Swarm Body: Embodied Swarm Robots

Sosuke Ichihashi*
Georgia Institute of Technology
Atlanta, Georgia, USA
sichihashi3@gatech.edu

So Kuroki
OMRON SINICX Corporation
Tokyo, Japan
so.kuroki@sinicx.com

Mai Nishimura
OMRON SINICX Corporation
Tokyo, Japan
mai.nishimura@sinicx.com

Kazumi Kasaura
OMRON SINICX Corporation
Tokyo, Japan
kazumi.kasaura@sinicx.com

Takefumi Hiraki
Cluster Metaverse Lab
Tokyo, Japan
t.hiraki@cluster.mu

Kazutoshi Tanaka
OMRON SINICX Corporation
Tokyo, Japan
kazutoshi.tanaka@sinicx.com

Shigeo Yoshida
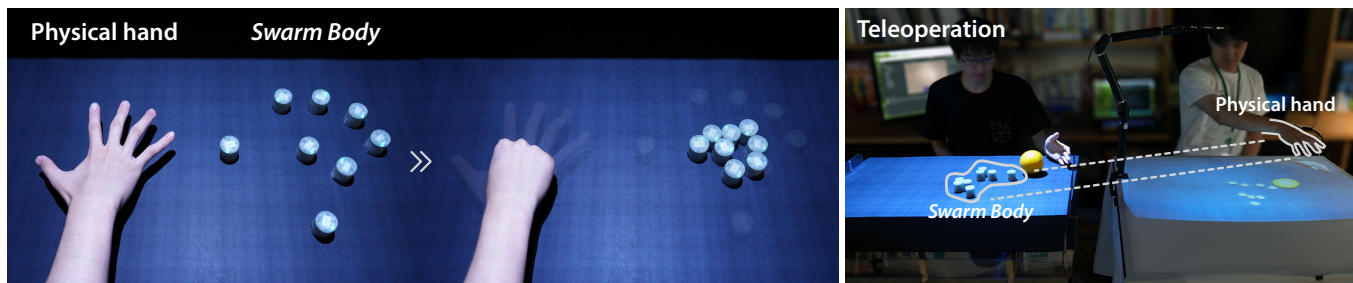OMRON SINICX Corporation
Tokyo, Japan
shigeo.yoshida@sinicx.com

Figure 1: Embodied swarm robots allow a human to dynamically shape and disassemble their bodies to physically and adaptively interact with the local and remote environments. In the right figure, the user on the right pushes a ball toward the other person's hand with embodied swarm robots. The user can see the swarm robots projected on the table and control them in an embodied manner with the hand tracker attached to the table.

ABSTRACT

The human brain's plasticity allows for the integration of artificial body parts into the human body. Leveraging this, embodied systems realize intuitive interactions with the environment. We introduce a novel concept: embodied swarm robots. Swarm robots constitute a collective of robots working in harmony to achieve a common objective, in our case, serving as functional body parts. Embodied swarm robots can dynamically alter their shape, density, and the correspondences between body parts and individual robots. We contribute an investigation of the influence on embodiment of swarm robot-specific factors derived from these characteristics, focusing on a hand. Our paper is the first to examine these factors through virtual reality (VR) and real-world robot studies to provide essential design considerations and applications of embodied swarm robots. Through quantitative and qualitative analysis, we identified a system configuration to achieve the embodiment of swarm robots.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**.

KEYWORDS

embodiment, swarm robotics, tangible interaction

1 INTRODUCTION

The plasticity of the human brain allows it to recognize objects as body parts and manipulate them intuitively under certain conditions. This cognitive ability has led to the development of embodied systems that enhance intuitive manipulation and immersion, such as teleoperated robotic arms or virtual avatars. These systems have proven invaluable; however, there remains untapped potential in

flexibility, adaptability to different environments, and overall robustness. For example, it is difficult for conventional embodied systems to increase their levels of embodiment by dynamically changing their size to suit the individual, changing the shape to one suitable for a certain action, or for use on a tabletop or in an environment with obstacles. To address these challenges, we introduce the concept of *Embodied Swarm Robots*.

Swarm robots consist of many small robots of similar types, such as a school of fish or a swarm of insects.Unlike conventional robotic arms, swarm robots are expected to cooperate and exert significant effects. This unique quality endows them with robustness, flexibility, and scalability, enabling tasks such as navigation through narrow paths [56], pattern formation [1], self-assembly [45], and the collective transport of objects that are larger and more complex than themselves [33]. Thus, swarm robots have a wide range of applications such as environmental monitoring, space exploration, agriculture, emergency rescue, warehouses, industrial plants, entertainment, surveillance, and maintenance [60].

Imagine a human body composed of *swarms* of robots, offering unprecedented adaptability. By using swarm characteristics, individuals can dynamically shape and equip their bodies to perform specific tasks and situations. For example, a human with a swarm body can effortlessly traverse through narrow paths accessible to the constituent robots of the swarm. Moreover, if some individual robots within a swarm are lost, the remaining individuals compensate for the loss of functionality. Embodying swarm robots holds promise for combining intuitive and immersive interactions with adaptability. However, the realization of swarm robot body systems remains challenging because of our limited understanding of the conditions to support the sense of embodiment and subsequent system design considerations.

Although previous studies explored the embodiment of robot arms and virtual avatars, the embodiment of swarm robots introduces additional complexities. For example, the size, density, and position distribution of robots, as well as the algorithms assigning the robots to the positions, can influence the levels of embodiment. Unlike robot arms, in which each moving link is constrained by joints, relative positions (*i.e.*, position distributions) of swarm robots can change without geometrical constraints other than collisions, which gives them flexibility. Thus, when representing a hand in a particular posture, various position distributions are possible, such as placing robots at the joint positions of the hand or placing them such that they are equally distributed within the hand shape. In addition, the algorithm assigning robots to this position distribution influences the embodiment. When moving a hand, the hand state (*i.e.*, posture and position) changes dynamically. Therefore, the robots need to follow each hand state while constantly assigning themselves to the position distribution for the current hand posture. One possible assignment method is to statically assign a particular robot to a particular position on a hand, whereas another method is to dynamically update the assignments so that the total sum of the travel distances is minimized. Therefore, identifying an appropriate algorithm for position distribution generation and assignment to follow the dynamically changing body states, as well as the size and density of robots, is crucial for the successful embodiment of swarm robots.

Similar to the many previous studies on embodiment introduced in section 2, this study investigated the embodiment of swarm robots by focusing on the hand, which is the part of the body where people most frequently interact with the environment. In addition, we focused on tabletop swarm robots because we intended to explore various everyday interactions that people have with their hands at a table. To evaluate the level of embodiment, we measured the sense of body ownership and agency [12]. The factors examined were robot size, density, position distribution generation algorithm, and assignment algorithm. VR and real-world robot experiments were conducted. In the VR experiment, all the aforementioned factors were explored using simulated swarm robots. This shows that swarm robots can be embodied and provides various insights into the embodiment of swarm robots with ideal swarm robot behavior. However, actual robots may behave differently from those in VR environments. To check whether swarm robots can be embodied and if similar embodiment characteristics are obtained in the real world, a similar embodiment experiment was conducted with the physical swarm robots. Based on the results of the VR and real-world experiments, we demonstrated the characteristics of swarm robot embodiment under ideal and actual robot behaviors, as well as design considerations for embodied swarm robot systems. By comparing these results, we discuss how the characteristics of a real-world system affect the embodiment of swarm robots, thereby providing useful information for the future design of new embodied swarm robot systems. Our contributions encompass:

(1) Proposing a framework for the embodiment of swarm robots in the hand.
(2) An algorithm to determine position distribution relatively from the hand skeleton and dynamically assign them to robots, as determined through a series of VR and physical experiments, enhanced the sense of body ownership, sense of agency, and overall usability compared with other conditions.
(3) Suggesting practical implementations and applications for embodied swarm robots that integrate these findings.

2 RELATED WORK
2.1 Body Ownership and Agency

The sense of embodiment is a fundamental product of the human mind. A key aspect of embodiment is the sense of body ownership. Body ownership is *"the sense that I am the one who is undergoing an experience"* [12]. Botvinick and Cohen [5] and others revealed that humans can perceive the sense of body ownership toward artificial objects with synchronous visuotactile stimulation. In addition to synchronous visuotactile stimulation, synchronous visuomotor feedback also affects embodiment [57]. Synchronous visuomotor feedback sheds light on another key aspect of embodiment: the sense of agency. The sense of agency is *"the sense that I am the one who is causing or generating an action"* [12]. From the perspective of the sense of body ownership and agency, researchers have expanded their knowledge of various aspects of human embodiment, such as its temporal [59, 62, 63, 72] and spatial [18, 22, 41, 53] characteristics, as well as the embodiment possibilities of artificial bodies.

With advancements in VR technology, many scholars have explored the embodiment possibilities of various non-biological bodies, including elongated [31], invisible [28], discontinued [50, 70], scrambled [27], re-associated [29, 30], shared [11, 13, 69], multiple [44], and supernumerary [2, 20] body parts, as well as noncorporeal objects [42], through synchronous visuomotor feedback using VR. These studies revealed that humans can embody not only their innate body parts with their original configurations but also those augmented or modified in certain ways as well as noncorporeal objects. This knowledge can be applied in robotics to develop embodied robots that realize intuitive interactions, similar to human interactions with innate bodies.

2.2 Robot Embodiment

Robot embodiment has been studied for on-site and remote intuitive robot manipulations to interact with the environment physically. For example, Aymerich-Franch *et al.* examined whether non-human–looking humanoid robot arms could be perceived as the user's arms and reported that the levels of embodiment were similar between human- and non-human–looking arms [4]. Takada *et al.* showed that a single user could feel a sense of agency toward multiple robot arms while playing ping-pong with two opponents [68]. Most studies on virtual and robot embodiments have focused on body configuration similar to the human biological body. Recently, researchers have begun to explore the embodiment of Supernumerary Robotic Appendages (SRAs) or unconventional body parts, such as additional limbs [58, 71, 76], hands [24, 77], fingers [51, 61], joints [36], and tails [20]. The embodiment of an SRA or an unconventional body has the potential to augment human abilities and interactions with the environment, which are currently limited by the innate body configuration. For example, tentacle limbs offer greater freedom than innate human limbs. Similarly, swarm robots have the potential to provide robustness, flexibility, and scalability to a human body that conventional humans do not possess. Therefore, this study is the first to investigate the embodiment of a swarm to realize embodied swarm robots.

2.3 Swarm Robots in HCI

Swarm robots operate such that many small robots of similar types cooperate with each other to exert significant effects. This unique property endows the devices with robustness, flexibility, and scalability. In HCI, actuated objects were used as tangible user interfaces (TUIs) to allow users to interact with the digital world through physical tabletop objects [10, 48, 49, 52, 54, 65]. Zooids [34] expanded these tangible interaction possibilities by introducing tabletop swarm robots and introduced the concept of Swarm User Interfaces (Swarm UIs). This design space has been further extended by developing new robots [8, 35], adding new action possibilities to robots [40, 45, 66, 67, 78, 79], building mixed reality system cooperative with projected images [15, 16], and introducing interfaces that seamlessly connect the digital and physical worlds [19, 21, 38, 46]. Fundamental applications include sensing, visual feedback, haptic feedback, shape presentation, object representation, object actuation, environmental adaptation, and collaborative actions. As such, the multi-agent nature of swarm robots has realized a flexible and scalable interface between the physical and digital worlds.

This multi-agent nature makes user control of swarm robots challenging. Three main approaches have been adopted to control swarm robots: predefined, physical, and synchronized. Many previous works have used predefined control in which a certain action or state triggers a certain robot's program. Kim *et al.* compared various human-control modalities and strategies for predefined swarm robot control and provided guidelines [25]. Another approach is to apply physical laws to robots so that users can expect and control their behavior [21, 38]. Synchronized control, where swarm robots act in sync with corresponding remote or virtual objects, is also widely used, particularly in the context of TUIs, VR, and physical telepresence. For example, HoloBots [19] move such that their positions match the positions of the corresponding robots or the user's body part in the remote physical world. These methods enabled physical interactions with the digital or remote physical worlds. However, these swarm robots do not aim for embodiment or seek to represent human body parts. By embodying swarm robots, the robot user can interact with the environment more intuitively and precisely while maintaining the power of the swarm robots. In addition, the observer can understand the intentions of robot users more intuitively and precisely while feeling more humanness or affection toward the robots. In human-robot interactions, an early attempt to control swarms in an embodied manner using fingers has already been made [23]; however, the robots do not represent the hand well, and their embodiment and interaction opportunities have not been investigated. Inspired by these, this study explores the interaction opportunities of embodied swarm robots.

3 FRAMEWORK FOR SWARM ROBOTS EMBODIMENT

We first introduce our framework to represent a hand with swarm robots in real-time. As our body moves dynamically, navigating a swarm of robots has dynamic destinations associated with each timestep rather than a single static goal. We refer to these dynamic destinations as subgoals that correspond to specific timesteps. Assuming that the hand skeleton coordinates are tracked at regular intervals, we define the robot destination coordinates for each hand state as *subgoal positions*. If each robot moves to each subgoal position instantly, we can focus on how to represent each hand state with subgoal positions.

However, swarm robots often do not reach the subgoal positions before the subgoals are updated. Therefore, it is necessary to control swarm robots to follow the subgoal positions so that the robots' movements represent the movement of the body part. Thus, to realize embodied motion with swarm robots, we require a new framework that simultaneously realizes both the proper hand representation and the smooth collective follow of swarm robots.

Common steps to control swarm robots are getting a subgoal formation (a collection of subgoal positions), assigning the subgoal positions to the robots (assignment), and obtaining a path for each pair of robots and its subgoal position (path planning). Inspired by this, we took the following three steps to achieve embodied movements of swarm robots as shown in Figure 2:

(1) generate a subgoal formation based on the current hand position and shape;
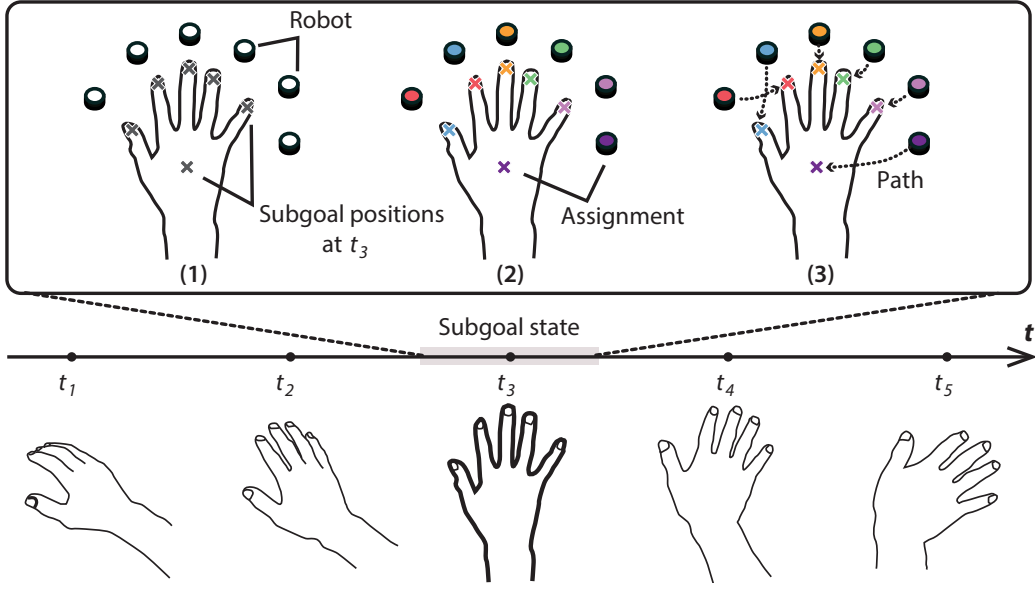(2) assign the generated subgoal positions to the robots;

Figure 2: Example of the (1) subgoal formation generation, (2) robot assignment, and (3) local path planning.

(3) obtain local paths and move the robots accordingly to avoid collisions with each other and obstacles.

3.1 Subgoal Formation Generation

Step (1) sets the subgoal formation for the swarm robots. Ideally, these positions are where the robots are; hence, this step affects the visual representation of the hand as well as the usability and interaction opportunity.

Because this step is similar to the virtual body representation, we took inspiration from that research. We explored abstract virtual body representation approaches that are applicable to two-dimensional representation because we applied them to tabletop swarm robots. Consequently, we identified two main approaches: point- and silhouette-based [74]. The point-based approach represents the body with a number of points corresponding to certain points on the body, whereas the silhouette-based approach represents a body with a solid shape in a single color. In other words, the former focuses on specific points, and the latter focuses on overall shapes. Based on these approaches, we designed two algorithms to obtain a set of subgoal positions from hand position, orientation, and shape, shown in Figure 3.

3.1.1 Bone-Based Subgoal Formation Generation. Using the point-based approach, a set of points on the hand is determined, and the subgoal positions are moved according to the relative positions and movements of those points. The set of points on the hand must be fixed with respect to the skeleton. Thus, the bones of the hand were used as references for the tracked points. Therefore, we refer to this as a bone-based algorithm. Because the swarm robots we examined only moved on the horizontal plane, the bone positions were projected on the horizontal plane before calculating the subgoal positions. For example, if the bone at the tip of the index finger is set as the subgoal, the robot moves to the bone position of

the hand projected onto a horizontal plane. The subgoal positions relative to the bones are predetermined (*e.g.*, Figure 6 in our study).

3.1.2 Silhouette-Based Subgoal Formation Generation. Using the silhouette-based approach, the subgoal formation should reflect the overall shape of the hand. To achieve this, a sensed hand skin mesh was obtained, and its vertices were clustered based on the number of required subgoal positions using the k-means algorithm [43]. The subgoal positions were set to the vertex positions closest to the centroid of each cluster ensuring they remained within the hand outline. The centroid itself may be placed outside the hand's outline, especially when clusters span multiple fingers, have gaps between fingers, or are around the proximal phalanges. We call this the silhouette-based algorithm.

We decided to examine both bone- and silhouette-based algorithms because they performed differently during our preliminary testing with two-dimensional hand representations using swarm robots. The bone-based algorithm can offer more predictable robot movements to the user than the silhouette-based algorithm because the subgoal positions are always fixed to certain locations of the hand. The silhouette-based algorithm constantly updates the subgoal positions on the hand. Therefore, it does not guarantee that the subgoal positions are on the parts the user expects. A potential advantage of the silhouette-based algorithm is its adaptability to various hand signs. For example, when the user closes their hand, in the bone-based algorithm, the distance between the bones when projected onto the horizontal plane becomes so small that the subgoal positions become close, and the robots may collide. However, in the silhouette-based algorithm, the subgoal positions are rarely too close to each other because clustering is conducted for that hand shape.
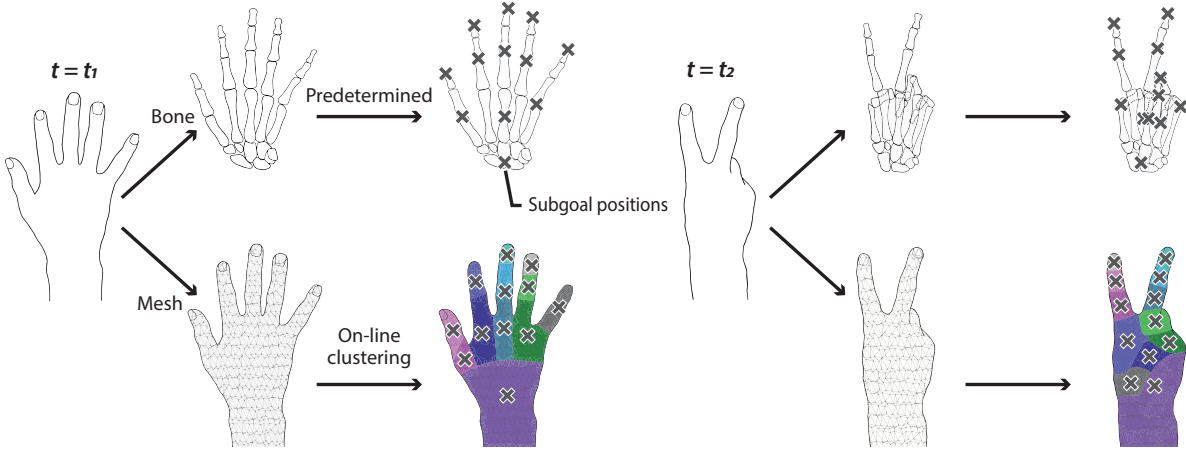
Figure 3: Bone- (top) and silhouette-based algorithms (bottom) to obtain subgoal formations. The bone-based algorithm obtains the subgoal formation based on the hand bone positions projected on the horizontal plane. The numbers in the figure correspond to the bone IDs provided by the Quest 2 hand tracking. The silhouette-based algorithm classifies the vertices of the hand mesh projected on the horizontal plane into clusters (shown in different colors in the figure) based on the number of robots. Then, the vertices closest to each cluster centroid become the subgoal positions.

3.2 Subgoal Position Assignment

Once the subgoal positions are obtained, they are assigned to the robots. Two assignment methods are considered: *static* and *dynamic*. The static method assigns a specific subgoal position to the same robot. In dynamic assignment, the subgoal positions are constantly reassigned to realize smoother transitions from one hand shape to another. This problem is defined as an assignment with variable subgoal formation. This was attributed to the linear sum assignment problem [7]. This problem can be solved by using the Hungarian algorithm [32].

The *bone-static* algorithm will offer more predictable robot movements than the other algorithm because the same robots always follow the same parts of the hand owing to the fixed subgoal formation and assignment. By contrast, the dynamic assignment can avoid potential collisions of the robots. For example, when the user turns the hand from facing upward to facing downward, dynamic assignment reassigns robots so that they do not have to flip their positions.

When the subgoal positions are constantly generated, and the current subgoal positions cannot be mapped to the past ones (*i.e.*, when the silhouette-based subgoal formation is used), the static assignment cannot be applied because the same subgoal position does not exist at the next instant. Therefore, the possible subgoal formation generation and assignment algorithms are: *bone-static*, *bone-dynamic*, and *silhouette-dynamic*.

3.3 Robot Control with A Local Path Planner

Once pairs of a robot and subgoal position are determined, path planning and following robot control are executed. To plan paths and move multiple robots to their assigned subgoal positions, the Reciprocal Velocity Obstacles (RVO) algorithm [73] was used. The RVO algorithm is an extension of the Velocity Obstacle concept,

which offers navigation among passively moving objects by treating them as obstacles in the velocity space. The RVO algorithm incorporates the assumption that other actively moving objects perform a similar collision avoidance behavior to the Velocity Obstacle and realized navigation among both passively and actively moving objects. More precisely, the RVO algorithm with nonholonomic constraints [64] was used as most of the swarm robots, including the tabletop swarm robots we use, are nonholonomic.

4 EMBODIMENT EXPERIMENT IN VR

We first conducted a VR experiment to examine how various factors, including robot size, affect the levels of embodiment. Then, to validate our findings in the real world, we conducted another embodiment experiment with fewer factors based on the VR results. This experiment aimed to determine when swarm robots give users a sense of body ownership and agency. Our experiment, designed based on prior embodiment studies, explored how specific swarm robot parameters affect body ownership, agency, and task load. Participants interacted with virtual swarm robots, varying in size, density, and control algorithms. After each trial, they answered questionnaires evaluating the sense of body ownership, sense of agency, and cognitive load.

4.1 Participants

The experiment involved 10 participants (6 males, 4 females; average age: 24.20 ± 2.57 SD). Participants were sourced from a recruitment post on social media. All participants were right-handed with normal or corrected vision and were unaware of the experiment's purpose. Half of the participants had minimal VR experience, while the other half had extensive experience. Participants signed a consent form regarding the experiment and were compensated with approximately $16 in Amazon gift cards. The ethics review board approved the experiment.

4.2 Apparatus and Setup

The experiment program, implemented using Unity, simulates and visualizes swarm robots and runs on a Windows-based computer. A Meta Quest 2 HMD[1] tracked participants' hands and provided visual feedback based on the Unity visualization. Participants wore noise-canceling headphones playing white noise to block external sounds. An iPad collected the post-trial questionnaire responses. In the VR, robots appeared on a table in front of the user as cylindrical bodies, a common form in HCI research [34]. The subgoal formations are located on the table surface, and they are not translated in the horizontal direction from the participant's actual hand; *i.e.*, there was no spatial discrepancy in the horizontal direction between the participant's hand and the robot's subgoal formation.
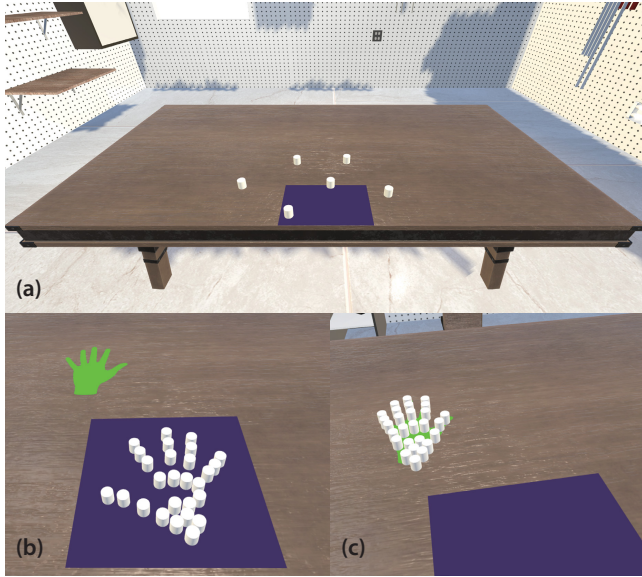
4.3 Experiment Design

Figure 4: The VR setup for the reaching tasks (a). The purple area is the starting area, and the green object is the hand-shaped reaching target. Participants were instructed to (b) reach the target with the specified hand sign and (c) fit the swarm robots in the green area.

We examined parameters (independent variables) specific to swarm robots. The potential design parameters include robot's latency, speed, acceleration, size, color, shape, density, and control algorithm. As it is not feasible to examine all the parameters, we focused on some of the parameters that are more unique for swarm robots: size, density, and control algorithm. We excluded latency from the examined parameters because the effect of delay on embodiment is not unique to swarm robots and has been investigated in visual-motor synchronicity [62]. In a system with program-robot communication, we should consider the effect on embodiment owing to the sum of the delay caused by the robot's
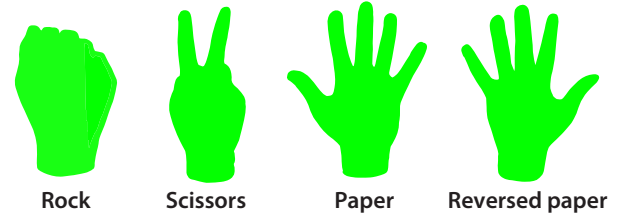

Figure 5: Hand signs used in the VR embodiment experiment. Participants make the rock, scissors, and paper shapes with their right hands, palms down. For the reversed paper sign, the palm should face up.

movement performance and the delay caused by this communication. The first delay was excluded from this experiment by setting the robot's wheel speed to 400 mm/s with which we did not see much latency in the pilot study. The second delay was excluded by removing the latency in the program-robot communication.

We used a factorial design with three factors: 2 levels of the robot's size, 3 levels of the density, and 3 levels of the control algorithm. The independent variables examined were robot's size (*30 mm* and *20 mm*), density (*sparse*, *medium*, and *dense*), and subgoal position generation and assignment algorithm (*bone-static*, *bone-dynamic*, and *silhouette-dynamic*). All variables were within participants.

The experiment was conducted in the virtual environment shown in Figure 4. Participants were tasked with guiding swarm robots to a target sheet while making specific hand signs using their right hand. The targets were green hand-shaped sheets with variations shown in Figure 5. Participants were instructed to change their hand signs to the target shapes when reaching the targets. A variety of hand signs were provided so that participants could explore the pros and cons of the subgoal position generation and assignment algorithms in various situations. For example, the silhouette-based algorithm may have fewer collisions than the bone-based algorithm for the rock and scissors hand signs, but they may not make a big difference for the paper hand sign. In the preliminary testing, the static and dynamic subgoal assignment algorithms resulted in very different robot behaviors (*i.e.*, the static assignment sometimes caused the robots to get stuck, while the dynamic assignment did not) when the hand was flipped over. Therefore, the reversed paper hand sign was also provided.

Participants were instructed to move the swarm robots representing their right hand to the purple starting area at the beginning of each task by moving their right hand. Once all the robots have stayed in the starting area for two seconds, a green hand-shaped reaching target appeared either at the left-front or at the right-front of the starting area. Specifically, the target appeared 300 mm to the rear and 173 mm to the side from the center of the starting area. The participants were instructed to reach for the object with the hand sign and to fit the robots in the target area as fast as possible. The target disappeared five seconds after the task started. The participants then return their hands and robots to the starting area.


Table 1: Number of Robots for Each Size and Density Condition.

	Sparse	Medium	Dense
20 mm	6	18	27
30 mm	6	8	12

Each trial consisted of eight tasks, *i.e.*, 4 hand signs × 2 reaching positions. The order of the hand signs and positions was randomized. At the end of the eighth task, the participants were asked to take off the HMD with a text instruction.

4.3.1 Robot's Size. The robot's size influences the visual feedback of the swarm robots. Two robot size conditions were prepared: 20 and 30 mm diameter. The bigger size was set to 30 mm using the size of Zooids [34], an open-source swarm robot often used in HCI research, as a reference. The smaller size was set to 20 mm because it is close to the average adult human finger width [6].
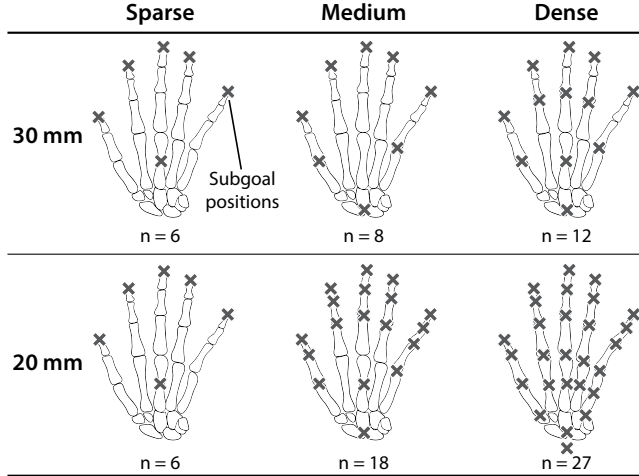
Figure 6: The predetermined subgoal positions relative to the hand bones. The gray crosses are the subgoal positions.

4.3.2 Robot's Density. The robot's density also affects the visual feedback of the swarm. As the density increases, the spatial resolution of the visual representation increases. However, this may complicate the representation and make control difficult as the number of robots increases. Three density levels were considered: *sparse*, *medium*, and *dense*.

We set the number of robots for each size and density condition based on preliminary testing. The results are summarized in Table 1. Through the preliminary testing of different densities, we concluded that at least six robots representing each finger and palm were required to represent a hand, regardless of the robot's size. Thus, we set the number of robots to six for the *sparse* condition for both the 20 mm and 30 mm robots. For the *dense* condition, the number of robots was set to twelve for 30 mm robots because that is the maximum number of 30 mm robots that fit in an average adult hand. To achieve the same density, the number of 20 mm robots was set

to 27. In the *medium* condition, the number of 20 mm robots was set to 18, which was the average of the numbers in the *sparse* and *dense* conditions. For 30 mm robots, the number for the *medium* condition was set to eight to achieve the same density as the 20 mm robots in the *medium* condition.

The subgoal positions relative to the hand bone need to be manually determined. Thus, we designed the subgoal position distributions shown in Figure 6 through the following step. For the *sparse* condition, we allocated subgoal positions to all the fingertips and the palm. For the denser conditions, we increased the number of subgoal positions on the palm and distributed them around the palm. If there are more subgoal positions, they were allocated to the proximal interphalangeal joints, to the metacarpophalangeal joints, and then, to the palm and the wrist.

4.3.3 Subgoal Formation Generation and Assignment Algorithm. As explained in subsection 3.1 and subsection 3.2, we examined three combinations of subgoal formation generation and assignment algorithms: *bone-static*, *bone-dynamic*, and *silhouette-dynamic*. The subgoal formation generation algorithm affects not only the visual feedback of the swarm but also the maneuverability of the subgoal formation. Users using the bone-based algorithm can predict subgoal formation's movements more precisely than those using the silhouette-based algorithm. This is because the silhouette-based algorithm continuously updates the correspondence between the hand and subgoal formation, and a certain hand movement does not necessarily result in the same subgoal formation's movements.

In addition, the assignment algorithm affects the robot's movements during hand shape transitions and hand movements. It may further influence the user's understanding of the correspondence between the hand and robot movements.

4.4 Measurements

We used a questionnaire to assess participants' sense of body ownership and agency. The questions related to the sense of body ownership and agency in the virtual embodiment questionnaire [55] were modified and used. The questionnaire consisted of eight items in two subsets of questions for the sense of body ownership and agency, as shown in Table 2. Each response was scored on a seven-point Likert scale (1 = strongly disagree; 7 = strongly agree). The scores for the sense of body ownership and agency were calculated by taking the average of the corresponding four questions as suggested by Roth *et al.* [55]. In addition, task load was measured using the NASA TLX questionnaire [14]. The pairwise comparisons of the factors were performed only after the first trial.

4.5 Procedure

Practice trials were conducted before the experiment to reduce learning effects. Alcohol disinfection was performed on the experimental apparatus and the hands of the experimenter and participant. Participants took a seat and were briefed on the experiment, procedures, data handling, risks, and rights and were instructed to sign a consent form if they agreed. The participants put on the HMD, and if they had no problems with the fit, they did two practice trials (one with the *30 mm*, *sparse*, and *bone-dynamic* settings and another with the *30 mm*, *dense*, and *bone-dynamic*). After the practice trials, the participants reviewed the questionnaires, and if they did

Table 2: Questionnaire used in the VR experiment.

Subscale	Questionnaire item
Body ownership	It felt like the swarm robot was my body. It felt like some of the robots were my fingers. It felt like the swarm robots belonged to me. The swarm robot felt like a human hand.
Agency	The movements of the swarm robot felt like they were my movements. I felt like I was controlling the movements of the swarm robot. I felt like I was causing the movements of the swarm robot. The movements of the swarm robot were in sync with my own movements.

not understand the question, an explanation was provided by the experimenter.

After the practice trials, the main experiment started. Participants wore the HMD while making the four hand signs in Figure 5 a random order, making each hand sign twice, once toward the left side of the table and once toward the right side of the table. Then, they removed the HMD and answered the questionnaires on an iPad. The participants repeated the task of making the hand signs and questionnaire responses a total of 18 times. To control for learning effects, the order of the experimental conditions (robot size, density, and subgoal position generation and assignment algorithm) was randomized. To control the interference effect of arm fatigue, the participants were asked to ensure that they were not fatigued before each task. A five-minute break was provided after the ninth task. After the 18th questionnaire response, the participants filled out a demographic questionnaire, and a semi-structured interview was conducted for approximately five to ten minutes. The entire experiment lasted approximately 90 min.
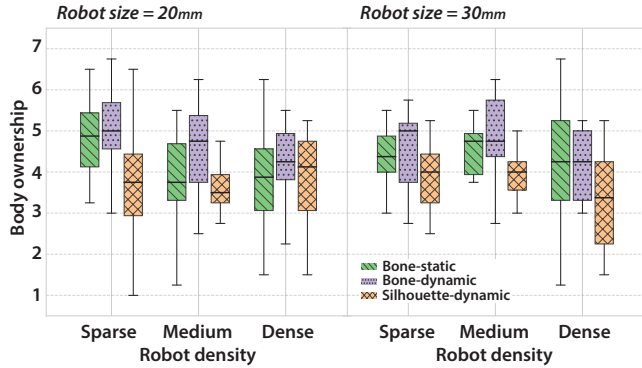
4.6   Results

Figure 7: Body ownership score obtained in the VR experiment for each size, density, and subgoal formation generation and assignment algorithm.

The body ownership score, agency score, and task load index were calculated for three factors: robot size, density, and subgoal formation generation and assignment algorithm. The results are
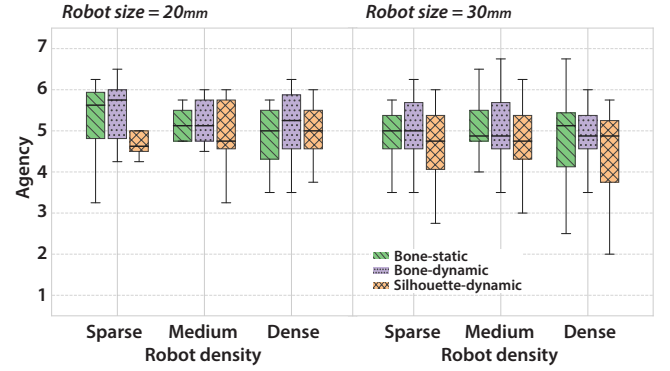

Figure 8: Agency score obtained in the VR experiment for each size, density, and subgoal formation generation and assignment algorithm.
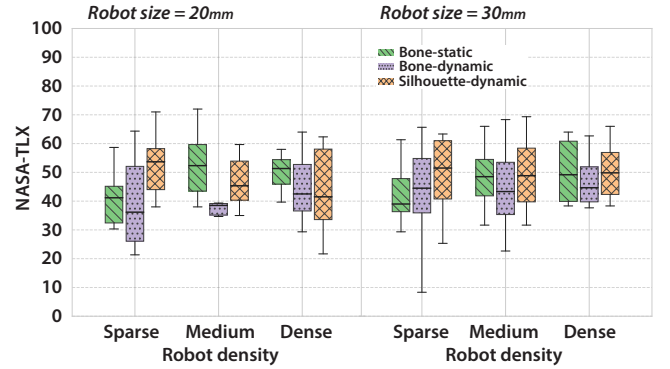

Figure 9: Task load index obtained in the VR experiment for each size, density, and subgoal formation generation and assignment algorithm.

shown in Figure 7, Figure 8, and Figure 9. As these were nonparametric data, we performed an Aligned Rank Transform (ART) [75] followed by a three-factor two-way repeated-measure ANOVA with Holm correction for each subscale (*i.e.*, body ownership, agency, and task load) to investigate the main effects and interactions.

4.6.1 Body Ownership. ANOVA revealed significant main effects of density ($F(2, 153) = 4.00$, $p = .020$, $\eta_p^2 = .05$) and subgoal formation generation and assignment algorithm on the body ownership score ($F(2, 153) = 15.28$, $p = .000$, $\eta_p^2 = .17$). There was a trend toward an interaction between size and density, but no significant interaction was found between any of the factors. Therefore, using the Holm-corrected ART-C [9], contrast tests were performed on the size, density, and subgoal formation generation, and assignment algorithm factors. The contrast test on the size factor revealed that there was no significant difference in body ownership score between *20 mm* and *30 mm* conditions. The *sparse* condition led to a significantly higher body ownership score than the *dense* condition ($p = .016$, cohen's $d = 0.516$). Finally, the *bone-dynamic* resulted in a significantly higher body ownership score than the other two (*bone-dynamic* and *silhouette-dynamic*: $p = .000$, cohen's $d = 1.002$, *bone-dynamic* and *bone-static*: $p = .033$, cohen's $d = 0.393$), and *bone-static* resulted in a significantly higher body ownership score than the *silhouette-dynamic* ($p = .002$, cohen's $d = 0.609$).

4.6.2 Agency. ANOVA revealed significant main effects of size ($F(1, 153) = 5.92$, $p = .016$, $\eta_p^2 = .04$) and subgoal formation generation and assignment algorithm on the agency score ($F(2, 153) = 5.07$, $p = .007$, $\eta_p^2 = .06$). No significant interactions were observed between any of the factors. Therefore, using the Holm-corrected ART-C, contrast tests were performed on size, density, and subgoal formation generation and assignment algorithm factors. The contrast test on the size factor showed that the *20 mm* condition led to a significantly higher agency score than the *30 mm* condition ($p = .016$, cohen's $d = 0.363$). In addition, the *bone-dynamic* algorithm resulted in a significantly higher agency score than the *silhouette-dynamic* algorithm ($p = .006$, cohen's $d = 0.576$). No significant differences in agency score were observed for the density factor.

4.6.3 Cognitive Load. The ANOVA revealed significant main effects of subgoal formation generation and assignment algorithm on cognitive load ($F(2, 153) = 6.11$, $p = .002$, $\eta_p^2 = .07$). However, a significant interaction between density and the algorithm was found ($F(4, 153) = 3.03$, $p = .019$, $\eta_p^2 = .07$). Therefore, multiple comparisons by ART-C (holm corrected) were performed. As a result, no significant differences were found among all groups, but there was a trend toward differences between *silhouette-dynamic + sparse* and *bone-static + medium* as well as *bone-static + medium* and *bone-static + sparse*. In addition, the difference in differences test for the density–algorithm interaction showed significant differences between the *silhouette-dynamic - bone-static* and *dense – sparse* ($p = .039$, cohen's $d = 1.278$) as well as the *silhouette-dynamic – bone-static* and *medium – sparse* ($p = .036$, cohen's $d = 1.307$).

4.6.4 Semi-Structured Interview. During the interview at the end of the experiment, a few common items were reported. Nine participants reported that the swarm robot felt like their hand at least once, and another participant reported that it felt like it was following their hand. This one participant noted that they often felt that the robot moved late relative to their hand, which led to the sensation that robots were following their hand. The other eight participants also noted that the swarm robots seemed slower under certain conditions and that increasing their speed would lead to a stronger embodiment.

Seven participants noted the importance of fingertips in feeling like the swarm robot as a hand. They reported that when the robot was positioned for each fingertip, they tended to recognize and move it as a hand. In addition, six stated that the hand-like appearance was lost when the robots collided or vibrated with each other, or when they moved, and when they did not fit well into the hand shape, coalesced around the palm, or were misaligned with the hand position.

All participants also mentioned the impact of the size of the robot. Two stated that the larger robot felt more like a hand or that they felt in control, while five stated that the smaller robots felt more like a hand or fingers.

Nine participants mentioned the influence of robot density. Eight stated that the lower density was more likely to cause the embodiment or lead to a higher maneuverability, but three of these stated that the higher density was felt as a hand when the robot's size was small, *i.e.*, 20 mm. Another participant felt like swarm robots were the hand regardless of density.

4.7 Discussion

In a VR psychophysical experiment, we studied how swarm robot factors like size, density, and algorithm impact embodiment. We focused on the sense of body ownership, sense of agency, and task load in the experimental results.

4.7.1 Embodiment Across Different Sizes of Swarm Robots. We first compared the embodiment scores for each size condition with the neutral level (*i.e.*, 4 point rating, which is a neutral response to the 7-point Likert scale questionnaire) to evaluate the level of embodiment. This neutral level was the null hypothesis of the tests. The body ownership scores were significantly higher than the neutral level for both *30 mm* ($p = .018$, cohen's $d = 0.413$) and *20 mm* ($p = .039$, cohen's $d = 0.351$), and the same was true for the agency scores (*30 mm*: $p = .000$, cohen's $d = 1.920$; *20 mm*: $p = .000$, cohen's $d = 2.172$). The cohen's d values for body ownership score indicate small to medium effect sizes, and those for agency score indicate large effect sizes. Thus, the questionnaire responses imply that the participants felt a higher level of embodiment for both *30 mm* and *20 mm* robots than the neutral state in which participants neither deny or affirm whether they feel the robots as their body parts or as if they are in control of the robots. Although these alone do not robustly show that Swarm Body was embodied by the participants for both conditions, they are consistent with the interview responses that many participants felt swarm robots like their hands at least once. Therefore, Swarm Body is likely to be embodied to some extent for both *30 mm* and *20 mm* robots with an appropriate algorithm and density.

When comparing the size conditions, no significant differences were found between the *20 mm* and *30 mm* conditions in the body ownership score and task load index, while the agency score was significantly higher for *20 mm* than for *30 mm* with a small to medium effect size (cohen's $d = 0.363$). A possible explanation for this is that larger perceived movements of robots when moving fingers led to a higher sense of agency. The movement of a robot relative to its size is larger for a smaller robot. Thus, the movements

of the 20 mm robots could be perceived as larger than those of the 30 mm robots. The reported preferences for robot size also varied in the interviews. Therefore, the impact of robot size on the embodiment and the resulting cognitive load is likely to vary among individuals, yet smaller robots tend to provide a higher sense of agency.

4.7.2 Sparse Swarm Robots Achieve A Higher Sense of Body Ownership. The body ownership score was significantly higher in the *sparse* than in the *dense* with a medium effect size (cohen's d = 0.516), suggesting that swarm robots with lower density is more likely to be felt as a hand. This is consistent with participants' reports that the lower density of the swarm robots felt more like hands. This may be due to the more frequent occurrence of collisions and vibratory movements of robots, their deviations from the hand shape, and getting stuck around the palm in the higher density conditions, which reduced the sense of embodiment as described by the participants. Another possible reason is that the participant could tell which robots represent fingertips better in the *sparse* condition; *i.e.*, it offers a better understanding of finger-robot correspondence with a simpler representation. This might result in a higher sense of body ownership as discussed in subsubsection 4.7.3.

The body ownership score for the *sparse* was also significantly higher than the neutral level (p = .004, cohen's d = 0.639). The agency scores were significantly higher than the neutral level in all the conditions (*dense*: p = .000, cohen's d = 1.676; *medium*: p = .000, cohen's d = 1.918; *sparse*: p = .000, cohen's d = 1.982), although there were no significant differences between the conditions. As stated in subsubsection 4.7.1, these comparisons with the neutral level do not guarantee the embodiment of Swarm Body; instead, they support that the level of embodiment is high in the *sparse*.

4.7.3 Bone-Dynamic Could Realize Higher Level of Embodiment. The sense of body ownership was found to be greatest for the *bone-dynamic*, followed by the *bone-static* and *silhouette-dynamic*. Similar to the sense of body ownership, the sense of agency showed a tendency that the *bone-dynamic* results in the highest level followed by the *bone-static*, and a significant difference between the *bone-dynamic* and *silhouette-dynamic* was found. Thus, it is suggested that the *bone-dynamic* algorithm results in the highest level of embodiment, while the *silhouette-dynamic* algorithm results in the lowest.

The higher level of embodiment in *bone*-based algorithms would come from the robot's ability to represent and respond to fingertip movements. In the interviews, participants reported that they were more likely to perceive the swarm robot as a hand when the robot responded to their fingertip movements. This suggests that the level of embodiment was improved when visual-motor synchronicity occurs even for local movements of the fingertips in addition to the whole hand movements. As the *bone*-based algorithm represents and responds to the participant's fingertip movements, it could enhance the level of embodiment through visual-motor synchronicity of fingertips.

In addition, similar to the *dense* condition, the lower level of embodiment in the static condition may be due to the collisions, and oscillatory movements between robots are more likely to occur with the static assignment. As such, the *bone-dynamic* is considered

to lead to the highest level of embodiment as it allows for visual-motor synchronization down to the fingertips while maintaining representation in various hand gestures and hand movements.

5 EMBODIMENT EXPERIMENT WITH ROBOTS

The VR experiment demonstrated that swarm robots could be embodied, offering insights into their ideal behavior. However, of course, there are differences between the VR environment and the real environment, and the use of real robots may have an effect on the results of the experiment. Therefore, we designed a similar embodiment experiment with fewer factors while considering the VR experiment results to examine whether similar embodiment characteristics can be observed in real-world settings.

5.1 Robot Implementation

We developed custom-made swarm robots to conduct an embodiment experiment in the real world. The robot design was inspired by Zooids [34], an open-source swarm robot platform, but the hardware and software were newly designed.
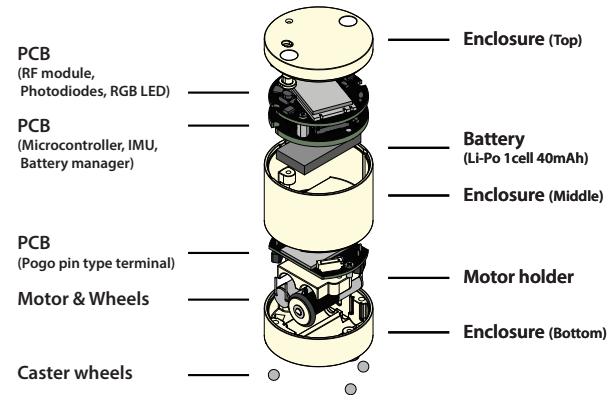
Figure 10: Exploded view of our custom-made swarm robot.

5.1.1 Hardware Design. The VR experiment results suggest that the embodiment of swarm robots occurs for both 20 mm and 30 mm robots and that they show similar embodiment characteristics. Also, owing to the limitations of currently available motor-based actuators and a communication module, the robot size cannot be reduced to 20 mm. Therefore, we assumed that 30 mm robots could be used to conduct the embodiment experiment and designed our robot with that size.

The hardware design is illustrated in Figure 10. The robot parts include a microcontroller unit (STM32G071KBU6[2] from STMicroelectronics), motor drivers (DRV8837DSGR[3] from Texas Instruments), RF module (RF2401F20[4] from NiceRF), motors with a 26:1 planetary gearbox (Pololu 2357[5]), photodiode (PD15-22C/TR8[6] from Everlight Electronics), and a 40mAh Li-Po battery.

---

[2]https://www.st.com/en/microcontrollers-microprocessors/stm32g071kb.html
[3]https://www.ti.com/product/DRV8837/part-details/DRV8837DSGR
[4]https://www.nicerf.com/item/nrf24l01-module-rf2401f20
[5]https://www.pololu.com/product/2357
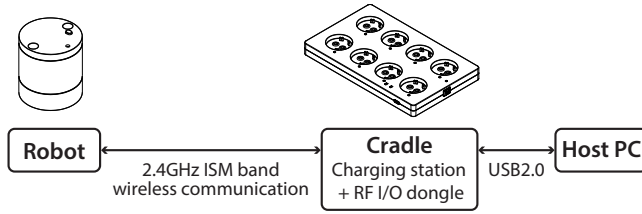[6]https://everlighteurope.com/ir-detectors/2374/PD1522CTR8.html

Figure 11: The robots communicate with the cradle through a 2.4 GHz ISM band wireless communication. The cradle communicates with the host PC with USB 2.0.

5.1.2 Communication and Projector-based Tracking System. The communication between the robots and the host computer is described in Figure 11. The methods are similar to the ones used for the Zooids. The robots and cradles are each equipped with an RF module and communicate through a 2.4GHz ISM band wireless communication. A projection-based localization system used for the Zooids was used to track the robots. A high-speed projector (DLP LightCrafter 4500[7] from Texas Instruments) was used to project a sequence of gray-coded patterns onto the table; the two photodiodes on the robot received the projected coded-pattern light, and the microcontroller of the robot decoded its pattern into position information. Then, the robot calculated its orientation from the positions of two photodiodes and broadcasted its position and orientation information to the host computer.

5.1.3 Simulation-Based Robot Control. The robots are controlled based on a simulation using the framework described in section 3. In particular, we employed the same simulation for the real-world experiment as was used in the VR experiment. Subgoal position generation and assignment are conducted based on tracked hand data, and the robot positions are given by an RVO simulation with nonholonomic constraints. Real robots are commanded to move to the current simulation robot positions every 100 ms. In this manner, the robots can obtain incremental subgoal positions along their paths. Then, the robot controls the rotation of the wheels according to the control law described in subsection A.1 to reach the subgoal position.

5.2 Participants

A total of 10 participants (4 males and 6 females; 28.89 ± 13.83 (SD) years old) participated in the experiment. Participants were recruited through a social media post. All the participants were unaware of the purpose of the experiment, had normal or corrected vision, and were right-handed. The participants signed a consent form regarding the experiment and were compensated with approximately $16 on Amazon gift cards. The ethics review board approved this study.

5.3 Apparatus and Setup

Figure 12 displays the experimental setup, which is similar to the VR experiment. In this study, a hand tracker (Leap Motion Controller 2[8] from Ultraleap) was used to track hands instead of Meta Quest 2

---

[7]https://www.ti.com/tool/DLPLCR4500EVM
[8]https://leap2.ultraleap.com/leap-motion-controller-2/

to create the system without an HMD. The hand tracker is located under the table, as shown in the Figure 12 (bottom).
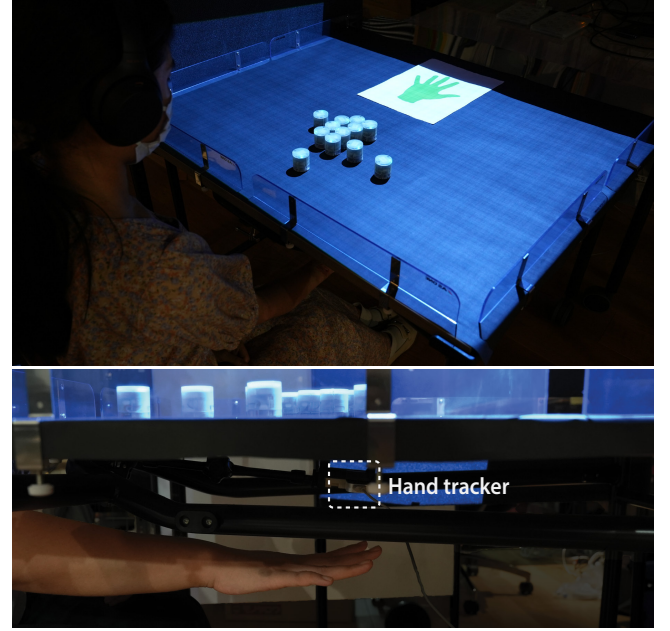

Figure 12: The experiment setup of the real-world robot study. The participant is wearing noise-canceling headphones and is controlling the swarm robots on the table with their right hand under the table (top). The hand is tracked with the hand tracker located under the table (bottom).

5.4 Experiment Design and Conditions

This experiment was similar to the VR experiment. However, the robot size was fixed at 30 mm, followed by a 3×3 factorial design. The independent variables examined were density (*sparse*, *medium*, and *dense*) and subgoal position generation and assignment algorithm (*bone-static*, *bone-dynamic*, and *silhouette-dynamic*). All variables were within the subject.

The task was similar to that in the VR experiment, which involved reaching a target with a specified hand shape. One difference was that in the VR experiment, the targets were positioned in the right and left fronts, but in the real-world experiment, the target position was limited to the front, and accordingly, the number of tasks per trial was halved to four (*i.e.*, four hand signs). This is because the swarm robots' battery capacity would not hold a charge until the end of the experiment, with eight tasks per trial.

In addition, instead of the targets automatically (dis)appearing on the desk in the VR, the experimenter manually placed and removed the targets printed on paper on the desk. To avoid the potential slipping of the robots on the target paper, the task was changed from moving the swarm robots to fit in the target to moving the swarm robots to a specified position in front of the target and making a specific hand shape.

5.5 Measurements

The same subjective measurements as those used in the VR experiment (*i.e.*, the modified embodiment questionnaire and NASA TLX) were used to evaluate the sense of body ownership, sense of agency, and cognitive load.

5.6 Procedure

This procedure is similar to that used in the VR experiments. Practice trials were conducted before the experiment to reduce the learning effects. After signing the consent form, participants started the practice trials (one under *sparse* and *bone-dynamic* and another under *dense* and *bone-dynamic*).

During the main experiment, the participants wore headphones with white noise and put their hands under the table to control the swarm robots. The participants were instructed to move the swarm robots representing their right hand under the table to the starting area at the beginning of each task. When all the robots returned to the starting area, the experimenter specified a hand sign by posting handshapes on paper on the desk. The participants were instructed to move their swarm robot hand forward with a specified hand sign. The hand sign sheet was removed after five seconds, and the participants moved their hands back to the starting area. Participants repeated the reaching task and completed the questionnaire nine times.

To control the interference effect of arm fatigue, the participants were asked to ensure that they were not fatigued prior to each task. A five-minute break was provided after the fifth task. After the ninth questionnaire response, the participants answered a demographic questionnaire, and a semi-structured interview was conducted for approximately five to ten minutes. The entire experiment took approximately one hour.
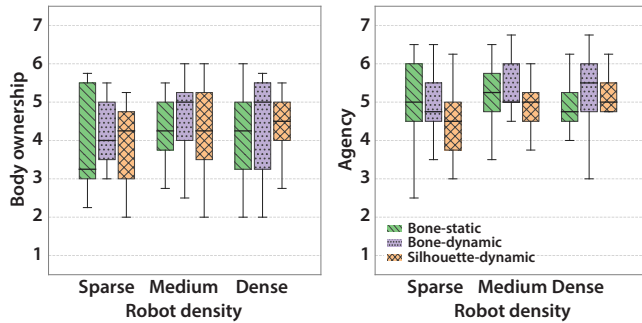
5.7 Results

Figure 13: The body ownership and agency scores obtained in the robot experiment for each density and subgoal position generation and assignment algorithm.

The body ownership score, agency score, and task load index were calculated for each combination of robot density and subgoal position generation and assignment algorithm. The hand tracker did not work properly and frequently lost track during the experiment of one of the participants, owing to the reflection of its own infrared light at the bottom of the table. As this might have strongly affected
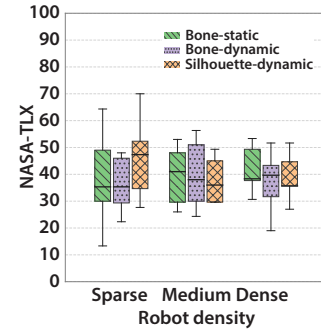

Figure 14: The NASA Task Load Index obtained in the robot experiment for each density and subgoal position generation and assignment algorithm.

their data, we excluded them from the results and analysis. The results are shown in Figure 13 and Figure 14. Similar to the analysis of the VR experiment results, we performed an ART procedure followed by a three-factor, two-way repeated-measures ANOVA with holm correction for the body ownership score, agency score, and task load index.

5.7.1 Body Ownership. ANOVA revealed a significant main effect of density ($F(2, 64) = 3.31$, $p = .043$, $\eta_p^2 = .09$). No significant interaction was found between the density and algorithm factors. Therefore, contrast tests were performed on the density and algorithm factors using the holm-corrected ART-C. The contrast test on the algorithm factor showed no significant differences in body ownership score between any of the conditions. The test on the density factor showed that the *medium* condition led to a significantly higher body ownership score than the *sparse* condition ($p = .047$, cohen's $d = 0.676$).

5.7.2 Agency. ANOVA revealed a significant main effect of density ($F(2, 64) = 3.55$, $p = .034$, $\eta_p^2 = .10$) and a moderate main effect of algorithm ($F(2, 64) = 2.89$, $p = .063$, $\eta_p^2 = .08$). No significant interaction was found between the density and algorithm factors. Therefore, contrast tests were performed on these factors using the holm-corrected ART-C. The contrast test on the algorithm factor showed no significant differences in agency score between any of the conditions, but there was a trend that the *bone-dynamic* condition led to a higher agency score than the *silhouette-dynamic* condition ($p = .068$, cohen's $d = 0.635$). The contrast test on the density factor also showed no significant differences in agency score between any of the conditions, but there were trends that *medium* condition led to a higher agency score than the *sparse* condition ($p = .052$, cohen's $d = 0.665$) and that the *dense* condition led to a higher agency score than the *sparse* condition ($p = .072$, cohen's $d = 0.583$).

5.7.3 Cognitive Load. ANOVA revealed no significant main effects of density and subgoal position generation and assignment algorithm. No significant interactions were observed. Therefore, contrast tests were performed on the density and algorithm factors using the holm-corrected ART-C. The contrast tests on these factors revealed no significant difference in the task load index

between *sparse*, *medium* and *dense* conditions, as well as between the *bone-static*, *bone-dynamic*, and *silhouette-dynamic* conditions.

5.7.4 Semi-Structured Interview. A few common items were identified during the interview at the end of the experiment. Four participants reported that the swarm robots felt like their hands in some trials, while the other five participants, including the participant under the poor hand tracking condition, reported that the robots were following their hands rather than being their own hands. One of the five participants noted that it was difficult to distinguish between the hand signs, making it feel like a mass following the hand. The other participant reported that the robots sometimes felt like their hand, but they were not fully convinced that it was their hand. Three participants mentioned that the control was intuitive when the robot was positioned for each fingertip.

Eight participants mentioned the influence of density on the embodiment. Six stated that the dense robots felt more like the hand or something they could control better; one preferred medium and sparse, and one preferred sparse. One of those who preferred the dense condition stated that the dense swarm presented a sense of oneness and coherence. Another participant reported that the swarm robots were embodied regardless of their density.

5.8 Discussion

To investigate the influence of swarm robot-unique factors on embodiment in real-world settings, we conducted a psychophysical experiment with the real swarm robots we developed, analyzing three aspects: sense of body ownership, sense of agency, and cognitive load. The results were also compared with those of the VR embodiment experiment in section 4, discussing the unique characteristics of the real-world system's effect on swarm robot embodiment.

5.8.1 Embodiment of Swarm Body with Bone-Dynamic Algorithm. No significant difference was found between any of the algorithm conditions. When compared with the neutral level (4 point rating in the 7-point Likert scale for body ownership and agency; 50 point rating in the 100-point NASA TLX for cognitive load), the *bone-dynamic* condition resulted in significantly higher body ownership and agency scores, and a significantly lower cognitive load score than the neutral levels with large effect sizes (body ownership: $p = .042$, cohen's $d = 0.751$; agency: $p = .000$, cohen's $d = 2.354$; cognitive load: $p = .000$, cohen's $d = 1.722$, where the neutral levels were the null hypotheses). Thus, the questionnaire responses suggest that the participants felt a higher sense of embodiment for the *bone-dynamic* algorithm than the neutral state in which participants neither deny or affirm whether they feel the robots as their body parts or as if they are in control of the robots. This is consistent with the VR study finding that the *bone-dynamic* condition achieved a higher level of embodiment than neutral levels. This is also consistent with the interview responses from three participants that they could control swarm robots intuitively when the robots were positioned for each fingertip (*i.e.*, bone-based algorithms). The interview responses further suggest that some of the participants felt swarm robots as they would their hands in some trials though we cannot tell which algorithm conditions they talk about. Overall, these questionnaire and interview results indicate that Swarm Body

was possibly embodied in some trials, and if so, the embodiment probably occurred with the *bone-dynamic* algorithm.

5.8.2 Shift in Preference toward Denser Swarm. When comparing the size conditions, the *medium* density condition resulted in a significantly higher body ownership score than the *sparse* condition. The *medium* and *dense* conditions tended to result in a higher agency score than the *sparse* condition. These results are consistent with the interview results that six out of eight participants who mentioned the influence of density preferred denser conditions. However, these results seem to be inconsistent with the VR experiment's result that the *sparse* condition led to a higher sense of body ownership than the *dense* condition. There are three possible causes for this shift in preference for denser conditions in the real world: the increased importance of visual similarity, sense of accomplishment, and collisions.

First, visual similarity may be more important than understanding the correspondence between the body parts and robots in the embodiment in real-world settings. As discussed in subsubsection 4.7.2 and subsubsection 4.7.3, sparse swarm robots with bone-based algorithms achieved a higher level of embodiment by offering a better understanding of the finger-robot correspondence through a simpler hand representation. However, as suggested in [3], the visual similarity of an object to a hand affects its level of embodiment. This effect might become more dominant in real-world settings. This hypothesis was supported by the participants' comments that the dense swarm had a greater sense of oneness and coherence and that the denser the robot, the more it felt like a hand shape.

Second, the sense of agency may be influenced by the expected amount of effort required to move the object to be embodied. Four participants reported a stronger feeling of controlling the robots and a stronger sense of accomplishment in the *dense* condition.

Third, real-world systems have more collisions between robots, which might have reduced the level of embodiment in the *sparse* condition. As discussed in subsubsection 4.7.2, the lack of collisions likely contributed to the high level of embodiment in the *sparse* condition in the VR study. Real-world systems cause more collisions owing to the robot position errors between the simulation and the real world. As a result, our real-world system had collisions even in the *sparse* condition, which potentially decreased the level of embodiment. It is also important to note that some robots turned off when they were stuck during the real-world study. This is because our control program turns off a robot when the torque applied on its motors exceeds a certain threshold to protect the motors. When that happened, the experimenter quickly turned them on, but this could affect the study results.

Thus, the participants' preferences shifted toward denser swarm robots, and the levels of embodiment in the *dense* and *medium* conditions were relatively higher in the real-world study. And the shift was possibly caused by the increased importance of visual similarity, a sense of accomplishment, and collisions. This finding is somewhat limited by the fact that the preference shift might come from increased collisions which is partly dependent on our implementation. However, as most real-world swarm robots have more collisions than their simulations, our finding is valuable for designing real-world embodied swarm robots though the degree

of this shift may vary. In addition, several consistent trends were observed throughout the VR and real-world experiments. For example, higher embodiment levels were reported when the robots were located at the fingertips, and a lower cognitive load was measured in the *sparse* condition.

6   APPLICATIONS

Swarm Body expands the design space of tangible and embodied interaction, offering unique characteristics such as robustness, flexibility, and scalability to the human body. The components can move without geometrical constraints other than collision with each other. Additionally, although the current implementation requires a projector, the robot itself is not anchored to a specific environment; thus, our system is versatile and can be used in a variety of locations, including ordinary desks.

The main application of Swarm Body is physical telepresence, where embodied swarm robots facilitate physical interaction with remote people and environments, as shown in Figure 1 (right). The operator can control Swarm Body projected on their table as if manipulating their own hands. The other person can physically interact with the operator through Swarm Body.

Our physical telepresence system is inspired by Physical Telepresence Workspace by Leithinger *et al.* [37], particularly in the physical representation of the user's hand using hand sensing and spatially-aligned visual feedback. Our work extends their interaction capabilities through swarm robots characteristics, such as swarm splits, mergers, transfers, and obstacle avoidances. Below, we outline scenarios that showcase new interaction opportunities in physical telepresence enabled by the characteristics of swarm robots.

6.1   Multipliable Body

Swarm Body can split from a single swarm into multiple swarms, each representing different body parts, and then merge back into a single swarm. These splits and mergers allow the user to adjust the number of independent swarms and the number of robots constituting each swarm as shown in Figure 15. This enables the user to seamlessly switch between one- and two-handed telepresence in a single interface. For example, when organizing a desktop remotely, the user can employ both hands to efficiently collect objects and then marge the robots to the dominant hand for precise organization. The user can duplicate one hand into two, enabling the performance of two similar tasks simultaneously through parallel embodiment as in [68]. Additionally, unlike existing embodied robots that require one system for each user, Swarm Body supports multiple users manipulating the robots through a single interface. For example, while one remote user engages in an activity such as rolling a ball with a local person, another remote user can allocate half of the robots to form a new hand and participate.

6.2   Form-Giving to Transformable Body

The malleability of Swarm Body enables the transformation of the swarm into various body parts of different sizes and unconventional forms (*e.g.*, a small hand, elongated fingers, and a tentacle-like limb) (Figure 16). This transformation provides enhanced interaction freedom while preserving embodiment features, such as intuitive
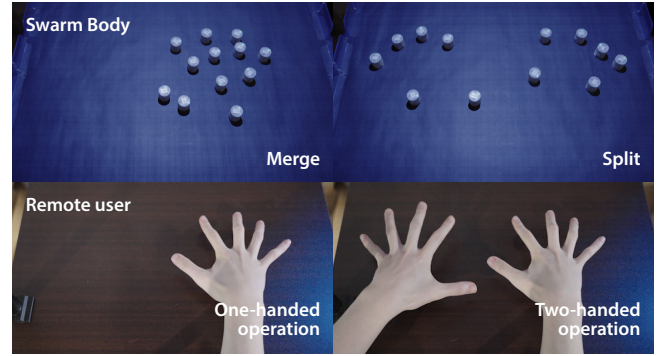

Figure 15: The remote user initially uses all the robots with their right hand (left). Then, they split them into two to use both of their hands (right).

swarm manipulation. For example, one can experience the affordance of objects from a child's perspective by interacting with them using Swarm Body, which simulates a smaller hand, as in [47]. Swarm Body can extend its fingers or transform them into tentacle shapes to reach and grasp objects at a distance or in narrow gaps. Similar to how pixels on a screen represent a range of embodied avatars, Swarm Body physically embodies diverse avatars through its ability to transform.
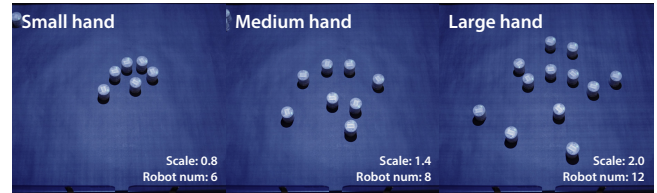

Figure 16: Swarm Body represents hands at various scales, allowing the user to experience different affordances of the environment.

6.3   Adaptability to the Environment

Swarm Body adapts to its environment by avoiding obstacles, resizing, or transforming itself as needed. As shown in Figure 17, our control method introduced in section 3 enables the swarm robots to not only mimic the body movements but also avoid obstacles. In a tabletop environment with obstacles, the user can interact with the environment without the need to intentionally avoid the obstacles. For example, when picking up a pen on the other side of a mug, the user can reach for it as if the mug did not exist. We believe that Swarm Body could develop into a system that seamlessly adapts to various settings, easing interactions even in cluttered spaces giving it the potential to eliminate physical barriers in interactions, enabling smoother engagements than what our own bodies can typically achieve.

6.4   Emotional Haptic Notification

Swarm Body allows the user to exert horizontal forces on remote objects or individuals in an embodied manner. This further expands
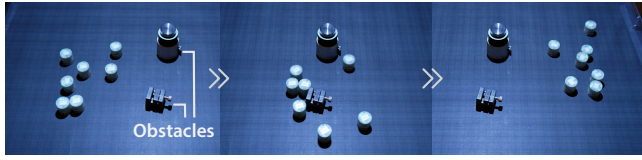
Figure 17: Swarm Body moves from left to right while passing through obstacles.

Figure 19: Swarm Body conveying the user's gestures. The user pointed at the black pen (left), then is now pointing at the pink pen (right).

the design space of physical telepresence with vertical actuation previously explored by Leithinger *et al.* [37]. Its embodiment aspect can also introduce intuitive and affective interactions, taking advantage of swarm characteristics in swarm user interface (SUI). Swarm Body achieves haptic communication by touching people Figure 18. Specifically, they can naturally get a person's attention or express emotions to an intimate partner with various forces and touch. For example, by gently tapping a person's arm engrossed in desk work, the robot can capture the person's attention and initiate communication. Although previous studies have explored haptic feedback using swarm robots [26], our system enables haptic feedback with the embodiment of the user. Thus, Swarm Body has the potential to haptically mediate the emotions and intentions of the user to the notified person.

Figure 18: The Swarm Body user lightly taps the person with the finger to casually draw the person's attention.

6.5 Gesture Presentation

Another interaction modality of Swarm Body is vision. When embodied as hands, swarm robots can communicate through gestures with a physical presence in remote environments, as illustrated in Figure 19. A remote individual can utilize physical gestures during their online presentations to boost engagement. In a remote collaboration scenario, a remote user can point to specific objects or convey simple reactions. For example, a remote craft instructor can point to the tools the students need to use at each step, direct their hand movements, and send a physical thumbs-up reaction upon task completion.

7 LIMITATIONS AND FUTURE WORK

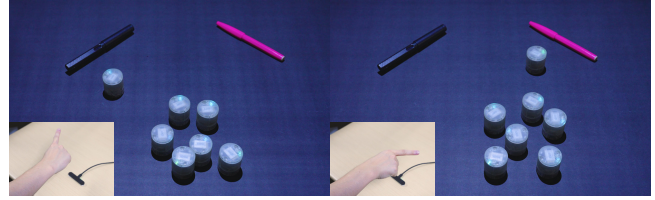Our study did not comprehensively investigate the embodiment characteristics of swarm robots and their applications.

7.1 More Extensive Investigation on Design Parameters

Our investigation focused on a subset of the factor levels that should be examined to understand the embodiment of swarm robots on the hand. Therefore, further studies should explore other possible factor levels. For example, it is possible to investigate the embodiment characteristics of robots smaller than 20 mm in size or under even denser conditions than the *dense* condition in the current work. Studies in VR on these conditions may show how much embodiment is possible in theory. The obtained level of embodiment for this theoretical condition will be a baseline when evaluating real-world systems.

We did not examine some design parameters and complex conditions to make the experiments feasible. Future research should explore various parameters including the robot's latency, speed, acceleration, color, and shape. Additionally, dynamic changes of the parameters seem to be promising approaches. For example, applying the bone-based algorithm to the fingertips and a silhouette-based algorithm to the other parts of the hand could integrate the advantages highlighted in our discussion. Also, robots in a swarm can have different sizes, shapes, densities, and functions and change them dynamically as Li *et al.* demonstrated [39]. Therefore, further investigation on the embodiment characteristics and applications of such swarm robots and algorithms is expected.

7.2 Recognition as a Body Part by an External Observer

While we revealed the embodiment characteristics of swarm robots for the operator, we did not investigate whether an external observer could recognized the robots as someone's body parts. In our preliminary testing, an observer was able to differentiate hand signs, although they were aware that the robots were representing a hand in advance. Since visual feedback is the only information source for an observer, a silhouette-based subgoal generation algorithm, which reduced the embodiment level for the operator in our study, might enhance their recognition of the robots as a hand. Further study on how an external observer recognizes the swarm robots is expected.

7.3 Beyond Tabletop Robots and Embodiment of the Hand

Swarm robots moving in 3D space, such as swarm drones, should be explored. Although some of our methods and findings will be

applicable to them, swarm robots moving in a 3D space have unique design parameters (*e.g.*, subgoal positions on the skin surface only vs. those on the entire volume, including the interior of the hand). This exploration also provides insight into how dimensional matching of the user's hand movements to the robot's subgoal formation affects a level of embodiment. If 3D formations give a higher level of embodiment, then restricting the user's hand movements to the 2D plane (*i.e.*, avoiding supinations and pronations) may also improve the level of embodiment in our 2D system. In addition, their 3D formation and movements will open up a further interaction design space. For example, embodied 3D swarm robots could pick up objects or enter space that is not accessible for wheeled swarm robots. Note that the projection-based robot localization method used in our study could be adapted to estimate the position and orientation of such swarm robots in 3D space [17]. These can be determined by solving the Perspective-n-Point problem, which involves considering the sensors mounted on the robots as points.

The lack of investigation on body parts besides hands also limits this study. We focused on hands as they are the most commonly used body parts for interaction with the environment. Although we revealed the embodiment characteristics of swarm robots for the hand, it remains unclear if the same tendencies apply to other body parts. We believe that our findings in the hand can be used to verify future studies on other body parts. Moreover, to expand the design space of Swarm Body, future research focusing on the effects of embodiment in body parts that differ from the user's actual body size and shape could be highly beneficial. This would open up new capabilities and applications for Swarm Body, allowing the user to manipulate objects at micro or macro scales with larger or smaller hands or using different shapes such as tentacle-like limbs.

7.4 Exploration of the Applications

Although we have presented several application scenarios, their effectiveness has yet to be evaluated. In future work, we plan to showcase our applications through interactive demonstrations and videos, thereby collecting direct feedback from participants. We also expect that this feedback will allow us to discover new application possibilities of Swarm Body beyond our initial scope.

7.5 Swarm Control Algorithm Dedicated for Embodied Behavior

Lastly, collisions and misalignments between our robots might have affected the participants' evaluation of embodiment. During the embodiment experiments in both VR and the real world, the experimenter sometimes observed collisions and misalignments of the robots. These undesirable behaviors potentially come from conditions, such as a large robot size or a high density, that make it difficult for swarm robots to avoid collisions while following a hand. As mentioned in subsubsection 5.8.2, our control program shuts down the robot in the case of excessive torque to protect the motors. This might affect the participants' sense of embodiment, leading to a different result from the VR study, even though the experimenter turned them on immediately. In other words, the participants' preference toward the denser swarm robots might be the characteristics of our real-world implementation not our approach. However, our real-world experiment and the interview responses

provided some consistent and generalizable insights, such as a higher level embodiment coming from fingertip representations and a lower cognitive load in sparse conditions. To address the collision issues, a specialized swarm control algorithm tailored for embodied behavior could potentially reduce collisions and enhance the embodied experience. Further psychophysical embodiment experiments with a control algorithm with fewer collisions will help us understand whether the differences in our VR and real-world study results are due to the approach or our implementation. It will also deepen our understanding on the embodiment characteristics we observed in both VR and real-world settings (*i.e.*, a higher level of embodiment coming from fingertip representations and a lower cognitive load in sparse conditions).

8 CONCLUSION

We proposed a new embodied system concept, *embodied swarm robots*, a group of robots collectively acting as a human body part. We presented a framework for the embodiment of swarm robots, investigating their characteristics in both VR and real-world environments. Our results offer two key insights into the embodiment of swarm robots.

(1) Swarm robots are likely to be embodied in VR and real-world scenarios using a suitable algorithm and density though there are some individual differences.

(2) The choice of swarm body control algorithm influences the level of embodiment, impacting both the visual-motor synchronicity of fingers and the frequency of robot collisions.

Additionally, we explored applications of our system, demonstrating how embodied swarm robots can enrich tangible and embodied interactions between humans and the environment.

REFERENCES

[1] Javier Alonso-Mora, Andreas Breitenmoser, Martin Rufli, Roland Siegwart, and Paul Beardsley. 2011. Multi-Robot System for Artistic Pattern Formation. In *2011 IEEE International Conference on Robotics and Automation*. 4512–4517. https://doi.org/10.1109/ICRA.2011.5980269

[2] Ken Arai, Hiroto Saito, Masaaki Fukuoka, Sachiyo Ueda, Maki Sugimoto, Michiteru Kitazaki, and Masahiko Inami. 2022. Embodiment of Supernumerary Robotic Limbs in Virtual Reality. *Scientific Reports* 12, 1 (June 2022), 9769. https://doi.org/10.1038/s41598-022-13981-w

[3] Ferran Argelaguet, Ludovic Hoyet, Michael Trico, and Anatole Lecuyer. 2016. The Role of Interaction in Virtual Embodiment: Effects of the Virtual Hand Representation. In *2016 IEEE Virtual Reality (VR)*. 3–10. https://doi.org/10.1109/VR.2016.7504682

[4] Laura Aymerich-Franch, Damien Petit, Gowrishankar Ganesh, and Abderrahmane Kheddar. 2017. Non-Human Looking Robot Arms Induce Illusion of Embodiment. *International Journal of Social Robotics* 9, 4 (Sept. 2017), 479–490. https://doi.org/10.1007/s12369-017-0397-8

[5] Matthew Botvinick and Jonathan Cohen. 1998. Rubber Hands 'Feel' Touch That Eyes See. *Nature* 391, 6669 (Feb. 1998), 756–756. https://doi.org/10.1038/35784

[6] Bryan Buchholz and Thomas J. Armstrong. 1991. An Ellipsoidal Representation of Human Hand Anthropometry. *Human Factors* 33, 4 (Aug. 1991), 429–441. https://doi.org/10.1177/001872089103300405

[7] Rainer Burkard, Mauro Dell'Amico, and Silvano Martello. 2012. *Assignment Problems: Revised Reprint.* Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611972238

[8] Artem Dementyev, Hsin-Liu (Cindy) Kao, Inrak Choi, Deborah Ajilo, Maggie Xu, Joseph A. Paradiso, Chris Schmandt, and Sean Follmer. 2016. Rovables: Miniature On-Body Robots as Mobile Wearables. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16).* Association for Computing Machinery, New York, NY, USA, 111–120. https://doi.org/10.1145/2984511.2984531

[9] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21).* Association for Computing Machinery, New York, NY, USA, 754–768. https://doi.org/10.1145/3472749.3474784

[10] Phil Frei, Victor Su, Bakhtiar Mikhak, and Hiroshi Ishii. 2000. Curlybot: Designing a New Class of Computational Toys. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00).* Association for Computing Machinery, New York, NY, USA, 129–136. https://doi.org/10.1145/332040.332416

[11] Rebecca Fribourg, Nami Ogawa, Ludovic Hoyet, Ferran Argelaguet, Takuji Narumi, Michitaka Hirose, and Anatole Lécuyer. 2021. Virtual Co-Embodiment: Evaluation of the Sense of Agency While Sharing the Control of a Virtual Body Among Two Individuals. *IEEE Transactions on Visualization and Computer Graphics* 27, 10 (Oct. 2021), 4023–4038. https://doi.org/10.1109/TVCG.2020.2999197

[12] Shaun Gallagher. 2000. Philosophical Conceptions of the Self: Implications for Cognitive Science. *Trends in Cognitive Sciences* 4, 1 (Jan. 2000), 14–21. https://doi.org/10.1016/S1364-6613(99)01417-5

[13] Takayoshi Hagiwara, Gowrishankar Ganesh, Maki Sugimoto, Masahiko Inami, and Michiteru Kitazaki. 2020. Individuals Prioritize the Reach Straightness and Hand Jerk of a Shared Avatar over Their Own. *iScience* 23, 12 (Dec. 2020), 101732. https://doi.org/10.1016/j.isci.2020.101732

[14] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology,* Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[15] Takefumi Hiraki, Shogo Fukushima, Yoshihiro Kawahara, and Takeshi Naemura. 2018. Phygital Field: An Integrated Field with Physical Robots and Digital Images Using Projection-Based Localization and Control Method. *SICE Journal of Control, Measurement, and System Integration* 11, 4 (2018), 302–311. https://doi.org/10.9746/jcmsi.11.302

[16] Takefumi Hiraki, Shogo Fukushima, Yoshihiro Kawahara, and Takeshi Naemura. 2019. NavigaTorch: Projection-Based Robot Control Interface Using High-Speed Handheld Projector. In *SIGGRAPH Asia 2019 Emerging Technologies* (Brisbane, QLD, Australia) *(SA '19).* Association for Computing Machinery, New York, NY, USA, 31–33. https://doi.org/10.1145/3355049.3360538

[17] Takefumi Hiraki, Yoshihiro Kawahara, and Takeshi Naemura. 2018. Projection-based localization and control method of robot swarms for swarm user interfaces. In *Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* 584–589. https://doi.org/10.1145/3267305.3267307

[18] The Vu Huynh, Robin Bekrater-Bodmann, Jakob Fröhner, Joachim Vogt, and Philipp Beckerle. 2019. Robotic Hand Illusion with Tactile Feedback: Unravelling the Relative Contribution of Visuotactile and Visuomotor Input to the Representation of Body Parts in Space. *PLOS ONE* 14, 1 (Jan. 2019), e0210058. https://doi.org/10.1371/journal.pone.0210058

[19] Keiichi Ihara, Mehrad Faridan, Ayumi Ichikawa, Ikkaku Kawaguchi, and Ryo Suzuki. 2023. HoloBots: Augmenting Holographic Telepresence with Mobile Robots for Tangible Remote Collaboration in Mixed Reality. https://doi.org/10.1145/3544548.3606727 arXiv:2307.16114 [cs]

[20] Ryota Ito, Nami Ogawa, Takuji Narumi, and Michitaka Hirose. 2019. Do We Have to Look at the Mirror All the Time? Effect of Partial Visuomotor Feedback on Body Ownership of a Virtual Human Tail. In *ACM Symposium on Applied Perception 2019 (SAP '19).* Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3343036.3343139

[21] Hiroki Kaimoto, Kyzyl Monteiro, Mehrad Faridan, Jiatong Li, Samin Farajian, Yasuaki Kakehi, Ken Nakagaki, and Ryo Suzuki. 2022. Sketched Reality: Sketching Bi-Directional Interactions Between Virtual and Physical Worlds with AR and Actuated Tangible UI. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22).* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3526113.3545626

[22] Andreas Kalckert and H. Henrik Ehrsson. 2014. The Spatial Distance Rule in the Moving and Classical Rubber Hand Illusions. *Consciousness and Cognition* 30 (Nov. 2014), 118–132. https://doi.org/10.1016/j.concog.2014.08.022

[23] Florian Kennel-Maushart, Roi Poranne, and Stelian Coros. 2023. Interacting with Multi-Robot Systems via Mixed Reality. In *2023 IEEE International Conference on Robotics and Automation (ICRA).* 11633–11639. https://doi.org/10.1109/ICRA48891.2023.10161412

[24] Paulina Kieliba, Danielle Clode, Roni O. Maimon-Mor, and Tamar R. Makin. 2021. Robotic Hand Augmentation Drives Changes in Neural Body Representation. *Science Robotics* 6, 54 (May 2021), eabd7935. https://doi.org/10.1126/scirobotics.abd7935

[25] Lawrence H. Kim, Daniel S. Drew, Veronika Domova, and Sean Follmer. 2020. User-Defined Swarm Robot Control. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376814

[26] Lawrence H. Kim and Sean Follmer. 2019. SwarmHaptics: Haptic Display with Swarm Robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19).* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300918

[27] Ryota Kondo, Maki Sugimoto, Masahiko Inami, and Michiteru Kitazaki. 2019. Scrambled Body: A Method to Compare Full Body Illusion and Illusory Body Ownership of Body Parts. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* 1028–1029. https://doi.org/10.1109/VR.2019.8798346

[28] Ryota Kondo, Maki Sugimoto, Kouta Minamizawa, Takayuki Hoshi, Masahiko Inami, and Michiteru Kitazaki. 2018. Illusory Body Ownership of an Invisible Body Interpolated between Virtual Hands and Feet via Visual-Motor Synchronicity. *Scientific Reports* 8, 1 (May 2018), 7541. https://doi.org/10.1038/s41598-018-25951-2

[29] Ryota Kondo, Maki Sugimoto, Kouta Minamizawa, Masahiko Inami, Michiteru Kitazaki, and Yamato Tani. 2018. Illusory Body Ownership Between Different Body Parts: Synchronization of Right Thumb and Right Arm. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* 611–612. https://doi.org/10.1109/VR.2018.8446282

[30] Ryota Kondo, Yamato Tani, Maki Sugimoto, Kouta Minamizawa, Masahiko Inami, and Michiteru Kitazaki. 2020. Re-Association of Body Parts: Illusory Ownership of a Virtual Arm Associated With the Contralateral Real Finger by Visuo-Motor Synchrony. *Frontiers in Robotics and AI* 7 (2020).

[31] Ryota Kondo, Sachiyo Ueda, Maki Sugimoto, Kouta Minamizawa, Masahiko Inami, and Michiteru Kitazaki. 2018. *Invisible Long Arm Illusion: Illusory Body Ownership by Synchronous Movement of Hands and Feet.* The Eurographics Association. https://doi.org/10.2312/egve20181310

[32] H. W. Kuhn. 1955. The Hungarian Method for the Assignment Problem. *Naval Research Logistics Quarterly* 2, 1-2 (1955), 83–97. https://doi.org/10.1002/nav.3800020109

[33] So Kuroki, Tatsuya Matsushima, Jumpei Arima, Hiroki Furuta, Yutaka Matsuo, Shixiang Shane Gu, and Yujin Tang. 2023. Collective Intelligence for 2D Push Manipulations With Mobile Robots. *IEEE Robotics and Automation Letters* 8, 5 (May 2023), 2820–2827. https://doi.org/10.1109/LRA.2023.3261751

[34] Mathieu Le Goc, Lawrence H. Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, and Sean Follmer. 2016. Zooids: Building Blocks for Swarm User Interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16).* Association for Computing Machinery, New York, NY, USA, 97–109. https://doi.org/10.1145/2984511.2984547

[35] Yujin Lee, Myeongseong Kim, and Hyunjung Kim. 2020. Rolling Pixels: Robotic Steinmetz Solids for Creating Physical Animations. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '20).* Association for Computing Machinery, New York, NY, USA, 557–564. https://doi.org/10.1145/3374920.3374987

[36] Sang-won Leigh and Pattie Maes. 2016. Body Integrated Programmable Joints Interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* Association for Computing Machinery, New York, NY, USA, 6053–6057. https://doi.org/10.1145/2858036.2858538

[37] Daniel Leithinger, Sean Follmer, Alex Olwal, and Hiroshi Ishii. 2014. Physical Telepresence: Shape Capture and Display for Embodied, Computer-Mediated Remote Collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14).* Association for Computing Machinery, New York, NY, USA, 461–470. https://doi.org/10.1145/2642918.2647377

[38] Jiatong Li, Ryo Suzuki, and Ken Nakagaki. 2023. Physica: Interactive Tangible Physics Simulation Based on Tabletop Mobile Robots Towards Explorable Physics Education. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23).* Association for Computing Machinery, New York, NY, USA, 1485–1499. https://doi.org/10.1145/3563657.3596037

[39] Shuguang Li, Richa Batra, David Brown, Hyun-Dong Chang, Nikhil Ranganathan, Chuck Hoberman, Daniela Rus, and Hod Lipson. 2019. Particle Robotics Based on Statistical Mechanics of Loosely Coupled Components. *Nature* 567, 7748 (March 2019), 361–365. https://doi.org/10.1038/s41586-019-1022-9

[40] Ting-Han Lin, Willa Yunqi Yang, and Ken Nakagaki. 2023. ThrowIO: Actuated TUIs That Facilitate "Throwing and Catching" Spatial Interaction with Overhanging Mobile Wheeled Robots. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23).* Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3544548.3581267

[41] Donna M. Lloyd. 2007. Spatial Limits on Referred Touch to an Alien Limb May Reflect Boundaries of Visuo-Tactile Peripersonal Space Surrounding the Hand. *Brain and Cognition* 64, 1 (June 2007), 104–109. https://doi.org/10.1016/j.bandc.

2006.09.013

[42] Ke Ma and Bernhard Hommel. 2015. Body-Ownership for Actively Operated Non-Corporeal Objects. *Consciousness and Cognition* 36 (Nov. 2015), 75–86. https://doi.org/10.1016/j.concog.2015.06.003

[43] J Macqueen. [n. d.]. SOME METHODS FOR CLASSIFICATION AND ANALYSIS OF MULTIVARIATE OBSERVATIONS. *MULTIVARIATE OBSERVATIONS* ([n. d.]).

[44] Reiji Miura, Shunichi Kasahara, Michiteru Kitazaki, Adrien Verhulst, Masahiko Inami, and Maki Sugimoto. 2021. MultiSoma: Distributed Embodiment with Synchronized Behavior and Perception. In *Proceedings of the Augmented Humans International Conference 2021 (AHs '21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3458709.3458878

[45] Ken Nakagaki, Joanne Leong, Jordan L. Tappa, João Wilbert, and Hiroshi Ishii. 2020. HERMITS: Dynamically Reconfiguring the Interactivity of Self-propelled TUIs with Mechanical Shell Add-ons. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 882–896. https://doi.org/10.1145/3379337.3415831

[46] Ken Nakagaki, Jordan L Tappa, Yi Zheng, Jack Forman, Joanne Leong, Sven Koenig, and Hiroshi Ishii. 2022. (Dis)Appearables: A Concept and Method for Actuated Tangible UIs to Appear and Disappear Based on Stages. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3491102.3501906

[47] Jun Nishida, Soichiro Matsuda, Hiroshi Matsui, Shan-Yuan Teng, Ziwei Liu, Kenji Suzuki, and Pedro Lopes. 2020. HandMorph: A Passive Exoskeleton That Miniaturizes Grasp. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 565–578. https://doi.org/10.1145/3379337.3415875

[48] Gian Pangaro, Dan Maynes-Aminzade, and Hiroshi Ishii. 2002. The Actuated Workbench: Computer-Controlled Actuation in Tabletop Tangible Interfaces. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology (UIST '02)*. Association for Computing Machinery, New York, NY, USA, 181–190. https://doi.org/10.1145/571985.572011

[49] Esben Warming Pedersen and Kasper Hornbæk. 2011. Tangible Bots: Interaction with Active Tangibles in Tabletop Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2975–2984. https://doi.org/10.1145/1978942.1979384

[50] Daniel Perez-Marcos, Maria V. Sanchez-Vives, and Mel Slater. 2012. Is My Hand Connected to My Body? The Impact of Body Continuity and Arm Alignment on the Virtual Hand Illusion. *Cognitive Neurodynamics* 6, 4 (Aug. 2012), 295–305. https://doi.org/10.1007/s11571-011-9178-5

[51] Domenico Prattichizzo, Monica Malvezzi, Irfan Hussain, and Gionata Salvietti. 2014. The Sixth-Finger: A Modular Extra-Finger to Enhance Human Hand Capabilities. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. 993–998. https://doi.org/10.1109/ROMAN.2014.6926382

[52] Jan Richter, Bruce H. Thomas, Maki Sugimoto, and Masahiko Inami. 2007. Remote Active Tangible Interactions. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (Baton Rouge, Louisiana) *(TEI '07)*. Association for Computing Machinery, New York, NY, USA, 39–42. https://doi.org/10.1145/1226969.1226977

[53] Daniele Romano, Elisa Caffa, Alejandro Hernandez-Arieta, Peter Brugger, and Angelo Maravita. 2015. The Robot Hand Illusion: Inducing Proprioceptive Drift through Visuo-Motor Congruency. *Neuropsychologia* 70 (April 2015), 414–420. https://doi.org/10.1016/j.neuropsychologia.2014.10.033

[54] D. Rosenfeld, M. Zawadzki, J. Sudol, and K. Perlin. 2004. Physical objects as bidirectional user interface elements. *IEEE Computer Graphics and Applications* 24, 1 (2004), 44–49. https://doi.org/10.1109/MCG.2004.1255808

[55] Daniel Roth and Marc Erich Latoschik. 2020. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (Dec. 2020), 3546–3556. https://doi.org/10.1109/TVCG.2020.3023603

[56] Dibyendu Roy, Arijit Chowdhury, Madhubanti Maitra, and Samar Bhattacharya. 2019. Virtual Region Based Multi-robot Path Planning in an Unknown Occluded Environment. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 588–595. https://doi.org/10.1109/IROS40897.2019.8968177

[57] Maria V. Sanchez-Vives, Bernhard Spanlang, Antonio Frisoli, Massimo Bergamasco, and Mel Slater. 2010. Virtual Hand Illusion Induced by Visuomotor Correlations. *PLOS ONE* 5, 4 (April 2010), e10381. https://doi.org/10.1371/journal.pone.0010381

[58] Tomoya Sasaki, MHD Yamen Saraiji, Charith Lasantha Fernando, Kouta Minamizawa, and Masahiko Inami. 2017. MetaLimbs: Multiple Arms Interaction Metamorphism. In *ACM SIGGRAPH 2017 Emerging Technologies (SIGGRAPH '17)*. Association for Computing Machinery, New York, NY, USA, 1–2. https://doi.org/10.1145/3084822.3084837

[59] Atsushi Sato and Asako Yasuda. 2005. Illusion of Sense of Self-Agency: Discrepancy between the Predicted and Actual Sensory Consequences of Actions Modulates the Sense of Self-Agency, but Not the Sense of Self-Ownership. *Cognition* 94, 3 (Jan. 2005), 241–255. https://doi.org/10.1016/j.cognition.2004.04.003

[60] Melanie Schranz, Martina Umlauft, Micha Sende, and Wilfried Elmenreich. 2020. Swarm Robotic Behaviors and Current Applications. *Frontiers in Robotics and AI* 7 (2020).

[61] Ali Shafti, Shlomi Haar, Renato Mio, Pierre Guilleminot, and A. Aldo Faisal. 2021. Playing the Piano with a Robotic Third Thumb: Assessing Constraints of Human Augmentation. *Scientific Reports* 11, 1 (Nov. 2021), 21375. https://doi.org/10.1038/s41598-021-00376-6

[62] Satoshi Shibuya, Satoshi Unenaka, and Yukari Ohki. 2018. The Relationship Between the Virtual Hand Illusion and Motor Performance. *Frontiers in Psychology* 9 (2018).

[63] Sotaro Shimada, Kensuke Fukuda, and Kazuo Hiraki. 2009. Rubber Hand Illusion under Delayed Visual Feedback. *PLOS ONE* 4, 7 (July 2009), e6185. https://doi.org/10.1371/journal.pone.0006185

[64] J Snape, J Van Den Berg, S J Guy, and D Manocha. 2010. Smooth and Collision-Free Navigation for Multiple Robots under Differential-Drive Constraints. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Taipei, 4584–4589. https://doi.org/10.1109/IROS.2010.5652073

[65] Masanori Sugimoto, Tomoki Fujita, Haipeng Mi, and Aleksander Krzywinski. 2011. RoboTable2: A Novel Programming Environment Using Physical Robots on a Tabletop Platform. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology* (Lisbon, Portugal) *(ACE '11)*. Association for Computing Machinery, New York, NY, USA, Article 10, 8 pages. https://doi.org/10.1145/2071423.2071436

[66] Ryo Suzuki, Eyal Ofek, Mike Sinclair, Daniel Leithinger, and Mar Gonzalez-Franco. 2021. HapticBots: Distributed Encountered-type Haptics for VR with Multiple Shape-changing Mobile Robots. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*. Association for Computing Machinery, New York, NY, USA, 1269–1281. https://doi.org/10.1145/3472749.3474821

[67] Ryo Suzuki, Clement Zheng, Yasuaki Kakehi, Tom Yeh, Ellen Yi-Luen Do, Mark D. Gross, and Daniel Leithinger. 2019. ShapeBots: Shape-changing Swarm Robots. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 493–505. https://doi.org/10.1145/3332165.3347911

[68] Kazuma Takada, Midori Kawaguchi, Akira Uehara, Yukiya Nakanishi, Mark Armstrong, Adrien Verhulst, Kouta Minamizawa, and Shunichi Kasahara. 2022. Parallel Ping-Pong: Exploring Parallel Embodiment through Multiple Bodies by a Single User. In *Proceedings of the Augmented Humans International Conference 2022 (AHs '22)*. Association for Computing Machinery, New York, NY, USA, 121–130. https://doi.org/10.1145/3519391.3519408

[69] Ryo Takizawa, Takayoshi Hagiwara, Adrien Verhulst, Masaaki Fukuoka, Michiteru Kitazaki, and Maki Sugimoto. 2021. Dynamic Shared Limbs: An Adaptive Shared Body Control Method Using EMG Sensors. In *Proceedings of the Augmented Humans International Conference 2021 (AHs '21)*. Association for Computing Machinery, New York, NY, USA, 10–18. https://doi.org/10.1145/3458709.3458932

[70] G. Tieri, E. Tidoni, E. F. Pavone, and S. M. Aglioti. 2015. Mere Observation of Body Discontinuity Affects Perceived Ownership and Vicarious Agency over a Virtual Hand. *Experimental Brain Research* 233, 4 (April 2015), 1247–1259. https://doi.org/10.1007/s00221-015-4202-3

[71] Kohei Umezawa, Yuta Suzuki, Gowrishankar Ganesh, and Yoichi Miyawaki. 2022. Bodily Ownership of an Independent Supernumerary Limb: An Exploratory Study. *Scientific Reports* 12, 1 (Feb. 2022), 2339. https://doi.org/10.1038/s41598-022-06040-x

[72] Loes C. J. van Dam and Joey R. Stephens. 2018. Effects of Prolonged Exposure to Feedback Delay on the Qualitative Subjective Experience of Virtual Reality. *PLOS ONE* 13, 10 (Oct. 2018), e0205145. https://doi.org/10.1371/journal.pone.0205145

[73] Jur van den Berg, Ming Lin, and Dinesh Manocha. 2008. Reciprocal Velocity Obstacles for Real-Time Multi-Agent Navigation. In *2008 IEEE International Conference on Robotics and Automation*. 1928–1935. https://doi.org/10.1109/ROBOT.2008.4543489

[74] Markus Wirth, Stefan Gradl, Georg Prosinger, Felix Kluge, Daniel Roth, and Bjoern M. Eskofier. 2021. The Impact of Avatar Appearance, Perspective and Context on Gait Variability and User Experience in Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 326–335. https://doi.org/10.1109/VR50410.2021.00055

[75] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[76] Nahoko Yamamura, Daisuke Uriu, Mitsuru Muramatsu, Yusuke Kamiyama, Zendai Kashino, Shin Sakamoto, Naoki Tanaka, Toma Tanigawa, Akiyoshi Onishi, Shigeo Yoshida, Shunji Yamanaka, and Masahiko Inami. 2023. Social Digital Cyborgs: The Collaborative Design Process of JIZAI ARMS. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Hamburg</city>, <country>Germany</country>, </conf-loc>) *(CHI '23)*.

Association for Computing Machinery, New York, NY, USA, Article 369, 19 pages. https://doi.org/10.1145/3544548.3581169

[77] Shigeo Yoshida, Tomoya Sasaki, Zendai Kashino, and Masahiko Inami. 2023. TOMURA: A Mountable Hand-Shaped Interface for Versatile Interactions. In *Proceedings of the Augmented Humans International Conference 2023* (Glasgow, United Kingdom) *(AHs '23)*. Association for Computing Machinery, New York, NY, USA, 243–254. https://doi.org/10.1145/3582700.3582719

[78] Lilith Yu, Chenfeng Gao, David Wu, and Ken Nakagaki. 2023. AeroRigUI: Actuated TUIs for Spatial Interaction Using Rigging Swarm Robots on Ceilings in Everyday Space. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3544548.3581437

[79] Yiwei Zhao, Lawrence H. Kim, Ye Wang, Mathieu Le Goc, and Sean Follmer. 2017. Robotic Assembly of Haptic Proxy Objects for Tangible Interaction and Virtual Reality. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. Association for Computing Machinery, New York, NY, USA, 82–91. https://doi.org/10.1145/3132272.3134143

A APPENDIX

A.1 Swarm Robot Control for Following Subgoal Position

The robot receives the subgoal (target) position from the host computer via RF communication and moves to follow it. It is desirable for the robot to follow a smooth path to the subgoal position. In controlling wheeled robots, using a Bézier, spline, or cross-oid curve as the path is common. This curve passes through the current and subgoal positions to avoid sudden changes in angular velocity. However, due to the limited computational resources of the robot's microcontroller, it has been difficult to implement control that sequentially calculates and follows these curves. On the other hand, since the robot can get its absolute position information with the projection-based method, and the subgoal position is updated by the host computer every 100 ms, the distance between the current position and the subgoal position is considered to be close, and we thought that sudden changes in angular velocity would be rare even on a path connecting these two points by a straight line. Thus, we implemented a control model that follows the straight path.
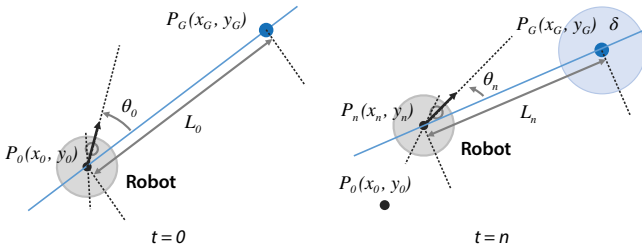

Figure 20: Concept of robot control model for following goal position.

The concept of the control model is shown in Figure 20. Let $P_0(x_0, y_0)$ be the robot's initial position at time $t = 0$ given the subgoal (target) position and $P_G(x_G, y_G)$ the subgoal position. Next, define $\overrightarrow{P_0P_G}$ as the vector connecting points P_0 and P_G, with L ($L = L_0$ at $t = 0$) as the distance between them, and θ ($\theta = \theta_0$ at $t = 0$) as the angle between the robot's direction of motion and the vector $\overrightarrow{P_0P_G}$. (Counterclockwise is defined as positive.) In this case, the following law controls the velocities V_l and V_r of the robot's left and right motors.

$$V_l = V - \Delta V$$
$$V_r = V + \Delta V$$
$$V = K_L L$$
$$\Delta V = K_\theta \theta + K_{\dot\theta} \dot\theta,$$

where $K_L, K_\theta, K_{\dot\theta}$ are constants that represent the control gain. Based on the above control law, $\overrightarrow{P_0P_G}$, the distance L_n, and the angle θ_n are updated from the robot's current position $P_n(x_n, y_n)$ to the subgoal position, and V_l and V_r are calculated and output. However, if this control law is followed, the point P_G is theoretically unreachable and will never fully converge, since $L = 0$ and the velocity will converge to 0 as the robot approaches P_G. Therefore, the robot is judged to have converged when it is within a certain distance σ from P_G, and the robot stops at that point. In addition, the minimum velocities $V_{l,min}$ and $V_{r,min}$ are set for V_l and V_r, respectively. The maximum speed of motors V_{max} is determined by the maximum speed value of the slower of the two motors. This is based on the actual measured speeds of the left and right motors as an upper limit. This can be written in the formula as follows:

$$V_l = \begin{cases} V_{l,min} & \text{if } V_l < V_{l,min} \\ V_{max} & \text{if } V_l > V_{max} \\ V_l & \text{otherwise.} \end{cases}, V_r = \begin{cases} V_{r,min} & \text{if } V_r < V_{r,min} \\ V_{max} & \text{if } V_r > V_{max} \\ V_r & \text{otherwise.} \end{cases}$$

$$V_{max} = \begin{cases} V_{l,max} & \text{if } V_{r,max} \geq V_{l,max} \\ V_{r,max} & \text{if } V_{r,max} < V_{l,max} \end{cases},$$

where $V_{l,max}$ and $V_{r,max}$ are the maximum speed value of left and right motors.

If a new subgoal position P_G is given before convergence, P_G is updated, and the subgoal following continues. Here, the dimension (unit) of the calculated V_l and V_r is [mm/s], but the dimension of the PWM duty of the voltage value, which is a control value that can be input to the motor, is [rpm]. Therefore, the calculated V_l and V_r are converted to the PWM duty value by the following formula:

$$Duty_l = f(V_l)$$
$$Duty_r = g(V_r),$$

where $Duty_l$ and $Duty_r$ are the PWM values of left and right motors, and $f(x)$ and $g(x)$ are functions corresponding to the PWM value calculated by the robot's calibration mode. We confirmed that this control rule can be used to control the robot to successfully reach the goal position even with a microcontroller with limited computational resources.