CollaboVR: A Reconfigurable Framework for Creative Collaboration in Virtual Reality

Zhenyi He*
New York University

Ruofei Du†
Google

Ken Perlin‡
New York University

(a) Discussing travel schedules in *integrated layout* with remote participants. (b) Presentation on the topic of hyper dimension in *Mirrored layout*. (c) Sketching a baroque-style pattern in *projective layout* to remote users. (d) Collaborative design session of furnitures and apartment arrangements.

Figure 1: CollaboVR is a reconfigurable VR framework that combines the abilities of animated sketching, collaborative scene editing, and multi-user communication in real-time. We showcase four use cases in custom *layouts*: (a) shows an *integrated* layout of a business meeting, (b) shows a *mirrored* layout of a math class presentation, (c) shows a third-person perspective of the *projective* layout where the user draws at hands and projects his sketches to remote participants on the shared interactive board, and (d) shows two roommates discussing the apartment design. Please refer to the supplementary video for live demos.

ABSTRACT

Writing or sketching on whiteboards is an essential part of collaborative discussions in business meetings, reading groups, design sessions, and interviews. However, prior work in collaborative virtual reality (VR) systems has rarely explored the design space of multi-user layouts and interaction modes with virtual whiteboards. In this paper, we present CollaboVR, a reconfigurable framework for both co-located and geographically dispersed multi-user communication in VR. Our system unleashes users' creativity by sharing freehand drawings, converting 2D sketches into 3D models, and generating procedural animations in real-time. To minimize the computational expense for VR clients, we leverage a cloud architecture in which the computational expensive application (Chalktalk) is hosted directly on the servers, with results being simultaneously streamed to clients. We have explored three custom layouts – integrated, mirrored, and projective – to reduce visual clutter, increase eye contact, or adapt different use cases. To evaluate CollaboVR, we conducted a within-subject user study with 12 participants. Our findings reveal that users appreciate the custom configurations and real-time interactions provided by CollaboVR. We have open sourced CollaboVR at `https://github.com/snowymo/CollaboVR` to facilitate future research and development of natural user interfaces and real-time collaborative systems in virtual and augmented reality.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Collaborative and social computing—Collaborative and social computing systems and tools

1 INTRODUCTION

During this COVID-19 pandemic, remote communication is becoming a crucial component of many people's daily lives. Not only many classes and business meetings are held online, but also international conferences, such as IEEE VR and HTC V^2EC, are being hosted in virtual formats. Despite recent advances in collaborative work in virtual reality (VR), exchanging ideas between users is mostly achieved through direct media such as audio [36] and video [58], or indirect media such as gestures [83] and scene editing [37]. Sketching, one of the most natural and fun ways to express ourselves, has rarely been explored in collaborative VR. Additionally, it is an open question what is the best layout and interaction mode for creative collaboration: An in-air shared canvas between users? A whiteboard in front of users? A notebook or a tabletop to be shared by users? Motivated by these alternate metaphors, we investigate the following research questions: What if we could bring *sketching* to *real-time* collaboration in virtual reality? If we can convert raw sketches into *interactive animations*, will it improve the performance of remote collaboration? Are there best *user arrangements* and *input modes* for different use cases, or is it more a question of personal preferences?

To answer these questions, we have developed CollaboVR, an end-to-end system for both distributed and co-located multi-user communication in virtual reality. Our system employs a cloud architecture in which applications such as Chalktalk [63] (a software system to convert raw sketches to digital animations) are hosted on the server. This architecture allows geographically dispersed clients to talk with each other, sketch on virtual sketching boards, and express ideas with interactive 3D animations with low-end VR headsets. Furthermore, CollaboVR allows real-time switching between different user arrangements and input modes. Whether the user intends to draw on a notebook, sketch in the air, or have a discussion in front of a whiteboard, CollaboVR can instantly and seamlessly switch context to support the user arrangement and input mode of choice. Our main contributions are summarized as follows:

1. Developing CollaboVR, an end-to-end collaboration system using a cloud-based computing architecture to support multi-user sketch, audio communication, and collaboration in 3D.
2. Real-time techniques to share freehand sketches, convert 2D sketches into 3D models, and interact with animations in collaborative virtual reality.
3. Designing custom configurations for real-time user arrangements and input modes for multi-user sketching scenarios inspired by real-world metaphors.

---

*e-mail: zhenyi.he@nyu.edu
†e-mail: me@duruofei.com
‡e-mail: ken.perlin@gmail.com

4. Quantitative and qualitative evaluation of CollaboVR with 12 participants to discuss its advantages, limitations, and potential impacts on collaborative VR systems.
5. Open-sourcing[1] our software to facilitate future development in collaborative VR systems with multi-modal inputs.

2 RELATED WORK

CollaboVR is a framework to assist communication in collaboration. By definition, communication is the act of expressing and understanding among a group. Similarly, *sensemaking* is the understanding of the meaning of a communicative action [60]. Sensemaking is a widely researched concept in information visualization. Dervin [14] describes sensemaking as using ideas, emotions, and memories to bridge a gap in understanding in a group. Learning how collaborative sensemaking is supported through different design considerations is very useful for multi-user communication. In this section, we first introduce collaborative sensemaking approaches. We then summarize how workspace awareness has positive effects on collaboration and how previous studies enhance workspace awareness. Last, we introduce immersive collaboration and communication and assess their advantages and limitations.

2.1 Collaborative Sensemaking

Prior arts have researched sensemaking [52] in HCI and computer-supported collaborative work (CSCW) area [1, 5, 47, 60]. Given that sensemaking involves data analysis [89], different designs of 2D displays and digital tabletop are frequently discussed. Prior work has shared two observations. First, large and shared displays have been shown to benefit sensemaking groups in several contexts. Paul and Morris [61] designed CoSense with a shared display, conducted an ethnographic study, and examined to support collaborative sensemaking. Vogt *et al.* [84] found that the large display facilitated the paired sensemaking process, allowing teams to spatially arrange information and conduct individual work as needed. Moreover, multiple digital tabletops were used for sensemaking tasks [39, 55]. Second, personal displays may lead to decreased collaboration in co-located settings [12, 86]. When designing CollaboVR, we considered the idea of "multiple" displays, displays with "different" angles, as well as adding "personal" displays into the mix, which leads to the design of different input modes and the placement of visual aids.

2.2 Workspace Awareness

Workspace awareness is the collection of up-to-the-minute knowledge a participant has of other participants' interaction with the workspace [27]. It includes the awareness of others' locations, activities, and intentions to the task and to space. Maintaining workspace awareness enables participants to co-work more effectively [28, 29]. Workspace awareness plays a crucial role in simplifying communication, taking turns, and action prediction [29]. Thus, maintaining and enhancing workspace awareness is beneficial to collaboration [65].

One trend is the use of see-through displays for distributed collaboration. The idea started with Tang and Minneman, who designed VideoDraw [78] and VideoWhiteBoard [77]. Both were two-user experiences. On each side, a camera was placed to capture the local user and the drawing. A projector was attached to present the remote user and the drawing. ClearBoard [40] extended the idea and used digital media and monitor. Similarly, KinectArms [21] used a tangible tabletop as the media and rendered the arm of the remote user for mixed presence. Furthermore, Li *et al.* [51] developed FacingBoard with two-sided transparent displays. Analogous to ClearBoard, the entire upper-body was displayed to other participants so gaze awareness was supported. To maintain gaze interaction, FacingBoard reversed the graphics on the display. Consequently, column-sensitive content, such as text and maps then became incorrect. To solve this problem, FacingBoard selectively flipped the column-sensitive content individually and adjusted the content position. However, when people pinpointed a specific sub-area within the content, the gaze and the place being pinpointed were inconsistent for both users. Considering flipping the content, Bork *et al.* [6] showed that the flipped version of Magic Mirror has better usability than the non-flipped version. In our system, we proposed different user arrangements to enhance workspace awareness, from which, there is a similar face-to-face experience. Differently, we manipulate the users' locations to maintain gaze awareness rather than flipping the content. That keeps the content in the original and correct format. Meanwhile, we support collaboration with more than two people.

2.3 Immersive Collaboration

Collaboration was pointed out as one of the important topics in a recent survey [41]. During the past, co-located and remote immersive collaboration systems have been developed. Multi-user entertaining experiences is one trend. For example, Popovici and Vatavu [67] examined users' preferences for AR television scenarios. Increasing engagement for single user [34] and for sharing museum experience was widely discussed [19]. Haptic feedback is investigated for remote collaboration [33]. Remote guidance is popular for AR and VR collaboration, such as exploring visual communication cues [42], creating virtual replicas of local objects for remote experts [18], updating remote objects based on local users' actions [81], and providing multiple view sharing techniques [49].

Developing telepresence experiences for bridging the gap between the physical and virtual worlds plays a vital role for remote collaboration. Teo *et al.* [79] explored mixing 360 video and 3D reconstruction for remote collaboration. MetaSpace [74] performed full-body tracking. Young and Cook [90] provided a hand overlay on a panoramic reconstruction. Holoportation [57] demonstrated real-time 3D reconstructions of an entire space with a comprehensive setup of eight cameras and gigabyte-level bandwidth. Beck *et al.* [2] implemented immersive group-to-group telepresence, which allowed distributed groups of users to meet in a shared virtual 3D world through two coupled projection-based setups. Similarly, Pejsa *et al.* [62] presented Room2Room, a telepresence system that leverages projected AR to enable life-size, co-present interaction between two remote participants. SharedSphere [50] was implemented to investigate how Mixed Reality (MR) live panorama reconstruction affects the remote collaborative experience with non-verbal cues.

In addition, collaborative tools, such as editing [4], manipulation [23], modeling [87], and information analysis [8, 9], were proposed for productive work in immersive environments. Hsu *et al.* [36] developed an architecture design discussion system that allows members to visualize, discuss, and modify the architectural models. Members communicate via voice, object manipulations, and mid-air sketching as well as on-surface sketching. Object manipulation and navigation were under research for decades. T(ether) [46] was a spatially-aware display system for co-located collaborative manipulation and animation of objects. T(ether) attached trackable markers on the pads so participants with gloves can interact with the objects through gestures. Kunert *et al.* [44] designed an application to support object manipulation tasks and scene navigation. Oda *et al.* [56] developed a distributed system for remote assistance. Geollery [16, 17] focused on social experiences by creating an interactive MR social media platform. Mahmood *et al.* [54] presented a remote collaborative visualization system through providing co-presence, information sharing, and collaborative analysis functions to discuss complex problems like environmental pollution.

For collaborative purposes like social networking and telepresence, engagement and a sense of being there are the most important qualities. In those scenarios, communication performance is not the focus. While for collaborative purposes, such as productive work,

---

[1] CollaboVR GitHub: `https://github.com/snowymo/CollaboVR`.

games, assistance, and object manipulation, which require complicated and specific operations and information exchange, communication performance becomes more important. In CollaboVR, our goal is to build a reconfigurable framework to fit different purposes of creative collaboration, including side-by-side whiteboarding, face-to-face demonstration, and lectures with a presentation.

2.4 Communication in Immersive Environments

We researched the trends of communication in immersive environments. Asymmetrical communication was under discussion for scenarios that not all the participants use the same device [24]. ShareVR [25] enabled the communication between an HMD user and a non-HMD user. Through floor projection, the non-HMD user can interact with the HMD user and become part of the VR experience. Mutual human actuation [10] ran pairs of users at the same time and had them provide human actuation to each other. Communication between the pair was through the shared interactive props. Avatar representation plays an important role [48]. Mini-Me [66] was an adaptive avatar representing the remote user's gaze direction and body gestures. Chow *et al.* [11] identified several challenges for time-distributed collaborators in asynchronous VR collaboration. Maintaining workspace awareness is one challenge.

Interacting with digital content in shared space also triggers a line of in-depth research. Kiyokawa *et al.* [43] have researched the communication behavior for two participants in collaborative AR. They found that placing the task space between participants led to the most active behaviors through an icon design task. "Three's Company" [76] explored three-way collaboration over a shared visual workspace. They illustrated the utility of multiple configurations of users around a distributed workspace. TwinSpace [68] supported deep interconnectivity and flexible mappings between virtual and physical spaces. Sra *et al.* [73] proposed "Your Place and Mine" to explore three ways of mapping two differently sized physical spaces to shared virtual spaces and to understand how social presence, togetherness, and movement are influenced. Irlitti *et al.* [38] discussed how to design and provide spatial cues to support spatial awareness in immersive environments for remote collaboration. Likewise, Volmer *et al.* [85] provided projector-based predictive cues to improve performance and to reduce mental effort for procedural tasks. Tan *et al.* [75] built a face-to-face presentation system for remote audiences. Lukosch *et al.* [53] pointed out that face-to-face collaboration increased social presence and allowed remote collaborators to interact naturally. Tele-Board [26] described a groupware system focused on creative working modes using a traditional whiteboard and sticky notes in digital form for distributed users. Benko *et al.* [3] proposed a unique spatial AR system that enables two users to interact in a face-to-face setup. Thanyadit *et al.* [82] presented ObserVAR to discuss gaze awareness and visual clutter for VR classroom.

Inspired by prior arts in one-one communication systems with two participants or at most three-participant cases, we further extend the scalability and design space of creative collaboration in VR. Previously, we performed preliminary exploration of face-to-face setup specifically for remote presentation task [31, 32]. In CollaboVR, we extend our system to enable group collaboration with more than two participants and allow them to seamlessly switch the layouts in real-time. Additionally, we carefully designed a collaborative task for a group of four participants to see how our system may help to resolve conflicts and reach consensus by sketching and negotiating.

3 SYSTEM SCOPE AND OVERVIEW

Our overarching goal is to propose a reconfigurable framework for creative collaboration in VR, which can adapt to different use cases and optimize virtual spaces depending on selected task. We restrict our scope to teamwork with whiteboards and visual information. We next describe our use cases and system architecture.

3.1 Creative Collaboration Use Cases

We envision the following potential use cases for CollaboVR.

Travel planning and brainstorming. CollaboVR can be used for trip schedule as presented in Fig. 1(a). Multiple remote users are rendered as virtual avatars in front of a large virtual interactive board. With freehand drawing, users can write and draw their desired travel plans and coordinate with friends via both audio communication and sketches. When they have different ideas, users can easily duplicate the current interactive board and iterate on the prior one to express new alternatives.

Interactive lectures. CollaboVR can also be used for interactive classes as shown in Fig. 1(b). In this case, CollaboVR places the presenter and the audience on opposite sides of the interactive board. Sketches are shown identical to both the audience and the presenter, so that the presenter and the audience observe the same scene. With face-to-face remote communication, the presenter may pay more attention to the audience's focus, while the audience can simultaneously follow the presenter's gestures and content.

Presenting live designs on a sketchpad. Writing directly on a whiteboard is not always preferred in creative collaboration sessions. Many users feel more comfortable writing on a notepad or tabletop while sitting in a chair. Hence, we enable CollaboVR to support a projection input mode as shown in Fig. 1(c). In this example, the lead designer can focus on sketching a baroque pattern on a small, flat, private interactive board. The experience is similar to drawing on a digital tablet with a pen while the contents will be projected to the large, shared interactive board to other audiences. Upon finishing, the lead designer may look at the audience and ask for questions and suggestions. Other participants can contribute by pointing or sketching onto the projected design draft.

Designing spatial layout. CollaboVR can also help with designing spatial layouts, especially in 3D. Imagine that a user has just moved into a new apartment and needs to remotely discuss the placement of the furniture with other roommates. As Fig. 1(d) demonstrates, the user can draw furniture with a combination of primitive 3D objects and place them directly at preferred locations. Spatial layout is difficult to describe clearly through words and gestures, and it often requires freehand drawings, multiple iterations, and multiple perspectives in 3D. CollaboVR satisfies users' needs by offering them a rich set of interaction tools and real-time sketch-to-object techniques via cloud-apps.

3.2 System Architecture and Workflow

CollaboVR aims to offer a reconfigurable architecture for creative collaboration in VR with lightweight software on the client side and low-latency services on the server side. Hence, we leverage a cloud-based architecture where the computational expensive applications are hosted on the servers and the rendering results are streamed to all clients.

As a proof-of-concept, we employ Chalktalk [63] as the server application to enable creative collaboration in VR. While there are many smart sketch-based online software programs – such as Autodraw [22], sketch2code [45], and Miro [72] – that can assist creative collaboration, Chalktalk is an open-source software with a rich set of sketch-based communication language and digital animations. It allows a presenter to create and interact with animated digital sketches on a blackboard-like interface. We chose to use Chalktalk because it is an open-source platform, so we can easily define the data-flow between the application and CollaboVR.

We designed an extendable protocol in the CollaboVR framework so it can work with other applications as long as the input and output are accessible. The protocol serializes input and display data from each user, routes that data through a network, and then de-serializes and interprets the data to correctly render the results into graphics.

The CollaboVR server is written in Node.js and C#. It synchronizes data across devices and supports custom data formats. For

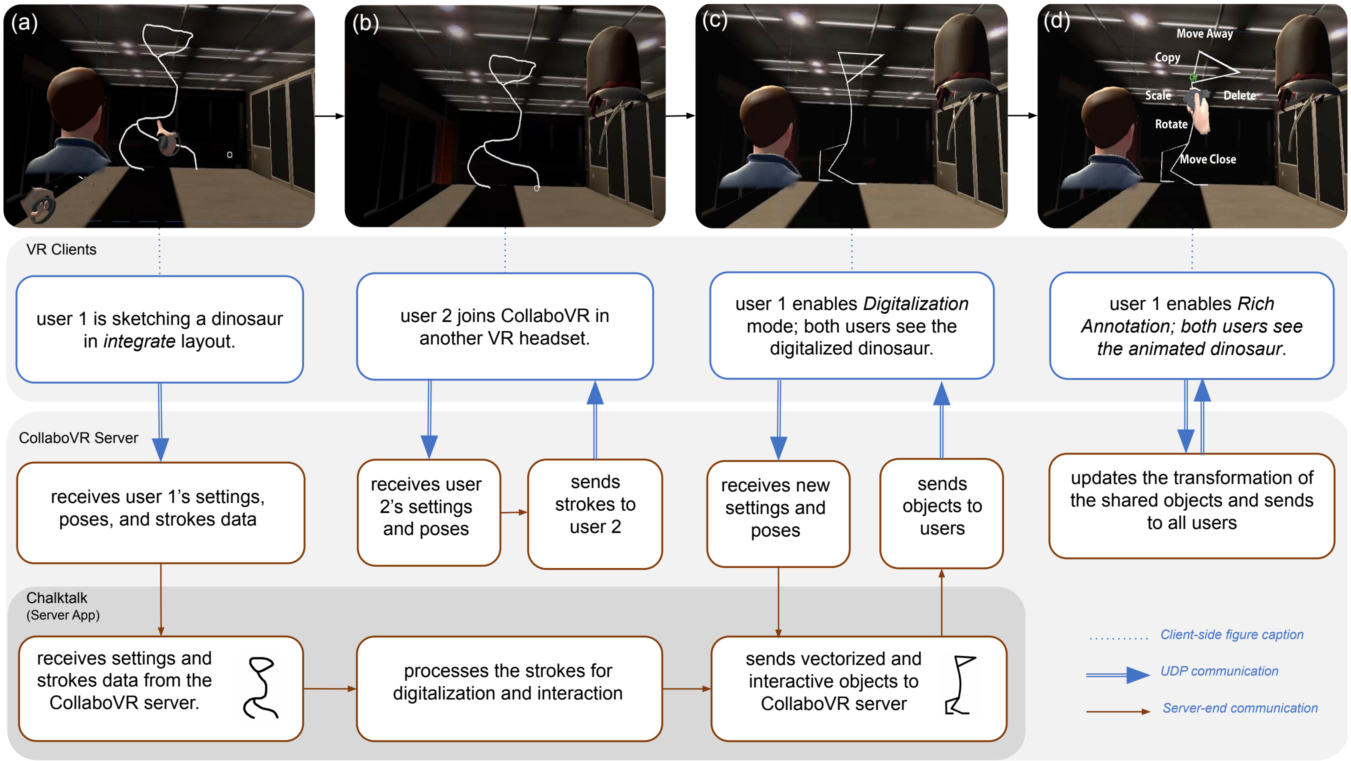


Figure 2: The workflow of CollaboVR: (a) As user 1 sketches in CollaboVR, the server receives the aggregated data of settings, poses and strokes and sends the strokes data to Chalktalk for further processing; (b) When user 2 joins CollaboVR, the server broadcasts the poses from user 1 as well as the latest stroke data so that both users see the sketches and each other; (c) After user 1 triggers the *Digitalization* mode and notifies the CollaboVR server, the server queries Chalktalk. In less than 16 milliseconds, Chalktalk converts the strokes into interactive objects. Then both users see digital objects (in this case, a triangle and several spline curves) from the CollaboVR server; (d) When user 1 performs *Rich Annotation* on the sketch, the CollaboVR server alone handles commands for scene editing tasks.

CollaboVR, we have two kinds of information: rendering data and user data. Fig. 2 demonstrates the workflow of ColalboVR. For rendering data, we first pass the user input from each client to the server. Then, the server transmits the user input together with its user identifier to the application. Next, the server receives the serialized display data from the application (Chalktalk). Finally, the server broadcasts the display data to each client for rendering. For user data, we broadcast the user's avatar, poses, and audio stream to each client after it has been received.

To unleash the users' creativity, we design "Rich Annotation" mode to empower CollaboVR clients to manipulate sketches and objects. After the clients receive and render the display data from the application, the display data are considered as interactive objects in a 3D world. This manipulation includes duplication, linear transformation (rotating, scaling, and translation), deletion, and colorization.

4 CUSTOM CONFIGURATIONS

As motivated in the Introduction, we designed CollaboVR as a reconfigurable framework to investigate the best configuration for creative collaboration tasks. Previous work has great insights on one specific user arrangement or input mode. We investigate three user arrangements (integrated, mirrored, and hybrid) and also offer two input modes (direct and projection). "Projection mode" is designed to see whether it is more effective for expressing ideas in remote presentations. Inspired by prior art in workspace awareness [28, 29], we focus on maintaining and enhancing workspace awareness, to empower participants to work together more effectively. CollaboVR allows users to alter their views of other participants. In other words, they can manipulate the spatial layout by which they see other users.

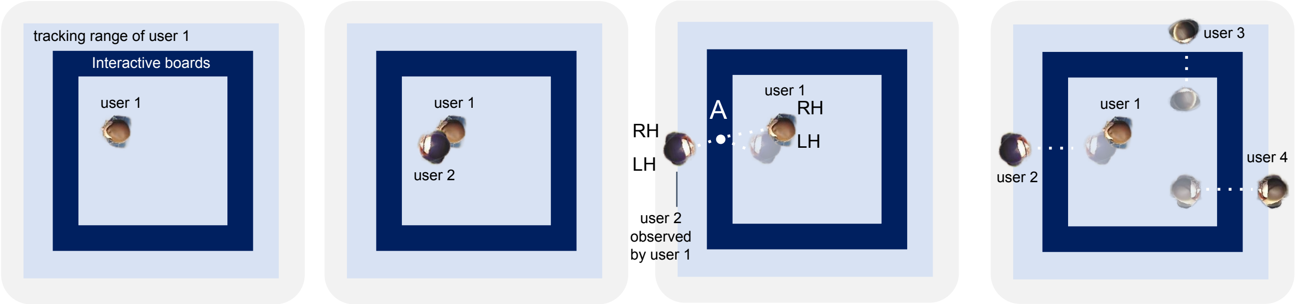
Concluded by previous work on collaborative sensemaking [39, 55, 60], we notice that multiple and shared large displays are useful for collaborative work in terms of 2D information. CollaboVR is an immersive 3D graphics world. Instead of "display", we set up multiple "interactive boards" in the virtual environment.

4.1 User Arrangements

We offer three user arrangements for CollaboVR: (1) side-by-side, (2) face-to-face, and (3) hybrid arrangement.

The side-by-side arrangement places each remote user into a shared virtual space, which is defined within the tracking range of the VR headset as shown in Fig. 3(a). The side-by-side arrangement enables multiple users to collaborate side-by-side in front of the same interactive board. All users may focus on the contents during the creative collaboration. However, two user avatars may be occluded with each other as illustrated in Fig. 3(b).

The face-to-face arrangement solves the occlusion issue by mirroring all the other avatars' locations to the other side of their currently activated interactive board. In Fig. 3(c), user 1 enables the face-to-face arrangement so user 2 in user 1's view is mirrored to the other side of the left interactive board which user 2 is looking at. Now let's take a look at the gaze interaction. Spot A is the same content that both users are looking at. After the mirroring operation, the gaze direction of users is maintained. Moreover, users are aware of each other's focus while gazing at spot A at the same time. In contrast to FacingBoard [51], we did not mirror-reverse the content so the content is still correct to each viewer. We then consider how spatial



(a) User 1 joined CollaboVR; walking range is limited within the interactive boards.

(b) After user 2 joined CollaboVR, side-by-side arrangement may cause occlusion or collision.

(c) Face-to-face arrangement solves this issue by mirroring user 2 by the shared interactive board.

(d) Face-to-face arrangement enables one-to-many communication and reduces visual clutter.

Figure 3: Comparison between *side-by-side* and *face-to-face* arrangements. (a) shows one user in CollaboVR. Interactive boards are depicted as dark blue rectangles. (b) shows two users in the side-by-side arrangement. This arrangement is intuitive to users and supports side-by-side whiteboarding tasks, but may cause occlusion or collision of virtual avatars. (c) shows how a face-to-face arrangement can solve this problem. For each user, the face-to-face arrangement mirrors the other user, so direct eye contact is preserved and both users can see each other while sketching on the same interactive board. Spatial direction remains the same, see 'LH' and 'RH' indicators of user 1 and user 2 observed by user 1. (d) shows an extended version with four users from user 1's perspective. Each user sees the others mirror reversed through their respective boards. Compared with the side-by-side arrangement, our face-to-face arrangement reduces visual clutter while maintaining eye contact.

instruction looks like. User 2 with transparent shading indicates the original position of user 2. From Fig. 3, we know that user 2 is on the left side to user 1 originally. Equivalently user 1 is on the right side to user 2. After enabling face-to-face arrangement in user 1's perspective, the spatial relationship remains the same for all the users. Face-to-face arrangement is like a mirror. Users only need to consider the spatial relationship from their own perspective, see the left hand (LH) and right hand (RH) indicators in Fig. 3(c) for user 1 and user 2 observed by user 1. In this user arrangement, the users can see each other for better workspace awareness. Fig. 3(d) illustrates the multi-user scenario when users are looking at different interactive boards. Compared with the integrated layout where every user is restricted within the shared virtual space, our mirrored layout greatly reduces the visual clutter and maintains users' eye contact.

The hybrid arrangement inherits the "teaching in a classroom" metaphor, where the teacher uses the face-to-face arrangement to observe students, and the students use the side-by-side arrangement for classmates and a face-to-face arrangement for the teacher. We envision that this arrangement may be useful for online education with a large audience.

4.2 Input Modes

Motivated by the two metaphors of writing on a whiteboard and sketching in a notebook, we offer two input modes in CollaboVR to support different use cases: direct mode and projection mode.

Direct mode adapts the metaphor of writing on a whiteboard (Fig. 1(a)). This may be best used where the user experience is similar to a brainstorming or interview session in the meeting room.

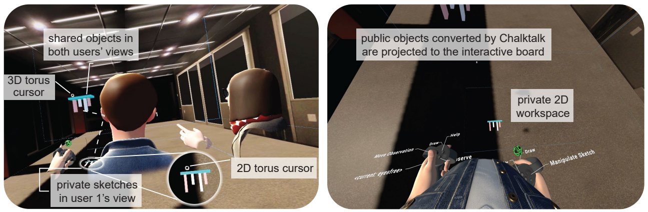
In addition to the direct mode where the user sketches on the interactive board, CollaboVR enables projection mode, where the user may sketch on a private workspace at the hands and project the contents onto the shared interactive board. We present both a third-person and a first-person perspective of the projection mode in Fig. 4. The private workspace is placed at an approximately 1-meter height, lower than users' hands, so the drawing won't be displayed above users' arms whether sitting or standing. For the other users who are not sketching, the content is duplicated and rendered on the shared interactive board, see Fig. 4(a). By doing this, we avoid a situation whereby the content is not readable for all the users around a table. When the user is writing in the private workspace, he/she is free to look at personal workspace or the shared interactive board (see Fig. 4(b).) Moreover, content on the private workspace is different from the content on the shared interactive board in two ways: scale and dimension. Given that the reach distance when writing on private workspace is smaller than on the shared one, we adjust the scale of the private workspace. Regarding the dimension of the content on the private workspace, we squeeze the content and render the content in 2D. (See how the table looks in Fig. 4.) The reason we implement squeezing is that we prefer to simulate a tablet-style input and keep the designing space clean as well. To enhance the awareness of where the user is writing, we render the projection point of the user's controller as a 3D/2D cursor (Fig. 4).

5 TECHNICAL IMPLEMENTATION

We implement CollaboVR with an extendable networking protocol, a calibration approach for co-located users, a client software for free-hand sketching and object manipulation, and a server-end software, Chalktalk to digitalize the sketches and generate animations.

5.1 Networking Protocol

For each creative collaboration session (like client session or server-end application session), CollaboVR establishes a UDP network for



(a) third-person perspective of the projection mode with two users.

(b) first-person perspective of the presenter modifying private sketches.

Figure 4: Projection mode. (a) demonstrates user in blue drawing a table in projection mode from a third-person perspective. There is a private sketch that only the person who is drawing can see. It is laid out in 2D at the user's waist height, meanwhile a 3D object is displayed in the interactive board for all the users to see. (b) shows the first-person perspective when the user looks down and creates his 2D sketch.

low-latency and real-time performance. The user data and rendering data need to be transmitted every frame. The server is written in Node.js and the client is written in Unity C# and Node.js.

We defined a *synchronizable object* as an object that needs to be synchronized each frame for the client who registered it. Each synchronizable object has a label and data stream. The label is a unique id for the client to register. The data stream includes the sending frequency and real-time data.

We provide two frequency values in the system: *one-time* and *per-frame*. A *one-time* synchronizable object is designed for sending commands including `join CollaboVR`, `switch to certain board`, `select objects`, etc. It does not happen for each frame. For a *one-time* object, we use two-way handshaking metaphor. The client sends the object to the server, the server returns an object including acknowledgement back to the client, then the client deregisters the object with this local label. The *per-frame* synchronizable object includes avatar representation, the audio data, and the display data from the Chalktalk application. We design a protocol to wrap all the display data. The data protocol includes information of all the rendered lines and meshes by encoding their attributes. Each client deserializes the data from the server and renders the deserialized data as strokes or meshes. Fig. 5 shows how CollaboVR performs with an increased number of clients. We evaluate a four-client case in user study while the system can afford at least 10 clients simultaneously. Overall, the networking latency is under 10 ms, rendering performance stays above 60 frames per second even when there are 10 clients discussing a full living room scene, and the throughput per frame is quite stable when the number of clients increases.

5.2 Calibration for Co-located Scenarios

CollaboVR works for both co-located and physically distributed scenarios. For distributed users, we simply overlap their virtual environments because they do not have any spatial relationship in reality. For co-located users, we need to carefully calibrate their relative locations, so their avatars are rendered in the same coordinate system. The key idea for calibration is that different clients should have a shared trackable proxy by their camera systems.

In Fig. 6, we present an example with HTC Vive Pro headsets in the co-located modes of CollaboVR. We enabled the mixed-reality mode to capture the co-located user setup. The shared proxy in Vive system is the base station. Each machine running Vive can retrieve the transformation of the base station. Because all machines (assuming N machines) have their own coordinate systems, we have N pairs of the transformation of the base station. We choose one base station as the proxy based on the unique serial number. Then, we treat the first connected client as the reference node. Later, all the following $N-1$ clients apply the inverse matrix between the base station of the reference node and their own base station. Fig. 6 shows user 1 drawing a physics model. Fig. 6(a) presents the front view and Fig. 6(b) presents user 2's view. With this co-located setup, users are unlikely to collide with each other and have occlusion.

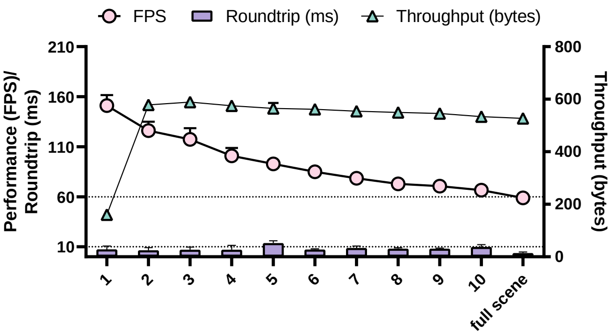
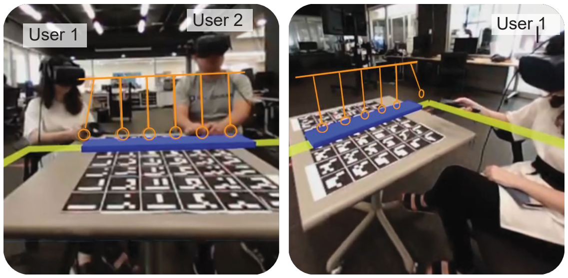


Figure 5: Chart of network latency and rendering performance as the number of clients ranges from 1 to 10. The last column shows the results with 10 clients as well as a fully designed living room. Networking latency remains around 10ms consistently; rendering performance drops from 160fps to 60fps with 10 clients; throughput does not change appreciably with an increased number of clients, but depends rather on the complexity of the displayed scene.



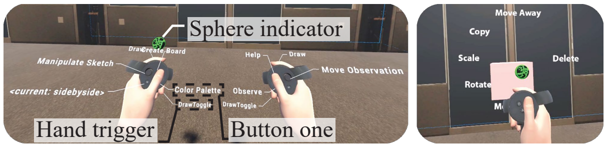
(a) Two users discussing an interactive Newton's cradle.

(b) First-person view of user 2 when user 1 manipulates the cradle.

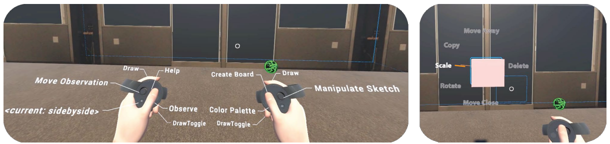
Figure 6: An example of our co-located user setup using HTC Vive Pro with accurate calibration. (a) shows two users discussing Newton's cradle in CollaboVR. (b) shows user 1 dragging a virtual ball to interact with objects in CollaboVR.

5.3 Client Software

CollaboVR includes UI for users to convert raw sketches into digital objects and manipulate them after freehand sketching. We provide the functionality of duplication, transforming, deletion, and colorization. To achieve this, we designed a pie menu triggered by the controller. The following is the workflow for a user's manipulation: first, place the controller so it hovers over the drawing of interest; second, press the thumbstick of the dominant controller; and then, the pie menu appears as Fig. 7(c); later, move the thumbstick to select the specific menu (see Fig. 7(d)); afterward, apply the corresponding movement in terms of the command and release the thumbstick. The color palette is toggled by button one, illustrated in Fig. 7(a). The user can drag the color from the palette to any drawing like world builder [69].



(a) button manual for left-handed users. (c) Pie menu.



(b) button manual for right-handed users. (d) Select menu scale

Figure 7: User interfaces for sketching and scene editing in CollaboVR clients. (a) and (b) present the button manual for left-handed and right-handed users, respectively. A small green sphere indicates which hand is currently enabled for drawing. (c) shows the interface when the user selects the color palette function. (d) shows the interface for scene editing.

The controller's trigger switches the commands of the two controllers for left-handed and right-handed users (see Fig. 7). As the user's view might be blocked by other users' avatar, we implement a spectator mode. Users can see the view from different users in the lower right corner. To encourage all users to work on the task together, we implement a permission strategy. Only one user can draw at one time. Once the user who is drawing releases the permission, other users can grab permission to draw, see Fig. 7(a). Deploying CollaboVR requires only a VR device running Unity for each client, a server machine running Node.js, and an optional router for ensuring low latency for data transmission.

5.4 Cloud-hosted Software: Chalktalk

To save the rendering and computational budget on the client side and reduce communication overhead, our networking protocol supports communication and synchronization between clients and server-end software. As a proof-of-concept, we employ a variant of an open sourced presentation and communication language, Chalktalk [63, 64]. Chalktalk allows a presenter to create and interact with animated digital sketches in order to demonstrate ideas and concepts in the context of a live presentation or conversation. For each raw sketch, Chalktalk first matches its strokes with the most similar one in a library of 150 glyph. We designed our own glyph for experiment use. Based on the recognized glyph pattern, it further converts the raw sketch into digital objects that the user can manipulate. We illustrate examples of real-time conversion from raw sketch to animated objects in Fig. 8 and the supplementary video.

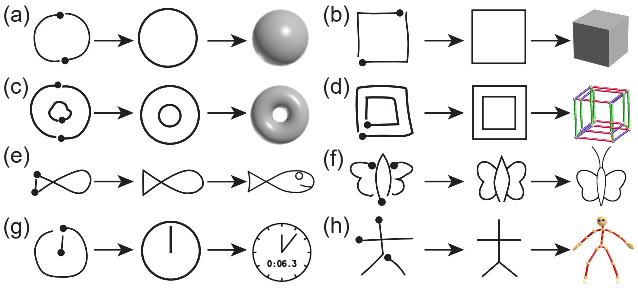


Figure 8: Examples of sketch recognition and object/animation generation in Chalktalk. Each subfigure shows three parts: 1) the raw sketch. Black dots indicate the starting positions of the raw strokes; 2) the intermediate conversion from the raw sketch to one of the 150 vectorized glyph; 3) the resulting instantiated object or animation. The user may translate, rotate, or scale the object as well as interact with it. These examples feature the generation of a) a sphere, b) a cube, c) torus, d) a hypercube, e) an animated fish, f) a butterfly, g) a running timer, and (h) a rigged avatar skeleton.

6 USER STUDY: SYSTEM EVALUATION

We evaluate the interaction cycles, design variables, and collaborative effectiveness of CollaboVR through a within-subject study to answer the following research questions: how does sketching affect real-time VR collaboration; how does interactive animations impact individual's behaviors, will it improve the performance of remote collaboration; are there best *user arrangements* and *input modes* for different use cases, or is it more a question of personal preferences? During the study, we collected qualitative feedback to gain insight into the potential benefits and impacts of CollaboVR, and quantitative data to research the mostly preferred layout with a collaborative design task.

6.1 Participants and Apparatus

We recruited a total of 12 participants at least 18 years old with normal or corrected-to-normal vision (5 females and 7 males, 1 left-handed and 11 right-handed; age range: $20-30$, $M = 23.58$, $SD = 3.45$) via campus email lists and flyers. None of the participants had been involved with this project before. The participants have reported various VR experiences in a questionnaire (rating scale: 1 (less) to 7 (more experienced), $Mean = 4.08$, $SD = 1.83$).

We deployed CollaboVR using Unity on workstations running Windows 10 with Nvidia GTX 1060 GPU, Intel Core i7 2.80 GHz CPU, and 16GB of RAM. We used Oculus Rift CV1 with two Touch controllers. Computers were connected to the router through Ethernet cables. For the duration of the study, participants' behavior, including their interaction patterns, body language, and strategies for collaboration in the shared space were observed and recorded.

In the study, four participants were grouped as a team. We instructed each group with one training session and three design sessions to perform a collaborative design task. After the design session, the researcher conducted semi-structured interviews to obtain additional insights into the most salient elements of the users' experience, challenges, and potential user scenarios. Next, we detail the training stage, design sessions, and interview stage.

6.2 Training Stage

At the beginning of each study session, we first introduced the project to the participants and collected consent forms for screen recording and video recording. Next, we gave the group a 10-minute lecture on Chalktalk and taught the participants how to create freehand drawings and convert them to 3D objects.

In the next 10 minutes, participants were given a demo on how to use CollaboVR. As part of the demo, a researcher put on the headset, mirrored the VR content with a regular monitor, and described how to use each button as well as each function, including sketching on the interactive board, obtaining permission to draw, and manipulating drawings and objects. Afterwards, all participants were placed in physically distant locations with an Oculus Rift running a CollaboVR client. We instructed the participants to try in-air sketching and object manipulation until all participants were familiar with the interaction paradigms. Finally, we put all participants into a shared virtual environment and started design sessions. Overall, the entire training session took approximately 30 minutes.

6.3 Design Sessions

Next, all the groups were asked to experience three 10-minute sessions in randomized orders. Each session featured a different condition motivated by real-world scenarios as follows:

C1: **integrated layout** which inherits the "physical side-by-side white-boarding" metaphor. This condition places all participants into a shared virtual space without any further arrangement. However, remote users have to rearrange their avatars to avoid visual clutter and occlusion.

C2: **mirrored layout** which inherits the "face-to-face communication" metaphor. This condition resolves the former clutter and occlusion issues by using the face-to-face arrangement as introduced in Sect. 4.1.

C3: **projective layout** which inherits the "lecture with a presentation" metaphor. In this condition, users can draw their design in their private workspace (as explained in Sect. 4.2) and then project it into the shared whiteboards to the audience at the opposite side. This may allow users to focus on individual design without too much distraction of the shared white boards.

To explore the use of the CollaboVR system for creative collaboration in the shared virtual space and motivated by the "building block" task in Holoportation [57], we further designed a "living room design" task. In each session, the participants were asked to design a living room containing only three pieces of furniture: a table, a chair, and a couch. To simulate conflicts and encourage discussion as in normal meetings, we asked each participant to pick one piece of furniture, sketch an original 3D design, and write down the layout of the three furniture before entering CollaboVR. We instruct the participants to be creative in color, shape, and textures of the selected furniture. Since only three items are assigned to four participants, the participants would have to resolve conflicts and come to a consensus through CollaboVR. After the individual ideation phase, the researcher instructed each participant to wear the VR headsets, enter CollaboVR, express their original ideas, and attempt to reach an agreement for the living room design. After each design session, they took off the headset and wrote down their final decisions for the design in a text file. After a five-minute break, they entered the next 10-minute session.

6.4 Semi-structured Interview

Afterwards, the researcher presented the participants with a set of statements (adapted from System Usability Scale [7] on CollaboVR and each session on a 7-point Likert scale). Next, the researcher conducted a semi-structured interview asking about their experience, trying to gain insight into usability and use cases of the system.

6.5 Data Collection and Analysis

We conducted one-way repeated measurements analysis of variance (RM-ANOVA) statistical analysis to examine the variations between different conditions on user preference, usability, and collaboration effectiveness for each participant, and the *task performance* for each group. *Task performance* is defined as the details of the living room design for each session. We analyzed what they wrote before and after each session by calculating the quantity of the details, such as color, shape, and texture. For example, "a *yellow triangle-based* table with *flower texture*" is counted as 3 points, "a chair with *wood material*" is counted as 1 point. The collaborative task is aiming at how participants discuss and come to a consensus of a topic requiring visual description, rather than how well their final design appears. Therefore, we observed the final design they completed, yet did not take its aesthetics into performance evaluation. The level of RM-ANOVA significance was set at $p < 0.05$.

7 RESULTS AND ANALYSIS

In this section, we analyzed CollaboVR in general, compared each condition for individual behaviors, and evaluated the effectiveness of collaboration for three conditions. In brief, we examined that CollaboVR is helpful to express ideas with high usability. Out of three conditions, the majority of the participants preferred mirrored layout and found it good for task completion and partner connection.

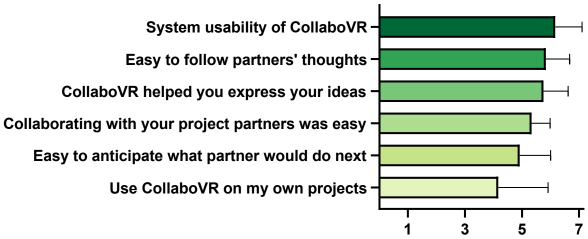


Figure 9: Overview of subjective feedback on CollaboVR. On the SUS, participants categorized CollaboVR as a "good and acceptable" system, $M = 6.17$. It was moderately easy to follow others' thoughts ($M = 5.83$), to express the ideas ($M = 5.75$), and to collaborate with partners ($M = 5.33$) with CollaboVR. Participants were positive about anticipating partners' next movement ($M = 4.92$) and using CollaboVR on their own projects in the future, $M = 4.17$.

Participants showed the willing of using CollaboVR in daily life and shared the thoughts of ideal scenarios for three conditions.

CollaboVR in general. We analyzed the result of CollaboVR usability ($M = 6.17$, $SD = 0.94$), how helpful is CollaboVR to express ideas to the group ($M = 5.75$, $SD = 0.87$), and whether the participant wants to use CollaboVR in their own project in the future ($M = 4.17$, $SD = 1.75$) (in Fig. 9). From the observation, CollaboVR's pipeline was quickly mastered by all participants during the training session. P9(F) commented *"it is intuitive to do the drawing in 3D."*. Moreover, P11(M) responded, *"it's totally a great prototyping idea/prototyping system. Can't say it'll replace AutoCAD, but in a few years it will do that."*.

Individual behaviors among conditions. We conducted RM-ANOVA tests to compare the effect of three conditions – integrated, mirrored, and projective layout – on ratings, how helpful for performing tasks, in-sync with other partners, connected with other partners, and easiness to use. We found a significant effect of the three layouts on ratings, $F(2,22) = 5.73$, $p = 0.01$. Post hoc comparisons using Bonferroni test indicated that the mean score for mirrored layout ($M = 6.08$, $SD = 0.79$) was significantly different from the projective layout ($M = 4.42$, $SD = 1.56$). However, the integrated condition ($M = 5.42$, $SD = 0.99$) did not significantly differ from mirrored condition and projective layout. In brief, these results suggest that the mirrored layout yields better ratings of the "living room design" task (Fig. 11(a)).

Additionally, we found a significant effect of the conditions on easiness to use, $F(2,22) = 11.76$, $p < 0.01$. Post hoc comparisons using Holm test indicated that the mean score for projective layout ($M = 4$, $SD = 1.71$) was significantly lower than the integrated condition ($M = 6.08$, $SD = 0.79$) and mirrored condition ($M = 6$, $SD = 1.04$)(Fig. 12(b)).

A significant effect of the conditions on "helpfulness in performing tasks" was found, $F(2,22) = 7.03$, $p = 0.004$. Post hoc comparisons indicated that the mirrored condition ($M = 6.17$, $SD = 0.72$) had significantly higher mean values than the integrated ($M = 5.17$, $SD = 1.03$) and projective layout ($M = 4.5$, $SD = 1.38$)(Fig. 11(b)).

We also asked participants on the rankings of the layouts. 58.3% of the participants (7 out of 12) preferred the mirrored layout most, while 25% (3 out of 12) of the participants thought the integrated layout is their favorite and two participants preferred the projective layout (see Fig. 10). One sample Kolmogorov-Smirnov test indicated that the user preferences did not follow a normal distribution, $D(12) = 0.3$, $p = 0.004$ (see Fig. 10).

Those who preferred mirrored layout mentioned: *"In mirrored it is easy and convenient to communicate with others."*(P3,F). *"People didn't block my view, and I could see the content clearly."*(P5,M). *"[It is] more helpful when working on a group project. Feels like I have enough space to draw."* (P9, F).

Participants who preferred integrated layout explained that *"because it is comparable to reality."*(P2,M). P1(M) had a similar opin-

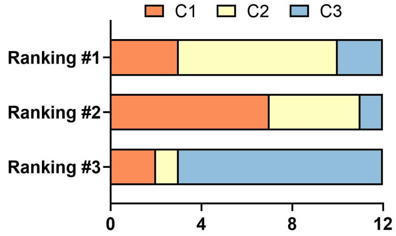


Figure 10: Rankings of user preferences among integrated layout (C1), mirrored layout (C2), and projective layout (C3). Mirrored layout is preferred the most for the "living room design" task.

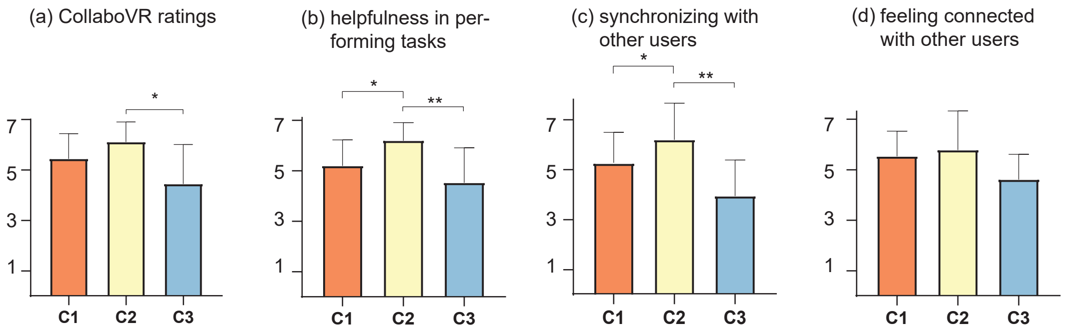


Figure 11: CollaboVR's ratings, degree of helpfulness users in performing tasks, in synchronizing with partners, and in connecting with partners using the integrated layout (C1), mirrored layout (C2), and projective layout (C3). *: $p < 0.05$, **: $p < 0.01$. We found a significant difference in ratings between C2 and C3; in degree of helpfulness between C1 and C2, C2 and C3; in synchronizing with partners between C1 and C2, C2 and C3. In terms of feeling connected with partners while using CollaboVR, the statistical results differed significantly among the three conditions. However, we did not find significant differences between each pair of conditions from post hoc tests.

ion *"because the real world is more similar to integrated layout."*

Two participants preferred projective layout emphasized that *"I could sit sketching and had more control." (P7,F).* P10(M) commented: *"[it] allows drawing on the table, more intuitive to draw."*

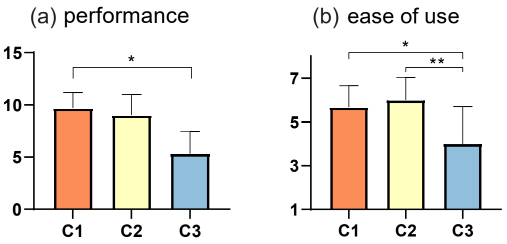


Figure 12: Comparison of performance and ease of use among integrated layout (C1), mirrored layout (C2), and projective layout (C3). *: $p < 0.05$, **: $p < 0.01$. We found a significant difference in performance between C1 and C3; ease of use between C1 and C3, C2 and C3. $p_{\text{performance}}(C2,C3) = 0.67$.

Effectiveness for Collaboration. Task performance of each group and questions about remote collaboration were analyzed through RM-ANOVA method. We found a significant effect of the conditions on task performance, $F(2,4) = 98$, $p < 0.001$. Post hoc comparisons indicated that the mean score for projective layout ($M = 5.33$, $SD = 2.08$) was significantly different than the integrated ($M = 9.67$, $SD = 1.53$) and mirrored layout ($M = 9$, $SD = 2$). However, the integrated layout did not significantly differ from mirrored layout. Therefore these results indicate that using projective layout has a negative effect for task performance. P8(M), a designer for 3D models who frequently used tablet for drawing, shared some feedback: *"this is like using a tablet. I preferred to spend more time on drawing the details and polishing my work when I was in this layout."*. Taking statistical results and subjective feedback into account, we think the projective layout may encourage participants to focus more on the details and better express themselves (Fig. 12(a)).

Participants thought it was easy to follow what partner was doing during the task ($M = 5.83$, $SD = 0.83$), easy to collaborate with others using CollaboVR ($M = 5.33$, $SD = 0.65$), and moderately easy to anticipate what partner planned to do next ($M = 4.92$, $SD = 1.08$). P3(F) commented that *"when [another user] started to draw the legs for the table, I quickly get his idea about the design of the legs, so he doesn't need to say what kind of legs he wants."* (Fig. 9). Furthermore, we ran RM-ANOVA test to compare different conditions on participants' feelings of connection and in sync with during the task. There was a significant effect of the condition on how connected do you feel to task partners, $F(2,22) = 3.89$, $p = 0.036$. Post hoc comparisons using Bonferroni test indicated that no significant effects among three conditions: projective layout has lowest score ($M = 4.58$, $SD = 0.99$), mirrored layout has the best result ($M = 5.75$, $SD = 1.54$) and integrated layout is in the middle ($M = 5.5$, $SD = 1$) (Fig. 11(d)).

Regarding how in-sync with the task partner during the experiment, we found a significant effect based on the RM-ANOVA results, $F(2,22) = 9.40$, $p = 0.001$. Post hoc comparisons showed that the mean score for each condition was significantly different from each other. The mirrored layout has the best results, with $M = 6.17$ and $SD = 1.47$. The integrated condition has a better-than-neural score in average ($M = 5.25$, $SD = 1.22$), while the projective layout has an average score ($M = 3.92$, $SD = 1.44$) (Fig. 11(c)).

Subjective Feedback. We asked participants what scenarios they would like to use CollaboVR and in which layout. The integrated layout is good for explanation in general. P8(M) commented, *"there could be merit once you're doing something more complex."*. P2(M), who rated himself as a novice VR user, thought, *" I like integrated layout because it is very easy to understand, just like reality."*.

Mirrored layout may be the best option for presentation. P4(M) recommended it because *"you can better control your drawing, meanwhile keep an eye on people's reaction."*. P5(M) considered it from a student's perspective, *"felt like Khan Academy [15] in 3D vision."*. P9(F) thought she can benefit from mirrored layout when brainstorming because no one is blocking the view, *"you can see everybody but you have your own space."*.

When discussing the suitable scenarios for the projective layout, P8(M) thought a VR live demo or presentation could be beneficial from projective layout, especially for a time-consuming one. He described *"himself giving a presentation to other people while an audience was looking at the large monitor-like board."* and *"just want to focus on the board."*. Meanwhile, P11(M) thought it would be helpful for collaborative design and suggested us to use a pen rather than the controller.

In general, the participants found it an engaging experience and love to spend more time with friends. *"It's definitely a fun environment, entertaining."*(P7,F).

Observations. When using mirrored layout, participants were confused about the spatial relationship in the beginning although researchers had explained it before the task. Then they quickly understood that other participants were in the "mirror". We also noticed that participants were willing to move one step aside when they were watching other participants and the content blocked the view between them and the others no matter in which condition. That suggests other alternatives should be associated with the face-to-face concept for maintaining eye contact. We also found that some of the participants preferred to look at the private workspace when in projective layout and others preferred to watch the shared board. For participants who were working on the content, providing the option to have eye contact or not for the participant is valuable.

8 DISCUSSION AND LIMITATIONS

With CollaboVR, we aim to explore opportunities and challenges for creative collaboration, explore the impacts of different layouts, and better comprehend the needs and challenges for multi-user communication in VR.

Improving remote creative collaboration. We consider the effectiveness of remote collaboration from two perspectives: how CollaboVR fosters communication among participants and how CollaboVR helps collaborative work. Our user studies showed that participants felt strongly connected with task partners when using their preferred condition ($M = 6.25$, $SD = 0.86$). In addition, they managed to follow their partners' work ($M = 5.83$, $SD = 0.83$) and anticipate their partners' behavior to some extent ($M = 4.92$, $SD = 1.08$). Participants felt highly in-sync with task partners while using CollaboVR ($M = 6.33$, $SD = 0.98$). Moreover, CollaboVR was greatly helpful to users for completion of the design task ($M = 6.33$, $SD = 0.65$). We concluded that CollaboVR can foster communication and help collaboration when participants are geographically dispersed.

User preferences. We found that the mirrored layout (C2) had better usability and task performance, and received the highest ratings from the participants. Many participants mentioned that the mirrored layout helped them focus on both the content and the other participants simultaneously and that their views were not blocked because of the layout design. The integrated layout received moderate scores from participants. Participants found it to be closest to real-life scenarios. That is to say that although the integrated layout did not solve certain issues, for example, participants' arms may block the sight of the audience, participants were able to alleviate those issues as they usually did in real life while having better communication and collaboration through CollaboVR. Projective layout was rated lowest but also showed the greatest potential in detail sketching and in being a good fit for long-term work.

We envision other user scenarios for CollaboVR. For example, CollaboVR could be used to communicate with others for non-expert use such as brainstorming and presentation. Different tasks may lead to different preferences in configurations. If the collaborative task is focused on object manipulation [70], floor plan design [81], or navigation [71, 88], participants may want to form a circle around the object. In that case, the mirrored layout is not very effective since the focus is not on the other participants of the group but rather on the objects to be manipulated. When giving a presentation, the presenter and audience may prefer different layouts. Mirrored layout maintains the gaze between the user and the others from the user's perspective, while sacrificing the gaze among others. However, integrated layout keeps this information. Although we evaluated each layout individually, CollaboVR is a reconfigurable framework that supports real-time layout switching and easy to scale to new layouts to meet various and changing requirements.

Miscellaneous User movement and tracking capability are usually constrained within the small space around the user's desk. Even if the user is not bounded by physical space, mirrored layout may be preferred for face-to-face collaboration; otherwise, the sketches appear reversed to the observer. Hence, the customization of user arrangements can greatly improve the overall user experience.

Projection mode leverages consistent mid-air user interaction as the direct mode. Supporting touchpads will be a nice extension for CollaboVR. However, the form factor of the current-generation touchpads may not be suitable for complex shapes.

Limitation. As a proof of concept and an example opensourced framework, one limitation of CollaboVR is that we currently only support one application, Chalktalk. Connecting various cloud-based applications will bring more possibilities and greater capability for CollaboVR. With recent advances in neural rendering [80], one may integrate GauGAN [59], SketchCOCO [13], and Text-based editing of talking-head [20] into CollaboVR.

Because our main contribution is the design and implementation of CollaboVR, and the exploration into different user arrangements and input modes, our user study focuses on comparison among the three layouts on a specific task, "designing a living room". A future study may enrich these results by allowing users to freely switchlayouts in real-time while assigning multiple collaborative tasks for different purposes, to study how the choice of layout for a given task may affect the results.

Potential Impacts. We envision that CollaboVR will be useful for collaborative scenarios such as remote presentations and virtual conferencing. For example, web conferencing software such as Google Meet and Zoom is widely used for meetings and 2D presentations. However, it is sometimes difficult for presenters to notice who in the audience is raising hands or asking questions, while also posing a challenge for audience turn taking. CollaboVR can help with such scenarios by providing workspace awareness. In virtual reality settings, Mozilla Hubs has been used to hold multi-user conferences with virtual avatars, yet provides very little support for creative collaborative work. CollaboVR may further extend the interaction capabilities of VR meetings by empowering participants to change meeting layouts and freely express their ideas by sketching or writing on virtual whiteboards.

9 CONCLUSION AND FUTURE WORK

We presented CollaboVR, an open-source reconfigurable framework for distributed and co-located creative collaboration in immersive environments. Our system was motivated by real-world metaphors such as side-by-side whiteboarding, face-to-face lecturing, and designing on sketchpads. We described the cloud-based system architecture, two design variables (user arrangement and input mode), rich interactive user interface, and corresponding technical details. We conducted a within-subject user study to quantitatively and qualitatively evaluate CollaboVR and compared three conditions: integrated, mirrored, and projective layouts. Our experimental results indicate that all participants can easily interact with CollaboVR and we found a significant difference in performance and ease of use in integrated layout v.s. projective layout and mirrored layout v.s. projective layout. Feedback from our interviews further suggested that CollaboVR is entertaining for communication and very helpful to foster collaboration. A few participants suggested that they would consider CollaboVR as a daily-life tool and can envision its potential for creative collaboration. Overall, the mirrored layout was mostly preferred by participants for our "design a living room" task, as it encourages more eye contact, and participants found it easy to reach a consensus when conflicts occur.

As we opensource CollaboVR as an extendable collaborative VR framework, we hope it will facilitate future research in collaborative work in VR, including extending the design space of sketch-based interaction, exploring effects of non-verbal cues in multi-user communication, and adding deep-learning-based models as cloud-hosted applications in CollaboVR. Eventually, virtual communication can in some ways be more effective than physical collaboration by giving remote participants the superpower to visualize ideas with speech and sketching [59], by transmitting physical or digital contents with cross-device interaction [35], and see, hear, and even feel each other by real-time reconstruction [57] and powerful sensors [30].

REFERENCES

[1] S. Albolino, R. Cook, and M. O'Connor. Sensemaking, Safety, and Cooperative Work in the Intensive Care Unit. *Cognition, Technology & Work*, 9(3):131–137, 2007. doi: 10.1007/s10111-006-0057-5

[2] S. Beck, A. Kunert, A. Kulik, and B. Froehlich. Immersive Group-to-Group Telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):616–625, 2013. doi: 10.1109/TVCG.2013.33

[3] H. Benko, A. D. Wilson, and F. Zannier. Dyadic Projected Spatial Augmented Reality. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, pp. 645–655. ACM, 2014. doi: 10.1145/2642918.2647402

[4] O. Bergig, N. Hagbi, J. El-Sana, and M. Billinghurst. In-Place 3D Sketching for Authoring and Augmenting Mechanical Systems. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 87–94. IEEE, 2009. doi: 10.1109/ISMAR.2009.5336490

[5] D. Billman and E. A. Bier. Medical Sensemaking With Entity Workspace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 229–232. ACM, 2007. doi: 10.1145/1240624.1240662

[6] F. Bork, R. Barmaki, U. Eck, K. Yu, C. Sandor, and N. Navab. Empirical Study of Non-Reversing Magic Mirrors for Augmented Reality Anatomy Learning. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 169–176. IEEE, 2017. doi: 10.1109/ISMAR.2017.33

[7] J. Brooke et al. SUS-A Quick and Dirty Usability Scale. *Usability Evaluation in Industry*, 189(194):4–7, 1996. doi: 10.1080/10447310802205776

[8] S. Butscher, S. Hubenschmid, J. Müller, J. Fuchs, and H. Reiterer. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018. doi: 10.1145/3173574.3173664

[9] M. Cavallo, M. Dholakia, M. Havlena, K. Ocheltree, and M. Podlaseck. Dataspace: a Reconfigurable Hybrid Reality Environment for Collaborative Information Analysis. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 145–153. IEEE, 2019. doi: 10.1109/VR.2019.8797733

[10] L.-P. Cheng, S. Marwecki, and P. Baudisch. Mutual Human Actuation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 797–805. ACM, 2017. doi: 10.1145/3126594.3126667

[11] K. Chow, C. Coyiuto, C. Nguyen, and D. Yoon. Challenges and Design Considerations for Multimodal Asynchronous Collaboration in VR. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–24, 2019. doi: 10.1145/3359142

[12] C.-W. Chung, C.-C. Lee, and C.-C. Liu. Investigating Face-to-Face Peer Interaction Patterns in a Collaborative Web Discovery Task: the Benefits of a Shared Display. *Journal of Computer Assisted Learning*, 29(2):188–206, 2013. doi: 10.1111/j.1365-2729.2012.00493.x

[13] N. M. Davis. Human-Computer Co-Creativity: Blending Human and Computational Creativity. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2013.

[14] B. Dervin. From the Mind's Eye of the User: the Sense-Making Qualitative-Quantitative Methodology. *Sense-Making Methodology Reader*, 1992.

[15] J. A. Dijksman and S. Khan. Khan Academy: the World's Free Virtual School. In *APS Meeting Abstracts*, 2011.

[16] R. Du, D. Li, and A. Varshney. Geollery: a Mixed Reality Social Media Platform. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 685. ACM, 2019. doi: 10.1145/3290605.3300915

[17] R. Du and A. Varshney. Social Street View: Blending Immersive Street Views With Geo-Tagged Social Media. In *Proceedings of the 21st International Conference on Web3D Technology*, Web3D, pp. 77–85. ACM, Jul. 2016. doi: 10.1145/2945292.2945294

[18] C. Elvezio, M. Sukan, O. Oda, S. Feiner, and B. Tversky. Remote Collaboration in AR and VR Using Virtual Replicas. In *ACM SIGGRAPH 2017 VR Village*, pp. 1–2. 2017. doi: 10.1145/3089269.3089281

[19] J. Franz, M. Alnusayri, J. Malloch, and D. Reilly. A Comparative Evaluation of Techniques for Sharing AR Experiences in Museums. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–20, 2019. doi: 10.1080/19443994.2013.787953

[20] O. Fried, A. Tewari, M. Zollhöfer, A. Finkelstein, E. Shechtman, D. B. Goldman, K. Genova, Z. Jin, C. Theobalt, and M. Agrawala. Text-Based Editing of Talking-Head Video. *ACM Transactions on Graphics (TOG)*, 38(4):1–14, 2019. doi: 10.1145/3306346.3323028

[21] A. M. Genest, C. Gutwin, A. Tang, M. Kalyn, and Z. Ivkovic. KinectArms: a Toolkit for Capturing and Displaying Arm Embodiments in Distributed Tabletop Groupware. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, pp. 157–166. ACM, 2013. doi: 10.1145/2441776.2441796

[22] Google. AutoDraw Is a New Kind of Drawing Tool That Pairs the Magic of Machine Learning With Drawings From Talented Artists to Help Everyone Create Anything Visual, Fast. `https://www.autodraw.com/`, 2019.

[23] J. G. Grandi, H. G. Debarba, I. Bemdt, L. Nedel, and A. Maciel. Design and Assessment of a Collaborative 3D Interaction Technique for Handheld Augmented Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 49–56. IEEE, 2018. doi: 10.1109/VR.2018.8446295

[24] J. G. Grandi, H. G. Debarba, and A. Maciel. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 127–135. IEEE, 2019. doi: 10.1109/VR.2019.8798080

[25] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio. 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*, pp. 4021–4033. ACM, 2017. doi: 10.1145/3025453.3025683

[26] R. Gumienny, L. Gericke, M. Quasthoff, C. Willems, and C. Meinel. Tele-Board: Enabling Efficient Collaboration in Digital Design Spaces. In *Proceedings of the 2011 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 47–54. IEEE, 2011. doi: 10.1109/CSCWD.2011.5960054

[27] C. Gutwin and S. Greenberg. Workspace Awareness for Groupware. In *Conference Companion on Human Factors in Computing Systems*, pp. 208–209. ACM, 1996. doi: 10.1145/257089.257284

[28] C. Gutwin and S. Greenberg. Design for Individuals, Design for Groups: Tradeoffs Between Power and Workspace Awareness. 1998. doi: 10.1145/289444.289495

[29] C. Gutwin and S. Greenberg. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):411–446, 2002. doi: 10.1023/A:1021271517844

[30] D. Harley, A. Verni, M. Willis, A. Ng, L. Bozzo, and A. Mazalek. Sensory VR: Smelling, Touching, and Eating Virtual Reality. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 386–397, 2018. doi: 10.1145/3173225.3173241

[31] Z. He and K. Perlin. Exploring the Effectiveness of Face-to-face Mixed Reality for Teaching with Chalktalk. *ArXiv Preprint ArXiv:1912.03863*, 2018.

[32] Z. He, K. T. Rosenberg, and K. Perlin. Exploring configuration of mixed reality spaces for communication. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2019. doi: 10.1145/3290607.3312761

[33] Z. He, F. Zhu, and K. Perlin. Physhare: Sharing physical interaction in virtual reality. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 17–19, 2017. doi: 10.1145/3131785.3131795

[34] Z. He, F. Zhu, K. Perlin, and X. Ma. Manifest the invisible: Design for situational awareness of physical environments in virtual reality. *Arxiv Preprint Arxiv:1809.05837*, 2018.

[35] V. Heun, J. Hobin, and P. Maes. Reality Editor: Programming Smarter Objects. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, pp. 307–310, 2013. doi: 10.1145/2494091.2494185

[36] T.-W. Hsu, M.-H. Tsai, S. V. Babu, P.-H. Hsu, H.-M. Chang, W.-C. Lin, and J.-H. Chuang. Design and Initial Evaluation of a VR Based Immer-

sive and Interactive Architectural Design Discussion System. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 363–371. IEEE, 2020. doi: 10.1109/VR46266.2020.1581231362069

[37] K. Huo, T. Wang, L. Paredes, A. M. Villanueva, Y. Cao, and K. Ramani. SynchronizAR: Instant Synchronization for Spontaneous and Spatial Collaborations in Augmented Reality. In *The 31st Annual ACM Symposium on User Interface Software and Technology*, pp. 19–30. ACM, 2018. doi: 10.1145/3242587.3242595

[38] A. Irlitti, T. Piumsomboon, D. Jackson, and B. H. Thomas. Conveying Spatial Awareness Cues in XR Collaborations. *IEEE Transactions on Visualization and Computer Graphics*, 25(11):3178–3189, 2019. doi: 10.1109/TVCG.2019.2932173

[39] P. Isenberg, D. Fisher, M. R. Morris, K. Inkpen, and M. Czerwinski. Co-located Collaborative Visual Analytics around a Tabletop Display. In *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, pp. 689–702. IEEE, 2012. doi: 10.1109/TVCG.2011.287

[40] H. Ishii and M. Kobayashi. ClearBoard: a Seamless Medium for Shared Drawing and Conversation With Eye Contact. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 525–532. ACM, 1992. doi: 10.1145/142750.142977

[41] K. Kim, M. Billinghurst, G. Bruder, H. B.-L. Duh, and G. F. Welch. Revisiting Trends in Augmented Reality Research: a Review of the 2nd Decade of ISMAR (2008-2017). *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2947–2962, 2018. doi: 10.1109/TVCG.2018.2868591

[42] S. Kim, G. Lee, W. Huang, H. Kim, W. Woo, and M. Billinghurst. Evaluating the combination of visual communication cues for hmd-based mixed reality remote collaboration. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–13, 2019. doi: 10.1145/3290605.3300403

[43] K. Kiyokawa, M. Billinghurst, S. E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. Communication Behaviors of Co-located Users in Collaborative AR Interfaces. In *Proceedings. International Symposium on Mixed and Augmented Reality*, pp. 139–148. IEEE, 2002. doi: 10.1109/ISMAR.2002.1115083

[44] A. Kunert, T. Weissker, B. Froehlich, and A. Kulik. Multi-Window 3D Interaction for Collaborative Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 2019. doi: 10.1109/TVCG.2019.2914677

[45] M. A. Lab. Sketch 2 Code. Transform Any Hands-Drawn Design Into a HTML Code With AI. `https://sketch2code.azurewebsites.net`, 2019.

[46] D. Lakatos, M. Blackshaw, A. Olwal, Z. Barryte, K. Perlin, and H. Ishii. 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*, pp. 90–93. ACM, 2014. doi: 10.1145/2659766.2659785

[47] J. Landgren and U. Nulden. A Study of Emergency Response Work: Patterns of Mobile Phone Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1323–1332. ACM, 2007. doi: 10.1145/1240624.1240824

[48] K. Layng, M. Gold, B. Ahlbrand, Z. He, and K. Perlin. The outpost. In *ACM SIGGRAPH 2020 Immersive Pavilion*, SIGGRAPH '20. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3388536.3407893

[49] G. Lee, H. Kang, J. Lee, and J. Han. A User Study on View-Sharing Techniques for One-to-Many Mixed Reality Collaborations. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 343–352. IEEE, 2020. doi: 10.1109/VR46266.2020.1581166222244

[50] G. A. Lee, T. Teo, S. Kim, and M. Billinghurst. A User Study on Mr Remote Collaboration Using Live 360 Video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 153–164. IEEE, 2018. doi: 10.1109/ISMAR.2018.00051

[51] J. Li, S. Greenberg, E. Sharlin, and J. Jorge. Interactive Two-Sided Transparent Displays: Designing for Collaboration. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, pp. 395–404. ACM, 2014. doi: 10.1145/2598510.2598518

[52] F. Lu, D. Yu, H.-N. Liang, W. Chen, K. Papangelis, and N. M. Ali. Evaluating Engagement Level and Analytical Support of Interactive Visualizations in Virtual Reality Environments. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 143–152. IEEE, 2018. doi: 10.1109/ISMAR.2018.00050

[53] S. Lukosch, M. Billinghurst, L. Alem, and K. Kiyokawa. Collaboration in Augmented Reality. *Computer Supported Cooperative Work (CSCW)*, 24(6):515–525, 2015. doi: 10.1007/s10606-015-9239-0

[54] T. Mahmood, W. Fulmer, N. Mungoli, J. Huang, and A. Lu. Improving Information Sharing and Collaborative Analysis for Remote GeoSpatial Visualization Using Mixed Reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 236–247. IEEE, 2019. doi: 10.1109/ISMAR.2019.00021

[55] M. R. Morris, J. Lombardo, and D. Wigdor. WeSearch: Supporting Collaborative Search and Sensemaking on a Tabletop Display. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, pp. 401–410. ACM, 2010. doi: 10.1145/1718918.1718987

[56] O. Oda, C. Elvezio, M. Sukan, S. Feiner, and B. Tversky. Virtual Replicas for Remote Assistance in Virtual and Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on UIST*, pp. 405–415. ACM, 2015. doi: 10.1145/2807442.2807497

[57] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou, et al. Holoportation: Virtual 3d Teleportation in Real-Time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 741–754. ACM, 2016. doi: 10.1145/2984511.2984517

[58] K. Otsuka. MMSpace: Kinetically-Augmented Telepresence for Small Group-to-Group Conversations. In *Virtual Reality (VR), 2016 IEEE*, pp. 19–28. IEEE, 2016. doi: 10.1109/VR.2016.7504684

[59] T. Park, M.-Y. Liu, T.-C. Wang, and J.-Y. Zhu. GauGAN: Semantic Image Synthesis With Spatially Adaptive Normalization. In *ACM SIGGRAPH 2019 Real-Time Live!*, pp. 1–1. 2019. doi: 10.1145/3306305.3332370

[60] S. A. Paul. Understanding Together: Sensemaking in Collaborative Information Seeking. 2009. doi: 10.1145/1718918.1718976

[61] S. A. Paul and M. R. Morris. CoSense: Enhancing Sensemaking for Collaborative Web Search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1771–1780, 2009. doi: 10.1145/1518701.1518974

[62] T. Pejsa, J. Kantor, H. Benko, E. Ofek, and A. Wilson. Room2room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pp. 1716–1725, 2016. doi: 10.1145/2818048.2819965

[63] K. Perlin, Z. He, and K. Rosenberg. Chalktalk: a Visualization and Communication Language-As a Tool in the Domain of Computer Science Education. *ArXiv Preprint ArXiv:1809.07166*, 2018.

[64] K. Perlin, Z. He, and F. Zhu. Chalktalk vr/ar. *International SERIES on Information Systems and Management in Creative eMedia (CreMedia)*, (2017/2):30–31, 2018.

[65] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6:5, 2019.

[66] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–13, 2018. doi: 10.1145/3173574.3173620

[67] I. Popovici and R.-D. Vatavu. Understanding Users' Preferences for Augmented Reality Television. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 269–278. IEEE, 2019. doi: 10.1109/ISMAR.2019.00024

[68] D. F. Reilly, H. Rouzati, A. Wu, J. Y. Hwang, J. Brudvik, and W. K. Edwards. TwinSpace: an Infrastructure for Cross-Reality Team Spaces. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, pp. 119–128. ACM, 2010. doi: 10.1145/1866029.1866050

[69] reklamistcomua. World Builder (virtual Reality). `https://www.youtube.com/watch?v=FheQe8rflWQ`, 2019.

[70] H. Salimian, S. Brooks, and D. Reilly. MP Remix: Relaxed WYSIWIS Immersive Interfaces for Mixed Presence Collaboration With 3D Content. *Proceedings of the ACM on Human-Computer Interaction*,

3(CSCW):1–22, 2019. doi: 10.1145/3359207

[71] K. A. Satriadi, B. Ens, M. Cordeil, B. Jenny, T. Czauderna, and W. Willett. Augmented Reality Map Navigation With Freehand Gestures. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 593–603. IEEE, 2019. doi: 10.1109/VR.2019.8798340

[72] T. C. Software. Miro Collaboration Without Constraints. `https://Miro.com/`, 2019.

[73] M. Sra, A. Mottelson, and P. Maes. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. In *Proceedings of the 2018 Designing Interactive Systems Conference*, DIS '18, pp. 85–97. ACM, New York, NY, USA, 2018. doi: 10.1145/3196709.3196788

[74] M. Sra and C. Schmandt. Metaspace: Full-Body Tracking for Immersive Multiperson Virtual Reality. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 47–48. ACM, 2015. doi: 10.1145/2815585.2817802

[75] K.-H. Tan, D. Gelb, R. Samadani, I. Robinson, B. Culbertson, and J. Apostolopoulos. Gaze Awareness and Interaction Support in Presentations. In *Proceedings of the 18th ACM International Conference on Multimedia*, MM '10, pp. 643–646. ACM, New York, NY, USA, 2010. doi: 10.1145/1873951.1874041

[76] A. Tang, M. Pahud, K. Inkpen, H. Benko, J. C. Tang, and B. Buxton. Three's Company: Understanding Communication Channels in Three-Way Distributed Collaboration. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, pp. 271–280. ACM, 2010. doi: 10.1145/1718918.1718969

[77] J. C. Tang and S. Minneman. VideoWhiteboard: Video Shadows to Support Remote Collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 315–322. ACM, 1991. doi: 10.1145/108844.108932

[78] J. C. Tang and S. L. Minneman. VideoDraw: a Video Interface for Collaborative Drawing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 313–320. ACM, 1990. doi: 10.1145/123078.128729

[79] T. Teo, L. Lawrence, G. A. Lee, M. Billinghurst, and M. Adcock. Mixed reality remote collaboration combining 360 video and 3d reconstruction. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–14, 2019. doi: 10.1145/3290605.3300431

[80] A. Tewari, O. Fried, J. Thies, V. Sitzmann, S. Lombardi, K. Sunkavalli, R. Martin-Brualla, T. Simon, J. Saragih, M. Nießner, R. Pandey, S. Fanello, G. Wetzstein, J.-Y. Zhu, C. Theobalt, M. Agrawala, E. Shechtman, D. B. Goldman, and M. Zollhöfer. State of the Art on Neural Rendering. *Eurographics*, 39(2), 2020.

[81] S. Thanyadit, P. Punpongsanon, and T.-C. Pong. Efficient Information Sharing Techniques Between Workers of Heterogeneous Tasks in 3d Cve. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–19, 2018. doi: 10.1145/3274441

[82] S. Thanyadit, P. Punpongsanon, and T.-C. Pong. ObserVAR: Visualization System for Observing Virtual Reality Users Using Augmented Reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 258–268. IEEE, 2019. doi: 10.1109/ISMAR.2019.00023

[83] B. Tversky, M. Suwa, M. Agrawala, P. Hanrahan, D. Phan, J. Klingner, M. Daniel, P. Lee, and J. Haymaker. Human Behavior in Design: Individuals, Teams, Tools. *Sketches for Design and Design of Sketches*, pp. 79–86, 2003. doi: 10.1007/978-3-662-07811-2

[84] K. Vogt, L. Bradel, C. Andrews, C. North, A. Endert, and D. Hutchings. Co-located Collaborative Sensemaking on a Large High-Resolution Display With Multiple Input Devices. In *IFIP Conference on Human-Computer Interaction*, pp. 589–604. Springer, Springer, 2011. doi: 10.1007/978-3-642-23771-3_44

[85] B. Volmer, J. Baumeister, S. Von Itzstein, I. Bornkessel-Schlesewsky, M. Schlesewsky, M. Billinghurst, and B. H. Thomas. A Comparison of Predictive Spatial Augmented Reality Cues for Procedural Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2846–2856, 2018. doi: 10.1109/TVCG.2018.2868587

[86] J. R. Wallace, S. D. Scott, T. Stutz, T. Enns, and K. Inkpen. Investigating Teamwork and Taskwork in Single-And Multi-Display Groupware Systems. *Personal and Ubiquitous Computing*, 13(8):569–581, 2009. doi: 10.1007/s00779-009-0241-8

[87] C. Weichel, M. Lau, D. Kim, N. Villar, and H. W. Gellersen. MixFab: a Mixed-Reality Environment for Personal Fabrication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3855–3864, 2014. doi: 10.1145/2556288.2557090

[88] T. Weissker, A. Kulik, and B. Froehlich. Multi-Ray Jumping: Comprehensible Group Navigation for Collocated Users in Immersive Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 136–144. IEEE, 2019. doi: 10.1109/VR.2019.8797807

[89] J. S. Yi, Y.-a. Kang, J. T. Stasko, and J. A. Jacko. Understanding and Characterizing Insights: How Do People Gain Insights Using Information Visualization? In *Proceedings of the 2008 Workshop on BEyond Time and Errors: Novel EvaLuation Methods for Information Visualization*, p. 4. ACM, 2008. doi: 10.1145/1377966.1377971

[90] J. Young, T. Langlotz, M. Cook, S. Mills, and H. Regenbrecht. Immersive Telepresence and Remote Collaboration Using Mobile and Wearable Devices. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1908–1918, 2019. doi: 10.1109/TVCG.2019.2898737