Improving Visibility of Remote Gestures in Distributed Tabletop Collaboration

Naomi Yamashita[1], Katsuhiko Kaji[2], Hideaki Kuzuoka[3], Keiji Hirata[1]

[1] NTT Communication Science Labs.
Kyoto and Atsugi, Japan.
naomiy@acm.org
hirata@brl.ntt.co.jp

[2] Nagoya University
Nagoya, Aichi, Japan.
kaji@nuee.nagoya-u.ac.jp

[3] University of Tsukuba
Tsukuba, Ibaraki, Japan.
kuzuoka@iit.tsukuba.ac.jp

ABSTRACT

Collaborative distributed tabletop activities involving real objects are complicated by invisibility factors introduced into the workspace. In this paper, we propose a technique called "remote lag" to alleviate the problems caused by the invisibility of remote gestures. The technique provides people with instant playback of remote gestures to recover from the missed context of coordination. To examine the effects of the proposed technique, we studied four-person groups who engaged in two mentoring tasks using physical objects with and without remote lags. Our results show that remote lags effectively alleviated the invisibility problems, resulting in fewer questions/confirmations and redundant instructions during collaboration. The technique also decreased the overall workload of workers as well as the temporal demands for both helpers and workers.

Author Keywords

Video-mediated communication, remote gesture, tabletop, group-to-group collaboration

ACM Classification Keywords

H.5.3 Information interfaces and presentation (e.g., HCI): Group and organization interfaces-Computer-supported cooperative work

General Terms

Human Factors, Design, Experimentation

INTRODUCTION

Tangible tabletops play a significant role in many small-group co-located activities [11, 21]. One of the typical activities performed over them is collaboration dealing with physical objects. For example, people might share and discuss documents, a helper might guide a group of workers performing mechanic repair, and a group of experts might work together for future urban planning. Such tabletop collaboration dealing with physical objects has received extensive attention in recent CSCW research. In line with such trend in CSCW, our interest is in supporting physical collaboration by a group of people over distributed tabletops.

CSCW 2011, March 19-23, 2011, Hangzhou, China

Figure 1. Two workers and two helpers at a remote site working on a mentoring task using physical objects.

To achieve rich physical interactions over and around tables in distributed environments, a variety of systems have been developed and examined (e.g., [2, 4, 14, 16]). Results have so far demonstrated the importance of providing users with a shared visual space and embodying gestures on a tabletop surface. Arm embodiments have shown to be particularly useful because they enable all users to directly access artifacts using natural hand gestures [13, 23].

Based on previous findings, we developed a system called t-Room (Figs, 1, 5). By merging the utility of video conferencing and a shared tabletop, t-Room provides a group of remote collaborators with equal and direct access to a shared tabletop workspace as well as a coherent environment in which to accomplish interaction. So far, we have shown how the changes in seating position across different sites affect video-mediated communication in terms of the unity and satisfaction with the group's decision [27].

This paper reflects our subsequent experiences and observations acquired with this system. In our preliminary experiments with t-Room where distant users engaged in mentoring tasks using physical objects, we found many instances where collaboration appeared inefficient due to the invisibility of remote gestures. Even though gestures play a significant role in collaborative physical tasks [13, 15], remote gestures were often invisible or/and missed. For example, remote gestures displayed on the tabletop were frequently occluded by the real objects themselves ("*occlusion*"). Remote gestures were also easily missed when a worker was concentrating on the work at hand or his/her attention was directed elsewhere ("*diverted attention*"). Such invisibility problems were aggravated when helpers supposed that their gestures were seen by the remote worker ("*supposition of visibility*").

The problems concerning the visibility of remote gestures are frequently found in many distributed tabletop systems using video technology. To date, however, they have not received much attention, and the problem remains unresolved. In fact, the problem appears to be quickly resolved in dyadic collaborations, which have been studied extensively in previous research. It becomes more acute as the number of participants increase and/or the table gets larger.

In order to improve the visibility of remote gestures, we propose a visual augmentation technique (hereafter, *remote lag*). Remote lag is a time-lagged visualization that provides users with richer cues of the remote user's recent motion in shared workspaces. By overlaying remote lags on the immediate feedthrough (i.e., live video), the technique provides users with a second chance to perceive the remote gestures.

To investigate the effects of remote lag, we studied four-person groups divided into two pairs and across two locations who engaged in two mentoring tasks using physical objects with and without remote lags. Using a combination of qualitative and quantitative research methods, we show that remote lag substantially supported the perception of remote gestures by improving their visibility. As a result, remote lag helped to significantly reduce the number of questions and confirmations issued by the workers as well as the number of the helper's utterances. The technique also decreased the overall workload of workers.

In the remainder of this paper, we first explain the motivation of our study and describe how our observations led to the development of remote lag. Next, we review related research and show how the topic is relevant to many tabletop systems. Then we describe a laboratory study in which we tested our technique in a remote mentoring task using physical objects. We conclude with a discussion about why remote lags led to effective collaboration as well as issues raised by our study.

MOTIVATION

In our preliminary experiments with t-Room where four-person groups engaged in a mentoring task using physical objects, we found many instances where collaboration appeared inefficient. In particular, workers asked many questions, and helpers repeated the same instructions. Through casual observations over video recordings of the experiments, we realized that users were having trouble understanding remote gestures that were inevitably invisible or difficult to see from their site. We identified two factors that caused the problem and one factor that exacerbated the problem by coexisting with those factors. The first two factors mainly related to workers and the last one mainly related to helpers:

- *Occlusion*: Remote gestures displayed on the tabletop were frequently occluded by a local member's arms or/and objects placed on the table. In particular, helper gestures toward remote objects were often occluded by the actual objects themselves (Figure 2). Although some workers lifted the objects to see the hidden gestures, they had by then already disappeared.

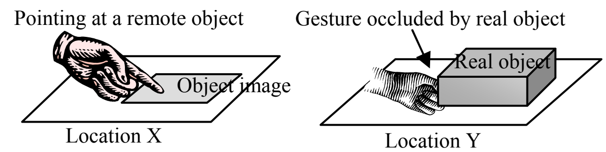


Figure 2. Example of occlusion

- *Diverted Attention*: Users also seemed to miss remote gestures when their attention was directed elsewhere. For example, workers tended to miss remote gestures when they were concentrating on the work at hand or their attention was directed to other ongoing activities by other members. Remote gestures appeared to be particularly unnoticeable when they were transected by the edge of the table, for example, leaving only the fingertip visible on the tabletop.
- *Supposition of Visibility*: Regardless of the invisibility of the above remote gestures, users tended to be unaware of the problems and supposed that their actions/gestures were receiving the attention of the remote users. For example, helpers tended to misinterpret that their gestures were receiving worker's attention when the workers were silent and/or seemed to be looking at the table.

To alleviate these problems, we improved the visibility of remote gestures by accompanying the live gestures with lagged ones. This idea was conceived from the intriguing fact that, in many cases, users were only slightly late in looking at the missed remote gestures. They often seemed to be aware of the existence of the missed gestures, presumably because the gestures were typically accompanied by such speech as "this one," "over here," etc. In the case of occlusion, users sometimes lifted the objects to see the remote gestures hidden behind them, which had just been displayed but were already gone when the object was lifted. These observations convinced us that the users had a fair chance of viewing the lagged gestures.

RELATED WORK

Embodiment of Remote Gestures

While various technologies are used to represent remote gestures (e.g., pointers [7, 8], pen-based drawings [4, 19], and human avatars [1]), gestures in distributed tabletop systems are typically represented through video-based embodiments [14, 16, 23, 25]. Such embodiments at workspaces were initially centered around collaborative drawing, although recent research (including this study) has put more focus on collaboration dealing with real objects [4, 14, 15, 17, 26] or/and tangible interfaces.

Studies along these lines have demonstrated the importance of supporting unmediated natural hand gestures [13] and constructing a coherent environment in which to accomplish interaction [9, 17]. A typical design for constructing such an environment fuses remote and local spaces into one hybrid workspace by directly projecting remote gestures into a physical workspace [13, 24].

Although significant progress has been made in the efficiency of distant collaboration, there are several invisibility problems to be solved in the understanding of remote gestures, which are typically found in distributed tabletop activities using video technology.

Invisibility of Remote Gestures

To consider technologies that support distant tabletop activities with video embodiment, we focus on the invisibility problems typically found in such systems. In this paper, we concentrate on three problems: occlusion, diverted attention, and the supposition of visibility.

Occlusion

Occlusion is a common issue in many distributed tabletop activities. The experimental findings by Tang et al. reported that the local user's hand often occluded the image of the remote collaborator's hand and complicated awareness of the current activity [24].

The problem could be observed more frequently when rear projection (including LCD) is employed. Yet, due to image clarity and small installation space, most tabletop systems adopt rear projection type display.

To avoid such a problem, some researchers introduced transparent objects into their tabletop systems [20]. However, this solution is difficult to implement on tabletops that use everyday objects as input devices. Since transparent objects cannot be substituted for ordinary everyday objects, the problem still occurs when they are brought on to the tabletop.

Diverted Attention

Video inevitably degrades the visual quality of three-dimensional structures in remote scenes, resulting in limited awareness of remote user activities. Gaver explicitly expressed this issue in relation to the flatness of displays [5]: "*movements in the remote space are largely two-dimensional and do not seem to impinge on the local scene.*" Consequently, a remote user's "*gestures appear to pass unnoticed by the potential recipient* [9]."

As the number of collaborators increase, remote gestures may pass unnoticed more frequently since user's attentions could be distracted by other activities.

Supposition of Visibility

Previous studies point out that when a user produces a gesture, he/she first determines whether the recipient can see and interpret it [10, 18]. If a gesture's producer knows that his/her gesture was not seen by the recipient, he/she either waits until the recipient sees it or exaggerates it to make it visible to the recipient [9].

Unfortunately, however, a gesture's producer using distributed tabletops often has problems determining whether his/her gesture was noticed by the recipient. The detection of diverted attention is particularly difficult because distributed tabletops usually do not convey the gaze awareness of remote users. Even inside t-Room where remote users' upper bodies were displayed, helpers sometimes misinterpreted the distant workers' locus of attention and started gesturing without the attention of the workers. According to Kraut, the producer of a gesture in distributed tabletops (or shared workspaces) typically assesses the observer's state by looking at the tabletop activities and listening to his/her speech. Such inference of the observer's state, however, will increasingly become difficult as the number of collaborators increases and tables get larger.

In the case of occlusion, the producer of a gesture tends to assume that gestures are visible to remote observers. According to Heath and Luff, "*Speaker shapes his/her gesture as if the potential recipient will view it in the way in which it is designed* [9]." For example, Yarosh implemented a shared play rug for distributed children whose technology resembles shared tabletops [28]. In their experiment using this system, children confused the above-beneath relationship (e.g., Fig. 2) when distributed objects were placed at the same position. Such confusion was also observed among adults in our experiment.

REMOTE LAG

To alleviate the problems caused by the invisibility of remote gestures, we propose a visual augmentation technique called *remote lag* that visualizes a remote user's recent motion in the shared workspace.

Remote lag is a technique similar to *traces* [7] and *local lags* [8]; like traces and local lags, remote lag visualizes the past motion of a workspace embodiment. While local lag introduces delay into the visual "feedback" (i.e., notification to local users as a result of their actions), remote lag introduces delay into the visual "feedthrough" (i.e., notification of their actions to remote users). By overlaying remote lags on the present workspace, the technique enables users to watch an instant playback of the

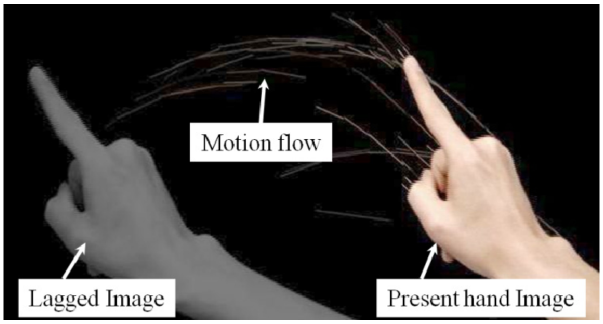


Figure 3. Representation of remote lag

remote gestures, providing them with a chance to recover from the missed context of coordination.

Gutwin [7] compared several possible representations for telepointer traces and determined that relatively short, low-contrast, fading motion lines are best for showing motion without adding excessive clutter to the display. We also compared several types of representations (varying time intervals between present and lagged images, the number of lagged images, their color and contrast, etc.) to determine the optimal representation of remote lags in the t-Room environment [12]. Based on our exploration, we adopted the following representation: 1.1-second interval between present and past images; lagged images in grayscale for easy discrimination; overlapped lagged images beneath current images so that the lagged images do not interfere with the current images in static scenes; instead of displaying multiple lagged images, we drew a set of fading line segments (hereafter, "*motion flow"*) between the current and lagged images to express the movement of gestures without making the workspace too cluttered. An example of remote lag representation is shown in Figure 3.

EVALUATION

Using a system called "t-Room," we carried out an experiment to investigate the effects of remote lags on collaborative physical tasks. Specifically, we sought to answer the following questions:

- Does remote lag affect collaborator interactions? Does it help them mitigate invisibility problems?
- Do remote lags improve conversational efficiency by reducing unnecessary questions/confirmations and redundant instructions?
- Do remote lags impose a greater mental load (e.g., effort, frustration) on users to perceive remote gestures?

Video System: t-Room

t-Room is a room-duplication system that supports distributed tabletop activities. It provides remote collaborators with equal and direct access to a shared tabletop workspace as well as a coherent environment in which to accomplish interaction. While other systems (e.g., [24]) restrict each user's position (i.e., do not allow various arrangements or walking around the table), t-Room supports flexible user arrangements around the table [27], which is essential in co-located tabletop activities [21].

t-Room consists of multiple cameras, screens, and speakers (Figure 4); its walls consist of eight 65-inch LCD panels, HDV cameras placed above them, and speakers placed below them; a table is located at the center of the room and includes two 40-inch LCD panels and a HDV camera.

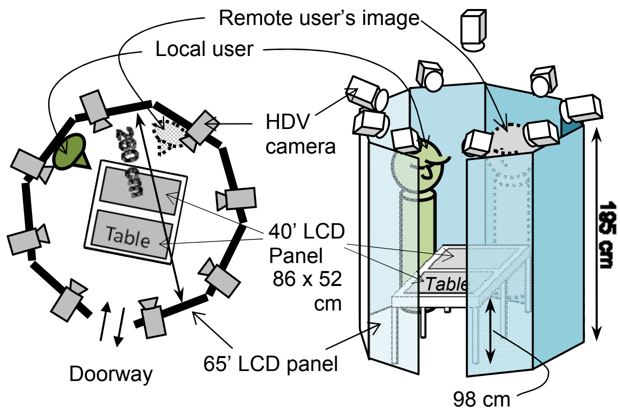


Figure 4. Hardware design of t-Room: Top-view (left) and Bird's eye-view (right)

Local activities above and around the table are captured through these cameras and displayed on LCD panels in the remote t-Room. The arrangements of video cameras and screens are configured to maintain spatial relationships between distant sites [27, 22]. Local sounds are also captured with wireless microphones attached to each user and played on the speakers at remote t-Rooms. Audio, however, is not spatialized in this study, since we haven't tracked each user's movement (or head position).

Method

We based our study on four-person collaboration because t-Room is designed to support group-to-group collaboration across distances. Although tables are inherently designed to support group work, the vast majority of studies on distributed tabletops have explored pairs, with such exceptions as [24].

Conditions

We compared conditions between feedback+feedthrough (i.e., normal version) vs. feedback+feedthrough+remote lags. Note that lagged gestures were only provided to remote participants but not to local participants. Participants in both conditions were provided with immediate feedback of their own gestures.

Apparatus

Two identical t-Rooms were installed in the cities of Atsugi and Kyoto, which are approximately 150 miles apart, and connected by a gigabit network. The network delay for video and audio transmission between the two cities was around 0.3-0.4 and 0.2-0.3 seconds; audio and video were not synchronized.

Participants

Twenty groups of four adults were recruited for this study and were paid for their participation. To screen out extraneous factors that may influence the results, we recruited only females who did not know each other; we chose females to avoid gender factors that might have interfered with participant positions around the table; strangers avoided familiarity factors that might interfere with collaboration efficiency. Participants had never used a video-conferencing system or a shared tabletop system prior to the experiment.

Task

Each task consisted of four participants, with two participants collocated in one location (Kyoto) and two in another location (Atsugi). Each collocated group was assigned a separate role: the Kyoto group was the worker role and the Atsugi group was the helper role. Lego Duplo pieces [1] were randomly placed on the Kyoto tabletop; workers had physical access to the actual pieces but helpers could only access them virtually. Helpers were handed out with a final completion map (Figure 5) and asked to instruct the workers to assemble the initially scattered pieces (assembly subtask), and place them precisely in relation to the white loop displayed on the tabletop surface (placement subtask). One helper and one worker were then randomly paired to form two helper-worker pairs who worked with different pieces: rectangular or non-rectangular.

Figure 5. An example of final completion map

As described above, the task employed in this study consisted of two subtasks: assembly and placement. Assembly subtask, which is characterized by physical manipulation (rotating, attaching, etc.) of the items, is mainly used in studies focusing on cooperative physical activities (e.g., [4, 14, 19]). In contrast, placement subtask requires high precision in placing the items and mimic the collaborative designing tasks (e.g., urban planning) used in many tabletop studies [26]. The task as a whole encompasses a variety of generic task elements, such as item selection, pattern matching, physical manipulation and error checking. In this respect, the task offers the opportunity to explore some of the demands that may be placed on real world applications in tabletop activities using real objects.

Design

Each group participated in two tasks with and without remote lags; the design formed a counter-balanced within-subjects design, comparing conditions with and without remote lags.

Procedure

Step (1): On arrival, participants in the same group were placed in separate rooms so that those at the same location did not get acquainted before the experiment. Next they completed experimental consent forms and moved to the t-Room. Participants were given a brief introduction to the study and the t-Room system, introduced to each other, and randomly divided into two helper-worker pairs.

Step (2): The following procedure was repeated two times: with and without remote lags. In each condition, participants were given a brief explanation of the system or/and the remote lag technique, and each pair was given an opportunity to practice giving and receiving instructions using the Lego pieces. Participants were then given 15 minutes to complete each task and were told the following: a) their task was to assemble and place the Lego pieces so that it matches the final completion map; b) they should basically work within pairs but they can help others as required; c) they could use any strategy to accomplish the task; and d) they should complete the task as quickly as possible. Following the completion of each task, participants completed post-task questionnaires about the work they had just experienced.

Step (3): Following the completion of the two tasks, participants completed a final questionnaire on the effects of remote lags. Finally, they were interviewed about the differences in the work between the two conditions. Participants in Kyoto were interviewed face-to-face and participants in Atsugi were interviewed through the t-Room.

Measures

We employed both qualitative and quantitative measures to examine the activities carried out in our study. Each session was video-taped, and we logged participant activities within the workspace. Speech data were transcribed from video recordings to analyze communicative efficiency, and video data were used for more detailed analysis on how remote lags influenced collaboration. NASA task load index (TLX) questionnaires [29] were included in the post-task questionnaires to measure the mental load remote lags imposed on each participant. Finally, interview results identified problems with remote lags and were used to develop ideas for improvement.

RESULTS

We present the results in three parts: first we investigate the effects of remote lags on distant collaboration by looking closely at how they actually helped mitigate the invisibility problems addressed in this paper; then to see how much, in total, remote lags helped reducing unnecessary utterances,

[1] Lego Duplo pieces are approximately twice as large as regular Lego pieces.

we examined the effects of remote lags on communication efficiency; finally, we examined the questionnaire results.

Invisible Gestures

Starting with the condition without remote lags, we carefully reviewed each tape and looked for instances of the invisibility issues raised in this paper. Each instance was categorized into one of the two categories: "occlusion" and "diverted attention." Since "supposition of visibility" coexisted in each of those instances, this was not included in the coding category. We then reviewed the video recordings of the remote lag condition and looked for instances in which problems once occurred but was resolved by using the remote lag. Easily understandable cases are selected and provided below.

Occlusion

Occlusions seemed to cause redundant interactions between collaborators, particularly when the helper's verbal instructions accompanying the occluded gestures were not informative enough to follow (e.g., "Do it like this," "Put it over here"). Such scenes were often found when the actions to be made were difficult to explain in words (i.e., when complicated gestures were used) or when collaborators were making fine adjustments in placing the assembled components.

FRAGMENT 1

```
01   H: kono shikaku no ah midori no kado ga::
        this square         green      corner
     W:                                          hai
                                                 yes
02   H: konoatari made  sagete  moratte iidesuka?
        till around here move   could you
     W:
03   H: (3.0) hai so::de::s
              yes  right.
         *Fig.6
     W:
```

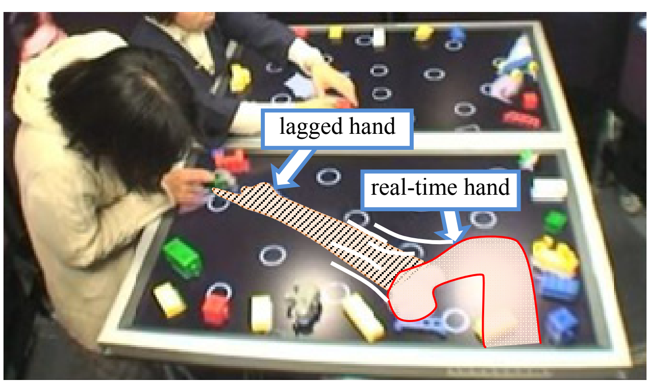


Figure 6. Worker using lagged gestures (for occlusion). For print legibility, we traced the remote gestures and motion flows.

As shown in the following instance (Fragment 1), remote lags improved the visibility of the occluded gestures by providing collaborators (workers) a chance to view the hidden gestures as they lifted up the actual objects. For verbal utterances in the fragment, the transcript gives the Japanese original with a literal English translation below.

In this fragment, the helper-worker pair in the foreground is finely adjusting the position of a green cube; the worker wearing a white jacket has placed it in the approximate location, and the helper is requesting that it be moved slightly downward until its corner touches her index finger. The pointing gesture, however, is occluded by the green block itself, so the worker cannot see the pointed location. The worker thus lifts up the green block to see the pointed location. This worker's action, however, is taken as a sign of comprehension (i.e., *supposition of visibility*); as the worker lifts the block, the helper quickly withdraws her hand from the block. As a result, the pointing gesture has already vanished when the worker looks beneath the block. She manages to determine the pointed location by looking at the lagged gesture (Figure 6) and then successfully puts the block at the specified position.

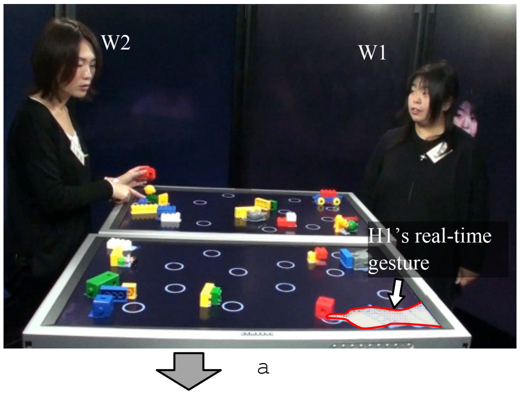
Diverted Attention

In our study, workers sometimes had their attention caught by activities unrelated to their current work (e.g., the other pair's interaction) by some immediate concern that arose from their work (e.g., the blocks did not fit tightly) or by conversation with another local participant. However, once the helpers started gesturing, it appeared that they rarely noticed the absence of the worker's attention (i.e., *supposition of visibility*); when a worker's attention was directed back to the current work, the helper's instructions were typically going a step ahead, leaving the workers behind. As shown in Fragment 2, remote lags helped the workers recover from such missed gestures and catch up with the real-time instructions. Note that '<X>' marks the point at which the worker became aware of the lagged gesture.

In Fragment 2, the workers on the right (W1) and left (W2) were discussing a possible mistake in the construction of the blocks. Fig.7a shows where the conversation has just stopped. At this moment, W1 is attending to W2 and not paying attention to the remote helper (H1). H1, however, seems to be unaware of W1's locus of attention, and starts pointing at a red block while uttering "red square". After uttering "red square," she pauses and glances at W1 to see if W1 understood her instruction. Meanwhile, W1 is staring at the table, presumably wondering about the conversation she just had with W2. Again, however, H1 is not aware of W1's locus of attention. Since W1remains silent, H1 supposes that W1 is paying attention. She proceeds with her instruction; she says "bring," as she moves her hand to the destination where the red block should be placed. Just then, W1 looks at H1 (Fig. 7b) and realizes that H1 was giving an instruction to her. She quickly shifts her attention to the tabletop (Fig. 7c) and notices the lagged gesture, which is still pointing at the red block and some motion flows. W1 then shows comprehension by uttering "red block" and quickly grabs the red block (Fig. 7d).

Overall, remote lags seemed to help the workers recover from the missed gestures by providing them a second chance to view those gestures. By following the lagged

FRAGMENT 2

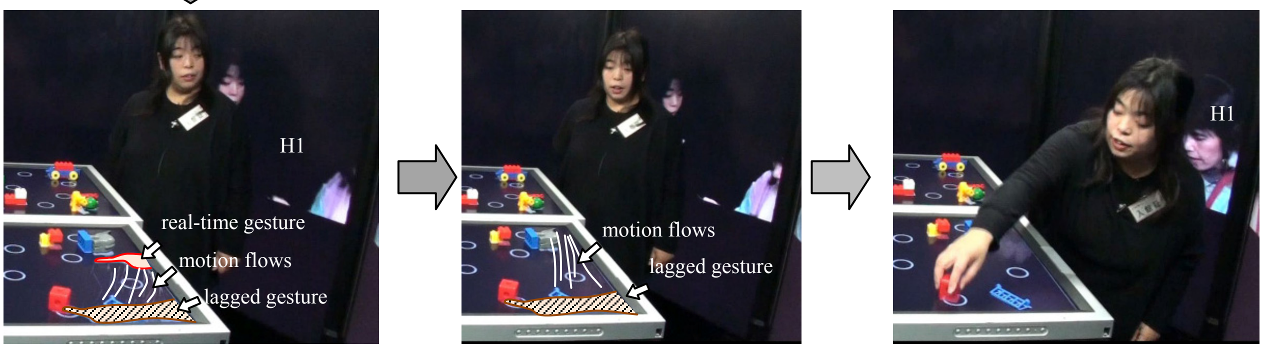



```
01   H1:aka no shikaku wo   (2.7)
        red     square
                  *a
     W1:

02   motteitte itadaite  (0.6)  kokorahen oite moraemasuka?
        bring it to                here     put  please?
                    *b      *c         *d
     W1:                 <X>akano burokku
                             red   block
```

<X>: point at which worker (W1) sees helper's (H1's) lagged hand.

a



b c d

Figure 7 Worker using lagged gestures (for diverted attention). For print legibility, we traced the remote gestures and motion flows.

gesture, workers managed to catch up with the real-time instruction without causing a big loss in their collaboration process.

Conversational Efficiency

To investigate how much remote lags decrease the overall redundancy of communication between collaborators, we compared the mean number of utterances per task by condition for each experimental role (Figure 8) [4, 6]. Because participants were nested within pairs within groups, they were analyzed using a repeated measure mixed model analysis of variance. Results indicated that the amount of helper talk was significantly less in the remote lag condition than the normal condition ($F[1,39]=4.51$, $p<.05$). The amount of worker talk did not differ significantly between the two conditions.

To further investigate how much remote lags contributed in decreasing unnecessary questions and confirmations, we counted the number of utterances that asked a question or made a confirmation for each participant in each task. Figure 9 shows the mean proportion of utterances about questions and confirmations per task by condition for each experimental role. Results indicate that workers asked questions and confirmations less frequently in the remote lag condition than in the normal condition. A repeated measures mixed model ANOVA indicated a significant difference between the two conditions ($F[1,39]=5.32$, $p<.05$). The frequency of questions and confirmations

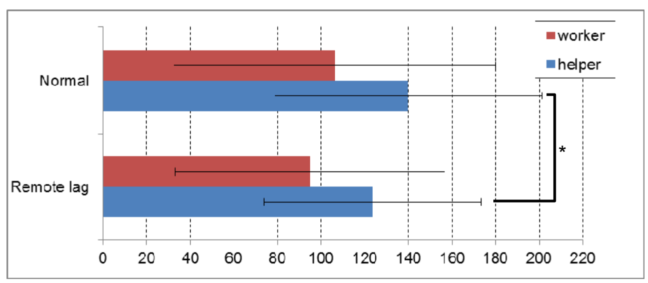


Figure 8. Mean number of utterances per task by condition for each experimental role; "*" denotes significant difference at $p < .05$ level.

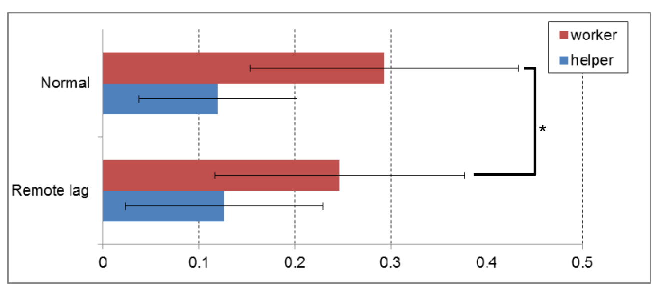


Figure 9. Mean proportion of utterances on questions and confirmations per task by condition for each experimental role; "*" denotes significant difference at $p < .05$ level.

issued by helpers did not differ significantly between the two conditions.

In sum, it appears that remote lags helped the workers achieve better understanding of a helper's instructions by increasing the visibility of the helper's gestures, resulting in a lower frequency of asked questions and made confirmations. We infer that fewer helper utterances in the remote lag condition resulted from the decreased need for answering questions and confirmations from the workers.

Subjective Perception of Workload

Although results have so far indicated that remote lags facilitate distant collaboration, it remains unclear how much extra effort is needed to notice the remote lags. For example, one apparent defect of introducing remote lags in a shared workspace is the added screen clutter to the workspace, which may require more concentration and effort during collaboration.

To see the effects of remote lags on participant perception of effort during collaboration, we measured participant subjective workload using NASA TLX, which is a multi-dimensional rating procedure that derives an overall workload score ranging from 0 to 100 (the lower the better) based on a weighted average of participant ratings on six subscales: mental, physical, and temporal demand as well as overall effort, frustration, and perceived performance. This index is widely used to measure the subjective workload of people working with various human-machine interfaces.

Figures 10 and 11 show the average rating scores for each subscale of workers and helpers. Overall, the workloads for workers in the normal and remote lag conditions were 43.4 (SD=23.3) and 38.6 (SD=19.9), respectively. A repeated measures mixed model ANOVA indicated that workers perceived significantly lower workload in the remote lag condition ($F[1, 39]=5.26$, $p<.05$). Similarly, the overall workloads for helpers in the normal and remote lag conditions were 61.8 (SD=19.3) and 58.3 (SD=22.5), respectively. Although helpers seemed to perceive less workload in the remote lag condition, the difference was not significant ($F[1, 39]=2.38$, $p=.13$).

A repeated measures mixed model MANOVA was conducted to test the differences on each TLX subscale. The results indicated both workers and helpers perceived a significantly lower temporal demand in the remote lag condition than in the normal condition (worker: $F[1, 39]=5.65$, $p<.05$; helper: $F[1, 39]=4.40$, $p<.05$). However, results indicate that helpers perceived a significantly higher physical demand ($F[1, 39]=6.92$, $p<.05$) in the remote lag condition than in the normal condition. We speculate that helpers perceived higher physical demand in the remote lag condition because they finished the task faster, making them busy instructing the workers one after another. The scores of other subscales did not differ significantly between the two conditions.

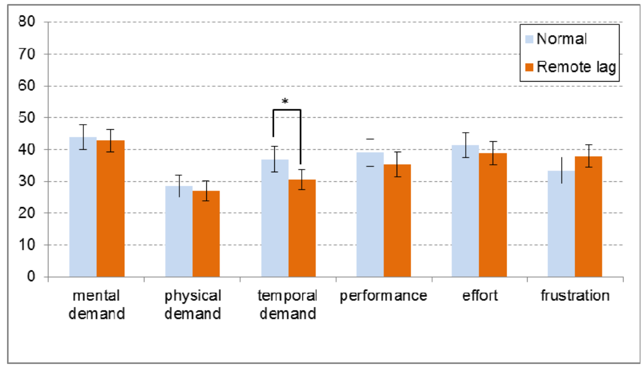


Figure 10. Mean TLX ratings of workers; "*" denotes significant difference at *p* < .05 level.

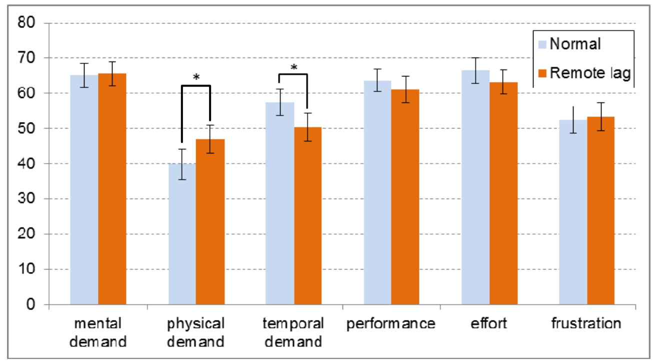


Figure 11. Mean TLX ratings of helpers; "*" denotes significant difference at *p* < .05 level.

DISCUSSION

In the following, we discuss two issues arising from our study: why remote lags appeared useful, and whether the benefits can be generalized to other distributed tabletop systems and other tasks.

Why Remote Lags Appeared Useful

Previous research demonstrated the synchronous nature between speech and action. According to Clark and Krynch [3], speech is particularly useful when it is precisely timed to actions. Additionally, Gergle et al. showed that even a small delay in updating visual space reduced the value of visual information and degraded performance [6].

Given that the simultaneity of speech and action is crucial, it is interesting to consider why the participants understood the lagged gesture of 1.1 second delay in our experiment.

The post-experimental interviews provide a hint: from them, we found that most participants normally heeded the live gestures and only looked at the lagged gesture when they had trouble viewing the live gestures. For example, one participant mentioned in her interview, "*I normally ignored the lagged gestures. I only followed them when I missed the live gestures.*"

As suggested by her comment, participants exploited the lagged gestures when they *knew* that they had missed the live gestures. Conversely, the lagged gestures appeared ineffective when participants did not realize that they had missed the remote gestures (e.g., when they mistook the other helper's gesture as a corresponding gesture to the current speech).

Thus, we infer that users can exploit the lagged gestures when they miss a gesture by a mere touch but can estimate its existence from verbal or/and visual cues. As for the verbal cues, it is not surprising if a user understands a lagged pointing gesture that arrives a moment after he/she hears "this one." Such speculation reminds us of Clark's "projective pair," which suggests that people typically anticipate a forthcoming gesture when they hear deictic expressions [3].

As for the visual cues, the workers in our experiment seemed to expect lagged gestures from a partially occluded live gesture, a passing live gesture, or a motion flow that remained on the tabletop display. Note that the lagged gestures were not presented independently; they were presented as an awareness of the live gestures. In fact, we speculate that lagged gestures would not improve collaboration if presented independently of the live gestures. Both live and lagged gestures must be presented as a *unit* that allows users to flexibly shift their attention between them based on their needs.

Applying Remote Lags to Other Systems/Tasks

In our study, we examined the effects of remote lags by implementing the technique in a video system that merges video conferencing with a shared tabletop. Aligned with the argument in [17], the video conferencing (i.e., vertical displays) in our experiment allowed the appearance of cues to anticipate the forthcoming tabletop gestures. Since most existing distributed tabletop systems are not equipped with such vertical displays, the participants using such systems would probably miss the remote gestures more frequently and require more effort to recover the invisible gestures. Therefore, we believe that remote lags will become more useful when implemented in distributed tabletop systems without vertical displays.

As for the tasks, we showed the advantages of remote lags on a mentoring task using physical objects. Although the invisibility problems caused by occlusion are inherent in tasks with physical objects, the other two problems (i.e., diverted attention and the supposition of visibility) may occur in a wider range of tabletop activities. Furthermore, remote lags may become more operative when the number of participants and the table size increase, because invisibility problems tend to occur more frequently in such situations.

CONCLUSIONS AND FUTURE WORK

In this study, we investigated three problems concerning the invisibility of remote gestures: occlusion, diverted attention, and the supposition of visibility. Although these problems do not appear significant in distant collaboration between pairs, they become serious as the number of collaborators increases and/or the table gets larger. The problem is exacerbated when collaborators change their arrangements during work. Although such situations match more practical tabletop activities [21], virtually no studies have investigated such collaboration or have suggested support for such collaboration. Based on our observations, we proposed a technique (i.e., remote lag) to mitigate the invisibility problems. Our experiment compared four-person group collaboration with and without remote lags and found that they greatly reduced the negative effects of the invisibility problems.

Previous studies on distributed tabletops have typically transmitted finger/pen contacts with a fading trail to support physical collaboration across distance [4, 24]. Using such techniques, users could make simple line drawings on a tabletop surface. Compared to remote lags, those techniques may be useful when there are only small movements during collaboration (as in the assembly subtask in our experiment) since it could avoid excessive screen clutter (explained below). Meanwhile, remote lags may be more useful when there is a relatively large shift in the locus of attention over the table (e.g., when a helper reaches out to point an object) because people tend to pull their hands off the tabletop surface in such a scene.

From the post-experimental interviews, we further identified several issues that may affect the actual deployment of the remote lag technique:

Screen clutter: Some participants mentioned that remote gestures with small finger movements were at times difficult to follow because the lagged gestures and the motion flows disturbed their appearances by adding too much clutter to a small space. One way to avoid such problems is to increase the transparency of the remote lag when the hand motion is more minute than a threshold.

Color of lagged gestures: In our study, lagged gestures were shown in grayscale. Although grayscale-lagged gestures enabled the participants to easily distinguish between current and lagged images, several participants complained about the difficulties of distinguishing lagged gestures between different people. Workers seemed to face particular difficulties when both helpers were standing close to each other and gestured toward relatively close objects. One way to avoid this problem is to show the lagged gestures in inconspicuous colors rather than the real-time gestures so that they could preserve clues (e.g., color of sleeves) to distinguish between different collaborators at a remote site.

Our next step is to explore the improvements raised here and carry out the technique in actual use.

ACKNOWLEDGMENTS

This work is partially supported by Grant for Scientific Research (C) 20500122. We thank Shigemi Aoyagi for help in running the experiment. We also thank Prof. Paul Luff and the reviewers for their constructive comments.

REFERENCES

1. Benford, S., Bowers, J., Lennart, E. F., Greenhalgh, C., and Snowdon, D., 1995. User embodiment collaborative

virtual environments. *Proceedings of CHI'95,* ACM Press (1995), 242-249.

2. Brave, S., Ishii, H., and Dahley, A., Tangible Interfaces for Remote Collaboration and Communication. *Proceedings of CSCW'98,* ACM Press (1998), 169-178.
3. Clark, H. H. and Krych, M. A., Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50(1)*, (2004), 62-81.
4. Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E. and Kramer, A., Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks. *Human-Computer Interaction, 19*, (2004), 273-309.
5. Gaver, W. The affordances of media spaces for collaboration. *Proceedings of CSCW'92*, ACM Press (1992), 17-24.
6. Gergle, D., Kraut, R.E., and Fussell, S.R., Language Efficiency and Visual Technology: Minimizing Collaborative Effort with Visual Information. *Journal of Language and Social Psychology, 23(4)*, (2004), 491-517.
7. Gutwin, C., Improving interpretation of remote gestures with telepointer traces. *Proceedings of CSCW'02*, ACM Press (2002), 49-57.
8. Gutwin, C., The Effects of Local Lag on Tightly-Coupled Interaction in Distributed Groupware. *Proceedings of CSCW'08*, ACM Press (2008), 447-456.
9. Heath, C. and Luff, P., Disembodied Conduct: Communication Through Video in a Multi-media Office Environment. *Proceedings of CHI'91,* ACM Press, (1991), 99-103.
10. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., and Greenhalgh, C., 1998. Fragmented Interaction: establishing mutual orientation virtual environments. *Proceedings of CSCW'98,* ACM Press, (1998), 217-226.
11. Jordà,S., Geiger, G., Alonso, M. and Kaltenbrunner, M., The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. *Proceedings of TEI'07,* ACM Press, (2007), 139-146.
12. Kaji, K., Yamashita, N. and Hirata, K. A Method for Supporting Recognition of Remote Gestures in a Video Communication System. *Proceedings of ASIAGRAPH'09*, (2009), 84-89.
13. Kirk, D., Crabtree, A., and Rodden, T., Ways of the hands. *Proceedings of ECSCW'05*, Springer (2005), 1-21.
14. Kirk, D. Fraser, D., Comparing Remote Gesture Technologies for Supporting Collaborative Physical Tasks. *Proceedings of CHI'06,* ACM Press, (2006), 1191-1200.
15. Kraut, R. E., Gergle, D., and Fussell, S. R., The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-Presence. *Proceedings of CSCW'02*, ACM Press (2002), 31-40.
16. Kuzuoka, H., Yamashita, J., Yamazaki, K., and Yamazaki, A. Agora: A Remote Collaboration System that Enables Mutual Monitoring, *Extended Abstracts of CHI'99*, ACM Press (1999), 190-191.
17. Luff, P., Heath, C., Kuzuoka, H., Yamazaki, K., and Yamashita, J., Handling documents and discrimination objects in hybrid spaces. *Proceedings of CHI'06,* ACM Press (2006), 561-570.
18. Moore, R., Ducheneaut, N., and Nickell, E., Doing Virtually Nothing: Awareness and Accountability in Massively Multiplayer Online Worlds, *JCSCW, 16, 3,* (2006), 265-305.
19. Ou, J., Fussell, S. R., Chen, X., Setlock, L. D., and Yang, J., Gestural communication over video stream: Supporting multimodal interaction for remote collaborative physical tasks. *Proceedings of ICMI'03,* ACM Press (2003), 242-249.
20. Sato, T., Mamiya, H., Koike, H. and Fukuchi, K., PhotoelasticTouch: transparent rubbery interface using a LCD and photoelasticity, *Proceedings of UIST'09,* ACM Press (2009), 43-50.
21. Scott, S. D., Grant, K. D., and Mandryk, R. L., System guidelines for co-located collaborative work on a tabletop display. *Proceedings of ECSCW'03,* Springer (2003), 159-178.
22. Sellen, A. Buxton, B., and Arnott, J. Using spatial cues to improve videoconferencing. *Proceedings of CHI'92*, ACM Press (1992), 651-652.
23. Tang, A., Neustaedter, C., and Greenberg, S., VideoArms: embodiments for mixed presence groupware. *Proceedings of British-HCI'06*, Springer (2006), 85-102.
24. Tang, A., Pahud, M., Inkpen, K., Benko, H., Tang, J. C., and Buxton, B., Three's Company: Understanding Communication Channels in Three-way Distributed Collaboration. *Proceedings of CSCW'10,* ACM Press, (2010), 271-280.
25. Tang, J. and Minneman, S., VideoDraw: a video interface for collaborative drawing. *ACM Transactions on Information Systems 9,* 2 (1991), 453-469.
26. Underkoffler, J. and Ishii, H., Urp: A luminous-tangible workbench for urban planning and design. *Proceedings of CHI'99*, ACM Press (1999), 386-393.
27. Yamashita, N, Hirata, K, Aoyagi, S, Kuzuoka, H, Harada, Y., Impact of Seating Positions on Group Video Communication, Proceedings of *CSCW'08*, ACM Press (2008), 177-186.
28. Yarosh, S. Inkpen, K., and Brush, A. J. B., Video Playdate: Toward Free Play across Distance. *Proceedings of CHI'10,* ACM Press, (2010), 1251-1260.
29. http://www.nrl.navy.mil/aic/ide/NASATLX.php