Received: 14 March 2018 | Accepted: 21 March 2018
DOI: 10.1002/cav.1838

SPECIAL ISSUE PAPER

WILEY

Empowerment and embodiment for collaborative mixed reality systems

Ye Pan | David Sinclair | Kenny Mitchell

Disney Research Los Angeles, 521 Circle 7 Drive, Glendale, CA 91201, USA

Correspondence
Kenny Mitchell, Disney Research Los Angeles, 521 Circle 7 Drive, Glendale, CA 91201, USA.
Email:
kenny.mitchell@disneyresearch.com

Abstract

We present several mixed-reality-based remote collaboration settings by using consumer head-mounted displays. We investigated how two people are able to work together in these settings. We found that the person in the AR system will be regarded as the "leader" (i.e., they provide a greater contribution to the collaboration), whereas no similar "leader" emerges in augmented reality (AR)-to-AR and AR-to-VRBody settings. We also found that these special patterns of leadership only emerged for 3D interactions and not for 2D interactions. Results about the participants' experience of leadership, collaboration, embodiment, presence, and copresence shed further light on these findings.

KEYWORDS

augmented reality and virtual reality, telecollaboration

1 | INTRODUCTION

Collaboration at a distance has long been an important research goal of networked or multiuser augmented reality (AR) and virtual reality (VR) systems. With the launch of low-cost head-mounted displays (HMDs), networked mixed reality (MR) environments have rapidly increased in prevalence and popularity as a form of remote collaboration.[1]

We present several collaborative MR settings, allowing multiple users to visualize and edit a planet in an MR environment. Table 1 gives an overview of the settings and technologies used, and these are detailed in Section 3. For AR-to-AR setting, we provided each participant an AR system on the basis of HTC Vive headset coupled with Ovrvision Pro stereo camera. The HTC Vive was chosen because Vive base stations provide space tracking so we can realize markerless MR easily. In addition, the Vive controllers allow for high-quality user-friendly interaction experiences. These two AR systems were then networked, enabling two users to interact with a shared virtual scene and with each other in a face-to-face arrangement. This setting allowed the establishment of a common ground for our study. For AR-to-VRBody setting, the VR system is a Vive headset. The participants were physically in two separate rooms while working together. Each user's body could be represented by a jointed self-avatar

TABLE 1 Scenarios, labels, and technology used

Label	Site A	Site B
AR-to-AR	AR	AR
AR-to-VRBody	AR	VR with virtual body
AR-to-VR	AR	VR
AR-to-Desktop	AR	Desktop

Note. AR = augmented reality; VR = virtual reality.

that was dynamically controlled by head and hand controllers. For the AR-to-VR setting, each user was represented only by models of controllers. This representation is common in consumer VR applications at the moment. For AR-to-desktop setting, an AR system was linked with a desktop computer.

To evaluate the effectiveness of our settings, we conducted a user study to investigate how people interact with each other in MR environments, especially for spatially complex 3D environments. Our hypotheses are thus as follows:

- Hypothesis 1: We expected that the more immersed participant was singled out as the leader. The AR-to-desktop setting will have the highest leadership effect; next comes the AR-to-VR setting, then the AR-to-VRBody setting, and finally, this advantage will be lost in the AR-to-AR setting.
- Hypothesis 2: We further expected that this leadership effect only emerged in 3D interactions but not in 2D interactions.

The results revealed that Hypothesis 1 is only partially supported. We did not find the leadership effect in the AR-to-VRBody setting. Hypothesis 2 is supported. Our system demonstration and results thus motivate the further study of collaborative MR.

2 | RELATED WORK

2.1 | MR systems

Milgram et al.'s virtuality continuum is the seminal taxonomy of the field.[2] It classifies systems of MR and virtual visual content from pure real environments (e.g., video) at one end to a purely synthetic virtual environment at the other. MR occupies the range of the continuum between these extremes, merging both real and virtual objects together. MR systems use a range of technologies, including projection displays, situated displays, and head-mounted AR displays.[3]

Inspired by these recently developed systems, we developed four MR-based telecommunication settings with different levels of immersion and examine how these cutting-edge systems can be used in collaborative interactions.

2.2 | Collaboration and leadership

A previous study of a puzzle-solving task with three participants found that leadership varies between a virtual setting, in which the more immersed participant is singled out as the leader, and a real setting, where no one is singled out as the leader, both of which have the same task performed.[4]

Most of these previous work has focused on different types of VR systems.[4,5] Because AR systems provide with different levels of immersion, there is a need for a closer examination of the leadership/contribution to the task and different types of MR systems. Furthermore, it is unclear whether the leadership/contribution depends on the nature of the task to be performed so we include both 2D and 3D interaction types in our task design.

2.3 | Avatars

The impact of a self-avatar has been investigated in many ways (e.g., the visual embodiment of the user).[6,7] The self-avatar in a collaborative MR has crucial functions in addition to those of single-user MR environments, as the avatar is used for communication, including determining position, identification, visualization of focus of attention, and recognition of gesture and actions.[8,9]

Various papers demonstrated that avatars exhibiting higher levels of visual quality or tracking quality (e.g., eye tracking, facial expression, and finger tracking) can potentially communicate more subtleties of human nonverbal communication, enhancing the perceived authenticity of the interaction.[10–12] However, there are problems in providing a self-avatar because of uncanny valley and different discrepancies.[13–15]

The general thrust of these works indicates that self-avatars are important and that animation of the avatar can improve the effect of the self-avatar for most tasks. In this study, we aim to grow the existing knowledge on how the self-avatar (e.g., realistic and nonrealistic, avatars, and no avatar at all) alters users' behavior in collaborative MR.

FIGURE 1 Photos taken from third-person views for different conditions. (a) Augmented reality (AR)-to-AR @ site A. (b) AR-to-virtual reality (VR) @ site A. (c) AR-to-VR @ site B. (d) AR-to-Desktop @ site B

3 | SYSTEM DESIGN

In this section, we discuss the system design and implementation of the experiment application. The experiment was conducted at two very similar cubicles with a size of 2.5 × 2.5 m on the same floor of a building. These two sites were networked so that the users were physically separated while working together in the MR environment (see Figure 1).

Each participant was supported by an application on a computer. Each application ran on a Windows 7 computer with an Intel Xeon processor, 16-GB RAM, and a GeForce TitanX graphics card. We provided each user with the HTC Vive headset (combining with Ovrvision Pro stereo camera for AR system) and controllers to view the virtual world (and control the self-avatar for VRBody system). The MR environment was created using Unity 5.6.2f and written in C#. The application rendered the gameworld at a minimum frame rate of 90 Hz. Audio extension cables were used, and we also ensured that HMD cables were long enough to not obstruct the participants' movements.

The scene consisted of three elements: a background scene, a planet, and self-avatars. The background scene was a model of the cubicle.

3.1 | Planet

The planet's appearance is of a textured sphere, procedurally generated using Unity shader code. The structure of the planet landmass is determined by a set of nodes (points on the surface with associated radii, representing continental landmasses) and links (terrain "bridges" between nodes). This graph-like structure is used to create a distance field cube-map texture representing the shortest distance to the nearest node or link. A few noise functions based upon simplex noise,[16] fractal Brownian motion,[17] and ridged multifractals[18] are used to perturb the distance field and simulate more realistic terrain boundaries; the terrain is then colorized according to the perturbed distances, and some basic lighting effects are added to create a more pleasing visual appearance. The terrain generation is performant enough that discrete edits to the terrain can be smoothly interpolated and animated in real time on consumer-grade desktop computers. For instance, a newly created node will appear to "grow" outward from the center until it reaches the appropriate radius.

3.2 | Avatars

Some participants had a self-avatar. We provided both male and female avatars in generic clothing taken from the Rocketbox Complete Characters HD set. We used each participant's height information to scale the height of the avatar. The participant held the two Vive controllers and wore the Vive HMD with tracking. This gave three points of tracking to animate the self-avatar. We linked these tracking points to the avatar's hands and head, respectively. We then used the VR IK solver from the Final IK plug-in to map the participant's movements in real-world space to the self-avatar's movements.

3.3 | Interaction techniques

There are four main editing operations used by participants to edit the planet. Participants in AR and VR settings use the Vive controller to complete the operations, and desktop participants use the mouse and keyboard. When in VR or AR mode, the virtual Vive controller appears almost identical to the real controller, except that a laser-like beam is emitted from the front of the controller to show the user which objects are being pointed at by the controller, and a color picker dial is superimposed over the Vive controller's touchpad. The operations are as follows.

3.3.1 | Change color

In the AR and VR settings, a vertical board was placed in the background scene close to both users; the board contains a series of colored rectangles, labeled with the planet's terrain types. To change the color of the planet's terrain, the user aims the Vive controller at the colored rectangle and manipulates the touchpad; the color of the rectangle and that of all the corresponding terrain on the planet are changed to the color that matches the color picker overlay on the virtual Vive controller.

Desktop users have an inset with the same board as that shown to the AR and VR users. Holding down the left mouse button with the pointer over one of the colored rectangles turns the colored rectangle into a color picker overlay. Moving the mouse pointer over a color on the picker changes the rectangle's color and that of the corresponding terrain to the corresponding color, and releasing the mouse button removes the picker, with the terrain color changed to the appropriate color.

3.3.2 | Create node

Nodes are terrain points where landmasses are centered; these are signified by a yellow node marker. There is a terrain radius associated with these nodes. Temporary nodes, which are created in an ongoing edit operation, are signified by a cyan marker until the operation is either completed (in which case it turns yellow) or aborted (in which case the marker disappears).

To create a node in the AR and VR settings, the user aims the controller beam at the planet and holds down the trigger; the radius of the terrain expands, with a real-time animation, until the user releases the trigger.

In the desktop setting, the operation takes a similar form. Holding down the left mouse button with the mouse button over the planet creates a node with terrain, which expands until the user releases the mouse button.

3.3.3 | Create link

Links are strips of terrain along the geodesic lines between two nodes, at least one of which is newly created. To create a link, the user first creates a node by either using the Vive controller trigger or the left mouse button as above. Then, while holding down the button or trigger, the user drags the controller pointer or mouse pointer to another point on the planet and releases the trigger or mouse button. If the release point is not an already existing node, two nodes will be created, one at the position initially pointed at by the user and the other at the position when the user released the trigger or mouse button, and there will be line of terrain between them. Both nodes would have the same terrain radius, which is determined by the length of time that the trigger or mouse button was held down.

If the user drags the pointer over another node while creating a link, then a geodesic link is created between the newly created node and the node dragged over. Only the newly created node's radius will be determined by the length of time that the trigger or mouse pointer is held down; the node terrain radius of the node dragged over remains constant. The width of the geodesic terrain line is linearly interpolated so that it matches the node terrain radii at either end, and there are no sharp edges or discontinuities in the resulting landmass.

3.3.4 | Delete node

Nodes can be deleted in the AR/VR settings by aiming the controller at an already created node marker and by pulling the trigger. Desktop users delete nodes by left-clicking on a node marker. In both cases, there is an animation showing the terrain radius decreasing and any geodesic terrain links receding until the terrain vanishes, and the node marker is removed.

3.4 | Networking

To ensure that all participants were receiving the same state for the virtual environment, we implemented a client–server system using Unity's built-in multiplayer networking system. We first tracked each participant's physical movement and behavior, obtaining 3D coordinate frames for all the tracked objects to animate the self-avatar at the local of each client. Then, these 3D coordinates were submitted to the server and propagated to all the remote clients. At the remote client, the corresponding avatar would be animated based on these 3D coordinate frames. Aural communication was supported using Skype. We identify spatialized 3D audio as an area of future work.

4 | EXPERIMENT

The goal of the study was to investigate leadership and collaboration for several MR settings. We manipulated the levels of immersion to examine the users' performance.

4.1 | Method

4.1.1 | Participants

We recruited 16 participants from Disney Research and made them work in pairs to complete the "Build Your Own Earth" task in the four conditions. The average age of the participants was 25.94 years, ranging between 21 and 33 years old; 50% were men. All participants reported some familiarity with AR or VR. They were naïve to the purposes of the study.

4.1.2 | Material

The goal of our "Build Your Own Earth" task was to ask participants to work collaboratively and visualize an ancient Earth, that is, at times, a giant hot molten ball of rock, and at other times, a frozen planet completely covered in snow and ice. A participant received two images, which illustrated about one of the stories (coal forests, desert earth, ice age, and snowball earth) of ancient Earth. Our task required participants to make an agreement on how would they like to paint the Earth and achieve the task goal. This trial was carried out by each group four times for four different MR settings.

This task was chosen because it demonstrates our MR settings supporting multiple users to visualize and edit a planet in real time. Moreover, it requires collaboration between the users because it is difficult for one participant to remember various characteristics of the planet. The task can be divided by each participant creating a different part of the planet such as the one working on the continent and the other working on color.

4.1.3 | Design

A repeated measures design was used. There were two independent variables: sites (site A or site B) and settings (AR-to-AR, AR-to-VRBody, AR-to-VR, or AR-to-desktop; see Figure 2). Each group of two participants took part in all four conditions. Note that non-AR participants did not know the other participants' system and vice versa. To minimize any practice or carry-over effects, the order of the settings was counterbalanced using a Latin square.

4.1.4 | Procedure

Before beginning the experiment, participants at both sites were asked to fill out a brief demographic survey and a consent form. The experimenters in both sites gave the participants an overview of the "Build Your Own Earth" task that the participants would engage in. The experimenters calibrated self-avatars of matched size for them for some VR and AR

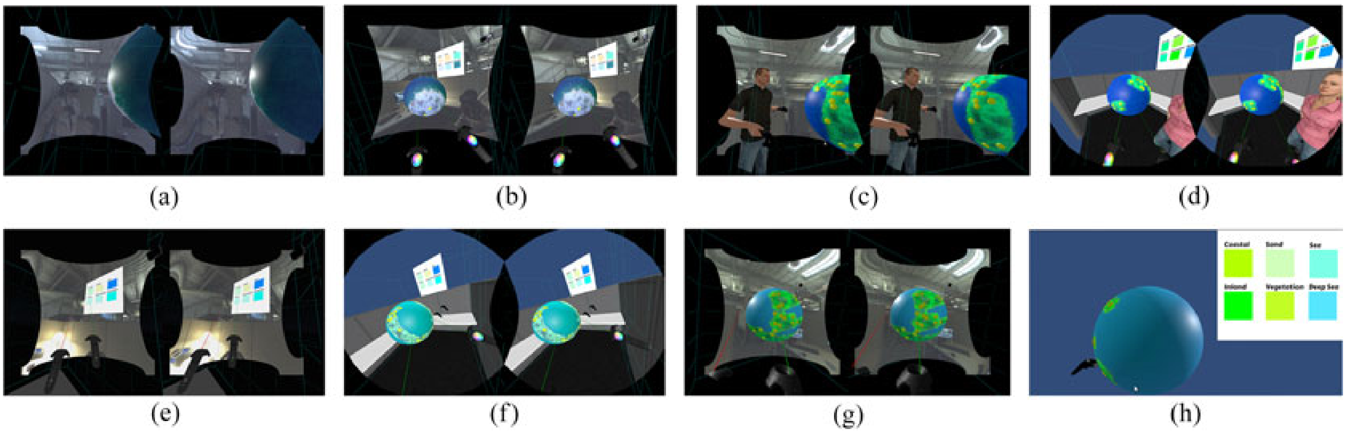


FIGURE 2 Screenshots for different conditions. Each pair of screenshots was simultaneously captured from the first-person view of each participant within the dyad. (a) Augmented reality (AR)-to-AR @ site A. (b) AR-to-AR @ site B. (c) AR-to-VRBody @ site A. (d) AR-to-VRBody @ site B. (e) AR-to-virtual reality (VR) @ site A. (f) AR-to-VR @ site B. (g) AR-to-Desktop @ site A. (h) AR-to-Desktop @ site B

TABLE 2 Post-questionnaire (seven-point Likert scale)

No.	Questionnaire Item
Q1	How would you evaluate your and your partner's level of activity in solving the task? Please rate YOUR level of activity.
Q2	How would you evaluate your and your partner's level of activity in solving the task? Please rate YOUR PARTNER's level of activity.
Q3	To what extent did you and your partner contribute to editing the terrain? Please rate YOUR level of contribution.
Q4	To what extent did you and your partner contribute to editing the terrain? Please rate YOUR PARTNER's level of contribution.
Q5	Who talked the most, you or your partner? Please rate YOUR the amount of verbal contribution.
Q6	Who talked the most, you or your partner? Please rate YOUR PARTNER's the amount of verbal contribution.
Q7	To what extent did you experience that you and your partner collaborated while editing the terrain?
Q8	During the experience, I felt that the body I saw when looking down toward myself was my own body (even though it did not look like me).
Q9	During the experience, I tried to avoid the virtual planet while performing the task.
Q10	There was a sense of being in the room that has the planet.
Q11	I think the virtual place is somewhere I visited rather than just images I saw.
Q12	There were times during the experience when the real world of the laboratory, in which the experience was really taking place, was forgotten.
Q13	The experience was more like working with other people rather than interacting with a computer.
Q14	There was a sense of being with the other people.

conditions. They then guided the participant on how to create continents and change colors using controllers in the VR and AR conditions or using a mouse in the desktop conditions.

For each trial, participants were asked to complete the “Build Your Own Earth” task. Then, the experimenters at both sites terminated the connection, and participants were taken to a nearby computer, where they completed a questionnaire featuring the questions relating to the experience in private.

Finally, when the participants completed all trials, an experimenter conducted an interview with the participants individually to collect general comments on their experience during the experiment. Participants received chocolates as compensation. The experiment took about 40 min.

4.1.5 | Post-questionnaire

Participants were presented with a post-experimental questionnaire that consisted of 14 randomized statements (see Table 2). The greatest part of the questionnaire is based on a previous work[4] because it has been shown to be a reliable indicator for leadership, collaboration, embodiment, presence, and copresence. Participants responded to a set of statements each with an associated 1–7 Likert scale, where an answer of 1 indicated complete disagreement and 7 indicated complete agreement.

4.1.6 | Data analysis

Because our experiment involved pairs of participants rather than individuals, we were unable to assume independence in measurements from participants in two sites. Therefore, we employed dyadic analysis methods to compare data across experiment conditions while taking the potential interdependencies in data from members of dyads into consideration.

4.2 | Results and discussion

4.2.1 | The 2D interactions

The primary measurement was the participants' interactions with 2D surface area and color change (see leftmost column in Figure 3). The mean number of 2D interactions is similar at both sites for all settings.

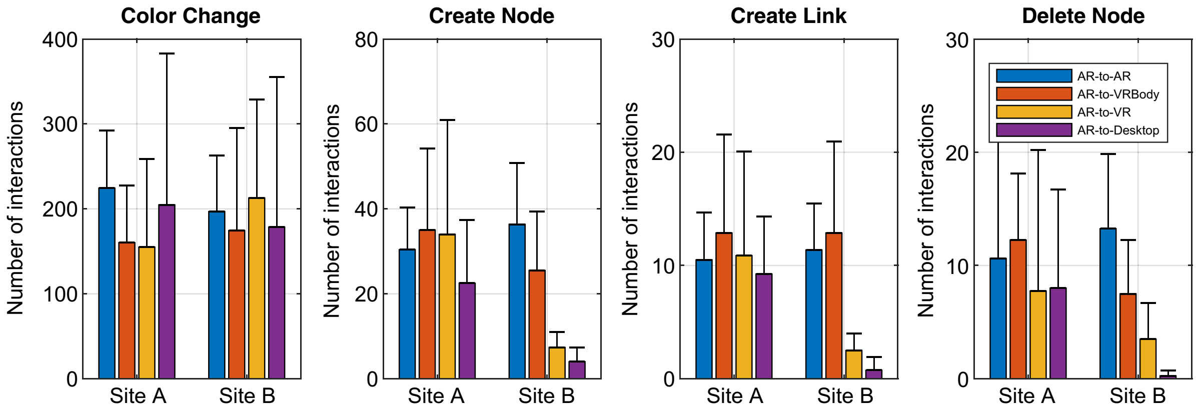


FIGURE 3 Bars showing the number of interactions for each setting and site. AR = augmented reality; VR = virtual reality

A two-way repeated measures analysis of variance was run to determine the effect of different settings at two sites on the number of 2D interactions. There were no outliers, and the data were normally distributed for each conditions as assessed by boxplot and Shapiro–Wilk test ($p > 0.05$), respectively. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the main effect of conditions, $\chi^2(5) = 4.695$, $p = 0.458$, but not for the two-way interaction, $\chi^2(5) = 12.192$, $p = 0.034$. Results revealed that there were no statistically significant differences for the two-way interaction: $F(3, 24) = 3.356$, $p = 0.036$; the main effect of settings: $F(1.775, 14.2) = 0.495$, $p = 0.689$; and the main effect of sites: $F(1, 8) = 0.007$, $p = 0.936$.

4.2.2 | The 3D interactions

We then looked at the participants' interactions with 3D surface area (see three right-hand columns in Figure 3). Results reveal that the AR-to-desktop and AR-to-VR settings clearly make a difference to the "equality" of the contribution at two sites, including the number of interactions for create node, create link, and delete node, whereas the AR-to-AR and AR-to-VRBody settings allow equal participation.

We define the overall 3D interactions as the union of the number of interactions for create node, create link, and delete node. A two-way repeated measures analysis of variance was run to determine the effect of different settings at two sites on the number of 3D interactions. There were no outliers, and the data were normally distributed for each conditions as assessed by boxplot and Shapiro–Wilk test ($p > 0.05$), respectively. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $\chi^2(5) = 1.738, p = 0.885$. There was a statistically significant two-way interaction between sites and settings: $F(3, 24) = 3.356, p = 0.036$. Therefore, simple main effects were run. The number of 3D interaction between two sites was not statistically significantly different in the condition AR-to-AR $t(8) = -1.101, p = 0.303$, and the condition AR-to-VRBody $t(8) = 1.151, p = 0.283$. However, there was a statistically significant mean difference in the condition AR-to-VR $t(8) = 2.239, p = 0.044$, and condition AR-to-desktop $t(8) = 3.594, p = 0.007$.

4.2.3 | Post-questionnaire

Leadership

Three pairs of questions were asked to allow the participants to evaluate their own and their partners' contribution to the task. The Q1 and Q2 are concerned about the contribution to the task in general, the Q3 and Q4 the contribution in editing the terrain, and the Q5 and Q6 the amount of verbal communication.

We first looked at the estimation of contribution regarding themselves and their partners for participants at site B (see Figure 4, blue box in Q1 and Q3, and red box in Q2 and Q4). We can see a clear downward trend from AR-to-AR, AR-to-VRBody, and AR-to-VR, to AR-to-desktop. A Friedman test was run to determine if there were differences in these four conditions. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. Adjusted p values are presented, and only significant results are shown. At site B, only the participant in the AR-to-desktop condition was evaluated, with their contribution being statistically significant less than that of the participant in the AR-to-AR condition ($p = 0.028$) and, in this respect, both to their contribution in solving the task and to editing the terrain ($p = 0.016$). The verbal contribution, however, was regarded as equal in all cases. These results were not surprising inasmuch

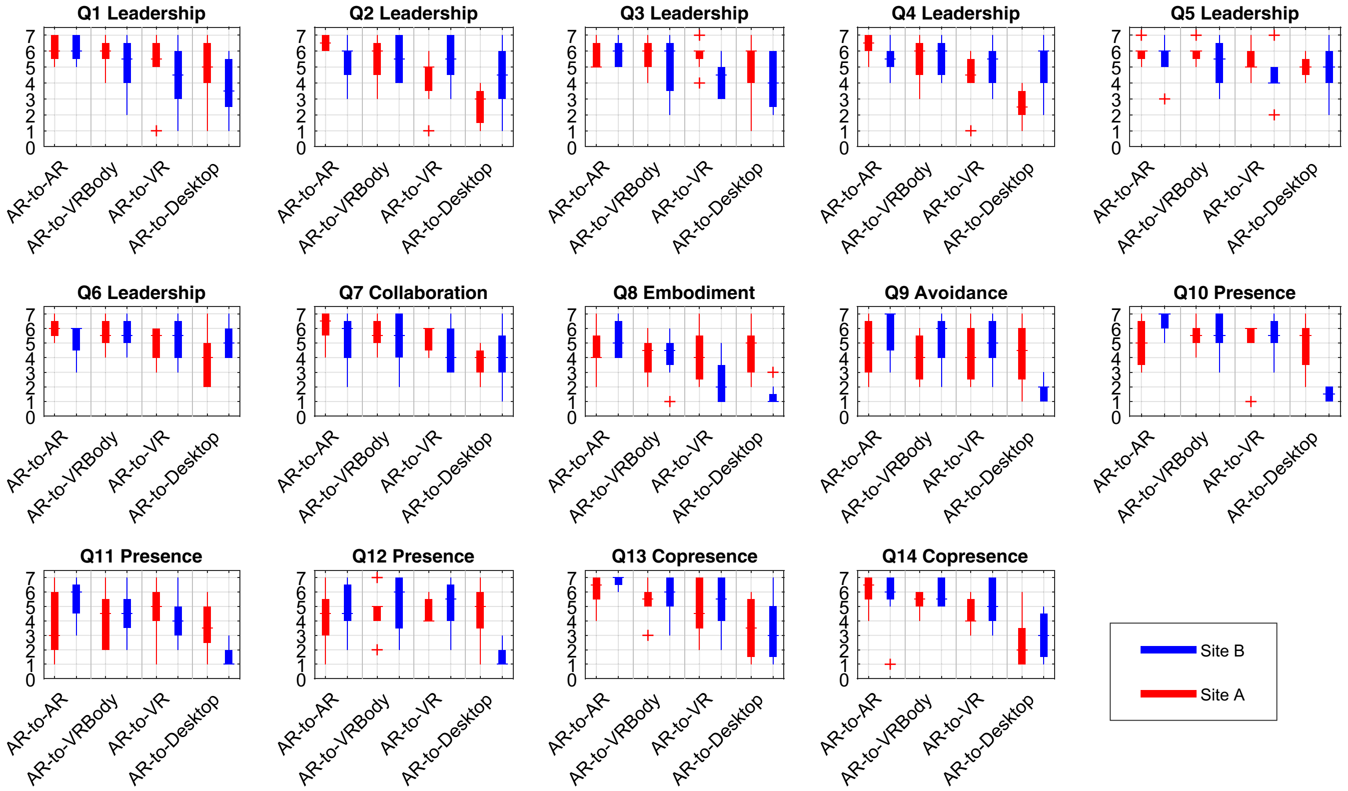


FIGURE 4 Boxplots for questionnaire items associated with Table 2. Medians, interquartile ranges, and full ranges are shown. AR = augmented reality; VR = virtual reality

as we would not expect there to be any difference in the verbal contribution, but we would expect differences for the spatial part of the task.

Figure 4 also showed that participants at site A were evaluated by both partners (red box in Q1 and Q3, and blue box in Q2 and Q4) as being more active in the task generally and contributing more to editing the terrain. This point can be spelled out in more detail for emphasis: Both partners agreed about the difference in their contributions, and there was agreement that this difference applied in terms of contribution to overall contribution, editing the terrain, and verbal communication.

Collaboration

We also asked the participants to evaluate collaboration (Q7). A Friedman test showed that there was a significant difference ($\chi^2(3) = 15, p = 0.002$) among multiple conditions at site A, but no such difference was found at site B. Post hoc analysis revealed that participants at site B from the AR-to-desktop condition reported a lower degree of collaboration than the AR-to-VRBody condition ($p = 0.04$) and the AR-to-AR condition ($p = 0.003$).

From the observations of all trials, it appears that some groups maintained a conversation while collaborating, constantly updating each other on the choices of what color might be and strategies for editing the terrain. Some groups did not feel a need to constantly update the partner verbally on progress as a quick glance was sufficient for sharing the partner's work. One participant in the AR-to-VRBody condition commented:

> *“We can see each other, we don't necessarily have to communicate verbally all the time.”*

Participants in the AR-to-VR setting gave detailed instructions. They often asked for “confirmation” to ensure the other partner could clearly understand while pointing, for example,

> *“Can you see my controller at least?”*

In contrast to the AR-to-VRBody setting and the AR-to-AR setting, deictic references such as "here" or "there" were more frequently observed.

We also looked at verbal communication during while collaborating, for example, in AR-to-Desktop condition,

> *"You are ruining my drawing!" "I am sorry. I am using a desktop. I cannot see your drawing. I am going to rotate the Earth and make it facing us again."*

Thus, this indicated that the desktop system introduces a possibility of interference and confusion, where one participant's actions potentially disturb the productivity of others.

If we look at leadership and collaboration together, for participants at site A only, we can see that, in the AR-to-desktop and AR-to-VR settings, where participants assessed their contributions unequally, they also reported a lower degree of collaboration. In the AR-to-AR and AR-to-VRBody settings, on the other hand, they assessed their respective contributions equally and reported more collaboration.

Embodiment

For Q8, we find a rank order: For participants at site B, the AR-to-AR has the highest reported embodiment, next comes the AR-to-VRBody, then AR-to-VR, and finally AR-to-desktop. A Friedman test showed that there was a significant difference ($\chi^2(3) = 20.186$, $p < 0.001$) among multiple conditions at site B. Post hoc analysis revealed that participants from the AR-to-AR and AR-to-VRBody reported a higher degree of embodiment than the AR-to-desktop condition, ($p = 0.04$) and ($p = 0.001$), respectively. Moreover, the difference between the AR-to-AR and the AR-to-VR was significant, ($p = 0.022$).

Presence

In relation to presence at site B, our findings are as expected, namely, desktop participants in the AR-to-desktop setting report a lower degree of presence. A Friedman test showed that there was a significant difference among multiple conditions for Q10, $\chi^2(3) = 21.286, p < 0.001$, Q11, $\chi^2(3) = 19.875, p < 0.001$, and Q12, $\chi^2(3) = 16.757, p = 0.001$, respectively. Post hoc analysis revealed that participants from the AR-to-desktop condition reported a lower degree of presence than the AR-to-AR condition for Q10 ($p < 0.001$), Q11 ($p = 0.001$), and Q12 ($p = 0.004$), and the AR-to-VRBody condition for Q10 ($p = 0.03$), Q11 ($p = 0.04$), and Q12 ($p = 0.016$).

Copresence

By copresence, we mean the subjective sense of being together or being colocated with another person in a computer-generated environment.

At site A, the Friedman test showed that there was a significant difference among multiple conditions for Q13, $\chi^2(3) = 12.785$, $p = .005$, and Q14, $\chi^2(3) = 19.708$, $p < 0.001$, respectively. Post hoc analysis revealed AR participants from the AR-to-desktop condition reported a lower degree of copresence than the AR-to-AR condition for Q13 ($p = 0.008$) and Q14 ($p = 0.001$). Moreover, the difference between the AR-to-desktop and the AR-to-VRBody condition was significant for Q14 ($p = 0.03$).

Interestingly, at site B, the Friedman test also showed that there was a significant difference among multiple conditions for Q13, $\chi^2(3) = 14.304, p = 0.003$, and Q14, $\chi^2(3) = 12.422, p = 0.006$, respectively. Post hoc analysis revealed that participants from the AR-to-desktop condition reported a lower degree of copresence than the AR-to-AR condition for Q13 ($p = 0.012$) and Q14 ($p = 0.03$).

We can note, from the post-interview, that the participants at site B sometimes misperceived what type of system their partner was working on, that is, participants using the VRBody or VR system tended to think that their partner was also using a VRBody or VR system like their own, and participants using desktop thought their partners were also using a desktop system.

5 | CONCLUSION

We have presented several prototype MR systems. Our AR-to-AR setting makes it possible for multiple users to walk around a virtual planet, collaboratively edit various characteristics of the planet, and see what the climate of that planet would be like. The AR-to-VRBody setting supports multiple users to visualize the virtual planet remotely, communicating through voice and gestures. As consumer HMDs and stereo cameras are now becoming very cheap, the low cost and ease of use make the systems an attractive means of viewing almost any spatial data.

We empirically evaluated how well our system can support collaboration. (a) Some of our findings are expected in the light of the previous studies,[4,5] which found a similar asymmetry between the more immersed and the less immersed partner in the VR environment. In our case, there was a stronger leadership effect in the AR-to-Desktop setting than the AR-to-VR setting and no leadership effect in AR-to-AR setting. The desktop setup has the lowest leadership score, which might be because of less immersive or less interaction ability. It extends the previous studies on applying them to MR environment. (b) What is surprising here is that there was no significant difference in terms of leadership in the AR-to-VRBody condition. This might be explained by a study on the impact of avatar realism on illusion of virtual body ownership, which indicated that the feeling of power was higher with nonrealistic strong-looking avatars.[19] (c) Another important finding is that the leadership effect only emerged for 3D interaction but not in the 2D interaction.

Altogether, our study has important implications for the design of collaborative MR systems, including both the technical features and the way to collaborate.

ORCID

Ye Pan http://orcid.org/0000-0003-0355-989X
Kenny Mitchell https://orcid.org/0000-0003-2420-7447

REFERENCES

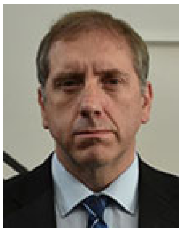
1. Ye P, Steed A. The impact of self-avatars on trust and collaboration in shared virtual environments. PloS ONE. 2017;12(12):e0189078.
2. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst. 1994;77(12):1321–1329.
3. Benko H, Ishak EW, Feiner S. Collaborative mixed reality visualization of an archaeological excavation. Paper presented at: ISMAR '04 Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality; 2004; Arlington, VA. Washington, DC:IEEE Computer Society; 2004. p. 132–140.
4. Steed A, Slater M, Sadagic A, Bullock A, Tromp J. Leadership and collaboration in shared virtual environments. Paper presented at: Proceedings IEEE Virtual Reality; 1999; Houston, TX. New York, NY: IEEE; 1999. p. 112–115.
5. Axelsson A-S, Abelin Å, Heldal I, Schroeder R, Wideström J. Cubes in the cube: A comparison of a puzzle-solving task in a virtual and a real environment. Cyberpsychol Behav. 2001;4(2):279–286.
6. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. Front Robotics AI. 2016;3:74.
7. Zünd F, Lancelle M, Ryffel M, Sumner RW, Mitchell K, Gross M. Influence of animated reality mixing techniques on user experience. Paper presented at: MIG '14 Proceedings of the Seventh International Conference on Motion in Games; 2014; Playa Vista, CA. New York, NY:ACM; 2014. p. 125–132.
8. Roth D, Lugrin J-L, Galakhov D, et al. Avatar realism and social interaction quality in virtual reality. Paper presented at: 2016 IEEE Virtual Reality (VR); 2016; Greenville, SC. New York, NY:IEEE; 2016. p. 277–278.
9. Roth D, Lugrin J-L, Latoschik ME, Huber S. Alpha IVBO - construction of a scale to measure the illusion of virtual body ownership. Paper presented at: CHI EA '17. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems; 2017; Denver, CO. New York, NY:ACM; 2017. p. 2875–2883.
10. Bodenheimer B, Fu Q. The effect of avatar model in stepping off a ledge in an immersive virtual environment. Paper presented at: SAP '15. Proceedings of the ACM SIGGRAPH Symposium on Applied Perception; 2015; Tübingen, Germany. New York, NY:ACM; 2015. p. 115–118.
11. Young MK, Rieser JJ, Bodenheimer B. Dyadic interactions with avatars in immersive virtual environments: high fiving. Paper presented at: SAP '15. Proceedings of the ACM SIGGRAPH Symposium on Applied Perception; 2015; Tübingen, Germany. New York, NY:ACM; 2015. p. 119–126.
12. Hodgins J, Jörg S, O'Sullivan C, Park SI, Mahler M. The saliency of anomalies in animated human characters. ACM TAP. 2010;7(4):22.
13. Kätsyri J, Förger K, Mäkäräinen M, Takala T. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. Front Psychol. 2015:6.
14. Piwek L, McKay LS, Pollick FE. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. Cognition. 2014;130(3):271–277.
15. McDonnell R, Breidt M, Bülthoff HH. Render me real?: investigating the effect of render style on the perception of animated virtual humans. ACM TOG. 2012;31(4):91.
16. Olano M. Modified noise for evaluation on graphics hardware. Paper presented at: HWWS '05. Proceedings of the ACM SIGGRAPH/EUROGRAPHICS Conference on Graphics Hardware; 2005; Los Angeles, CA. New York, NY: ACM; 2005. p. 105–110.
17. Fournier A, Fussell D, Carpenter L. Computer rendering of stochastic models. In: Tutorial: Computer graphics; image synthesis. New York, NY: Computer Science Press, Inc., 1988; p. 114–127.

18. Ebert DS, Musgrave FK, Peachey D, Perlin K, Worley S. Texturing and modelling: A procedural approach. 3rd ed. San Francisco, CA: Morgan Kaufmann; 2003.

19. Lugrin J-L, Landeck M, Latoschik ME. Avatar embodiment realism and virtual fitness training. Paper presented at: 2015 IEEE Virtual Reality (VR); 2015; Arles, France. New York, NY:IEEE; 2015. p. 225–226.

Ye Pan is a Postdoctoral Associate at Disney Research Los Angeles. She received the PhD in Virtual Environment and Computer Graphics from the University College London (UCL) in 2015. She previously graduated from UCL in 2011 with an MRes in Virtual Environments, Imaging and Visualisation, and from Purdue/UESTC in 2010 with a BSc in Communication and Information Engineering. Her research interests include: avatars, computer human interaction, virtual environments, CSCW, computer graphics, social interaction.

David Sinclair is a software engineer with a background in Mathematics, currently working at Disney Research as a Lab Coordinator. He has interests in scripting and automation, as well as Virtual and Augmented Reality.



Professor Kenny Mitchell is head of Interactive Graphics at Disney Research and chair of Game Technology at Edinburgh Napier University. He has published in real-time topics of rendering, vision and visualisation, and he is an organiser of many international conferences, including papers co-chair CVMP 2016 and papers co-chair of ACM I3D 2018.

How to cite this article: Pan Y, Sinclair D, Mitchell K. Empowerment and embodiment for collaborative mixed reality systems. *Comput Anim Virtual Worlds*. 2018;29:e1838. https://doi.org/10.1002/cav.1838