ScalableBody: A Telepresence Robot that Supports Face Position Matching using a Vertical Actuator

Akira Matsuda
The University of Tokyo
Bunkyo-ku, Tokyo, Japan
akira.matsuda.ut@gmail.com

Takashi Miyaki
The University of Tokyo
Bunkyo-ku, Tokyo, Japan
miyaki@acm.org

Jun Rekimoto
The University of Tokyo, Sony CSL
Bunkyo-ku, Tokyo, Japan
rekimoto@acm.org

Figure 1: Outline of ScalableBody. (A) ScalableBody can change its height dynamically. (B) It has an omnidirectional camera at the top of its head and the operator can control ScalableBody with a wide field of view. (C) Adjusting its height based on communication contexts of sitting or standing.

ABSTRACT

Seeing one's partner's face during remote conversation is one of the most important factors for effective communication. When using a telepresence robot, matching face positions with one's partner is sometimes difficult, because face position varies in different situations (e.g., standing or sitting). However, existing telepresence robots cannot change their height. Moreover, due to limited camera angle, the conversation partner's face is often partly cut off in the camera view. Therefore, users cannot communicate while seeing each other's faces. To overcome these problems, we designed a telepresence robot called ScalableBody. ScalableBody has a vertical actuator that allows it to change its height and an omnidirectional camera that provides a wide view. The robot facilitates communication for different contexts using vertical actuation to match the conversation partners' face positions. Furthermore, the operator can see a partner's face in any direction through an omnidirectional camera. This approach can also provide users with the experience of being a different height, as if a giant or a child. In this paper, we describe the vertical actuator mechanism and report our user study on the telepresence robot.

AH '17, March 16-18, 2017, Mountain View, CA, USA

DOI: http://dx.doi.org/10.1145/3041164.3041182

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); *User interface design;* **•Computer systems organization → Robotics; External interfaces for robotics; •Hardware → Communication hardware, interfaces and storage;**

Keywords

telepresence, remote communication, surrogate robots, eye-contact, face position matching

1. INTRODUCTION

Looking at a conversation partner's face while talking gives a good impression and produces various beneficial effects in communication. Eye contact is an especially important factor in communication [25]. In recent years, though, many remote communication methods that use digital technologies have been proposed. These include not only conventional text-based chatting or video conference software, but also have extended to telepresence robots that enable us to communicate with a remote person and to move freely around a remote location [3, 5, 6, 8, 7]. Telepresence robots have been used in myriad situations, including participating in international conferences [2, 4].

Typical telepresence robots have wheels for moving around, a screen for video communication, speakers, and microphones [1, 3, 7]. Their physical characteristics, such as height, are decided by the manufacturers, forcing the user to communicate with others from a fixed height, which may affect their relationship positively or negatively [9, 13]. Moreover, the user cannot follow changes in a conversation partner's eye or face position for various situations like standing or sitting, and it can be difficult to talk while looking at one's partner's face because of the limited view angle of the camera. Furthermore, Rae et al. reported that the height of the robot makes the operator less persuasive when the robot is shorter than the partner [23].

In this paper, we propose a novel telepresence robot, ScalableBody, which can dynamically change its height and freely look around (Figure 1). We develop a vertical actuator for ScalableBody that enables the operator to change the robot's height dynamically. Moreover, the robot has an omnidirectional camera so that the remote user can see in all directions without mechanically operating the robot. We investigate the usability of gaze/facial position matching using a vertical actuator and an omnidirectional camera.

2. RELATED WORK

2.1 Telepresence system and robot for human representation

Many telepresence robots and systems [1, 3, 7, 19, 20, 21, 22, 24] and teleoperation systems [12] have been studied, demonstrating their initial feasibility. The appearance of typical telepresence robot is almost always the same: a static body that contains a pole that is connected to a wheeled base lifting a display that shows the remote user's face. For example, the Double Robotics can change its height dynamically, but with a range of only 30 cm (120 cm to 150 cm) [3]. Therefore, current telepresence robots cannot match face positions for any context because of their short height; 150 cm is too short for conversation with the average person who is standing.

Sirkin et al. stated a relation for robot motion [26]. They proposed a telepresence robot that can move its head or hand to make a nonverbal gesture. Tobita et al. developed an alternative telepresence robot with a blimp [28]. We explore a telepresence robot variation, called ScalableBody, focusing on its height and face position.

Rae et al. investigated the influence of height in telepresence robot communication [23]; the operator of a telepresence robot feels dominated by a conversation partner when the robot is shorter than the partner. Desai et al. found that the local user preferred that the robot's height was the same as theirs [10]. These studies also suggest that the height of the robot is an important social cue for natural communication. However, they investigated the effect of robot height for the dialogue partner; an effect on the operator has not yet been determined.

Our research focuses on communication factors in a telepresence robot, not only its height but also the face or eye position of the operator.

2.2 Robot for human augmentation

Changing the height of a telepresence robot can also be seen as an example of human augmentation. *Big Robot Mk1* is a special exoskeleton robot that is 5-m tall, in which a pilot can ride and move [15]. This installation allows us to investigate how we feel when we become that tall. *Flying head* is a drone control system [14]. The user controls a drone with a motion capture system and looks at a drone image through a head mounted display. Our system also provides a telepresence view from a higher point of view than the users.

3. SCALABLEBODY

To address the issue of the face position mismatch due to different contexts, we develop a novel telepresence robot called ScalableBody. We implement ScalableBody with the following features.

3.1 Dynamic Height Changing using Vertical Actuation

As described in the introduction, existing telepresence robots cannot change their height to accommodate the context of a conversation partner. For example, when the partner sits down, his/her face position matches the robot's camera. However, when the partner stands up, his/her face will be out of the camera's view. Such situations have the effect of hindering communication. To address this issue, we propose dynamic height change actuation. This vertical actuation enables three features: (1) face position matching, (2) reflecting the operator's actual height, and (3) human height augmentation. We describe these features below.

3.1.1 Face position matching

ScalableBody has a high-extension ratio vertical actuator that enables the operator to move the position of the camera mounted on top of the robot up and down by remote control (Figure 2). This feature enables communicating partners to match each other's face position.

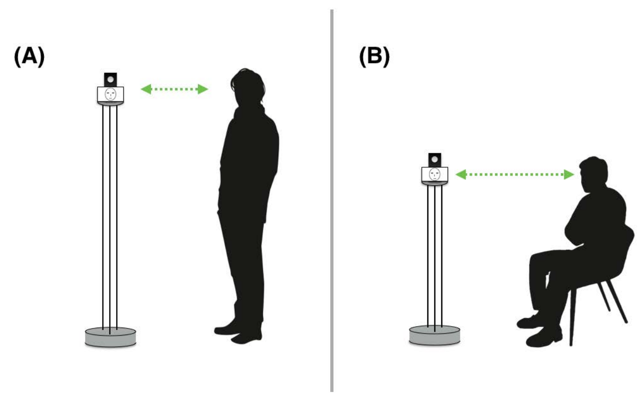

Figure 2: Changing robot height dynamically for each context. (A) When the partner is standing, the robot can stretch to the height of the partner. (B) When the partner is sitting down, the robot can become short.

3.1.2 Reflecting the operator's actual height

The operator can adjust the robot's height to the operator's height (Figure 3). For instance, if the operator is short, the robot can also become shorter. This feature solves the actual height difference problem between the operator and conversation partner.

3.1.3 Human height augmentation

The operator can have an extraordinary experience, as if they had become a very short or very tall person. Thus, vertical actuation also enables human augmentation (Figure 4).

It is possible to be taller than the operator by vertical actuation, allowing the operator to see the world through another person's

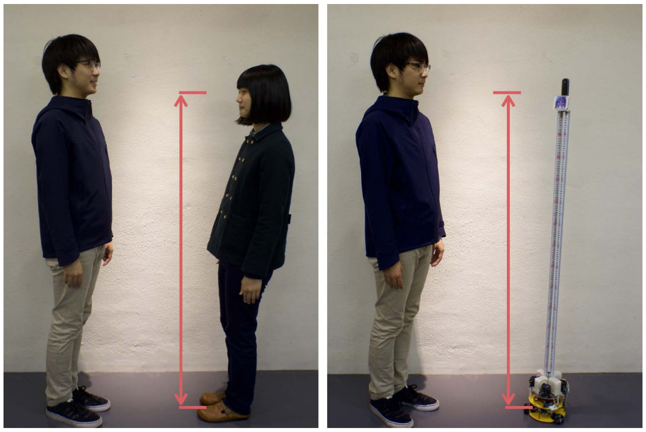

Figure 3: The robot becomes the same height as the operator.

view. This is an important function from the viewpoint of accessibility. For example, it would make it possible to design facilities according to a more realistic children's point of view. Moreover, by looking around a facility through ScalableBody, it is possible to evaluate its design for a variety of users.

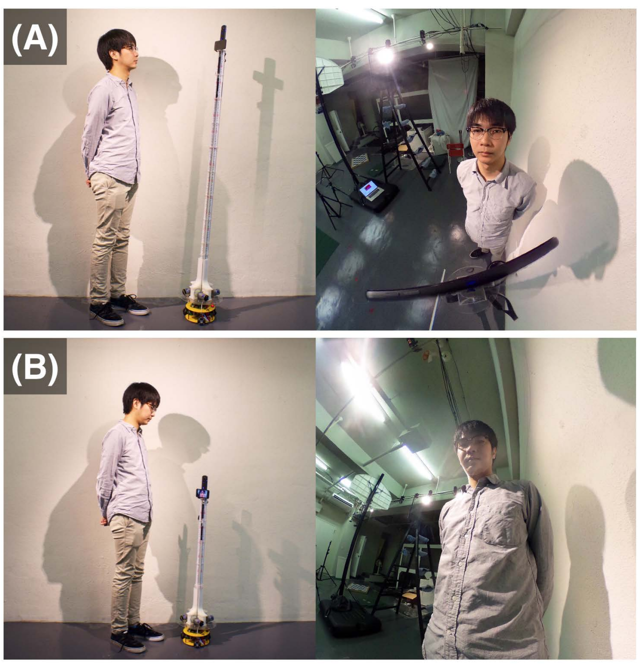

Figure 4: Difference in viewpoint that depends on height. (A) The view from a high position makes you a tall person. (B) The view from a low position is an extraordinary experience.

3.2 Omnidirectional Camera View

An omnidirectional camera is installed on the top of ScalableBody. Through the omnidirectional camera, the operator can move his/her viewpoint freely, so that it is easy to understand the local environment, such as the condition of ground or the height of the ceiling.

It is important to understand the local environment in remote collaboration work [11], as demonstrated with an omnidirectional camera by Kasahara et al [17]. In addition, it is useful in a telepresence robot [18]. Therefore, we added an omnidirectional camera to better understand the local environment according to occlusion of the height, which provides the following benefits:

- The wide omnidirectional view allows the operator to look over the local environment easily.
- When the robot stretches to a high position, the operator can see his/her conversation partner from the top of his/her head to the bottom of his/her legs.

We developed an operation interface (client) for operating the robot while watching the entire surrounding image. The operator controls the robot using a mouse and a keyboard. By moving the mouse, the operator can change the viewpoint in any direction (Figure 5), enabling the operator to see up, down, left, and right without changing the robot's orientation. Existing telepresence robots, such as Double and BeamPro, have a camera for seeing feet, although this requires switching camera images.

In addition, it is possible to extend, retract, rotate, or move forward and backward by keyboard input. When the robot moves forward, it always moves forward in the direction of current viewpoint. Therefore, it is not necessary for the operator to consider the orientation of the robot.

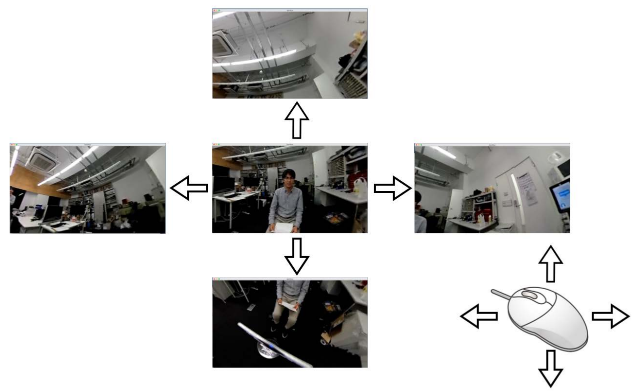

Figure 5: The omnidirectional camera interface makes it easy to see in any direction using mouse movement. When the operator moves the mouse to the right, the field of view moves to the right. Likewise, by moving in another direction, it is possible to move the field of view in various directions.

4. SYSTEM CONFIGURATION

ScalableBody consists of a vertical actuator, an omnidirectional camera, an omniwheel robot, a screen for displaying the operator's face, and an operation interface. The system configuration is shown in Figure 6.

4.1 Vertical Actuator

4.1.1 Mechanism

The vertical actuator's mechanism was inspired by the high extension ratio rod actuator used in KineReels [27]. Although they use the proposed rod mechanism to express three-dimensional structures, we applied this mechanism to telepresence robots and added a ring mechanism to prevent bending.

Compared to other expansion and contraction mechanisms, this mechanism is advantageous in that it is compact and has a high extension ratio. In a telepresence robot, the mechanism needs to be

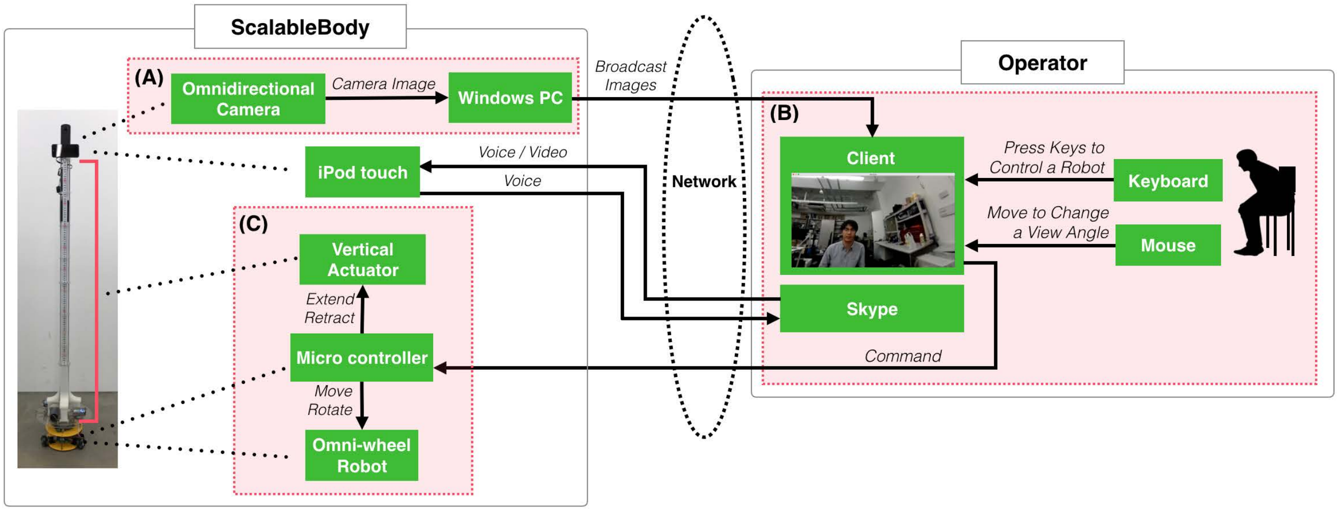

Figure 6: System configuration. (A) An omnidirectional camera at the top of ScalableBody, connected to a Windows PC. The Windows PC broadcasts the camera image to the client on the operator's PC. (B) On the operator's PC, the client shows the camera image and the operator can change the view angle. The client also sends keyboard commands to control ScalableBody. Moreover, Skype sends the operator's voice and video to an iPod Touch on ScalableBody. (C) When a microcontroller on ScalableBody receives the command, it controls an omniwheel robot at the bottom of ScalableBody or a vertical actuator.

small because of the robot's payload and mobility. Other expansion mechanisms include rack-and-pinion, leadscrew, and scissor lift. Rack-and-pinion can only be extended to at most twice the minimum stature. Leadscrew will stay in the field of view of the perimeter camera because the rod that supports it as it is short and will remain as it is. The size of the scissor lift is smaller than any of the above, and as the height at maximum extension increases, the mechanism becomes huge. Since our mechanism is completed only with an iron reel that becomes small by winding on a shaft and DC motor, it does not become huge compared to these methods. Therefore, to be able to become extremely tall and short, we found it necessary to adopt our mechanism.

4.1.2 Implementation

To change the robot's height dynamically, we developed a high extension ratio vertical actuator. At the bottom of the actuator, there are four iron reels. These are bundled by *rings* and make one rod. Each reel is connected to a DC motor (TSUKASA Electric, TG-05D-SG-300-HA). When the motor runs, a shaft of the iron reel is rotated and pushes up the iron reel. Therefore, four iron reels make a long rod. The rings are connected by *pole* and are lifted when the rod becomes tall to prevent the iron reels from bending. ScalableBody can stretch up to 235 cm from a minimum height of 115 cm. The maximum and minimum heights depend on the length of the iron reels, the power of the DC motors, the number of rings, and the stability of the rod, which could swing during the robot's movement. Figure 7 shows a mechanism for lifting rings using poles.

4.2 Omnidirectional Camera

We used the Ricoh Theta S[1] for an omnidirectional camera. Ricoh Theta S is a USB camera on a Windows PC mounted on ScalableBody; the camera image, which is a dual fisheye image, is transmitted to the client via WiFi using GStreamer for video transmission. The resolution of the camera image is 1280 $\times$ 720 and

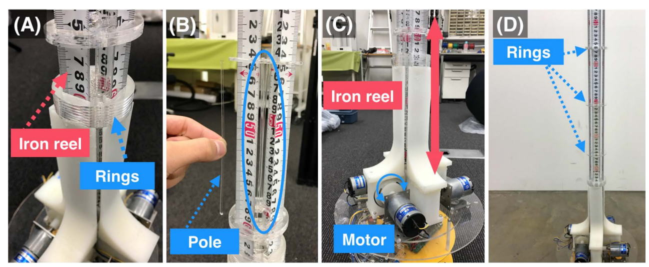

Figure 7: Mechanism of the vertical actuator. (A) Rings are supporting an iron reel and are stacked at the bottom of the actuator. (B) Rings are connected by a pole. (C) A DC motor rotates and pushes up an iron reel. Rings are pulled up and a pole hooks to pull up a lower ring. (D) Each ring is pulled up and supports an iron reel.

its frame rate is 15 fps. The client converts the received image into an equirectangular image and maps the equirectangular image as a texture onto the spherical 3D model. By rotating the sphere with mouse movement, the operator can change the viewpoint. To convert the camera image to an equirectangular image, we used OpenGL shader; we used openFrameworks to implement the user interface.

4.3 Omniwheel robot

We used an omniwheel robot (Nexus Robot, 3 WD 48 MM OMNI WHEEL MOBILE ARDUINO ROBOT KIT 10019) [2] as the base of ScalableBody. To remotely operate this robot, we used ESP8266 [3] as a micro controller. The ESP8266 is connected to the control board (Arduino) of the omniwheel robot. The client sends data to the ESP8266 using the OSC protocol (OpenSound Control proto-

[1]https://theta360.com/en/about/theta/s.html

[2]http://www.nexusrobot.com/product.php?id_product=87

[3]http://espressif.com/en/products/hardware/esp-wroom-02/overview

col). ESP 8266 controls the Arduino via the UART per the commands received from the client.

4.4 Operator's Face Display

We used an iPod Touch (5th Gen) as the screen to display the operator's face. Skype connects to the iPod Touch and the PC used by the operator; the video of the operator are displayed on its screen. The voice also is sent to the operator via Skype on the PC side.

4.5 Operation User Interface

The operator controls ScalableBody through the client remotely. Audio from partners and the operator's face and voice are transmitted through Skype. The surrounding image is received via WiFi. He/she can change viewpoint by using mouse movement to rotate the omnidirectional camera image. The input from the keyboard is transmitted as commands to the ESP8266, and the robot moves forward or backward, rotates left or right, and extends or retracts its actuator. The key bindings are the same as those of Double.

5. EXPERIMENT: VERTICAL ACTUATOR

To investigate the performance of the vertical actuator, we investigated the maximum payload and the extension and retraction speeds.

5.1 Payload size

To measure payload, the actuator was set to expand and contract. For force measurements, a force gauge (IMADA DS2-200N) was used. We apply a voltage of 12V for several seconds and investigate the value when the force gauge value does not increase. We tried this ten times. The experimental environment and result are shown in Figure 8. We found the average force to be 61.7 N, the maximum to be 70.5 N, and the minimum to be 52.2 N (SD = 6.13). The average value almost agrees with the 6 kg of the motor's torque. The value fluctuation is a problem of the actuator design. The shaft winding the iron reel is connected to the motor shaft directly. In this design, the iron reel loosens inside the case when receiving force from above. Therefore, it is impossible to provide upward propulsive power while it is loose. In response to this problem, instead of rotating the shaft that winds the iron reel, it can be improved by making the mechanism push the reel by pinching it with the roller and pushing the reel by rotating the roller.

5.2 Actuation Speed of the Vertical Actuator

To investigate the drive speed of the actuator, we measured how many seconds it took to stretch 10 cm while stretching between 70 cm and 170 cm.

The results of this experiment are shown in Figure 8. The graph shows that the speed when starting to extend (the speed at 80 cm), is the slowest. It is likely that the motor needs time to perform in its best condition. The graph at the bottom right of Figure 8 shows the average time at which each height was reached. This graph shows that it takes about 2.5 seconds to reach from 80 cm to 70 cm, then it takes about 2 seconds to reach 90 cm. Therefore, the speed is not stable unless it is applied for about 5 seconds. Moreover, the speed at the end of shrinking (at 70 cm) is the slowest. This is the timing at which the iron reel fits inside the case and it is assumed that the speed decreases due to the friction inside the case.

The graph at the top right of Figure 8 shows that the speed is slowest when the actuator is retracting, possibly because the iron reel is wound around the shaft as it is loosened when foreshortening and friction occurs between the reel and case and the speed decreases.

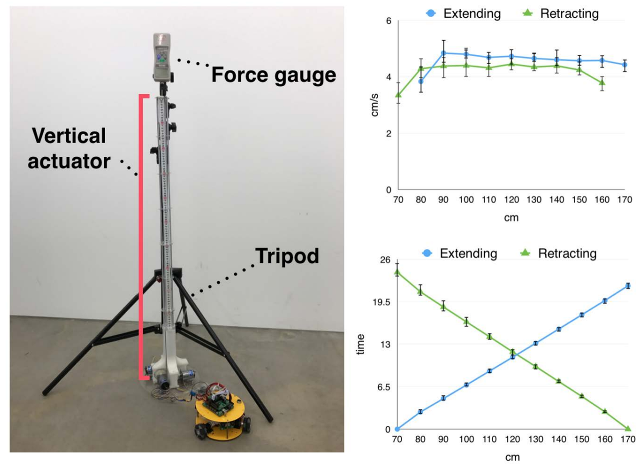

Figure 8: Performance of our vertical actuator. Error bars show the maximum or minimum value of each condition. Left: experimental setup; Bottom right: time at which each height was reached; Top right: speed of vertical actuation.

6. USER STUDY

We conducted a user study to compare multiple conditions and verify the effectiveness of adjusting face position by adjusting the robot's height and the operation of the surrounding image.

6.1 Participants

We recruited seven participants (mean age: 24.7 years old, SD = 4.02). Five participants were students, and two were not. The students are engineering majors and the other participants are a programmer and an engineer. Two participants had some experience with telepresence robots and all participants had used videoconferencing tools such as Skype for various applications, ranging from meetings to conversation with a friend to a job interview or presentation at a remote site event. All participants were using our robot for the first time.

6.2 Procedure

The experience task was divided into two parts and carried out in three conditions. The participant listened to the description of various research projects from his/her conversation partner through a telepresence robot. First, the partner sat down and talked about the research. In this part, the height of the robot was adjusted so that both ScalableBody and Double matched the face position of the experiment collaborator. Then, the partner would stand up and stand beside a large screen (described in the environment setup). Then, they would continue to describe the research with pictures on the screen. We instructed the participants to see the face of their partner as much as possible when the partner was talking to confirm the effect of conversation when matching face position on remote communication, which is the backbone of this research. The three conditions were as follows: (1) use ScalableBody without stretching, (2) use ScalableBody freely, and (3) use another telepresence robot. In this case, we used Double. The experiment time was about 3 to 4 minutes for each task. We randomized the task execution order for each participant.

We used a questionnaire to investigate its attributes before starting the experiment. After that, we instructed the participants on how to operate the robot and allowed them practice for about 5 minutes. Each time the task finished, we asked a simple question by questionnaire with a 7-point Likert scale. The questions in the questionnaire used for each condition are as follows:

Q1 The video was sufficiently clear.

Q2 The audio was sufficiently clear.

Q3 The presentation was intelligible.

Q4 I felt that the presenter was friendly.

Q5 I was nervous while listening.

Q6 I felt as if I was close to the presenter in the same room.

Q7 I felt as if I were talking with the presenter in the same room.

After completing the three tasks, we presented a questionnaire using the System Usability Scale (SUS) with a 7-point Likert scale and the free description question about the operation feeling of the robot. The SUS questions are as follows:

Q1 I think that I would like to use this system frequently.

Q2 I found the system unnecessarily complex.

Q3 I thought the system was easy to use.

Q4 I think that I would need the support of a technical person to be able to use this system.

Q5 I found the various functions in this system to be well-integrated.

Q6 I thought there was too much inconsistency in this system.

Q7 I would imagine that most people would learn to use this system very quickly.

Q8 I found the system very cumbersome to use.

Q9 I felt very confident using the system.

Q10 I needed to learn a lot of things before I could get going with this system.

The questions in the questionnaire after all the task are as follows:

Q1 Did you feel that it was hard to control the ScalableBody? (Yes / No)

Q2 Describe the reason why if you answered yes in Q1. (Free description)

Q3 Did you feel that it was hard to see the partner face? (Yes / No)

Q4 Describe the reason why you answered yes in Q3. (Free description)

Q5 Did the impression of the partner change as the height of the robot changed? (Yes / No)

Q6 Describe the reason why you answered yes in Q5. (Free description)

Q7 Please tell us other points you noticed. (Free description)

Finally, after answering all the questionnaires, we conducted a semi-structured interview about the responses for about 10 minutes.

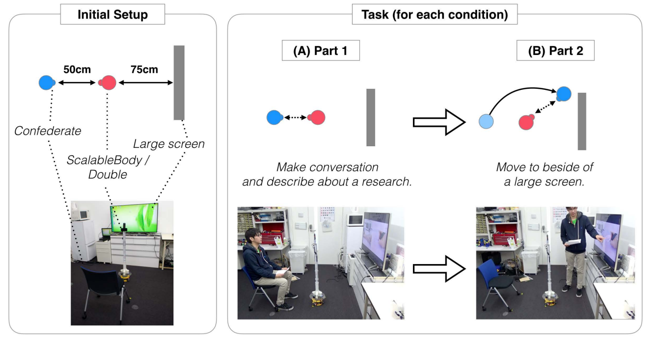

Figure 9: User study environment. In the initial setup, the participant and partner face each other. Each condition has two parts. (A) Part 1, the partner sits down while presenting a summary of research. (B) Part 2, the partner moves beside the large screen and presents more about the research with some pictures.

6.3 Environment Setup

The robot operated by the participant and the conversation partner begin a task by facing each other. We used a 50-inch SHARP LC- 50U30 as the large screen. The robot and the partner started the experiment a distance away from each other's face. We arranged the robot, large screen, and partner as shown in Figure 9. During the user study, we recorded the participant, the PC screen, and the partner.

6.4 Result

In this section, we describe the results of our user study.

6.4.1 Face position matching and social interaction

Four participants answered that their impression of their partner changed. By stretching the robot's height to fit his/her partner, Participant 1 said, *"I felt as if seeing the screen with the same person as the partner."* Participant 2 said, *"The intimidation disappeared between when I used Double. I was able to talk with the usual impression and feel equal."* Participant 4 gave a similar answer. In addition, Participant 2 explained that he felt as though he were talking in the real world, even though the height of the robot was lower than his height when his partner's face position was matched: *"When the robot's height becomes the same as the partner and the height of the position of the face, it was the same feeling when I was talking with the partner."*

A negative response was given by multiple participants when the height was lower than the partner's. Participant 4 said, *"It is hard to speak to when I was shorter than partner."* Participant 5 said, *"When I was listening the research description, I wanted to adjust the height to see the face and I did it."* All participants could not see the face of the partner.

6.4.2 Usability of controlling an omnidirectional camera image

Some opinions were gained from the participants regarding the interface for viewing the omnidirectional camera images. Participant 1 said, *"I felt that it is surprisingly convenient to be able to control the viewpoint of 360-degree camera image."* Participant 2 stated the following on the view angle: *" It is good to look around, but the feet could not be seen without moving the viewpoint. This robot is better than Double. Because humans are grasping obstacles in the peripheral vision, I felt it difficult to check the floor when*

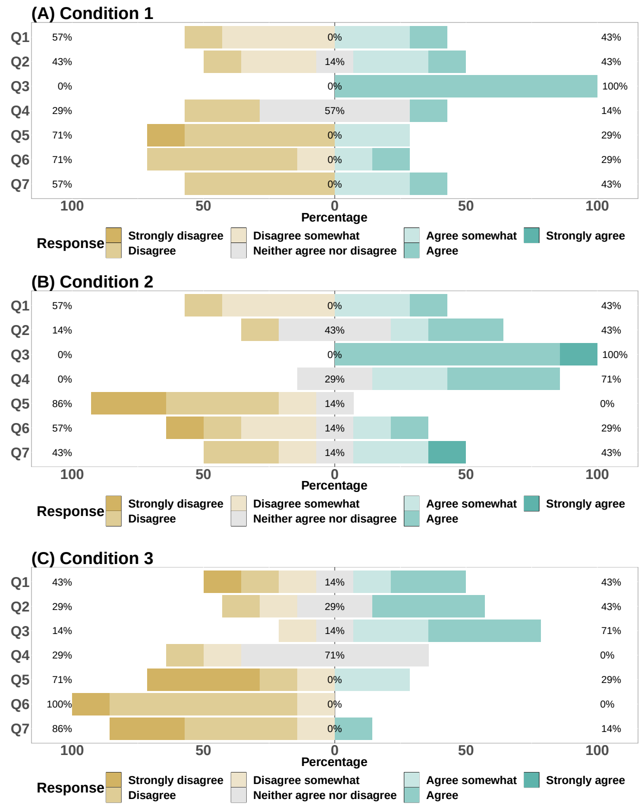

Figure 10: Result of the questionnaire for each condition. (A) In Condition 1, participant uses ScalableBody with disabled extend/retract feature. (B) In Condition 2, participant uses ScalableBody without any limitation. (C) In Condition 3, participant uses Double.

moving.“ Participant 2 described Double's viewpoint movement as follows: *“ScalableBody is very convenient to be able to freely move the point of view whereas stress is accumulated when using Double.“* Participant 3 found it hard to adjust viewpoint switching because *“The trackpad sensitivity was high, so it 's hard to change the field of view. Also, the distance to the partner is hard to understand.“* Participant 6 responded as follows: *“By letting the image of the omnidirectional camera freely, it was convenient to look around. It is no need to move the robot to see in all direction.“* Participant 7 answered *“It is hard to rotate the view to the left and right by mouse operation when looking down.“*

6.4.3 System usability

Figure 11 shows the result of the SUS. We provided the questionnaire after answering questions about what felt difficult when using ScalableBody. Participant 1 said, *“I did not know how high the robot grew. I did not know whether I could extend to the height I wanted.“* Participant 2 said that, *“It (an omnidirectional camera image) was convenient to look around the surroundings.“* Participant 5 said that, *“It was difficult to adjust the position of the robot with the partner because it cannot move the robot and stretch in parallel.“* Participant 7 had *“Video latency; there was a gap between when the button was pressed and when the robot moved.“* In the experimental environment, the image latency was almost the same between ScalableBody and Double. Also, Participant 7 replied,“Agree somewhat“ to the question that they needed to learn a lot to use this system: *“ I thought that practicing to maneuver the robot without colliding was necessary.“*

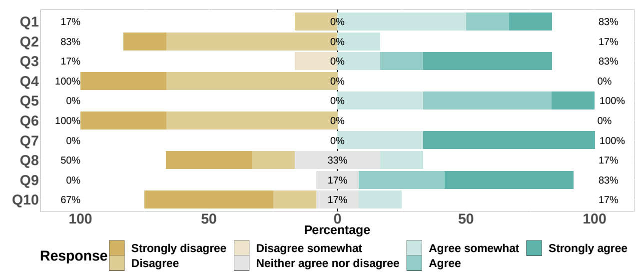

Figure 11: Result of the system usability scale questionnaire.

6.4.4 Observation

From the observation from recorded video of the user study, in Condition 3 in Part 2, the participant adjusted the orientation and position as much as possible so that the user could see both the partner and the screen, so that they fit into the field of view of the robot's camera. With Conditions 1 and 2, some participants turned their direction of view to the partner and screen alternately. Other users did not change gaze direction at all. As the partner stood up, a difference in height with Double was created, and participants could not see the large display while looking at the face of the partner. Many participants always faced the direction of the screen without facing the direction of their partner. In addition, the participant moved position to where the screen could be seen well.

In Condition 3, all participants retreated to see the screen. However, four had an issue where the partner would hit the chair. Participants who collided with the chair included both participants who were checking their feet and those who did not.

Four participants were watching a space that did not include either the partner or screen while listening in Conditions 1 and 2. This behavior was not seen in Condition 3, which used Double.

7. DISCUSSIONS

7.1 Communication with free field of view and face position matching

As shown in Figure 10, the participants found that in Condition 2, they felt more friendly toward their partner than in other conditions. It is assumed, as is clear from the previous interview, that the impression of the operator changes positively by matching height with the dialogue partner. In addition, it is assumed that the reduction in intimidation makes the participant feel friendly to the partner in Condition 2.

From Figure 10, the participants felt as though they were interacting in the same room as the partner in Conditions 1 and 2. While using ScalableBody, most participants watched the screen and their partner at the same time, whereas when using Double, only the screen could be seen. Being able to see the conversation partner's face freely is considered to enhance engagement. In Condition 2, several participants said that they felt equal by talking at the same height in the interview. Therefore, seeing a face and matching face position enhances the engagement between partners.

Participant 2 said that s/he did not know the height of the robot. However, s/he successfully changed its height to his/her desired height. This feedback suggests that there is no need to show the current robot's height to the user.

7.2 Usability for the omnidirectional camera interface

As shown in Figure 11, it seems that ScalableBody can be easily operated even by first-time users. In the interview, all participants said that our robot is much convenient than Double, even though most participants had never used a telepresence robot. Johnson et al. stated that it was difficult to use the telepresence robot if the field angle was too wide [16]. However, the SUS and interview support the idea that participants manipulate the robot while watching the omnidirectional camera image. Rather, some participants said that the wider angle of view is better than Double.

The omnidirectional camera image is thought to facilitate the perception of the local environment by allowing the viewpoint to be changed freely. As described by Participant 5, it is possible to freely see the omnidirectional camera image, so that the direction of the robot disappears. Therefore, rotation of the robot itself is unnecessary, and it is considered that the robot should move in parallel to the left and right.

7.3 Augment human perception with vertical actuation

We suggest that it is possible to expand human ability using telepresence robots. It may be possible to manipulate the operators' feelings and their impression on their conversation partner by changing height freely. It has been pointed out in previous studies that height is a major factor that determines the impression of oneself. Therefore, a person who is short feels intimidated, a problem that it is difficult to dispel. Although it is possible to change your height by wearing equipment like an exoskeleton, this is difficult to arrange. ScalableBody is thought to be effective a method to solve this problem, although it is for remote communication.

7.4 Implications

The following implications are conceivable from our experimental results.

- Improving the impression of the operator compared to the conversation partner by matching the face or eye position with the partner. When designing a telepresence robot, it should be possible to change its height to match the height of the partner. By adjusting your height, you can make your partner feel closer to you and reduce your sense of intimidation.
- Usability can be improved by texture-mapping the image of the omnidirectional camera to a sphere. Moreover, since it becomes possible to look around and move freely with an omniwheel robot, the robot's orientation disappears. Therefore, it is not necessary to rotate the robot, but rather an operation that enables parallel translation should be implemented.

7.5 Limitations

Our system has some limitations.

First, because of weight limitations, there is a performance limit to the camera and/or display. When the robot has too large a screen or camera, the vertical actuation mechanism breaks when the robot is changing its height or moving.

Second, the display sways when the body is extended, since its mounted on the top of the robot. We noticed that the swaying pattern and natural frequency differ depending on the height. This may cause motion sickness, though image stabilization technologies can potentially resolve this issue.

Third, ScalableBody has a screen displaying the face of the operator installed only in one direction, so the partner does not know where the operator is looking. This may be solved by rotating the face of the operator on the display using a spherical or flexible display. Additionally, we need additional user study because of our study was limited in its conditions, the number of participants, its specific tasks, and a smaller display size than a typical telepresence robots.

8. CONCLUSION

ScalableBody is a telepresence robot designed to adjust its height dynamically and provide an omnidirectional view, allowing communication while watching a conversation partner's face in a remote place regardless of situation. In this paper, we reported our implementation of the vertical actuator for changing the robot's height and investigated its performance. By changing height according to the situation of the partner, it is possible to make conversation with the same eye level or face position as we meet and talk. In addition, taking the viewpoint of a very short or very tall person makes it possible to observe the world from a viewpoint that you cannot usually experience. The user study shows that matching the eye or face position creates a good impression of one's partner, enhancing the engagement between the partner and operator. Moreover, an omnidirectional camera view supports the usability of a telepresence robot for understanding the environment around the robot. ScalableBody is focused on the effect of communication of a human height. We have to investigate the effect of communication of a one's characteristic such as a figure, various non-verbal cues, and an expression in telepresence system.

9. REFERENCES

[1] Anybots 2.0 inc. - new qb2.0 webrtc. http://www.anybots.com/.

[2] Attending cscw | cscw 2016. https://cscw.acm.org/2016/attend/telepresence.php.

[3] Double robotics - blended learning / hybrid classroom. http://www.doublerobotics.com/education/.

[4] Remote attendance « chi 2016. https://chi2016.acm.org/wp/telepresence/.

[5] Should i attend a conference via a telepresence robot? http://spectrum.ieee.org/automaton/robotics/industrial-robots/attending-conference-via-telepresence-robot.

[6] Suitable technologies beam smart presence system. https://www.suitabletech.com/beam/.

[7] Suitable technologies beam smart presence system. https://suitabletech.com/beampro/.

[8] University of wisconsin telestroke network. http://www.uwhealth.org/telehealth/university-of-wisconsin-telestroke-network/20572.

[9] J. K. Burgoon and N. E. Dunbar. An interactionist perspective on dominancesubmission: Interpersonal dominance as a dynamic, situationally contingent social skill. *Communication Monographs*, 67(1):96–121, 2000.

[10] M. Desai, K. M. Tsui, H. A. Yanco, and C. Uhlik. Essential features of telepresence robots. In *2011 IEEE Conference on Technologies for Practical Robot Applications*, pages 15–20, April 2011.

[11] S. R. Fussell, R. E. Kraut, and J. Siegel. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, CSCW '00, pages 21–30, New York, NY, USA, 2000. ACM.

[12] S. Gauglitz, B. Nuernberger, M. Turk, and T. Höllerer. World-stabilized annotations and virtual scene navigation for

remote collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pages 449–459, New York, NY, USA, 2014. ACM.

[13] R. Gorawara-Bhat, M. A. Cook, and G. A. Sachs. Nonverbal communication in doctor-elderly patient transactions (NDEPT): Development of a tool. *Patient Education and Counseling*, 66(2):223–234, 2007.

[14] K. Higuchi and J. Rekimoto. Flying head: Head-synchronized unmanned aerial vehicle control for flying telepresence. In *SIGGRAPH Asia 2012 Emerging Technologies*, SA '12, pages 12:1–12:2, New York, NY, USA, 2012. ACM.

[15] H. Iwata. Big robot mk1. http://www.aec.at/postcity/en/big-robot-mk1/, 2015.

[16] S. Johnson, I. Rae, B. Mutlu, and L. Takayama. Can you see me now?: How field of view affects collaboration in robotic telepresence. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2397–2406, New York, NY, USA, 2015. ACM.

[17] S. Kasahara and J. Rekimoto. Jackin head: Immersive visual telepresence system with omnidirectional wearable camera for remote collaboration. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, VRST '15, pages 217–225, New York, NY, USA, 2015. ACM.

[18] D. A. Lazewatsky and W. D. Smart. A panorama interface for telepresence robots. In *Proceedings of the 6th International Conference on Human-robot Interaction*, HRI '11, pages 177–178, New York, NY, USA, 2011. ACM.

[19] M. K. Lee and L. Takayama. "now, i have a body": Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 33–42, New York, NY, USA, 2011. ACM.

[20] D. Leithinger, S. Follmer, A. Olwal, and H. Ishii. 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, pages 461–470, New York, NY, USA, 2014. ACM.

[21] I. Rae, L. Takayama, and B. Mutlu. One of the gang: Supporting in-group behavior for embodied mediated communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 3091–3100, New York, NY, USA, 2012. ACM.

[22] I. Rae, L. Takayama, and B. Mutlu. In-body experiences: Embodiment, control, and trust in robot-mediated communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1921–1930, New York, NY, USA, 2013. ACM.

[23] I. Rae, L. Takayama, and B. Mutlu. The influence of height in robot-mediated communication. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, HRI '13, pages 1–8, Piscataway, NJ, USA, 2013. IEEE Press.

[24] I. Rae, G. Venolia, J. C. Tang, and D. Molnar. A framework for understanding and designing telepresence. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, pages 1552–1566, New York, NY, USA, 2015. ACM.

[25] S. E. Scherer and M. R. Schiff. Perceived intimacy, physical distance and eye contact. *Perceptual and motor skills*, 36:835–841, 1973.

[26] D. Sirkin and W. Ju. Consistency in physical and on-screen action improves perceptions of telepresence robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '12, pages 57–64, New York, NY, USA, 2012. ACM.

[27] S. Takei, M. Iida, and T. Naemura. Kinereels: Extension actuators for dynamic 3d shape. In *ACM SIGGRAPH 2011 Posters*, SIGGRAPH '11, pages 84:1–84:1, New York, NY, USA, 2011. ACM.

[28] H. Tobita, S. Maruyama, and T. Kuzi. Floating avatar: Telepresence system using blimps for communication and entertainment. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 541–550, New York, NY, USA, 2011. ACM.