Parallel Eyes: Exploring Human Capability and Behaviors with Paralleled First Person View Sharing

Shunichi Kasahara[1,2]
kasahara@csl.sony.co.jp

Mitsuhito Ando[3]
ando@ycam.jp

Kiyoshi Suganuma[3]
suga@ycam.jp

Jun Rekimoto[1,2]
rekimoto@acm.org

[1]SonyCSL, [2]Tokyo Univ.
Tokyo, Japan

[3]YCAM Interlab
Yamaguchi, Japan

ABSTRACT
Our research explores how humans can understand and develop viewing behaviors with mutual paralleled first person view sharing in which a person can see others' first person video perspectives as well as their own perspective in realtime. We developed a paralleled first person view sharing system which consists of multiple video see-through head mounted displays and an embedded eye tracking system. With this system, four persons can see four shared first person videos of each other. We then conducted workshop based research with two activities, drawing pictures and playing a simple chasing game with our view sharing system. Our results show that 1) people can complement each other's memory and decisions and 2) people can develop their viewing behaviors to understand their own physical embodiment and spatial relationship with others in complex situations. Our findings about patterns of viewing behavior and design implications will contribute to building design experience in paralleled view sharing applications.

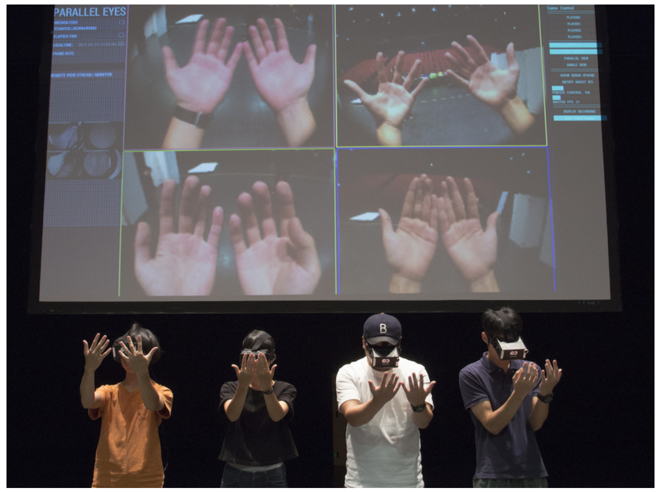


Figure 1. Parallel Eyes: a technology probe for creating a mutual visual shared parallel experience. With Parallel Eyes, users wear a head mounted display which shows shared videos of each other's perspectives as well as one's own perspective in realtime.

Author Keywords
Video Communication; Shared Experiences; Wearable; Workshop; First Person Video; Parallel Experiences

ACM Classification Keywords
H.5.3. Information interfaces and presentation: Group and Organization Interfaces - Computer Supported Cooperative Work

INTRODUCTION
As small wearable networked cameras emerge as common technologies, first person video is becoming a significant form of media for sharing one's experiences. The combination of a wearable camera and a head mounted display (HMD) will provide significant benefits for parallel experiences in various applications. However, first person video sharing is neither naturally used in daily life nor professional domains. Why not? In addition to the need for technological improvement and the lack of social acceptance for wearable cameras, we think that one of the missing pieces is the exploration of design sharing configurations beyond the "local-remote" (a.k.a. shooter-watcher) model. Current applications using first person video streaming are commonly designed to have users take on explicit roles as a shooter (e.g., using a head mounted camera) and a watcher (e.g., using a mobile screen or a computer screen). We think that mutual view sharing is one of the possible configurations beyond them. It will make a practical impact upon remote activities, education, and more professional mechanical operations, service professionals for providing hospitality.

Jason Procyk et. al [30] explored paired first person video exchanges with head mounted cameras and mobile screens. Their work showed that paired mutual video streaming achieved parallel experiences in public settings. However, there are still a lot of uncertain design challenges in more complex symmetric configurations where two or more persons exchange their own perspectives with each other. We especially focus on the research question of whether humans can perceive and understand multiple perspectives, including those of one's self and others. Although those visual situations seem to be highly complicated, findings and implications from this exploration can be generalized to future settings where one will be able to stream one's own perspective and to see others' perspectives simultaneously with wearable devices.

DOI: http://dx.doi.org/10.1145/2858036.2858495

To explore this scenario, we developed a technology probe [12] called "Parallel Eyes," which is a system in which users wear a HMD which shows multiple shared first person video perspectives as well as one's own perspective in realtime. This technology probe generates a mutual visual shared parallel experience among four persons in realtime (Fig. 1).

Our aim is to investigate and analyze how humans acquire and develop their viewing behaviors in such paralleled visual sharing situations. Therefore, we designed two activities. The first one is "faint memory drawing," where participants draw something on the basis of only their memory, with the exception of seeing others' perspectives. The second one is "game of tag," where participants play a simple game with Parallel Eyes. These two activities provide a general basis for future settings where multiple persons may exchange their perspectives on intellectual work, physical cooperation, and even technology mediated sports.

As an approach of this exploration, we conducted a study through a workshop to investigate how future prospective users use the technology and develop their own behaviors and abilities in the wild. We then analyzed an observation of our workshops, user behaviors, and recorded data. Our results from the drawing activities show that our technology probe dynamically complemented the collective memory and decisions taken among four persons during the drawing process. We also found that participants in the game of tag remarkably developed several strategies to understand their own physical embodiment and spatial relationship with each other.

The revealed patterns of viewing behaviors indicate design factors for both conscious and unconscious action. Those findings led to design implications which will contribute to building design experience in mutual paralleled view sharing applications.

RELATED WORK

Video as communication media

Video mediated communication is an active research domain. Previous research studied the possibilities and limitations of video mediated communication for face-to-face video conferencing [15, 26, 10, 7] or shared spatial workspaces for remote collaboration [21, 2].

Towards achieving a useful video mediated communication system for daily life, Brubaker et al. studied video communication across personal and professional domains by focusing on shared experiences[1]. The study yielded various insights into social practices and mentioned the possibility of multiple video streams for shared experiences. Experiences2Go [14] studied video mediated sharing experiences by using a mobile device and an in-home display. Both studies have provided insight into various user behaviors and revealed new findings in regard to video mediated communication.

Gaver et al. have presented the possibility of multiple views in a media space for remote collaboration; however, a common issue has arisen regarding an increase in the complexity of video mediated communication[4]. As wearable devices emerge, we should also consider the human capability to perceive the complexity of multiple video media in a wearable setting. Our work especially explores whether humans can adopt and develop their ability for visual complexity in a wearable setting to achieve a mutual parallel experience.

Sharing first person video

Recent advances in wearable technologies have increased the importance of research in the sharing of first person video, especially in the application domains of tele-presence[31], remote assistance [13, 19] and remote collaboration [5, 2]with various configurations. In some research, augmented reality graphics played an important role in conveying information from a remote person to a local person [11, 3, 17, 28].

These emerging technologies enable novel methods for performing live entertainment with first person video streaming that generate social engagement in public [32]. These related studies about first person video indicate the possibilities of broader applications for sharing experiences through first person video, which would require broader implications for designers and researchers.

Parallel Experience

Various studies have been conducted as to how technology can encourage shared physical activities with others. For instance, Mueller et al. proposed sharing jogging activity over the Internet with sound communication [25]. Most closely related to our study, Jason Procyk et. al [30] explored a parallel experience in which a pair of people performed an outdoor activity while video streaming from a head mounted camera and using a smartphone attached to an armband to view exchanged first person video. This research yielded various findings about parallel experience sharing in a social context. Our research focuses on parallel experiences with a more intimate, complex visual exchange to examine human capability as well as to explore social contexts.

Interconnected, unbound, transformed perspective

The relationship between visual perspective and embodiment is tightly connected. Therefore, unbinding and reconstructing the relationship forms the basis for an interesting field of research ,art and technical development to provide brand-new experiences. Hachiya[6] created an art performance involving the swapping of visual perspectives between a pair of humans for the purpose of unbinding body and perspective to achieve a unique form of inter-personal communication. The Machine to be Another project [27] achieved illusional embodiment by means of swapped visual perception enabled by well designed procedures. Kasahara et al. achieved visual tele-presence by means of omnidirectional video[18], which enables independent observation by remote persons.

These projects indicate a prospective research agenda that revolves around inter-connected human communication. Especially considering how a third person view of one's own body image provides an extended understanding of one's own body [24, 8]. Accordingly, our technology probe will bring about an inter-connected visual experience between humans which provides a different perspective - one that enables the viewing of one's own body from the perspective of others.

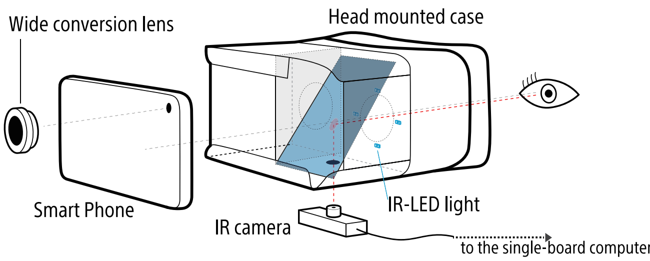


Figure 2. Parallel Eyes goggles: a video see-through HMD with an embedded eye tracking system.

TECHNOLOGY PROBE: PARALLEL EYES

To create a mutual visual shared parallel experience and to investigate and analyze how humans acquire and develop their behaviors, we developed the Parallel Eyes system as a technology probe. With Parallel Eyes, users wear a HMD which shows each other's shared perspectives as well as one's own perspective in realtime. Furthermore, the embedded eye tracking camera reveals how users watch displayed paralleled shared first person video.

HMD view sharing system

The Parallel Eyes system includes a HMD (Google Cardboard) with a smartphone screen as shown in Fig. 2. The smartphone also has a camera that captures the wearer's field of vision so that we can create a video see-through HMD. However, the field of view of a typical smartphone (for instance, the iPhone 5s we used) is not wide enough for a video see-through HMD. We also attached a conversion lens to acquire a wider view (approx. 100 deg).

In the Parallel Eyes system, four head mounted smartphones are connected via a wireless connection. We developed specially designed applications on the smartphone for the Parallel Eyes system. The video captured from the camera of the smartphone is displayed on the smartphone screen, and it is additionally transmitted to other head mounted smartphones. The video transmission is achieved with jpeg compression of the video image with the Turbo-jpeg library and the ZeroMQ Protocol [1] over Wi-Fi. The data bandwidth for video streaming is 0.6-2.4 Mbps for each video transmission, thus 9Mbps-40Mbps for the total system. The video resolution can be adjusted from 80 x 60 pix to 320 x 240 pix to keep the minimum frame rate at 20 fps and the maximum frame rate at 30 fps.

On each smartphone display, the wearer can see a video stream of the others' fields of view as well as one's own perspective. These video images are displayed with a four split layout as shown in Fig. 4 and Fig. 5.

We designed the number of video connections and the image layout with the following design rationale: (1) To produce complicated situations, more than two other perspectives should be displayed in addition to one's own perspective. (2) There is a limitation of wireless bandwidth for interconnected video transmission. Interconnected video streaming among four smartphones requires 12 sessions at the same time. (3) There is a limitation in the field of view of the head mounted smartphone display. (4) The eye tracking data that indicates which view the user is watching needs to be acquired (as described in detail in a later section). A four split screen layout is suitable for detecting eye movement. Note that cases 2, 3, and 4 are basically technical issues and will be overcome with future technical improvement.

[1]ZeroMQ zeromq.org

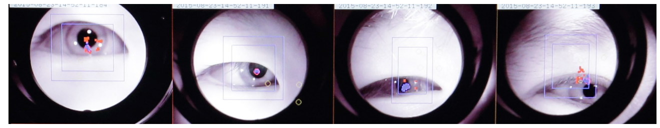


Figure 3. Captured eye images of four participants from Parallel Eyes goggles.

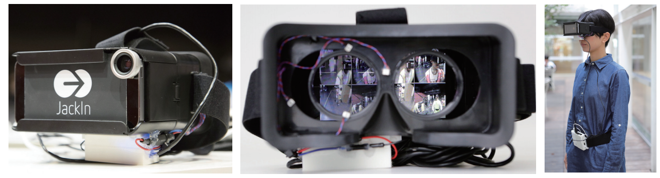


Figure 4. A wearable system for Parallel Eyes: a video see-through HMD, a small board computer, and a small battery module attached to the body.

Eye tracking system

We hypothesize that each user would demonstrate unconscious watching behaviors that could not be captured during post-interview. For realtime observation, the reflection upon behaviors, and post analysis, we embedded a small IR camera and IR-LED into the head mounted smartphone goggles to capture the single eye image (Fig. 3). Then we implemented an eye tracking system with standard computer vision pupil detection. The small IR camera and IR-LED are connected with a Raspberry Pi (a small, single-board computer), and the captured eye image is streamed toward the master computer via a wireless network. The master machine processes the detection and tracking of eye movement in four received eye images in parallel.

As a well known technical issue, the slide of the position between the user's eye and the eye camera is often caused by the movement of the user's head, and it requires calibration when these gaps occur. The focus of our study is to understand when and how the user watches which perspective view, rather than to overcome these technical issues. Therefore, we deal with this technical issue with manual calibration by the operator who is observing the trace of the eye position.

As another technical design rationale, our study only requires the eye tracking to be precise enough to determine which view area the user is watching. This is to avoid a dedicated calibration process for users each time they wear the goggles. In our study, we did not show the trace of eye movement in the Parallel Eyes goggles because the eye tracking is for reflection and analysis of viewing behavior.

Design consideration for technology probe

Even though the system has been built upon iterative development efforts, this Parallel Eyes system has been developed as a technology probe to explore how humans behave in designed situations rather than for the purpose of evaluating the system itself. Thus, there are several criteria for the technology probe in our study method.

1) Feasibility: This wearable system will be worn by a large variety of participants including children so that the system should be fixed in an instant or replaced when some problem occurs. The wearable components are composed of two device units, the small board computer and the small battery module attached to the body with a flexible belt and HMD as we mentioned. These modules can be altered easily and rebooted with a simple step (Fig. 4).

2) Reproducibility: This system is built upon an open source software library and commercially available technical components so that other researchers and educators are allowed to replicate this system as well as the workshop procedure.

3) Engagement: To ensure the success of our study, ensuring an appropriate video streaming quality is a significant factor for maintaining the engagement of participants in the workshop activity. Any delays during video streaming diminish the experience. Therefore, the video resolution and frame rate of the video streaming should be adjusted to provide a better experience for each activity with limited wireless network bandwidth.

Furthermore, the paralleled shared video streaming is recorded with an overlaid visualization of the eye tracking, which then can be played back immediately for reflection and activities in the workshop. (This is also quite an important part of system design for maintaining the engagement of participants in the workshop.)

STUDY METHODOLOGY

We conducted a study throughout the workshop to explore how humans acquire and develop their behaviors with shared parallel first person view situations created by the Parallel Eyes technology probe. We then analyzed the observation of our workshops, user behaviors, and recorded data to yield findings and design implications.

For this method, we designed workshop activities which are conducted with participant engagement rather than through a controlled experiment, and at the same time, beyond being just fun, should be arranged to foster insightful discussion among researchers and participants. In this section, we provide an in-depth description of the design of the workshop in our study.

Workshop design

We designed two main workshop activities in our study with the Parallel Eyes system that produce a shared parallel view experience. The first one is "faint memory drawing," where participants draw something on the basis of only their memory. The second one is a "game of tag in a specially designed field," where participants engage in a simple game with complex parallel view conditions.

There are different design considerations based on unique aspects of Parallel Eyes in each activity. First, in the static workspace activity, the first person video reflects one's output of intellectual work or creative tasks. In the Parallel Eyes situation, those shared visuals from others are provided to users. Therefore, we hypothesize that the intervention by the Parallel Eyes will provide the opportunity to observe how each person influences each other in terms of the drawing process as well as the output of the work. Thus we designed the "faint memory drawing" activity to be performed with a large variety of participants regardless of age or sex, without special knowledge in advance.

Another significant aspect of the Parallel Eyes system is that users are allowed to acquire other perspectives to possibly understand the environment, objects, and spatial coordination of others. Moreover, when users are in proximity to each other, users can observe their own bodies from others' perspectives. With Parallel Eyes, determining how humans acquire and develop their behaviors in understanding their own physical embodiments would provide significant implications for physical activities that include mutual sharing experiences. Thus we designed the "game of tag" activity in which participants play a simple chasing game in a specially designed field (Fig. 7).

The procedure of the workshop is roughly described in three stages: 1) an initial explanation, 2) two main activities, and 3) reflection and discussion after each workshop activity.

Participants and recruitment

We recruited eight participants in each workshop via an announcement at a public educational media center. We conducted five workshops throughout the study, thus the total number of participants were 40 (26 males and 14 females). Participant ages ranged from 8-52. In addition to participants, our workshop was opened for observation by an audience.

Our participants were mostly novices or beginners in regard to technology like virtual reality, HMDs, and wearable computers. All participants don't have experience in regard to HCI research and technology development.

In the workshop activities, eight participants were divided into two teams of four members each. Then each member of each team performed activities in turns, and other members observed the activities. In the workshop space, four paralleled view video streams were displayed on a large projected screen with an overlaid indication of the eye movements of each participant (Fig. 6 and 8). Note that they could not see the trace of the eye movement in the Parallel Eyes goggles.

Pre-activity

In the beginning of the workshop, we first provided an initial explanation of the HMD and then demonstrated the simple experience of first person view sharing (i.e., participants watched the transmitted first person video of the instructor through the HMD). The experience that they were watching was from the perspective of another person, and they eventually realized that. For the next orientation, four participants wore the HMD of the Parallel Eyes system (the Parallel Goggles) and were asked to find their own perspective from the four displayed parallel video streams. They were also asked to shake hands with each other while using the Parallel Eyes (Fig. 5). Through this process, participants understood the

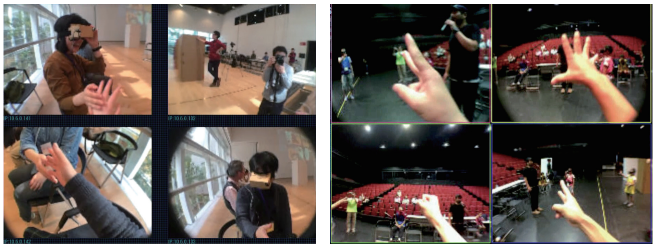


Figure 5. Introductory demonstration of Parallel Eyes system. (left) Participants tried to shake hands with each other while experiencing a shared parallel view. (right) Participants played rock-paper-scissors through a shared paralleled view.

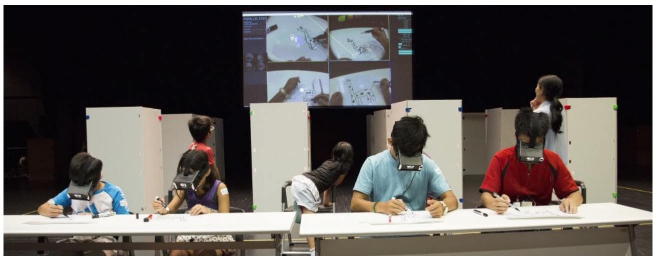


Figure 6. Workshop activity: Drawing with faint memory. Drawing described illustrations without any hints expect others' perspectives.

basic idea behind first person view sharing and understood what was happening with the Parallel Eyes system.

Drawing with faint memory

For the first activity of the workshop, we performed "drawing with faint memory." Participants were asked to draw something (e.g., "draw a giraffe") without any hints expect others' perspectives. It could be something that almost all participants have seen or know about, such as famous historical persons portrayed on paper money, well-known landmarks, (the Statue of Liberty), popular animals (giraffes), and traffic signs. They sat at a desk within a certain distance so that they could hear the other participants' voices, but could not see other drawings from "their own perspective." Instead of that, the Parallel Eyes goggles provided four perspectives (one for one's own, three for others) in parallel (Fig. 6).

One drawing session took approx. two minutes. After participants performed two sessions, they could see the answer image (example image) of drawings to compare with their own drawings. Participants also compared each other's drawings in terms of specific details and characteristics. We and the participants then reflected on the process of their drawings with a recorded paralleled view video overlaid with an indication of the eye movement of each participant. This procedure fostered a discussion about how participants actually drew the pictures and how they were influenced by each other both consciously and unconsciously.

Game of tag with Parallel Eyes

We then moved on to the second activity of the workshop: A "game of tag with Parallel Eyes." Participants were asked to play a game of tag with the Parallel Eyes on a specially designed stage. On the game stage, several cardboard columns were placed at regular intervals (Fig. 8). We describe the person to chase another person (i.e., "it") as a chaser, and we

Figure 7. Workshop activity: Game of tag. Participants were asked to play a game of tag with Parallel Eyes on a specially designed stage.

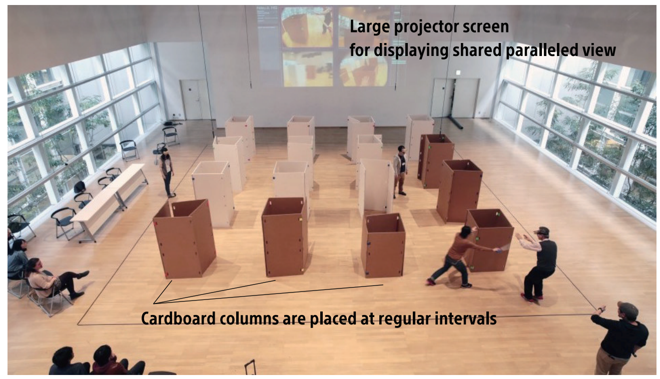


Figure 8. Stage design of chasing game. Cardboard columns are placed at regular intervals.

describe the others as chased. All participants in the game could see four perspectives (one for one's own, three for others) through the Parallel Eyes goggles. When the chased were touched by the chaser, the roles were exchanged. Thus there was only one chaser at any one time. In the Parallel Eyes goggles, the view from the chaser's perspective was highlighted with a color (Fig 11).

With the Parallel Eyes goggles, participants would have a different feel for distance that could present a potential risk during the activity. Therefore, we designed several rules to prevent accidents. 1) The chaser must use a soft sponge sword to touch (i.e., tag) the chased. This is to avoid excessively close physical contact between the chaser and the chased. 2) Participants were not allowed to run, only to walk (similar to the rules of speed walking). 3) When a participant felt any symptoms related to cyber-sickness [35] or when the operator felt sickness was evident, the participant had to withdraw from the activity.

One game session was approx. four minutes. After two teams performed one game session, we researchers and all of the participants held a short conversation to share the strategies the participants developed during the game session. This reflection is quite important for yielding how participants develop their behaviors in the initial stage. Then, sharing those findings among the participants would produce further behavioral changes in the next session. When the participants finished the second game session, we again reflected on interesting moments with the recorded parallel view with eye movement traces. The discussion focused on how participants themselves develop their behaviors throughout the activities to highlight unconscious reactions which they did not notice as well as observing strategies they consciously developed.

RESULTS

In this section, we present the results from observation and analysis of the workshop in terms of overall workshop design, the drawing activity, and the game of tag activity. In general, the initial induction worked well to give the basic idea of first person view sharing through the video see-through HMD (Parallel Eyes goggles). We found that some participants made statements such as: *"What I see now is... myself. OK, I can see myself from the perspective of others behind me."* When seen by others, a female participant said: *"Hmm, somebody is watching me. This is funny."* Those were typical initial reactions when they tried using the system, and accordingly, those reactions led to an understanding of the situation. When we had to decide who would be the first chaser in the activity, one of the participants discovered a smarter way of playing rock-paper-scissors: *"We can play rock-paper-scissors without coming into close contact; just show your hand from your own perspective. We can see all hands."* In terms of engagement, all participants enjoyed the activities throughout the workshop. Some of the young participants commented after the workshop: *"This is my first time to use this (HMD), I (we) want to play the games again!"* Our workshop design invited participants to consider applying our technology probe to their own daily lives. Some of the young participants made comments such as:

P11: *"I want to try using Parallel Eyes goggles with my family. I want to know what my father and mother are looking at. Oh, but... they also can see what I'm watching..."*

P13: *"I usually play soccer, so I think it would be cool if you can use this in soccer."* From these feedback and comments, we found that the design of our workshop worked well for even beginners with the technology, including children.

All sessions of the workshop including activities and discussions were video recorded. For the post analysis, four paralleled participant perspective videos, captured videos of the eye, and eye movement trace data were recorded. Our analysis includes analyzing transcriptions of discussions and analyzing participants' behaviors in each session with video material and eye tracking data. In the following section, we present our results on the basis of the observation and analysis of each workshop activity.

Results of drawing activity

As a general observation, we could find active viewing behaviors in which participants looked at each other's view and created interactions among drawing outcomes. Not all participants had a clear idea of what they were being asked to draw. For instance, P21, who is an elementary school student, and P26, who is a college student from another country, did not recognize the historical person portrayed on the paper money in our study. However, they surprisingly collected the information from the perspectives of the other participants and completed their own drawings.

The Parallel Eyes as a technology probe generated a perceptually high cognitive load and an overload of information. Even in this demanding situation, most of the participants were able to develop their viewing behaviors and complete their drawings. Here, we analyze those developed human behaviors to reveal patterns of behaviors in paralleled first person view sharing.

The viewing behaviors developed by participants roughly consist of two factors: **who to watch**, and **how to watch** through the vision of others. We describe our findings in accordance with two factors along with an example sequence which includes emerged behavior patterns in our analysis (Fig. 9).

Who to watch

A significant aspect of this factor is a **viewing target**, which a user mainly watches. The viewing target is determined in the earlier stage of the drawings on the basis of what a user and others are drawing at that time. As we can see in Fig. 9: P2, at an earlier time, periodically watched P1 as a viewing target. P2 looked at P1's view with almost every drawing stroke approx. before the 60:00 sec mark. From our empirical observation, this results in a similar drawing outcome. In fact, P6 commented *"I think those (drawings by P1 and P2) look similar."*

However, a person can also change the viewing target on the basis of his or her own drawing situation and the drawing situation of others. For instance, the moment at Fig. 9 - *4, P1 was the viewing target of P2 before this moment, but P1 erased his or her own drawing. Then P2 changed the target to P3 and started to draw the "dot" that P2 found in P3's drawing.

We also frequently found that a person usually set a viewing target who is drawing the same or the next element of the drawing. We hypothesize that the person could build a contextual relationship implicitly so that the person could follow the other's perspective in parallel.

How to watch

single focus: When participants already found their own main viewing target, they frequently looked at it, as we mentioned. While viewing the target, their focal point was already determined, thus participants usually only looked at one single view. The duration of viewing others' views depended on the purpose for watching (e.g., confirming an element or recognizing the structure). We can see two shorter and longer viewing behaviors in Fig. 9 for each of the participants.

passing seek: Another viewing behavior is to look over each drawing in one successive action (Fig. 9-*2, 3 and 5.) This would be for determining the next drawing elements, observing the progress of others, or comparing one's own drawing with others.

Interwoven reference

Throughout the observation, we also could find that participants referred to each other and complemented the elements of each other's drawings. At the moment in Fig 9-*1, first P1 drew a "head," which P3 then observed and copied. P3 also drew a "horn," which is the first thing to be drawn among participants. Then P1 looked at P3 and drew a "horn" at the next

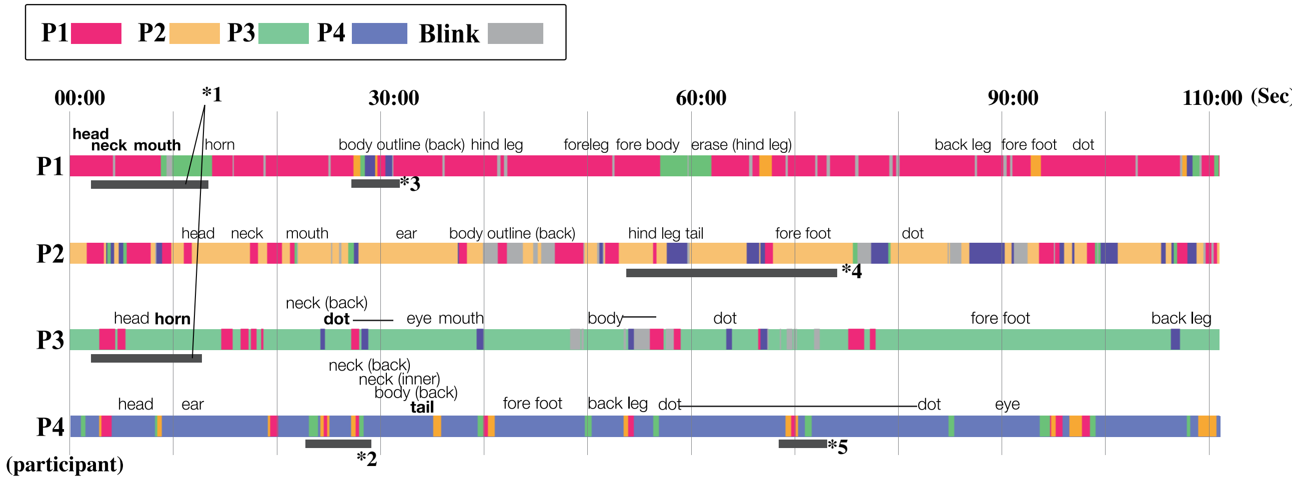


Figure 9. An example sequence in the drawing activity (giraffe) which includes emerged behavior patterns. (The same sequence as Fig. 10) The color bar shows what each participant was watching. Text labels indicate when each drawing element appeared.

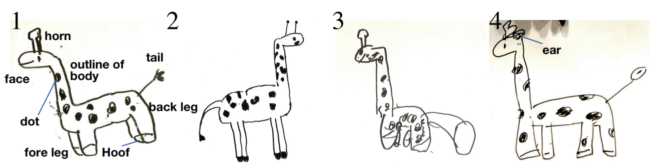


Figure 10. Example outcome from faint memory drawing (giraffe). Almost all drawing elements were drawn by each participant.

moment. As a result, both P1 and P3 could complete drawing a giraffe. This sort of interwoven reference among participants eventually produced a propagation of outcomes for each person during the paralleled view drawing. Fig 9 shows how drawing elements appeared through this propagation in each participant's outcomes.

Results of game of tag activity

Our aim in this activity is to study how humans develop behaviors in this complex paralleled view condition. These results will lead to significant implications for physical activities with mutual sharing experiences. Most participants remarkably developed several strategies as the game proceeded and played the game with a high level of engagement.

Participants commented about the general impression of the game after the activities. From these comments, we confirmed that the level of the game was adequately designed for young to adult participants and provided high engagement, resulting in a slightly stressful thought process.

P4 : *"I think I enjoyed this game. I feel mentally tired. I think that this was a psychic game rather than a physical game."*

P16 : *"If I practice more, I'm sure I can do much better. I feel that I'm getting some sense of how the game works..."*

P9 : *"That was quite an unusual situation beyond my daily perspective, but I could enjoy it. If these kinds of (Parallel Eyes) sports existed, they would be really interesting."*

Next we report the initial reactions of the participants during the game and the several strategies that participants established in the game. We then describe two notable factors that are evident in several strategies.

Initial reaction

In the beginning, some participants found it very confusing to walk around; however, by guiding them to find their own perspective (for instance, having them look down to find their own foot), they found a way to recognize their own body and eventually could even walk a bit faster.

In the earlier stage of the game, several chased participants were astonished when they found themselves in the perspective of the chaser. One of the participants was also astonished because he thought that he was being watched from the chaser's perspective when in fact the chaser was actually watching another player. However, as the game proceeded, they got used to Parallel Eyes and started to develop viewing strategies.

Developed viewing strategies

Through the reflective discussion between and after the game, we collected comments about developed strategies such as:

P4: *"When I found myself in the chaser's perspective, I knew it was time to escape!"*

P12: *"What I did was I used the chaser's perspective when I escaped from the chaser. Because I could see myself from the chaser's point of view."* (Fig. 11 - *5)

P3: *"I could use multiple views... When I got closer to the chaser, I could see the situation from another player's view. So I decided where to escape to."*

P10: *"I was never caught; I frequently checked the chaser's perspective. I could roughly understand the position of the chaser. So I just went as far as possible."* (Fig. 11 - *3)

P8: *"When I escaped from the chaser, I felt some spatial distance between myself and the chaser's perspective."*

P6: *"My particular tip is to hide your vision. I kept facing the floor, but I was watching the others' views."* (Fig. 11 - *2)

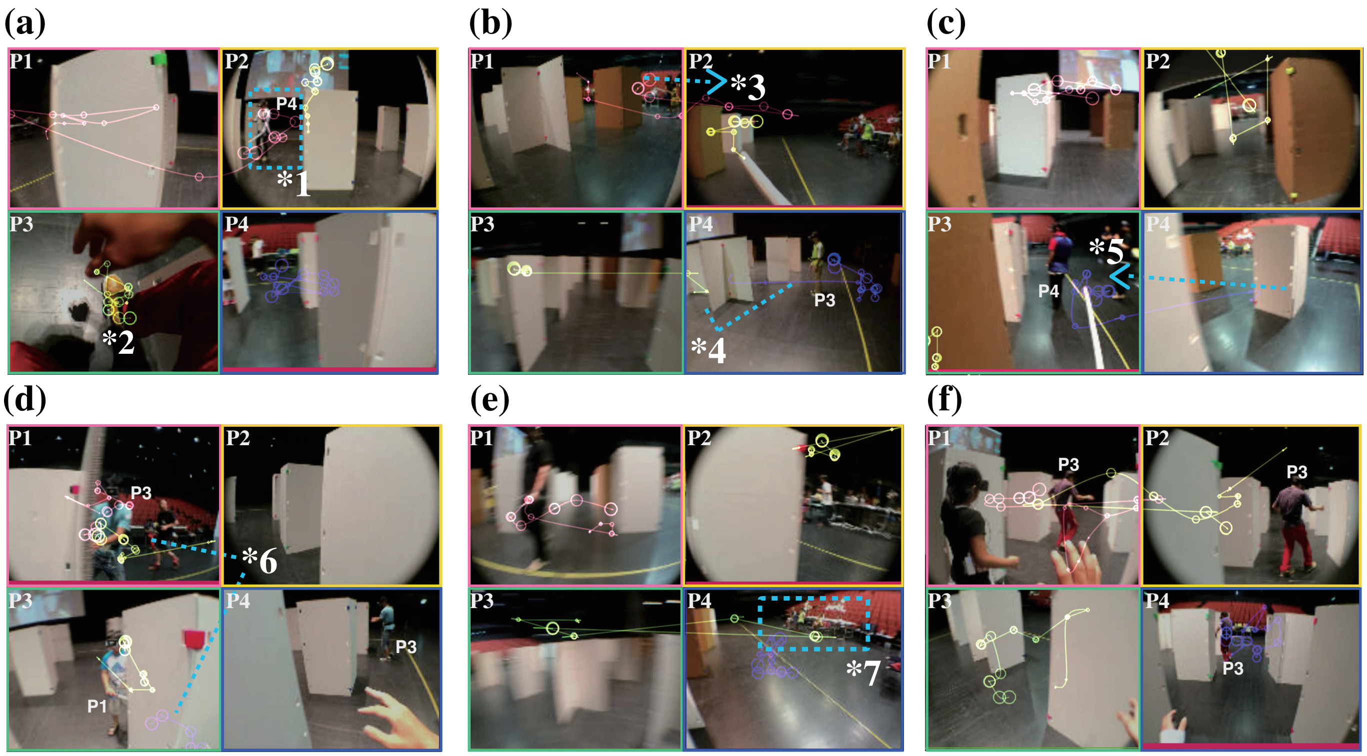


Figure 11. Example of moments during game of tag with Parallel Eyes. Traces of color line indicate eye movement history of each participant. (a) shows moment when P1 found chaser (*1) in P2's view, P3 is in the "stealth" mode (*2). (b) shows that P1 is looking at P2 (chaser's) view (*3), P3 found his own (*4) in P4's view (c) shows that P4 is escaping using the perspective of the chaser P3 (*5). (d) shows that P2 and P4(*6) are engaged in watching the event between P1 and P3. (e) shows the moment when P3 unconsciously a found item (*7) with a similar color to their own clothing. (f) shows the moment when P3 could be found in other 3 views.

P7: *"I also hid myself behind someone and kept facing that person's back. And I was watching that person's view instead of mine."*

We therefore observed these moments again with the recorded video and the analysis data from the eye tracking. We carefully observed what they did as they consciously stated and what was happening unconsciously. From this observation process, especially in regard to the viewing strategies explained by P12, P3, and P8 where the participants recognized both their own and others' perspectives at the same time, we found that viewing strategies were pursued in accordance with two successive factors. The first one is a visual "attention cue," which made the person find something that is related to their own body. The second is described as "context building," where the person understands the relationships between the attention cue and their own body, then produces information to determine the next action.

attention cue: It could be argued that human perceptual reaction had been performed to achieve this viewing behavior. Even the participant did not watch or pay attention to another view area; they could immediately focus on some visual items which could be related to their own body. For instance:

1) finding one's own body image even if it was quite a small region. (Fig. 11 - (b) *4)

2) finding a moving object with a similar color to the chaser. (Fig. 11 - (a) *1)

3) finding items with a similar color to one's own body color. (Fig. 11 - (d) *7)

We can see the sudden movement of the eye tracking trace in those situations as shown in Fig. 11. Note that they did not comment on these behaviors in the workshop reflection. While this should be examined in a more controlled situation using psychological theory, such as the Common coding theory[29], we hypothesize that peripheral vision performs a significant role in aiding human unconscious perceptual ability.

context building: After viewing the reaction toward the attention cue, participants performed a higher level of processing by:

1) confirm whether the visual attention cue is oneself or not.

2) understand where others are looking at me from.

3) understand the relative or global positions of others.

Contrary to the reaction to the attention cue, these behaviors were explicitly mentioned by participants. For instance:

P8: *"As a chaser, I sometimes found somebody watching me, and I started to chase him."* Therefore, this can be argued as a conscious behavior for participants.

Humans would have an ability to recognize their own physical embodiments and other spatial relationships in paralleled perspective sharing. We also found two key factors, an attention cue for achieving a conscious reaction and context building for understanding a high level spatial context with others' perspectives.

DESIGN IMPLICATION AND DISCUSSION

While our findings are yielded from our specific technology probe, the "Parallel Eyes system," we suggest that design implications from them could generally be applied to various view sharing applications with head mounted cameras and HMDs, such as those used in remote assistance, field research, emergency operations, creative processes, sports activities and professional service. In this section, we extend discussion by connecting several research domains to derive design implications, and we raise further research questions and hypotheses that should be considered in future exploration.

Adaptation for extended perspective

Parallel Eyes can be considered relevant to rfesearch and explorative art involving altered or transformed visual perception. Various psychological studies and experimental art projects have been conducted. In psychological research, Stratton et al. [36] and Kohler et al. [20] studied the visual perceptual adaptation of transformed perspectives using reversing glasses (i.e., "upside-down" goggles). The aim of the research was to investigate how perception can be adapted for transformed visual stimuli. They reported that humans could successfully adapt to upside-down glasses after nine days of exposure. While we did not perform our experiment with a long exposure period, we hypothesize that humans would be able to extend visual perception with mutual perspective sharing through the process of visual adaptation in conjunction with transformed perspective.

Some experimental art projects explored transformed visual perception by exchanging perspectives between participants [6, 27] or by transforming the position of cameras to acquire a first person view [33]. These projects were designed to foster discussion about human perspective and embodiment by disconnecting their perspective and embodiment. Therefore, people who tried these systems had difficulty walking and even moving their heads. On the other hand, participants using Parallel Eyes recognized their own embodiments after a few minutes and could even walk a bit faster and run. Moreover, they developed viewing strategies to understand spatial context in regard to others' perspectives, as we observed.

From these observations and analysis, we suggest that design implication for the rapid adaptation of extended perception is guaranteeing own perspective in the visual display with own embodiment. This helps the user to retain a required minimum perception and to develop viewing behaviors on top of their own perspective with other capabilities of visual perception, such as detecting motion with peripheral vision [23] and finding objects related to their own body.

Losing and restoring one's own body

In the game of tag, we found interesting body behaviors during the game. Some participants forgot to move their own bodies and kept standing during certain moments because they were engaged with watching the others' perspectives. Several participants even crashed their own bodies into the cardboard wall of the game stage when they were focusing on the chaser's perspective. This would be explained as a broken up action-perception feedback loop of embodied cognition. When participants focused on another person's view, their own body motion and motion of view focusing conflicted, leading them to lose their own embodiment. In some cases, a certain duration for re-focusing on one's own perspective in the game was required.

Because we noticed these phenomena throughout preliminary exploration, we also experimentally shuffled the layout of four views suddenly before the game of tag. At that very moment, participants lost their embodiment. Or in the worst case, when participants were doing similar actions, shuffling without notification made the situation more confusing in that they continued to misunderstand that they were focusing on their own perspective, even though it was already swapped with the other's perspectives. This type of confusion was also observed in the game.

When participants found it shuffled, they tried to restore their own bodies with roughly two types of strategies, moving their own heads or waving their own hands in front of their own faces. However, moving their own heads sometimes did not work well when more than one person tried a similar action. Showing one's own hand with particular gestures worked well. Especially after a few trials of shuffling, some participants used a strategy of defining their own unique gestures similar to those used in "rock," "paper," "scissors" and maintained the gestures in front of their own cameras.

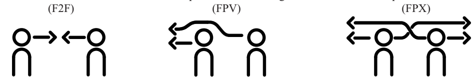
Such phenomena of losing one's own embodiment would be a potential risk for paralleled view sharing with wearable devices. Therefore, we suggest the following design implications: 1) the layout of one's own perspective should be in a stable position, 2) if a dynamic layout of perspectives is required, graphics to indicate one's own and others' perspectives should be designed, and 3) the provision of a UI for hand gestures for retaining one's own embodiment would significantly improve experience.

Distributed cognitive process

What the drawing activity of the study revealed could be a form of distributed cognition, where participants use each other to jog their memory or imagination. Distributed cognition has a psychological framework in which knowledge exists not only within the individual, but also among multiple persons within a social and physical environment [9].

As we found with the interwoven reference in the drawing process, it produced a propagation of each memory, and in some cases, it eventually formed an instant collective outcome from four persons. From this point of view, mutual shared perspectives by means of Parallel Eyes formed distributed cognition among participants without any verbal communication; this is a notable difference in contrast to conventional distributed cognition with verbal communication.

In addition, it should be noted that our configuration fosters an interwoven reference to form distributed cognition. This



	Face to face (F2F)	First person video sharing (FPV)	Mutual first person video exchange (FPX)
Configuration	symmetric	asymmetric	symmetric
Metaphor	sitting / standing together	looking over someone's shoulder	beyond physical metaphor
Use cases	tele-conferencing, collaboration for creation	remote collaboration, assistant, and sharing experience	distributed collaboration, operation, and parallel experience sharing

Figure 12. Comparison of video mediated communication.

is because Parallel Eyes enables socially blind referring, i.e., all players do not know whether another person is looking at their view, and it removes social barriers that cause hesitation toward one another. However, a mutual shared perspective would also produce an unexpected social relationship due to the nature of socially blind referring. Therefore, researchers and designers need to carefully consider how this shared perspective configuration should be used and designed for specific purposes.

Comparison of video mediated communication

For illustrating how this kind of mutual perspective sharing would be beneficial in actual use cases, we took a step back to compare video mediated communication. Although video mediated communication has been studied from various research perspectives, here we roughly describe three configurations: A) face to face video conferencing and the extension of it (F2F), B) remote-local first person video sharing (FPV), and C) mutual first person video exchange (FPX).

As a well known configuration, A) face to face video conferencing (F2F) aims to achieve a sense of physical co-presence, workspace sharing, gesture sharing, and visual conduct for use cases including teleconferencing, spatial collaboration, discussion, and creation accompanied by remotely distributed persons [22, 34, 10]. This configuration basically forms a model of a physical setting, such as one in which people are sitting face to face or standing around a table, and each person can be treated as an individual actor as they are in reality. This well known metaphor helps to boost user expectations, but at the same time, it sometimes limits use cases to an imitation of one medium of communication [10].

As for proliferation of wearable technology, B) first person video sharing (FPV) provides an important configuration where the first person video of the local person is shared with the remote person. FPV also enables various use cases including remote collaboration and assistance from remote persons [21, 2, 18, 28, 11, 3] and experience sharing with the local person [16, 32]. In this configuration, we can apply the metaphor of looking over someone's shoulder, and this forms an explicit communication model between the "watcher" and the "watched." Therefore, this might prevent the fluid exchange of roles between the "watcher" and the "watched" both locally and remotely, subsequently producing a social asymmetric relationship.

Parallel Eyes forms C) a mutual first person video exchange (FPX) where humans can see their own and others' shared perspectives at the same time, in a parallel manner or in a selective manner. We note that the way of perceiving FPX is not limited to the current form of Parallel Eyes. This configuration does not define an explicit role for FPX because of the symmetric nature in this communication model. We suggest that FPX plays an important role in use cases involving professional service providers, emergency operations, security operations in which distributed cognition is required to understand massive information in paralleled situations, and in the coordination of one's own and others' activity. The FPX configuration basically forms a communication model that is beyond the metaphor of physical existence. A remarkable aspect of this communication model is that we do not necessarily require explicit verbal communication to understand each other. In other words, a mutual shared perspective would extend one's own perception as well as establish communication with multiple persons. We believe that such an inter-connected perspective will extend human perception and encourage interesting future work.

CONCLUSION

Mutual first person view sharing will be a promising form of communication as wearable cameras and wearable displays emerge as common technologies. However, there are still uncertainties about the human capacity for such kinds of mutual paralleled first person video sharing. Therefore, our focus was to explore how humans can understand and develop viewing behaviors with mutual paralleled first person video sharing. We described the paralleled first person view sharing system as a technology probe that we developed. It consisted of multiple video see-through HMDs with embedded eye tracking. With this system, four persons were allowed to see four shared first person videos of each other. Then we described the design of our workshop in detail for this study. It was designed to investigate how future prospective users develop their own behaviors and abilities, and it included two activities, drawing pictures and playing a simple chasing game. Our results of the analysis showed that people developed behaviors such as the dynamic mutual complementing of memory and decision during the drawing activities. We also found that people in the game of tag remarkably developed several strategies to understand their own physical embodiment and spatial relationship with each other, even in highly complex situations. In addition to our findings in the development of human behavior, we also found perceptual phenomena, behaviors of the body in mutual visual sharing, and a derived communication model of mutual visual sharing. Those findings led to design implications and further research questions which will contribute to building design experience and fostering further research in mutual parallel experiences.

ACKNOWLEDGEMENTS

We thank all of our study participants. This work was supported by the Agency for Cultural Affairs Government of Japan.

REFERENCES

1. Jed R. Brubaker, Gina Venolia, and John C. Tang. 2012. Focusing on Shared Experiences: Moving Beyond the Camera in Video Communication. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*.

ACM, New York, NY, USA, 96–105. DOI: http://dx.doi.org/10.1145/2317956.2317973

2. Susan R. Fussell, Leslie D. Setlock, and Robert E. Kraut. 2003. Effects of Head-mounted and Scene-oriented Video Systems on Remote Collaboration on Physical Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 513–520. DOI: http://dx.doi.org/10.1145/642611.642701
3. Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. In Touch with the Remote World: Remote Collaboration with Augmented Reality Drawings and Virtual Navigation. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology (VRST '14)*. ACM, New York, NY, USA, 197–205. DOI: http://dx.doi.org/10.1145/2671015.2671016
4. William W. Gaver, Abigail Sellen, Christian Heath, and Paul Luff. 1993. One is Not Enough: Multiple Views in a Media Space. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93)*. ACM, New York, NY, USA, 335–341. DOI: http://dx.doi.org/10.1145/169059.169268
5. K. Goldberg, Dezhen Song, and A. Levandowski. 2003. Collaborative teleoperation using networked spatial dynamic voting. *Proc. IEEE* 91, 3 (Mar 2003), 430–439. DOI: http://dx.doi.org/10.1109/JPROC.2003.809209
6. Katsuhiko Hachiya. 1993. Inter Dis-communication Machine. *Prix Ars Electronica* 96 (1993), 138–139.
7. Christian Heath and Paul Luff. 1991. Disembodied Conduct: Communication Through Video in a Multi-media Office Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. ACM, New York, NY, USA, 99–103. DOI: http://dx.doi.org/10.1145/108844.108859
8. Keita Higuchi and Jun Rekimoto. 2013. Flying Head: A Head Motion Synchronization Mechanism for Unmanned Aerial Vehicle Control. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 2029–2038. DOI: http://dx.doi.org/10.1145/2468356.2468721
9. James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed Cognition: Toward a New Foundation for Human-computer Interaction Research. *ACM Trans. Comput.-Hum. Interact.* 7, 2 (June 2000), 174–196. DOI: http://dx.doi.org/10.1145/353485.353487
10. Jim Hollan and Scott Stornetta. 1992. Beyond Being There. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. ACM, New York, NY, USA, 119–125. DOI: http://dx.doi.org/10.1145/142750.142769
11. Tobias Höllerer, Steven Feiner, Tachio Terauchi, Gus Rashid, and Drexel Hallaway. 1999. Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. *Computers & Graphics* 23, 6 (1999), 779–785.
12. Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology Probes: Inspiring Design for and with Families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 17–24. DOI: http://dx.doi.org/10.1145/642611.642616
13. Hiroyuki Iizuka, Daisuke Kondo, Hiroki Kawasaki, Hideyuki Ando, and Taro Maeda. 2011. Coordinated Behavior Between Visually Coupled Dyads. In *Proceedings of the 2Nd Augmented Human International Conference (AH '11)*. ACM, New York, NY, USA, Article 23, 4 pages. DOI: http://dx.doi.org/10.1145/1959826.1959849
14. Kori Inkpen, Brett Taylor, Sasa Junuzovic, John Tang, and Gina Venolia. 2013. Experiences2Go: Sharing Kids' Activities Outside the Home with Remote Family Members. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 1329–1340. DOI: http://dx.doi.org/10.1145/2441776.2441926
15. Ellen A. Isaacs and John C. Tang. 1993. What Video Can and Can'T Do for Collaboration: A Case Study. In *Proceedings of the First ACM International Conference on Multimedia (MULTIMEDIA '93)*. ACM, New York, NY, USA, 199–206. DOI: http://dx.doi.org/10.1145/166266.166289
16. Shunichi Kasahara, Shohei Nagai, and Jun Rekimoto. 2015. First Person Omnidirectional Video: System Design and Implications for Immersive Experience. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 33–42. DOI: http://dx.doi.org/10.1145/2745197.2745202
17. Shunichi Kasahara and Jun Rekimoto. 2014. JackIn: Integrating First-person View with Out-of-body Vision Generation for Human-human Augmentation. In *Proceedings of the 5th Augmented Human International Conference (AH '14)*. ACM, New York, NY, USA, Article 46, 8 pages. DOI: http://dx.doi.org/10.1145/2582051.2582097
18. Shunichi Kasahara and Jun Rekimoto. 2015. JackIn Head: Immersive Visual Telepresence System with Omnidirectional Wearable Camera for Remote Collaboration. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology (VRST '15)*. ACM, New York, NY, USA, 217–225. DOI: http://dx.doi.org/10.1145/2821592.2821608
19. H. Kawasaki, H. Iizuka, S. Okamoto, H. Ando, and T. Maeda. 2010. Collaboration and skill transmission by first-person perspective view sharing system. In *RO-MAN, 2010 IEEE*. 125–131. DOI: http://dx.doi.org/10.1109/ROMAN.2010.5598668

20. Ivo Kohler. 1963. The formation and transformation of the perceptual world. *Psychological issues* (1963).

21. Hideaki Kuzuoka. 1992. Spatial Workspace Collaboration: A SharedView Video Support System for Remote Collaboration Capability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. ACM, New York, NY, USA, 533–540. DOI: http://dx.doi.org/10.1145/142750.142980

22. Paul K. Luff, Naomi Yamashita, Hideaki Kuzuoka, and Christian Heath. 2015. Flexible Ecologies And Incongruent Locations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 877–886. DOI: http://dx.doi.org/10.1145/2702123.2702286

23. Suzanne P. Mckee and Ken Nakayama. 1984. The detection of motion in the peripheral visual field. *Vision Research* 24, 1 (1984), 25 – 32. DOI: http://dx.doi.org/10.1016/0042-6989(84)90140-8

24. monobanda. 2014. 3RD, Transforming space through perspective. (An interactive installation). (2014).

25. Florian 'Floyd' Mueller, Shannon O'Brien, and Alex Thorogood. 2007. Jogging over a Distance: Supporting a "Jogging Together" Experience Although Being Apart. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems (CHI EA '07)*. ACM, New York, NY, USA, 2579–2584. DOI: http://dx.doi.org/10.1145/1240866.1241045

26. Claire O'Malley, Steve Langton, Anne Anderson, Gwyneth Doherty-Sneddon, and Vicki Bruce. 1996. Comparison of face-to-face and video-mediated interaction. *Interacting with Computers* 8, 2 (1996), 177–192. DOI: http://dx.doi.org/10.1016/0953-5438(96)01027-2

27. Bertrand Philippe, Daniel González-Franco, Arthur Pointeau, and Christian Cherene. 2014. The Machine to be Another - Embodied Telepresence using human performers. *Prix Ars Electronica* (2014).

28. Ronald Poelman, Oytun Akman, Stephan Lukosch, and Pieter Jonker. 2012. As if Being There: Mediated Reality for Crime Scene Investigation. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 1267–1276. DOI: http://dx.doi.org/10.1145/2145204.2145394

29. Wolfgang Prinz. 1990. *A common coding approach to perception and action*. Springer.

30. Jason Procyk, Carman Neustaedter, Carolyn Pang, Anthony Tang, and Tejinder K. Judge. 2014. Exploring Video Streaming in Public Settings: Shared Geocaching over Distance Using Mobile Video Chat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2163–2172.

31. Irene Rae, Gina Venolia, John C. Tang, and David Molnar. 2015. A Framework for Understanding and Designing Telepresence. In *Proc. CSCW 2015*. ACM Association for Computing Machinery.

32. Stuart Reeves, Christian Greiffenhagen, Martin Flintham, Steve Benford, Matt Adams, Ju Row Farr, and Nicholas Tandavantij. 2015. I'D Hide You: Performing Live Broadcasting in Public. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2573–2582.

33. Piatza Sebastian, Zöllner Christian, and Adenauer Julian. 2013. EYEsect. (2013).

34. Abigail Sellen, Bill Buxton, and John Arnott. 1992. Using Spatial Cues to Improve Videoconferencing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. ACM, New York, NY, USA, 651–652. DOI: http://dx.doi.org/10.1145/142750.143070

35. Kay M Stanney, Robert S Kennedy, and Julie M Drexler. 1997. Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 41. SAGE Publications, 1138–1142.

36. George M. Stratton. 1896. Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review* 3, 6 (1896), 611–617. http://dx.doi.org/10.1037/h0072918