

HoloBots: Augmenting Holographic Telepresence with Mobile Robots for Tangible Remote Collaboration in Mixed Reality

Keiichi Ihara

University of Tsukuba
Tsukuba, Japan
University of Calgary
Calgary, Canada
kihara@iplab.cs.tsukuba.ac.jp

Mehrad Faridan

University of Calgary
Calgary, Canada
mehrad.faridan1@ucalgary.ca

Ayumi Ichikawa

University of Tsukuba
Tsukuba, Japan
aichikawa@iplab.cs.tsukuba.ac.jp

Ikkaku Kawaguchi

University of Tsukuba
City, Japan
kawaguchi@cs.tsukuba.ac.jp

Ryo Suzuki

University of Calgary
Calgary, Canada
ryo.suzuki@ucalgary.ca

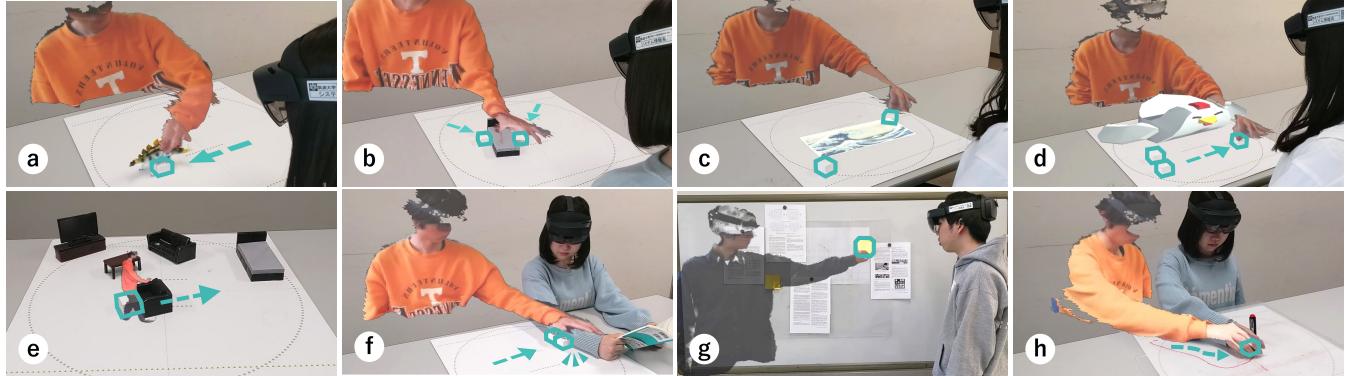


Figure 1: HoloBots explores augmenting holographic telepresence with tabletop mobile robots for remote collaboration. The remote user can interact with the local user through various methods, such as (a, b) actuating objects, (c, d) sharing tangible user interfaces, (e) representing the body, and (f) providing haptic feedback. By using attachments, HoloBots is adaptable in situations that involve (g) vertical surface and (h) drawing scenarios.

ABSTRACT

This paper introduces HoloBots, a mixed reality remote collaboration system that augments holographic telepresence with synchronized mobile robots. Beyond existing mixed reality telepresence, HoloBots lets remote users not only be visually and spatially present, but also *physically* engage with local users and their environment. HoloBots allows the users to touch, grasp, manipulate, and interact with the remote physical environment as if they were co-located in the same shared space. We achieve this by synchronizing holographic user motion (Hololens 2 and Azure Kinect) with tabletop mobile robots (Sony Toio). Beyond the existing physical

telepresence, HoloBots contributes to an exploration of broader design space, such as object actuation, virtual hand physicalization, world-in-miniature exploration, shared tangible interfaces, embodied guidance, and haptic communication. We evaluate our system with twelve participants by comparing it with hologram-only and robot-only conditions. Both quantitative and qualitative results confirm that our system significantly enhances the level of co-presence and shared experience, compared to the other conditions.

CCS CONCEPTS

- Human-centered computing → Mixed / augmented reality.

KEYWORDS

Mixed Reality; Remote Collaboration; Physical Telepresence; Mobile Robots; Actuated Tangible UI;

ACM Reference Format:

Keiichi Ihara, Mehrad Faridan, Ayumi Ichikawa, Ikkaku Kawaguchi, and Ryo Suzuki. 2023. HoloBots: Augmenting Holographic Telepresence with Mobile Robots for Tangible Remote Collaboration in Mixed Reality. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23, October 29–November 1, 2023, San Francisco, CA, USA)*. © 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 979-8-4007-0132-0/23/10...\$15.00 https://doi.org/10.1145/3586183.3606727

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UIST '23, October 29–November 1, 2023, San Francisco, CA, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0132-0/23/10...\$15.00

https://doi.org/10.1145/3586183.3606727

1 INTRODUCTION

Today's mixed reality telepresence still falls short of replicating the rich tangible experiences that we naturally enjoy in our everyday lives. In real-world collaboration, for example, we casually grasp and manipulate physical objects to facilitate discussions, employ touch for social interactions, spatially arrange physical notes for brainstorming, and guide others by holding their hands. However, such tangible interactions are not possible with current mixed reality telepresence systems, as the virtual remote user has no way to *physically* engage with the local user and environment.

In this paper, we introduce HoloBots, a mixed reality interface to achieve *tangible remote collaboration* by synchronously coupling holographic telepresence with an actuated physical environment. Beyond existing holographic telepresence like *Holoportation* [51], HoloBots lets remote users not only visually and spatially co-present but also physically touch, grasp, manipulate, and interact with remote tangible objects, as if they were co-located in the same shared space. We achieve this by synchronizing the remote user's motion rendered in a mixed reality headset (Hololens 2 and Azure Kinect) with physical actuation enabled by multiple tabletop mobile robots (Sony Toio).

Our idea builds up on the existing *physical telepresence* [40] and other related approaches [37, 68], but we make two key contributions beyond them. First, we explore a *broader design space* of tangible remote collaboration with holographic telepresence, which are not fully investigated in the prior work [19, 37, 68]. For example, we showcase various interactions, such as object actuation, virtual hand physicalization, world-in-miniature exploration, shared tangible interfaces, embodied guidance, and haptic communication. We also demonstrate use cases and application scenarios for each interaction, such as physical storytelling, remote tangible gaming, and hands-on instruction.

Second, we contribute to a *holistic user evaluation* to better understand how mobile robots can enhance holographic telepresence in different application scenarios. To this end, we compare our approach (hologram + robot) with hologram-only and robots-only conditions through a within-subject user study with twelve participants. Both quantitative and qualitative results suggest that our system significantly enhances the level of co-presence and shared experience for mixed reality remote collaboration, compared to the other two conditions. Based on the insights, we also discuss the future of tangible remote collaboration that leverages robotic environments.

Finally, this paper contributes the following:

- (1) A design space exploration and application demonstrations that showcase a set of possible interactions and use cases enabled by our system.
- (2) Results and insights from our user study that confirm the benefits of our approach over hologram-only and robots-only conditions.

2 RELATED WORK

2.1 Remote Collaboration

2.1.1 Mixed Reality Remote Collaboration. Recent advances in mixed reality technologies have enabled immersive remote collaboration that was not possible with traditional desktop interfaces. Prior research has explored various approaches for immersive telepresence, such as holographic teleportation (e.g., *Holoportation* [51], *Virtual Makerspaces* [60], *Loki* [75]), virtual avatars (e.g., *CollaboVR* [18], *Mini-Me* [57], *Shoulder of Giant* [58], *ARTEMIS* [15]), and projected video stream (e.g., *Room2Room* [56], *3D-Board* [82]). These systems allow remote users to be spatially co-located in the same shared space, which greatly enhances collaborative experiences [4, 8]. For example, by showing virtual hands and bodies in 3D space, the local users can more easily understand the intention of the remote users for various physical tasks such as block assembly [79], origami [31, 32], mechanical tasks [52, 53], and physiotherapy education [12]. However, current holographic telepresence lacks the physical embodiment of the remote user, which significantly reduces the sense of co-presence [37]. This limitation also constraints rich physical affordances which we naturally employ in co-located physical collaboration [40, 68].

2.1.2 Robotic Telepresence. To address this limitation, past research has explored robotic telepresence that aims to physically embody remote users by adding a robotic body to a 2D video screen (e.g., *MeBot* [1], *RemoteCode* [66]) or by replicating the remote user with a humanoid or non-humanoid robot (e.g., *TELESAR V* [13], *Telenoid* [50], *You as a Puppet* [65], *GestureMan* [34], *Geminoid* [64]). The robotic telepresence can greatly enhance user engagement by enabling physical interactions such as gestures [1] and body movement [38, 48, 61]. For example, mobile robots allow remote users to move freely around a table to interact with local users and objects for remote education (e.g., *RobotAR* [76], *ASTEROIDS* [42]). Beyond a screen-based representation, *VROOM* [22, 23] overlays a holographic avatar on a telepresence robot that enriches non-verbal communication such as gestures or eye-contact.

2.1.3 Physical Telepresence. An alternative approach to adding physical embodiment to remote users is using *synchronized distributed physical objects* [6], rather than embodying users themselves with robotic telepresence. Such an approach was originally explored through *InTouch* [5], *ComTouch* [9], and *PsyBench* [6], in which synchronized tangible tokens embody the remote user's motion and behavior. This idea has evolved into a concept of *physical telepresence* [40], which synchronizes physical shape rendering with the remote users' visual appearance. For instance, Leithinger et al. [40] uses a shape-changing display [14] to physically render a remote user's hand and surrounding objects with screen-based visual feedback. Recent works have also expanded this concept by combining a virtual avatar with a motorized X-Y plotter to actuate a single token (e.g., *Physical-Virtual Table* [37]). However, the existing approach using shape displays lacks deployability due to the dedicated hardware requirement, and X-Y plotters lack scalability and generalizability due to a single point actuation and limited interaction area. More closely related to our work, a few researchers have explored the use of mobile robots for tangible remote collaboration in VR (e.g., *PhyShare* [19]) and mixed reality

environments (e.g., Siu *et al.* [68]). However, this approach of using multiple mobile robots has not been fully explored yet, as these prior works do not present the comprehensive design space and have not conducted any user evaluation to understand the benefits and limitations of this approach. Beyond these prior works, we contribute to 1) an exploration of the broader design space with a demonstration of comprehensive applications, and 2) a holistic user evaluation through condition experiments.

2.2 Bi-Directional Virtual-Physical Interaction

Outside the context of remote collaboration, past HCI research has also explored bi-directional virtual-physical interaction by leveraging augmented reality and actuated environments [70]. For example, systems like *Kobito* [3], *Augmented Coliseum* [33], and *InreTable* [41] explore the synchronous coupling between AR and actuated physical objects, which can enrich visual feedback and affordances of robots and actuated tangible interfaces. These interfaces typically employ robot motion (e.g., *exTouch* [25]), actuated tangible tokens (e.g., *PICO* [55], *Reactile* [71], *Actuated Workbench* [54]), IoT devices (e.g., *MechARSpace* [81], *WIKA* [21], Kim *et al.* [28]) to synchronize between virtual and physical outputs in a bi-directional manner. Similar to our work, *Sketched Reality* [24] and *Physica* [43] explores bi-directional interaction between embedded virtual objects and tabletop robots. Our system extends their work in the context of holographic tangible remote collaboration in mixed reality environments.

2.3 Actuated Tangible User Interfaces

Actuated tangible user interfaces were originally developed to address the challenge of digital-physical discrepancies in conventional tangible interfaces [59]. Towards this goal, HCI researchers have explored a variety of actuated tangible user interfaces [59] and shape-changing user interfaces [2, 10, 62], using magnetic actuation [55], ultrasonic waves [44], magnetic levitation [36], and wheeled and vibrating robots [49]. Rosenfeld *et al.* [63] introduced the concept of using physical mobile robots as an actuated tangible user interface. This concept has been expanded through various systems such as *Zooids* [35], *ShapeBots* [74], *HERMITS* [46], *Rolling Pixels* [39], and *(Dis) Appearables* [47]. Swarm user interfaces can also provide haptic sensations [30, 72, 73, 80] and actuate everyday objects [11, 29]. Inspired by these works, we also leverage multiple tabletop robots for our actuated interfaces.

3 HOLOBOTS: SYSTEM DESIGN

This section introduces HoloBots, a system that augments holographic telepresence with multiple tabletop robots. As illustrated in Figure 2, HoloBots consists of three main components: 1) *capturing a remote user* with the Azure Kinect depth camera, 2) *holographic rendering and hand tracking* with Microsoft Hololens 2 headset, and 3) *synchronized actuation* with Sony Toio tabletop mobile robots.

Capturing a Remote User with a Depth Camera. The Azure Kinect RGB-D camera is used to capture the remote user's body. The camera is positioned in front of the remote user with a tripod stand. The Kinect camera is connected to the local PC (G-Tune, Intel Core i7-11800H 2.30GHz CPU, NVIDIA GeForce RTX 3060 GPU, 64GB RAM) via a USB cable. The depth information is captured through

the Azure Kinect Sensor SDK running on the local PC. The depth sensor first generates a point cloud with a resolution of 640 x 576, which is then converted into a real-time colored 3D mesh using the Azure Kinect Examples for Unity package¹. Mesh data is captured with 30 FPS and the size of each mesh data is approximately 20 MB.

Holographic Rendering and Hand Tracking. In our setup, both local and remote users wear the Microsoft Hololens 2 mixed reality headset, which has a diagonal field of view of 52 degrees. The remote user's holographic mesh generated by the local PC is rendered in Hololens 2 through Holographic Remoting², enabling high-quality and low-latency (60 FPS) rendering over an Ethernet connection, allowing the local user to view the mesh. Hololens 2 is also used to track the user's hand movements using the MRTK hand-tracking library. Tracked hands are used to 1) grab virtual robots to manipulate and synchronize the physical one in the remote environment, or 2) move robots based on the finger position to physicalize the virtual hand. These processes are executed on Unity running on the remote PC (G-Tune, Intel Core i7-11800H 2.30GHz CPU, NVIDIA GeForce RTX 3070 GPU, 64GB RAM) and the local PC, respectively, connected with Hololens 2 through Holographic Remoting.

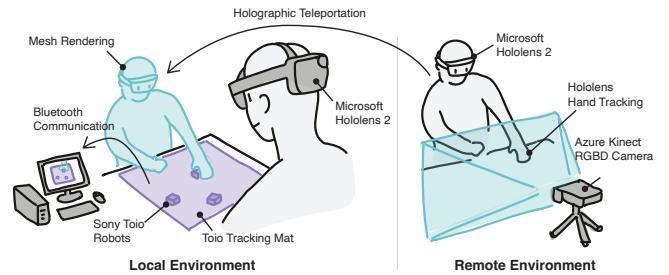


Figure 2: System Setup: The local user can see the remote user's avatar and interact Toios or virtual objects with the remote user through Hololens. The remote user's body is tracked by Azure Kinect and the hands are tracked by Hololens.

Synchronized Actuation with Tabletop Mobile Robots. Our system uses Sony Toio³ as tabletop mobile robots. Each robot measures 3.2 cm × 3.2 cm × 2.5 cm and can move at a speed of up to 35 cm/sec for straight-line movement and 1500 deg/sec for rotation. The robot has a built-in camera that can scan patterns printed on a mat (Toio Tracking Mat) to detect their position and orientation. The size of the tracking mat has 55 cm × 55 cm of covered area, but it can be extended by aligning multiple mats. Each Toio robot is controlled using the Toio SDK for Unity⁴ on a PC and continuously sends its position and orientation to the PC via Bluetooth® standard Ver. 4.2 every 10 ms. For the controlling algorithm, we adapt to the open-source library [46] and rewrite the algorithm for our Unity application. To start using our application,

¹<https://assetstore.unity.com/packages/tools/149700>

²<https://learn.microsoft.com/en-us/windows/mixed-reality/develop/native/holographic-remoting-player>

³<https://www.sony.com/en/SonyInfo/design/stories/toio/>

⁴<https://github.com/morikatron/toio-sdk-for-unity>

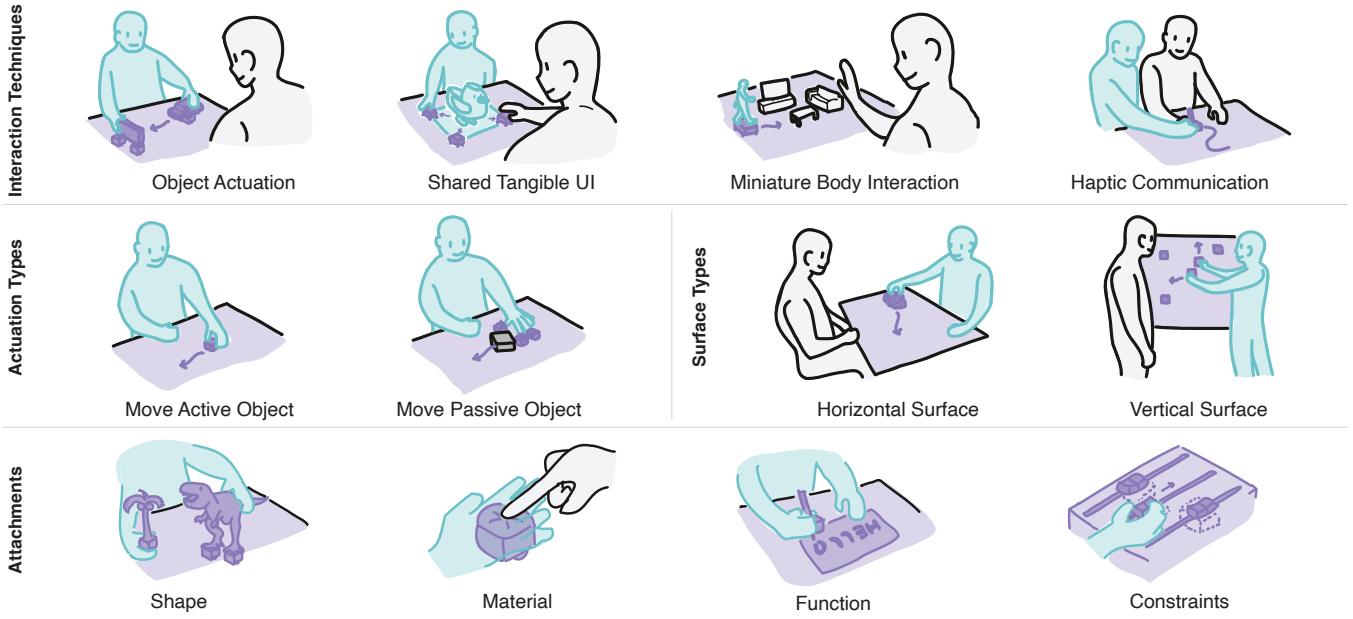


Figure 3: Design Space of HoloBots.

the local user first performs a manual calibration to align the remote user’s holographic mesh with the Toio mat. This alignment can be bypassed in subsequent uses saving the relative position between the Toio mat and the avatar mesh. By placing a QR code on the Toio mat to acquire the mat’s position and leveraging the relative position, we can display the avatar mesh in the appropriate position.

After the calibration, each robot’s position is controlled through the following three ways: 1) physical Toio movement in the remote environment, 2) virtual object movement in the remote user’s Hololens, or 3) finger position movement of the remote user. When both users have a physical Toio setup, the system can simply synchronize the position of each environment. On the other hand, when only the local user is equipped with the Toio robot, then the remote user can manipulate virtual Toios by grasping and manipulating virtual Toio objects rendered in the Hololens, while the local user manipulates physical Toios. Alternatively, the remote user can manipulate these Toio robots with hand and finger tracking. For the finger binding, we use the thumb, index, and/or pinky finger positions, depending on the available number of robots. The position data for each robot is sent between the remote and local PCs through UDP communication. In our implementation, we set the Toio robot’s speed up to 17.5 cm/sec taking into account the balance between the speed and accuracy. Therefore, if the remote user attempts to move the local Toio robot at speed higher than this, it may lead to positional errors. Considering the tradeoff between precise movements and jittering, we set the default tolerance to 1.1 cm for all interactions, except for the miniature body interaction, where we set it to 0.4 cm since accuracy with the avatar’s body was more crucial than some small jittering. Finally, to increase the reproducibility, we make our software open source⁵.

⁵<https://github.com/KeiichiHara/HoloBots>

4 HOLOBOTS DESIGN SPACE

In this section, we explore the design space of HoloBots in the following four dimensions: 1) interaction techniques, 2) actuation types, 3) surface types, and 4) physical attachment (Figure 3).

4.1 Interaction Techniques

Object Actuation. HoloBots offers various ways for remote users to interact with the local user. The object actuation enables remote users to move and manipulate objects in the local environment. For example, remote users can directly grab the Toio robot to move its location, or the attached object for more expressive engagement.



Figure 4: Storytelling

The object actuation can be used for different use cases, such as storytelling, gaming, and drawing. For storytelling, HoloBots allows both local and remote users to participate in creating a story together with tangible objects. The local user can either observe as an audience member or actively engage with the story-creation process. Figure 4 illustrates a remote user physically moving a dinosaur toy on a stage to narrate a story to the local user. This provides, for example, engaging tangible storytelling for children and their remote parents or friends.

Shared Tangible UI. Another interaction technique is the shared tangible user interface, which allows both local and remote users to manipulate virtual object properties through tangible tokens. Toio

robots can be represented as various tangible UI elements, such as control points, buttons, sliders, and knobs, so that by controlling the same UI, local and remote users can manipulate the UI together. For example, Figure 5 illustrates users changing the position and scale of the virtual picture by manipulating robots, which are represented as a control point of the image.



Figure 5: Shared Tangible UI

In Figure 6, three Toios are used as a tangible UI for manipulating a 3D object. Two of the Toios represent sliders and adjust the width and depth of the object, while the third Toio represents a knob and alters the object's height.



Figure 6: Collaborative Design

Miniature Body Interaction. The robot can also embody the remote user through a miniature body. Our system also facilitates the collaborative world-in-miniature exploration, by representing as the miniature user. Similar to the prior work that explores tangible world-in-miniature exploration (e.g., *miniStudio* [27], *Does it Feel Real?* [45], *Shoulder of Giants* [58], *ASTEROIDS* [42]), the tangible embodiment of the miniature user facilitates rich physical affordances for the world-in-miniature interaction, while providing effective visual feedback through holographic representation. The remote user can walk around on a real-size environment, which is captured and tracked through Azure Kinect body tracking. For example, the remote user can visually instruct the local user using gestures and physically move objects in the local environment by pushing them with Toios (Figure 7).



Figure 7: Miniature Body Interaction

Taking inspiration from the immersive interior and architectural design (e.g., *DollhouseVR* [20]), this could be used for the collaborative world-in-miniature exploration, in which the robot can embody the physical representation of the miniature user. For example, Figure 8 illustrates an application for collaborative interior design. This application uses miniature furniture to facilitate discussion and decision-making between remote and local users. The remote user is visually represented as a miniature avatar, with a

Toio representing the remote user's physical body. The remote user can visually instruct the local user using gestures and physically push the miniature furniture to arrange the position.



Figure 8: Interior Design

Haptic Communication. Haptic communication is another interaction technique that enables the remote user to provide haptic feedback to the local user. There are various ways to provide haptic communication. For example, the user can guide the Toio robot to navigate the remote user based on the actuation, similar to *de-Pend* [77], as if they were holding their hands. This technique can be used for hands-on instruction. Alternatively, the remote user can physically touch the local user by moving and touching the local user's body using Toios, similar to *SwarmHaptics* [30]. This can be used for remote social interaction.

Figure 9 shows a remote user controlling the movement of a red pen to draw on the physical canvas. By attaching a physical pen to a Toio, the remote user can move the pen and draw on a physical canvas. Local and remote users can therefore collaborate in real-time to create drawings and illustrations together.



Figure 9: Haptic Communication

This can also provide haptic notifications, enabling remote users to physically notify local users using Toios. By attaching Toios to the remote avatar's hand, the remote user can touch the local user and initiate communication. In Figure 10, the remote user touches the local user who is reading a book to start a conversation.



Figure 10: Notification

4.2 Actuation Types

Move Active Object. In HoloBots, the remote users can actuate physical objects in two ways. First, the user can simply grasp and move the Toio robot itself. By moving the Toio, which is attached to the various object, HoloBots enables the remote user to actuate physical objects (Figure 4).

One possible application is the remote gaming experience. By attaching Toios to game objects, the local user can physically interact with the remote user through the tangible game. Figure 11 depicts a table hockey game application that utilizes three Toios—two for the mallets and one for the puck, similar to [24, 47]. This application allows users to play and compete with each other in real-time, creating an engaging and immersive gaming experience.

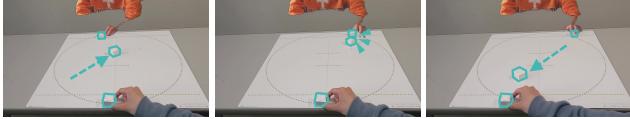


Figure 11: Table Hockey

Move Passive Object. Alternatively, the user can also actuate everyday passive objects by pushing these objects with the Toio. This allows actuating objects without attaching robots in advance. Similar to [26], by making the robots follow the user’s fingers, the remote user can physicalize their hands and fingers, so that pushing the other passive objects (Figure 12). This method allows an intuitive way of interacting with physical objects, as the remote user can use hand gestures to control objects. In the current setup, each Toio can push an object up to 32 grams.



Figure 12: Move Passive Object

4.3 Surface Types

Horizontal Surface. HoloBots also supports two different surface types that the robot can move around. The first one is the horizontal surface, such as a tabletop surface where the users sit down together to manipulate objects on the table.

Vertical Surface. Alternatively, by attaching a small magnet at the back of the Toio, Toio can move on a vertical surface such as a whiteboard or a magnetic wall. By moving Toios on a vertical surface can be useful for applications that require standing up, such as brainstorming or presentations. In our prototype, we attach an N35 neodymium magnet (8 mm × 3 mm, 1 mm thickness) to the bottom of the Toio robot with tape, which has a strong attraction force to be attached to the whiteboard, while weak enough to move on a wall. For the tracking of the vertical surface, we use a thinner tracking mat (Toio Developer Mat, 0.1 mm thickness) that can be attached to the whiteboard. With the vertical surface, we can also expand the application domains, such as collaborative discussion and brainstorming with the post-it notes on a whiteboard (Figure 13).



Figure 13: Vertical Surface

4.4 Attachments of the Robot

HoloBots is also designed to be versatile and adaptable to various applications by allowing the user to attach different components to the robot. These attachments provide additional functionalities and enable the robot to perform a wider range of tasks, making it suitable for a variety of applications.

Shape Props. Shape props can modify the robot’s shape and physical appearance. As illustrated in Figure 4, attaching a dinosaur toy to the robot can be used to represent a dinosaur, expanding its interactive potential. By attaching Toios to physical objects such as puppets, stuffed animals, toy figures, and LEGO blocks, both local and remote users can move the objects, crafting the story and narrative, as we do in physical space.

Material Props. The addition of material props such as soft materials, fur, and fabric enables the local user to enhance the sensation of remote objects and users. For example, by attaching soft materials, mobile robots can represent remote users’ hands to improve haptic communication. Also, the use of fabric materials enables the mobile robots to represent portions of the remote user’s arm that are clothed.

Functional Props. Attachments can supplement the robot with added functionalities. For example, Figure 14 illustrates remote users drawing on a transparent sheet using a robot equipped with a pen, which facilitates visual communication between users. As shown in Figure 13, attaching post-it notes to the mobile robots enables the remote user to highlight specific parts in the local user’s environment. Also, by attaching magnets to the robots, users can extend their mobility from horizontal to vertical surfaces.



Figure 14: Pen Attachment for Drawing

Constraints. Mechanical constraints, such as rings and rubber bands, can be employed to restrict the movements of mobile robots as PICO [55]. This provides both the remote and local users to move the robots within a specific range of movement. For example, by using a straight ring, the movement of mobile robots can be limited to a straight line, which could help the creation of a precise slider UI. Also, confining all the mobile robots within a ring can help limit the area in which the remote user can influence the local environment.

5 USER STUDY

To evaluate the effectiveness of incorporating both virtual and physical representations in holographic remote collaboration, we conducted a user study comparing our system with Hologram-Only and Robots-Only conditions across four distinct interactions. We gathered both quantitative and qualitative measurements for various aspects, such as social presence, system usability, and cognitive workload, with a within-subject user study.

5.1 Method

5.1.1 Participants. We recruited 12 participants (11 male, 1 female) from our local university, with an age range of 21-24 years ($M = 22.1$, $SD = 1.16$). Participants were surveyed on their familiarity with VR/AR using a 7-point Likert scale from 1 (novice) to 7 (expert), and the average score was 2.75 ($SD = 1.71$).

5.1.2 Study Setup. We present the setup used in our study in Figure 15. One of the authors acts as a remote collaborator (referred to as "the experimenter") for each participant to reduce differences in interaction between groups. The participant and the experimenter are situated in separate rooms and communicate remotely. The dimensions of the participant's room were approximately 11.3 m by 5.2 m, while the experimenter's room measured approximately 7.5 m by 5.9 m. Both the participant and experimenter were equipped with Hololens 2 headsets. On the participant's desk, we placed the Toios and a Toio mat. To enable the experimenter to view the participant's workspace, we used an iPad to capture the video image and transmitted it to the experimenter's display.

For audio communication, we used Discord⁶, a voice chat application. To mitigate any potential interference from the Toio's sound, the participant wore noise-canceling headphones (Sony WH-1000XM4).



Figure 15: Study Setup. Left: Experimenter's room, Right: Participant's room

5.1.3 Study Design. We designed our study with a within-subject design that compares the following three conditions:

C1. Hologram + Robots : Participants interacted with the remote experimenter via hologram and voice chat with using mobile robots.

C2. Hologram-Only : Participants interacted with the remote experimenter via hologram and voice chat without using mobile robots.

C3. Robots-Only : Participants interacted with the remote experimenter via voice chat without a hologram with using mobile robots.

To evaluate the difference in these conditions across various interactions, we used four application scenarios that best represent each interaction technique in our design space:

D1. Object Actuation : We used the physical storytelling application (Figure 4). Participants were instructed to create a short story with the remote experimenter by manipulating virtual or physical dinosaur toys. The remote experimenter could also move the dinosaur toys. The fundamental elements of the stories shared similarities, including dinosaurs fighting, making up, walking around, and talking to each other. However, participants chose how the dinosaurs fight, where they make up or rest, and which directions they walk. We displayed virtual toys for C2 and used physical toys for C1 and C3.

D2. Shared Tangible UI : We used virtual image manipulation (Figure 5). Participants were instructed to adjust the size of a virtual picture until it matched the target size printed on paper, collaborating with the remote experimenter. The virtual cubes or mobile robots were attached to the upper left and bottom right of the virtual picture, and participants and a remote experimenter could adjust the size by moving them. We used virtual cubes for C2, and used physical cubes for C1 and C3. The virtual picture was displayed in all conditions.

D3. Miniature Body Interaction : We used the interior and architectural design application (Figure 8). The remote experimenter was presented as a miniature body, similar in size to miniature furniture. Participants were instructed to move the furniture and determine furniture placement in discussion with the remote experimenter. The remote experimenter could also physically move the furniture in the conditions with a mobile robot.

D4. Haptic Communication : We used the haptic notification application (Figure 10). Participants were instructed to read a book and engage in conversation with the remote experimenter when they were contacted. The remote experimenter initiated contact through virtual or physical touch, with mobile robots following the remote experimenter's fingers to physically touch the participants.

5.1.4 Measurements. We measured four different aspects: 1) **Social Presence** based on the Networked Mind Measure of Social Presence questionnaire [16], 2) **Cognitive Workload** based on NASA Taskload Index (NASA-TLX) [17], 3) **System Usability** based on System Usability Scale (SUS) [7], and 4) **Preference** based on the questionnaire in which the participants were asked which condition they preferred the most. In addition to these measurements, we conducted an interview after the study to gather qualitative feedback from participants.

5.1.5 Procedure. After participants signed a consent form, we provided them with instructions on how to use the Hololens 2 and Toio robot. Participants then conducted a task involving 12 sessions (4 applications \times 3 conditions), each lasting 3 minutes. Participants used four applications in the following order: Shared Tangible UI, Object Actuation, Miniature Body Interaction, and Haptic Communication. Participants conducted each application in all three conditions (C1, C2, and C3). The order of the three conditions was

⁶<https://discord.com/>

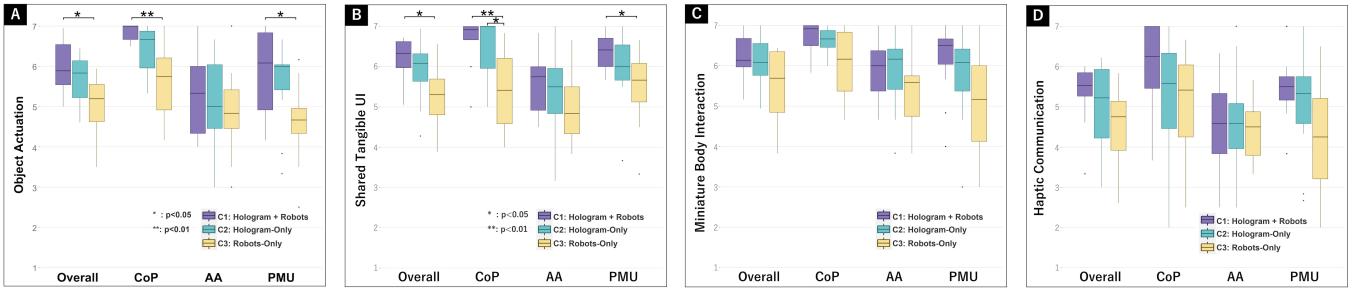


Figure 16: Social Presence Questionnaire Results. A: Object Actuation, B: Shared Tangible UI, C: Miniature Body Interaction, D: Haptic Communication. CoP: Co-Presence, AA: Attentional Allocation, PMU: Perceived Message Understanding.

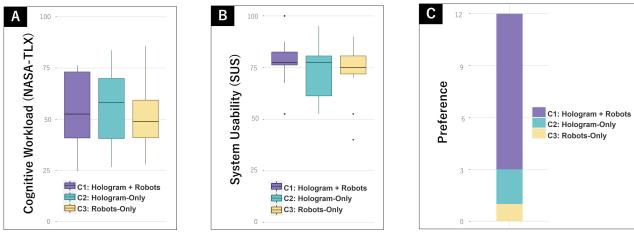


Figure 17: A: Cognitive Workload (NASA-TLX), B: System Usability (SUS), C: Preference

counterbalanced across participants to control for order effects. After each application, participants answered the social presence questionnaire. After each condition, participants answered the SUS and NASA-TLX questionnaire to compare the three conditions. In total, we asked the participants to complete 12 social presence questionnaires and 3 SUS and NASA TLX questionnaires. After the participants finished all of the sessions, we conducted a brief open-ended interview for 10–15 minutes. The study took approximately 90 minutes in total, and each participant was compensated with 10 USD (The actual currency has been removed for anonymity).

5.2 Results

To analyze the data collected in our study, we employed a Friedman's test for each measurement. To assess pairwise differences between conditions, we conducted multiple pair-wise comparisons using the Wilcoxon signed-rank test with Bonferroni correction. We set the significant level at 5 %.

5.2.1 Social Presence. The Social Presence Questionnaire consisted of three sub-scales: Co-Presence (CoP), Attentional Allocation (AA), and Perceived Message Understanding (PMU). Figure 16 shows the result of the social presence questionnaire for a total of 12 sessions (4 applications × 3 conditions for each). In addition, we calculated an overall score by averaging the three sub-scales. We checked the internal consistency with Cronbach's alpha for each sub-scale: $\alpha_{CoP} = 0.90$, $\alpha_{AA} = 0.78$, $\alpha_{PMU} = 0.93$.

For Object Actuation (D1) and Shared Tangible UI (D2), Hologram + Robots (C1) condition had significantly higher overall social

presence scores than Robots-Only (C3) condition. For both Object Actuation (D1) and Shared Tangible UI (D2), pairwise comparisons revealed that Hologram + Robots (C1) condition was significantly higher scores than Robots-Only (C3) condition for CoP (D1: $Z = 3.68$, $p = 0.0007 < 0.001$, D2: $Z = 3.29$, $p = 0.003 < 0.01$), PMU (D1: $Z = 2.46$, $p = 0.042 < 0.05$, D2: $Z = 2.61$, $p = 0.027 < 0.05$), and Overall (D1: $Z = 2.63$, $p = 0.025 < 0.05$, D2: $Z = 2.86$, $p = 0.013 < 0.05$).

In the interviews, participants made comments that suggested that Hologram + Robots (C1) condition resulted in a stronger sense of presence compared to Hologram-Only (C2) condition. Specifically, one participant noted that “*Hologram + Robots clearly felt the presence of the other party, whereas Hologram alone was less present.*” (P1), while another participant mentioned that “*Hologram-only conditions were difficult to react to when the other person was out of sight*” (P2). These comments suggest that combining mobile robots with holographic telepresence could help users better understand the remote user’s actions and movements, even when the holographic user is out of sight. Furthermore, for all four applications, the graph of the data suggested that Hologram + Robots (C1) had the highest scores, followed by Hologram-Only (C2) and Robots-Only (C3).

5.2.2 Cognitive Workload. The results for the cognitive workload are shown in Figure 17 (A). A lower score indicates a lower workload. The average score for each condition was 54.0 ($SD = 19.4$) for Hologram + Robots (C1) conditions, 55.3 ($SD = 18.9$) for Hologram-Only (C2), and 51.6 ($SD = 16.9$) for Robots-Only (C3). The Friedman test showed no significant difference ($\chi^2(2) = 0.30$, $p = 0.86$).

5.2.3 System Usability. The results for the system usability scale are shown in Figure 17 (B). A higher score indicates higher usability. The average score for each condition was 77.9 ($SD = 11.3$) for Hologram + Robots (C1) conditions, 73.3 ($SD = 14.2$) for Hologram-Only (C2), and 72.9 ($SD = 14.1$) for Robots-Only (C3). The Friedman test showed no significant difference ($\chi^2(2) = 1.64$, $p = 0.44$). During the interviews, participants provided feedback on the usability. One participant noted that “*Conditions which use Toio were easy to manipulate*” (P3), and another participant noted that “*It was easy to adjust the size of the virtual picture using Toio*” (P8). Although Hologram + Robots (C1) had a higher average usability score (77.9) than the average score (68) [67], it was not significantly better than Hologram-Only (C2). One participant reported, “*The coupling between the actual movements and the robot was slow and misaligned,*

which sometimes make it difficult to understand” (P3). This feedback suggests that the low ability of coupling between the hologram and mobile robots may have negatively impacted usability.

5.2.4 Preference. The results for the preference are shown in Figure 17 (C). 75 % of the participants preferred Hologram + Robots (C1) as the best, followed by Hologram-Only (C2) (17 %) and Robots-Only (C3) (8 %). Chi-squared goodness of fit test revealed a significant difference from random choice ($\chi^2(2) = 9.5, p = 0.009 < 0.01$). In our study, participants preferred Hologram + Robots (C1) over Hologram-Only (C2) and Robots-Only (C3). Five out of nine participants mentioned social presence as a key factor in their preference for Hologram + Robots (C1), while the remaining four participants mentioned usability as a determining factor. Therefore, the high social presence and usability in Hologram + Robots (C1) can enhance the overall user experience.

5.3 Limitations and Design Implications

5.3.1 Precise Coupling between Holographic Users and Robot Movement. In the applications used in the study, the coupling between the virtual body movements and mobile robots was occasionally slow and misaligned due to the Toio’s maximum speed and the calibration error between the avatar and Toios. Upon testing the start latency, the average latency was 0.483 s, 0.262 s, 0.443 s, and 0.615 s in D1, D2, D3, and D4, respectively. This issue could potentially impact both the social presence and usability of the Hologram + Robots (C1) condition. Employing faster mobile robots and implementing a more accurate position calibration method between the avatar and Toios could alleviate this problem.

5.3.2 Noise of Robot Movement. Several participants reported that the sound generated by the Toios could be distracting and interfere with their ability to concentrate on the task. For example, one participant commented “*Toios sound was sometimes a little loud, and it was difficult to concentrate on the task.*” (P2), while another mentioned “*I was distracted by the noise of Toios*” (P3). Upon testing the noise levels generated when moving the Toio 45 cm in 4 seconds, the maximum recorded noise was 64.5 dB, 60.3 dB, 65.0 dB, and 70.0 dB in D1, D2, D3, and D4, respectively. This issue could potentially impact the user experience and social presence. To address this problem, we could improve the system to make Toios travel to their destination by the shortest route, reducing travel time and the duration of sound generation.

5.3.3 Bi-Directional Collaboration between Participants. Additionally, the collaboration in our study was between a participant and an experimenter. To gain further insights, it may be beneficial to set up an environment where participants can collaborate with other participants without the presence of an experimenter. This can provide insights on more realistic collaboration scenarios.

5.3.4 Group Size. In our study, collaboration was limited to only two people, one participant and one experimenter. Using larger groups could potentially increase the number of interactions and affect the social presence and user experience. However, this could also increase conflicts and misunderstandings. Therefore, conducting studies with larger groups could help us understand how these factors influence our system.

5.3.5 Number of Robots. In our study, we used two Toios, but it is possible to use more. One participant noted that “*I thought it would be good if the picture application could increase the number of manipulable objects (Toio) and allow more complex UI manipulation*” (P8). This comment suggests that using more Toios for UI manipulation could affect usability and user experience. Additionally, we could use more Toios for body or hand representation, which could enhance the resolution of the remote user’s movements and gestures, which could enhance the social presence.

5.3.6 Enhancing Holographic Visualization. In this study, we used a single Kinect camera to capture the remote user’s body movements for holographic avatar generation. Future work could expand this setup by adding more Kinect cameras to capture the user’s hologram from multiple angles. This could improve the remote user’s clarity and accuracy via multi-directional coverage. Through these improvements, the local user would better comprehend the remote user’s intentions and interactions with the physical environment and overall body language.

6 FUTURE WORK

In this paper, our design space exploration is limited to the form factor of HoloBots. In this section, we suggest several directions to expand the design space of holographic tangible remote collaboration with different form factors, interaction modality, and user representations.

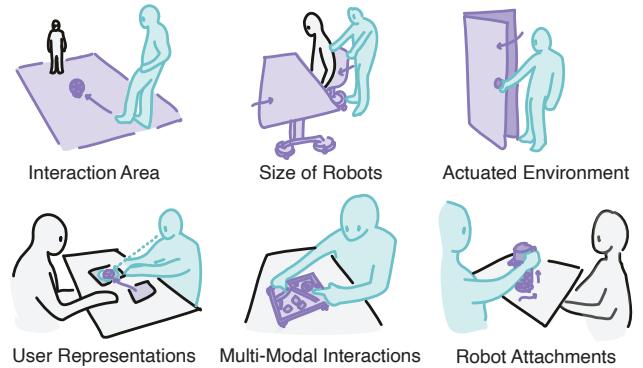


Figure 18: Future Work

Scalability of Interaction Area. In our current implementation, the interaction area was limited to one Toio mat, which restricted users to using the limited space of the table or whiteboard. However, by aligning multiple Toio mats, we can expand the interaction area to cover larger surfaces, such as larger tables or whiteboards. This expansion would allow for the use of additional mobile robots and the accommodation of more users. Furthermore, incorporating fiducial markers and tracking cameras, as seen in ASTEROIDS [42], could further increase the scalability of the interaction area to encompass an entire table or even a whole room. This approach is not limited to Toio robots, meaning that various robotic platforms could also be incorporated.

Different Robotic Form Factors. In this study, we focused on the combination of a holographic avatar with tabletop mobile robots. However, future work could expand on this area by incorporating different types of robots. For example, larger robots, such as robotic vacuum cleaners, could be utilized. By synchronizing large robots' movements with the remote user, the holographic user could physically impact the local environment on a larger scale. This could facilitate physical gaming and sports, and/or enable real-scale interior design by moving actual furniture, similarly to *RoomShift* [69].

IoT and Actuated Environments. In future work, we could integrate IoT devices, such as lights, fans, and curtains to enable users to affect the local environment in a different way. For example, by synchronizing the remote user's movements with a door's movements, the remote user could interact with the door in the local environment. Also, by using reeling mechanisms like AeroRigUI [78], remote users can raise and lower the curtains in the local environment.

Different User Representation. In addition to the current finger and miniature body representations, there are other potential ways to embody remote users using Toio robots. We are interested in exploring the following alternative representations in future work. *Whole Arm:* Mobile robots could represent the entire arm of the remote user, enabling them to physically move multiple objects simultaneously or to visualize the joints in the user's arm such as their wrist and elbow. *Gaze and Eye:* Mobile robots could follow the remote user's gaze or eye movement, allowing the local user to physically understand where the remote user is looking, an insightful metric for gauging a user's intent or thought processes. *Shadows:* Mobile robots could represent the shadow of the remote user, providing additional information about their body position beyond what an avatar can offer.

Multi-Modal Interaction. Another potential direction for future work would be to explore multi-modal interaction methods that incorporate different modalities of communication. One such method could involve the use of a tablet device. By connecting mobile robots to the tablet and sharing the screen between the tablets of both remote and local users, the remote user can manipulate the tablet's position and engage with its contents. This interaction enables functionalities such as scaling and rotating maps shown on the tablet, as well as drawing pictures on the tablet.

Versatile Robot Attachments. In future work, the capabilities of HoloBots could be enhanced by incorporating versatile physical attachments, inspired by the *HERMITS* [46] concept. These on-demand attachments could significantly expand the functionality and flexibility of the tabletop robots and enable them to adapt to a wider range of tasks and interaction scenarios for tangible remote collaboration. Here are some possible directions for versatile physical attachments: *2.5D Shape Display:* By adopting the technique presented in *HapticBots* [72], we can enhance our mobile robots to enable height adjustments, allowing them to actuate in 2.5 dimensions. *Attaching Tangible Controllers:* By attaching complex tangible controllers, such as joysticks, sliders, or knobs, our system could enable remote users to perform intricate manipulations using the shared tangible UI. *Attaching Grippers:* Equipping mobile robots with grippers could allow the remote user to manipulate

small objects, thereby making it possible for them to actuate a wider range of object types. *Force Aggregation:* Combining multiple robots within a single shell could aggregate their force, enabling the remote user to move heavier objects for object actuation or provide stronger haptic feedback for haptic communication.

7 CONCLUSION

In this paper, we presented HoloBots, a novel mixed reality interface that augments holographic telepresence through synchronized tabletop mobile robots. With HoloBots, we demonstrated that the remote users can physically engage with local users and the environment, enabling them to touch, grasp, manipulate, and interact with tangible objects as if they were co-located in the same space. This paper expands upon existing physical telepresence by presenting more comprehensive design space and interaction techniques, such as object actuation, virtual hand physicalization, miniature body interaction, shared tangible interfaces, embodied guidance, and haptic communication. We demonstrated various applications for HoloBots, such as physical storytelling, remote tangible gaming, and hands-on instruction. A user study with twelve participants revealed that HoloBots significantly enhances co-presence and shared experiences in mixed reality remote collaboration, proving its scalability, deployability, and generalizability for a wide range of remote tangible collaboration scenarios. Additionally, we have outlined several potential avenues for future work that could extend the design space and uncover new opportunities for HoloBots.

ACKNOWLEDGMENTS

This research was funded part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2021-02857, NSERC RTI Grant RTI-2023-00582, and Mitacs Globalink Research Award. We also thank all of the participants for our user study.

REFERENCES

- [1] Sigurdur Orn Adalgeirsson and Cynthia Breazeal. 2010. MeBot: A robotic platform for socially embodied telepresence. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 15–22.
- [2] Jason Alexander, Anne Roudaut, Jürgen Steimle, Kasper Hornbæk, Miguel Bruns Alonso, Sean Follmer, and Timothy Merritt. 2018. Grand challenges in shape-changing interface research. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.
- [3] Takafumi Aoki, Takashi Matsushita, Yuichiro Iio, Hironori Mitake, Takashi Toyama, Shoichi Hasegawa, Rikiya Ayukawa, Hiroshi Ichikawa, Makoto Sato, Takatsugu Kuriyama, et al. 2005. Kobito: virtual brownies. In *ACM SIGGRAPH 2005 emerging technologies*. 11–es.
- [4] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
- [5] Scott Brave and Andrew Dahley. 1997. inTouch: a medium for haptic interpersonal communication. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems*. 363–364.
- [6] Scott Brave, Hiroshi Ishii, and Andrew Dahley. 1998. Tangible interfaces for remote collaboration and communication. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. 169–178.
- [7] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).
- [8] Yuanzhi Cao, Xun Qian, Tianyi Wang, Rachel Lee, Ke Huo, and Karthik Ramani. 2020. An exploratory study of augmented reality presence for tutoring machine tasks. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
- [9] Angela Chang, Sile O'Modhrain, Rob Jacob, Eric Gunther, and Hiroshi Ishii. 2002. ComTouch: design of a vibrotactile communication device. In *Proceedings of the*

- 4th conference on Designing interactive systems: processes, practices, methods, and techniques. 312–320.
- [10] Marcelo Coelho and Jamie Zigelbaum. 2011. Shape-changing interfaces. *Personal and Ubiquitous Computing* 15, 2 (2011), 161–173.
- [11] Samin Farajian, Hiroki Kaimoto, and Ryo Suzuki. 2022. Swarm Fabrication: Reconfigurable 3D Printers and Drawing Plotters Made of Swarm Robots. *arXiv preprint arXiv:2202.10978* (2022).
- [12] Mehrad Faridan, Bheesha Kumari, and Ryo Suzuki. 2023. ChameleonControl: Teleoperating Real Human Surrogates through Mixed Reality Gestural Guidance for Remote Hands-on Classrooms. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [13] Charith Lasantha Fernando, Masahiro Furukawa, Tadatoshi Kurogi, Sho Kamuro, Kouta Minamizawa, Susumu Tachi, et al. 2012. Design of TELESAR V for transferring bodily consciousness in telexistence. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5112–5118.
- [14] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. 2013. inFORM: dynamic physical affordances and constraints through shape and object actuation. In *Uist*, Vol. 13. 2501–988.
- [15] Danilo Gasques, Janet G Johnson, Tommy Sharkey, Yuanyuan Feng, Ru Wang, Zhuoqun Robin Xu, Enrique Zavala, Yifei Zhang, Wanze Xie, Xinming Zhang, et al. 2021. ARTEMIS: A collaborative mixed-reality system for immersive surgical telementoring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [16] Chad Harms and Frank Biocca. 2006. Internal Consistency and Reliability of the Networked Minds Social Presence Measure.
- [17] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [18] Zhenyi He, Ruofei Du, and Ken Perlin. 2020. Collabovr: A reconfigurable framework for creative collaboration in virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 542–554.
- [19] Zhenyi He, Fengyuan Zhu, and Ken Perlin. 2017. Physhare: Sharing physical interaction in virtual reality. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology*. 17–19.
- [20] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. 2015. Dollhouse vr: a multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Emerging Technologies*. 1–2.
- [21] Yunwoo Jeong, Han-Jong Kim, Gyeongwon Yun, and Tek-Jin Nam. 2020. WIKA: A projected augmented reality workbench for interactive kinetic art. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 999–1009.
- [22] Brennan Jones, Yaying Zhang, Priscilla NY Wong, and Sean Rintel. 2020. Vroom: virtual robot overlay for online meetings. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–10.
- [23] Brennan Jones, Yaying Zhang, Priscilla NY Wong, and Sean Rintel. 2021. Belonging there: VROOM-ing into the uncanny valley of XR telepresence. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–31.
- [24] Hiroki Kaimoto, Kyzyl Monteiro, Mehrad Faridan, Jiatong Li, Samin Farajian, Yasuaki Kakehi, Ken Nakagaki, and Ryo Suzuki. 2022. Sketched Reality: Sketching Bi-Directional Interactions Between Virtual and Physical Worlds with AR and Actuated Tangible UI. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–12.
- [25] Shunichi Kasahara, Ryuma Niizuma, Valentin Heun, and Hiroshi Ishii. 2013. exTouch: spatially-aware embodied manipulation of actuated objects mediated by augmented reality. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. 223–228.
- [26] Florian Kennel-Maushart, Roi Poranne, and Stelian Coros. 2023. Interacting with Multi-Robot Systems via Mixed Reality. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE.
- [27] Han-Jong Kim, Ju-Whan Kim, and Tek-Jin Nam. 2016. Ministudio: Designers' tool for prototyping ubicomp space with interactive miniature. In *Proceedings of the 2016 CHI Conference on human factors in computing systems*. 213–224.
- [28] Kangsoo Kim, Luke Boelling, Steffen Haesler, Jeremy Bailenson, Gerd Bruder, and Greg F Welch. 2018. Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 105–114.
- [29] Lawrence H Kim, Daniel S Drew, Veronika Domova, and Sean Follmer. 2020. User-defined swarm robot control. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [30] Lawrence H Kim and Sean Follmer. 2019. Swarmhaptics: Haptic display with swarm robots. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [31] Seungwon Kim, Gun Lee, Mark Billinghurst, and Weidong Huang. 2020. The combination of visual communication cues in mixed reality remote collaboration. *Journal on Multimodal User Interfaces* 14, 4 (2020), 321–335.
- [32] Seungwon Kim, Gun Lee, Weidong Huang, Hayun Kim, Woontack Woo, and Mark Billinghurst. 2019. 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*. 1–13.
- [33] Minoru Kojima, Maki Sugimoto, Akihiro Nakamura, Masahiro Tomita, Hideaki Nii, and Masahiko Inami. 2006. Augmented coliseum: An augmented game environment with small vehicles. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'06)*. IEEE, 6–pp.
- [34] Hideaki Kuzuoka, Shinya Oyama, Keiichi Yamazaki, Kenji Suzuki, and Mamoru Mitsuishi. 2000. GestureMan: A mobile robot that embodies a remote instructor's actions. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 155–162.
- [35] Mathieu Le Goc, Lawrence H Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, and Sean Follmer. 2016. Zoids: Building blocks for swarm user interfaces. In *Proceedings of the 29th annual symposium on user interface software and technology*. 97–109.
- [36] Jinha Lee, Rehmi Post, and Hiroshi Ishii. 2011. ZeroN: mid-air tangible interaction enabled by computer controlled magnetic levitation. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. 327–336.
- [37] Myungho Lee, Nahal Norouzi, Gerd Bruder, Pamela J Wisniewski, and Gregory F Welch. 2018. The physical-virtual table: exploring the effects of a virtual human's physical influence on social interaction. In *Proceedings of the 24th ACM symposium on virtual reality software and technology*. 1–11.
- [38] Min Kyung Lee and Leila Takayama. 2011. "Now, i have a body" uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 33–42.
- [39] Yujin Lee, Myeongseong Kim, and Hyunjung Kim. 2020. Rolling Pixels: Robotic Steinmetz Solids for Creating Physical Animations. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*. 557–564.
- [40] Daniel Leithinger, Sean Follmer, Alex Olwal, and Hiroshi Ishii. 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*. 461–470.
- [41] Jakob Leitner, Michael Haller, Kyungdahn Yun, Woontack Woo, Maki Sugimoto, and Masahiko Inami. 2008. IncrTable, a mixed reality tabletop game experience. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*. 9–16.
- [42] Jiannan Li, Mauricio Sousa, Chu Li, Jessie Liu, Yan Chen, Ravin Balakrishnan, and Tovi Grossman. 2022. ASTEROIDS: Exploring Swarms of Mini-Telepresence Robots for Physical Skill Demonstration. In *CHI Conference on Human Factors in Computing Systems*. 1–14.
- [43] Jiatong Li, Ryo Suzuki, and Ken Nakagaki. 2023. Physica: Interactive Tangible Physics Simulation based on Tabletop Mobile Robots Towards Explorable Physics Education. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 1485–1499.
- [44] Mark Marshall, Thomas Carter, Jason Alexander, and Sriram Subramanian. 2012. Ultra-tangibles: creating movable tangible objects on interactive tables. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2185–2188.
- [45] Thomas Muender, Anke V Reinschluessel, Sean Drewes, Dirk Wenig, Tanja Döring, and Rainer Malaka. 2019. Does it feel real? Using tangibles with different fidelities to build and explore scenes in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [46] Ken Nakagaki, Joanne Leong, Jordan L Tappa, João Wilbert, and Hiroshi Ishii. 2020. Hermits: Dynamically reconfiguring the interactivity of self-propelled tuis with mechanical shell add-ons. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 882–896.
- [47] Ken Nakagaki, Jordan L Tappa, Yi Zheng, Jack Forman, Joanne Leong, Sven Koenig, and Hiroshi Ishii. 2022. (Dis) Appearables: A Concept and Method for Actuated Tangible UIs to Appear and Disappear based on Stages. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [48] Hideyuki Nakanishi, Kei Kato, and Hiroshi Ishiguro. 2011. Zoom cameras and movable displays enhance social telepresence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 63–72.
- [49] Diana Nowacka, Karim Ladha, Nils Y Hammerla, Daniel Jackson, Cassim Ladha, Enrico Rukzio, and Patrick Olivier. 2013. Touchbugs: Actuated tangibles on multi-touch tables. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 759–762.
- [50] Kohei Ogawa, Shuichi Nishio, Kensuke Koda, Giuseppe Balistreri, Tetsuya Watanebe, and Hiroshi Ishiguro. 2011. Exploring the natural reaction of young and aged person with telenoid in a real world. *J. Adv. Comput. Intell. Intell. Informatics* 15, 5 (2011), 592–597.
- [51] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L Davidson, Sameh Khamis, Mingsong Dou, et al. 2016. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th annual symposium on user interface software and technology*.

- 741–754.
- [52] Eime Oyama, Kohei Tokoi, Ryo Suzuki, Sousuke Nakamura, Naoji Shiroma, Norifumi Watanabe, Arvin Agah, Hiroyuki Okada, and Takashi Omori. 2021. Augmented reality and mixed reality behavior navigation system for telexistence remote assistance. *Advanced Robotics* 35, 20 (2021), 1223–1241.
- [53] Eime Oyama, Motoki Yodowatari, Sousuke Nakamura, Kohei Tokoi, Arvin Agah, Hiroyuki Okada, and Takashi Omori. 2021. Integrating AR/MR/DR technology in remote seal to maintain confidentiality of information. *Advanced Robotics* 35, 11 (2021), 704–714.
- [54] Gian Pangaro, Dan Maynes-Aminzade, and Hiroshi Ishii. 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*. 181–190.
- [55] James Patten and Hiroshi Ishii. 2007. Mechanical constraints as computational constraints in tabletop tangible interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 809–818.
- [56] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. 2016. 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*. 1716–1725.
- [57] Thammathip Piumsomboon, Gun A Lee, Jonathon D Hart, Barrett Ens, Robert W Lindeman, Bruce H Thomas, and Mark Billinghurst. 2018. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.
- [58] Thammathip Piumsomboon, Gun A Lee, Andrew Irlitti, Barrett Ens, Bruce H Thomas, and Mark Billinghurst. 2019. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–17.
- [59] Ivan Poupyrev, Tatsushi Nashida, and Makoto Okabe. 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*. 205–212.
- [60] Iulian Radu, Tugce Joy, and Bertrand Schneider. 2021. Virtual makerspaces: merging AR/VR/MR to enable remote collaborations in physical maker activities. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–5.
- [61] Irene Rae, Bilge Mutlu, and Leila Takayama. 2014. Bodies in motion: mobility, presence, and task awareness in telepresence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2153–2162.
- [62] Majken K Rasmussen, Esben W Pedersen, Marianne G Petersen, and Kasper Hornbaek. 2012. Shape-changing interfaces: a review of the design space and open research questions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 735–744.
- [63] Dan Rosenfeld, Michael Zawadzki, Jeremi Sudol, and Ken Perlin. 2004. Physical objects as bidirectional user interface elements. *IEEE Computer Graphics and Applications* 24, 1 (2004), 44–49.
- [64] Daisuke Sakamoto, Takayuki Kanda, Tetsuo Ono, Hiroshi Ishiguro, and Norihiro Hagita. 2007. Android as a telecommunication medium with a human-like presence. In *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 193–200.
- [65] Mose Sakashita, Tatsuya Minagawa, Amy Koike, Ippei Suzuki, Keisuke Kawahara, and Yoichi Ochiai. 2017. You as a puppet: evaluation of telepresence user interface for puppetry. In *Proceedings of the 30th annual ACM symposium on user Interface software and technology*. 217–228.
- [66] Mose Sakashita, E Andy Ricci, Jatin Arora, and Fran ois Guimbret iere. 2022. RemoteCoDe: Robotic Embodiment for Enhancing Peripheral Awareness in Remote Collaboration Tasks. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–22.
- [67] J. Sauro. 2011. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Measuring Usability LLC.
- [68] Alexa F Siu, Shenli Yuan, Hieu Pham, Eric Gonzalez, Lawrence H Kim, Mathieu Le Goc, and Sean Follmer. 2018. Investigating tangible collaboration for design towards augmented physical telepresence. In *Design thinking research*. Springer, 131–145.
- [69] Ryo Suzuki, Hooman Hedayati, Clement Zheng, James L. Bohn, Daniel Szafrir, Ellen Yi-Luen Do, Mark D. Gross, and Daniel Leithinger. 2020. RoomShift: Room-Scale Dynamic Haptics for VR with Furniture-Moving Swarm Robots. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [70] Ryo Suzuki, Adnan Karim, Tian Xia, Hooman Hedayati, and Nicolai Marquardt. 2022. Augmented Reality and Robotics: A Survey and Taxonomy for AR-enhanced Human-Robot Interaction and Robotic Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–32. <https://doi.org/10.1145/1122445.1122456>
- [71] Ryo Suzuki, Jun Kato, Mark D Gross, and Tom Yeh. 2018. Reactile: Programming swarm user interfaces through direct physical manipulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [72] Ryo Suzuki, Eyal Ofek, Mike Sinclair, Daniel Leithinger, and Mar Gonzalez-Franco. 2021. HapticBots: Distributed Encountered-type Haptics for VR with Multiple Shape-changing Mobile Robots. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 1269–1281.
- [73] Ryo Suzuki, Abigail Stangl, Mark D Gross, and Tom Yeh. 2017. Fluxmarker: Enhancing tactile graphics with dynamic tactile markers. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. 190–199.
- [74] Ryo Suzuki, Clement Zheng, Yasuaki Kakehi, Tom Yeh, Ellen Yi-Luen Do, Mark D Gross, and Daniel Leithinger. 2019. Shapebots: Shape-changing swarm robots. In *Proceedings of the 32nd annual ACM symposium on user interface software and technology*. 493–505.
- [75] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. 2019. Loki: Facilitating remote instruction of physical tasks using bi-directional mixed-reality telepresence. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 161–174.
- [76] Ana M Villanueva, Ziyi Liu, Zhengzhe Zhu, Xin Du, Joey Huang, Kylie A Peppler, and Karthik Ramani. 2021. Robotar: An augmented reality compatible teleconsulting robotics toolkit for augmented makerspace experiences. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [77] Junichi Yamaoka and Yasuaki Kakehi. 2013. dePEND: augmented handwriting system using ferromagnetism of a ballpoint pen. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 203–210.
- [78] Lilith Yu, Chenfeng Gao, David Wu, and Ken Nakagaki. 2023. AeroRigUI: Actuated UIs for Spatial Interaction using Rigging Swarm Robots on Ceilings in Everyday Space. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [79] Xiangyu Zhang, Xiaoliang Bai, Shusheng Zhang, Weiping He, Peng Wang, Zhuo Wang, Yuxiang Yan, and Quan Yu. 2022. Real-time 3D video-based MR remote collaboration using gesture cues and virtual replicas. *The International Journal of Advanced Manufacturing Technology* (2022), 1–23.
- [80] Yiwei Zhao, Lawrence H Kim, Ye Wang, Mathieu Le Goc, and Sean Follmer. 2017. Robotic assembly of haptic proxy objects for tangible interaction and virtual reality. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. 82–91.
- [81] Zhengzhe Zhu, Ziyi Liu, Tianyi Wang, Youyou Zhang, Xun Qian, Pashin Farsak Raja, Ana Villanueva, and Karthik Ramani. 2022. MechARspace: An Authoring System Enabling Bidirectional Binding of Augmented Reality with Toys in Real-time. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–16.
- [82] Jakob Zillner, Christoph Rhemann, Shahram Izadi, and Michael Haller. 2014. 3D-board: a whole-body remote collaborative whiteboard. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 471–479.