

Figure 1: Three Configurations for MR Communication.

Inspiration for configurations: 1) side-by-side whiteboard brainstorming, 2) daily face-to-face conversations and 3) drafting boards

KEYWORDS

Mixed reality; communication; collaboration;

Exploring Configuration of Mixed Reality Spaces for Communication

Zhenyi He

New York University
New York, NY, USA
zh719@nyu.edu

Karl Toby Rosenberg

New York University
New York, NY, USA
ktr254@nyu.edu

Ken Perlin

New York University
New York, USA
ken.perlin@gmail.com

ABSTRACT

Mixed Reality (MR) enables users to explore scenarios not realizable in the physical world. This allows users to communicate with the help of digital content. We investigate how different configurations of participants and content affect communication in a shared immersive environment. We designed and implemented side-by-side, mirrored face-to-face and eyes-free configurations in our multi-user MR environment and conducted a preliminary user study for our mirrored face-to-face configuration, evaluating with respect to one-to-one interaction, smooth focus shifts and eye contact within a 3D presentation using the interactive Chalktalk system. We provide experimental results and interview responses.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5971-9/19/05.

<https://doi.org/10.1145/3290607.3312761>

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; *Collaborative interaction*.

Table 1: Terminology Table

| Term | Definition |
|-------------------------|--|
| Content Creation Server | The content creation server is an external server which takes raw drawing point data as the input, recognizes the drawing, and creates interactive objects. |
| Content Board | The content board is the transparent area on which content is displayed. The user's input will be projected to the content board and sent to the content creation server for processing. Afterwards, the result will be displayed in 3D in the transparent area. |
| Configuration | The configuration refers to the placement and orientation of users and the content board with respect to each other. |

INTRODUCTION

Virtual Reality and Mixed Reality (VR, MR) are being explored increasingly, spurred by the availability of high quality consumer headsets in recent years. VR and MR enable rich design spaces in HCI by providing 3D input and immersive experiences. Decades ago, the "Office of the future" was proposed to allow remotely located people to feel as though they were together in a shared office space [10], via a hybrid of modalities including telepresence, large panoramic displays and shared manipulation of 3D objects. The core idea was that VR/MR had the potential to enhance communication among groups of people. Since then, significant progress has been made in exploring techniques for communication [4, 8]. However, less studied is the configuration (see our full definition in table 1) of people and shared manipulable objects in the environment, which may lead to different communication experiences.

In daily life while speaking to others, we commonly use gestures or visual aids to help present ideas, either subconsciously or purposefully. Visual aids can be drawn on paper, a whiteboard, or a screen via video chat. A key task for collaborators is the shifting of focus between spoken words, gestures and visual aids such as notes and drawings. Smooth transitions in conversations have been found to be important for collaboration [1]. Prior work addressed this with alternative configurations of spaces for communication. Tan et al. built a face-to-face presentation system for remote audiences [11]. ClearBoard [4] created a shared workspace in which two users collaborate remotely without losing all the advantages of in-person face-to-face interactions. One such advantage relates to learning. When a teacher looks away from the audience, Lanir et al. [6] observed that audiences in a classroom would not focus on the presenter, which might "create a learning environment in which there is no interpersonal engagement between the presenter and the audience, thus reducing learning outcomes." We find that personal engagement, such as one-to-one interaction and eye contact [3], in addition to focus shift [1], is therefore relevant to evaluating communication experiences.

Still, it is unclear how the configuration of users and content in an MR space affects communication for co-located and distant people. We have implemented a multi-user MR workstation to explore how configuration impacts communication by providing three different configurations: 1) side-by-side, 2) mirrored face-to-face and 3) eyes-free. We conducted a preliminary user-study of the mirrored face-to-face configuration. Afterwards, we proceeded to user interviews in which participants spoke about the face-to-face experience, one-to-one interaction, focus shift and eye contact.

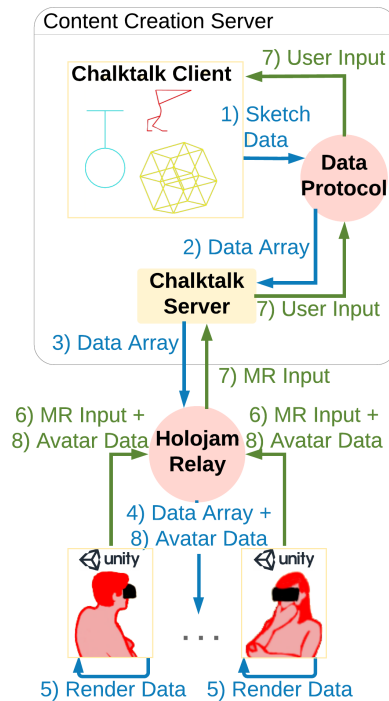


Figure 2: Data Pipeline

Chalktalk's content ("sketches") is composed of graphical elements such as point data. 1) Sketch data are serialized on the Chalktalk client side every frame and 2) sent as a data array to the Chalktalk server, 3) to the Holojam relay, and then 4) to all Unity clients, where 5) the data are rendered on content board(s). 6) MR user input is sent back through the pipeline to the Chalktalk client and 7) translated into HTML canvas mouse events. 8) Avatar synchronization data are also sent between clients using the Holojam relay.

RELATED WORK

Much work has been done in collaborative applications in VR/MR. T(ether) is a spatially-aware display system for co-located collaborative manipulation and animation of objects [5]. Trackable markers on pads and digital gloves allow participants to use gestures to manipulate objects in space. Virtual Replicas for Remote Assistance is a remote collaboration system, allowing a remote expert to guide local users to assemble machine parts by using virtual replicas [7]. SpaceTime is a scene editing tool supporting multi-user collaboration in VR, either co-located or remote [12]. To support conflict resolution when two users wish to work on the same object simultaneously, SpaceTime creates per-user branches of the object, which can later be merged. Rather than enhancing experiences for specific tasks, our system aims to enhance the experience of general communication, which can benefit various kinds of more specific collaborative work.

Some previous work contributed to communication in VR/MR too. ClearBoard allows a pair of users to shift easily between interpersonal space and a shared workspace [4]. The key metaphor of ClearBoard is "talking through and drawing on a big transparent glass board." No gaze or eye contact information is lost while working on the content. ShareVR enables communication between an HMD user and a non-HMD user [2]. By using floor projection and mobile displays to visualize the virtual world, the non-HMD user is able to interact with the HMD user and become part of the VR experience. The work discusses how people with different devices communicate with each other. MMSpace allows face-to-face social interactions and telepresence in the context of small group remote conferences [8]. It uses custom-built mechanical displays on which images of remote participants are projected, and which move in response to users' movements. Pairs of participants can maintain eye contact with each other and remain aware of each other's focus. Instead of designing a configuration to fit one specific communication use case, we provide three different configurations for general-purpose communication in VR/MR.

DETAILED DESIGN AND IMPLEMENTATION

Three Configuration Designs. We implemented three *configurations* for communication in our system: 1) side-by-side, 2) mirrored face-to-face and 3) eyes-free (see figure 1). For 1), our implementation has users facing a *content board* (see full definition in table 1) from the same side.

For 2), the users are face-to-face in the virtual environment with the content board placed between them, so that each sees the other on the opposite side of the content board, left-right reversed as if reflected in a mirror. The challenge is ensuring all the content is consistent for everyone on each side of the board. Mirror reversal allows, for example, text to be readable and asymmetric objects to appear correct for each participant. ClearBoard [4] implemented "mirror reversal" via video capture and projection techniques to solve a similar problem for 2D displays. Inspired by that, we implemented a

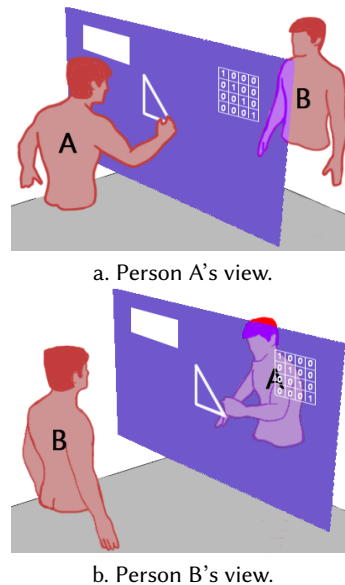


Figure 3: Mirrored Face-to-face Configuration Implementation. Person A is drawing a triangle with right hand from his perspective, whereas person B sees person A drawing with the left hand. The content of the drawing appears the same to both.

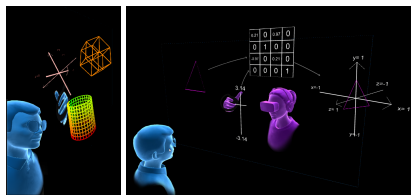


Figure 4: Left: The content board is 3D. Right: Screenshot from experiment with matrix presentation.

3D immersive mirror reversal for our MR configuration. We place the users physically on the same side of the content board and mirror all other users (from one user's perspective) to the other side. This way, information such as gaze and gesture direction is preserved, so participants can know where each other is looking and pointing (see figure 3).

For 3), instead of writing or drawing in mid-air, users can create content in MR while resting their arms atop a horizontal surface in the real world (e.g. table), thereby avoiding the potential fatigue of drawing in mid-air for long periods of time. For this configuration, we now have two boards in the MR world: a horizontal drafting board used for writing and drawing, and a vertical board that displays all information to be seen by all people in a group. All content is duplicated onto the vertical board. A cursor on the vertical board corresponds to the position of the user's hand on the horizontal board. This way, the person using the horizontal board need not look downward, remaining free to look at the vertical content board and other people. This allows users to pay more attention to the environment and each other in a shared experience. The configuration can extend 1) and 2) since we can choose where to place the vertical duplicate board.

The MR System. Our system comprises 1) a content creation server, 2) an internal network framework and 3) VR/MR clients (see figure 2 for a detailed description). 1) To enable interactive content during communication, we chose Chalktalk as our content creation server (see figure 4). Chalktalk is a web browser-based 3D presentation and communication tool in which the user draws interactive "sketches" for presentation. We designed a generic data serialization protocol to connect and decouple the content creation server and the VR/MR clients, so alternative content could easily be plugged in. 2) To ensure communication between different devices, we used Holojam [9], a shared space network framework designed at our lab. It synchronizes data across devices and supports custom data formats. 3) We implemented the VR/MR clients with Unity to support multiple VR/MR devices.

PRELIMINARY EXPERIMENT

We plan to conduct experiments for each configuration design and to compare the configurations. So far, we have completed a preliminary user study on our mirrored face-to-face configuration, conducted with 8 participants ($F=4$) between the ages of 22 and 26 ($M=23.71$, $SD=1.50$), recruited via email and word-of-mouth. Participants were required to have taken a linear algebra class and to have prior experience with VR/MR. First, we introduced Chalktalk to participants who were unfamiliar with it to reduce the novelty effect. Then we gave a presentation on matrix transformations (a concept in the computer graphics curriculum, see figure 4 and sidebar 1) using our system in the mirrored face-to-face configuration. For each session, we had 1 presenter and 2 study participants in the audience, all physically remote. The presentation was repeated 4 times for 8 participants in all. We ran the experience with Oculus Rift headsets. Upon completion of all sessions, participants answered a questionnaire that gauged their opinions on one-to-one interaction [6], eye contact [3] and focus

To create a realistic presentation, we invited a computer graphics professor to present a lesson on matrices. He permitted us to use his presentation as the basis for our experiment. We chose the topic of matrix transformations so participants could see and interact with 3D moving content. For content within the 3D immersive environment, we chose visualization of the way a matrix translates and rotates geometry. During the lecture the presenter demonstrated matrix transformations, including translation and rotation, by using 3D interactive visualizations from the content server. Then she showed that matrix operations are non-commutative. The following lists the steps in the presentation:

- (1) The presenter creates a 4x4 rotation matrix object and links geometry to it.
- (2) Modifying the matrix values rotates the geometry in 3D space. Students are invited to walk around the MR environment to observe from any angle.
- (3) The presenter creates a second 4x4 matrix for translation, which she composes with the rotation matrix.
- (4) To demonstrate that matrices are non-commutative, she shows that by changing the order in which the translation and rotation matrices are applied, the geometry's position and rotation change visibly with the same matrix values.

Sidebar 1: Matrix Lecture for Experiment

shift [1] during the experience. Participants then joined a semi-structured exit interview. All factors were evaluated using the 7-point Likert scale. The study was recorded with users' permission.

RESULTS AND FINDINGS

The three main discussion topics during the post-test interviews were "the feeling of one-to-one," "ability to shift focus smoothly" and "eye contact" (see questions in Table 2 and results in Figure 5).

One-to-one Experience. Most participants felt the face-to-face experience created a feeling of one-to-one interaction with the presenter (6/8 agree more than moderately). P1(F): *"It felt like it was a one-on-one lesson even though there was more than one person there. It felt like the person [the presenter] was right in front of me."* P1(F) contrasted the experience with a college lecture format in which the lecturer stands to the side, which she considered *"more distant."* Similarly, P4(F) thought it *"really felt like a private lesson"* and that *"It was one-to-one-like we were together in this."*

Focus Shift. Participants responded differently to the question (Q4) concerning how often they shifted focus between content and the presenter (2/8 very rarely, 2/8 moderately rarely, 3/8 slightly frequently and 1/8 moderately frequently). Those who shifted focus least often (P2,M and P4,F) thought the presenter was always in the field of view and felt they did not need to shift focus while looking at the content. Most reported that they could follow both content(Q3) (5/8 strongly agree) and the presenter(Q2) (6/8 more than moderately agree) well. This suggested that the focus shift between content and the presenter was smooth to some extent in the mirrored face-to-face configuration.

Eye Contact. Some participants did not always want eye contact. P5(M): *"you don't want the presenter to be always looking at you."* Others like P4(F) felt it was necessary: *"I look at the professor the whole time. That's the only way I can pay attention."* Those who preferred less eye contact (6/8 reported eye contact less than slightly rarely) admitted they felt the presenter was looking at them frequently, but chose to look at the content instead of the presenter. Participants who preferred to have eye contact reported having it more often (2/8 more than half the time). This suggests that our mirrored face-to-face configuration supports eye contact well and that the users chose whether to make eye contact or not.

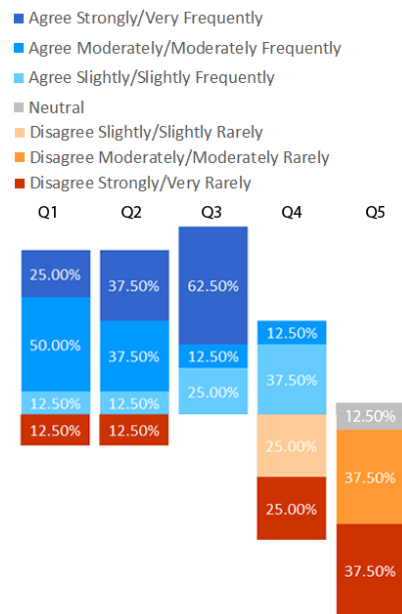
Findings. Participants also thought the face-to-face format helped them concentrate during the presentation. Referring to the front-and-center presence of the presenter's avatar, P3(M) said *"I don't think I'd be able to concentrate if I was just listening to somebody. I'd need someone to actually be there."* P6(F) also found it easier to concentrate: *"It felt one-to-one, so you won't be distracted by other people."* P3(M), P4(F) and P5(M) noted the format also made the experience feel interactive, *"unlike a video."*

CONCLUSIONS AND FUTURE WORK

We have presented our ongoing work, a multi-user MR system for communication. We designed three configurations for MR communication and evaluated the mirrored face-to-face configuration with

Table 2: Questions in Questionnaire

Q1: To what degree do you feel it is a one-on-one lecture?
 Q2: Is it easy to follow the presenter?
 Q3: Is it easy to follow the content?
 Q4: How often did you switch your focus between the presenter and the content?
 Q5: How often did you have eye contact with the presenter?

**Figure 5: Results for the 7-point Likert Scale Questionnaire**

respect to one-to-one interaction, smooth focus shifts and eye contact. The preliminary study suggests that the face-to-face configuration facilitates a feeling of one-to-one collaboration. Participants responded positively to being given attention and feeling as though they were working directly with the lecturer. Positioning the presenter in-front helped some participants concentrate.

In the near future, we plan to conduct multiple user studies in which we will compare side-by-side, mirrored face-to-face and eyes-free configurations. We will investigate which configurations are best suited to communication experiences involving presentations, group discussions and collaborative tasks. Additional questions remain to be explored. For example, how does the feeling of one-to-one interaction change with the configuration of the users and content?

REFERENCES

- [1] William Buxton. 1992. Telepresence: Integrating shared task and person spaces. In *Proceedings of graphics interface*, Vol. 92. Canadian Information Processing Society Toronto, Canada, 123–129.
- [2] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 4021–4033.
- [3] David Insa, Josep Silva, and Salvador Tamarit. 2016. Where You Sit Matters How Classroom Seating Might Affect Marks. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '16)*. ACM, New York, NY, USA, 212–217. <https://doi.org/10.1145/2899415.2899444>
- [4] Hiroshi Ishii, Minoru Kobayashi, and Jonathan Grudin. 1993. Integration of interpersonal space and shared workspace: ClearBoard design and experiments. *ACM Transactions on Information Systems (TOIS)* 11, 4 (1993), 349–375.
- [5] David Lakatos, Matthew Blackshaw, Alex Olwal, Zachary Barryte, Ken Perlin, and Hiroshi Ishii. 2014. T (ether): spatially-aware handhelds, gestures and proprioception for multi-user 3D modeling and animation. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. ACM, 90–93.
- [6] Joel Lanir, Kellogg S. Booth, and Leah Findlater. 2008. Observing Presenters' Use of Visual Aids to Inform the Design of Classroom Presentation Software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 695–704. <https://doi.org/10.1145/1357054.1357165>
- [7] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. 2015. Virtual replicas for remote assistance in virtual and augmented reality. In *Proceedings of the 28th Annual ACM Symposium on UIST*. ACM, 405–415.
- [8] Kazuhiro Otsuka. 2016. MMSpace: Kinetically-augmented telepresence for small group-to-group conversations. In *Virtual Reality (VR), 2016 IEEE*. IEEE, 19–28.
- [9] Ken Perlin. 2016. Future Reality: How emerging technologies will change language itself. *IEEE computer graphics and applications* 36, 3 (2016), 84–89.
- [10] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. 1998. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. ACM, 179–188.
- [11] Kar-Han Tan, Dan Gelb, Ramin Samadani, Ian Robinson, Bruce Culbertson, and John Apostolopoulos. 2010. Gaze Awareness and Interaction Support in Presentations. In *Proceedings of the 18th ACM International Conference on Multimedia (MM '10)*. ACM, New York, NY, USA, 643–646. <https://doi.org/10.1145/1873951.1874041>
- [12] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. 2018. Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality. In *The 31st Annual ACM Symposium on UIST*. ACM, 853–866.