IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 20 April 2024, accepted 19 June 2024, date of publication 28 June 2024, date of current version 8 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3420724

RESEARCH ARTICLE

A Study of Spatial Perception in Virtual Space via Display and the Intervention Effects of Haptic Feedback

KOTARO OKADA[1], KAZUHIRO MATSUI[1], (Member, IEEE), KEITA ATSUUMI[1,2], KAZUHIRO TANIGUCHI[1,3], HIROAKI HIRAI[1], (Member, IEEE), AND ATSUSHI NISHIKAWA[1]

[1]Graduate School of Engineering Science, Osaka University, Toyonaka 560-8531, Japan
[2]Graduate School of Information Sciences, Hiroshima City University, Hiroshima 731-3194, Japan
[3]Faculty of Human Ecology, Yasuda Women's University, Hiroshima 731-0153, Japan

Corresponding author: Kazuhiro Matsui (matsui.kazuhiro.es@osaka-u.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP20K14693; in part by the Future Social Value Co-Creation Project, Osaka University; and in part by the Hattori Hokokai Foundation.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee for Research Involving Human Subjects at Graduate School of Engineering Science, Osaka University, under Application No. R3-17-1, and performed in line with the Declaration of Helsinki.

ABSTRACT In recent years, technologies for presenting users with virtual spaces have been rapidly developing. Some applications allow users to observe three-dimensional spaces via screens and influence those virtual spaces using certain devices. In such cases, it is necessary not only to rely on the presentation of depth information but also to adapt to depth in the virtual space to completely execute tasks. In this study, we investigated the spatial perception characteristics of commonly used two-dimensional (2D) displays and advanced spatial reality (SR) displays with spatial presentation capabilities. The research participants performed tasks of tracing a trajectory on a wall surface that changes angles in a three-dimensional virtual space. In addition, we examined the effect of haptic feedback intervention on spatial perception during task execution and its persistence. For 2D displays, we realized that haptic feedback improves task accuracy and that the effect persists even after the feedback is removed. However, this applied only to tasks under feedback conditions, and no broad effect on spatial perception was observed. By contrast, for SR displays, we realized that haptic feedback may have a detrimental effect on spatial perception. Moreover, we quantitatively proved that the use of SR displays improves spatial perception accuracy compared with 2D displays and that the relationship between the line of sight and display angle is critical for spatial perception. In conclusion, the following two points are inferred from this study. (1) To improve spatial perception, it is necessary to consider methods that directly intervene in the body schema and the peripersonal space in the future. (2) Feedback by multiple modalities is not necessarily effective in presenting information on virtual space and obtaining spatial perception.

INDEX TERMS Haptic feedback, multimodal, multiple modalities, spatial perception, spatial reality display, virtual reality, visual feedback.

I. INTRODUCTION

With advances in computer graphics (CG) technology in recent years, technologies for presenting users with virtual spaces have been rapidly developing. Applications that allow users to observe three-dimensional (3D) spaces via screens and influence those virtual spaces through some devices are prevalent in various aspects of our lives, ranging from entertainment purposes, such as games, to industrial uses, such as computer-aided design. Virtual reality technology,

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaogang Jin.

which enables immersive experiences in virtual spaces via head-mounted displays (HMDs), has also advanced, allowing users to experience virtual spaces in 3D, which were previously only perceivable via two-dimensional (2D) displays. In addition to HMD-based methods, full-surround 3D display technologies, such as Cave automatic virtual environment (CAVE) [1] and Telexistence wide-angle immersive stereoscope (TWISTER) [2], which incorporate motion capture systems, have been used as interactive communication tools that allow users to act upon real spaces in remote locations. For instance, the da Vinci surgical robot (Intuitive Surgical, Inc.) has a system in which images captured by a stereo 3D camera are provided to the surgeon through a 3D monitor on a console located away from the patient, and the surgeon controls the surgical robot via 3D images using a controller on the console [3].

In addition, ELF-SR1 Special Reality Display (SONY, hereinafter "SR display") is an example of a device that presents 3D images. The SR display is capable of reproducing images in 3D space by allowing viewers to see 3D-CG images with approximately 4-K resolution in stereoscopic vision without the need for glasses. The display surface is covered with micro-optical lenses, which are a type of lenticular lens, and the lens structure allows different images to be projected to the viewer's left and right eyes. Moreover, the SR display is equipped with an infrared camera on its top, which tracks the 3D position of the viewer's eyes. Based on these coordinates, images suitable for the left and right eyes are rendered and projected onto the SR display according to the surface lens structure, enabling a natural 3D imaging experience [4].

Despite the promising applications of virtual space presentation technology in various fields, the introduction of specialized, expensive, and large equipment is often required for advanced 3D space presentations, and such technology is not yet widely disseminated to the public. Displays that only possess 2D planar rendering capabilities are commonly used as interfaces with electronic devices; however, they cannot present depth information. Various methods have been devised to present depth information on 2D displays, such as those using perspectives [5], blur effects [6], and parallax effects [7], [8]; however, they still have limitations with visual information alone. Thus, considerable research has been conducted on methods to complement depth information by presenting it through information other than visual cues. For example, Clare et al. [9] verified that looming sounds and changes in volume can lead to depth perception that enhances visual information.

However, in cases such as endoscopic surgery through displays, it would be more desirable to intuitively perceive spaces without consciously thinking about depth as if one were observing the surgical site [10]. Thus, it is necessary not only to rely on the presentation of depth information but also to adapt to depth in the virtual space to fully execute tasks. Therefore, in this study, we investigate the characteristics of spatial perception on two types of displays and analyze the effects of haptic feedback using linear vibration actuator as an aid or intervention in-depth perception, as well as the intervention effect persistence. Particularly, we first synchronize the real-world workspace with the virtual space and evaluate task accuracy within the workspace through a 2D display to investigate spatial perception. The effect of haptic feedback on spatial perception is then examined. Spatial perception aids using haptic feedback have been reported in the literature, such as Yoshimoto et al.'s haptic augmented reality navigation system using electrical stimulation [11]; however, these studies have merely demonstrated the auxiliary effects of feedback and do not consider feedback removal. We further investigated the extent to which the effects of adaptation to spatial perception persist after removing the haptic feedback, i.e., the intervention effect persistence. In addition, to verify the differences in the effect of haptic feedback on spatial perception through visual depth presentation, validation using the aforementioned SR display is performed.

II. MATERIALS AND METHODS

A. EXPERIMENTAL ENVIRONMENT

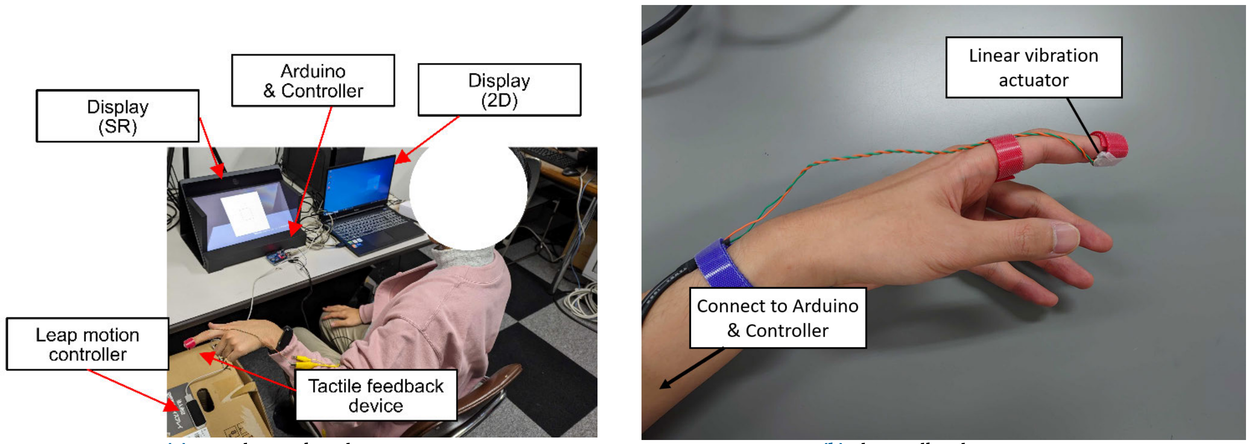
Based on our previous research [12], [13], [14], we constructed the experimental environment. Moreover, compared to previous research, we expanded the experiment by adding participants and completely reanalyzed the data.

Fig. 1 shows the experimental environment. Participants were seated in a chair, and they executed tasks while focusing on the workspace projected onto the display in front of them. 2D and SR displays were used. A participant did not use the 2D and SR displays simultaneously, and sufficient break time (approximately 15 min) was provided between switching displays so that the display position could be adjusted to the most visible position for a participant each time.

In this experiment, we used a leap motion controller (Leap Motion, Inc.) and an infrared sensor to measure the fingertip position and synchronize it with a hand model in virtual space. The controller comprised three infrared light-emitting diodes (LEDs) and two infrared cameras, which captured the fingers illuminated by the infrared LEDs and measured the 3D position of the fingers in space through image analysis.

In the experimental environment, we placed the sensor on a stand to the left of the participant. The participant's hand model was displayed within the workspace shown on the front display according to the fingertip data obtained from the sensor, and the space on the stand where the sensor was placed was virtually synchronized with the workspace. In addition, by positioning the sensor outside the participant's direct line of sight (LOS), the perception of fingertip position in the real space was limited to the participant's proprioception and the imagery on the front display alone.

For depth perception feedback, vibratory stimulation was applied to the fingertips. As shown in Fig. 1, a linear vibration actuator was attached to the participant's left index finger. Based on the spatial relationship between the objects in the workspace and the hand model, vibration control was performed using an Arduino UNO microcontroller (Arduino).



(a) Experimental environment (b) Linear vibration actuator

FIGURE 1. Linear vibration actuator for haptic feedback and experimental environment. Notably, the 2D and SR displays were moved in front of a participant according to the experimental environment.

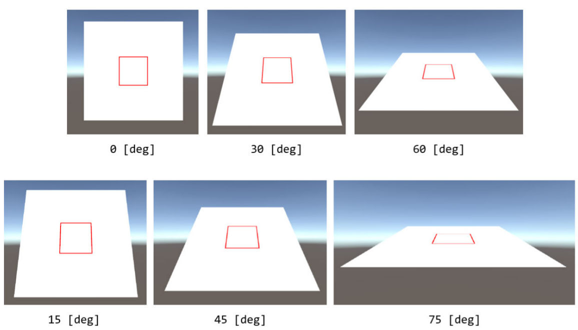


FIGURE 2. Virtual wall in virtual space.

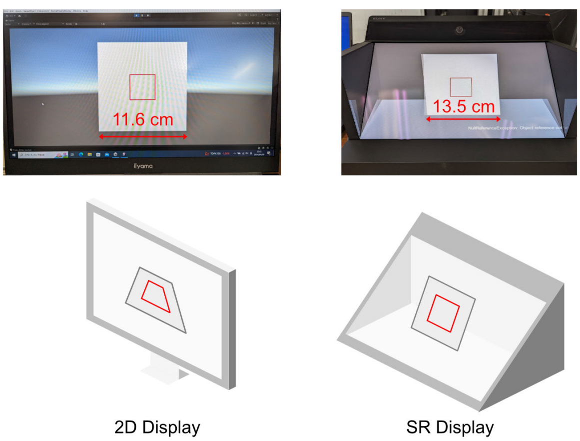


FIGURE 3. Appearance and size of each display.

We built a virtual space as the workspace using the Unity game engine (Unity Technologies) to integrate and control the displays, sensors, microcontrollers, and feedback devices.

B. EXPERIMENTAL TASK

In the aforementioned environment, we examined spatial perception mediated through a display and the assistance effect on spatial perception using tactile feedback.

Participants were asked to accurately trace a shape projected onto the virtual space on the display with their fingertips, which included depth information. In the virtual space, there was a square wall surface and a square marker drawn on that wall. This wall can function as an object with different depth information by rotating around its horizontal center axis that passes through the center of the plane. The wall is a square with each side measuring 30 cm, and the markers are squares with each side measuring 10 cm. However, these lengths represent the actual distances that fingers must move, and the lengths visible on the display may not necessarily correspond to these measurements. Fig. 2 presents an example of a wall object displayed on an actual display. The wall tilt angle varied according to the experimental conditions.

As aforementioned, the virtual space within the display shows the hand object acquired by the leap motion controller. The position adjustment was made such that the wall was directly above the controller.

For haptic feedback, contact detection with the wall was used. When the fingertip coordinates of the left index finger in the workspace penetrated beyond the wall surface, a linear vibrator attached to the fingertip vibrated to provide feedback on the wall position. Here, the vibration intensity was kept constant without any changes.

As aforementioned, we used two types of displays, 2D and SR displays, and the effects of haptic feedback were compared. Therefore, subsequent experiments for each feedback condition were performed with a change in the display.

The wall tilt angle was varied as 0 [deg], 30 [deg], 60 [deg], 15 [deg], 45 [deg], and 75 [deg]. In this study, we defined 0 [deg], 30 [deg], and 60 [deg] as *"standard angle conditions"* (SAC) and 15 [deg], 45 [deg], and 75 [deg] as *"additional angle conditions"* (AAC).

To investigate the effect of intervention on spatial perception by haptic feedback, task accuracies with and without feedback conditions were compared. After performing a task

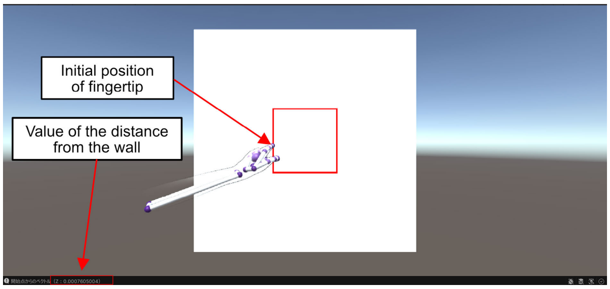


FIGURE 4. Overview of calibration procedure.

once under feedback-enabled conditions, the task was performed again under feedback-disabled conditions to compare accuracies before and after feedback intervention. In addition, if spatial perception itself was improved, task accuracy would have been better even when performing an inexperienced task. AAC was also tested to verify this effect(AAC not performed under feedback-enabled conditions).

C. EXPERIMENTAL PROTOCOL

The experiment was conducted as one set based on the following protocol.

1) Participants sat down so that the display position coincided with the midline of the trunk on the frontal plane. The spatial relationship with the leap motion controller was adjusted such that the shoulder joint position on the sagittal plane matched the virtual wall's origin, and the distance from the leap motion controller's origin on the frontal plane was set to allow the participant to trace the task trajectory without maximum elbow extension or flexion.
2) We set the initial position for the marker tracing task to the left-side center of the square marker, aligning the fingertip with the start position. Thus, we visually aligned the wall in both the planar and depth directions by looking at the on-screen value of the distance to the wall and adjusting it to as small a value as possible (less than 0.01). Fig. 4 shows the screen image of this system during calibration.
3) On the operator's signal, the tracing task began. We instructed the participants to trace the marker clockwise five times in 10 s with their left index finger while focusing on the display (Fig. 4), doing this continuously for five cycles. A countdown voice was played every second for 50 s, matching each cycle to help the participants keep pace with the task.
4) We ended the experiment after 50 s had elapsed. The passage of 50 s was indicated to the participants through the aforementioned countdown voice, enabling them to spontaneously recognize it.

This protocol varied according to the wall tilt angle, display, and feedback conditions.

Eleven healthy adults (23.3 ± 1.9 years old, all right-handed) participated in the study. They have no medical history of vision-related issues, except for the correction of myopia using glasses or contact lenses. The tasks are summarized in Table 1. The tasks were performed by all participants in order from A to F.

TABLE 1. Types of experimental tasks.

Task Name	Conditions		
	Display	Feedback	Wall tilt angle
Task A	2D	OFF	SAC, AAC
Task B	2D	ON	SAC
Task C	2D	OFF	SAC, AAC
Task D	SR	OFF	SAC, AAC
Task E	SR	ON	SAC
Task F	SR	OFF	SAC, AAC

For participants 1–6, AAC was not performed for tasks A and D; therefore, the number of participants who performed both SAC and AAC for tasks A and D was 5 (23.4 ± 1.5 years old). Although tasks A–C and D–F were performed consecutively, there were sufficient breaks between tasks C and D, and the seat position was moved in front of the SR display to minimize the influence of spatial perception from tasks A–C on tasks D–F.

The reason for using the left hand in the experiment was our belief that targeting the nondominant upper limb would enable effective measurement of the learning effect. Moreover, as all participants are right-handed, conducting the experiment targeting the left hand ensures no differences in the experimental conditions.

Participants were included in this study with the approval of the Research Ethics Committee for Research Involving Human Subjects at Graduate School of Engineering Science, Osaka University (R3-17-1). Written informed consent was obtained from all participants.

D. SIGNIFICANCE OF TASKS

Both experiments involved performing tasks A–C with a 2D display and tasks D–F with an SR display, with other variables kept constant. In task A/D, the task was first performed without feedback to verify spatial perception tendencies that were uninfluenced by feedback. In task B/E, the tendency of spatial perception when using feedback was examined. In task C/F, the feedback was turned off again to verify the persistent effects of the feedback. During this stage, by also conducting experiments on the angle conditions (AAC) that were not performed in task B/E, we verified whether the persistent effects of feedback appeared even for tasks that were not performed under the feedback application.

E. EVALUATION METHOD

In studies related to robot arms, the evaluation of control accuracy is often performed by measuring the error against the target trajectory or angle [15], [16]. Therefore, in this study, referencing the literature [17], we quantitatively evaluated the tracing movement accuracy by calculating the distance participant's motion deviated from the marker. This approach separated the assessment of the tracing accuracy of the wall surface marker from the ability to grasp the depth information of the wall itself.

First, we centered the measurement data obtained from the infrared sensor on the sensor's position. Therefore, it was necessary to recenter the origin to the marker's center and perform a coordinate transformation in which the plane containing the square marker represented the xy-plane and the axis perpendicular to the xy-plane through the square marker's center represented the z-axis. In addition, because the coordinate units were unique units of Unity, we normalized them based on the length of one side of the square marker during the coordinate transformation. Fig. 5 shows an example of the experimental results displayed after coordinate transformation for task A of participant 3.

Then, based on (1)–(4), we calculated the tracing accuracy error index $D_{xy}(n)$ for the square marker and the depth perception error index $D_z(n)$ for the transformed measurement data. n denotes the step number, and x_n, y_n, z_n are the fingertip coordinate data at each step. We calculated $D_{xy}(n)$ and $D_z(n)$ for each step (hereinafter, unless specifically needed, $D_{xy}(n)$ and $D_z(n)$ are denoted as D_{xy} and D_z, respectively). Therefore, the mean values of these error indexes for each task, $\overline{D_{xy}}$ and $\overline{D_z}$, can be considered planar and depth perception capability indicators, respectively.

$$D_{xy}(n) = |\sqrt{x_n^2 + y_n^2} - r_{\theta_n}| \tag{1}$$

$$D_z(n) = |z_n| \tag{2}$$

$$\theta_n = \cos^{-1} \frac{x_n}{\sqrt{x_n^2 + y_n^2}} \tag{3}$$

$$r_{\theta_n} = \begin{cases} |\frac{1}{2\cos\theta_n}| & (0 \le \theta_n < \frac{\pi}{4}, \frac{3\pi}{4} \le \theta_n \le \pi) \\ \frac{1}{2\sin\theta_n} & (\frac{\pi}{4} \le \theta_n < \frac{3\pi}{4}) \end{cases} \tag{4}$$

F. STATISTICAL TESTING METHOD AND EFFECT SIZE

The Steel–Dwass method was used for statistical analysis. In the experiment, we evaluated the average tracing errors $\overline{D_{xy}}$ and $\overline{D_z}$ for each participant as separate samples for each angle condition, but there was not necessarily any correspondence in the data between groups due to differences in angle conditions. Moreover, from the aforementioned equation ((1)–(4)), errors were evaluated as absolute values; thus, data distributions in each group were not assumed to follow a normal distribution. Therefore, we used the Steel–Dwass method [18], a nonparametric, unpaired multiple comparison test. The significance level was set to 5%.

Note: The analysis in a previous study [12] was conducted assuming a normal distribution; however, a review of the characteristics of the values presented in this study shows that it is reasonable to assume a non-normal distribution. Therefore, the method of analysis was changed as described above.

In comparisons between tasks, because the sample size was small for tasks A and D, in addition to significance testing

TABLE 2. Effect size for each task in tasks A, B, and C.

Tasks	Angle Condition	$\overline{D_{xy}}$ Effect Size r	$\overline{D_z}$ Effect Size r
A, B	SAC	0.19	0.35
B, C	SAC	0.04	0.03
A, C	SAC	0.25	0.29
A, C	AAC	0.12	0.11

using the Steel–Dwass method, effect size was also used for discussion. Effect size is a standardized measure of effect that does not depend on sample size. Significance testing with p-values is often reported to be affected by sample size. Thus, in this study, we combined p-values from significance testing and effect size for a comprehensive discussion. The effect size was calculated using (5).

$$r = \frac{Z}{\sqrt{n}} \tag{5}$$

where Z is the test statistic calculated during the Steel–Dwass procedure, and n represents the sample size.

Generally, for the Steel–Dwass test, an effect is considered negligible if $r \le 0.1$, small if $0.1 < r \le 0.3$, moderate if $0.3 < r \le 0.5$, and large if $r > 0.5$ [19].

III. RESULTS

A. UNDER THE 2D DISPLAY CONDITION (TASKS A–C)

First, the results under the 2D display condition are described. Fig. 6 shows the mean and standard deviation (SD) of $\overline{D_{xy}}$ and $\overline{D_z}$ for each angle condition in tasks A–C, averaged across all participants. The error bars represent SD.

From the above data, the following can be inferred.

Result 1-1 For tasks A–C, significant differences were confirmed between widely separated angles.

Next, the relationships among tasks A–C are described. Fig. 7 shows the relationship of $\overline{D_{xy}}$ and $\overline{D_z}$ among tasks A–C. SAC and AAC were separated, and the mean and SD were calculated for all samples from all participants for each condition.

Table 2 shows the effect sizes.

From the presented data, the following observations can be made.

Result 1-2 n tasks A, B, and C, there was statistically nonsignificant difference ($p > 0.05$) observed in $\overline{D_{xy}}$ between the tasks.

Result 1-3 For $\overline{D_z}$ under SAC, there was a significant difference ($p < 0.05$) between tasks A and B as well as between tasks A and C.

Result 1-4 Under AAC, no significant difference ($p > 0.05$) was observed either in $\overline{D_{xy}}$ or $\overline{D_z}$.

Result 1-5 The effect size between tasks A and B in SAC was 0.19 for $\overline{D_{xy}}$ and 0.35 for $\overline{D_z}$, with $\overline{D_z}$ showing the largest effect.

Result 1-6 The effect size between tasks B and C in SAC was small for both $\overline{D_{xy}}$ and $\overline{D_z}$.

Result 1-7 The effect size between tasks A and C in AAC was 0.12 for $\overline{D_{xy}}$ and 0.11 for $\overline{D_z}$, which were roughly the same.

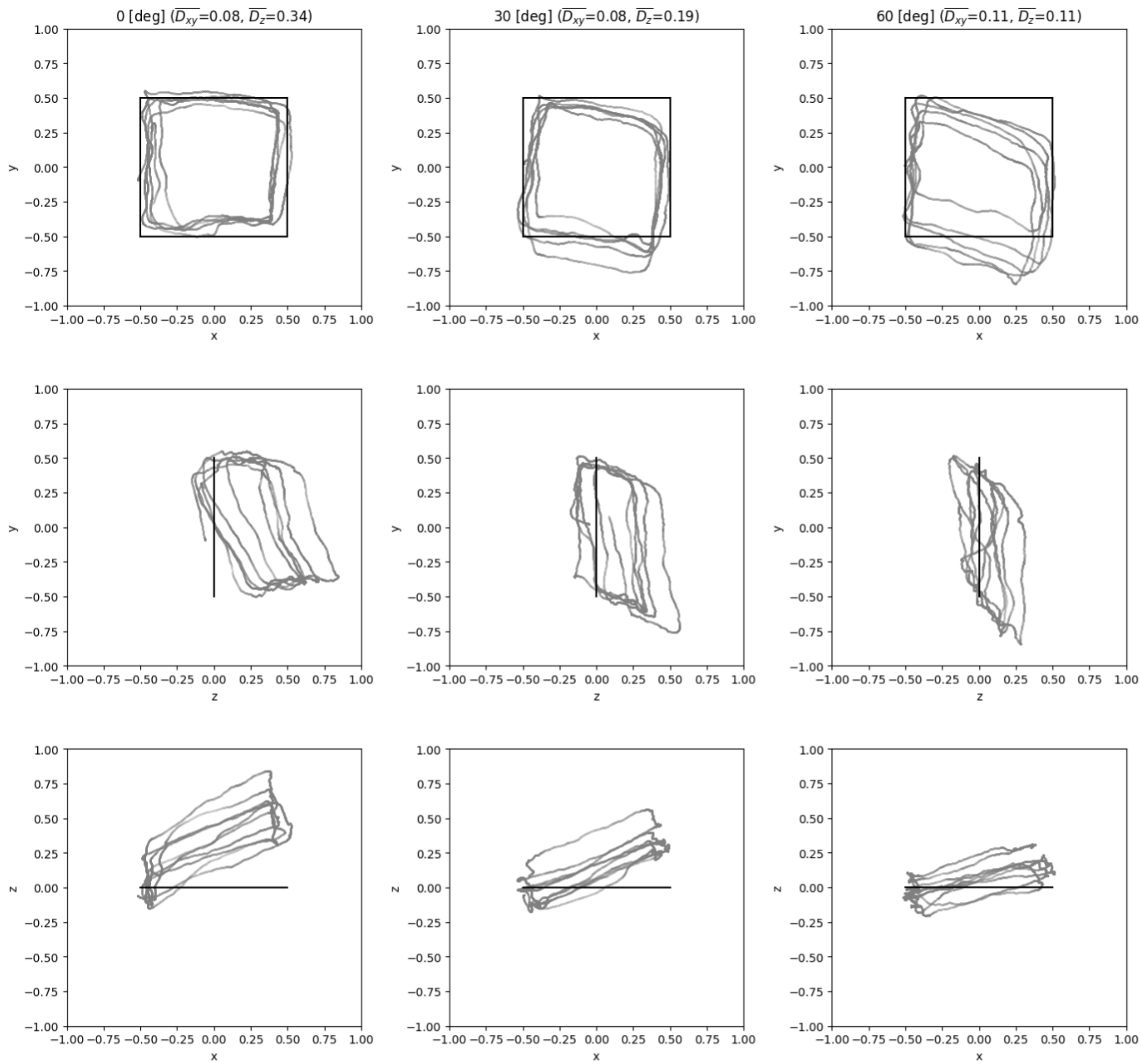


FIGURE 5. Participant 3's fingertip trajectory for task A.

B. UNDER THE SR DISPLAY CONDITION (TASKS D–F)

Further, the results under the SR display condition are described. Fig. 8 shows the mean and SD of $\overline{D_{xy}}$ and $\overline{D_z}$ in tasks D–F for all participants. The error bars represent SD.

From the above data, the following can be inferred.

Result 2-1 For the DEF tasks, significant differences were confirmed ($p < 0.05$) between widely separated angles for $\overline{D_{xy}}$.

Result 2-2 For tasks D–F, no significant differences were observed ($p > 0.05$) between angles for $\overline{D_z}$.

In addition, the relationships among tasks D–F are discussed. Fig. 9 shows the relationship of $\overline{D_{xy}}$ and $\overline{D_z}$ among tasks D–F. SAC and AAC were separated, and the mean and SD were calculated for all samples from all participants for each condition.

TABLE 3. Effect size for each task in tasks D–F.

Tasks	Angle Condition	$\overline{D_{xy}}$ Effect Size r	$\overline{D_z}$ Effect Size r
D, E	SAC	0.03	0.12
E, F	SAC	0.09	0.04
D, F	SAC	0.08	0.15
D, F	AAC	0.06	0.14

Table 3 lists the effect sizes.

From the data presented, the following observations can be made:

Result 2-3 For $\overline{D_{xy}}$, there were no significant differences ($p > 0.05$) between any of the groups, and the effect size was negligible.

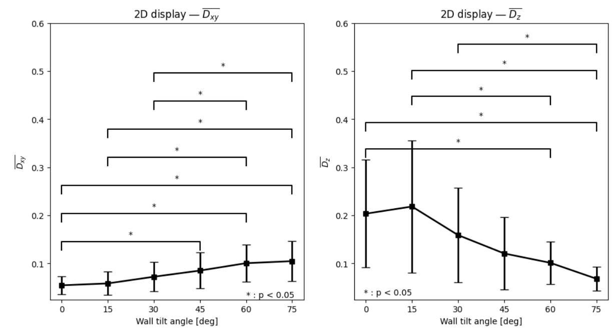


FIGURE 6. $\overline{D_{xy}}$, $\overline{D_z}$ **for each angle condition in tasks A–C. Significant differences confirmed by the Steel–Dwass method are marked with an asterisk.**

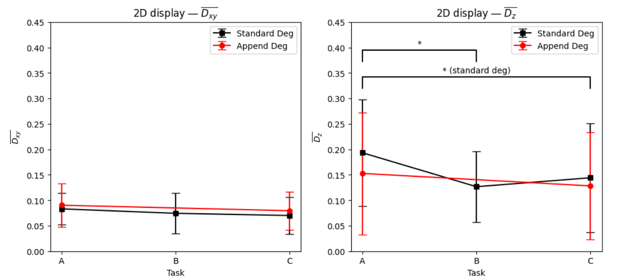


FIGURE 7. **Relationship of** $\overline{D_{xy}}$, $\overline{D_z}$ **among tasks A–C under SAC and AAC.**

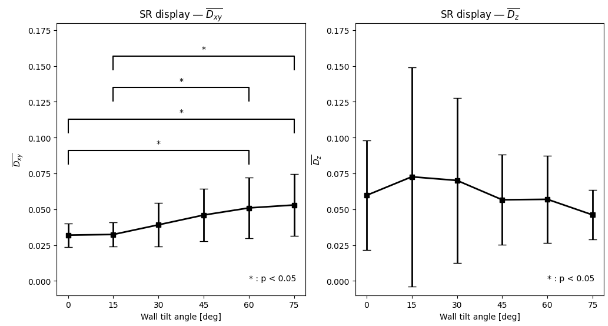


FIGURE 8. $\overline{D_{xy}}$, $\overline{D_z}$ **for each angle condition in Tasks D–F. Significant differences confirmed by the Steel–Dwass method are marked with an asterisk.**

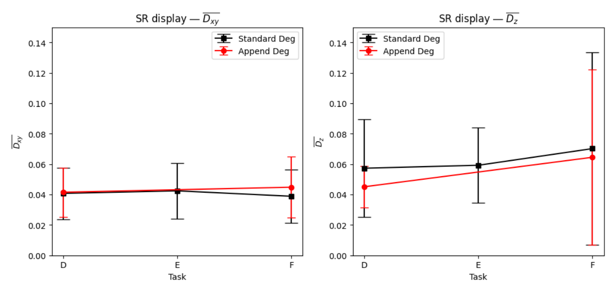


FIGURE 9. **Relationship of** $\overline{D_{xy}}$, $\overline{D_z}$ **among tasks D–F under SAC and AAC.**

Result 2-4 For $\overline{D_z}$ under SAC, no significant differences were observed ($p > 0.05$) between any of the groups. However, there was an increase in $\overline{D_z}$ with an effect size of approximately 0.12 between tasks D and E and approximately 0.15 between tasks D and F.

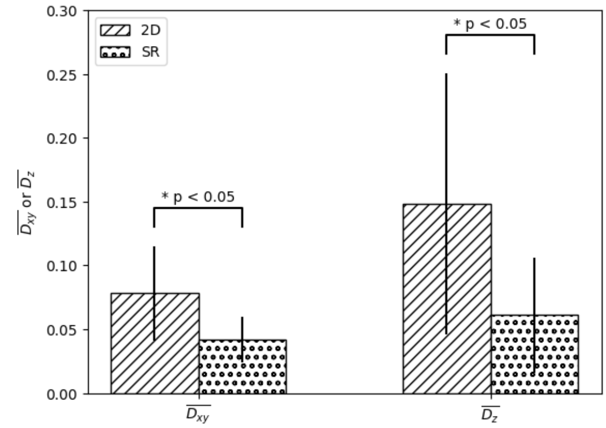


FIGURE 10. $\overline{D_{xy}}$ **and** $\overline{D_z}$ **for each display condition. The asterisks indicate significant differences detected using the Steel–Dwass method.**

TABLE 4. **Effect sizes for each display condition.**

Display Conditions	$\overline{D_{xy}}$ Effect Size r	$\overline{D_z}$ Effect Size r
2D, SR	0.54	0.58

Result 2-5 For $\overline{D_z}$ under AAC, no significant difference was observed ($p > 0.05$) between tasks D and F. However, there was an increase in $\overline{D_z}$ with an effect size of approximately 0.14.

C. COMPARISON OF RESULTS BETWEEN 2D AND SR DISPLAYS

Comparison results between the 2D and SR displays are described. Fig. 10 depicts a summary of $\overline{D_{xy}}$ and $\overline{D_z}$ for the 2D display condition (tasks A–C) and the SR display condition (tasks D–F).

The effect sizes for each display condition are summarized in Table 4.

From these results, the following insights can be deduced.

Result 3-1 For $\overline{D_{xy}}$, the SR display showed a decrease with an effect size of 0.54 compared with the 2D display.

Result 3-2 For $\overline{D_z}$, the SR display showed a decrease with an effect size of 0.58 compared with the 2D display.

IV. DISCUSSION

As aforementioned, $\overline{D_{xy}}$ and $\overline{D_z}$ can be considered indicators of a participant's ability to accurately grasp a plane and perceive depth (planar and depth perception capabilities, respectively). Larger values indicate larger errors, indicating less accurate perception. Taking this into account, we discuss each of the results.

A. TRENDS IN SPATIAL PERCEPTION WITH 2D DISPLAY

From **Result 1-1**, with the 2D display regardless of feedback conditions, as the wall tilt angle increases, $\overline{D_{xy}}$ increases and $\overline{D_z}$ decreases, i.e., the accuracies of planar and depth perception decrease and increase, respectively. Possible reasons for this include the following. 1) When the wall tilt angle is 0 [deg], meaning the wall is perpendicular to the floor, there is no object in the display to index depth, making it extremely difficult to grasp depth. As the wall tilts more, the wall object shows more spread in the depth direction, possibly serving as a cue for depth perception, thereby decreasing $\overline{D_z}$. 2) When the wall is perpendicular to the floor, the plane that contains the marker to be traced can be viewed perpendicularly, making the plane easier to grasp. As the wall tilts, the plane containing the marker to be traced appears distorted on the display, making it difficult to grasp planar perception, thereby increasing $\overline{D_{xy}}$.

In summary, tasks in the 2D display context showed that with increasing wall tilt angle, depth perception improved but planar perception became more challenging. This suggests that presenting objects with an obvious depth spread in virtual environments can enhance depth perception even without additional sensory feedback.

B. EFFECTIVENESS OF HAPTIC FEEDBACK WITH 2D DISPLAY

First, we consider the effectiveness of haptic feedback in the experimental setup. From Result 1-2, there is virtually no change in planar perception capability across tasks A–C. Meanwhile, Result 1-3 indicates that depth perception capability improves under feedback conditions. The improvement in spatial grasping capability with task progression is attributable to two factors: adaptation to the experimental system through repeated experience and adaptation through haptic feedback. Based on Result 1-5, $\overline{D_{xy}}$ shows some effect size between tasks A and B, attributable to adaptation through repeated experience rather than the feedback intervention effect. By contrast, $\overline{D_z}$ shows a larger effect size than $\overline{D_{xy}}$ under the same conditions, indicating that haptic feedback enhanced depth perception capability. Our experimental system provides simple feedback by vibrating the actuator when the fingertip penetrates behind the wall surface, directly reducing the error D_z related to depth but not the error D_{xy} in the plane. Therefore, haptic feedback is effective in improving depth perception capability but not planar perception capability.

Next, regarding the effect persistence, there is no significant difference between $\overline{D_z}$ in tasks B and C under SAC; also, as shown in Result 1-6, the effect size becomes negligible, indicating that depth perception capability improved by haptic feedback persists after removing the feedback. In addition, there is a significant difference between tasks A and C, which suggests an improvement in-depth perception capability even when compared under conditions without feedback.

As shown in Result 1-4, no significant difference in-depth perception capability was observed under AAC. In addition, based on Result 1-7, because the effect sizes are almost the same for $\overline{D_{xy}}$ and $\overline{D_z}$, it is thought that the decrease in mean values is due to adaptation through repeated experience, as mentioned earlier.

Thus, haptic feedback is suggested to be effective in enhancing depth perception capability and maintaining the intervention effect after removing feedback. However, it is not effective for feedback-inexperienced tasks, suggesting that spatial perception has not been improved by adaptation to space itself.

When considering interventions in spatial perception itself, research on various body enhancement techniques might serve as a reference. To improve spatial perception in the motor tasks of operating an avatar different from oneself, as in this study, it is important to update one's body schema and peripersonal space (PPS) [20]. Umezawa et al. [21] investigated whether changes occur in the body schema by attaching an extended finger to the real body using an obstacle avoidance task. Alternatively, Buck et al. [22] confirmed whether experiencing avatars of different body sizes changes PPS. In both studies, it was concluded that the body schema and PPS were not updated. This result resonates with our findings, suggesting that updating the body schema is key to improving spatial perception. In the future, it will be necessary to consider methods that directly intervene in the body schema and PPS.

C. TRENDS IN SPATIAL PERCEPTION WITH SR DISPLAY

From Result 2-1 and Result 2-2, we observe that regardless of feedback conditions, as the wall tilt angle increases, $\overline{D_{xy}}$ increases while $\overline{D_z}$ does not change. This means that as the wall tilt angle increases, the planar and depth perception accuracies decrease and remain unchanged, respectively. The differences from Result 1-1 under 2D display conditions are attributable to the angle between the LOS and display. As shown in Fig. 1, the SR display is placed at a slightly lower position than the 2D display. Thus, the SR display is viewed from a slightly downward-looking angle. Because the SR display allows objects to be viewed in 3D, it is possible that due to the angle, the wall may be presented perpendicularly to the LOS, similar to when the wall tilt angle is 0 [deg] under the 2D display, making depth perception difficult. The angle at which the wall appears to be presented perpendicularly may vary from person to person because of differences in stature, leading to significant variability and the absence of significant differences, attributable to the characteristics of the SR display, which allows stereoscopic viewing of images.

In summary, the SR display's ability to present objects in 3D led to variations in-depth perception accuracy due to the viewing angle, resulting in different spatial perception capabilities compared with the 2D display. This study highlights the importance of considering physical setup and ergonomics when designing virtual environments for tasks that require accurate spatial perception.

D. EFFECTIVENESS OF HAPTIC FEEDBACK IN SR DISPLAY

From Result 2-3, haptic feedback does not affect the ability to perceive plane surfaces, similar to the case with the 2D display.

Result 2-4 suggests that the use of haptic feedback can deteriorate depth perception ability by an effect size of approximately 0.15 and that this impairment persists, attributable to the high spatial representation capability of SR displays. SR displays reproduce space in 3D, making them an extremely intuitive output medium for spatial perception. Under the 2D display conditions, simple, incomplete feedback, which exhibited depth scale discrepancies, proved effective because the workspace information in 3D was reduced through the use of a 2D display, which was insufficient for spatial understanding. However, for the SR display that outputs the 3D workspace as is, the simple, incomplete feedback system used in our experiment may function as an error that leads to confusion in spatial comprehension. From Result 2-4, the fact that depth perception ability deteriorates even more when feedback is disabled in task F suggests that the feedback system used in our experiment affects depth perception. Contrary to the improved depth perception ability on 2D displays due to the maintained effect of feedback, in SR displays, the confusion caused by the feedback continues, implying that noise has been learned.

According to Wada et al. [23], improper auditory information can negatively affect time perception based on appropriate visual information. As such effects of sensory influence due to the coupling of multisensory stimuli have been reported, it can be considered that the decrease in spatial perception ability due to SR displays is caused similarly.

In addition, Result 2-5 highlights the differences in tendencies between 2D and SR displays. As mentioned earlier, in 2D displays, there was no intervention effect for tasks without previous feedback experience (AAC), and it did not improve depth perception capability. However, for AAC on the SR display, despite being inexperienced in task E, the depth perception capability deteriorated by an effect size roughly equivalent to that between tasks D and F in SAC before and after the feedback experience. Considering the performance differences between 2D and SR displays, presented below, the visual dominance in spatial perception in this system, which integrates visual and haptic feedback, is evident. Thus, haptic feedback plays only an auxiliary role in spatial perception. Based on the effectiveness of haptic feedback on 2D displays, when the primary visual information for spatial perception is incomplete, a plausible spatial perception can be constructed by integrating auxiliary haptic information. However, when sufficient visual information for perfect spatial perception is provided, attempts to learn and forcibly alter the already ''correct'' spatial perception through haptic feedback may adversely affect general spatial perception. In other words, simple, low-precision haptic feedback, such as that used in our experiment, may have intervention and persistent effects that inhibit spatial perception.

The reason of the aforementioned confusion may lie in scale discrepancies. As shown in Fig. 3, the markers projected onto the display are not necessarily 10 cm squares. The length of one side is approximately 3.9 cm in 2D display and approximately 4.5 cm in SR display. Instead, the virtual space is projected to ensure that the distance measured using the involved leap motion controller is 10 cm. The actual movement distance of the fingertip, inferred from proprioception, differs from the visual information provided by the avatar's fingertip movement distance. Although depth information is not directly rendered on the 2D display, it is understood as visual information on the SR display, allowing for the perception of distance from the fingertip to the wall surface. In this scenario, two depth scales simultaneously exist in the SR display: the distance inferred from proprioception (real-world scale) and the distance inferred from visual information (virtual-space scale). However, haptic feedback in this experiment contributes to the real-world scale. This misalignment between real-world and virtual-space scales may cause confusion. This phenomenon is considered an inherent issue of 3D displays, where different depth scale spaces can be simultaneously perceived by representing a realistic virtual space within the real world.

In summary, haptic feedback, such as that used in our experiment, conducted in the real world with a different depth scale than that of the involved virtual space, may have intervention and persistent effects that inhibit spatial perception in 3D displays where multiple depth scales simultaneously coexist.

One method to test this hypothesis is to experiment in a manner that avoids conflicts between visual and haptic information by aligning the scales of the wall surface and markers displayed on the screen with real-world scales.

In particular, to accurately represent the size of objects as stereoscopic images, information such as parallax, viewing distance, and vergence angle is necessary [24]. The SR display used in this experiment is presumed to focus on the display surface and may not accurately evaluate the vergence angle [25]. Therefore, images formed at positions other than the display surface may not be accurate. For a more accurate representation of stereoscopic images, methods such as using eye tracking to calculate vergence angles [26], subjectively adjusting the magnification of images based on some criterion, or applying a gain to the leap motion tracking information can be considered. The challenge ahead is to eliminate the differences in scale between visual and haptic information using these methods and verify the effectiveness of using a more accurate feedback environment.

Alternatively, consideration of the type of stimulus, for example, other than vibration, is necessary.

E. COMPARISON OF 2D AND SR DISPLAYS

From Result 3-1 and Result 3-2, for both planar and depth perception capabilities, the SR display has a higher accuracy by an effect size of $r = 0.5$ than the 2D display. Further, Fig. 10 shows that the SD is smaller for the SR display in both cases. This difference is thought to be because the amount of change in response to angle variation is smaller for the SR display (Fig. 7 and Fig. 9).

V. CONCLUSION

In this study, we investigated the spatial perception characteristics of two types of displays and examined the effects and persistent influence of interventions using haptic feedback to assist depth perception.

First, it was shown that the use of simple haptic feedback on a 2D display can improve task accuracy, and the effect of the intervention persists even after the feedback is removed. However, the improvement was limited to the feedback-experienced tasks and not an intervention on spatial perception itself. Further development of intervention methods is required. Meanwhile, under SR display conditions, simple haptic feedback could have negative effects on spatial perception.

Further, we quantitatively demonstrated that the use of advanced spatial presentation methods, such as SR displays, improves spatial perception accuracy over that of commonly used 2D displays. It was also shown that depth and planar perception abilities in virtual spaces are mainly due to the angular relationship between the gaze and the images of objects in the virtual space. 2D displays align the screens' normal direction with the LOS, whereas SR displays do not project onto a screen; thus, in addition to the angles of objects in virtual space, the direction of the gaze due to a participant's head position has an influence.

This study suggests that feedback from multiple modalities is not always effective in presenting information in virtual spaces and obtaining spatial perception. The combination of low-precision information can improve the overall accuracy of spatial perception; however, in situations such as those involving 3D displays, where multiple depth scales of information are present, low-precision elements can cause cognitive confusion. Because humans are visually dominant, combining other forms of feedback with visual cues requires caution.

The limitations of this study are the lack of examination using the dominant hand and the absence of a control group. In the future, it is necessary to address these limitations and increase the sample size for more detailed validation.

REFERENCES

[1] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, "Surround-screen projection-based virtual reality: The design and implementation of the cave," in *Proc. Seminal Graph. Papers, Pushing Boundaries*, vol. 2, 2023, pp. 51–58.

[2] S. Tachi, "Twister: Immersive omnidirectional autostereoscopic 3D booth for mutual telexistence," in *Proc. ASIAGRAPH*, vol. 1, 2007, pp. 1–6.

[3] J. Bodner, H. Wykypiel, G. Wetscher, and T. Schmid, "First experiences with the da vinci^{TM} operating robot in thoracic surgery," *Eur. J. Cardio-Thoracic Surg.*, vol. 25, no. 5, pp. 844–851, May 2004.

[4] (2023). *Spatial Reality Display Developer Site*. Accessed: Feb. 11, 2024. [Online]. Available: https://www.sony.net/Products/Developer-Spatial-Reality-display/en/

[5] K. Nakayama and S. Shimojo, "Experiencing and perceiving visual surfaces," *Science*, vol. 257, no. 5075, pp. 1357–1363, 5075.

[6] N. Asada and M. Baba, "A unified camera model of zoom, focus and IRIS parameters for camera-calibrated computer graphics," in *Proc. Comput. Graph. Imag.*, 2000, pp. 1–18.

[7] D. Frederick, "The effects of parallax scrolling on user experience in web design," *J. Usability Stud.*, vol. 10, pp. 87–95, Aug. 2015.

[8] J. Geng, "Three-dimensional display technologies," *Adv. Opt. Photon.*, vol. 5, no. 4, p. 456, 2013.

[9] C. A. M. Sutherland, G. Thut, and V. Romei, "Hearing brighter: Changing in-depth visual perception through looming sounds," *Cognition*, vol. 132, no. 3, pp. 312–323, Sep. 2014.

[10] K. Schwab, R. Smith, V. Brown, M. Whyte, and I. Jourdan, "Evolution of stereoscopic imaging in surgery and recent advances," *World J. Gastrointestinal Endoscopy*, vol. 9, no. 8, p. 368, 2017.

[11] S. Yoshimoto, Y. Kuroda, M. Imura, Y. Kagiyama, and O. Oshiro, "Development of a spatially transparent electrotactile display and application to hand tool navigation," *Trans. Jpn. Soc. for Med. Biol. Eng.*, vol. 49, no. 1, pp. 54–61, 2011.

[12] K. Okada, K. Matsui, K. Atsuumi, K. Taniguchi, H. Hirai, and A Nishikawa, "Research on spatial perception in virtual space via display-verification of intervention effect on spatial perception by tactile feedback," in *Proc. Robot. Mechatronics Conf.*, 2023, pp. 2–16.

[13] K. Okada, K. Matsui, K. Atsuumi, K. Taniguchi, H. Hirai, and A. Nishikawa, "Study of spatial perception in virtual space-comparison of the two types of displays and verification of intervention effect by tactile feedback," in *Proc. 28th Annu. Conf. Virtual Reality Soc. Jpn.*, 2023, pp. 1–14.

[14] M. Okuno, K. Matsui, T. Shimoshiro, K. Atsuumi, K. Taniguchi, H. Hirai, and A. Nishikawa, "Analysis of human depth perception for two-dimensional visual information and significance of tactile feedback," in *Proc. 39th Annu. Conf. Robot. Soc. Jpn.*, 2021, pp. 1–4.

[15] O. M. Omisore, S. Han, Y. Al-Handarish, W. Du, W. Duan, T. O. Akinyemi, and L. Wang, "Motion and trajectory constraints control modeling for flexible surgical robotic systems," *Micromachines*, vol. 11, no. 4, p. 386, Apr. 2020.

[16] M. Hofer, L. Spannagl, and R. D'Andrea, "Iterative learning control for fast and accurate position tracking with an articulated soft robotic arm," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 6602–6607.

[17] C. Ikeda, S. Nakajima, T. Ohyanagi, S. Goto, and Y. Sengoku, "The developmental tendency about drawing performance and handwriting motion," *Jpn. J. Occupational Therapy Pediatrics*, vol. 4, no. 1, pp. 39–47, 2016.

[18] R. G. D. Steel, "Some rank sum multiple comparisons tests," *Biometrics*, vol. 17, no. 4, p. 539, Dec. 1961.

[19] G. M. Sullivan and R. Feinn, "Using effect size—Or why the P value is not enough," *J. Graduate Med. Educ.*, vol. 4, no. 3, pp. 279–282, Sep. 2012.

[20] G. Ganesh, T. Yoshioka, R. Osu, and T. Ikegami, "Immediate tool incorporation processes determine human motor planning with tools," *Nature Commun.*, vol. 5, no. 1, p. 4524, Jul. 2014.

[21] K. Umezawa, Y. Suzuki, G. Ganesh, and Y. Miyawaki, "Bodily ownership of an independent supernumerary limb: An exploratory study," *Sci. Rep.*, vol. 12, p. 2339, Sep. 2022.

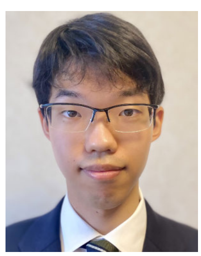
[22] L. E. Buck, S. Chakraborty, and B. Bodenheimer, "The impact of embodiment and avatar sizing on personal space in immersive virtual environments," *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 5, pp. 2102–2113, May 2022.

[23] Y. Wada, N. Kitagawa, and K. Noguchi, "Audio–visual integration in temporal perception," *Int. J. Psychophysiology*, vol. 50, nos. 1–2, pp. 117–124, Oct. 2003.

[24] N. Wade and M. Swanston, *Visual Perception: An Introduction*, 3rd ed. London, U.K.: Psychology Press, 2012.

[25] (2023). *Sony Corporation—-Spatial Reality Display | Design | Scene*. Accessed: Ap 3, 2024. [Online]. Available: https://www.sony.net/Products/Developer-Spatial-Reality-display/en/develop/Design/Scene.html

[26] I. T. C. Hooge, R. S. Hessels, and M. Nyström, "Do pupil-based binocular video eye trackers reliably measure vergence?" *Vis. Res.*, vol. 156, pp. 1–9, Mar. 2019.



KOTARO OKADA received the B.E. degree from Osaka University, in 2023, where he is currently pursuing the master's degree with the Graduate School of Engineering Science.

KAZUHIRO MATSUI (Member, IEEE) received the B.H.Sc. degree from Hiroshima University, in 2008, and the M.E. and Ph.D. degrees from Osaka University, in 2010 and 2015, respectively. From 2010 to 2019, he was an Engineer with OMRON Healthcare Company Ltd. Since October 2019, he has been an Assistant Professor with the Graduate School of Engineering Science, Osaka University. He is an RPT. His current research interests include rehabilitation engineering, health technology, robotics, and extended reality. He is a member of the Japanese Society for Medical and Biological Engineering (JSMBE), the IEEE Engineering in Medicine and Biology Society (EMBS), the Virtual Reality Society of Japan (VRSJ), the Robotics Society of Japan (RSJ), the Japanese Physical Therapy Association (JPTA), the Institute of Systems, Control and Information Engineers (ISCIE), the Japanese Society for Clinical Biomechanics, and others.

KEITA ATSUUMI received the Ph.D. degree from Osaka University, in 2023. Since April 2007, he has been an Assistant Professor with the Graduate School of Information Sciences, Hiroshima City University. His current research interests include instrumentation and control systems, human–machine interface, and mechatronics. He is a member of Japan Society of Mechanical Engineers (JSME), The Society of Instrument and Control Engineers (SICE), and the Institute of Electrical Engineers of Japan (IEEJ).

KAZUHIRO TANIGUCHI received the Ph.D. degree from Osaka University, in 2008. Since April 2020, he has been a Professor with the Faculty of Human Ecology, Yasuda Women's University. His current research interests include wearable computers, human interfaces, and aesthetics of design. He is a member of Japan Society of Computer Aided Surgery (JSCAS).

HIROAKI HIRAI (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in mechanical engineering from Osaka University, Japan, in 1997, 1999, and 2004, respectively. He was an Engineer for electronic musical instruments with Yamaha Corporation. He was a Visiting Researcher with The 77 Lab, MIT. He is currently an Associate Professor with Osaka University. His current research interests include human motor control and human enhancement, in particular: for athletes, children, and patients. He is a member of RAS and EMBS.

ATSUSHI NISHIKAWA received the B.E., M.E., and Ph.D. degrees from Osaka University, in 1990, 1992, and 1995, respectively. From 1995 to 1996, he was a Visiting Researcher with the Institute for Robotics and Intelligent Systems, University of Southern California. From 1996 to 1997, he was an Assistant Professor with the Faculty of Engineering Science, Osaka University. From 1997 to 2005, he was an Assistant Professor with the Graduate School of Engineering Science, Osaka University, where he was an Associate Professor, from 2005 to 2010. From 2010 to 2019, he was a Professor with the Faculty of Textile Science and Technology, Shinshu University. Since April 2019, he has been a Professor with the Graduate School of Engineering Science, Osaka University. His current research interests include medical robotics, human interface, and computer-assisted surgery. He is a member of the Japanese Society for Medical and Biological Engineering (JSMBE), Japan Society of Computer Aided Surgery (JSCAS), the Robotics Society of Japan (RSJ), and the Virtual Reality Society of Japan (VRSJ).