Journal of Universal Computer Science, vol. 19, no. 8 (2013), 1158-1174
submitted: 30/11/12, accepted: 29/3/13, appeared: 28/4/13

Gesturing in the Air: Supporting Full Mobility in Remote Collaboration on Physical Tasks

Weidong Huang
(CSIRO ICT Centre, Marsfield, Australia
Tony.Huang@csiro.au)

Leila Alem
(CSIRO, ICT Centre, Marsfield, Australia
Leila.Alem@csiro.au)

Abstract: Many collaborative situations require that a remote helper guides a local worker in performing manipulations of physical objects in the real world (physical tasks). Existing systems supporting such collaboration often confine collaborators in fixed desktop settings. Therefore they have limited usefulness in situations in which collaborators are mobile and/or desktop settings are not feasible. In this paper, we present HandsInAir, a wearable system for remote guidance. This system is designed to support mobility of the collaborators and provide easy access to remote expertise. HandsInAir draws on the richness of hand gestures for remote guiding and implements a novel approach that supports unmediated remote gestures and allows the helper to perform natural gestures by hands without the need of a physical support. We review related work, describe technical implementation, and present a usability study demonstrating the usefulness and usability of HandsInAir.

Keywords: Wearable computing, Remote collaboration, Mobile collaboration, Remote gestures
Categories: H.5.2

1 Introduction

Rapid advancements in technology have made it possible for remotely located people to break barriers of distance in space and collaborate with each other in similar ways as they do when they are co-located. Nowadays collaboration between individuals across the globe and organizations has become an essential part of our daily activities. The past decades have seen a fast growing interest among researchers and engineers in developing systems to support remote collaboration [Gauglitz, 12]. However, most of these systems aim to support collaborations in which individuals play similar roles. Relatively less attention has been given to collaborative activities in which collaboration partners have distinct roles, particularly with one partner playing the role of *helper* and the other playing the role of doer/*worker*.

More specifically, as technologies become increasingly complex, our dependence on external expertise to understand and use the technology is growing rapidly. There is a range of real world situations in which assistance from a remote helper is needed for a local novice worker to accomplish collaborative physical tasks [Fussell, 04]. Such tasks require remote collaborators to work together manipulating physical objects in the real world, involving complex coordination between verbal communication and physical actions. For example, an ultrasound examination is one of medical checks that require specific expertise to conduct. However, such expertise is often limited in supply and not always available locally. In some cases, there is a need for a remote radiologist to guide a non-specialist doctor or nurse operating an ultrasound machine to conduct a quality diagnostic ultrasound scan. Other examples include a remote expert providing technical support for an onsite technician to maintain or repair a piece of equipment; a remote instructor helping a disabled student at home to complete art and craft homework.

It has been widely agreed that one of the main issues with remote collaboration is that there is no longer common ground for collaboration partners to communicate the same way as they do when they are co-located [Clark, 91]. A series of studies have been conducted demonstrating that providing remote collaborators with access to a shared visual space helps to achieve common ground and can be beneficial to the completion of collaborative tasks (e.g., [Fussell, 04]). According to Tang et al. [Tang, 04], "a shared visual workspace is one where participants can create, see, share and manipulate artefacts within a bounded space". For remote collaboration, shared visual spaces are often provided in the form of video views of the workspace of the worker [Gergle, 06].

Further, prior research has indicated that the reason why face-to-face communication is more efficient than video-mediated communication is mainly because in the face-to-face condition, collaboration participants are able to perform gestures over the task objects and these gestures are visually available to all participants [Kirk, 05b]. This suggests that it is important to support gesturing in the shared visual space for effective remote collaboration.

A number of systems have been developed to support remote guidance by providing a shared visual space and using the space for gesturing (e.g., [Alem, 11a]). However, existing systems often confine collaborators to fixed desktop settings. The value of remote guiding technology in supporting mobility of the collaboration and providing easy access to remote expertise has not been fully explored:

Mobility: During the guiding process, workers may be required to walk around to fetch tools and inspect machines, while helpers may need to go to different locations to look for materials or information. For example, in a call centre, a service provider often deals with customers who may use a range of devices or machines of different models. He/she often needs to walk around to look for the specific manual for that device, or have a closer look at the sample machines for various purposes such as identifying the right model.

Accessibility: When remote expertise is required it is often urgent and helpers may be out of their office and on the move. For example, in a manufacturing factory, when a sophisticated machine suddenly breaks down, onsite maintenance technicians require urgent input from a remote expert, as the time lost in the machine not being running translates into a loss in productivity.

In an attempt to address the mobility and accessibility issues, we developed HandsInAir, a wearable system for mobile remote guidance, which was briefly reported in a poster [Huang, 13]. In the remainder of this paper, we first present a theoretic background for our research, with a focus on shared visual spaces and the role of remote gestures for collaborative physical tasks. Then we briefly review approaches of supporting remote gestures that have been used in previous research, followed by the presentation of the "hands-in-the-air" approach. Next we introduce our HandsInAir system with detailed user interfaces and system specifications, followed by a usability study. Finally we conclude the paper with a brief discussion and a short summary.

2 Background

2.1 Shared Visual Space

In performing collaborative physical tasks, people interact with each other via various communication channels. The interpersonal communication can be more effective when collaborators share a greater amount of common ground, which includes mutual knowledge, beliefs, attitudes, expectations. Previous research has demonstrated the value of shared visual space in achieving common ground (e.g., [Fussell, 00]). In particular, according to Kraut et al. [Kraut, 03], shared views of a workspace play at least three interrelated roles:

- Maintain situational awareness
- Aid conversational grounding
- Promote sense of co-presence

First, to have a successful collaboration, collaborators need to have on-going awareness of the task and their partner. This awareness can be used to plan what to say and what to do next, serving as a mechanism to coordinate between their verbal utterances and physical actions. Such awareness can be obtained through shared visual views of workspace because collaborators can see what is happening. Second, effective communication largely depends on how much mutual knowledge they have about the task and their partner. More specifically, as stated by Gergle et al. [Gergle, 06], "speakers form utterances based on an expectation of what a listener is likely to know and then monitor whether the utterance was understood. In return, listeners have a responsibility to demonstrate their level of understanding". Information needed for building such mutual knowledge can be obtained from shared views of workspace. Third, when collaboration takes place among individuals who are physically distributed, it is important to help collaborators to feel connected. Enhancing sense of co-presence has proved to be beneficial to the success of collaboration, and shared views of workspace help to promote such sense of "being together" [Alem, 11b; Kraut, 02].

2.2 Remote Gestures

Although a shared visual space is helpful for grounding, or establishing common ground between collaboration partners, it is not feasible, if not impossible, to provide all visual information that is available to co-located collaborators to remote collaborators, due to bandwidth limitations and limited cognitive capacity of humans [Fussell, 04; Kraut, 03]. Therefore in developing systems for remote collaboration, it is important to determine what visual information is the most important and make sure this information is provided in an appropriate way.

Observational studies of remote collaboration on physical tasks have revealed that collaborators speak and act in relation to the position and status of objects in the workspace and on-going activities of each other in the environment. Their speeches and actions are intricately dependant on each other; while speaking, they constantly use hand gestures to clarify and enhance their messages. Fussell and her colleagues conducted a series of studies on collaborative physical studies and found that not only speech, but also gestures and actions were used for grounding and that the use of gestures improved task performance (e.g., [Fussell, 03; Fussell, 00]). With access to the shared visual space, helpers allocate most of their attentions on workers' hands and task objects [Gergle, 06; Kirk, 05a]. All these findings indicate that it is important to support remote gestures for developing tools for remote collaboration on physical tasks.

Further, Fussell et al. [Fussell, 03] classified hand gestures into two groups: pointing gestures and representational gestures. The former is used to indicate the direction of movement or the locations of task objects, while the latter is to represent the form and nature of task objects or actions to be taken with the objects. The authors conducted two studies to investigate the role of these two types of gestures and how the gestures could be effectively conveyed to the remote site. The first study used a system that was mouse-based and supported remote pointing only, while the second study used a system that used pen-based drawings to represent hand gestures. The results indicated that only a simple cursor pointing was not enough for effective collaboration, while pen-based drawings of remote gestures resulted in communication and performance being as good as that in co-located collaboration.

More comprehensively, Kirk et al. [Kirk, 05b; Kirk, 06] conducted a series of studies that compared different ways of conveying remote gestures, including projected hands, video presented hands and sketches. These studies investigated the effects of gesture formats on both immediate task performance and longer-term knowledge development (learning). They found that gesturing with an unmediated representation of hands led to significantly better performance of collaborative physical tasks.

2.3 Supporting Remote Gestures

A number of systems have been proposed or developed in the literature, supporting remote gestures using various technologies for remote guidance. In this subsection, we briefly review prior approaches with a focus on how remote gestures are performed by the helper.

2.3.1 Agent based remote gestures

In this approach, helper gestures are delivered by an agent located at the workspace of the worker such as a laser pointer or a stick. For example, in the WACL system of Sakata et al. [Sakata, 03], the worker wears a steerable camera attached with a laser head. The helper can independently control the camera to see the workspace and point to the real object via the laser pointer. In this setting, the helper is sitting at a desk operating the laser pointer. The GestureMan systems of Kuzuoka et al. [Kuzuoka, 16] also employed the agent approach. In these systems, the helper uses joystick to control a mobile robot that is located on the worker site. The helper points to the object on the screen and this gesture is conveyed by the mobile robot through the use of a pointing stick and a laser pointer.

2.3.2 Digital annotations/gestures

In this approach, digital sketching is used to represent gestures. For example, Ou et al. [Ou, 03] developed a DOVE (Drawing over Video Environment) system that integrates gestures of helper into the live video of the worker's workspace. In this system, the helper uses a digital pen and performs gestures by drawing on the video streams of the work environment while providing verbal instructions. Palmer et al. [Palmer, 07] and Gauglitz et al. [Gauglitz, 12] also used this approach in developing their remote guidance systems.

2.3.3 Projected hands

In this approach, the helper hands are directly projected into the worker's workspace and aligned with the associated objects. For example, Kirk and Fraser [Kirk, 05b; Kirk, 06] presented a mixed ecology system. This system requires the helper to gesture at the desk while looking at the monitor in the front. His hands are captured by a video camera, and the captured hands are directly projected onto the desk of the worker. Yamashita et al. [Yamashita, 11] developed an immersive system called T-room. This system also uses projected hands to support remote gestures, but with additional images of helpers shown on the vertical walls of T-room.

2.3.4 Hands over workspace videos

In this approach, the helper performs gestures over the workspace videos showed on a computer display. The hand movements of the helper are captured by a camera and displayed to the worker. For example, in the SharedView system of Kuzuoka [Kuzuoka, 92], the helper is required to stand at the side of a display that shows the video of the worker's workspace. He uses his hands to gesture on the objects over the display. The combined video of the helper hands and the worker workspace is captured by a camera and then sent to the worker side and displayed onto the head-mounted display worn by the worker. This approach was also used in the HandsOnVideo system of Alem et al. [Alem, 11a].

3 System Overview

Our HandsInAir system includes two parts: a helper station and a worker station. The two stations are connected through a wireless network. In this section, we first explain how remote gestures are supported in HandsInAir. Then the system's hardware and software implementtaion is presented, which is followed by the information of how the system works.

3.1 Hands in the Air

In order to support mobility of collaborators, we used wearable computers to run the system software and near-eye devices to display the visual information. In previous approaches of supporting remote gestures, helpers are often required to work within a fixed desktop setting and need to use, touch or control physical objects to perform gestures. This seeting is no longer feasible to set up when they are mobile. As a result, a major challenge we faced, among others (e.g., [Herskovic, 11, Ghiani, 10]), was how to support the richness of remote gestures when support of a physical display/screen or other operational objects for the helper was no longer available.

In developing HandsInAir, we implemented an approach that was to meet this challenge [Alem, 11c]. As shown in Figure 1, the helper wears a helmet mounted with a camera and a near-eye display and he is able to perform gestures in the air for guiding purposes. The near-eye display that we used in HandsInAir is a device with two small screens. The user can look into the screens and see a virtual display that is equivalent to a 67-inch screen as viewed from ten feet away. More details are introduced in the next subsection.

Figure 1: Helper performs gestures in the air (left) and the near-eye display (right)

3.2 Hardware and Software Implementation

Both the helper and worker stations have the same hardware configuration. As shown in Figure 2, the hardware used to implement each wearable unit consists of a helmet mounted with a Microsoft Lifecam Webcam on top of, and a Vuzix 920 wrap near-eye display beneath the brim. Both the camera and the near-eye display are connected to a wearable PC worn by the user. The camera of the worker station is used to capture the workspace of the worker, while the camera of the helper station is to capture the hands of the helper. The near-eye display is used to display the combined video of the worker's workspace and the helper's hands. There is also an audio connection between the two sides to support verbal communications.

Figure 2: User interface

A software application is running on each wearable PC to provide necessary functions of video processing/transmission and network communication between the two stations. The software is developed in C++ on Windows XP machines utilizing a number of open source libraries. Both the worker and helper stations simultaneously act as a video server and a video client. The worker station acts as a server sending local camera feeds of the workspace and as a client to the helper station receiving video feeds of the helper's hands. Likewise the helper station acts as a video client receiving workspace feeds and sending video feeds containing the helper's hands. The Intel OpenCV open source computer vision library is used to implement an Adaptive Skin Detector, which extracts the helper's hands from video feeds of the helper camera and combines them with corresponding video feeds of the workspace (see Figure 3). This detector is also used to display the combined videos on the near-eye displays of the helper and the worker.

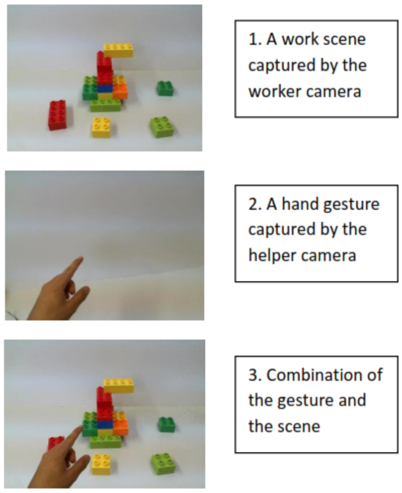

Figure 3: Illustration of combining a hand gesture and a workspace scene

Network connections are realized at the low level by opening up streaming connections as the wearable PCs on both sides simultaneously send and receive a sequence of images. The images are compressed with JPEG compression prior to sending, and decompressed upon receipt using the open source IJG (Independent JPEG Group) LibJPEG library to avoid sending costly raw image data and to maintain a real time frame rate on both sides.

3.3 How the System Works

How the system works is illustrated in Figure 4. Once a connection is established the system initializes two video streams between the stations. First the scene video from the worker camera is fed to the helper station and displayed on the near-eye display. The helper examines the video, talks to the worker and performs gestures which are captured by the helper camera. The hands are extracted without the background and combined with the scene video. What is shown on the helper's near-eye display is continuously updated with the combination. In other words, the helper is able to see his hands performing gestures at the task artefacts on the display. The extracted hand images are also sent to the worker side, combined with the scene video and displayed on the near-eye display. This allows the worker to see the unmediated hand gestures. The worker hears the audio instructions, sees the visual aids by looking up in the near-eye display, when necessary, and performs operations as instructed by the helper.

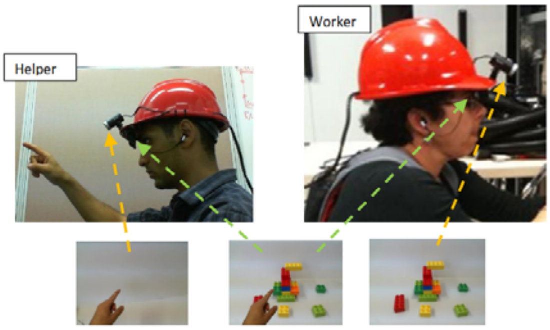

Figure 4: Illustration of Camera captures and the content of near-eye displays [Alem, 11c]

4 Usability Study

As described in the last section, HandsInAir not only allows collaborators to be mobile, but also enables helpers to perform pointing and more complex hand gestures over the remote objects shown on the virtual display. We followed a participatory design approach during the design and development of this system. In this section, we present a user study that we conducted to validate its usability.

4.1 Design

There were two separate rooms with one hosting the helper station, and the other hosting the worker station. Users were recruited to complete representative physical tasks collaboratively. The participants were randomly grouped in pairs with one playing the role of helper and the other playing worker. The participants of each pair were each located in one of the two rooms according to the role being assigned. The helper and worker could talk to each other through speaker/microphone headsets. The whole task process was video recorded on both helper and worker sides for further analysis.

There was also a questionnaire session at the end of the task: one for the helper and one for the worker. Both included questions asking participants to rate the system usability. In this particular study, the usability was evaluated from the following perspectives:

- Ease of learning
- Ease of use
- Usefulness
- Task satisfaction
- Mobility
- Perception of interaction
- Environment awareness

- Co-presence
- Perception of hand gestures

Questionnaires also included open questions. The open questions asked participants' experiences of using the system, and their opinions for possible further improvements.

4.2 Participants

Twenty people volunteered to participate in the study. Fifteen of them were male and the rest were female. They were aged between 20 and 40. By the time they participated in the study, none of them had experience of using HandsInAir, or any other systems of the same kind.

4.3 Task

In this study, the participants were asked to assemble a set of loose Lego toy pieces. Previous user studies in the literature have also used the assembly of toy blocks for similar purposes (e.g., [Huang, 11]). Since toy assembly has components that can be found in a range of real world physical tasks such as assemble, disassemble, select and rotate, this task is considered representative for remote guidance on physical objects [Fussell, 04]. During the task, the worker was asked to assemble the toy pieces into a pre-specified complex model under the instruction of the helper.

There was an instruction manual for the helper. The guiding manual was divided into three parts and the parts were placed separately in different locations of the room. The helper needed to go to the first place, pick up the manual and do the guiding, and then go to the next until the end of the task. The helper was instructed that he could provide verbal and gestural instructions to the worker at any time, but not allowed to show any part of the manual to the worker. The worker, on the other hand, had no idea about what steps were needed to complete the task.

To mimic the general workplace settings of workers, we used a workshop room for the location of the worker (see Figure 5). The workshop was full of equipments, tools and was composed of a number of work areas. The toy pieces were placed in different locations. The worker had to move around the workspace to collect them and get the task done. To test whether the worker was aware of the environment while he walked with a near-eye display, small obstacles were deliberately placed in the trajectory of the worker. The worker had to avoid them while moving around. To prevent workers from tripping over, only light empty boxes were used as obstacles. The helper room had tables and chairs and it was about twenty meters away from the worker room.

Figure 5: The workshop room setup

4.4 Procedure

Before the experiment, an introduction session was given to the two participants of each pair. They were gathered in a meeting room. First they were given a short tutorial and a brief demonstration on how the system worked. The helper interface and the worker interface were introduced. Then the task and the procedure of the study were explained. The two participants were also given a chance to get familiar with the system and try out the equipment. During this session, the participants could ask questions and answers were provided by two experimenters.

When they were ready, the two participants were randomly assigned roles with one as helper and the other as worker. Then, each of them was led to the corresponding room where the helper or worker station was located. On each site, there was also an experimenter providing further assistance to the participant, putting the wearable backpack on, recording videos, observing and taking notes of the collaboration behavior.

Once the connection was established on both sides, the participants started performing the guiding task on the Lego toys provided. After the task was completed, each participant was asked to fill the helper or worker questionnaire depending on his role. After finishing their questionnaires, the participants went back to the meeting room for a semi-structured interview. They were encouraged to ask questions, propose ideas and further improvements, debate on the issues and comment on the system. The whole experiment for each pair took about 40 minutes.

4.5 Results

4.5.1 Observations

All participants were able to perform and complete the assembly task without obvious delays. It seemed that the participants were comfortable with system. Helpers were able to gesture in the air while looking at the video in the near-eye display and giving verbal instructions to workers. Workers were able to assemble the toy pieces with their hands while receiving verbal instructions from helpers and looking at the visual aids shown in the near-eye display. It appeared that the participants were able to communicate with each other smoothly and effectively via both the visual and verbal channels provided by HandsInAir.

There were no apparent difficulties observed for the helpers to collect instruction materials and gesture in the air. While walking, they tended to slow down or stop to perform gestures. It seemed natural for them to perform pointing gestures using one hand, or perform representational gestures using two hands.

For workers, it was observed that they were able to avoid obstacles on the way. They generally became more careful when they were close to an obstacle. The workers could also easily locate and fetch the toy pieces required following the instructions given by the helper.

4.5.2 Usability ratings

The usability of HandsInAir was tested based on a mix of positive and negative statements. Each statement was to evaluate one specific aspect of the usability. These statements are listed in Table 1.

Usability	**Statement**
Ease of learning	I found that the system was easy to learn
Ease of use	I found that the system was difficult to use
Usefulness	I found that the system was useful for remote guiding tasks
Task satisfaction	I was disappointed with my task performance
Mobility	I felt that I was free to move around
Environment awareness	I felt that I was unaware of my physical surroundings
Co-presence	I felt that my partner and I were at the same location
Perception of interaction	I felt that I was engaged with my partner during the task
Perception of gestures (helper)	I found it difficult to point to objects
Perception of gestures (helper)	I found it easy to demonstrate assembly of objects
Perception of gestures (worker)	I found it difficult to understand which objects my partner was pointing to
Perception of gestures (worker)	I found it easy to follow my partner’s hand gestures to assemble objects

Table 1: Statements for usability ratings

The participants rated the extent to which they agreed with the statements, based on a scale of 1 to 7, with 1 being “strongly disagree”, 7 being “strongly agree” and 4 being “neutral”. For the purpose of analysis, user ratings were first transferred so that higher ratings meant better usability, and then averaged across the participants. The average of the obtained ratings was computed for the overall usability. The results are shown in Table 2.

	Average	**StDev**
Ease of learning	6.05	1.10
Ease of use	5.50	0.89
Usefulness	6.15	0.67
Task satisfaction	5.30	0.86
Mobility	5.70	0.92
Environment awareness	5.55	0.94
Co-presence	4.35	0.88
Perception of interaction	4.90	0.79
Perception of pointing gestures	5.75	0.97
Perception of representational gestures	5.25	0.85
Overall usability	**5.45**	**0.89**

Table 2: Average values and standard deviations of user ratings for individual and overall usability measures

As can be seen from Table 2, overall, the participants were positive about the usability of HandsInAir with a rating of 5.45 on the scale of 1 to 7. More specifically, the participants rated the usefulness of the system at the highest value of 6.15, indicating that HandsInAir was considered useful. The participants were also positive about ease of learning, ease of use, task satisfaction, mobility, environment awareness, perception of interaction and hand gestures, while co-presence was rated just above being neutral (4.35).

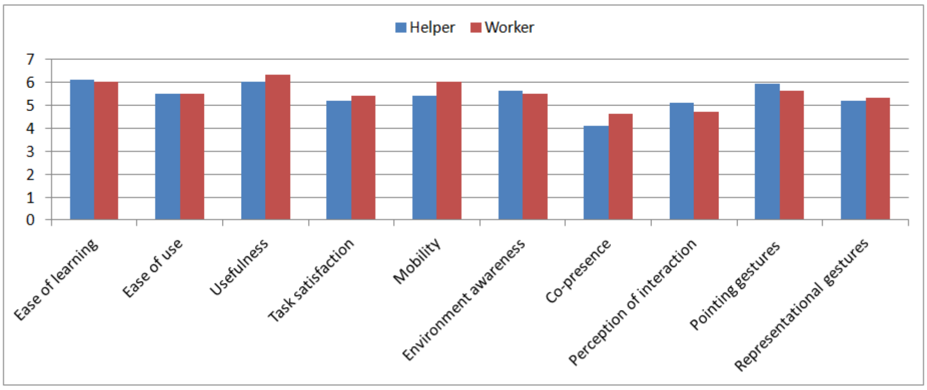

Figure 6: Average usability ratings from helpers and workers

To explore possible differences between helpers and workers, we looked at their user ratings separately and the results are shown in Figure 6. It can be seen that all usability measures were rated positive (above 4). Ease of use was rated the same. Helpers gave higher usability ratings for easy of learning, environment awareness, perception of interaction and pointing gestures, while workers rated higher for usefulness, task satisfaction, mobility, co-presence and representational gestures. However, statistical tests revealed that these differences in the ratings between helpers and workers were not statistically significant at the significance level of 0.05.

4.5.3 User comments

In regard to user responses to the open questions, the participants were generally positive about the system. User comments included "(the system is) very useful for remote guiding." "It is perfect when I do not know how to do and want to be guided." Helpers appreciated being able to perform hand gestures without any physical constraints and commented that gesturing with hands in air is a "cool factor". For example: "I can explain my intention with the help of my hands easily." "It is very helpful, especially for the complex tasks which is hard to instruct by speaking only." "Using my hands for this task is the most appropriate and the best way of guiding the remote worker." "I have a much better idea of what workers can see, much better remote awareness of working environment than other methods. Feel like you can nearly directly interact with the remote environment." "This is very useful and an intuitive way of guiding." "I can do anything I like with my hands (showing shapes and orientations)". However, it was also mentioned that "pointing in the air can be tiring after a while".

Workers also found helpful being able to see the helper's hands via the near-eye display. For example: "It is helpful to be able to see gestures." "It is good that both parties can see and hear the same 'scene' during the task." "It is easy to use." "Seeing the helper's hands is very helpful and instructive." "Hand gestures are useful for knowing which object they want you to interact with. Otherwise they would have to spend much time explaining what object they want you to use."

Further, the qualitative user feedback indicated that the system provided both an easy access to the helper and a good user experience with mobility. For example, "I enjoyed being able to move around in between giving instructions, this gives me the freedom to attend to other tasks if need be. I would not have this option if I was using a desktop computer." "Sure! I can have the helper pack in the boot of my car. Then I can use it at any time/anywhere." "Being able to move around is great." "I can sit or stand so that I can be in a comfortable position." "I don't feel restricted in anyway. I can be standing or sitting anywhere I want." "It was a great experience. I felt relaxed." "To me the value of this system is that you can access the expert/helper wherever they are as long as they are wearing the gear with them; you do not need for your expert to be in a specific room. Experts are highly mobile workers. Accessing them should be as easy as calling them on their mobile phone."

Two participants mentioned that they tried to perform hand gestures while walking and pointed out that this might not be a feasible thing to do. One commented that "I wanted to walk and guide at the same time, but I find that I can't do it. As soon as I try to gesture, I find it hard to keep walking because my attention is on guiding and subconsciously feel dangerous to walk at the same time, without looking at the floor/surrounding. As soon as I stopped guiding, I can walk again." The other participant commented that "I can monitor what my partner is doing while I am on the move. I can provide audio instructions while on the move, but the moment I need to give instructions requiring pointing and showing with my hand how to perform a task, I need to be static."

On the other hand, user comments also suggested some areas for further improvements. For example, some participants mentioned that during the task, they sometimes had to keep switching between the near-eye display and the workspace. This could make their eyes tired if they use the system for a long time. We are planning to use a see-through device to replace the near-eye display to avoid such frequent switches between the video and the real world. Other suggestions made by participants were related to the limitations of network bandwidth, system process capacity and the hardware that we currently could provide. We believe that these limitations could be removed when more powerful technologies and devices become available to us. For example, the quality of videos and images of hands could be further improved by using higher resolution displays and cameras; 3D cameras could be used so that users could have better understanding of the spatial relationships between objects (e.g., [Tecchia, 12]).

5 Discussion

5.1 Usability

The study results confirmed the usability and usefulness of HandsInAir in supporting a remote mobile helper who guides a mobile worker in performing physical tasks. Our observations revealed that participants were able to complete the tasks comfortably without apparent difficulties. The user feedback and usability ratings also indicated that HandsInAir can be useful and usable for real world use.

In particular, the study participants were positive about the mobility support provided by the system to the collaborators. According to their feedback, the mobility support allows a worker to access a remote helper more easily. Also helpers are enabled to continuously engage with the system and their partner when they move around during the guiding process. Participants who played the role of helper also considered gesturing in the air as being intuitive and effective.

5.2 Gesturing in the Air

To meet our goal of freeing helpers from a fixed position, we implemented an approach that enables helpers to perform hand gestures in the air. This is achieved by combining the videos of helper's hands and worker's workspace. The combined video is displayed on the near-eye display. Therefore, all the helper needs to do is to look at the objects shown in the video and gesture in the air, thus removing the reliance on a desktop display/screen.

The current approach also allows the system to convey unmediated hand gestures to the worker. It has been demonstrated that this type of gesture is associated with better task performance and co-presence (e.g., [Kirk, 06; Kraut 02]). Further, the use of a wearable computer and a near-eye display effectively frees helpers from traditional fixed desktop settings. However, it is important to note that the mobility support in this system is to ensure that the collaborators can continuously engage with each other and the system while moving around. It is not our purpose and expectation that the helper is required to *gesture* while walking and that the worker is required to *manipulate objects* while walking.

5.3 Limitation of the Studies

Due to our limitations in accessing real world resources, we used assembly of Lego toy pieces as the experimental task, recruited volunteers as the targeted users and conducted the studies in simulated workshop settings. However, these were done at the cost of realism and generalizability. As demonstrated in prior research (e.g., [Kuzuoka, 92]), testing the system with the real users in real world workplaces for an extended period of time would allow us to systematically examine usability issues and provide us with unbiased insights into the usability of the system, thus being more desirable.

6 Concluding Remarks and Future Work

We have argued that confining collaborators to a fixed position is a limiting factor. This is because in many real world situations, collaborators need to be away from a fixed position for various purposes. In this paper, we have presented HandsInAir, a wearable system for mobile remote collaboration to be used in these situations. The system has been tested for its overall usability and its ability to support mobility of collaborators and the test results are positive.

We are currently planning two field trials with one in an aircraft manufacturing factory and the other in a mining site. Our field partners are currently recording performance data for the current practices of remote guidance. They will also record data for the time period for which HandsInAir is used. This will allow us to conduct comparative studies to understand the benefits of the system. We will also conduct onsite observations to understand user behavior changes before and after the use of the system and to investigate research questions such as how users interact with each other and with the system and how their hands, visual focus, body and verbal communication coordinate together when mobility is an essential part of their collaboration.

References

[Alem, 11a] Alem, L., Tecchia, F. and Huang, W. (2011) HandsOnVideo: Towards a gesture based mobile AR system for remote collaboration. In Recent Trends of Mobile Collaborative Augmented Reality Systems. 127-138. Springer, NY, USA.

[Alem, 11b] Alem, L. and Li, J. (2011) A Study of Gestures in a Video-Mediated Collaborative Assembly Task. Advances in Human-Computer Interaction, vol. 2011, Article ID 987830, 7 pages.

[Alem, 11c] Alem, L. and Huang, W. (2011) Developing Mobile Remote Collaboration Systems for Industrial Use: Some Design Challenges. In Proceedings of The 13th IFIP TC13 Conference on Human-Computer Interaction (INTERACT'11), 442-445.

[Clark, 91] Clark, H. H. and Brennan, S. E. (1991) Grounding in communication. In Perspectives on socially shared cognition. Washington, DC: American Psychological Association.

[Fussell, 04] Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E. and Kramer, A. D. I. (2004) Gestures over video streams to support remote collaboration on physical tasks. Human-Computer Interaction, 19:273-309.

[Fussell, 03] Fussell, S. R., Setlock, L. D. and Parker, E. M. (2003) Where do helpers look? gaze targets during collaborative physical tasks. In Proceedings of extended abstracts on Human factors in computing systems, 768-769.

[Fussell, 00] Fussell, S. R., Kraut, R. E., and Siegel, J. (2000) Coordination of communication: effects of shared visual context on collaborative work. ACM Conference on Computer Supported Cooperative Work, 21-30.

[Gauglitz, 12] Gauglitz, S., Lee, C., Turk, M. and Höllerer, T. Integrating the physical environment into mobile remote collaboration. In Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services (MobileHCI'12), 241-250.

[Gergle, 06] Gergle, D., Kraut, R. E. and Fussell, S. R. (2006) The impact of delayed visual feedback on collaborative performance. ACM Human Factors in Computing Systems, 1303-1312.

[Ghiani, 10] Ghiani, G. and Paternò, F. (2010) Supporting Mobile Users in Selecting Target Devices. Journal of Universal Computer Science, vol. 16, no. 15, 2019-2037.

[Herskovic, 11] Herskovic, V., Ochoa, S.F., Pino, J.A., and Neyem, A. (2011) The Iceberg Effect: Behind the User Interface of Mobile Collaborative Systems. Journal of Universal Computer Science, vol. 17, no. 2, 183-202.

[Huang, 11] Huang, W. and Alem, L. (2011) Supporting Hand Gestures in Mobile Remote Collaboration: A Usability Evaluation. In Proceedings of The 25th BCS Conference on Human Computer Interaction.

[Huang, 13] Huang, W. and Alem, L. (2013) HandsinAir: a wearable system for remote collaboration on physical tasks. In Proceedings of the 2013 conference on Computer supported cooperative work companion (CSCW '13). 153-156.

[Kirk, 05a] Kirk, D. S., Crabtree, A., and Rodden, T. (2005) Ways of the Hand. Proceedings of the European Conference on Computer Supported Cooperative Work, 1-21.

[Kirk, 07] Kirk, D. S., Rodden, T. and and Fraser, D. S. (2007) Turn it this way: grounding collaborative action with remote gestures. ACM Human Factors in Computing Systems, 1039-1048.

[Kirk, 05b] Kirk, D. S., and Stanton Fraser, D. (2005) The Impact of Remote Gesturing on Distance Instruction. In Proceedings of the Conference on Computer Supported Collaborative Learning, 301-310.

[Kirk, 06] Kirk, D. S., and Stanton Fraser, D. (2006) Comparing Remote Gesture Technologies for Supporting Collaborative Physical Tasks. ACM Human Factors in Computing Systems, 1191-1200.

[Kraut, 03] Kraut, R. E., Fussell, S. R. and Siegel, J. (2003) Visual information as a conversational resource in collaborative physical tasks. Human-Computer Interaction, 18, 13-49.

[Kraut, 02] Kraut, R. E., Gergle, D. and Fussell, S.R. (2002) The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-Presence. ACM Conference on Computer Supported Cooperative Work, 31-40.

[Kuzuoka, 04] Kuzuoka, H.; Kosaka, J.; Yamazaki, K.; Suga, Y.; Yamazaki, A.; Luff, P. and Heath, C. (2004) Mediating dual ecologies. ACM Conference on Computer Supported Cooperative Work, 477-486.

[Kuzuoka, 92] Kuzuoka, H. (1992) Spatial workspace collaboration: a SharedView video support system for remote collaboration capability. ACM Human Factors in Computing Systems, 533-540.

[Ou, 02] Ou, J., Fussell, S. R., Chen, X., Setlock, L. D. and Yang, J. (2003) Gestural communication over video stream: supporting multimodal interaction for remote collaborative physical tasks. Proceedings of the 5th conference on Multimodal interfaces, 242-249.

[Palmer, 07] Palmer, D., Adcock, M., Smith, J., Hutchins, M., Gunn, C., Stevenson, D. and Taylor, K. (2007) Annotating with light for remote guidance. In Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI'07), 103-110.

[Sakata, 03] Sakata, N., Kurata, T., Kato, T., Kourogi, M. and Kuzuoka, H. (2003) WACL: supporting telecommunications using - wearable active camera with laser pointer. Proceedings of the seventh International Symposium on Wearable Computers, 53-56.

[Tang, 04] Tang, A., Boyle, M. and Greenberg, S. (2004) Display and presence disparity in Mixed Presence Groupware. Proceedings of the fifth conference on Australian user interface, 73-82.

[Tecchia, 12] Tecchia, F., Alem, L. and Huang, W. (2012) 3D helping hands: a gesture based MR system for remote collaboration. In Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '12). 323-328.

[Yamashita, 11] Yamashita, N., Kaji, K., Kuzuoka, H. and Hirata, K. (2011) Improving Visibility of Remote Gestures in Distributed Tabletop Collaboration. ACM Conference on Computer Supported Cooperative Work, 95-104.