MultiSoma: Motor and Gaze Analysis on Distributed Embodiment With Synchronized Behavior and Perception

Reiji Miura[1]*, Shunichi Kasahara[2,3], Michiteru Kitazaki[4], Adrien Verhulst[3], Masahiko Inami[3] and Maki Sugimoto[1]

[1] Department of Information and Computer Science, Keio University, Tokyo, Japan, [2] Sony Computer Science Laboratories, Inc., Tokyo, Japan, [3] Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan, [4] Department of Computer Science and Engineering, Toyohashi University of Technology, Toyohashi, Japan

Human behavior and perception are optimized for a single body. Yet, the human brain has plasticity, which allows us to extend our body schema. By utilizing technology like robotics or virtual reality (VR), we can modify our body parts or even add a new body to our own while retaining control over these parts. However, the update of body cognition when controlling multiple bodies has not been well examined. In this study, we explore the task performance and body cognition of humans when they have multiple full bodies as an extended embodiment. Our experimental system allows a participant to control up to four bodies at the same time and perceive sensory information from them. The participant experiences synchronizing behavior and vision perception in a virtual environment. We set up three tasks for multiple bodies and evaluated the cognition of these bodies with their gazing information, task performances, and subjective ratings. We found that humans can have the sense of body ownership and agency for each body when controlling multiple bodies simultaneously. Furthermore, it was observed that people manipulate multiple bodies by actively switching their attention in a static environment and passively switching their attention in a dynamic environment. Distributed embodiment has the potential to extend human behavior in cooperative work, parallel work, group behavior, and so on.

Keywords: virtual reality, augmented human, embodiment, multiple bodies, gaze analysis

1. INTRODUCTION

Due to the nature of the human body structure, our behavior and perception are optimized for a single body. Yet, the human brain allows us to extend our embodiment beyond our own body. In a previous study, it was shown that people can feel and manipulate a virtual body as if it were their own by synchronizing their vision and movements (Kilteni et al., 2012a; Braun et al., 2018). This is called the sense of embodiment, which includes the sense of self-location, the sense of body ownership, and the sense of agency. People can also have the sense of embodiment for a body with extended structures (Kilteni et al., 2012b; Kondo et al., 2018a) and partial representations (Kondo et al., 2018b). Research has explored various possibilities of the sense of embodiment with extended body structures, such as third arms (Saraiji et al., 2018),

long arms (Kilteni et al., 2012b; Kondo et al., 2018a), and partial representations (Kondo et al., 2018b). Furthermore, the sense of embodiment can be held in a body shared with another person (Fribourg et al., 2020) or swapped to another body (Petkova and Ehrsson, 2008). In terms of the sense of self-location, there have been studies on bi-location, which means to exist in two places at the same time (Wissmath et al., 2011; Furlanetto et al., 2013; Aymerich-Franch et al., 2016; Guterstam et al., 2020; Nakul et al., 2020). The question arises: Under the condition that the movements of multiple bodies are synchronized as extended bodies, how do people operate them?

This study focuses on the sense of embodiment with directly synchronized vision and movements with multiple bodies. In the field of robotics, interaction with multiple robots by a single operator has been studied (Glas et al., 2011; Kishore et al., 2016), but the sense of parallel embodiment into multiple bodies has not been investigated. On the other hand, techniques of visual sharing with others have been studied (Kawasaki et al., 2010; Kurosaki et al., 2011; Mitchell et al., 2015; Kasahara et al., 2016; Pan et al., 2017; Sypniewski et al., 2018). We applied the method of splitting the visual field to share visions with multiple bodies.

The human body is unique in the physical environment. We envision that, if we can control multiple bodies simultaneously, we can interact with multiple environments in parallel, which may improve efficiency in tasks and show the possibility of cooperative work and group behavior by one person.

In this study, we attempt to provide direct control for distributed embodiment and explore how body cognition changes with multiple bodies and how the user controls them.

We used a Virtual Environment (VE) to make distributed avatars. In our system, the user's movements are synchronized to multiple bodies, and the user can perceive visual sensory information from them simultaneously. We can synchronously control up to four virtual bodies and use split views to display the First Person Perspective (FPP) of the bodies (similar to Kasahara et al., 2016).

In Experiment 1, we asked participants to touch balls while changing the number of avatar bodies. We evaluated the differences in body cognition between single embodiment and multiple embodiment based on the task performances and subjective ratings. Then, in Experiment 2, we asked participants to catch flying balls that were thrown toward the avatar bodies under conditions with four distributed embodiments to see the possibility of controlling multiple bodies simultaneously. Then in Experiment 3, we asked participants to avoid flying balls that were thrown toward the avatar bodies under four-body condition to see the ability to have a body image for each perspective. We used gaze information, task performance, and subjective ratings to evaluate how people manipulate multiple bodies and take avoidance actions in a dynamic environment.

The main contribution of this work is to explore the following items:

- The effects of increasing the number of bodies performing motor synchronization and visual sharing on task performance and the sense of embodiment.

- How people operate multiple bodies when their movements are synchronized with them.
- Analysis of gaze information when manipulating multiple bodies.

2. RELATED WORK

2.1. Bi-Location in Virtual Reality

Wissmath et al. (2011) conducted a user study in a VE displayed on a TV screen to examine whether it is possible to feel localized at two distinct places at the same time. Participants rated their self-localization in their "immediate environment" (i.e., the real room) and in their "mediated environment" (i.e., the VE). The authors compared the two ratings and showed that participants equally distributed their self-localization between both environments (after 30 s, about 50% immediate environment and 50% mediated environment).

Furlanetto et al. (2013) reviewed bi-location works and argued for decomposing bi-location into three components: self-localization; self-identification (i.e., a concept similar to body ownership); and FPP. These components are closely related to the sense of self-location and sense of body ownership from the sense of embodiment (Kilteni et al., 2012a). A related study was conducted by Aymerich-Franch et al. (2016), who developed a system to embody a robot and conducted a user study. Participants rated their self-location and self-identification as they observed themselves (i.e., their real body) from the robot. The authors showed that it is possible to be bi-located when embodied in two bodies.

Recently, Nakul et al. (2020) compared two self-location tasks in a Third Person Perspective (TPP) fully immersive VR setup: a so-called Mental Imagery Task (MIT) based on perception, where participants guessed when a ball moving at constant speed would hit them; and a Locomotion Task (LT) based on motion, where participants walked with their real body toward a fake body. The authors showed that bi-location was only felt in the MIT task.

2.2. Operating Multiple Robots

Glas et al. (2011) developed a user interface for operating with multiple social robots and evaluated the system. This system optimally sorts the operation requests from the robots to a single user. While the user operates one robot, the other robots are controlled in the background. The results of the evaluation experiment show that the system is useful for operating multiple social robots with a single operator. However, the user controls the robots by entering commands into the interface. Embodiment in multiple robots has not been considered.

Kishore et al. (2016) developed a system that allows a user to fully control a humanoid robot using full body tracking and switch a subject of control freely to control multiple robots. The user can feel an audiovisual perception of each robot with a Head Mounted Display (HMD) and two-way audio communication. The study showed the possibility of eliciting a full body ownership illusion over not one but multiple artificial bodies concurrently. We investigate the sense of embodiment with synchronized vision and movements in parallel with up to four multiple bodies.

2.3. Sharing Multiple Fields of View

Several view sharing systems have been developed, such as View Blending; View Swapping (i.e., you see another user's view); View Cycling (i.e., you cycle through all the users' views, including your own); and View Splitting (along the horizontal line, the vertical line, or in a grid).

Kawasaki et al. (2010) developed an HMD blended shared view system for two users. While their system supported the blended view, they conducted a user study with view swapping (i.e., your own view has 100% transparency [= not visible] and your partner's view has 0% transparency). This study showed view sharing was significantly better than no view sharing at transmitting/acquiring motor-skills. From the same team and with the same system, Kurosaki et al. (2011) conducted a user study with view blending (i.e., your own view has 50% transparency and your partner's view has 50% transparency). This also showed view sharing was significantly better at transmitting/acquiring motor-skills than no view sharing.

Mitchell et al. (2015) developed an HMD view swapping system for two participants where only one participant was able to see something (here, his/her partner's view). They tested their system in several workshops (running, throwing, etc.) where the participants had to stand back to back. The participants reported it was enjoyable but difficult to find their orientation.

Kasahara et al. (2016) developed an HMD split view sharing system for four participants. The four views were displayed side by side along a 2 × 2 grid. They tested their system in several workshops (drawing, shaking hands, etc.) and noted (among other things) that: (i) participants developed their own viewing behavior (i.e., looking at the others' views to know what to do); and (ii) participants lost sight of their own body. Similarly, Sypniewski et al. (2018) (from the same team as Mitchell et al., 2015) conducted several workshops with four participants, but they used an HMD with a view cycling system. The participants reported (among other things) that: (i) they developed their own viewing behavior (i.e., waited for their own view before moving); and (ii) they focused more on non-visual perceptions.

Pan et al. (2017) compared three HMD shared view implementations for two participants: blended views, split views along a horizontal line, and split views along a vertical line. They showed that blended and vertical split views significantly provided the best sense of body ownership.

3. CONCEPT

The main purpose of this study is to explore cognition and behavior changes on distributed multiple bodies with simultaneous motion and visual perception sharing. This section describes the concept of the MultiSoma with synchronized behavior and perception.

Our body is a single instance in our physical space, but it is possible to extend it with virtual reality or robotics technology (Waterworth and Waterworth, 2014). To have multiple bodies with a single user can be considered as an implementation of human augmentation. By having multiple bodies, we are able to consider various scenarios.
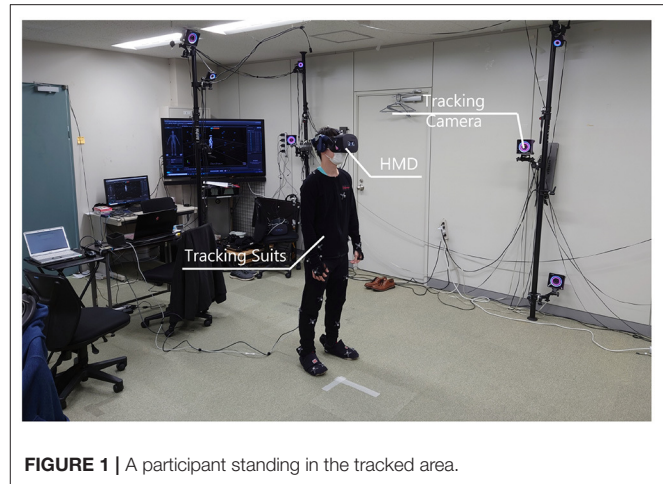

FIGURE 1 | A participant standing in the tracked area.

Working in multiple places: Multiple bodies allow the user to show their skill to other people as a trainer in multiple remote locations. Distributed embodiment allow us to existing and interact with multiple places in parallel.

Self co-working: A mirrored body is able to cooperate with the original one to held both ends of a heavy or long item and move it.

Maze: To find a route in a maze is a possible scenario for the user with the multiple bodies. The user is able to perform speculative search to use the advantage of the distributed embodiment.

We are able to think about various possible use cases. But, in this study, we focus on investigating user cognition and behavior on such multiple bodies with synchronized behavior and perception through task based user experiments.

4. IMPLEMENTATION

Our VR system allows a participant to be distributed in multiples bodies (1–4) in real-time.
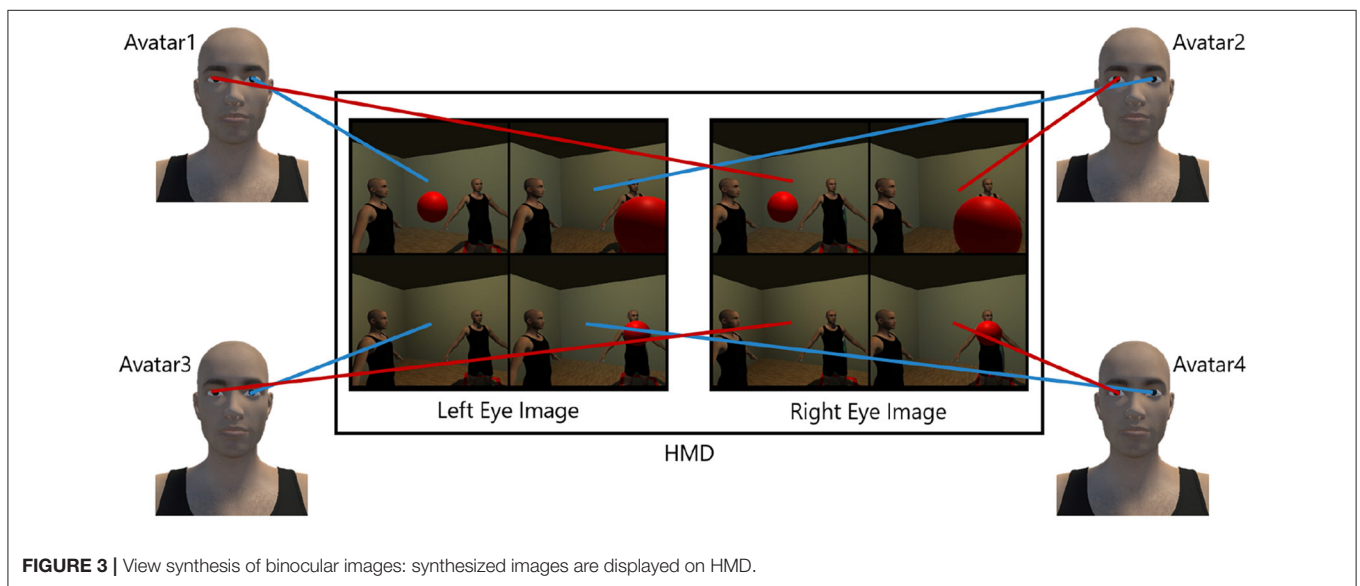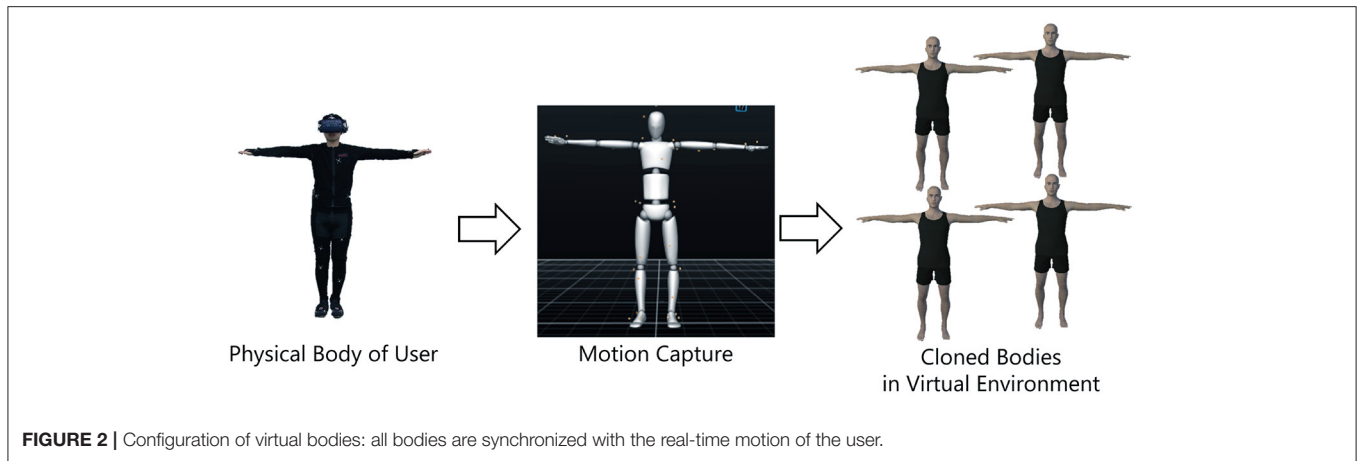
4.1. Material

The participant's body motions were tracked with 24 Optitrack cameras[1] in a tracked area of $\approx$ 4.37 × 2.37 m (**Figure 1**). We implemented the VR application in Unity3D[2]. To display the VE, we used the HTC Vive Pro Eye HMD[3].

It has 1,440 × 1,600 pixels per eye, a 90 Hz refresh rate and 110° of viewing angle. It performs eye tracking at 120 Hz with an accuracy of 0.5–1.1°. The eye tracking is done by a camera attached to the inside of the HMD, which can be used by calibrating to follow a point with the eyes. By Motive, OptiTrack control software, the participant's skeleton data was generated from the marker pattern of the tracking suit and streamed to Unity (**Figure 2**).

[1]https://www.naturalpoint.com/
[2]https://unity.com/
[3]https://www.vive.com/eu/product/vive-pro-eye/overview/

FIGURE 2 | Configuration of virtual bodies: all bodies are synchronized with the real-time motion of the user.

FIGURE 3 | View synthesis of binocular images: synthesized images are displayed on HMD.

4.2. Multiple Bodies

The avatar body used in the application was one of avatars of Daz Studio[4]. Since the bodies were copies of each others, they were indistinguishable from each others.

The bodies moved relatively to the participant's motion. In order to synchronize the motion of the multiple bodies and the participant, we used skeleton data streamed from the motion capture system. The skeleton data is segment information defined hierarchically. We copied the positions and rotations of the corresponding segments to the multiple virtual avatars. The position of each body was moved relative to the actual participant's movement from each initial position.

4.2.1. Views Sharing

There are several methods to share the views of multiples bodies, such as view cycling, view swapping, view splitting, etc. (c.f.,

Section 2). Here we used view splitting in a 1×2 or 2×2 matrix (respectively if there was two, or four avatars).

The user's view was split to the number of bodies and each split view is the first person view from each avatar. To make binocular cues, we combined visual information from avatars for left and right eye independently in consideration of view dependency. This approach allows us to make stereoscopic perception from synchronized bodies (**Figure 3**).
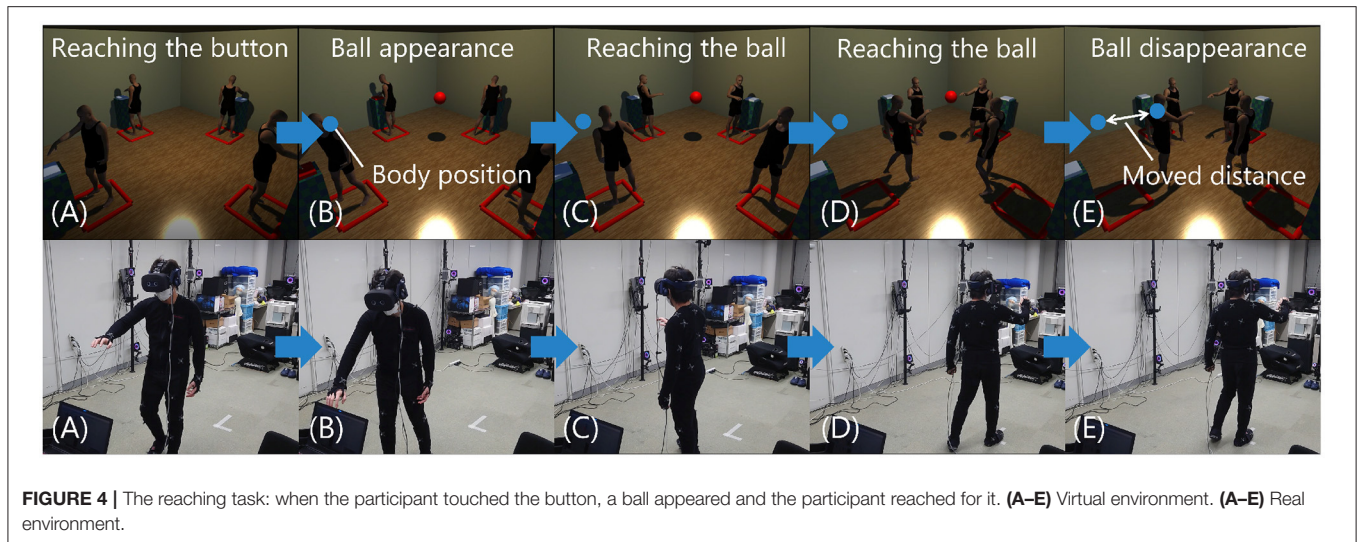
5. EXPERIMENT 1

Experiment 1 focused on the effects of multiple bodies on task performance and the sense of embodiment.

5.1. Method

We asked participants to complete a reaching task for virtual balls. There were three conditions: the one-body condition; the two-body condition; and the four-body condition. In the two-body condition, the bodies were rotated 180° in the horizontal

---

[4]https://www.daz3d.com/

FIGURE 4 | The reaching task: when the participant touched the button, a ball appeared and the participant reached for it. **(A–E)** Virtual environment. **(A–E)** Real environment.

plane (so they were facing each other), and in the four-body condition, the bodies were rotated 90° in the horizontal plane to face each other.

5.1.1. Participants

Twelve participants (all male, age: $M = 23.0, SD = 1.7$, 10 right-handed) joined the experiment. All of them had normal vision and motor ability.

5.1.2. Reaching Task

We set up a reaching task in a square room of $\approx 5.5 \times 5.5$ m (fitting within the tracked area). There were 20 trials. In each trial, participants were asked to press a button in front of the view where all views (four or two) were identical . Once the button was pressed, the virtual target ball appeared for the reaching task (**Figure 4**). Participants then looked back toward the center of the room (**Figure 4B**). In this situation, the four or two views were not identical since the virtual ball was not symmetrically located. Then participants reached for the virtual ball (**Figure 4C**).

We prepared 20 positions that were arranged evenly within 1.0m from the center of the square room, and a ball appeared at one of the positions in random order.

In this task, one trial was defined as the period between the appearance and disappearance of the ball, and the distance the participant's head moved and the task completion time taken during one trial were measured to evaluate task performance.

During the task, the position of the ball was not moved. Thus, we could observe the behavior and cognition with a static environment in this experiment.

5.1.3. Procedure

The participant wore the HMD and the tracking suit. Body tracking was calibrated at the beginning of the session. After receiving instructions for the experimental procedure, the participant performed the task under one of the three conditions. After completing the task, the participant answered a questionnaire on a 7-level Likert scale of -3 (very strongly

disagree) to 3 (very strongly agree)[5]. This was repeated for each condition. The order of the conditions was randomized and counter balanced between the participants. The participants were asked to respond to the following after each condition:

- Q1. I felt as if all the virtual bodies were my bodies.
- Q2. I did not feel as if some virtual bodies were my bodies.
- Q3. I felt as if I controlled all the virtual bodies.
- Q4. I felt as if I was not the one in control of some of the bodies.
- Q5. I was able to perform the task efficiently.
- Q6. It was difficult to operate.
- Q7. What did you think about when you performed the task? (Free Answer)
- Q8. What were your thoughts on the task? (Free Answer).

The following questions were asked only for multiple body conditions (two-body condition and four-body condition):

Q9 I felt as if I was controlling all the virtual bodies at the same time.

Q10 . I felt as if I was controlling a body of the view I was paying attention to.

5.2. Result
5.2.1. Task Performance

To see the effects of synchronization with multiple bodies on task performance, we evaluated the moving distance of the participant and the completion time in each trial. The result of moving distance of the participant is shown in **Figure 5**, left. As an objective measurement, we found that the moving distance in the task decreased as the body count increased. For the experimental setup, if a participant utilized distributed embodiment efficiently, moving distance in the multiple-body conditions was reduced compared to the one-body condition. The result suggests the participants took advantage of the distributed embodiment and manipulated multiple bodies to

---

[5]The original questionnaires were written in Japanese and translated in English for this publication.
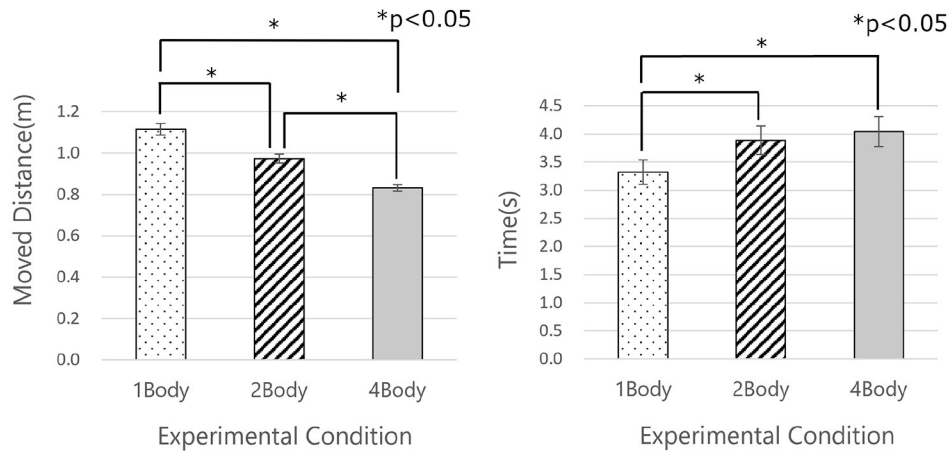
FIGURE 5 | Result of task performance. (Left) Average moving distance in each trial. (Right) Average task time in each trial. The asterisks in the figure indicate between-condition comparisons with P-value for significant differences smaller than 0.05.

complete the task. Consequently, the moving distance in the multiple-body conditions was reduced. The moving distance was shorter in the two-body condition than in the one-body condition ($p < 0.05$) and shorter in the four-body condition than in the one-body condition ($p < 0.05$) and two-body condition ($p < 0.05$). The Tukey-Kramer test was used as a statistical test for the performance analysis of this study.

On the other hand, task completion time in the multiple-body conditions increased more than in the one-body condition. **Figure 5**, right shows the result of task completion time. The task completion time was shorter in the one-body condition than in the two-body condition ($p < 0.05$) and four-body condition ($p < 0.05$). This could be due to the cognitive load of sharing multiple views and choosing an appropriate behavior, which increased the task completion time while the moving distance decreased in accordance with the number of bodies.

5.2.2. Sense of Embodiment

To see the effects of synchronization with multiple bodies on the sense of embodiment, we analyzed subjective responses from the participants.

We performed Steel-Dwass test on Q1–Q4 among the body conditions. **Figure 6**, left two blocks shows the results of Q1 and Q2 regarding the sense of body ownership. The score of Q1 was higher in the one-body condition than in the two-body condition ($p < 0.05$). The score of Q2 was lower in the one-body condition than in the two-body condition ($p < 0.05$).

Figure 6, center two blocks shows the results of Q3 and Q4 regarding the sense of agency. We compared the results of the sense of agency (Q3, Q4) between the body conditions. The Q3 score was higher in the one-body condition than in the two-body condition ($p < 0.05$) and four-body condition ($p < 0.05$). The Q4 score was lower in the one-body condition than in the two-body condition ($p < 0.05$).

Figure 6, right two blocks shows the result of the comparison among the conditions regarding the question about efficiency and difficulty(Q5, Q6). The score of the feeling of difficulty was lower in the one-body condition than in the two-body condition ($p < 0.05$) and four-body condition ($p < 0.05$). We also used a Steel-Dwass test for the statistical test.

There was no significant difference on the sense of body ownership for all bodies (Q1 and Q2) in between the one-body and four-body conditions. Also, there was no significant difference in between the one-body and four-body conditions on a question for the sense of agency (Q4). This result suggests that the participants kept the sense of embodiment close to the one-body condition even in the four-body condition.

From the results of the comparison among conditions, the sense of agency (Q3) decreased as the number of bodies increased, even though it was preserved in between the one-body and four-body conditions in the other questions (Q1, Q2, and Q4). This is similar to the sense of presence in a study of bi-location by Wissmath et al. (2011). In that study, the self-localization of two environments converged to 50%. In the same way, the sense of agency in each body is considered diminished due to the participant becoming the object of an action multiplied by distributed embodiment.

In the two-body condition, the two bodies faced each other by 180 degrees of rotation so the participants could see both the bodies' movements from a single perspective. Also, in the participants' free responses (Q7), there were several comments that the task "could be completed by using only one of the two perspectives in the two-body condition." These facts suggest that the perception of one body as "my body" and the other as "another body synchronized to me" in the two-body condition may have reduced the sense of body ownership. It can be said that the sense of body ownership may be retained in conditions where the two bodies are not visible to each other, even in the two-body condition. At the same time, these facts support the possibility of spatial perspectives by taking perspectives other than one's own, which has been proposed in several studies (Kawasaki et al., 2010; Kurosaki et al., 2011; Becchio et al., 2013; Kasahara et al., 2016).
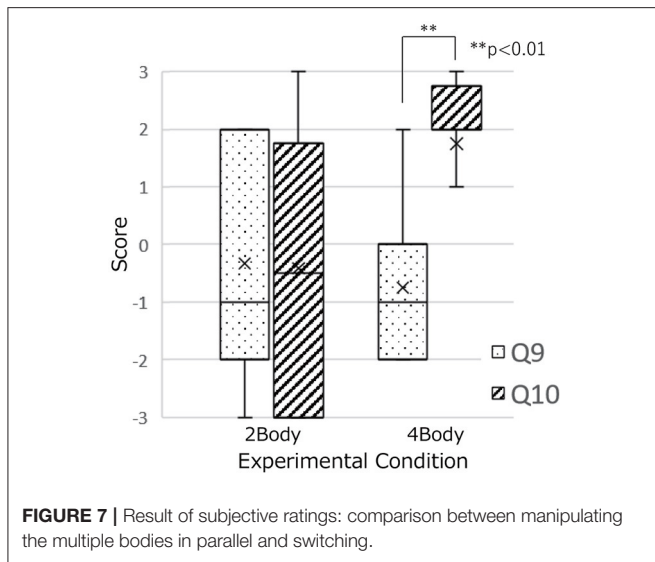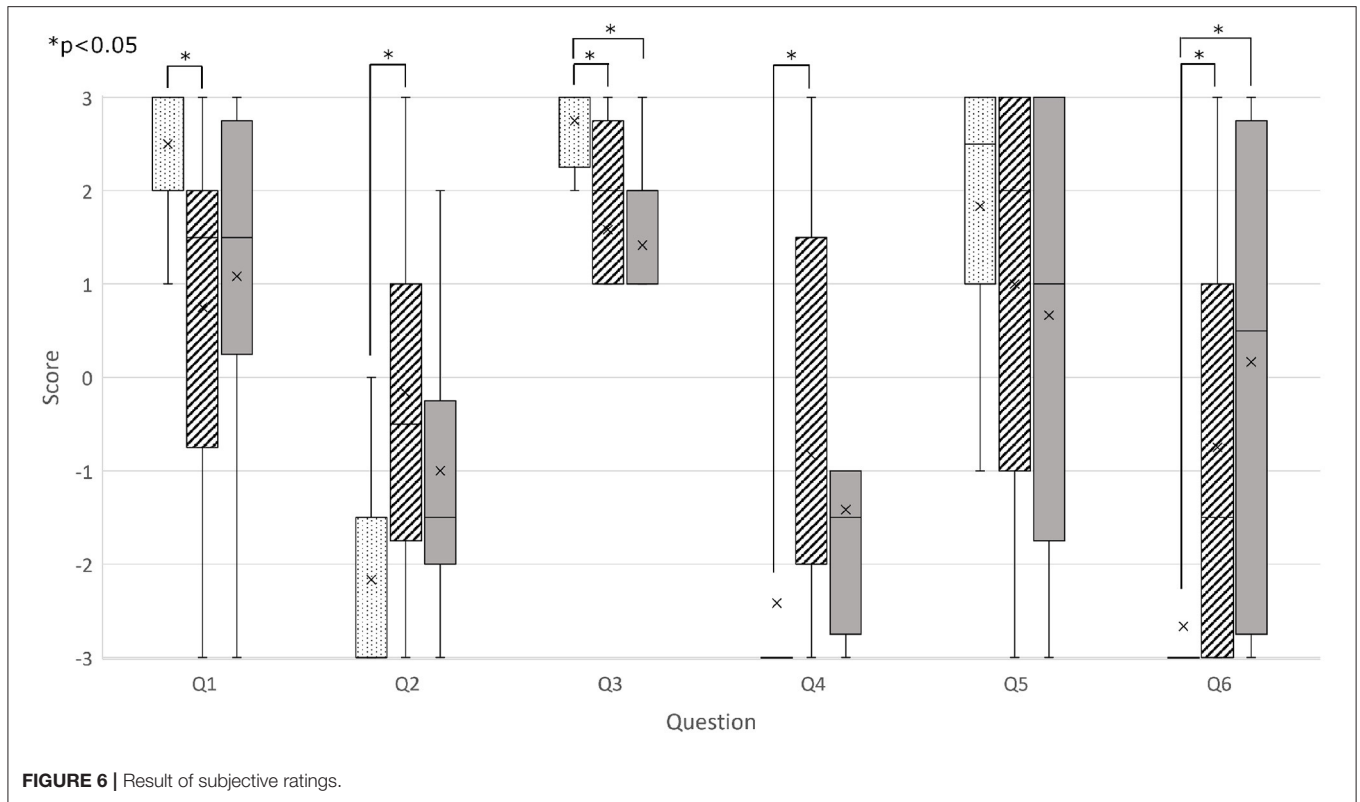
FIGURE 6 | Result of subjective ratings.

FIGURE 7 | Result of subjective ratings: comparison between manipulating the multiple bodies in parallel and switching.

Figure 7 shows the result of comparing the manipulation of bodies in parallel (Q9) and the switching of attention(Q10). In the four-body condition, the score of Q10 was higher than that of Q9 ($p < 0.01$). We used Mann–Whitney U-test for the statistical test.

When embodiment is distributed in multiple bodies, we can set up the two hypotheses that the operations of multiple bodies are performed in parallel or by switching the attention. From the comparison between Q9 and Q10 in each condition, it can

be said that the manipulation of the four bodies is not parallel but is performed by switching the body of attention in the consciousness. Even in the case of two bodies, this body switching is expected to occur in conditions in which the bodies are not visible to each other in light of the earlier discussion. In Q7 and Q8 (free response) of the 4-body condition, participants' comments such as "I looked for the body that seemed to be closest to the ball and focused my attention on that body to perform the task" and "I selected body used for reaching while looking at the ball" were observed. This suggests that participants switch their attention to the body.

The switching the body of attention was done actively by each participant's decisions in Experiment 1. In Experiment 2, we investigated the case where requests of body-switching were generated by external factors due to the task.

6. EXPERIMENT 2

In Experiment 2, to explore the limits of operating multiple bodies directly under the condition of synchronizing with four bodies in different places, we gave the participants a flow task that generated body-switching requests by the balls flying one after another. We found that when a participant is synchronized with multiple bodies, the switching of attention to sensory information from the distributed bodies was reported in Experiment 1. The switching of attention was caused by active behaviors to complete the reaching task. Thus, in this experiment, we tried to explore behavior and
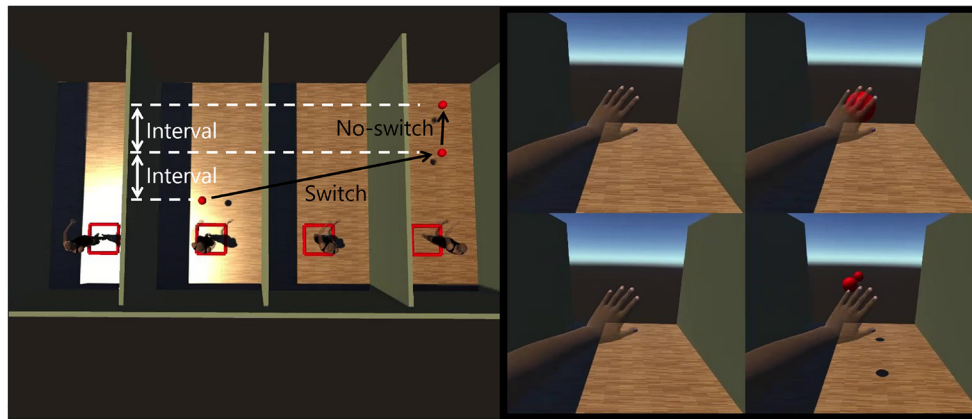
FIGURE 8 | The flow task. (Left) A view of the task from above. (Right) The participant's view.

cognition on passive behavior by having a flow stimulus in a dynamic environment.

6.1. Method

6.1.1. Participants

The same 12 participants participated in Experiment 2 after they performed Experiment1.

6.1.2. Flow Task

We set up a flow task with a ball flying continuously into each body under the conditions of synchronization with the four bodies (**Figure 8**). The balls flew to a range of 0.3 m on each side from the center of the target's body and 1.3–1.6 m in height. The participants were asked to touch the flying balls with their dominant hand. We induced body switching to occur by flying the next ball to a different body target than the previous ball. To observe switching performance when the body switching was requested faster, the intervals between the appearance of the ball could be changed. There were 12 patterns of switching from one body to another in the four bodies. In addition, to randomize the timing of the body switching, we prepared cases where the switching frequency was once per one trial, once per two trials, and once per three trials in each switching pattern. They were randomly combined for 72 trials [36 switches, 36 no-switches; 12 patterns × (1 trial + 2 trials + 3 trials)]. The participants were prevented from occlusion by having their hands made translucent and were informed whether the reaching was a success or a failure by sound. We measured the ball touch success rate in body switching and no-switching trials.

6.1.3. Procedure

Participants wore the HMD and the tracking suits. The skeleton tracking was calibrated. After the experiment was explained, the participants performed the flow task. The flow task was given six times (one practice phase and five performance phases) with varying intervals between ball appearances (0.6, 0.5, 0.4, 0.3, 0.2, 0.1 s). Then, the participants were asked to answer the questionnaire on a 7-level Likert scale of −3 (very strongly

disagree) to 3 (very strongly agree). The questions asked were mainly about the following:

- Q1. It felt like a ball was coming at me.
- Q2. It felt like the ball was flying at me from four directions.
- Q3. It felt as if a ball was flying to each of the four distributed me.
- Q4. What did you think about when you performed the task? (Free Answer)
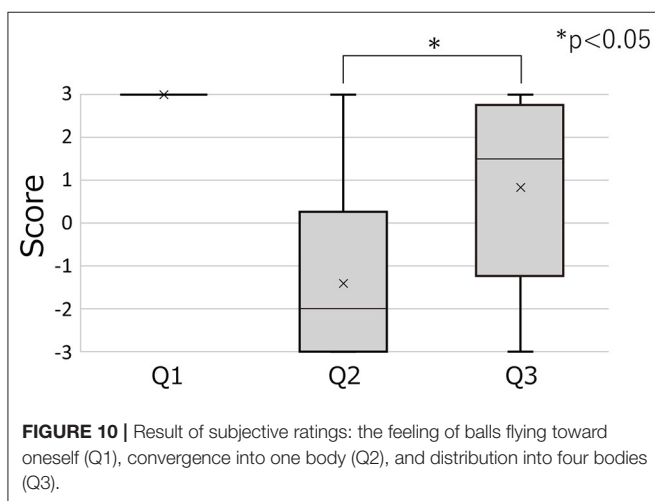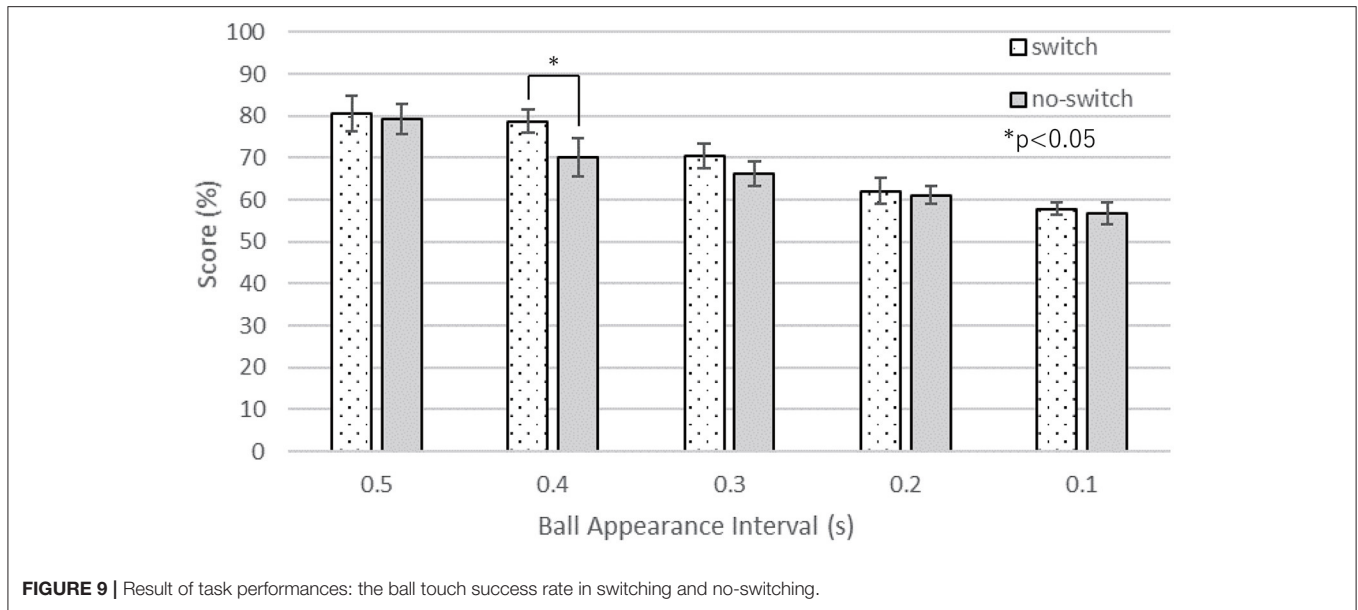- Q5. What were your thoughts on the task? (Free Answer).

6.2. Result

6.2.1. Task Performance

We measured the success rate of touching the flying balls in body switching and no-switching under each interval. The result is shown in **Figure 9**. We performed a paired t-test to compare the task performances between the body switching and no-switching conditions in each interval. When the intervals were 0.4 s, the success rate of the switching trials was higher than that of the no-switching trials ($p < 0.05$).

The task performance decreased as the intervals became shorter in both switching and no-switching trials. This is natural because the needed motor control faster as the interval gets shorter. Only when the intervals were 0.4 s was the success rate of the switching trials higher than that of the no-switching trials. There was no significant difference in the other intervals. Overall, We did not observe advantages in the no-switching trials. It suggests that the participants were able to react correspondingly to the stimuli for the distributed bodies in parallel/semi-parallel even it was supported that the multiple bodies were operated with the sense of switching attention as shown the subjective evaluation in the questionnaire of Experiment 1 (Q9–Q10). The occlusion between the balls in the no-switch trials due to the experimental design may influenced the fact that switching trials showed the better performance.

6.2.2. Subjective Ratings

Figure 10 shows the subjective ratings for the feeling of balls flying toward oneself (Q1), convergence into one body (Q2), and

FIGURE 9 | Result of task performances: the ball touch success rate in switching and no-switching.

FIGURE 10 | Result of subjective ratings: the feeling of balls flying toward oneself (Q1), convergence into one body (Q2), and distribution into four bodies (Q3).

the feeling of distribution into four bodies (Q3). The result of Q1 was answered by all participants as 3 (very strongly agree). We performed a Mann-Whitney U-test between Q2 and Q3. The score of Q3 was higher than that of Q2 ($p < 0.05$).

By looking at the result of Q1, it was observed that participants experienced a strong sensation of a ball coming at them. This suggests that humans can perceive all the presented sensory information from the distributed embodiments as their own.

From the comparison between Q2 and Q3, the sense of being distributed into four bodies was stronger among participants than the sense of being converged in one body. This indicates that motion synchronization and view sharing with multiple bodies induce the distribution of the sense of presence.

Next, we turn our attention to the open-ended questions. Regarding the participants' strategies for the task (Q4), some of them said, "I focused on the field of view where there were

more balls and gave up on the ones that were not in time" and "I thought about which order to move the view." This confirms again that the operation in consciousness is done by switching. In addition, multiple bodies have the potential to improve the efficiency of the task through speculative execution. In the feedback for the task (Q5), there were comments like "Even if I could recognize which field of view the ball was coming, my hand could not catch up" and "I felt a big gap between the position of my virtual hand and my actual hand when I switched the field of attention." These comments confirmed that there is a time lag from the switching of the attention view to the spatial grasp of that view. In addition, it is possible that the body switching may be caused by a passive request, which leads to the perception of a large gap between the body position and the actual hand.

7. EXPERIMENT 3

In Experiment3, we gave a task to avoid balls flying one by one under a four-body condition. In the flow task of Experiment2, it was able to be done with only hand agency. On the other hand, in the task of Experiment3, the participants were asked to avoid balls considering collision with the whole body parts. Thus, this experiment required the body image of full embodiment. We analyzed the participants' gazing information, task performance, and subjective ratings in the avoidance behavior and investigated how multiple bodies are manipulated.

7.1. Method
7.1.1. Participants
Six participants (all male, age: $M = 23.5$, SD = 1.97) joined the experiment. All of them had normal vision and motor ability.
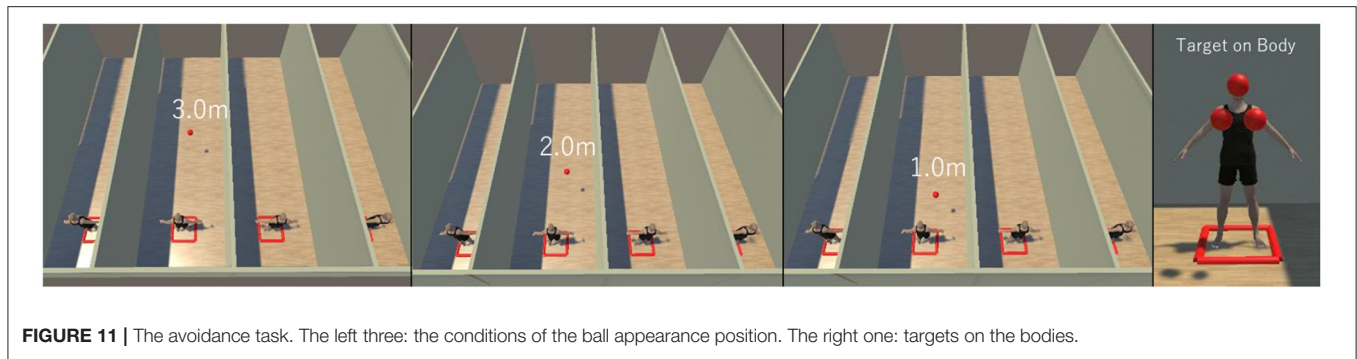
FIGURE 11 | The avoidance task. The left three: the conditions of the ball appearance position. The right one: targets on the bodies.

7.1.2. Avoidance Task

We set up an avoidance task. In this task, the participants were instructed to avoid balls that fly to each body one by one in succession under the four-body-condition. As in Experiment 2, We induced body switching by changing the target of the next ball to a different body than the previous ball. In the four bodies, there were 12 patterns of switching from one body to another. And, to randomize the timing of the body switching, we prepared cases where the switching frequency was once per one trial, once per two trials, and once per three trials in each switching pattern. We randomly combined 72 trials in which the ball flew toward the same body as the previous one and trials in which the ball flew toward a different body than the previous one in order to induce a body switch [36 switches, 36 no-switches; 12 patterns × (1 trial + 2 trials + 3 trials)]. In each trial, the ball was generated randomly toward the head or shoulders of the target body (**Figure 11**), and the flying path of the ball was kept straight. We fixed the speed of the ball 0.5 m/s and the interval between the appearance of the balls at 2.0 s. And we adjusted the required reaction time by varying the distance between the appearance position of the balls and multiple bodies. The ball disappeared when it had passed the target's body. When the ball hit the body, the participants were informed by sound. We measured the ball avoidance success rate and participants' gaze.

7.1.3. Procedure

Participants wore HMD and tracking suits. Then, the body tracking and the eye tracking systems were calibrated. After the instruction of the experiment, the participants performed a practice of the avoidance task (72 trails in 4.0 m condition). Then, the participants performed the ball avoidance task, and they were asked to answer the questionnaire on a 7-level Likert type scale of −3 (very strongly disagree) to 3 (very strongly agree)[6]. The questionnaire consist of the following items:

- Q1. It felt like a ball was coming at me.
- Q2. I could not recognize the ball in time to avoid it.
- Q3. I could not move in time to avoid the ball.
- Q4. I felt a gap between the center of my body and the center of multiple bodies.

[6]The original questionnaires were written in Japanese and translated in English for this publication.

The avoidance task and questionnaire were conducted in three conditions with different ball appearance locations (3.0, 2.0, 1.0 m condition). The order of implementation of the conditions was randomized for each participant.

7.2. Result
7.2.1. Task Performance

As an objective measurement, We obtained the success rate of avoiding the balls. The results of the success rate of switch and no-switch conditions at each appearance position were shown in the **Figure 12**. We conducted a two-factor repeated measures ANOVA on the three ball appearance conditions, and the result showed a significant difference in the ball appearance location factor ($p < 0.01$). We did not find any significant difference in the switch factor ($p > 0.5$), and the interaction effects ($p = 0.98 > 0.50$). Then, we performed a Tukey-Kramer test to compare the task performances between the ball appearance conditions, and the result showed significant differences between 3.0 and 1.0 m conditions ($p < 0.05$) and between 2.0 and 1.0 m conditions ($p < 0.05$) in both switch and no-switch.

The task performance decreased as the intervals became shorter in both switching and no-switching trials. This is reasonable because the required motor control became faster as the appearance position was getting closer. In 3.0 m condition, participants successfully avoided an average of 90%. This means that we can perform avoidance tasks on each body while simultaneously controlling four bodies. We did not observe the difference in task performance between switching and no-switching trials. It suggests that the participants were able to react correspondingly to the stimuli that require the body image of the distributed embodiment, regardless of whether the body switch occurs or not.

7.2.2. Gaze Information

From the gaze information, we analyzed the gaze field the participant was looking at during the task, one of four fields of the split view from each body. First, we focused on the timing when the gaze point shifted from one field of view to another. The number of times the gaze point shifted in each conditions were shown in the **Figure 13**. We conducted a two-factor repeated measures ANOVA on the three ball appearance conditions. We
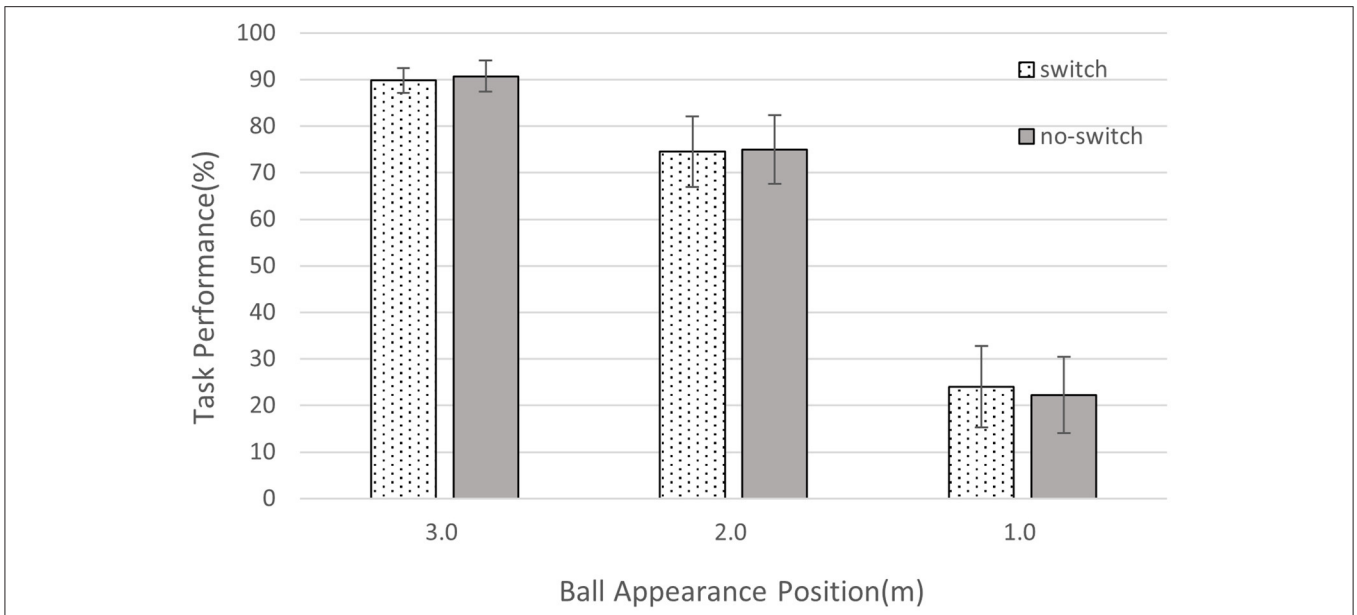
FIGURE 12 | Result of task performances: the ball avoidance success rate in switching and no-switching.
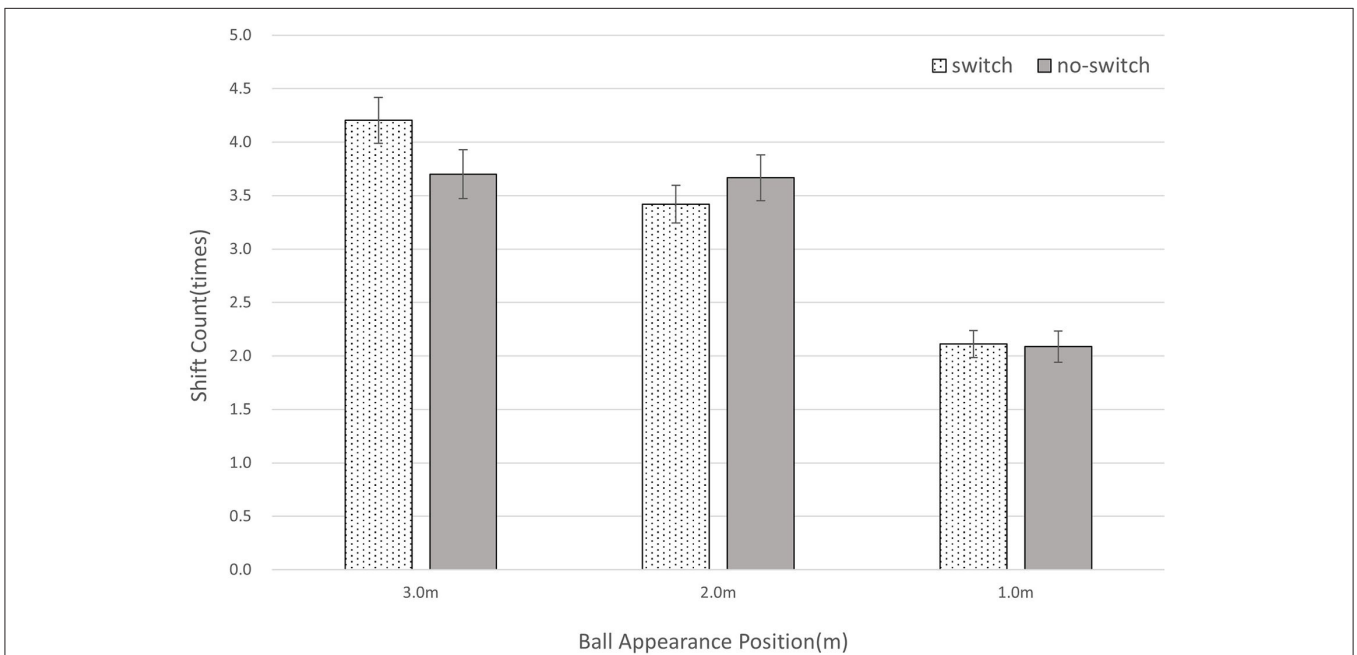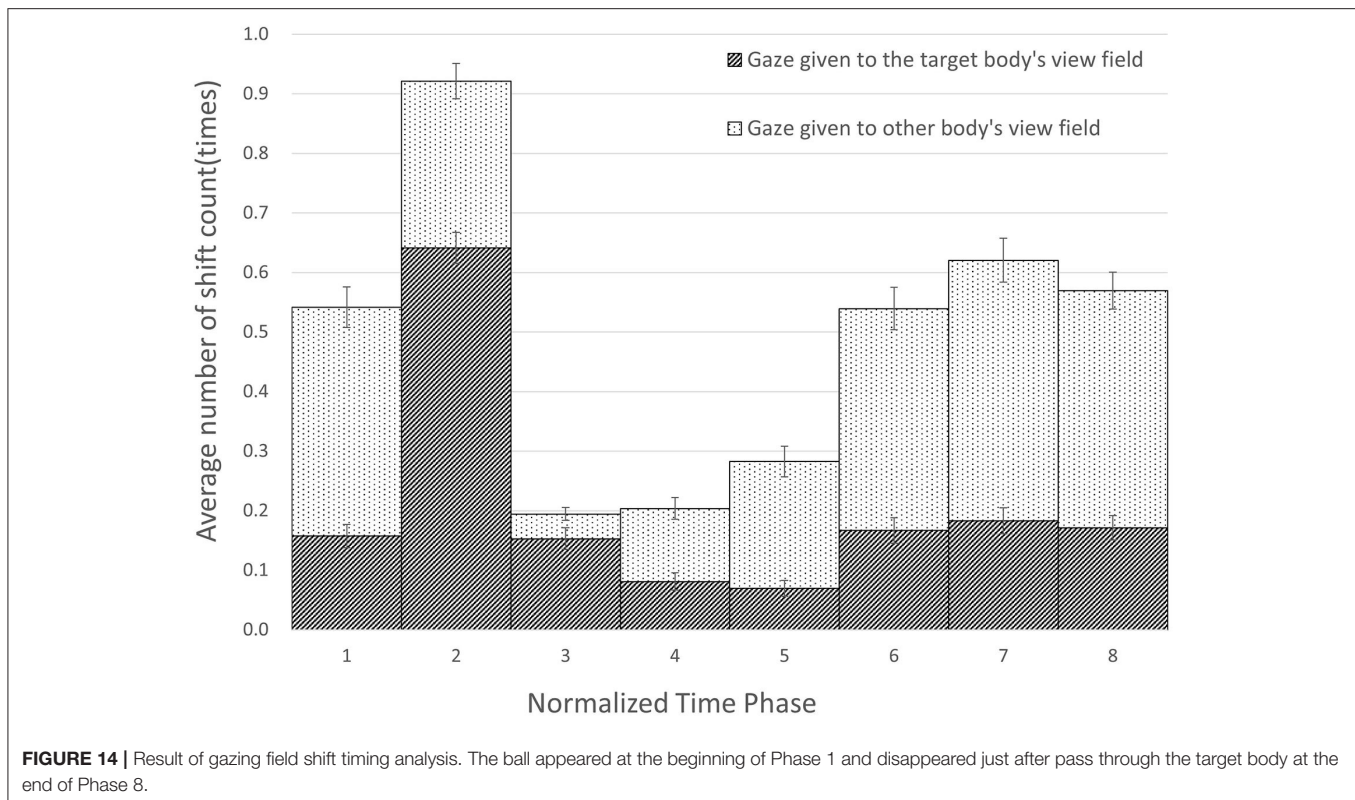

FIGURE 13 | Result of gaze information: the number of times the gazing field shifted in switching and no-switching.

did not find any significant difference in the switch factor ($p > 0.1$), and the interaction effects ($p > 0.1$).

The result shows that participants changed their gaze field and monitored the four views regardless of the switch or no-switch condition. The fact that no significance between conditions on the task performance result can be attributed to this multiple view monitoring.

Then, we focused on how the gaze field shifted during a single trial in which a ball appeared, flied toward one body, and disappeared. The time axis of each trial was normalized and divided into eight phase, with the time when the ball appeared being 0 and the time when it disappeared being 1, and the number of gaze field changes in each period was analyzed. Furthermore, we analyzed whether the gaze field changed to the

FIGURE 14 | Result of gazing field shift timing analysis. The ball appeared at the beginning of Phase 1 and disappeared just after pass through the target body at the end of Phase 8.

view of the target body of the ball in each trial. The result was shown in the **Figure 14**. The Tukey-kramer test was used to compare the number of gaze-field changes for each eighth phase. There were significant differences between one-eighth and two-eighths to five-eighths phases, between two-eighths and three-eighths to eight-eighths phases, and between three-eighths to five-eighths and six-eighths to eight-eighths phases ($p < 0.05$). According to the results, we can observe increases of the gaze field changes at just after the ball appearance and at before the ball disappearance. By analyzing the field of getting gaze and losing gaze, a large percentage of the gaze was given to the target body field immediately after the ball appeared in the two-eighths phase when changes were most frequent. Also, a large percentage of the gaze loss before the disappearance of the ball was observed.

This suggests that, in a single trial, the gaze field shifts occur in response to the ball appearance stimulus, followed by avoidance behavior, and when the ball is modeled to be avoidable, the gaze is changed to monitoring the rest of the distributed bodies.

7.2.3. Subjective Ratings
Figure 15 shows the subjective ratings for the feeling of balls flying toward oneself (Q1), not being able to recognize the ball in time (Q2), not being able to avoid the ball in time (Q3), the feeling of the gap between the center of your body and the center of multiple bodies (Q4). We performed Steel-Dwass test on Q1–Q4 among the appearance conditions. There was no significant difference in Q1 and Q4 ($p > 0.05$). In Q2, there were significant differences between the 3.0 and 1.0 m conditions and between the 2.0 and 1.0 m conditions. In Q3, there were significant difference

between the 3.0 and 1.0 m condition ($p < 0.05$). By looking at the result of Q1, as in Experiment 2, it was observed that a strong sensation of a ball coming at them. This supports that we can perceive each of the presented sensory information from the distributed embodiment as our own. By looking at the results of Q2 and Q3, in the 1.0 m condition, the sensation of not being able to recognize and move in time was stronger than in the 3.0 m condition and only the sensation of not being able to recognize in time was stronger than in the 2.0 m condition. From this, we can say that there is a sense that even if the ball is perceived in time, the motor control can be fail in time as the required avoidance speed increases. By looking at the result of Q4, in the 3.0 m condition, the sensation of the shift between the center of my body and the center of each body was weak. And, although the mean values between the conditions increased, there was no obvious difference. This suggests the avoidance task was done with less feeling of the shift of the center of body even with the split view for the distributed embodiment. This can be a proof of our ability to adapt to the coordinate system of each field of view.

8. DISCUSSION

To summarize the results of Experiments 1, 2, and 3 under the condition of synchronizing with four bodies, the manipulation of the bodies by switching the attention body was performed whether it is the reaching task, the flow task, or the avoidance task. This suggests that the switching of attention to sensory information among multiple distributed bodies is caused by an
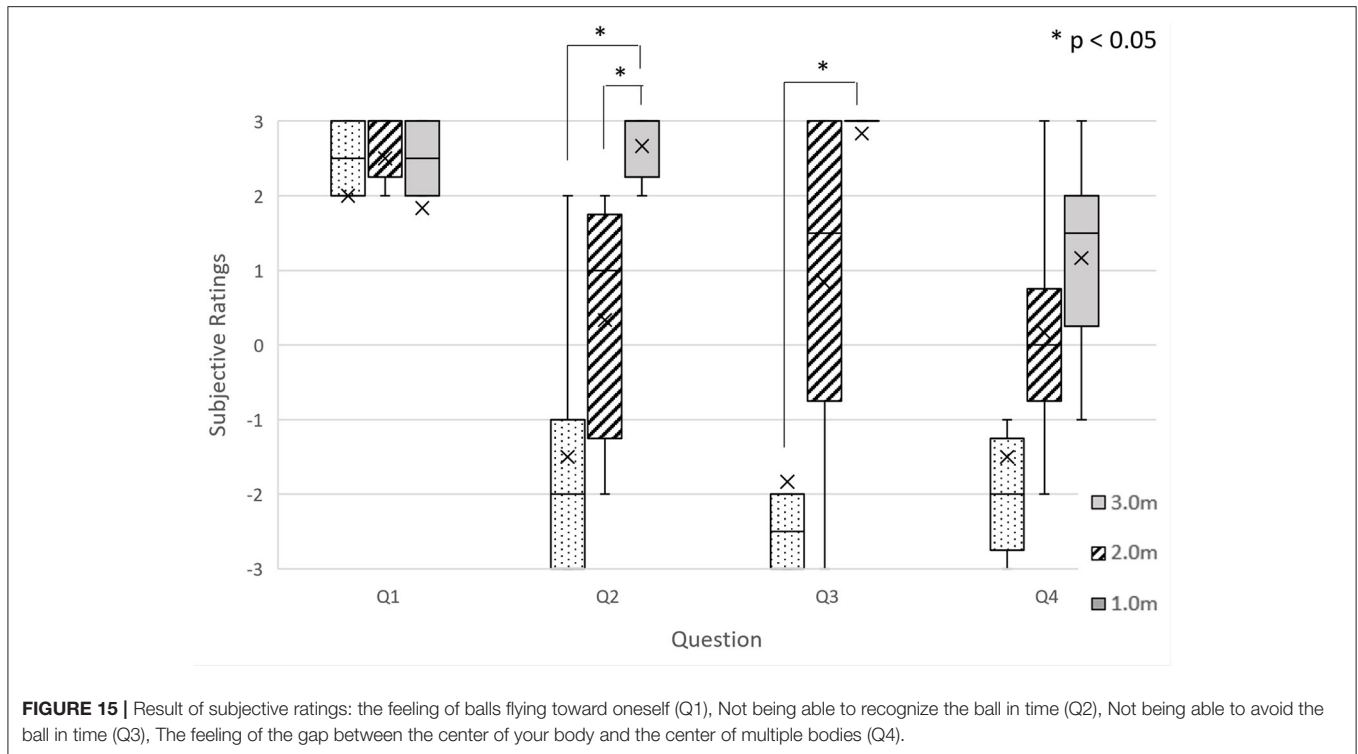
FIGURE 15 | Result of subjective ratings: the feeling of balls flying toward oneself (Q1), Not being able to recognize the ball in time (Q2), Not being able to avoid the ball in time (Q3), The feeling of the gap between the center of your body and the center of multiple bodies (Q4).

active transition to the most appropriate body to complete a task in a static environment and a passive transition to correspond to stimuli for each body in a dynamic environment. Furthermore, in the gaze analysis of the avoidance task, it was observed that the participants monitored multiple views by frequently switching the gaze field immediately after stimulus onset and before passing.

The participants manipulated the four bodies by switching the attention to the sensory information among these bodies while feeling the sense of presence, body ownership, and agency in parallel to the four bodies. This is a very interesting fact. The sense of embodiment can be retained in parallel, but a switch may be needed to perform an intentional behavior. Further investigation is required to verify this.

9. LIMITATION AND FUTURE WORK

In this study, the Experiments were conducted with imbalanced numbers of participants in Experiments 1, 2, and 3. Furthermore, the sample sizes were small. Thus, we applied statistical tests in consideration of this point.

Experiment 1 was conducted in the environment where the heads of multiple bodies were visible to each other. Guterstam et al. (2020) related to distributed embodiment conducted their experiment in a condition where the heads of each other were not visible. The effect of this head visibility on embodiment was not considered in this study.

The questionnaires in the experiment were given to the participants in Japanese. It should be noted that "I" in Japanese does not distinguish between singular and plural.

The sense of embodiment is partly due to the internal forward model and partly due to the reasoning process (Blakemore et al., 2000). The results of questionnaire for participants' reflection include the sense from the reasoning process. Thus, further studies are needed to investigate body sensations close to the internal forward model, for example, using sensory decay by an efference copy of the motor command.

To unify each body, we did not provide any differences for the multiple bodies and things around them in this study. However, adding some differences may help in gaining perceptions of positional relationships with each other. Also, the induction of attention may make the switch more efficient.

We tested a case where all the bodies were similarly synchronized with the participants, but in adapting to multiple bodies, there are other possible synchronization methods, such as mirrored methods and methods involving different body shapes, such as a long arm (Kondo et al., 2018b). There are also other possible operating methods that can be used to switch the body a user is moving by using gaze information.

In this study, there were two types of body sensations: those that could hold four bodies in parallel and those that were switched. We discussed the possibility that the user's intention may influence the difference.

We examined user behavior and cognition on distributed embodiment up to four bodies in the experimental setup. As a future study, we would like to explore the maximum limit for

body cognition by increasing the number of bodies beyond the current configuration.

10. CONCLUSION

In this study, we examined participants' body cognition when experiencing synchronized movements and shared views with multiple bodies through the VE. We set up three tasks and evaluated them by the task performances, the subjective ratings, and the gaze analysis. We found that people can have the sense of embodiment for each body even with the four-body condition when synchronized with multiple bodies. Also, we observed that the moving distance in the reaching task decreased by choosing the proper body in Experiment 1. Our experiment demonstrated that participants are able to manipulate multiple virtual bodies to complete the task. We also observed the sense of body switching in the active reaching with multiple bodies (Experiment 1) and the task performance which suggests the participants were able to react to passive triggers in parallel/semi-parallel with stimuli for each distributed body (Experiment 2 and Experiment 3). In addition, we observed that the participants monitored multiple bodies by shifting the gaze field frequently immediately after stimulus onset and before passing. (Experiment 3).

We believe that distributed embodiment has a potential to bring various extensions such as cooperative work, parallel work, and group behavior by one person for future studies.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the ethics application on this study does not allow the release of the data. Queries regarding the datasets should be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of the Faculty of Science and Technology, Keio University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RM, SK, MK, MI, and MS contributed to the conception and design of the study. RM and AV implemented the system. RM wrote a draft of the paper. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

REFERENCES

Aymerich-Franch, L., Petit, D., Ganesh, G., and Kheddar, A. (2016). The second me: Seeing the real body during humanoid robot embodiment produces an illusion of BI-location. *Conscious. Cogn.* 46, 99–109. doi: 10.1016/j.concog.2016.09.017

Becchio, C., Del Giudice, M., Dal Monte, O., Latini-Corazzini, L., and Pia, L. (2013). In your place: neuropsychological evidence for altercentric remapping in embodied perspective taking. *Soc. Cogn. Affect. Neurosci.* 8, 165–170. doi: 10.1093/scan/nsr083

Blakemore, S.-J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport* 11, R11–R16. doi: 10.1097/00001756-200008030-00002

Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H. H., and Philipsen, A. (2018). The senses of agency and ownership: a review. *Front. Psychol.* 9:535. doi: 10.3389/fpsyg.2018.00535

Fribourg, R., Ogawa, N., Hoyet, L., Argelaguet, F., Narumi, T., Hirose, M., et al. (2020). Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals. *IEEE Trans. Visual. Comput. Graph.* 27, 4023–4038. doi: 10.1109/TVCG.2020.2999197

Furlanetto, T., Bertone, C., and Becchio, C. (2013). The bilocated mind: new perspectives on self-localization and self-identification. *Front. Hum. Neurosci.* 7:71. doi: 10.3389/fnhum.2013.00071

Glas, D. F., Kanda, T., Ishiguro, H., and Hagita, N. (2011). Teleoperation of multiple social robots. *IEEE Trans. Syst. Man Cybern. Part A* 42, 530–544. doi: 10.1109/TSMCA.2011.2164243

Guterstam, A., Larsson, D. E., Szczotka, J., and Ehrsson, H. H. (2020). Duplication of the bodily self: a perceptual illusion of dual full-body ownership and dual self-location. *R. Soc. Open Sci.* 7:201911. doi: 10.1098/rsos.201911

Kasahara, S., Ando, M., Suganuma, K., and Rekimoto, J. (2016). "Parallel eyes: exploring human capability and behaviors with paralleled first person view sharing," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, (San Jose, CA), 1561–1572. doi: 10.1145/2858036.2858495

Kawasaki, H., Iizuka, H., Okamoto, S., Ando, H., and Maeda, T. (2010). "Collaboration and skill transmission by first-person perspective view sharing system," in *19th International Symposium in Robot and Human Interactive Communication* (Viareggio), 125–131. doi: 10.1109/ROMAN.2010.5598668

Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence* 21, 373–387. doi: 10.1162/PRES_a_00124

Kilteni, K., Normand, J.-M., Sanchez-Vives, M. V., and Slater, M. (2012b). Extending body space in immersive virtual reality: a very long arm illusion. *PLoS ONE* 7:e40867. doi: 10.1371/journal.pone.0040867

Kishore, S., Muncunill, X. N., Bourdin, P., Or-Berkers, K., Friedman, D., and Slater, M. (2016). Multi-destination beaming: apparently being in three places at once through robotic and virtual embodiment. *Front. Robot. AI* 3:65. doi: 10.3389/frobt.2016.00065

Kondo, R., Sugimoto, M., Minamizawa, K., Hoshi, T., Inami, M., and Kitazaki, M. (2018a). Illusory body ownership of an invisible body interpolated between virtual hands and feet via visual-motor synchronicity. *Sci. Rep.* 8, 1–8. doi: 10.1038/s41598-018-25951-2

Kondo, R., Ueda, S., Sugimoto, M., Minamizawa, K., Inami, M., and Kitazaki, M. (2018b). "Invisible long arm illusion: illusory body ownership by synchronous movement of hands and feet," in *ICAT-EGVE* (Limassol), 21–28.

Kurosaki, K., Kawasaki, H., Kondo, D., Iizuka, H., Ando, H., and Maeda, T. (2011). "Skill transmission for hand positioning task through view-sharing system," in *Proceedings of the 2nd Augmented Human International Conference on - AH'11* (Tokyo: ACM Press). doi: 10.1145/1959826.1959846

Mitchell, R., Bravo, C. S., Skouby, A. H., and Möller, R. L. (2015). "Blind running," in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '15* (New York, NY: ACM Press), 649–654. doi: 10.1145/2793107.2810308

Nakul, E., Orlando-Dessaints, N., Lenggenhager, B., and Lopez, C. (2020). Measuring perceived self-location in virtual reality. *Sci. Rep.* 10:6802. doi: 10.1038/s41598-020-63643-y

Pan, R., Singhal, S., Riecke, B. E., Cramer, E., and Neustaedter, C. (2017). "MyEyes: the design and evaluation of first person view video streaming for long-distance couples," in *Proceedings of the 2017 Conference on Designing Interactive Systems* (New York, NY: ACM), 135–146. doi: 10.1145/3064663.3064671

Petkova, V. I., and Ehrsson, H. H. (2008). If I were you: perceptual illusion of body swapping. *PLoS ONE* 3:e3832. doi: 10.1371/journal.pone.0003832

Saraiji, M. Y., Sasaki, T., Kunze, K., Minamizawa, K., and Inami, M. (2018). "Metaarms: body remapping using feet-controlled artificial arms," in *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (New York, NY), 65–74. doi: 10.1145/3242587.3242665

Sypniewski, J., Klingberg, S. B., Rybar, J., and Mitchell, R. (2018). "Towards dynamic perspective exchange in physical games," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, Vol. 2018* (New York, NY: ACM), 1–6. doi: 10.1145/3170427.3188674

Waterworth, J., and Waterworth, E. (2014). Distributed embodiment: real presence in virtual bodies. 589–601. doi: 10.1093/oxfordhb/9780199826162.013.024

Wissmath, B., Weibel, D., Schmutz, J., and Mast, F. W. (2011). Being present in more than one place at a time? Patterns of mental self-localization. *Conscious. Cogn.* 20, 1808–1815. doi: 10.1016/j.concog.2011.05.008

Conflict of Interest: SK was employed by the company Sony.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.