Preprint

Investigating Tangible Collaboration for Design Towards Augmented Physical Telepresence

Alexa F. Siu, Shenli Yuan, Hieu Pham, Eric Gonzalez, Lawrence H. Kim, Mathieu Le Goc, and Sean Follmer

Abstract While many systems have been designed to support collaboration around visual thinking tools, much less work has investigated how to share and collaboratively design physical prototypes—an important part of the design process. We describe preliminary results from a formative study on how designers communicate and collaborate in design meetings around physical and digital artifacts. Addressing some limitations in current collaboration platforms and drawing guidelines from our study, we introduce a new prototype platform for remote collaboration. This platform leverages the use of augmented reality (AR) for rendering of the remote participant and a pair of linked actuated tabletop tangible interfaces that acts as the participant's shared physical workspace. We propose the use of actuated tabletop tangibles to synchronously render complex shapes and to act as physical input.

1 Introduction

The need to collaborate with remote partners to accomplish joint tasks has risen over the years. However, most current technology limits individuals to passively watch video feeds which do not allow for any interaction with the remote physical environment. In particular, they do not support collaborative physical tasks in which two or more individuals work together to perform actions on 3-Dimensional (3D) objects in the world.

For designers, sharing physical objects is particularly important. Design collaboration involves talk, gestures, and the joint creation of design ideas through representations, which include not only sketches but also physical prototypes. Physical artifacts produced throughout the design process serve not only for functional testing but also for communication purposes.

They are helpful in establishing a common ground among those involved in the design and in externalizing or supporting what a designer is relaying verbally to the others (Fussell et al. 2004; Clark and Brennan 1991; Buxton 2009).

A.F. Siu (✉) • S. Yuan • H. Pham • E. Gonzalez • L.H. Kim • M. Le Goc • S. Follmer
Stanford University, Stanford, CA, USA
e-mail: afsiu@stanford.edu

As effective design collaboration involves sharing, exploring, referencing, and manipulating the physical environment, tools for remote collaboration should provide support for these interactions. Buxton (2009) defined three spaces to be considered at the microlevel of collaboration: person space (where one sees the remote person's face, expressions and voice), task space (where the work appears), and reference space (where body language and gesturing can be inferred). Most standard video systems support person space through video feed. Task space is usually abstracted from the physical workspace and only exists in the digital form e.g. collaborative online whiteboard apps. Lastly, reference space is mostly unsupported.

In this work, we address the aforementioned limitations leveraging the use of augmented reality (AR) for rendering of the remote participant and a pair of linked actuated tabletop tangible interfaces for rendering the participant's shared physical workspace or task space. For our implementation, we use Zooids, a tabletop swarm interface with many robotic tangibles.

Our telepresence system aims to brings together person space, task space, and reference space (Buxton 2009) with the hope of increasing the sense of co-presence and the ability of participants to communicate naturally using gaze, gesture, posture, and other nonverbal cues. We hope to contribute not only to the understanding of how physical prototypes are used as part of the design thinking process, but also to enable these benefits over distance.

2 Related Work

Remote collaboration is for most people primarily associated with video chat tools with screen-sharing capabilities (e.g., Microsoft Skype, Google Hangouts, Apple FaceTime, and Cisco Webex). Shared workspaces have, however, long been explored for video and audio across locations. Researchers have investigates collaboration for a wide variety of tasks, such as sharing physical documents using video feeds on a screen, as in TeamWorkStation (Ishii 1990), or projected, as in Video Draw (Tang and Minneman 1991c). Video Whiteboard explored wall-scale shared workspaces, while providing feedback on remote user's presence (Tang and Minneman 1991a). Clearboard went further by allowing for proper gaze estimation of remote users (Ishii and Kobayashi 1992). More recently, these techniques have been applied to applications, such as collaborative website development (Everitt et al. 2003a), remote board games (Wilson and Robbins 2007) and family communications (Yarosh et al. 2013).

2.1 Gestures in Video-Mediated Communication

Past research has shown that co-located partners work more efficiently than distributed partners due to participant's ability to use gestures and support their conversation (Kraut et al. 2002; Fussell et al. 2004). Several projects investigate video-mediated collaboration using surrogates for hand gestures (Tang and Minneman 1991b; Fussell et al. 2004) while others have used video streams and computer vision to register hand gestures and transmit them (Coldefy and Louis-dit Picard 2007; Wood et al. 2016). In our work, we focus on showing user's gestures with low-fidelity but in their 3D spatial context.

2.2 Mixed Reality

As Virtual Reality (VR) and Augmented Reality (AR) have matured, collaborative tools have been extended that can support more than video mediated interaction, drawing on a variety of different display and interaction techniques. Raskar presented a vision of what future collaborative workspaces could look like with AR technology (Raskar et al. 1998). This has also been explored in the context of CAVE based VR, where two CAVES can be linked allowing for VR collaboration (Gross et al. 2003). MirageTable utilized projected AR to give users a view corrected steroview of a collaborative 3D workspace and remote collaborator (Benko et al. 2012).

Beyond head-worn and projected AR, other display technologies have the opportunity to give remote users strong local presence, such as TeleHuman's Cylinder display (Kim et al. 2012) or BeThere (Sodhi et al. 2013), which focuses on handheld AR where remote users can point into a remote 3D scene.

Some of these systems also allow for spatial annotations in the world (Fakourfar et al. 2016; Gauglitz et al. 2014). However, most of the work on this topic has not included physical feedback or tangible interaction.

2.3 Telerobotics

A vibrant research community is addressing engineering and design challenges in representing remote people using telepresence robots (e.g., Tsui et al. 2011; Paulos and Canny 1998). The Personal Roving Presence concept (Paulos and Canny 1998), in particular, was an early exploration into tele-embodiment with different implementations using both screen-based robots and flying blimps. Much research has also explored how these devices can be used in the workplace, and how they influence user's sense of social presence (Lee and Takayama 2011). Our interactions focus on robotics to provide a shared workspace, with less emphasis on the mobility provided by robots for telepresence.

2.4 Tangible Remote Collaboration

Remote collaboration has also been explored through Tangible User Interfaces (TUIs) to address the aforementioned video streams limitations. Typically TUIs have been used for the manipulation of remote physical environments (Brave et al. 1998; Pangaro et al. 2002; Richter et al. 2007; Riedenklau et al. 2012; Leithinger et al. 2014) and the display of information (Everitt et al. 2003b; Ullmer and Ishii 1997; Leithinger et al. 2014). Recently, tabletop swarm user interfaces such as Zooids (Le Goc et al. 2016) have emerged as a promising collaboration platform. They can be used as an information display as well as a type of input/output interface where some agents act as controls or handles while others act as outputs. Few projects, however, have combined AR with remote tangible interaction.

3 User Evaluation

In face-to-face collaboration on physical tasks, people readily combine speech and gesture to support their communication. They can monitor one another's hands and jointly observe task objects and the environment. Although studies of the use of physical artifacts in design exist, less work has explored their specific use in co-located collaboration for design. We conducted a small-scale user study to understand how designers communicate and collaborate in co-located design meetings around representations. We want to understand differences in interaction when the representation takes different forms, specifically as physical and digital 3D models. From these results we wish to draw guidelines for creating physical interfaces for design collaboration which bring together aspects of both physical and digital manipulation.

3.1 Design

Five pairs of graduate student served as participants (ages 22–27, one female). No previous design experience was required. They each received $15 for their participation. Each pair of users completed a task under two different conditions. The task was to sketch a completed house model according to a set of given design goals.

The participants had asymmetrical roles: one was the instructor and the other was the sketcher. The instructor had the design goals as sketches of a completed house model, and his/her goal was to communicate this model to the sketcher. The sketcher on the other hand, had incomplete sketches of the same house and his/her goal was to complete the house model sketch based on the instructor's communication. Participants were seated facing each other but there was a physical separation on the table such that they could not show each other their respective sketches.

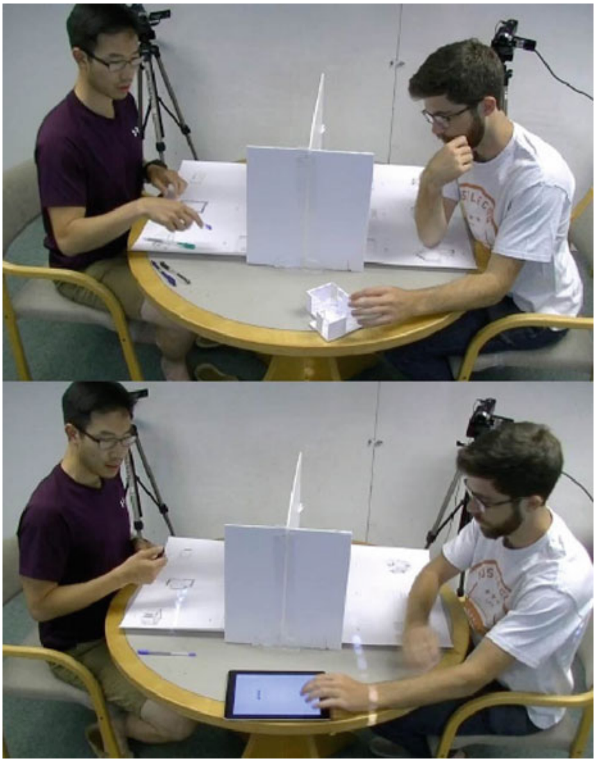

Fig. 1 User study setup. Left participant is the sketcher with the incomplete house model. Right participant is the instructor with the complete house model. On the *top figure*, participants share a physical model of the house. On the *bottom figure*, participants share a digital model of the house

To support communication, the users shared a representation of the incomplete house model. In one condition, the representation was a 3D physical model and in the other condition, the representation was a 3D digital model displayed on an iPad. Users had the freedom to rotate and change views of the model at any time. The complete setup is shown in Fig. 1.

3.2 Materials

Two different two-story house models (with equivalent number of components) were created using the Arckit[1] architectural modeling system. Digital models of the each house were made on Sketchup using the Arckit Digital library. The shared

[1]https://www.http://arckit.com/.

physical model was the incomplete house made with Arckit components, while the shared digital model was viewed in the Sketchup app for iPad.

Instructions for the instructor had all views (front, back, left, right, top, iso) of the completed house model and were printed on a $24'' \times 16''$ sheet of paper. The sketcher received a similar printed sheet with the same house model views (front, back, left, right, top, iso) but with only the first floor present.

Three online surveys were administered through Qualtrics. The first was a pretest questionnaire to collect demographic information (e.g. gender, age, field of study). A post-task questionnaire administered after each task, asked questions about the success of each collaboration. There were free response questions as well as responses made on a 5-point scale ranging from 1 (Very Hard) to 5 (Not very hard). Certain questions were reworded for the Sketcher roles (e.g. Is there anything your partner could have done differently in communicating the instructions?).

An exit interview was conducted upon completion of all tasks. Both instructor and sketcher were present during the same interview. Five questions were asked during the interview, relating to the overall experience as well as role-specific experience.

3.3 Procedure

Participants were given an overview of the task, were assigned roles, and given specific goals for each role. They then completed consent forms and the pretest questionnaire. They were also introduced to the different components that could make and Arckit house model (e.g. wall, window, ceiling). They were shown how they could interact with the Sketchup app for viewing digital models (e.g. pinch for zoom, drag to rotate).

Each task was timed for 20 min and participants were given a 5-min warning. Upon completion, each would complete a post-task questionnaire. They then moved on to the next task with the same time lapse. After both tasks and post-task questionnaires were done, the exit interview was conducted.

3.4 Preliminary Findings

Video and audio from the experiment was analyzed and encoded for gestures. A gesture coding scheme based on that from Fussell et al. (2004) was used and is summarized bellow:

Gesture coding scheme:

- **Exploratory**: Moving and/or rotating the representation
- **Iconic**: Forming hands to show what a piece looks like, or to show how two pieces should be positioned relative to one another.

- **Pointing**: Orienting a finger or hand toward a point in the environment. Reference to objects and locations.
- **Spatial**: Indicating through use of one or both hands how far apart two objects should be or how far to move a given object.

One of our task performance metrics was percentage of task completion. This was determined by analyzing the sketches participants made for each task. To compute percentage of task completion, we counted the number of correctly sketched parts and divided by the total number of components we expected participants to have sketched to achieve the full house model. Percentage of task completion was higher in the iPad condition ($\mu = 0.71$, $\sigma = 0.3$) than in the physical condition ($\mu = 0.59$, $\sigma = 0.26$).

One of the survey questions asked users to rate (on a five-point scale) how difficult they found the task and how much mental effort they had to put in (Fig. 2). For both conditions, sketchers reported having to put more effort when compared to instructors. Instructors rated their effort similarly for both conditions. Other questions asked participants to rate how difficult they found the collaboration and how confident they felt they understood correctly (Figs. 3 and 4). Similar to the previous, sketchers reported more difficulty and lower confidence for the digital condition. Instructors rated collaborating with the digital slightly more difficult but rated their confidence level similarly for both conditions.

From participants' comments, sketchers had a much stronger preference for the physical representation. Instructors and sketchers both commented on the ease of manipulation of the physical model. "Being able to touch the physical representation and rotating it to my will was very helpful." "It was much easier to work with than

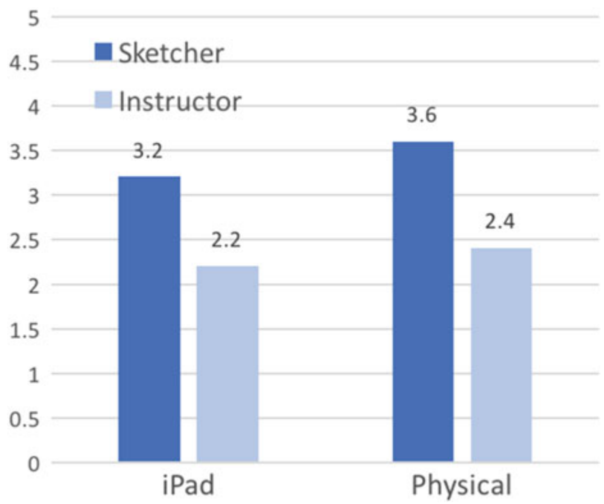

Fig. 2 How hard did you have to work (mentally) to accomplish your level of performance? 1 (Very Hard) to 5 (Not Very Hard)

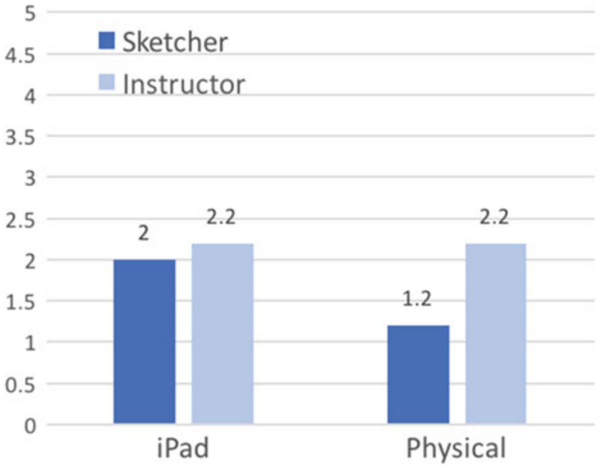

Fig. 3 How confident are you that you correctly understood your partner's design instructions (*or that your partner understood your instructions*)? 1 (Extremely Confident) to 5 (Extremely Insecure)

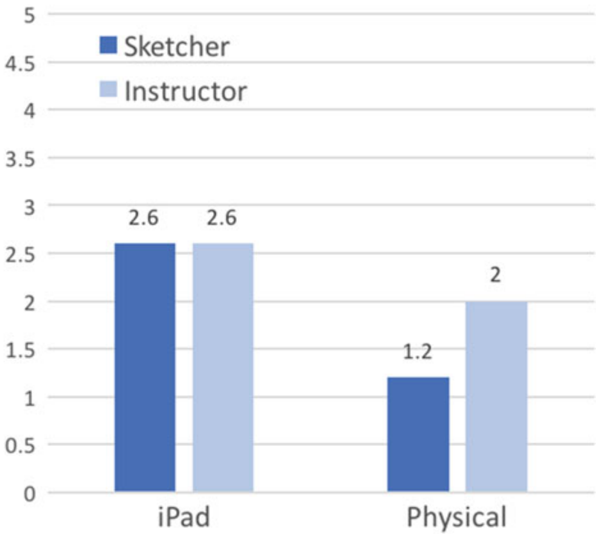

Fig. 4 How difficult was it collaborating with your partner on the design activity? 1 (Extremely Easy) to 5 (Extremely Difficult)

with the iPad for sure. iPad was just replicating what was on paper but 3D model was easier to visualize and communicate." Sketchers also highlighted how the digital made it more difficult for both to interact with it at the same time which may explain why instructors dominated most of the interaction with the representation. "There was one time where we both wanted to point something out on the iPad. There was also times where he wanted information from a certain perspective and I could only describe it to him from another perspective so that made things more difficult."

On the other hand, instructors highlighted how in sharing one perspective, the digital had its advantages. "...the iPad was much more helpful because you can do more with little effort; like zooming, panning and rotating by just moving your fingers and not your whole body." With the digital representation, both participants share one view of the screen and when pointing to a specific part, for example, it is clear what one is pointing at. With the physical, since users are seated facing each other, each has a different view of the representation and when pointing to a specific part, more verbal explanation may be required to clarify what one is referring to. We often found that once instructors had explained a new part of the model for sketching, the sketcher would often repeat the instruction to confirm his/her understanding.

3.5 Implications for Design

Instructors found the iPad to be easier for explaining since it did not require them to hold it in place to show a specific view. The instructor could use both of his/her hands for explaining and the sketcher would still have the same perspective view. However, the sketcher felt more confident in the physical model, noting how it was easier to find an orientation and manipulate the view they wanted. This difference highlights both an advantage and disadvantage of the physical model and directly relates to object orientation. Within the scope of remote collaboration, this could be addressed by providing flexible, user-controlled orientation of the shared workspace. If the remote user is rendered digitally, then physically, there will not be any interference from the local user such that both participant's hands could even overlap on the same physical space. The local user is therefore free to orient the workspace in any way.

Another implication we can draw from this is that there should be support for simultaneous actions from the participants involved. While the one perspective-view from the iPad is beneficial in providing a common ground, it limits simultaneous access to different views of the model. With the physical object, this is not an issue since the user can simply move around if he/she wants another view of the object without interrupting the other user.

Users found the physical model more engaging and easier to manipulate. The more-accessible manipulation allowed them to better understand the model. If both physical and digital representations are used in the workspace, it must enable easy and responsive input for manipulating the object's orientation.

Both sketchers and instructors said having the ability to manipulate and annotate the model would have made the task easier. Adding annotations or changes to the shared representation could be useful in complementing what the users want to convey.

4 Mixed Reality + Remote Collaboration Platform

Based on our user study findings, an interface for physical remote collaboration should be capable of (1) physical input to the remote participant's physical workspace, and (2) representation of the remote participant spatially linked to the shared workspace. In such an interface, designers could easily discuss, annotate, and physically modify prototypes (e.g., of cars or buildings) with remote collaborators all over the world.

Towards this goal, we introduce a low resolution prototype platform (Fig. 5) that uses a head-mounted AR display and many actuated tabletop tangibles, known as Zooids (Le Goc et al. 2016). The AR display is used to spatially render the remote user in the real world. We leverage the use of Zooids as both input and output controls that all together can render complex shapes as well as make changes to the physical space.

We introduce a prototype telepresence platform for remote collaboration that allows physical input/output and also spatial rendering of the remote user in the real world. We leverage the use of many actuated tangibles as both input and output controls that together can render complex shapes as well as make changes to the physical space.

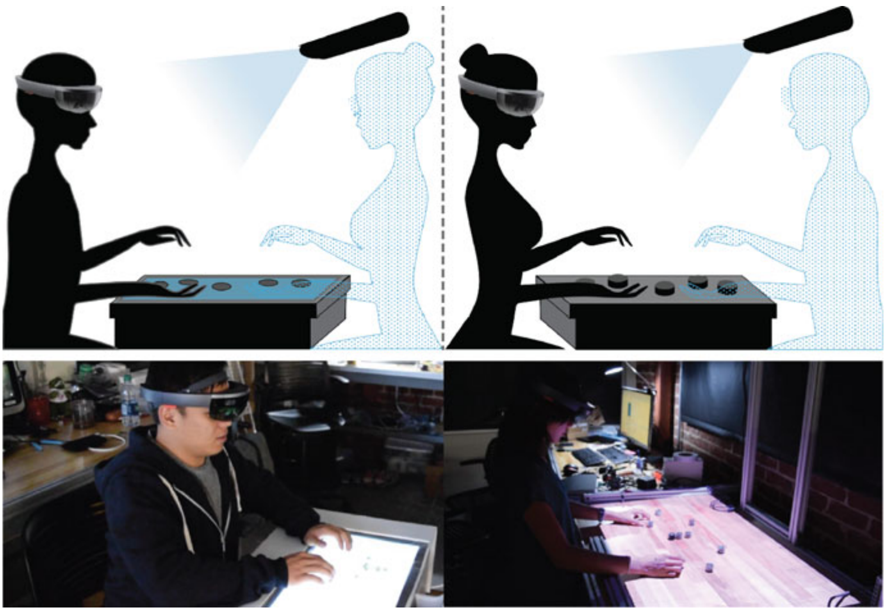

Fig. 5 System diagram (*top*) and actual setup (*bottom*) of the remote collaboration platform. The *left* shows the user with the tablet swarm interface while the *right* shows the user with the physical swarm platform. A depth sensor captures the local user and the Hololens renders the remote participant, seen as a *blue point cloud*

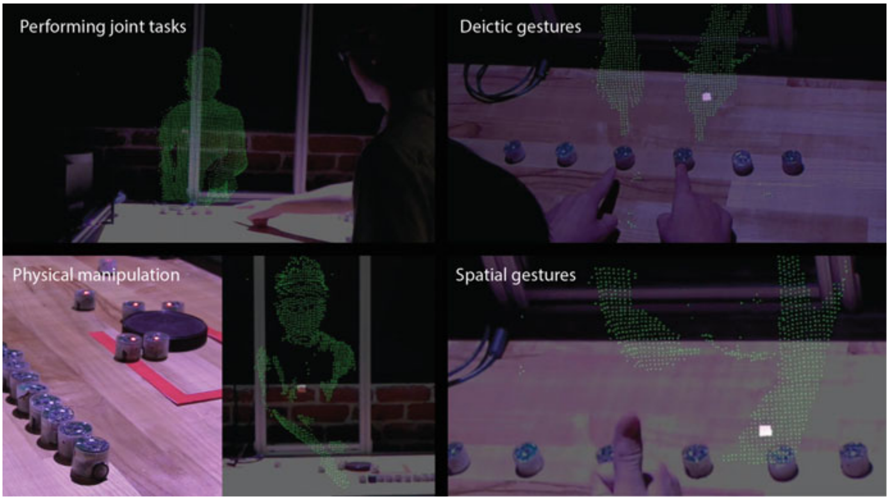

Fig. 6 Example scenarios of our remote telepresence platform

4.1 Preliminary Exploration

We show usage of the remote telepresence platform in four different scenarios (Fig. 6).

Performing joint tasks: In one case, both users jointly move a Zooid along a path. One user moves the physical Zooid, while the other user drags the Zooid on the tablet. This requires coordination and clear spatial understanding of where the remote user's hand is moving, such that the local user can closely follow.

Manipulating the physical environment remotely: In another case, the remote participant manipulates a group of Zooids to insert a disc into a slot. This shows a scenario where the platform enables users to make changes on the remote physical space.

Rendering shapes: Users are also able to jointly draw and modify shapes using the Zooids. This is a simpler analogy to the design collaboration scenario where two designers may be discussing and modifying a model.

Conveying meaning through gestures: Users are able to easily understand commonly used gestures for referencing objects and conveying spatial relationships, quantities, and shapes.

4.2 Implementation

Microsoft Kinect sensors are used to capture the geometry and appearance of the environment of the remote client. The captured Kinect frame was clipped to encompass a tight region where only the person is present and the information was then downsampled to under 60,000 points. Background subtraction is performed after clipping and down sampling based on a weighted sum from an initial calibration that came from taking a static image of the background and with the body tracking data available from the Kinect default libraries. This information is sent to the local client using connectionless network sockets following the standard User Datagram Protocol (UDP).

The local client consisted of a Hololens which received the remote client's data and used it to reconstruct the 3D geometry and appearance of the remote participant locally, thus obtaining the virtual copy. The data was rendered as a one-color 3D point cloud. This point cloud started as a simple flat mesh with zero depth at each vertex. At each frame rate, a shader updated the depth value of each of the point cloud's vertex based on the information received. The analogous procedure was applied on the remote client to obtain and render a virtual copy of the local participant. Thus the real person in each room was able to see and hear the virtual copy of the person from the other room in real time.

The shared workspace between the participants consists of the swarm user interface, Zooids (Le Goc et al. 2016). One participant's workspace includes the physical Zooids while the other participant accesses the Zooids through a tablet application (Fig. 5). The physical and digital Zooids are linked such that movement of Zooids on the local side is reflected on the remote side and vice versa.

5 Limitations and Future Work

Limited resolution and fidelity of the remote participant limits the amount of content that can be virtually rendered. Changing to a color point cloud and incorporating surface normals would result in a more realistic representation of the remote user.

Latency was not strictly investigated. However, there is some inherent latency in the robots physically moving from one place to another, i.e. "refreshing" the workspace. This could place limits in the kind of interactions possible. In addition, network latency was not strictly investigated. This also results in communication delays between client and server, and the graphics rendering pipeline for creating the virtual copy of the participant through the Hololens. Minimal latency is required for real-time interaction.

The scale and number of Zooids in our current system also limits the type of interaction and applications that can be created. With many more, more compelling shapes could be rendered. To better support the remote collaboration side, it will also be important to develop a number of interaction techniques around annotating and manipulating the shared workspace. This could be implemented using the same AR display.

AR resolution and field of view present issues for the system's use. Our current setup uses a one-color point cloud which would not allow the user to show distinguishable objects. Moreover, there is no support for facial expressions which are important in communication. Changing to a color point cloud and incorporating surface normals would result in a more realistic representation of the remote user. Depth perception could also be improved by adding shading and depth-of-field blur. The field of view of the Hololens is also limited, which causes issues for users trying to look both at the Zooids on the table as well as the face of the remote participant.

In future work, we would also like to explore the use of linked shape displays for displaying shared content. Unlike the Zooids platform, these types of input/output devices are not as limited in their vertical displacement. They are also capable of being smaller and having higher resolution (Poupyrev et al. 2007; Follmer et al. 2013).

However, shape displays have limited degrees of freedom of input; users can only push down and pull up on the pins. In addition, the display is continuous, and each pin is a single rigid object. This means that users cannot pick up or physically move parts rendered by the shape display. This limits the degrees of freedom and the expressiveness of input. We address this to some extent with the control of tangible objects, but future systems could also take inspiration from Modular Robotics to use small robotic blocks that could be snapped together and removed from the table (Gilpin et al. 2007). Beyond this even smaller particles, or fluids, could be imagined to give a full sense of shared objects.

6 Conclusion

We have provided preliminary design guidelines for the design of tangible collaboration interfaces. Based on these guidelines, we also introduced a low resolution prototype using actuated tangibles and head-mounted AR displays. Our preliminary findings show benefits from both physical and digital artifacts in design collaboration. This suggests a need to better understand how future work can harness the best of both the physical and digital world in designing platforms that support remote collaboration.

Acknowledgements This work is supported in part by the NSF GRFP, Stanford School of Engineering Fellowship, Hasso Plattner Design Thinking Research Program, and HP Inc.

References

Benko, H., Jota, R., & Wilson, A. (2012). MirageTable: Freehand interaction on a projected augmented reality tabletop. In *ACM CHI '12* (pp. 199–208). http://dl.acm.org/citation.cfm?id=2207676.2207704

Brave, S., Ishii, H., & Dahley, A. (1998). Tangible interfaces for remote collaboration and communication. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work* (pp. 169–178). New York: ACM.

Buxton, B. (2009). Mediaspace – meaningspace – meetingspace. In *Media space 20 + years of mediated life* (pp. 217–231). London: Springer.
Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 13–1991). Washington, DC: American Psychological Association.
Coldefy, F., & Louis-dit Picard, S. (2007). Digitable: An interactive multiuser table for collocated and remote collaboration enabling remote gesture visualization. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). New York: IEEE.
Everitt, K. M., Klemmer, S. R., Lee, R., & Landay, J. A. (2003a). Two worlds apart: Bridging the gap between physical and virtual media for distributed design collaboration. In *ACM CHI '03* (pp. 553–560). http://dl.acm.org/citation.cfm?id=642611.642707
Everitt, K. M., Klemmer, S. R., Lee, R., & Landay, J. A. (2003b). Two worlds apart: Bridging the gap between physical and virtual media for distributed design collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 553–560). New York: ACM.
Fakourfar, O., Ta, K., Tang, R., Bateman, S., & Tang, A. (2016). Stabilized annotations for mobile remote assistance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1548–1560). New York: ACM.
Follmer, S., Leithinger, D., Olwal, A., Hogge, A., & Ishii, H. (2013). Inform: Dynamic physical affordances and constraints through shape and object actuation. In *UIST* (Vol. 13, pp. 417–426).
Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E., & Kramer, A. D. I. (2004). Gestures over video streams to support remote collaboration on physical tasks. *Human–Computer Interaction, 19*(3), 273–309.
Gauglitz, S., Nuernberger, B., Turk, M., & Höllerer, T. (2014). World-stabilized annotations and virtual scene navigation for remote collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (pp. 449–459). New York: ACM.
Gilpin, K., Kotay, K., & Rus, D. (2007). Miche: Modular shape formation by self-disassembly. In *IEEE International Conference on Robotics and Automation 2007* (pp. 2241–2247). New York: IEEE. http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=4209417
Gross, M., Lang, S., Strehlke, K., Moere, A. V., Staadt, O., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., & Van Gool, L. (2003). blue-c: A spatially immersive display and 3D video portal for telepresence. *ACM Transactions on Graphics, 22*(3), 819–827. http://dl.acm.org/citation.cfm?id=882262.882350
Ishii, H. (1990). TeamWorkStation: Towards a seamless shared workspace. In *ACM CSCW '90* (pp. 13–26). http://dl.acm.org/citation.cfm?id=99332.99337
Ishii, H., & Kobayashi, M. (1992). ClearBoard: A seamless medium for shared drawing and conversation with eye contact. In *ACM CHI '92* (pp. 525–532). http://dl.acm.org/citation.cfm?id=142750.142977
Kim, K., Bolton, J., Girouard, A., Cooperstock, J., & Vertegaal, R. (2012). TeleHuman: Effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod. In *ACM CHI '12* (p. 2531). http://dl.acm.org/citation.cfm?id=2207676.2208640
Kraut, R. E., Gergle, D., & Fussell, S. R. (2002). The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (pp. 31–40). New York: ACM.
Le Goc, M., Kim, L. H., Parsaei, A., Fekete, J. D., Dragicevic, P., & Follmer, S. (2016). Zooids: Building blocks for swarm user interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (pp. 97–109). New York: ACM.
Lee, M. K., & Takayama, L. (2011). "Now, i have a body": Uses and social norms for mobile remote presence in the workplace. In *ACM CHI '11* (pp. 33–42). http://dl.acm.org/citation.cfm?id=1978942.1978950
Leithinger, D., Follmer, S., Olwal, A., & Ishii, H. (2014). Physical telepresence: Shape capture and display for embodied, computer-mediated remote collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (pp. 461–470). New York: ACM.

Pangaro, G., Maynes-Aminzade, D., & Ishii, H. (2002). The actuated workbench: Computer-controlled actuation in tabletop tangible interfaces. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology* (pp. 181–190). New York: ACM.

Paulos, E., & Canny, J. (1998). PRoP: Personal roving presence. In *ACM CHI '98* (pp. 296–303). http://dl.acm.org/citation.cfm?id=274644.274686

Poupyrev, I., Nashida, T., & Okabe, M. (2007). Actuation and tangible user interfaces: The Vaucanson duck, robots, and shape displays. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (pp. 205–212). New York: ACM.

Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., & Fuchs, H. (1998). The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *ACM SIGGRAPH '98* (pp. 179–188). http://dl.acm.org/citation.cfm?id=280814.280861

Richter, J., Thomas, B. H., Sugimoto, M., & Inami, M. (2007). Remote active tangible interactions. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (pp. 39–42). New York: ACM.

Riedenklau, E., Hermann, T., & Ritter, H. (2012). An integrated multi-modal actuated tangible user interface for distributed collaborative planning. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* (pp. 169–174). New York: ACM.

Sodhi, R. S., Jones, B. R., Forsyth, D., Bailey, B. P., & Maciocci, G. (2013). BeThere: 3D mobile collaboration with spatial input. In *ACM CHI '13* (pp. 179–188). http://dl.acm.org/citation.cfm?id=2470654.2470679

Tang, J. C., & Minneman, S. (1991a). VideoWhiteboard: Video shadows to support remote collaboration. In *ACM CHI '91* (pp. 315–322). http://dl.acm.org/citation.cfm?id=108844.108932

Tang, J. C., & Minneman, S. (1991b). Videowhiteboard: Video shadows to support remote collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 315–322). New York: ACM.

Tang, J. C., & Minneman, S. L. (1991c). Videodraw: A video interface for collaborative drawing. *ACM Transactions on Information Systems, 9*(2), 170–184. http://dl.acm.org/citation.cfm?id=123078.128729

Tsui, K. M., Desai, M., Yanco, H. A., & Uhlik, C. (2011). Exploring use cases for telepresence robots. In *ACM/IEEE HRI '11* (pp. 11–18). http://dl.acm.org/citation.cfm?id=1957656.1957664

Ullmer, B., & Ishii, H. (1997). The metadesk: Models and prototypes for tangible user interfaces. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology* (pp. 223–232). New York: ACM.

Wilson, A. D., & Robbins, D. C. (2007). Playtogether: Playing games across multiple interactive tabletops. In *IUI Workshop on Tangible Play: Research and Design for Tangible and Tabletop Games*.

Wood, E., Taylor, J., Fogarty, J., Fitzgibbon, A., & Shotton, J. (2016). Shadowhands: High-fidelity remote hand gesture visualization using a hand tracker. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces* (pp. 77–84). New York: ACM.

Yarosh, S., Tang, A., Mokashi, S., & Abowd, G. D. (2013). "almost touching": Parent-child remote communication using the sharetable system. In *ACM CSCW '13* (pp. 181–192). http://dl.acm.org/citation.cfm?id=2441776.2441798