before engaging in an analysis of how to best merge virtual worlds with interactive whiteboards we had to determine a hardware and software configuration that met some minimum requirements for capturing the user and their input. Feeling that tracking the head and hands brings the highest marginal value in representing overall physical embodiment, we chose a combination of vision-based head tracking and hand tracking in front of a large LCD panel. The prototype we developed consists of two 65" Sharp displays each mounted with multitouch overlays from NextWindow and Logitech QuickCam Pro 9000 cameras. For head tracking we used the FaceAPI library from SeeingMachines, which tracks the head location and orientation of a single user in three dimensions along with lips and eyebrows. Another commercial system, the ZCam from 3DV, had been demonstrated to track hands and recognize some simple gestures but became unavailable after Microsoft Corporation purchased it. As a fallback, we turned to an OptiTrack fiducial tracker with multiple cameras (6) focused on the interaction space in front of the LCD display. We deployed our virtual teamspaces using Sun Microsystems Wonderland open source Java library, which features full persistence and module-based customization. The base Wonderland system includes several features that support collaboration such as an interactive SVG whiteboard module and stereo directionalized voice conferencing. We developed a

VRPN client module for acquiring head and hand tracking data, and customized the avatar character class to perform arm inverse kinematics from hand position. We modified the SVG whiteboard module to include free-hand drawing and interactive client updates during gestures. We also added features that let us store and load documents using a WebDAV repository. We also created a Metadata module for tagging documents and a ProjectManager module that lets the user choose from a set of associated documents from within a HUD view.

4 Creating a Shared Collaborative Space Using Portals

One of our main goals was to allow scalability and flexibility in how the system is configured and used. Real world use of whiteboards in teamrooms provides the flexibility to leave content on multiple boards over time and to use any of those surfaces at any time. Like many other collaboration systems, the ClearBoard design creates a portal to the remote site on the other side of a display surface. While most systems use the display as a window into the virtual world we realized that this limits ability to scale to multiple portals. In order to implement an interactive whiteboard across the physical/virtual divide, the rendering into the virtual space has to be flipped across the vertical display plane. Using a window paradigm, a second portal on an adjacent wall in the room adds a second viewpoint into the virtual world from beyond another similarly placed portal in that space (Fig. 1a). The local portal user and a virtual collaborator can both turn in the same direction and transition their collaboration onto the second portal in a natural manner.

Unfortunately, using a window portal approach creates some problems. Given physical user P looking into a portal with another to their right, the arrangement of the virtual world will appear to be reversed with the second portal residing to the left of users in the virtual world. When the scene is later viewed through a virtual world browser (VWB), this inconsistency may interfere with workflow or even create confusion. A second issue arises when localized sound sources are added to the configuration. Sound is a critical part of distributed collaborations and directional sound has been shown to facilitate discriminating individual speakers amongst other competing voices [9]. If the sound of user V is directional, then its source will have to make an abrupt transition from one location to another as user P turns to the other view (Fig. 1a). Finally, having multiple viewpoints into the virtual space begs the following question: If a third user V_2 joins the collaboration through a VWB, do they find user P on the same side of the whiteboard surface as user V_1? If the user P viewpoint is indeed looking into the teamspace from outside then on which side should V_2 join the collaboration? If user P is indeed inside the teamspace with V_1 and V_2 his will inevitably lead to confusion as V_2 may believe they are standing next to user P when in fact they are across from him.

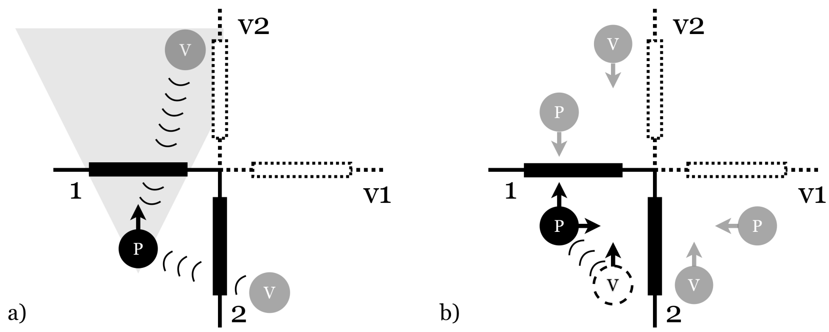


Figure 1. In a), a physical user P looks through window portal 1 at virtual user V for whom portal 2 appears to be on their left. If P turns to look through window portal 2 the audio for V must transition. In b), mirror portals avoid these issues since user P acknowledges that the portal is the right of V. As P turns the audio source for V inside the room remains consistent.

The solution to these problems lies in acknowledging that transitions between physical portal and similarly aligned virtual portal views behave in much the same way as collaborating in front of mirrors. The main implication being that a mirror view reflects back the virtual image of user P (Fig. 1b). Making this subtle change suggests that user P is actually sharing space with user V. Instead of acknowledging separate local space and remote spaces, the two spaces are fused into a common shared space. There is already some empirical support that virtual users respect personal space in much the same way as physical users do [10]. One can easily imagine a scenario where P and V_1 are collaborating and V_2 approaches from the rear. When P and V_1 see themselves and the approaching user V_2 sharing the same space they will naturally move aside, acknowledging the presence of user V_2 in their own space. A mirrored paradigm resolves questions about where each collaborator resides because all are collaborating in the same virtual teamspace side-by-side. A mirror paradigm allows gaze and gestures between users to work in a natural manner without software intervention. This approach also resolves discontinuities with directional sound. When user P acknowledges that V is standing to his side it becomes possible to locate the sound source of the collaborator within the same room (Fig. 1b). When turning to look into another portal, the sound location of V within the room stays consistent with their visual location in the other portal.

The foreshortening that happens close to a virtual camera makes showing activity within a virtual teamspace impractical with window portals. However, they can increase awareness of the virtual context for user P by creating views into the space surrounding the teamspace through virtual doors and windows. Analogous to their function in physical spaces, these apertures give feedback to virtual users indicating when they are visible to physical teamroom users. While head tracking can provide perspective corrected views in a mirror portal, window views are best rendered using a fixed viewpoint since they will be viewed from disparate viewpoints within the room. This general rule enforces another of our principles for the ClearSpace design, that user distance from portals should reduce the granularity of the information they provide.

5 Managing Teamroom Work Artifacts

Documents produced by portals include whiteboard drawings, affinity diagrams and introduced PDF documents, images and 3D models. A VWB user can ostensibly peel a finished document off of a portal and place it arbitrarily within the teamspace. Moreover, these VWB users can also travel arbitrarily to documents in order to access them. This raises an important question: Should all documents be portals in themselves? If each document is a portal in itself then two VWB users can conceivably collaborate on a document away from the physical portal. If physical portal users are prevented from joining in on the collaboration, then this lack of consistency in affordance will contribute to presence disparity. To avoid this, VWB users can travel to, access and manipulate documents in any location privately, but all collaborations must happen through portals. When users want to make their actions pubic and available for collaboration, they have to move a document onto a portal. This requires the user travel to a nearby portal, determine if another user is currently using it and coordinate placing a copy of the document into the portal. A VWB user accessing a document can also bring the document into a heads up display (HUD) mode. In HUD mode, the user has the option to see users publicly manipulating the document appear behind it. The decision to move to a portal can start by switching viewpoint to that of the portal that created the document. From there, the presence of other users, their activity and negotiation for portal space can begin.

A well-established design principle of Computer Supported Collaborative Work (CSCW) is the need to maintain awareness of collaborator activities. In many collaborative workspaces global actions are inherently visible to other users, while in virtual environments a common shared viewpoint cannot be guaranteed. Remaining aware of the viewpoint of virtual users via their location and head orientation helps in this regard. Some virtual environments such as Second Life do not enforce a linkage between user embodiment and viewpoint. In an effort to ensure consistency of both affordance and appearance, user viewpoint should always be tied to their embodiment. This is no longer the case when a user brings a document into a HUD view and some visual feedback to users becomes necessary to maintain awareness. In this situation, the VWB user can be automatically moved into a position in front of the associated document and returned to their position when the interaction ends. A semi-transparent trace avatar can also remain at the original location to avoid discontinuities in user awareness. This same technique can also be used when a VWB user decides to take HUD interactions to a portal for public collaboration. If the user decides to stay in front of the document or portal their trace avatar can move to join them there, making other users aware of the action.

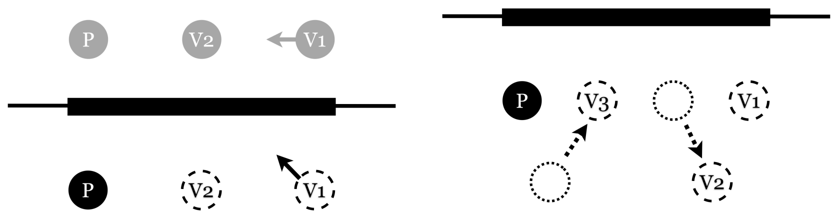
Although using a virtual world as a document repository makes user actions with those documents visible, forcing all users to travel to find and access documents is impractical. A useful HUD addition is an overview mode that includes thumbnails of associated work artifacts created by the same portal. When a VWB user employs this overview HUD they are no longer just interacting with a single document and a trace visualization of the overview HUD in front of the user is used. Access rights to user information determine the degree to which users can actually see the details of this behavior. To ensure consistency, this same overview mode is available to physical portal users as well. When a portal user takes his interactions private their view of the portal may change but those of other users at the portal remain the same. Portal users become aware of private actions of other users through a change in their appearance to semi-transparent and a trace of their activity. This trace can appear localized to an area around the virtual user in order to reduce any conflict between it and the document visible in front of the physical portal user.

A key ingredient in teamroom activities is the placement of documents in locations throughout the space. When a new document is introduced into the teamspace, a VWB user can ostensibly place that document anywhere in the environment. When introduced via a portal, either by whiteboarding or some other means, it is appropriate to place the document around or beyond that portal. Documents in the virtual world are associated with the portals that introduce them and should benefit from a visual indication by either restricting their location to a portal wall or visually indicated area around the portal. Physical portal users are somewhat more restricted in their ability to arrange documents within the scene. A simple heuristic for blank document creation is to position finished documents on the wall in sequence from left to right. When more explicit control over document position is required, a fixed third person view of the portal is appropriate. Using the same fixed view of the portal between physical and VWB users provides consistency. The fixed view around the portal should let users scale less prominent documents to a smaller size. This action can either result in scaling the document or only the appearance of scaling by pushing documents back in depth. Algorithmically adding depth to documents should be avoided if it begins to compromise the ability of teamroom users to appreciate document arrangement when viewed later through a VWB.

A portal view using a mirror paradigm will only include documents arranged behind the portal within its viewing frustum. And, a third person viewpoint will not facilitate arranging documents behind the user. One solution to this problem is to project a window portal view from a fixed perspective onto the wall containing the actual portal. This display around the portal creates a focus-and-context style display and gives the user arranging documents immediate feedback about their actions through peripheral vision. With hand tracking and some simple gesture recognition in front of the portal, a user can move documents off of the portal space and into position on the wall. Another way to make documents around a portal visible is to show the third person viewpoint instead of the mirrored view of the wall behind the user. Regardless of what is shown beyond the screen, some form of mode switching will be necessary to disambiguate between actions directed at the portal from those directed at content beyond it. A third person view of content on the wall around the user can also be created in a physical room by placing a mirrored wall at the back of the room. It may provide more spatial consistency to the user to use a virtual mirrored rear wall to construct the third person view that includes yet another view of the user within it.

6 Exploiting Model-Based Representations to Manage Awareness

One of the most significant advantages of a model-based user representation is the flexibility it affords. Procedural manipulation of user representation can be used to exaggerate awareness of user behaviors by amplifying their rendered signal [11]. This can also be extended to synthesizing behaviors such as gaze that have not been captured but are inferred from user behavior [12]. An example of how this might be exploited in a mirror portal setup involves awareness of gaze between two remote users. Given P side-by-side with virtual collaborators V_1 and V_2, when V_1 addresses V_2 he will only turn his head slightly to make eye contact with that user (Fig. 2a). When interpreted by the system, the awareness of this action to P can be enhanced visually by showing V_1 actually turning to face V_2. Once the underlying system is managing the gaze and gestures of users it also becomes feasible to have the system manipulate their apparent location. The space in front of portals will undoubtedly get crowded with the addition of multiple collaborators and the inclusion of mirrored avatars. Instead of having users explicitly coordinate their use of portal space, implicit heuristics can manage proxies of each user differently in each portal. Like some voice conferencing support systems, those users actively engaging in conversation or whiteboard activities can be represented near the front while those not participating are moved to a location around the periphery (Fig. 2b).



a) b)

Figure 2. In a), user V_1 turns slightly to look at V_2 but his representation turns to completely face V_2. In b), the representation of a less active V_2 moves to the periphery while a more active V_3 moves forward.

The straightforward way to ensure a consistent relationship between physical portals and virtual portals is for each to share the same size and relative location. More realistically, there will be mismatches between not only different physical teamroom setups but also between physical portals and their virtual equivalent. When the locations of portals do not align, the transformation of user position into those coordinate systems will create discontinuities as the user moves between portals. A model-based representation allows the system to leverage blending to generate consistent behaviors as users transition between different frames of reference. A simple heuristic that can be used in this context involves using the relative head orientation towards each portal. Following this approach, a user facing a portal would see virtual user V in the appropriate position relative to that portal (Fig. 3a). However, as user P turns to address another portal, the avatar of user V would appear to move as necessary to assume their appropriate position relative to the portal in view (Fig. 3b).

A more common discrepancy will be relative size, which results not only from physical portals of varying sizes but also from mobile devices. Transforming user behaviors between a small portal and a larger portal simply means applying them to an avatar sized appropriately for the larger context (Fig. 3c). Both user A interacting with a laptop and user B at a smaller portal will have their movement scaled up to fit their presentation. For laptop user A, gaze and mouse activity become a combination of gaze and body movement in the remote portal. Similar heuristics can also be applied to allow portals to pan and zoom into documents. The main result is that portal users no longer share the same viewing frustum of the portal or the users beyond it. When a remote user zooms into a document, the reduced viewing frustum is highlighted by a rectangle on the remote portal (Fig. 3c). When combined with their mirror reflection, this rectangle gives users visual feedback about their visibility to the remote user. Larger portals do not have to result in larger documents as sub-parts of whiteboard space can be saved off as individual documents. As portals increase in size to that of a wall, their effective work area becomes bounded by the size of the portal in the virtual teamspace and the viewing frustums of remote collaborators.

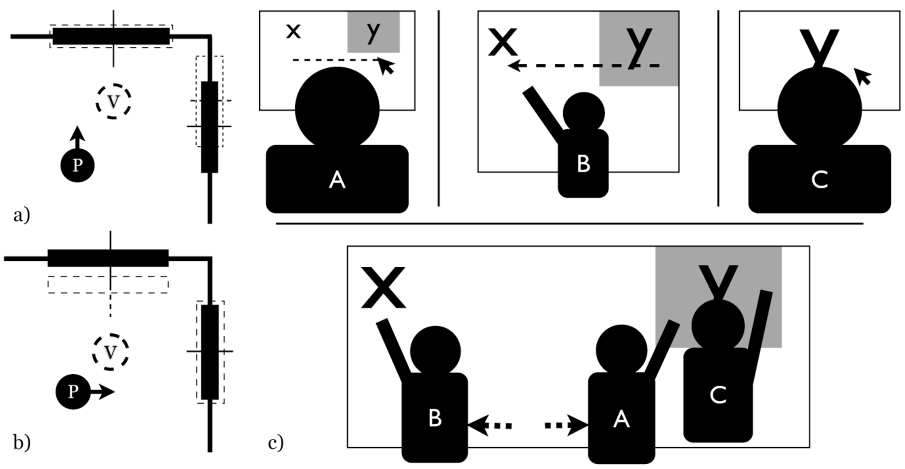


Figure 3. In a), user V appears in the actual position relative to the front portal and in b), appears relative to the side portal when P faces it. In c), laptop user A and remote user B both appear at the same scale in a larger portal. A reduced frustum of user C appears as a rectangle.

7 Discussion and Future Work

In this project, we developed a ClearBoard inspired distributed collaboration system based on the Wonderland Java library. Instead of using video, we used model-based representation in an effort to fully explore the potential of using mixed reality techniques and system-managed awareness to decrease display and presence disparity. In contrast to using a window into the virtual environment we found that a mirrored paradigm solved a number of problems with scalability and flexibility. Our analysis suggests that managing user awareness will help handle discrepancies in portal size and location. In addition to executing many of the unimplemented system features described here, our future work will include replacing our OptiTrack fiducial tracking system with one or more Microsoft Kinect controllers.

Acknowledgments. We would like to thank Sun Microsystems for their financial support, Jacob Schiefer, Aneesh Tarun and the Steelcase WorkSpace Futures team for their generous time spent inside Wonderland evaluating ClearSpace concepts.

References

1. Yankelovich, N., Simpson, N., Kaplan, J., Provino, J.: Porta-person: telepresence for the connected conference room. CHI '07 extended abstracts on Human factors in computing systems. pp. 2789–2794. ACM, New York, NY, USA (2007).
2. Ishii, H., Kobayashi, M.: ClearBoard. Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 525-532. , Monterey, California, United States (1992).
3. Tang, A., Boyle, M., Greenberg, S.: Understanding and mitigating display and presence disparity in mixed presence groupware. J. RES. AND PRACTICE IN INF. TECH. 37,
4. Regenbrecht, H., Ott, C., Wagner, M., Lum, T., Kohler, P., Wilke, W., Mueller, E.: An Augmented Virtuality Approach to 3D Videoconferencing. Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality. pp. 290–. IEEE Computer Society, Washington, DC, USA (2003).
5. Vertegaal, R.: The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 294–301. ACM, New York, NY, USA (1999).
6. Kauff, P., Schreer, O.: An immersive 3D video-conferencing system using shared virtual team user environments. Proceedings of the 4th international conference on Collaborative virtual environments. pp. 105–112. ACM, New York, NY, USA (2002).
7. Regenbrecht, H., Haller, M., Hauber, J., Billinghurst, M.: Carpeno: interfacing remote collaborative virtual environments with table-top interaction. Virtual Reality. 10, 95-107 (2006).
8. Tang, A., Neustaedter, C., Greenberg, S.: VideoArms: Embodiments for Mixed Presence Groupware. In: Bryan-Kinns, N., Blanford, A., Curzon, P., and Nigay, L. (eds.) People and Computers XX — Engage. pp. 85-102. Springer London, London (2007).
9. Yankelovich, N., Walker, W., Roberts, P., Wessler, M., Kaplan, J., Provino, J.: Meeting central: making distributed meetings more effective. Proc. 2004 ACM conference on Computer supported cooperative work. pp. 419–428. ACM, New York, NY, USA (2004).
10. Miwa, Y., Ishibiki, C.: Shadow communication: system for embodied interaction with remote partners. Proceedings of the 2004 ACM conference on Computer supported cooperative work. pp. 467–476. ACM, New York, NY, USA (2004).
11. Deng, Z., Ma, X.: Perceptually guided expressive facial animation. Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. pp. 67–76. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2008).
12. Steptoe, W., Oyekoya, O., Murgia, A., Wolff, R., Rae, J., Guimaraes, E., Roberts, D., Steed, A.: Eye Tracking for Avatar Eye Gaze Control During Object-Focused Multiparty Interaction in Immersive Collaborative Virtual Environments. Virtual Reality Conference, IEEE. pp. 83-90. IEEE Computer Society, Los Alamitos, CA, USA (2009).