IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 24, 2020, accepted December 6, 2020, date of publication December 10, 2020, date of current version December 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3043783

Hand-in-Air (HiA) and Hand-on-Target (HoT) Style Gesture Cues for Mixed Reality Collaboration

SEUNGWON KIM[1,2], ALLISON JING[1], HANHOON PARK[3], GUN A. LEE[1], (Member, IEEE), WEIDONG HUANG[4], AND MARK BILLINGHURST[1]

[1]School of Information Technology and Mathematical Sciences, University of South Australia, Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia
[2]Korea Institute of Science and Technology (KIST), Seoul 02792, South Korea
[3]Department of Electronic Engineering, Pukyong National University, Busan 48513, South Korea
[4]Faculty of Transdisciplinary Innovation, University of Technology Sydney, Sydney, NSW 2007, Australia

Corresponding author: Seungwon Kim (seungwon.kim@unisa.edu.au)

ABSTRACT In collaboration systems, collaborators often use hand gestures for showing information relating to a distant object. However, in viewer's perspective, it is sometimes difficult to know where the conventional gesture cue (Hands-in-Air style: HiA) refers to because the conventional HiA gesture appears away from the distant object. In this paper, we investigate how two factors, distance to the object and view angle difference between collaborators, influence the understanding of HiA gesture by comparing the use of it at 25 positions with 5 distances and 5 view angles. In a user study, we found that the distance to the target object and view angle difference negatively influenced HiA gesture communication. The influence of the distance and view angle was more serious between smaller angles (0∼30 degree) than larger angles (30∼60 degree), and between shorter distances (1m∼1.5m) than between longer distance (1.5m∼2m). As a solution, we propose the Hands-on-Target (HoT) style which positions the hand gesture cue on the surface of the target object. The HoT style gesture cue dramatically reduces the negative effect of the distance and view angle difference. Participants completed the task 29.3 percent faster, selected the correct object 2.75 times more, and felt 63.1 percent less mental effort. For further investigation, we discuss on the extensibility of the HoT interface that it can be used not only for object selection but also for diverse type of gesture communication including gesture for object manipulation because the HoT interface can support all possible real world hand gestures.

INDEX TERMS Augmented reality, collaboration, mixed reality, virtual reality.

I. INTRODUCTION

Hand gestures are a fundamental and natural communication tool [1] not only in the real world [1], but also in technology mediated collaborative environments such as Mixed Reality (MR) remote collaboration [2] and Collaborative Virtual Environments (CVEs) [3]. Many prior research [4], [5] have implemented hand gesture communication in CVE or MR collaboration in a way that is similar to how they are used in the real world by visualizing the hands at a relative position to the user's head and showing hand motions like in real world. This conventional Hands-in-Air (HiA) style hand gesture (Figure 1a and 1b) often provides realistic and natural interaction with objects [6], [7].

However, if using HiA gesture communication for a distant object, the viewer may have difficulty in understanding which object or location the hand gesture is referring to [8], [9] because the hand gesture appears away from the target object and is not aligned with it in the viewer's perspective (Figure 1c).

As a solution, Wong and Gutwin [10] used an arm extending mechanism adapted from Poupyrev's go-go technique [11], but their system used a rigid virtual hand model so did not support natural hand gesturing with hand and finger movement.

Recently, some researchers [4], [12], [13] implemented an interface supporting a pointer cue and HiA hand gestures so

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna Dulizia.

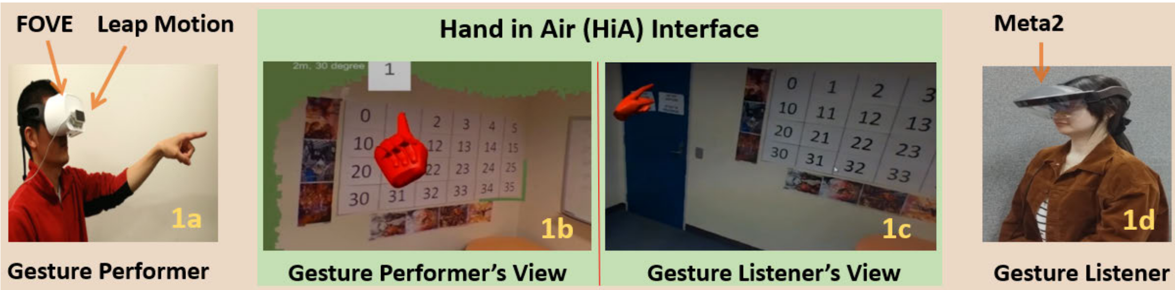


FIGURE 1. **The difficulty to understand gestures with the HiA interface. The gesture performer points at the number '1' (1b) and the Viewer should guess which number is being pointed at but from viewer's viewpoint it is ambiguous (1c). In our system, the performer wears a FOVE HMD with an attached Leap Motion (1a) and the viewer wears a Meta2 HMD (1d).**

collaborators used the pointer for indicating distant objects and hand gestures for other information. However, the pointer could not show the same rich communication information as hand gesture cues [14] and the hand gesture still could not reach to the distant objects. Similar to using a pointer, eye gaze also has been studied for indicating a distant object when a hand gesture cue is also available [15], [16], but eye gaze does also not show rich communication information like a hand gesture even if it was faster than hand gesture to indicate object.

Sousa *et al.* [8] and Mayer *et al.* [9] introduced another method for reducing the ambiguity in understanding hand gestures, by warping the gesture performer's arm to point at the target object from the viewer's perspective. However, the hand gesture in their system still appears away from the object so the viewer should need to guess where it refers to.

In this paper, we firstly investigate how two factors: the distance to the object and view angle difference between collaborators, influence the understandability of conventional hand gesture communication (i.e. HiA), then investigate whether an alternative Hands-on-Target (HoT) style, which moves the virtual hands onto the surface of the task object (Figure 2), can improve the understandability.

This research makes the following contributions:

(1) Investigating how two factors, the distance to the task object and the view angle difference between a gesture performer and a viewer, influences the understandability of HiA gestures
(2) Presenting a novel HoT technique in a collaborative virtual environment
(3) Conducting a user study to find out if the HoT technique performs better than the HiA under varying distances to the task object and view angle differences

In the following sections, we review previous work, describe the HiA and HoT interfaces and their implementation, present a user study design followed by the experiment results and discussion, then conclude with a summary and direction for future research.

II. RELATED WORK

Hand gesture communication has been studied in many MR interfaces for collaboration [17]. In this section, we review previous studies of hand gesture communication, then explain the research gap between these previous studies and our work.

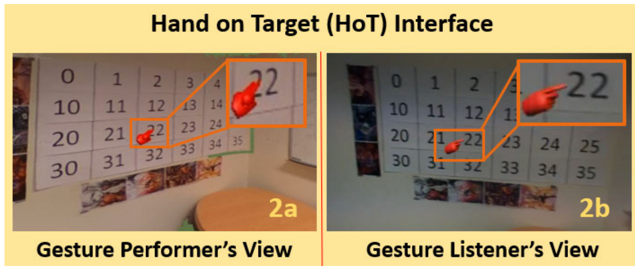


FIGURE 2. **Collaborators using the HoT interface. The gesture performer is pointing at the number '22' and the viewer should guess the number being pointed at.**

A. HAND GESTURE IN COLLABORATION SYSTEM

Hand Gesture interfaces in Virtual Reality (VR) and Augmented Reality (AR) environments allow the use of hand gestures as in the real world [18]. Hand gestures include fingers, hand and arm movements and are considered a natural and intuitive way to communicate [19], [20]. Using hand gesture for providing information while collaborating reduces the workload [21] and increases the collaborator's feeling of co-presence compared to a pointer interface [22].

In early remote collaboration system, Kirk *et al.* [23] used a simple method of sharing hand gestures: 1) sharing and projecting a live video of the task space on gesture performer's desk, 2) performing hand gestures on top of the task space live video, and 3) taking and sharing another live video with the hand gesture and sending back to viewer's side. Later, Alem *et al.* [24] adopted a similar mechanism but provided portability on the gesture viewer side by taking a live video of task space with a head own camera and displaying another live video of the hand gesture on a near eye display.

While early works [23], [24] included two live videos between gesture performer and viewer, some researchers only shared one live video by extracting and integrating the hand gestures into the task space live video [25]–[27]. In their study, participants could complete the task faster with hand gestures compared to a voice only condition [27], and had better awareness of which object the gesture performer was discussing about with the hand gestures [25]. However, there was no depth perception with the 2D video.

Researchers have started investigating 3D hand gesture cues in a shared 3D task space [5], [28], [29]. With the help of depth sensing cameras such as Leap Motion or Microsoft Kinect, 3D virtual hand gestures look almost identical to real hand gestures. Huang *et al.* [5] found that using 3D virtual hands improved the level of co-presence and participants had better spatial relations in the task space and a more immersive experience. Gao's system [28], [30] used 3D point-cloud data for sharing the 3D task space and capturing 3D gestures. In their pilot study, sharing a 3D task space and gesture increased the naturalness of using the hand gesture. In this paper, we call this conventional 3D hand gesture cue as a HiA style hand gesture because it is mostly in the air and displayed relative to the gesture performer's head and body depending on the hand tracking results.

With the HiA style hand gesture, we conducted two user studies [4], [14] and found that the use of the HiA is powerful if the task space is small (i.e. a desk) or collaborators have an identical view. With a small task space, task objects are mostly in hand reachable distance and the HiA hand gesture is near to the task object in the viewer's perspective, so it was easy to know what or where the hand gesture refers to. With an identical view, a viewer looks at the hand gesture from the same perspective as the gesture performer, so the performer's hand gesture is aligned with a target object in the viewer's perspective, regardless of how far away the task object is.

However, in recent collaboration systems, each collaborator mostly has a first-person view and they have different viewpoint positions and perspective rather than identical view in a large task space [4], [17]. In the second study of our previous work [4] where both the gesture performer and the viewer had different perspective views and the performer was away from the task objects in a large task space, the HiA hand gesture was difficult to use. In this paper, we confirm whether the view angle difference between collaborators and distance to objects are the factors influencing the use of hand gesture.

B. PREVIOUS SOLUTIONS

In previous studies, there were three main approaches to address the issue of the HiA interface. First, using pointer, gaze, or/and sketch cues rather than using HiA gestures. The pointer [31], gaze [32] and sketch [33] cues are displayed on the surface of the task objects and the gaze ray shows a line to the object [15], [32], so a viewer could easily know which object they are referring to. Additionally, the pointer cue is simple and can show a precise point information [31]. Gaze can move faster than hand gestures [16], and does not require any additional effort while hand gesture does (e.g. performing the required gesture), because a collaborator uses continual gaze activity to look around the task space anyway [38].

However, the pointer, eye gaze, or sketch cues cannot be prefect substitutes for hand gesture communication, because they have different characteristics compared to the hand gesture cues and have different effects on collaboration [4], [27]. Hand gesture cues look similar to the real hand gesture, so hand gesture cues provide more natural and realistic interaction compared to a pointer, gaze, or sketch [6], [7]. Hand gesture can show diverse types of information such as pointing, shape of an object, direction, volume (put hands apart according to the size of volume), and even social cues which could be difficult to present with pointer and sketch cues [39]. Additionally, in a previous remote collaboration study [14], participants preferred to use the hand gesture cue more than the pointer or sketch cues. Therefore, using pointer or sketch cues instead of the hand gesture communication may be not recommended.

Second, Sousa *et al.* [8] and Mayer *et al.* [9] introduced an interface to reduce the ambiguity in understanding where a hand gesture is referring to, by warping the gesture performer's arm to point at the target object from the viewer's perspective. This may help the viewer to easily guess which object the hand is pointing at. However, their approach has drawbacks such that the viewer still needs to continuously switch his/her focus between the pointing hand and the task object and still need to guess where it is pointing to.

The third approach to solve the issue of the HiA interface was extending arm technique. Extending arm and translating the virtual hand near to the target object was originally for a single user application to easily select and manipulate distant objects [41] and named Go-Go [11] and HOMER techniques [40]. Hindmarsh *et al.* [42], [43] and Wong and Gutwin [10] adopted the extending virtual arm technique for the gesture communication between collaborators, but their systems did not support hand and finger movement and so limited the use of hand gesture. They also did not explore the influence of distance to the task object and view angle difference between collaborators.

In this paper, we firstly investigate and confirm the influence of the distance and view angle difference in using the HiA interface, then secondly explore how our HoT technique can improve understanding of hand gesture cues.

C. HYPOTHESES

Based on literature review, we formulated the following hypotheses:

- **H1** The distance to the task object and the view angle difference between a hand gesture performer and viewer are the factors negatively influencing on the use of the HiA gesture.

To solve the issue in HiA, we suggest a HoT technique and formulate three more hypotheses:

- **H2** Collaborators have better hand gesture communication with the HoT interface than with the HiA interface.
- **H3** The HoT interface overcomes the negative impact of the distance to task objects.
- **H4** The HoT interface overcomes the negative impact of the view angle difference.

III. METHODOLOGY

To test the hypotheses, we conducted a user study with a MR remote collaboration system that supports both HiA and HoT interfaces for a hand gesture communication. The MR remote collaboration is a type of mediated collaboration between a local worker and a remote helper (a.k.a. remote exert) to solve a physical task [17], [44], [45]. The local worker is the person viewing the hand gesture to get a help from a remote helper and the remote helper is the person performing the hand gesture to help the local worker [46], [47]. The local worker has an AR view to see the physical task space overlaid with the remote helper's hand gestures while the remote expert has a VR view with a 3D reconstruction of the real task space [2].

A. SYSTEM DEVELOPMENT

We developed an MR remote collaboration prototype system for investigating different styles of hand gesture cues including HiA and HoT.

1) HARDWARE DEVICE

The prototype system used two computers with the same specification for the local worker and remote helper: Intel Core i7-7700K 4.2GHz quad core CPU, 16 GB RAM, and NVIDIA GeForce GTX 1070 graphics card. The computers were connected each other by a local area network via Ethernet cable. A Meta2[1] optical see-through AR HMD (Figure 1d), was worn by the local worker and connected to the local worker's computer. The Meta2 has a 90-degree FOV and 2550 × 1440 pixel resolution with 60 frames per second (fps) refresh rate. It supports real-time 3D mesh reconstruction and SLAM (Simultaneous localization and mapping) visual tracking [48]. Since the Meta2 does not provide textures on the reconstructed 3D mesh, we employed a 360-degree camera, the Ricoh Theta V[2] (Figure 3), to capture and map the real-time live texture onto the 3D mesh.

In the remote helper's system (remote system afterward), the FOVE VR display[3] (Figure 1a) was employed. This has a 100-degree FOV with a resolution of 2560 × 1440, and updated at 70 fps. It supports tracking user's head motion with an infrared camera and inertial measurement units consisting of a tri-axis gyroscope and accelerometer, so the FOVE display supports the remote helper's navigation according to his/her head movement. Since the FOVE does not support hand tracking, we attached a Leap Motion[4] hand tracker onto the front face of it (Figure 1a). The Leap Motion tracks hands within a 150-degree FOV at 300 fps with an accuracy of 0.7 mm [49].

[1]Meta2 augmented reality head mounted display website. https://www.metavision.com/. Accessed: 2020-07-28

[2]Ricoh Theta V. https://theta360. com/en/about/theta/v.html, Accessed: 2019-07-08

[3]FOVE head mounted display website. https://www.getfove.com/. Accessed: 2020-07-28

[4]LeapMotion. Leap Motion hand tracking website. https://www. leapmotion.com/. Accessed: 2020-07-28

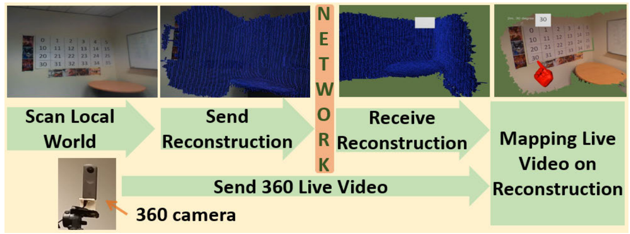


FIGURE 3. Sharing the task space view. A local worker scans the task space and sends its 3D reconstruction to the remote system. A live 360-degree video of the local task space from a 360 camera, Theta V, is also shared with the remote system and texture mapped on to the 3D reconstruction.

2) SHARING THE LOCAL TASK SPACE

All software development was done in the Unity 2019.3.0f3 game engine. To share the local task space, we exploited 3D reconstruction example code from the Meta2 Unity SDK 2.6. With the reconstruction function, a local worker scanned the local task space and created the 3D reconstruction, which was then sent to the remote system (Figure 3). Since the remote system was running the FOVE and Meta2 Unity SDKs, the remote system could render the received reconstruction in the FOVE view. The size and position of the shared 3D mesh reconstruction were the same for the local and remote systems, but the local reconstruction was transparent allowing the local worker to directly see the task space, while the remote helper saw the reconstruction.

Since the Meta2 Unity SDK 2.6 does not update the 3D mesh reconstruction, we used a live video texture to show real-time updates. The system shared the live video texture from a 360-degree camera, the Ricoh Theta V (Figure 3) and the live video was mapped onto the 3D reconstruction. To map the live video, we performed a manual registration process to find the position and orientation of the 360 camera in the Meta2 coordinate system by positioning a virtual circle at the lens of the camera and using its position as the centre of a spherical projective texture mapping which we implemented as a custom Cg[5] shader.

Within the shared 3D view, the remote helper and local worker could have an almost identical task space and navigate around with the help of the head tracking function which was provided by the Meta2 and FOVE SDKs. To help the local worker and remote helper understand where each other was, we added a virtual head representing the partner's head in the shared task space.

3) HiA AND HoT INTERFACES

The remote system tracked the remote helper's hands and shared the tracked hand data with the local system. To implement the HiA interface, we customized the Leap Motion Unity example scene.[6] The remote system tracked the remote helper's hands in real time and shared the hand tracking frame

[5] NVIDIA Cg Toolkit https://developer.nvidia.com/cg-toolkit

[6]Leap Motion example. 2014. Leap Motion Pinch Draw Demo. https://gallery.leapmotion.com/pinch-draw/. Accessed: 2020-07-28

with the local system. Then, both systems virtually rendered the hand gesture on both system views (Figure 1b and 1 c). In the HiA interface, the position and shape of the virtual hands were relative to the remote helper's head position with the results of the tracking hand (left picture of Figure 4), because the Leap Motion was attached to the front of the remote helper's FOVE HMD.

Our system has options to use pointer and/or sketch cues when using HiA, with the pointer and sketches placed on the task objects. However, the use of pointer and sketch cues was investigated in our user study because they have different effects on remote collaboration compared to the hand gesture cue [14], [27] (as described in section 2.B), and our study solely focused on sharing of hand gesture cues.

To implement the HoT interface (Figure 2), we extended the HiA interface to support translating the virtual hands onto the surface of the task object. To translate the virtual hand, we used a ray casting method and defined the ray casting direction from remote helper's head position to the tracked hand position (see right picture of Figure 4). The point where the ray casting intersected with the 3D mesh reconstruction (which was used for sharing the task space view) was set as the new position of the virtual hands working as a HoT interface. Since our system controlled two hands separately, the remote helper could use both left and right hands. Our HoT interface could still perform all types of hand gestures including pointing gesture, showing the shape of an object, direction, size (e.g. putting hands apart according to the size), required hand operation (e.g. a hand gesture turning clockwise to show how to open a valve) and even social cues (thumb up). Therefore, our HoT interface can convey all possible information that the conventional hand gesture cue can show, while it appears near to the object hence it looks smaller from a distance.

B. USER STUDY DESIGN

With the system, we conducted a user study. In this section, we describe the study conditions, task, and data collection.

1) STUDY SETUP

To investigate the influence of the distance to the task objects and view angle difference between remote helper and local worker, we prepared a room-size task space with task objects placed on a wall. We prepared five distances and five view angles, so there were 25 positions, from where a remote helper views the task objects (the blue circles in Figure 5). In this study, the remote helper physically stayed in the next room while wearing the FOVE HMD to look at the experimental space from one of the 25 positions. The height of the remote helper's view (i.e. the height of the 360 camera) was fixed at 175 cm to avoid the variance in height becoming a confounding factor.

We define these 25 viewing positions by five distances from the center of the task objects (1m, 1.25m, 1.5m, 1.75m, and 2m away) and five view angles (0, 15, 30, 45, 60 degrees rotated clockwise from a perpendicular line from the center of task objects to the local worker's position). While the position of the remote helper varied among the 25 positions, the local worker always stood at the same point which was 2m away perpendicularly (0 degree) from the center of the task objects (the orange circle in Figure 5). In the rest of the paper, we refer to these 25 remote helper positions as the 'experimental positions'.

By placing the local worker and remote helpers at a set of fixed positions, we could reduce the effect from users navigating around that is a confounding factor, and clearly investigate the influence of the distance and view angle that are systematically defined. Additionally, we can remove the risk of remote helpers bumping into the wall or other real-world objects while wearing the FOVE HMD. We note that in our user study, even though the remote helper's view position was fixed, the remote helper and local worker could rotate his/her view by turning their head or body.

2) TASK

Since our focus was on how well the local worker knew where the remote helper's hand gesture referred to, we prepared an object selection task (Figure 5) in which a local worker called out the object that the remote helper's hand gesture referred to. To complete the study with 25 experimental positions in a timely manner, the task did not include complex object manipulation. We note that object manipulation could not be performed without proper object selection and without knowing where the remote helper's hand gesture referred to.

FIGURE 4. Remote helper's hand gesture position with HiA and HoT interfaces.

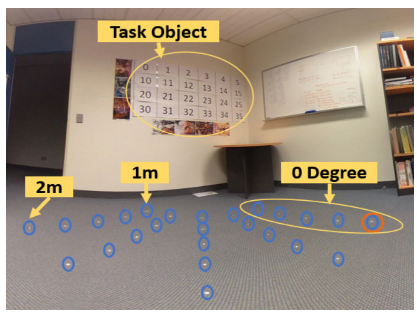


FIGURE 5. Experimental positions of the remote and local uses in the task space. The 25 white dots in the blue circles represent the remote user's virtual positions during the user study. The white dot in an orange circle is the point where a local user stays.

In preparation of the selection task, we considered the size and shape of the task object which may have affected the study results. Thus, we prepared objects with the same size and shape, i.e. 24 A4 papers with written numbers on them (Figure 5). To reduce the effect of the remote helper's object searching activities (which is not our focus), we decided to use a simple numeric coordinate system to label each task object location. The objects are arranged into a matrix of 4 rows and 6 columns, and the columns were numbered 0 to 5, from left to right, and the rows were numbered 0 to 3, from top to bottom. Combining the numbers of row and column we could identify task objects, for example object 23 was at row 3 column 4 (Figure 5).

We systematically prepared a sequence of selection tasks for the experimental task. Each pair of participants performed 600 selections in total (2 conditions $\times$ 25 experimental positions $\times$ 12 selections per condition at each position $=$ 600 selections). To maintain a similar level of task difficulty between the HiA and HoT conditions, we grouped 24 numbers into two groups (Group1: 0, 2, 4, 11, 13, 15, 20, 22, 24, 31, 33, 25; Group2: 1, 3, 5, 10, 12, 14, 21, 23, 25, 30, 32, 34) and had a similar level of object distribution between the two groups. During the experiment, we randomly shuffled the order of the target object for selection. Alternating the assigned object group in each condition across pairs of participants, all target numbers in the task space were selected once at each experimental position.

The participants in the user study were not allowed to correct their answer after an incorrect selection to complete the experiment in a timely manner (no more than 2 hours), and to reduce the effect of verbal correction (e.g. saying "two columns right from the number you incorrectly chose"). Additionally, the remote helper could not tell any number to encourage the use of hand gesture.

3) PROCEDURE AND DATA COLLECTION

We measured the task completion time, the number of correct selections where the remote helper's hand gesture referred to, and the participants' mental effort (SMEQ) [54], [55]) at each experimental position with each hand gesture condition. Additionally, after using each condition at all 25 experimental positions, we also collected participants' ratings from questionnaires asking about workload (NASA-TLX [50]), co-presence [51], understanding messages [51], realism (IPQ[7] [52]), and hand ownership [53].

The experiment started with a pair of participants signing a consent form and answering a demographic questionnaire asking about their age, gender, and the level of familiarity with VR/AR systems and hand gesture interaction. We randomly assigned the roles of the remote helper and local worker and explained the rules in the experiment to the participants (e.g., not being allowed to verbally communicate target numbers). We let them practice collaboration for five number selections face-to-face. This was to ensure that the participants understood the task before testing the conditions.

We separated the participants into two adjacent rooms connected through an open door, so they could talk to each other but could not see each other. Then, we explained and prepared the system to let them know how to use HoT and HiA interfaces. The local participants wore a Meta2 AR HMD, to see the remote participants' hand gesture in the task space. The remote participants wore a FOVE VR HMD, to see the task space showing the target numbers. When the system was ready, the participants tried out both conditions to get familiar with them before starting the experimental trials. We then started the experimental trial with the conditions in a counter balanced order.

In each condition, remote participants had 25 experimental positions and selected 12 numbered task objects at each experimental position. The target number was shown at the top of the remote helper's display (Figure 1b), and the remote helper indicated the number with a hand gesture and verbal explanation without telling the target number. The order of the experimental positions was randomised. After selecting 12 targets at each experimental position, we measured the task completion time and participants' mental effort (SMEQ) [54], [55]. For each condition, participant took around 25 minutes for the task.

After trying each condition with 25 experimental positions, both the local and remote participants answered questions in the questionnaire asking about the task load level (using NASA-TLX [50]), co-presence [51], understanding messages [51], realism (using two customized questions from IPQ [52]; (1) How real did the virtual hands seem to you?, (2) How much did your experience with the virtual hands seem consistent with your real world experience?), and hand ownership (three customized questions from [53]; (1) I felt as if I was looking at my (or partner's) own hand (2) I felt as if the virtual hand was part of my (or partner's) body, (3) I felt as if the virtual hand was my (or partner's) hand). All questions were answered on a 7-point Likert rating scale from 1 (strongly disagree) to 7 (strongly agree).

After finishing both conditions, participants chose the best and worst conditions according to their preference between HiA and HoT conditions, and answered two open-ended questions: what do you like and what do you dislike in using the HiA and HoT? Overall, the experiment took about 90 minutes.

4) PARTICIPANTS

We recruited 28 participants in 14 pairs: 15 male and 13 female university students with their ages ranging from 19 to 45 years old ($M = 26.2$, $SD = 8.3$). Participants had slightly below average level of familiarity on using hand gesture interaction ($M = 3.4$, $SD = 1.8$) and VR/AR environment ($M = 3.8$, $SD = 2.1$), rated on a 7-point rating scale (1: Novice – 7: Expert).

[7] igroup presence questionnaire (IPQ), http://www.igroup.org/pq/ipq/download.php, Accessed:2019-07-28

IV. USER STUDY RESULTS

The task completion time and number of correct selections were not normally distributed according to a Shapiro-Wilk test, and the other collected data were in an ordinal scale (Likert-scale ratings). So, we used the Wilcoxon signed rank test ($\alpha = .05$) to compare the HiA and HoT conditions, and the Friedman test ($\alpha =.05$) for comparing between five distances or five view angles. For the results showing a significant difference from the Friedman test, we ran post hoc tests for pair-wise comparison using the Wilcoxon signed rank tests with the Bonferroni correction applied ($\alpha = .005$).

The main results are summarized as below:

- With the HiA condition, the view angle difference between the local worker and remote helper and the distance to the task object negatively impacted understanding the hand gesture cue.
- With the HiA condition, the influence of the angle difference was more obvious between smaller angles (0, 15, and 30 degrees in our study) than between larger angles (30, 45, and 60 degrees)
- With the HiA condition, the influence of the distance was more serious between shorter distances (1m, 1.25m, and 1.5m in our study) than between longer distances (1.5m, 1.75m, and 2m).
- Compared to the HiA condition, the HoT condition showed significant benefits in task completion time, number of correct selections, understanding hand gestures, and higher user feeling of co-presence and reality with less task load and mental effort. However, there was no significant benefit of the HoT condition in local worker's feeling of the co-presence and hand ownership.
- Even though the HoT condition significantly reduced the negative impact of the view angle difference compared to the HiA, the view angle difference still influenced the HoT style gesture communication.
- The remote participants felt that they spent more mental effort not only in using the hand gesture cues, but also in looking at the objects from the side (with 45 and 60-degree views) rather than in front of them (with 0 and 15-degree views).

A. PERFORMANCE OF THE HiA INTERFACE

In this section, we compare the use of the HiA interface in various distances and angle differences with the data of task completion time, the number of correct selections, and the required mental effort (SMEQ) [54] (Figure 6 and Table 1). Friedman tests found significant differences in all measurements between distances and degrees (all *p*-values are less than.001; Table 1).

1) EFFECT OF ANGLES

The participants playing a role of local worker (local participant afterward) understood the HiA interface better when having 0 or 15 degree angle difference with remote participants compared to when having the other angle difference (Figure 6 and Table 1). When there was no angular difference (0 degree angle) between the local and remote participants, local participants felt that it took less mental effort to understand where the HiA referred to compared to when using it at the other angles (four left columns in Table 1; 0 – 15, 0 – 30, 0 – 45, 0 – 60 columns). The results of task completion time and the number of correct selections also showed similar trends with significances (except between 0 and 15 degree in task completion time).

Using the HiA at 15-degree angle difference had similar results. Local participants felt less mental effort in understanding where the HiA referred to at a 15-degree angle than at 30, 45, and 60 degree angles (Table 1; 15 – 30, 15 – 45, 15 – 60 columns). The same trend was also revealed in the result of the task completion time, but not with the result

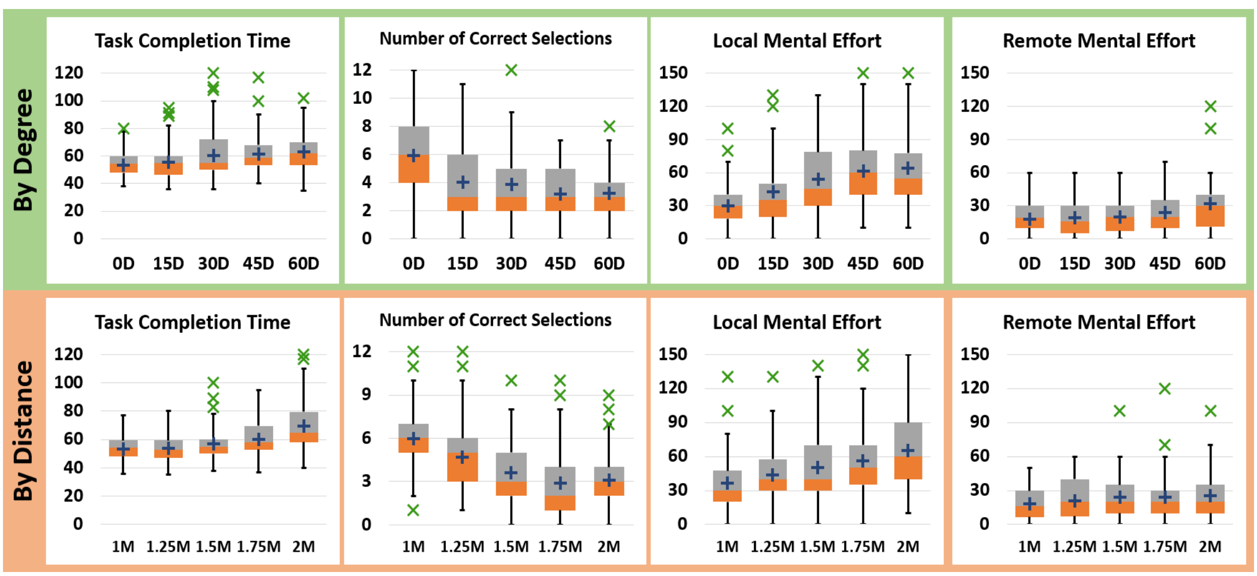


FIGURE 6. HiA Results: the average task completion time (seconds), the number of correct selections (0∼12), and the required mental effort by the local and remote participants (0: lowest∼150: highest mental effort) for each distance and degree (+: mean; x: outlier; D: degree).

TABLE 1. Results of the HiA condition by the Friedman tests and Wilcoxon signed rank tests (W in the table) with the data of task completion time, number of correct selections, local and remote participants' mental effort (significant results are in grey).

			Angle Difference in Degree									
Measurement	**Friedman Test**	***W***	**0 - 15**	**0 – 30**	**0 – 45**	**0 – 60**	**15 – 30**	**15 – 45**	**15 – 60**	**30 – 45**	**30 – 60**	**45 - 60**
Completion Time	$\chi^2(4) = 61.061$,	*Z*	-2.187	-3.964	-5.725	-5.947	-2.807	-4.673	-4.575	-1.484	-1.987	-0.925
	$p<.001$	*p*	.029	.000	.000	.000	.005	.000	.000	.138	.047	.355
Correct Selections	$\chi^2(4) = 59.159$,	*Z*	-4.510	-4.515	-5.966	-5.810	-0.329	-2.884	-2.453	-2.585	-1.878	-0.173
	$p<.001$	*p*	.000	.000	.000	.000	.743	.004	.014	.010	.060	.863
Local Mental Effort	$\chi^2(4) = 126.425$	*Z*	-4.163	-6.301	-6.761	-6.871	-3.792	-5.651	-6.188	-2.628	-3.358	-0.623
	$p<.001$	*p*	.000	.000	.000	.000	.000	.000	.000	.009	.001	.534
Remote Mental Effort	$\chi^2(4) = 69.874$,	*Z*	-0.860	-1.140	-4.078	-5.890	-0.345	-3.565	-5.119	-3.105	-5.884	-4.342
	$p<.001$	*p*	.390	.254	.000	.000	.730	.000	.000	.002	.000	.000
			Distance (in meter)									
Measurement	**Friedman Test**	***W***	**1 – 1.25**	**1 – 1.5**	**1 – 1.75**	**1 – 2**	**1.25–1.5**	**1.25–1.75**	**1.25 – 2**	**1.5–1.75**	**1.5 – 2**	**1.75 – 2**
Completion Time	$\chi^2(4) = 84.960$,	*Z*	-0.056	-2.661	-4.926	-6.497	-3.040	-4.817	-6.543	-2.611	-6.068	-4.654
	$p<.001$	*p*	.955	.008	.000	.000	.002	.000	.000	.009	.000	.000
Correct Selections	$\chi^2(4) = 85.776$	*Z*	-4.457	-5.820	-6.366	-6.127	-3.527	-4.748	-4.518	-2.370	-2.253	-0.761
	$p<.001$	*p*	.000	.000	.000	.000	.000	.000	.000	.018	.024	.447
Local Mental Effort	$\chi^2(4) = 88.133$,	*Z*	-3.379	-4.522	-5.728	-6.282	-2.878	-4.302	-5.850	-2.327	-4.772	-3.268
	$p<.001$	*p*	.001	.000	.000	.000	.004	.000	.000	.020	.000	.001
Remote Mental Effort	$\chi^2(4) = 27.212$,	*Z*	-2.407	-4.168	-3.932	-3.905	-2.378	-1.182	-2.185	-0.007	-0.906	-1.162
	$p<.001$	*p*	.016	.000	.000	.000	.017	.237	.029	.995	.365	.245

of number of correct selections, especially with the 30 and 60 degree angles.

The effect of the angle difference became less obvious at larger angle differences. The task completion time and the number of correct selections did not show a significant difference among the 30, 45, and 60 degree angles (Table 1; 30 – 45, 30 – 60, 45 – 60 columns). The local participants' mental effort in using the HiA had significant differences only between 30 and 60 degree angle differences.

Interestingly, the results of the remote participants' mental effort showed a different trend than the results of other measurements (task completion time, number of correct selections, and mental effort of local participants). All comparisons against either 45 or 60-degree angles showed significant differences in the remote participants' mental effort. This may be because the remote participants were looking at the objects from the side rather than in front of them.

Additionally, we also found that the numbers on left three columns were called more than the numbers on right three columns at all angles (all p-values are less than.005; Table 1 in Table 6 and 7) except 0-degree angle ($p = .138$). This may be because the remote participants' experimental positions were on the left side of the local participant except for the experimental positions at 0-degree angle, and used the HiA from the left. (Figure 5).

2) THE EFFECT OF DISTANCE

As the distance to the task object becomes longer, the local participants spent more time and needed more mental effort to understand which object the HiA hand referred to and had fewer correct selections.

The difference in using the HiA at five distances is clear in the results of the task completion time and local participants' mental effort (Figure 6 and the bottom of Table 1). All pair-wise comparisons showed significance in the results of the task completion time except the three comparisons between 1m and 1.25m, 1m and 1.5m, and 1.5m and 1.75m (we note the p-value is.008 and.009, close to the significant level, in the comparisons between 1m and 1.5m and between 1.5m and 1.75m respectively, Table 1). All comparisons with the data of local participants' mental effort showed significant differences except between 1.5m and 1.75m.

Interestingly, in the results of the number of correct selections, a significant difference between distances was only revealed in the pair-wise comparison with either 1m or 1.25m, but not among the three longer distances (1.5m, 1.75m, and 2m). Thus, the influence of the distance was more obvious between the shorter distances (1m, 1.25m, and 1.5 m) than between the longer distances (1.5m, 1.75m, and 2m).

Interestingly, the results of the remote participants' mental effort showed a different trend than the results of other measurements. The results of the remote participants' mental effort showed significant differences only between 1m and 1.5m, 1m and 1.75m, and 1m and 2m distances.

B. COMPARING HiA AND HoT INTERFACES

In this section, we compare HoT and HiA conditions by every distance and view angle difference first, then compare them with participants' rating results from questionnaires.

With the HoT condition, participants completed the task 29.3 percent faster (58.79 → 41.57 seconds), had 2.75 times more correct selections (33.75% → 93%), and felt 63.1 percent less mental effort (22.4 → 8.48) compared to when using the HiA condition. These benefits of the HoT condition were revealed at every distance and view angle difference (all p-values are less than 0.001). In addition,

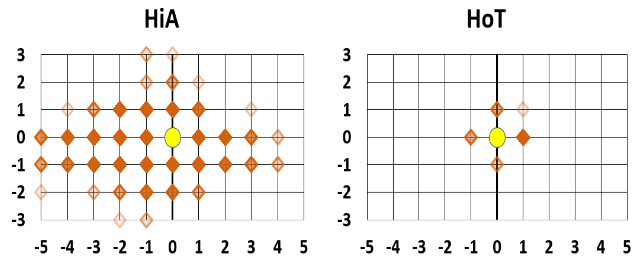


FIGURE 7. Distribution of incorrect selections with the HiA and HoT interfaces. The number at each axis means the difference (in row-column) between the target and the incorrect selection. The dot transparency is set by the frequency of the incorrect selections.

TABLE 2. Results of Wilcoxon signed ranks tests in comparing the HoT and HiA methods with the data of workload, co-presence, understanding hand gesture, reality and hand ownership (significant results are in grey).

	Remote Helper			Local Worker		
Measurement	*Z*	*p*	*r*	*Z*	*p*	*r*
Workload	-2.727	.006	.557	-3.296	.001	.673
Co-presence	-2.064	.039	.421	-0.210	.834	.043
Understanding Hand Gesture	-3.301	.001	.674	-3.301	.001	.674
Reality	-2.669	.008	.545	-2.663	.008	.544
Hand Ownership	-2.210	.027	.451	-1.086	.278	.222

with the incorrect selections, we calculated how far the local participants misunderstood the remote participants' hand gesture selection (Figure 7 and Table 4 and 5) and the results showed that the incorrect selections were significantly farther from the correct selection with the HiA condition than with the HoT condition at every distance and view angle (all *p*-values were less than 0.001).

According to the results of the participants' ratings (Table 2 and Figure 8), the remote participants felt less workload, more co-presence with the local partner, higher realism and hand ownership, and their hand gesture messages were better understood with the HoT condition than with the HiA condition (all *p*-values were less than 0.05). Local participants also felt less workload, had better understanding of hand gesture messages, and better realism with the HoT than with the HiA (all *p*-values were less than 0.05).

However, interestingly, the local participants felt no significant difference between the two conditions in the level of co-presence ($Z = -0.210$, $p = .834$) and hand ownership ($Z = -1.086$, $p = .278$). We found that the mean of the local participant's ratings about co-presence (HoT – *M*:4.85, *SD*:1.37; HiA – *M*:4.81, *SD*:1.02) and hand ownership (HoT – *M*:4.54, *SD*:1.92; HiA – *M*:4.07, *SD*:1.35) were higher with the HoT than with the HiA, so we may say that using the HoT did not reduce the local participants' feeling of co-presence and partner hand ownership compared to when using the HiA.

In their preference, all remote and local participants preferred the HoT interface over the HiA interface.

C. PERFORMANCE OF THE HoT INTERFACE

In this section, we compare the use of HoT interface in various distances and angle differences with the data of task completion time, number of correct selections, and required mental effort (SMEQ) [54] (Figure 9 and Table 3).

According to Friedman tests, the results with the data of all measurements in distances were not significant (all *p*-values are higher than.05). Thus, the distance to task object did not have a significant negative impact on understanding the remote helper's gesture with the HoT interface. However, among the five angles, the Friedman test showed significant results (Table 3).

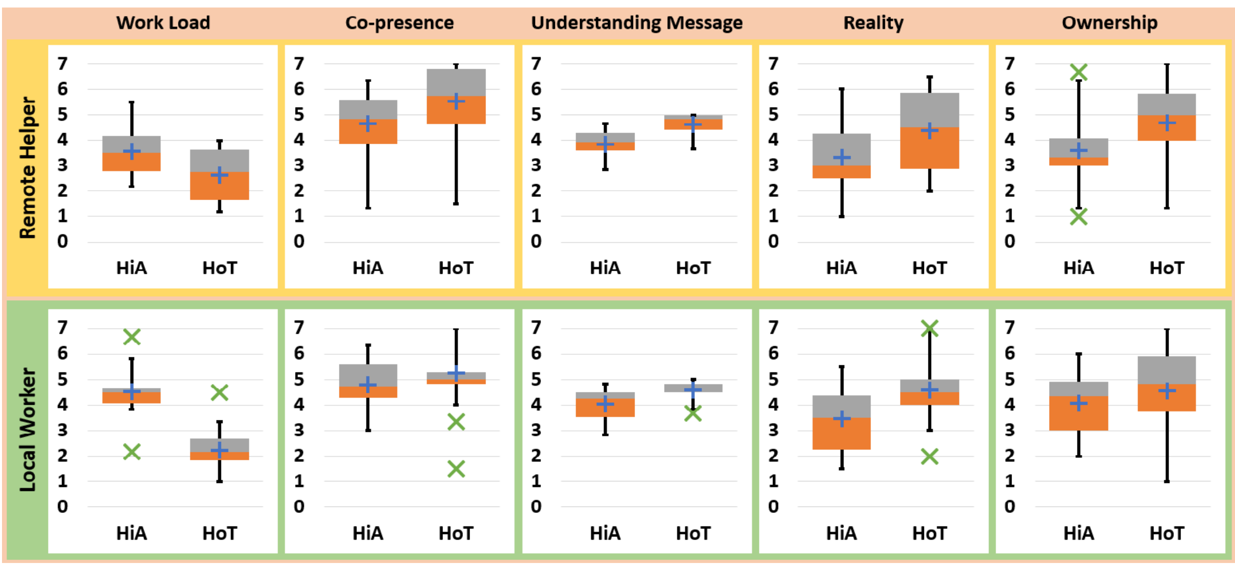


FIGURE 8. Questionnaire results asking about workload, ease of understanding the hand gesture message, level of co-presence, reality, and ownership when using the HiA and HoT interfaces (1: lowest~7: highest for all factors; +: mean; x: outlier).

TABLE 3. Results of the HoT condition by the Friedman test and Wilcoxon signed rank test with the data of task completion time, number of correct selections, local and remote participants' mental effort (significant results are in grey).

			Angle Difference in Degree									
Measurement	Friedman Test	W	0 - 15	0 – 30	0 – 45	0 – 60	15 – 30	15 – 45	15 – 60	30 – 45	30 – 60	45 - 60
Completion Time	$\chi^2(4) = 32.656$,	Z	-2.781	-2.800	-1.793	-2.684	-0.098	-1.204	-4.831	-1.324	-4.958	-4.266
	$p<.001$	p	.005	.005	.073	.007	.922	.229	.000	.186	.000	.000
Correct Selections	$\chi^2(4) = 113.285$,	Z	-3.584	-4.423	-4.622	-5.117	-1.882	-0.959	-5.538	-0.775	-6.035	-6.046
	$p<.001$	p	.000	.000	.000	.000	.060	.338	.000	.439	.000	.000
Local Mental Effort	$\chi^2(4) = 12.353$,	Z	-1.023	-1.282	-0.102	-1.316	-0.466	-0.922	-2.056	-2.236	-3.146	-1.662
	$p=.015$	p	.306	.200	.919	.188	.641	.357	.040	.025	.002	.097
Remote Mental Effort	$\chi^2(4) = 37.678$,	Z	-1.333	-0.382	-0.954	-3.678	-0.948	-2.147	-4.476	-1.496	-3.768	-3.787
	$p<.001$	p	.183	.702	.340	.000	.343	.032	.000	.135	.000	.000

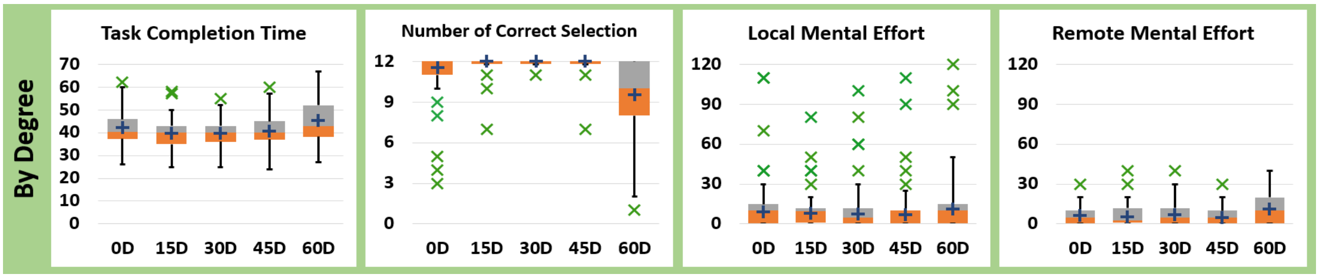


FIGURE 9. With the HoT condition, the average task completion time (seconds), number of correct selections (0~12), and required mental effort by local and remote participants (0: lowest~150 highest; 120 was highest data) (+: mean; x: outlier).

During the user study, we observed that the remote participants positioned their hands at the center of the target number, but local participants sometimes selected the number immediately right to the correct answer. This error mostly happened when using the HoT at a 60 degree angle (Table 4 and 5) where the remote participants' hand gesture looked like it was pointing towards the right in the local participants' perspective (Figure 10a). This influenced the results of the task completion time and the number of correct selections. In Figure 9 and Table 3, using the HoT at a 60 degree angle took more time to complete the task with fewer correct selections compared to using it at the other angles (Table 3; 45 – 60, 30 – 60, 15 – 60).

This error of selecting the target on the right side caused the local participants to call out (or select) the right most numbers more frequently (Table 6 and 7).

In addition, the pointing gesture at the 0 degree angle also influenced the HoT results. The remote participants' hand pose appeared to be pointing upward from time to time (Figure 10b) and this hand pose was more obvious when using HoT at 0 degrees than using it at the other angles (Figure 10c). Consequently, using the HoT at 0 degrees took more time to complete the task with fewer correct selections compared to the HoT at 15, 30, or 45 degrees.

Interestingly, the results of the remote participants' mental effort showed significance when comparing 60 degree angle to any other angles. This may be because looking at the task objects from the side (at a 60 degree angle) was more difficult than looking from in front of it.

D. COMMENTS AND OBSERVATION

Most remote (13 out of 14) and local (12 out of 14) participants commented on the benefit of the HoT ("*I like the hands on the number because I easily knew which number he was pointing at*"; "*My partner knew where I pointed (with the HoT)*"; "*I should look at the numbers and my friend's hand (by switching viewpoint with the HiA); "He was wrong (with the HiA")*).

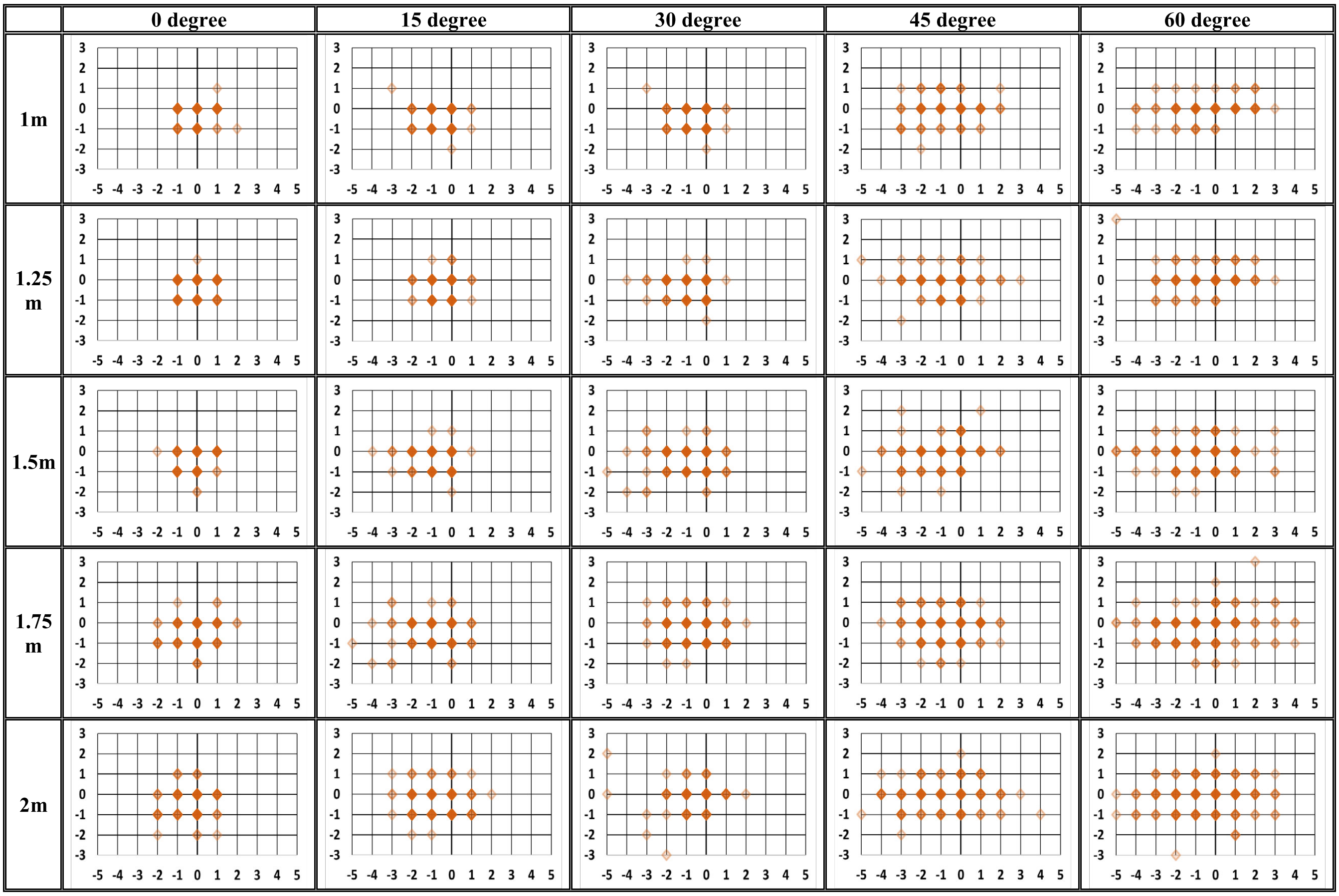
One interesting comment by a remote participant was that "*I wanted to move closer to the numbers, but I could not*". However, we report that providing navigation function according to his/her movement may have risk of bumping into the real-world objects.

Additionally, according to our observation, when the remote participant used the HiA, the local participant selected fewer numbers on the bottom row (Table 6 and 7). This may be because the HiA was displayed a bit higher within the local participants' view as the most local participants were shorter than 175cm which was the remote participants' view height.

V. DISCUSSION

In this paper, we explored the use of hand gesture cues in technology mediated collaborative environments with a large task space. We intensively investigated how the distance to task objects and the view angle difference between the gesture performer (i.e. the remote helper) and viewer (i.e. the local worker) negatively influenced the understanding of the conventional hand gesture cue (HiA). We proposed a solution, the HoT style hand gesture cue and the user study results

TABLE 4. The graph in each cell shows the distribution of incorrect selections at each experimental position with the HiA interface. The number at each axis of the graphs is the difference (in row-column) between the target number and the incorrect selection (left (−5)∼right (+5); below (−3)∼above (+3)). The transparency of the dots is set by the frequency of the incorrect selections (more incorrect selections: more occluded).



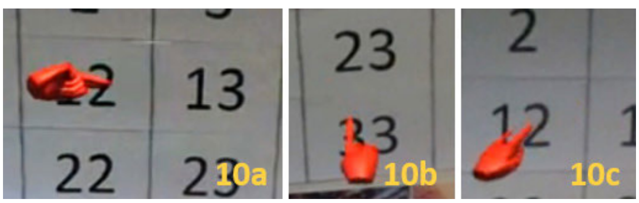


FIGURE 10. Hand pointing gestures with the HoT condition.

showed that the HoT significantly reduced the negative influence of the distance and view angle difference.

A. EXAMINATION OF THE HYPOTHESES

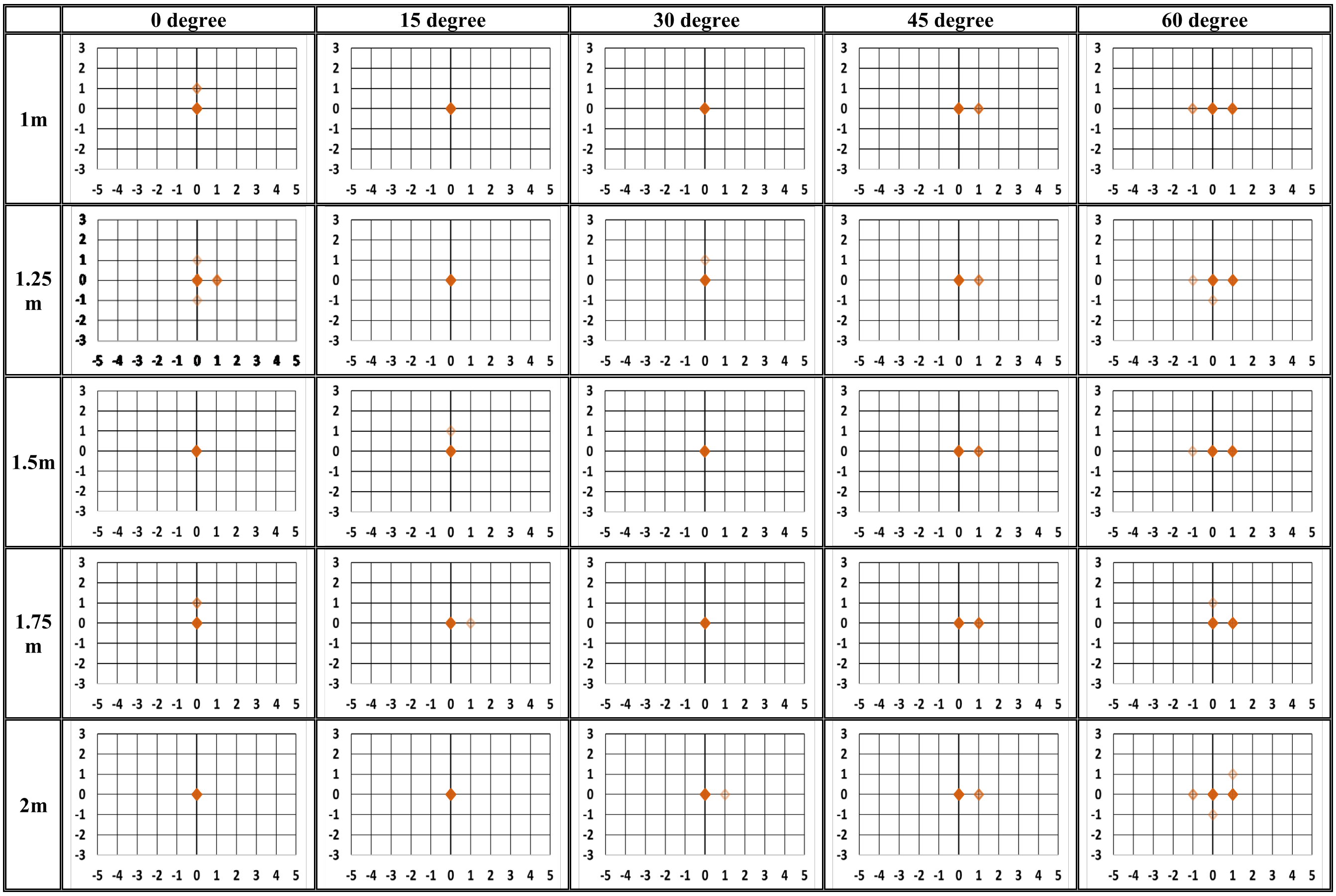
In the user study we found a negative influence of the distance and view angle difference in understanding where the remote participant's HiA interface was referring to. Thus, our hypothesis **H1** is supported. In addition, the influence of the angle difference is more obvious between smaller angles (0, 15, and 30 degrees in our study) than between larger angles (30, 45, and 60 degrees). The influence of the distance was more serious between shorter distance (1m, 1.25m, and 1.5m in our study) than between longer distance (1.5m, 1.75m, and 2m).

The HoT interface produced significantly better results in terms of task completion time, number of correct selections, task load, required mental effort, and participants' feeling of realism while keeping the participants' feeling of co-presence and hand ownership compared to the HiA interface. Therefore, hypothesis **H2** is confirmed.

With the HoT interface there was no negative influence of the distance to the task object on the local participants correctly understanding where the HoT style hand gesture referred, and no significant difference in mental effort or task load among various distances. Since the HoT interface placed the hand gesture on the surface of the task object, the distance to the task object would be meaningless. Therefore, hypothesis **H3** is supported.

The influence of the view angle difference between a gesture performer and viewer was significantly reduced with the HoT interface. Participants completed the task faster, had more correct selections and spent less mental effort when using the HoT compared to when using the HiA. Interestingly, even with this benefit, the HoT interface was not perfect. Since the HoT interface still kept the orientation of the hand

TABLE 5. The graph in each cell shows the distribution of incorrect selections at each experimental position with the HoT interface. The number at each axis of the graphs means the difference (in row-column) between the target and the incorrect selection (left (−5)∼right (+5); below (−3)∼above (+3)). The transparency of the dots is set by the frequency of the incorrect selections (more incorrect selections: more occluded).



gesture which mostly depended on the view angle of the remote helper, the hand gesture with the 60 degree angle may have resulted in the hand gesture appearing to point towards the right in the local participants' perspective. Therefore, **H4** is still disputable.

B. EXTENSIBILITY

Our study results could be applied to any hand gesture communication scenarios such as co-located collaboration and CVEs as far as they involve issues arising from the two factors: distance to an object and view angle difference. Even though we investigated the HiA and HoT within a MR remote collaboration system, the limitation of HiA with such issue could happen in any hand gesture communication. Therefore, the results of HoT interface could also be applied to other collaboration senario where such issue exists.

Additionally, even though we did not test the HoT interface with the physical task including complex object manipulation, the HoT interface can be used with the benefit of it (easy to know where or which object the hand gesture refers to), because the HoT interface still can show diverse types of information including rotating, flipping, and moving objects mostly at any axis.

C. HAND OWNERSHIP & CO-PRESENCE

Interestingly, the remote helper still felt the benefits of the HoT in co-presence and hand ownership, but the local worker did not (Figure 8 and Table 2). Since the remote helper could still control the hand gesture with the HoT interface even though it is away at a distance, he/she might have given positive ratings on the co-presence and hand ownership considering the benefit of the HoT. However, the local worker saw the hand appearing away from the remote helper and might have felt less as if it was controlled by remote helper, so the benefit of the HoT interface might have been less apparent in terms of co-presence and hand ownership, showing non-significant results.

D. ALTERNATIVES

Since the pointer, eye gaze, and sketch cues were alternatives to the HiA interface in previous studies [1], [12], [13], [32], our study may have been more comprehensive by comparing them to our HoT interface. The pointer, eye gaze, and sketch

TABLE 6. The grid in a cell shows how many more or less selections were made for each task objects at each experimental position (more selection (+): red~less selection (−): dark blue) with the HiA interface. Each grid has the same 4 × 6 format with the task objects format on the wall (see Figure 3), so the number in the grid means how many more or less selections at the matched task object. With 14 pairs of participants, each number object should be selected 7 times by each hand gesture condition, so '0' means the number object was selected 7 times.

	0 degree	15 degree	30 degree	45 degree	60 degree
1m	-4 2 3 3 2 -4 -3 3 3 3 5 -4 -3 4 1 6 3 -1 -5 0 -2 -4 -4 -4	0 2 3 6 -1 -4 1 3 0 5 2 -5 5 -1 7 6 0 -4 -7 0 -5 -3 -5 -5	-1 2 -1 0 -2 -5 2 0 3 3 0 -2 3 2 4 4 1 -6 0 -2 -1 -1 0 -7	3 -1 3 2 -6 -5 1 5 4 2 -2 -3 3 -1 6 3 -3 -4 1 2 -2 0 -4 -4	0 2 0 1 -1 -2 1 2 -2 1 1 -3 0 3 5 -1 -1 -2 0 0 0 4 -4 -4
1.25 m	-1 6 4 2 4 -2 -4 8 3 2 4 -4 -5 2 4 5 1 -1 -4 -3 -5 -7 -3 -6	1 3 2 5 -3 0 6 3 8 2 -1 -4 4 3 3 -1 0 -2 -5 -5 -4 -4 -5 -6	3 2 5 -4 -1 -4 5 7 3 3 -4 -3 4 2 4 1 0 -6 -1 1 -3 -3 -6 -5	1 4 0 3 -5 -5 4 2 5 5 -1 -5 6 5 0 0 -4 -3 2 -3 2 -4 -5 -4	2 -2 1 -3 -1 -5 -1 5 3 8 -3 -6 0 -1 3 5 3 -3 1 1 4 0 -5 -6
1.5 m	-1 4 4 4 4 -2 1 3 0 3 3 -1 -5 4 1 4 6 -4 -3 -5 -1 -7 -6 -6	10 4 2 6 -4 -7 10 3 -2 1 -4 -2 8 3 2 -2 1 -6 -3 -5 -4 -2 -2 -7	6 0 0 1 -5 -6 3 8 8 6 -4 -5 5 4 6 2 -5 -7 3 -3 -2 -2 -7 -6	6 0 1 0 -6 -7 9 4 10 -1 -3 -4 6 4 3 5 -4 -6 0 -3 0 -2 -6 -6	4 -4 5 4 -4 -4 3 4 8 2 -2 -6 4 5 9 1 -3 -5 0 -1 -2 -4 -7 -7
1.75 m	1 2 10 4 2 0 0 -1 -1 4 4 -5 1 2 1 1 3 -3 -6 -3 -4 -4 -7 -1	11 1 -1 5 -1 -4 10 4 4 -2 -2 -3 7 0 5 -2 -5 -5 4 -4 -5 -5 -6 -6	0 2 -2 1 -4 -6 16 1 7 1 -2 -5 7 5 7 -1 -6 -6 4 -3 3 -7 -5 -7	1 4 1 4 -3 -4 6 6 7 1 -1 -5 2 4 2 0 -5 -5 0 1 -3 -3 -5 -5	4 4 4 0 -6 -4 5 8 6 -2 -3 -1 6 2 2 -1 -5 -3 -6 0 1 -3 -4 -4
2m	2 2 5 5 5 1 2 3 -2 1 -5 -1 -2 1 1 0 -1 -1 0 -4 -1 -3 -4 -4	7 1 4 2 0 0 3 1 8 -4 -5 -4 6 3 1 4 -3 -6 1 -3 -3 -3 -4 -6	6 1 3 0 2 -4 6 8 2 -2 1 -5 -2 3 -3 -1 -1 -4 4 -4 -1 -2 -5 -3	3 1 5 -4 -5 -4 12 5 1 1 -3 -5 2 5 7 -2 -3 -6 3 1 -2 -2 -4 -6	0 5 0 0 -3 -5 0 6 9 3 -1 -2 0 4 3 4 -4 -7 2 -3 -1 -1 -5 -4

cues are displayed on top of the task objects, so they could also solve the issue of the HiA interface as the HoT interface does. The pointer cue is simple and precise to show point information [31], [37]. Gaze moves faster than mouse pointer and hand gesture [16], [36], [56], so it may show better performance in indicating an object compared to the hand gesture and mouse pointer. The sketch cue could show a sequence of required activities (e.g. remaining sketches and showing selection sequence) as drawn sketches remain until being erased [31], [33].

However, using pointer, gaze, and sketch cues has a limitation in terms of extensibility to support other types of tasks, such as object manipulation [4], while in contrast, the HoT interface is extensible for its use in object manipulation tasks (as described at in section 5.B). The HoT interface can show 3D information while the pointer, gaze, and sketch cues are mostly limited to showing a point or 2D information. Additionally, the hand gesture cue has better naturalness [20], intuitiveness [19], and level of co-presence [22] compared to the pointer, gaze, and/or sketch cue(s) because the hand gesture cue is also used in collaborations in the real world.

Although the gaze pointer or ray is mostly limited to showing a point information, they have additional advantages compared to other pointer type cues (e.g. mouse pointer [34], [35] or hand pointer [13], [14]). What gaze shows is not simply a point information but the collaborators' gaze activity, i.e. where the collaborator is looking at [56], [57]. Higuich *et al.* [16] and Bai *et al.* [15] both investigated interfaces that support both gaze and HiA hand gesture and found that they are well harmonious and using both supports better performance with higher level of co-presence. Therefore, exploring the effect of the gaze with the HoT interface could be an interesting topic for future work.

Displaying the partner's first person view in a small window at a corner may also solve the issue of the HiA interface because the local worker could see and understand the remote helper's hand gesture through the remote helper's first person view in the small window. However, this may clutter the view especially on the local worker's AR view because AR HMDs typically have smaller FOV compared to VR HMD. In our system, Meta2 AR HMD has 90 degrees FOV. Additionally, a local worker and a remote helper needs to perform additional interaction: switching their focus to the small window while also keeping their eyes on their own view. In the future, we would further explore this concept of small corner window interface in comparison to the HoT interface.

E. LIMITATIONS

The user study produced a number of interesting results, but there were some limitations as well. The participants in the

TABLE 7. The grid in a cell shows how many more or less selections were made for each task objects at each experimental position (more selection (+): red~less selection (−): dark blue) with the HoT interface. Each grid has the same 4 × 6 format with the task objects format on the wall (see Figure 3), so the number in the grid means how many more or less selections at the matched task object. With 14 pairs of participants, each number object should be selected 7 times by each hand gesture condition, so '0' means the number object was selected 7 times.

	0 degree	15 degree	30 degree	45 degree	60 degree
1m	0 -1 1 0 0 1	0 0 0 0 0 0	0 0 0 0 0 0	0 -1 1 -2 0 2	-1 1 -1 -1 -2 4
	0 0 0 0 0 -1	0 0 -1 0 -1 1	0 0 0 0 0 0	0 0 -1 0 -2 3	0 -1 1 -4 1 3
	0 0 0 0 0 0	0 1 0 0 0 0	0 0 -1 0 0 0	0 0 0 -2 1 1	-2 2 -3 0 0 2
	1 -1 0 0 -1 1	0 0 0 0 0 0	0 0 1 0 0 0	0 0 -1 1 -3 3	0 -1 1 -4 1 4
1.25 m	0 1 0 0 -1 1	0 0 0 0 -1 1	0 0 0 0 0 0	0 0 -1 1 -4 4	0 -1 1 -2 1 2
	0 -2 1 0 0 0	0 0 0 0 0 0	0 0 0 0 0 0	0 0 0 -2 1 1	-1 0 0 -2 0 2
	-1 0 -1 1 -1 1	0 0 0 0 0 0	0 0 0 0 0 0	-1 1 -1 1 -2 2	0 -1 1 -2 -1 4
	1 0 0 -1 1 0	0 0 0 0 0 0	0 0 0 0 -1 1	0 -1 -1 1 1 0	-1 1 -3 1 -1 2
1.5 m	0 0 0 0 0 0	0 0 0 -1 1 0	0 0 0 0 0 0	0 -1 1 -1 1 0	-1 1 -1 -1 -1 3
	0 0 0 0 1 -1	0 0 0 0 0 0	0 0 0 0 0 0	-1 0 -1 2 -1 1	0 -3 3 -2 -1 3
	-1 0 0 0 0 0	-1 0 0 0 0 0	-2 1 0 -1 0 0	0 -1 0 -1 0 2	-2 0 0 1 -3 5
	1 0 0 0 0 0	1 0 0 0 0 0	1 0 0 1 0 0	-1 0 0 0 -1 2	1 -3 3 -3 -2 3
1.75 m	0 0 0 0 1 -1	0 0 0 0 0 0	0 0 0 -1 1 0	0 0 0 1 -3 2	0 -4 5 -3 2 2
	0 0 0 1 -1 0	0 0 0 0 0 0	-1 1 0 1 -1 0	0 -1 1 -1 0 1	-2 1 -2 2 -4 3
	-2 0 0 0 0 0	0 0 0 0 0 0	0 0 0 -1 2 -1	-1 1 -1 0 -2 4	0 -3 2 -1 0 2
	2 0 0 0 0 0	0 0 0 0 -1 1	0 0 0 0 -1 1	0 -1 1 -1 -1 1	-3 3 -4 3 -4 5
2m	0 0 0 0 0 0	0 0 0 0 0 0	0 0 0 0 0 0	0 -1 1 -1 0 1	-1 1 -2 2 -1 1
	0 0 0 0 0 0	0 0 0 0 0 0	0 0 0 0 0 0	-1 1 -1 1 -2 2	0 -1 0 -2 0 3
	0 0 0 0 0 0	0 0 0 0 0 0	0 0 0 -1 1 0	0 -2 2 0 -2 2	-3 2 -1 1 0 1
	0 0 0 0 1 -1	0 0 0 0 0 0	0 0 0 0 0 0	-1 1 -1 1 0 0	0 -1 1 -4 -1 5

user study were not allowed to correct the misunderstanding of the hand gestures to reduce the effect of verbal correction and to complete the experiment in a timely manner. In real tasks, the local worker and remote helper should be able to use speech to correct misunderstandings when the hand gesture cue is difficult to understand. However, even if verbal correction was allowed, using the HiA would require more verbal correction than when using the HoT as the HiA interface had significantly more incorrect selections than the HoT interface. Thus, the benefit of HoT could be even more prominent. In addition, we note that the verbal correction would be much easier with our HoT interface than the HiA interface because the misunderstanding of where the hand gesture refers to is usually adjacent to the correct target object with the HoT, but further away when using HiA (Figure 7 and Table 4 and 5).

In our study, the experimental task was object selection (to see whether a local worker knows where remote helper's hand gesture refers to) and did not include object manipulation. However, our research was about solving the issue of difficulty in knowing where the hand gesture refers to, which is most relevant to selection tasks. In addition, object manipulation is impossible if the local worker does not know where the remote helper's hand gesture refers to and we reported the extensibility of the HoT for object manipulation task in the section 5.B. In the future, we would further explore the HoT interface with the object manipulation task.

Another concern when using our approach is that the HoT interface is limited to showing depth or volume information because it is always placed on the surface of the target. In the future, we will explore how the HoT interface could be displayed not on the object surface but near the object (within a hand reachable area), so that the virtual hands can move in depth direction, allow more natural hand gestures, and represent more types of information, such as showing information relating to volume.

If a remote collaboration system supports an easy and quick navigation function, issues with the HiA interface may be less prominent because the remote helper can navigate close to the task object before using hand gesture cues so the local worker could understand hand gesture nearby the task objects without difficulty. However, the remote helper needs to be able to navigate back and forth to the task objects which are distributed around the large task space. In addition, the navigation may require more user inputs. We note that the remote participants in our study could see the large task space from a distance and still use hand gesture cues that appears close enough to the task object without additional inputs. Moreover, the remote helper's navigation was limited because of the worry about bumping into the wall or objects when wearing a HMD.

VI. CONCLUSION

In recent collaboration systems, each collaborator mostly has a first-person view and they have different view positions in a large task space. However, collaboration systems still use the conventional Hands-in-Air (HiA) style hand gesture cue that have an issue of misunderstanding where it is referring to from the viewer's perspective. As a solution, in this paper, we proposed a new hand gesture cue: Hands-on-Target (HoT) which positions the remote helper's virtual hands on the surface of the task object.

We conducted a user study, in which we first investigated the influence of the distance to the task object and view angle difference between a gesture performer and a viewer when using the HiA style hand gesture cues, then secondly explored whether our HoT style overcomes the negative influence of the distance and view angle.

In the user study, we found that increasing the distance and view angle difference caused misunderstandings with the HiA hand gesture. The negative influence of the distance was more serious between shorter distance (1m, 1.25m, and 1.5m in our study) than between longer distances (1.5m, 1.75m, and 2m). Similarly, the negative influence of the view angle difference was more serious between smaller angles (0, 15, and 30 degrees in our study) than between larger angles (30, 45, and 60 degrees). The HoT interface reduced the negative impact of the view-angle difference and the distance to the objects, and significantly improved the use of the hand gesture cue.

In the future, we plan to extend our study with an object manipulation task and revise the HoT interface to place the virtual hand not on the surface of a task object but near the objects (within a hand reachable area), hence supporting 3D movement and more natural hand gestures. Moreover, we will explore one-to-many remote collaboration (e.g. one local user and two or more remote users) with different types of devices such as tablets, HMDs, and PCs.

REFERENCES

[1] M. W. Alibali, "Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information," *Spatial Cognition Comput.*, vol. 5, no. 4, pp. 307–331, Dec. 2005.

[2] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst, "Mini-me: An adaptive avatar for mixed reality remote collaboration," in *Proc. CHI Conf. Hum. Factors Comput. Syst. (CHI)*, Montreal, QC, Canada, 2018, pp. 1–13.

[3] S. Benford, C. Greenhalgh, T. Rodden, and J. Pycock, "Collaborative virtual environments," *Commun. ACM*, vol. 44, no. 7, pp. 79–85, 2001.

[4] S. Kim, G. Lee, M. Billinghurst, and W. Huang, "The combination of visual communication cues in mixed reality remote collaboration," *J. Multimodal User Interfaces*, vol. 14, no. 4, pp. 321–335, Dec. 2020, doi: 10.1007/s12193-020-00335-x.

[5] W. Huang, L. Alem, F. Tecchia, and H. B.-L. Duh, "Augmented 3D hands: A gesture-based mixed reality system for distributed collaboration," *J. Multimodal User Interfaces*, vol. 12, no. 2, pp. 77–89, Nov. 2017.

[6] P. Wang, X. Bai, M. Billinghurst, S. Zhang, S. Wei, G. Xu, W. He, X. Zhang, and J. Zhang, "3DGAM: Using 3D gesture and CAD models for training on mixed reality remote collaboration," *Multimedia Tools Appl.*, early access, Sep. 2, 2020, doi: 10.1007/s11042-020-09731-7.

[7] M. Billinghurst, T. Piumsomboon, and H. Bai, "Hands in space: Gesture interaction with augmented-reality interfaces," *IEEE Comput. Graph. Appl.*, vol. 34, no. 1, pp. 77–80, Jan. 2014.

[8] M. Sousa, R. K. D. Anjos, D. Mendes, M. Billinghurst, and J. Jorge, "Warping deixis: Distorting gestures to enhance collaboration," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Glasgow, U.K., May 2019, pp. 1–12.

[9] S. Mayer, J. Reinhardt, R. Schweigert, B. Jelke, V. Schwind, K. Wolf, and N. Henze, "Improving Humans' ability to interpret deictic gestures in virtual reality," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Honolulu, HI, USA, Apr. 2020, pp. 1–14.

[10] N. Wong and C. Gutwin, "Support for deictic pointing in CVEs: Still fragmented after all these years," in *Proc. 17th ACM Conf. Comput. Supported Cooperat. Work Social Comput. (CSCW)*, Baltimore, MD, USA, 2014, pp. 1377–1387.

[11] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa, "The go-go interaction technique: Non-linear mapping for direct manipulation in VR," in *Proc. 9th Annu. ACM Symp. User Interface Softw. Technol. (UIST)*, Seattle, WA, USA, 1996, pp. 79–80.

[12] T. Teo, L. Lawrence, G. A. Lee, M. Billinghurst, and M. Adcock, "Mixed reality remote collaboration combining 360 video and 3D reconstruction," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Glasgow, U.K., May 2019, pp. 1–14.

[13] G. A. Lee, T. Teo, S. Kim, and M. Billinghurst, "A user study on MR remote collaboration using live 360 video," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Munich, Germany, Oct. 2018, pp. 153–164.

[14] S. Kim, G. Lee, W. Huang, H. Kim, W. Woo, and M. Billinghurst, "Evaluating the combination of visual communication cues for HMD-based mixed reality remote collaboration," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Glasgow, U.K., May 2019, pp. 1–13.

[15] H. Bai, P. Sasikumar, J. Yang, and M. Billinghurst, "A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Honolulu, HI, USA, Apr. 2020, pp. 1–13.

[16] K. Higuch, R. Yonetani, and Y. Sato, "Can eye help you?: Effects of visualizing eye fixations on remote collaboration scenarios for physical tasks," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, San Jose, CA, USA, May 2016, pp. 5180–5190.

[17] S. Kim, M. Billinghurst, and K. Kim, "Multimodal interfaces and communication cues for remote collaboration," *J. Multimodal User Interfaces*, vol. 14, pp. 313–319, Oct. 2020, doi: 10.1007/s12193-020-00346-8.

[18] M. Deller, A. Ebert, M. Bender, and H. Hagen, "Flexible gesture recognition for immersive virtual environments," in *Proc. 10th Int. Conf. Inf. Visualisation (IV)*, 2006, pp. 563–568.

[19] S. W. Greenwald, W. Corning, and P. Maes, "Multi-user framework for collaboration and co-creation in virtual reality," in *Proc. 12th Int. Conf. Comput. Supported Collaborative Learn. (CSCL)*, Jun. 2017, pp. 1–2.

[20] R. A. Bolt, "Put-that-there: Voice and gesture at the graphics interface," in *Proc. SIGGRAPH*, 1980, pp. 262–270.

[21] T. Yamamoto, M. Otsuki, H. Kuzuoka, and Y. Suzuki, "Tele-guidance system to support anticipation during Communication," *Multimodal Technol. Interact.*, vol. 2, no. 3, p. 55, Sep. 2018.

[22] P. Wang, S. Zhang, X. Bai, M. Billinghurst, W. He, M. Sun, Y. Chen, and H. Ji, "DHANDS: A gesture-based MR remote collaborative platform," *Int. J. Adv. Manuf. Technol.*, vol. 102, pp. 1339–1353, Jan. 2019.

[23] D. Kirk, A. Crabtree, and T. Rodden, "Ways of the hands," in *Proc. 9th ECSCW*. Paris, France: Springer, 2005, pp. 1–21.

[24] L. Alem, F. Tecchia, and W. Huang, "Remote tele-assistance system for maintenance operators in mines," in *Proc. 11th Underground Coal Operators Conf.*, Wollongong, NSW, Australia, 2011, pp. 171–177.

[25] S. Chen, M. Chen, A. Kunz, A. E. Yantaç, M. Bergmark, A. Sundin, and M. Fjeld, "SEMarbeta: Mobile sketch-gesture-video remote support for car drivers," in *Proc. 4th Augmented Hum. Int. Conf. (AH)*, Stuttgart, Germany, 2013, pp. 69–76.

[26] A. Kunz, T. Nescher, and M. Kuchler, "CollaBoard: A novel interactive electronic whiteboard for remote collaboration with people on content," in *Proc. Int. Conf. Cyberworlds*, Oct. 2010, pp. 430–437.

[27] W. Huang, S. Kim, M. Billinghurst, and L. Alem, "Sharing hand gesture and sketch cues in remote collaboration," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 428–438, Jan. 2019.

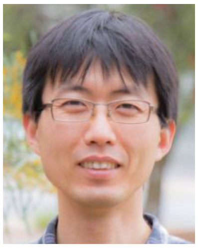
[28] L. Gao, H. Bai, G. Lee, and M. Billinghurst, "An oriented point-cloud view for MR remote collaboration," in *Proc. SIGGRAPH ASIA Mobile Graph. Interact. Appl. (SA)*, Macau, China, 2016, pp. 1–4.

[29] R. S. Sodhi, B. R. Jones, D. Forsyth, B. P. Bailey, and G. Maciocci, ''BeThere: 3D mobile collaboration with spatial input,'' in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Paris, France, Apr. 2013, pp. 179–188.
[30] L. Gao, H. Bai, R. Lindeman, and M. Billinghurst, ''Static local environment capturing and sharing for MR remote collaboration,'' in *Proc. SIGGRAPH Asia Mobile Graph. Interact. Appl. (SA)*, Bangkok, Thailand, 2017, pp. 1–6.
[31] S. Kim, G. Lee, N. Sakata, and M. Billinghurst, ''Improving co-presence with augmented visual communication cues for sharing experience through video conference,'' in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Munich, Germany, Sep. 2014, pp. 83–92.
[32] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst, ''The effects of sharing awareness cues in collaborative mixed reality,'' *Frontiers Robot. AI*, vol. 6, no. 5, p. 5, Feb. 2019.
[33] S. Kim, G. A. Lee, S. Ha, N. Sakata, and M. Billinghurst, ''Automatically freezing live video for annotation during remote collaboration,'' in *Proc. 33rd Annu. ACM Conf. Extended Abstr. Hum. Factors Comput. Syst. (CHI EA)*, Seoul, South Korea, 2015, pp. 1669–1674.
[34] S. Kim, G. A. Lee, and N. Sakata, ''Comparing pointing and drawing for remote collaboration,'' in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Adelaide, SA, Australia, Oct. 2013, pp. 1–6.
[35] T. Y. Wang, Y. Sato, M. Otsuki, H. Kuzuoka, and Y. Suzuki, ''Developing an AR remote collaboration system with semantic virtual labels and a 3D pointer,'' in *Proc. IHCI*, Copenhagen, Denmark, 2020, pp. 407–417.
[36] P. Wang, S. Zhang, X. Bai, M. Billinghurst, W. He, S. Wang, X. Zhang, J. Du, and Y. Chen, ''Head pointer or eye gaze: Which helps more in MR remote collaboration?'' in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Osaka, Japan, Mar. 2019, pp. 1219–1220.
[37] K. Gupta, G. A. Lee, and M. Billinghurst, ''Do you see what i see? The effect of gaze tracking on task space remote collaboration,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 11, pp. 2413–2422, Nov. 2016.
[38] W. Sato, T. Kochiyama, S. Uono, S. Yoshimura, and M. Toichi, ''Neural mechanisms underlying conscious and unconscious gaze-triggered attentional orienting in autism spectrum disorder,'' *Frontiers Hum. Neurosci.*, vol. 11, p. 339, Jun. 2017.
[39] S. Goldin-Meadow, ''The role of gesture in communication and thinking,'' *Trends Cognit. Sci.*, vol. 3, no. 11, pp. 419–429, Nov. 1999.
[40] D. A. Bowman and L. F. Hodges, ''An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments,'' in *Proc. Symp. Interact. 3D Graph. (SI3D)*, RI, USA, Apr. 1997, pp. 35–38.
[41] T. Feuchtner and J. Müller, ''Extending the body for interaction with reality,'' in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Denver, CO, USA, May 2017, pp. 5145–5157.
[42] J. Hindmarsh, M. Fraser, C. Heath, S. Benford, and C. Greenhalgh, ''Object-focused interaction in collaborative virtual environments,'' *ACM Trans. Comput.-Hum. Interact.*, vol. 7, no. 4, pp. 477–509, 2000.
[43] M. Fraser, S. Benford, J. Hindmarsh, and C. Heath, ''Supporting awareness and interaction through collaborative virtual interfaces,'' in *Proc. 12th Annu. ACM Symp. User Interface Softw. Technol. (UIST)*, Asheville, NC, USA, 1999, pp. 27–36.
[44] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsomboon, and M. Billinghurst, ''Revisiting collaboration through mixed reality: The evolution of groupware,'' *Int. J. Hum.-Comput. Stud.*, vol. 131, pp. 81–98, Nov. 2019.
[45] K. Kim, M. Billinghurst, G. Bruder, H. B.-L. Duh, and G. F. Welch, ''Revisiting trends in augmented reality research: A review of the 2nd decade of ISMAR (2008–2017),'' *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 11, pp. 2947–2962, Nov. 2018.
[46] S. Kim, M. Billinghurst, and G. Lee, ''The effect of collaboration styles and view independence on video-mediated remote collaboration,'' *Comput. Supported Cooperat. Work (CSCW)*, vol. 27, nos. 3–6, pp. 569–607, Dec. 2018.
[47] S. Kim, M. Billinghurst, C. Lee, and G. Lee, ''Using freeze frame and visual notifications in an annotation drawing interface for remote collaboration,'' *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 12, pp. 6034–6056, 2018.
[48] H. Durrant-Whyte and T. Bailey, ''Simultaneous localization and mapping: Part I,'' *IEEE Robot. Autom. Mag.*, vol. 13, no. 2, pp. 99–110, Jun. 2006.
[49] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, ''Analysis of the accuracy and robustness of the leap motion controller,'' *Sensors*, vol. 13, no. 5, pp. 6380–6393, May 2013.
[50] S. G. Hart and L. E. Staveland, ''Development of NASA-TLX (task load index): Results of empirical and theoretical research,'' *Adv. Psychol.*, vol. 52, pp. 139–183, Apr. 1988.
[51] C. Harms, and F. Biocca, ''Internal consistency and reliability of the networked minds measure of social presence,'' in *Proc. 7th Annu. Int. Workshop, Presence*, Valencia, Spain, 2004, pp. 1–7.
[52] T. Schubert, F. Friedmann, and H. Regenbrecht, ''The experience of presence: Factor analytic insights,'' *Presence, Teleoperators Virtual Environ.*, vol. 10, no. 3, pp. 266–281, Jun. 2001.
[53] A. Kalckert and H. H. Ehrsson, ''The moving rubber hand illusion revisited: Comparing movements and visuotactile stimulation to induce illusory ownership,'' *Consciousness Cognition*, vol. 26, pp. 117–132, May 2014.
[54] F. R. H. Zijlstra, ''Efficiency in work behaviour: A design approach for modern tools,'' Delft Univ. Press, Delft, The Netherlands, Tech. Rep., 1995.
[55] F. R. H. Zijlstra and L. V. Doorn, ''The construction of a scale to measure subjective effort,'' *Delft, Netherlands*, vol. 43, pp. 124–139, 1985.
[56] G. A. Lee, S. Kim, Y. Lee, A. Dey, T. Piumsomboon, M. Norman, and M. Billinghurst, ''Improving collaboration in augmented video conference using mutually shared gaze,'' in *Proc. ICAT-EGVE*, Adelaide, SA, Australia, 2017, pp. 197–204.
[57] G. Lee, S. Kim, Y. Lee, A. Dey, T. Piumsomboon, M. Norman, and M. Billinghurst, ''[POSTER] mutually shared gaze in augmented video conference,'' in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR-Adjunct)*, Nantes, France, Oct. 2017, pp. 79–80.

SEUNGWON KIM received the bachelor's and master's degrees from the University of Tasmania, in 2008 and 2010, respectively, and the Ph.D. degree from the HIT Lab NZ, New Zealand, in 2016, under the supervision of Prof. M. Billinghurst. During his Ph.D. study, he developed a remote collaboration system supporting stabilized sketch and pointer cue. He is currently a Postdoctoral Fellow with the Empathic Computing Laboratory, University of South Australia. He has more than 20 publications in remote collaboration study at several notorious journals and conferences. His research interests include remote collaboration using augmented virtual communication cues and sharing experience between distance users.

ALLISON JING received the master's degree in human-computer interaction from the University of St Andrews, in 2018. She is currently pursuing the Ph.D. degree with the Empathic Computing Laboratory, University of South Australia. Prior to her Ph.D. studies, she worked as a UX/UI Designer/Business Consultant for international clients, including Microsoft, Xbox, and Amazon Game Studios. She also works under Prof. M. Billinghurst and Dr. G. Lee supervision, focusing on non-verbal gesture sharing XR remote collaboration research.

HANHOON PARK received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Hanyang University, Seoul, South Korea, in 2000, 2002, and 2007, respectively. From 2008 to 2011, he was a Postdoctoral Researcher with the NHK Science and Technology Research Laboratories, Tokyo, Japan. In 2012, he joined the Department of Electronic Engineering, Pukyong National University, Busan, South Korea, where he is currently a Professor. His current research interests include augmented reality, human–computer interaction, and affective computing.



GUN A. LEE (Member, IEEE) received the Ph.D. degree in computer science and engineering from POSTECH, in 2009, researching immersive authoring methods for creating virtual reality (VR) and augmented reality (AR) contents, developing the first system in the world that allowed people to create AR applications within the AR environment. He had been worked as a Senior Researcher with ETRI from 2005 to 2011, where he developed VR and AR systems for industrial applications. He is currently a Senior Research Scientist with the Empathic Computing Laboratory, University of South Australia, researching interaction and visualization methods for mobile and wearable AR and VR systems. He has produced over 60 publications in the areas, such as augmented reality and virtual reality.

WEIDONG HUANG received the Ph.D. degree in computer science from The University of Sydney. He is currently an Associate Professor with the University of Technology Sydney. He also has formal training in experimental psychology and professional experience in psychometrics. He is the author of over 140 publications. His research interests include human-centred computing in general, visual data analytics, human–computer interaction, and visual perception in particular. He is also a Founding Chair of the Technical Committee on Visual Analytics and Communication for IEEE SMC Society and a Guest Editor of a number of SCI indexed journals. He has served as a conference chair, a PC chair, or an organization chair for a number of international conferences and workshops.

MARK BILLINGHURST received the Ph.D. degree in electrical engineering from the University of Washington, in 2002. He is currently a Professor of human-computer interaction with the University of South Australia, researching Empathic Computing. Prior to this, he was the Director of the HIT Lab NZ, University of Canterbury, from 2002 to 2015, developing innovative computer interfaces. He has produced over 400 publications in areas, such as augmented reality, multimodal interaction, and mobile interfaces. In 2013, his work in augmented reality received the IEEE VR Technical Achievement Award. He was elected as a Fellow of the Royal Society of New Zealand.