ACTA IMEKO
ISSN: 2221-870X
March 2020, Volume 9, Number 1, 1 - 10

3D Head Pointer: A manipulation method that enables the spatial position and posture for supernumerary robotic limbs

Joi Oh[1], Fumihiro Kato[2], Yukiko Iwasaki[1], Hiroyasu Iwata[3]

[1] *Waseda University, Graduate School of Creative Science and Engineering, Tokyo, Japan*
[2] *Waseda University, Global Robot Academia Institute, Tokyo, Japan*
[3] *Waseda University , Faculty of Science and Engineering, Tokyo, Japan*

ABSTRACT

This paper introduces a novel interface "3D Head Pointer" for the operation of a wearable robot arm in 3D space. The developed system is intended to assist its user in execution of daily chores while operating the robot arm in parallel. Previous studies have demonstrated the difficulty in simultaneously controlling a robot arm and one's own hands. The proposed method combines head-based pointing and voice recognition to manipulate the position, orientation, and switching between these two modes of the robot arm. In a virtual reality environment, the positional instructions of the proposed system, as well as its usefulness, were evaluated by measuring the accuracy of the instructions and time required using a fully immersive head-mounted display (HMD). In addition, the entire system, including posture instructions with two switching methods (voice recognition and head gestures), was evaluated using an optical see-through HMD supposed to use in actual robot arms. The obtained results displayed an accuracy of 1.25 cm and 3.56° with a 20-s time span necessary to communicate an instruction. These results, along with the confirmed ability of operation the actual robot arms using an optical see-through HMD, show that voice recognition is an effective switching method compared to head gestures.

Section: RESEARCH PAPER

Keywords: VR/AR, Hands-free interface, Polar coordinate system, Teleoperation, SRL

Citation: Joi Oh, Fumihiro Kato, Yukiko Iwasaki, Hiroyasu Iwata, Acta IMEKO, vol. A, no. B, article C, Month Year, identifier: IMEKO-ACTA-A (Year)-B-C

Section Editor: name, affiliation

Received month day, year; **in final form**, month day, year; **Published** month year

Corresponding author: Paul P. L. Regtien, e-mail: paul@regtien.net

1. INTRODUCTION

In recent years, a considerable amount of research and development have been done on the use of supernumerary robotic limbs (SRLs) for "body augmentation." In previous studies, robotic technology, especially wearable robotics, was developed as prosthesis for rehabilitation purposes. SRL aims to provide its users with additional capabilities, enabling them to accomplish tasks that a person would otherwise be incapable of performing. In this respect, SRL is different from other types of existing wearable robots. A lightweight, sufficient torque, and highly maneuverable SRL developed by Vernonia et al. [1] is one such classic example. Such robots were aimed for usage in any context, helping individuals to perform both household chores and improving industrial productivity.

To effectively assist in daily chores (opening an umbrella, stirring a pot, etc.), users require an interface that indicates the target point location to the end effector of the SRL without requiring him/her to interrupt his/her actions. However, such a method has not yet been established. Parietti et al. [2,3] developed a manipulation technique in which the operator's movements were monitored by a robot, following which the robot arm correspondingly performed those movements. Iwasaki et al. [4] proposed an interface that allowed the operator to actively control the SRL by using the orientation of the face. Furthermore, Sasaki et al. [5] developed a manipulation method that enabled more complicated operations of the robot arm by using the user's legs as the controllers. Previous studies have overlooked the balance between ensuring motion freedom of the operator's limbs and providing detailed instructions to the SRL. There are further challenges with respect to multitasking in the context of daily life. Therefore, in this study, we propose a method for manipulating SRLs, so that two parallel tasks do not interfere with each other, and evaluate their usefulness.

In this study, we conducted a two-stage experiment. This section describes the hypothesis of the whole method, and Section 2 presents the method of position instruction along with its experimental results and discussion. In Section 3, we propose a manipulation method that includes postural instructions and present the experimental results and discussion. Next, a discussion of the two experiments is provided. Finally, Section 4 presents the limitations and comparisons with other similar methods, and Section 5 presents the conclusions.

The following two elements are considered essential for achieving daily support for parallel works:

1) Undisturbed movement of the operator's limbs
2) Indication of spatial position and posture

Till date, several hands-free interfaces have been proposed to satisfy requirement 1). Some of them are operated by the tongue [6], eye movement [7], and voice [8], and are used for either screen control or robot manipulation (or both). Methods to control robotic limbs with brain waves [9] are also being investigated.

However, this study focuses on requirement 2) mentioned above and the construction of a more intuitive instructional method. When the operator provides directions related to a 3D space location, he/she must accurately indicate the target point. The range of the field-of-view, within which a person can perceive the shape and position of an object, is as narrow as 15° from the point of view [10]. Hence, a compensatory action of directing the face and gaze in the instructional space is necessary for spatial position instructions. Therefore, this interface takes advantage of the compensatory action of turning the face and uses it as an instruction method.

Methods for using the head as a joystick have already been proposed. One method involve the manipulation of the head for instruction in a 2D plane, such as on-screen operations [11]. Another method involves switching between the vertical and horizontal planes by nodding toward the plane to be manipulated. This supplements the plane manipulation by the head so that only the head is used to manage the 3D space [12]. However, these methods do not use compensatory motion of the head as a manipulation technique.

2. PROPOSAL OF A METHOD FOR POSITIONING USING HEAD BOBBING

The task of turning one’s face can be used to instruct the radial direction of the target point in polar coordinates. In this section, we propose a pointing interface that combines head bobbing with the orientation of the face in a polar coordinate system. Head bobbing is a small back and forth motion of the head that does not interfere with the operator’s movement.

This research was performed using the standard morphology of a Japanese man, as recorded by Kouchi et al. [13]. According to these data, the range of head bobbing was determined as approximately 9.29 cm, which allows the operator to keep the zero-moment point in the torso of the body and operate a robot arm without losing balance. A doughnut-shaped area set with the innermost and outermost radii of 30 and 100 cm, respectively, around the operator was defined as an example of an SRL operating range [14]. The depth change factor by head bobbing was 70/9.29 = 7.53 or more. The range of motion that can be performed using head bobbing is much lesser compared with that of the arms.

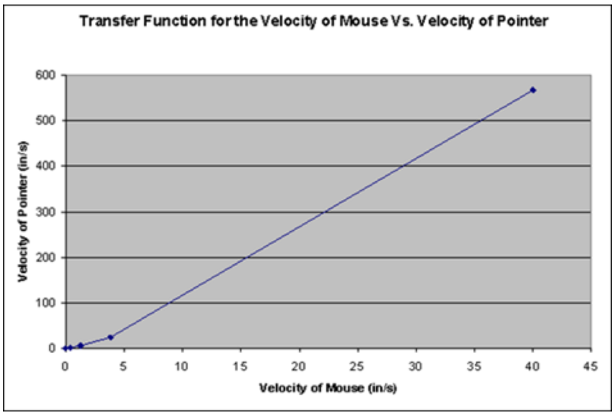


Figure 1. Microsoft's mouse-cursor speed-change settings [15].

Hence, preliminary experiments showed that at high magnification, the instructional accuracy of head bobbing was lower than that of other comparable methods. Additionally, the required instructions were shown to be longer. Therefore, an increase/decrease factor (IDF) that gradually changes the depth of the head bobbing task based on the head velocity was introduced. The IDF allows precise instructions while maintaining a high magnification. In this study, the IDF was constructed using the mouse-cursor change factor shown in Figure 1, set by Microsoft Windows [15].

2.1. Evaluation test with a fully immersive head-mounted display

This section examines the usefulness of the IDF and 3D head pointer as a whole. This study was conducted considering the previously developed robot arm proposed by Nakabayashi et al. [14] and Amano et al. [16], as shown in Figure 2. The arm has a reach of up to 1 m, and its jamming hand, shown in Figure 3, can be used as an end-effector to grasp an object with an error of up to 3 cm [16].

Figure 2. External view of the robotic arm used by proposed by Nakabayashi et al. [14] and Amano et al. [16].

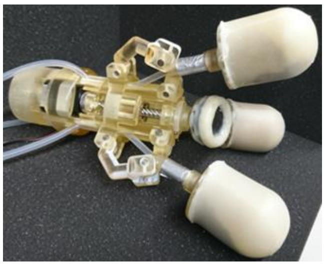


Figure 3. External view of the jamming hand.

Figure 4. 3D image of the head pointer operation.

Hence, the allowable indication error at the interface in this experiment was set to 3 cm. In this study, the validation was performed in a VR environment. The indication of radial direction by face orientation was measured from the front of the head-mounted display (HMD). The depth indicator was implemented by setting up a sphere with the operator at the center, as shown in Figure 4, and by changing the radius of the sphere by head bobbing.

The user can experience the proposed method using an HMD (HTC VIVE [17]). The experimental procedure is described as follows.

1) The participant wears the VIVE headset and grasps a VIVE controller in each hand holding them up in front of his or her chest, as shown in the right image in Figure 5. This is defined as the "rest position." The subject's avatar is displayed in the VR space, as shown in the left image of Figure 5.

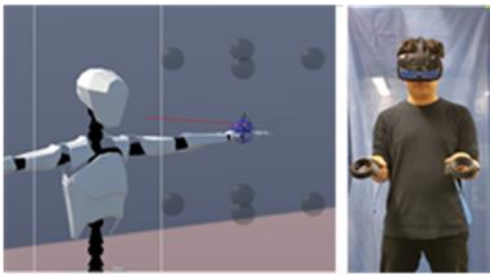


Figure 5. Experimental constitution of the interface operation (left: instructional goals and participants within the VR; right: participant wearing the HMD and holding the controllers).

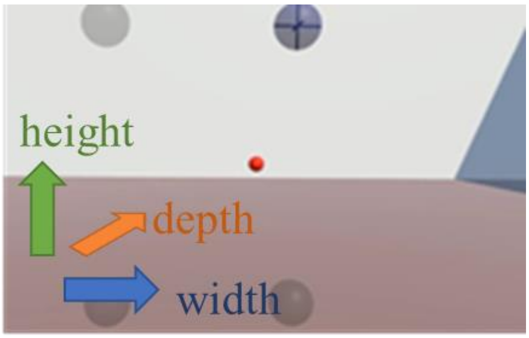


Figure 6. Subjective view of the subject's operation.

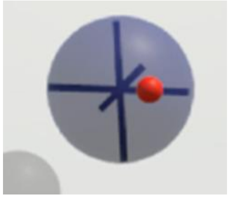


Figure 7. Target sphere and cursor visibility.

2) The 3D head pointer's control cursor (the red ball in the center of Figure 6) appears 65 cm in front of the eye. Simultaneously, the target sphere with a 10-cm diameter (the blue transparent sphere in the upper right corner of Figure 6) appears at any of the eight locations at a ±30-cm height, ±20-cm width, and ±20-cm depth, and positioned ±20-cm from the cursor.

3) The participant aligns the cursor to the center of the target sphere by using the 3D head pointer.

4) When the participant perceives that he/she has reached the center of the target sphere, they verbalize the completion of the instruction. As shown in Figure 7, the target sphere has a reference frame with its origin at the center of the sphere. Accordingly, the participant adjusts the position of the cursor.

5) Steps (1)–(4) are performed for all eight target sphere positions.

The above-mentioned procedure was performed by two groups of six participants each. The experiments were performed once under different conditions for each group. Table 1 shows the experimental conditions and group distribution. Group 1 was asked to perform the same tasks as described, but with a predefined time limit for instruction execution. Group 2 was asked to perform the experiment either with or without an IDF.

Figure 8 shows the relationship between head-bobbing speed and magnification. The "Not available IDF" is a condition in which the rate of change in depth due to head bobbing is fixed at 10 times without using the IDF.

Based on the aforementioned experiments, the usefulness of the 3D head pointer was evaluated using the average indication error condition (a) shown in Table 1, the relationship between the indication accuracy error and operation time in conditions (a)–(f), and the maximum arm sway of the subject measured by the VIVE controller according to condition (a).

Table 1. The experimental conditions and group distribution.

condition	Requirement	Group
(a)	No requirements	1, 2
(b)	2-s time limit for instruction	1
(c)	3-s time limit for instruction	1
(d)	4-s time limit for instruction	1
(e)	6-s time limit for instruction	1
(f)	8-s time limit for instruction	1
(g)	the rate of change in depth due to head bobbing is fixed at 10 times	2

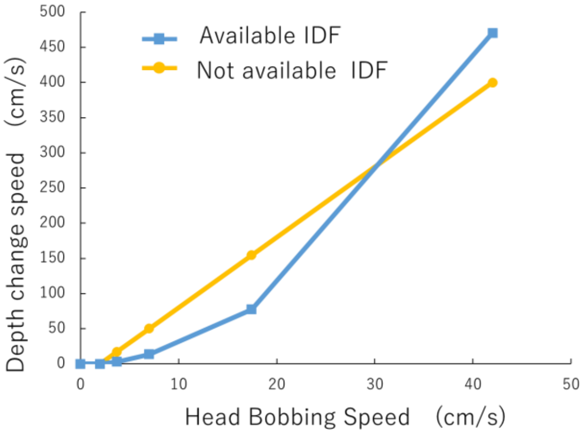


Figure 8. Change in head bobbing magnification with and without IDF.

At the same time, the usefulness of the IDF was tested by comparing the instructional error between conditions (a) and (g).

2.2. Results of and discussion on the fully immersive HMD

In this study, the Wilcoxon signed-rank-sum test was used to verify the significant differences between any two conditions. The Wilcoxon signed-rank-sum test is a nonparametric test used when the population does not follow a normal distribution. We obtained the difference of $Z_i = Y_i - X_i$ between the experimental values of two conditions X_i and Y_i performed on the i-th participant. Next, we arranged Z_i in order of decreasing absolute value and assigned rank R_i to the smaller value. The Wilcoxon signed-rank-sum test quantity of W is then calculated as follows:

$$W = \sum_{i=1}^{n} \emptyset_i \, R_i \,. \tag{1}$$

However, in this case, $\emptyset_i$ is calculated as

$$\emptyset_i = \begin{cases} 1(Z_i > 0) \\ 0(Z_i < 0) \end{cases} . \tag{2}$$

Significant differences were calculated by comparing test quantity W to Wilcoxon signed-rank-sum table [18]. In this experiment, instead of the table, the Excel statistics function (Microsoft Inc.) was used to calculate significant differences.

Table 2. Average instruction error.

Subject	Instructional error (cm)
1	1.20
2	2.50
3	1.54
4	2.19
5	2.41
6	1.06
7	1.06
8	0.882
9	0.757
10	0.905
11	0.695
12	0.668
Average	1.32

2.2.1. Indication Error

The instructional error of the distance from the center of the target ball to the control cursor was measured upon completion of the instruction.

This was done in VR by using an IDF-based 3D head pointer for 12 people, divided equally into two groups (1 and 2). The results are presented in Table 2.

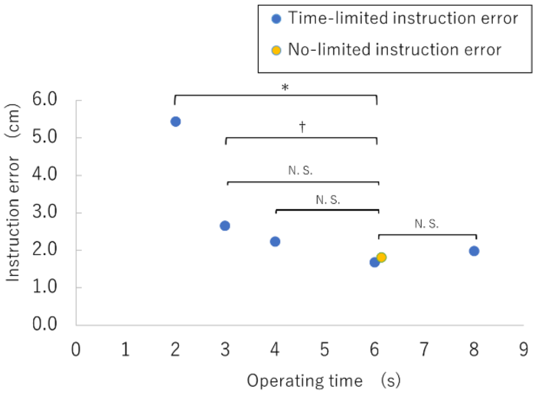
In this study, a jamming hand [16] capable of grasping an object with an error of up to 3 cm in target point indication, was used as a reference-index end effector. The average error of the instructions in this experiment was approximately 1.32 cm, with the highest instructional error of 2.5 cm. These results suggest that the indication error of the 3D head pointer is within the range of absorbable error in the case of grasping and manipulating an object with the specific end effector.

The standard deviation of the indication error is 0.65 cm, and the error varies widely from person to person. This result may be related to the familiarity level of each individual in the use of a VR space. The results must be validated by considering VR experience.

2.2.2. Change in indication error at each indication time

The experiment was conducted under conditions (a)–(f) for six members of Group 1. The relationship between the instruction error and instruction time is shown in Figure 9.

The average operation time under condition (a), with no time limit, was 6.2 s. When the operation time was limited, the indication error decreased rapidly with the increase in the time limit from 2 to 3 s. When the time was greater than 4 s, this error remained almost constant, regardless of the time taken. This suggests that the operation with the 3D head pointer itself was already completed by 4 s.



(Wilcoxon signed-rank , N=6, *:p<0.05, , †:p<0.1, N.S:not significant)

Figure 9. Instruction error per operating time in the evaluation test.

2.2.3. Maximum arm sway

Regarding the six participants in group 1, their maximum arm sway was measured from the movement of the VIVE controller while standing upright and compared to the maximum arm sway when the 3D head pointer was manipulated in condition (a). The results are presented in Figure 10.

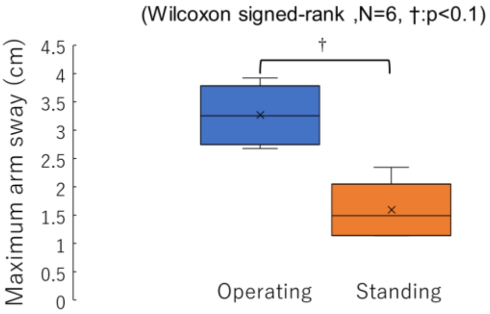


Figure 10. Maximum arm sway when standing upright and operating the 3D head pointer.

The comparison results showed that the maximum arm sway was found to be greater with a 3D head pointer. However, the Wilcoxon signed-rank-sum test did not show any significant difference in these two conditions ($N = 6$, $p < 0.1$), suggesting that the proposed method allows a user to continue performing regular arm motion, while performing according to the instruction indication. Because the proposed method requires visibility of the target space for performing tasks with SRL, multitasking is sometimes impossible, and interruption of the task being performed by the user is unavoidable. However, if the operator's hand position can be maintained while using the 3D head pointer, the interrupted task can be resumed quickly after instructions to the SRL; this is significantly more efficient than performing the two tasks separately.

2.2.4. Differences in Indication Error with and without IDF

We conducted the experiment under conditions (a) and (g) for the 6 members of group 2 and measured the instruction errors of the 3D head pointer and the depth-only instruction errors of head bobbing. The results are shown in Figures 11 and 12, respectively. The use of IDF reduced the average instruction error by approximately 77.6% for the depth instruction by head bobbing and approximately 67.0% for total error in three axes (x, y, z). Additionally, a significant difference was observed between the two conditions with and without IDF in the case of the Wilcoxon signed-rank-sum test ($n = 6$, $p < 0.05$). Therefore, it was confirmed that the introduction of the IDF greatly improved the accuracy and demonstrated its usefulness. Nevertheless, it is still necessary to verify whether the accuracy can be further improved with additional fine-tuning of the parameters related to the magnification change ratio.

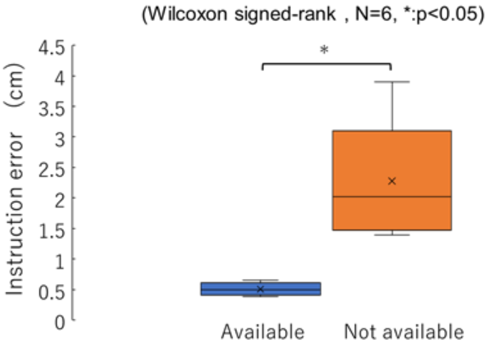


Figure 11. Depth error based on head bobbing with and without IDF.

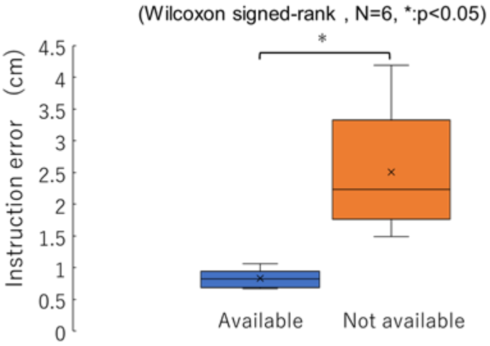


Figure 12. Total error in three axes due to 3D head pointer with and without IDF.

3. PROPOSAL OF COMBINING THE POSITIONING AND POSTURE INDICATION METHOD

The previous section showed the effectiveness of the position indications for SRL. However, without posture instructions at the interface, the SRL cannot perform complex daily chores (e.g., holding an umbrella at an angle to strong winds, pouring the contents of a bottle into a cup). Some objects can only be grabbed from certain directions. In this study, we propose a method of using the head for SRL to provide postural indications. Because it is difficult to provide stereotactic and postural instructions simultaneously with the head, we also propose "switching indication," which switches between position and posture indications.

3.1. Proposal of a posture indication method using isometric input

Figure 13 shows that the human head can rotate in three axes using Unity-Chan (humanoid model created by Unity Technologies Japan [19]). The use of the axes of the head rotation directly for SRL posture indication (yaw, pitch, and roll) facilitates intuitive instructions. However, the head has limiting angles of yaw, pitch, roll at (−60°)–(+60°), (−50°)–(+60°), and (−50°)–(+50°), respectively [20]. If the displacement of the head is used as an input device, the SRL cannot be instructed to posture at an angle beyond the limit angle of the head. In addition, according to the requirement definition (2) in Section1, if the head is moved more than 15°, the operation target will be out of the operator's effective field-of-view.

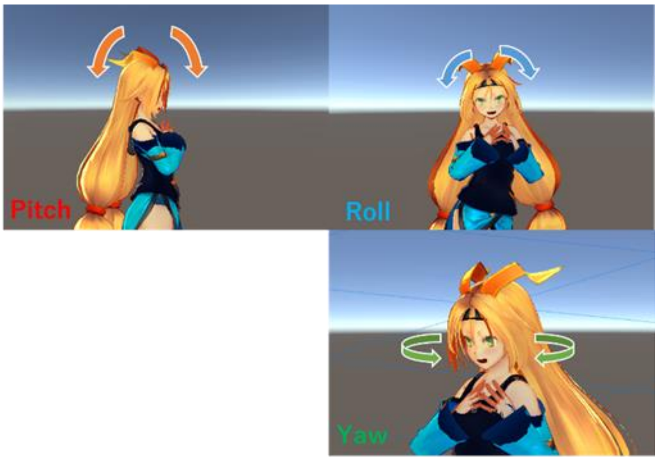


Figure 13. Three different rotation axes of the head.

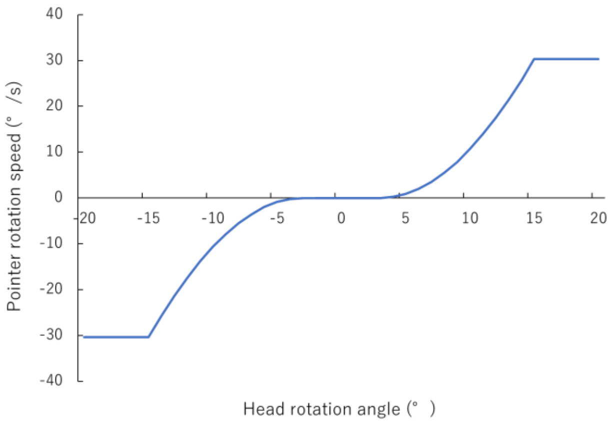


Figure 14. Relationship between head rotation angle and posture rotation speed.

In this study, the three-axis rotation of the head was used as an isometric input-device parameter that determines the rotational velocity of the pointer according to the rotation angle of the head [21]. The maximum input angle of the head was set to 15 °, which is the maximum limit angle of the effective field-of-view. To avoid incorrect input, head rotations of ≤3° were not detected as inputs. The changes in the rotational velocity were spherically interpolated using trigonometric functions. Figure 14 shows the relationship between the amount of rotation of the head and the rotation speed of the posture indicator. The reference angle for head rotation is the face direction when switching to the posture indication.

3.2. Proposal of a mode switching method using voice recognition

The increase of the number of body parts used for manipulation is undesirable because it leads to an increase in the body load. The switching method was constructed using head or voice. In this study, we propose two types of switching instruction methods and compare them in an evaluation test.

3.2.1. Voice-recognition-based switching indication method

A switching method based on voice recognition is less physically demanding and has less impact on the operator's limbs than physical operations. Table 3 lists the commands used for the voice indications.

3.2.2. Head-gesture-based switching indication method

We devised a method to switch between postural and positional instructions using head gestures. In this method, a "head tilt" motion was performed for switching from positional to postural instructions (top of Figure 15). For switching from posture instruction to position instruction, the "Head Bobbing" motion was performed (Bottom of Figure 15). This switching method was set up because the head tilt was not performed during position instruction, and conversely, head bobbing was not performed during posture instruction.

Table 3. Voice command list.

Voice command	Function
Indicate position	Switch from posture indication to position indication
Indicate posture	Switch from position instructions to posture instructions
Finish	Signals that the indication has been completed. (Used for evaluation tests)

Figure 15. Top: Switch to posture instruction;
Bottom: Switch to position instruction.

Because the user only has to indicate the operation mode to which he/she wants to switch, the head-gesture-based switching method requires little cognitive load, and switching can be done intuitively.

3.3. Evaluation test with optical see-through HMD

This section presents an evaluation of the usefulness of posture and switching instructions in the 3D head pointer, as well as the evaluation of the usefulness of the 3D head pointer in real space. To operate the SRL on a real machine, the tip of the SRL and target object must be visible. There are two ways to see the tip of the SRL on a real machine: by using a video see-through HMD or an optical see-through HMD [22]. The video see-through system may not be able to deal when the SRL malfunctions because of the delay in viewing the actual device. In this experiment, we constructed the proposed method using an optical see-through HMD (Hololens2 [23]) to evaluate the usefulness of the entire 3D head pointer. To provide posture instructions, the pointing cursor was changed from a red sphere to a blue-green bipyramid, as shown in Figure 16.

The indication of the radial direction based on face orientation was measured from the front of the HMD.

Figure 16. Pointer cursor corresponding to posture indication.

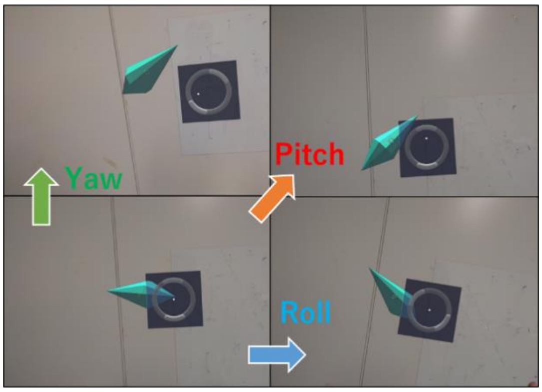


Figure 17. Auxiliary UI for posture instruction.

The depth indicator was implemented by changing the radius of the sphere by head bobbing, as described in Section 2.1. The amount of head rotation in the posture indication was determined by measuring the posture of the HMD. Compared to the position indication, it is difficult to know how much the operator inputs in the posture indication. To visually display the head-rotation amount of the user, the UI is displayed as shown in Figure 17, during posture instruction. The white point on the UI is based on the center and moves up, down, left, and right according to the amount of yaw and pitch fed as the input. The roll angle input is displayed as a white circle in the UI, and the circle rotates according to the amount of roll input. This UI allows the operator to visually understand how much he/she is moving the head to input. For speech recognition, we used Microsoft's Mixed Reality Toolkit [24].

In this experiment, we set up a pointing task for a target that appears in the air. The experimental procedure is described as follows.

1) The subjects stood upright while wearing the HMD and Bluetooth headset in a room with white walls.

2) The 3D head pointer point cursor (blue-green bipyramid in Figure 18) and the target (purple bipyramid in Figure 18) were displayed in front of the participant. The target appeared at a random position within 15° to the left and right of the subject's gaze direction, and within 30 to 100 cm in depth, as shown in Figure 19. The direction of the target was determined randomly among six directions: up, down, left, right, front, and back.

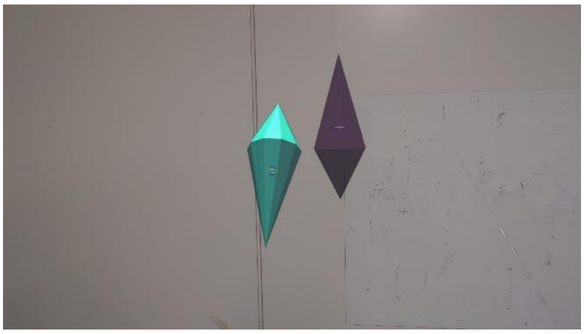


Figure 18. Cursor and target in the experiment.

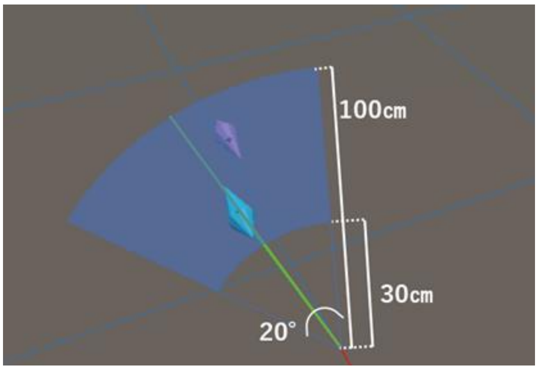


Figure 19. Area where the target appears (blue area in the Figure).

3) The participant moved the cursor to the same position and posture as the target using a 3D head pointer. When the subject perceived the completion of the operation, he/she verbalized "instruction complete" to the Bluetooth headset. Markers were displayed at the center of the cursor and the target position and rotation, as shown in Figure 18. These markers were always visible to the participant regardless of the position and posture of the cursor and target, and the operator relied on these markers for position and posture indications.

4) Steps 1)–3) were performed 12 times in succession in one experiment. The evaluation experiment was conducted under the following two conditions:
   A) Switching indications by voice recognition.
   B) Switching indications by head gestures.

   A verbal questionnaire was administered after operation completion.

The experiment was conducted using a total of 6 men and 6 women in their 20s and 30s, with the order of conditions A) and B) randomized. Procedures 1)–4) were performed at least 1 set as a proficiency time before conducting the experiment, and additional practice was conducted until the subject judged that he/she was proficient. Based on the above experiments, the usefulness of posture indication was verified according to the posture indication error and operation time. The usefulness of the switching instruction was verified by comparing the position error, posture error, and operation time in each condition. Finally, the usefulness of the 3D head pointer as a whole was verified based on the position indication error, posture indication error, and operation time. Section 3.4 describes these results.

3.4. Results and discussion on optical see-through HMD

3.4.1. Position indication error and posture indication error

The average values of the position and angle errors for each condition for the six subjects are shown in Figure 20. In this experiment, the tolerance was set assuming the same use of SRL as in the experiment discussed in Section 2.2.1, and the tolerance of the position indication was 3 cm. In the jamming hand of SRL, when reaching vertically to a cylindrical or spherical object, the success rate of grasping does not decrease if the angular error is within 30° [16].

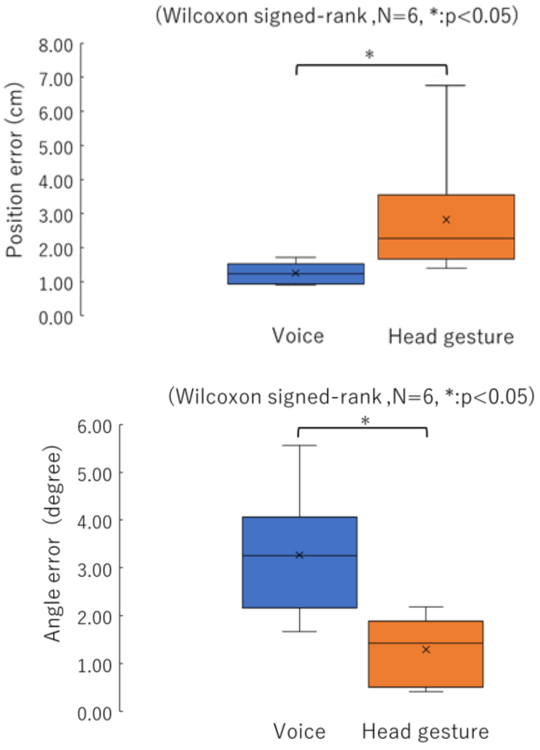


Figure 20. Top: Error in position indication;
Bottom: Error in posture indication.

The average position error of the instructions in this experiment was approximately 1.25 cm for the voice switching method and approximately 2.82 cm for the head gesture switching method, and a significant difference was observed between the two conditions in Wilcoxon signed-rank-sum test. This result depicts that the voice recognition method is more accurate in indicating the position. Since the instruction error of the position instruction alone in section 2.2.1 was 1.32 cm, this result shows that the head gesture switching method has a negative effect on the accuracy of the location instruction. The reason for the increased error in the head gesture is attributed to the shift in the position indication, because when the head is tilted to switch from position to posture instructions, the direction of the face moved accordingly. In addition, in the actual oral questionnaire, there were several comments that it was difficult to tilt the head without moving the direction of the face during the head gesture.

The average error of posture instruction was approximately 3.56° for the voice switching method and approximately 1.78° for the head gesture switching method, and a significant difference was observed between the two conditions in Wilcoxon signed-rank-sum test. This result shows that the accuracy of the posture indication is higher when switching by head gestures. This result was attributed to the fact that the posture instruction is an isometric input. In the posture instruction, as long as the head is rotated from the origin, the posture of the cursor will continue to rotate. If the operator uses head gestures, the instruction can be rapidly switched to stereotactic instructions, and consequently, the cursor posture can be fixed at the moment the continuously rotating cursor reaches the target posture. In the voice-based switching method, there is a delay between the time the voice command is uttered and the time the uttered voice is recognized as a command by voice recognition.

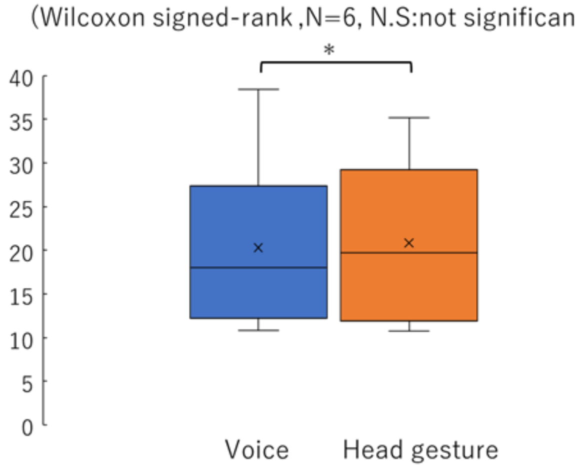


Figure 21. The mean values of the operation time.

Therefore, voice-based switching might have caused the cursor to rotate during the time when the user wanted to switch; however, a time delay occurred when the operation actually switched to the position instruction, resulting in a posture error. These results show that voice-based switching is effective in terms of position indication, and head-gesture-based switching is effective in terms of posture indication. Further, when switching by voice, the posture error increases; however, even for the subject with the largest error, the average error is 5.56°, which is within the acceptable range of 30°. However, the subject with the largest error in the case of head-gesture-based switching had an average position instruction error of 6.74 cm, which is far beyond the acceptable error of position instruction. Thus, we conclude that the voice-based switching method is more useful in terms of instructional accuracy, as all the values are within the acceptable error for the SRL assumed in this experiment.

3.4.2. Operation time

The mean values of the operation time for each condition for the six participants are shown in Figure 21. The average operating time was approximately 20.3 s for the voice switching method and approximately 20.8 s for the head gesture switching method. There was no significant difference between the two conditions in the Wilcoxon signed-rank-sum test. This indicates that there is no significant difference between the two switching methods in terms of operation time. When combined with the results of instructional accuracy, the results suggest that voice switching is more practical.

Moreover, the average operation time for position instructions alone, as discussed in Section 2.2.2, was 6.2 s. In this experiment, the operation time was higher by three times than the former owing to the addition of posture and switching indications. In addition, compared to the participant with the shortest average operation time, the participant with the longest average operation time had a three times longer operation time. When we asked the subjects about the cause of the increase in operation time in the verbal questionnaire, some of them said that the operation took longer when the posture indication did not go well. The causes of the delay for posture indication were as follows:

1) When giving posture instructions, rotation in unnecessary directions were mistakenly fed as input.
2) Compared to positional instructions, it is difficult to correct errors when they occur.

3) It is difficult to understand the posture of the cursor or target during rotation instructions.

Since postural manipulation of intentionally moving the neck along the three axes is not performed in daily life, the reason for cause 1) was verified. The reason for cause 2) was because, it took a long time to correct the error because the error had to be corrected by indicating the amount of displacement in the posture indication, as opposed to the position indication, which can directly specify the correct position when an error occurs. The reason for cause 3) was related to depth perception and size perception in peripheral vision. The permissible eccentricity for recognizing the position and shape of an object in the peripheral vision is 15° [10], but the perceptible eccentricity for depth is less than 12.5° and the perceptible eccentricity for size is less than 5° [25]. In addition, the accuracy of both depth perception and size perception decreased with eccentricity from the gazing point. Because the posture indication recognizes the posture of an object from changes in the size and depth of each side of the cursor or target, it required more visual information than the position indication. These reasons made it difficult to recognize the posture of the object when the face was turned away by up to 15° during posture manipulation.

3.4.3. Evaluation of the usefulness of the 3D head pointer as a whole

In the case of the voice switching method, the error in both position and posture indications was within the acceptable range, suggesting that the accuracy of the 3D head pointer is also effective for indications in real space through an optical see-through HMD. In terms of operation time, there was a large variation, and the indication time was not stable, indicating room for improvement. The improvement of the posture instruction, which is the most significant factor in the increase of operation time, is considered to be effective, and from the results of the oral questionnaire, the improvements to be made are as follows.

1) Construct with daily head movements.
2) Use isotonic input.
3) Does not leave the operator's gazing point.

Of these, 1) and 2) should be solved by using face orientation directly in the posture indication, but there is a potential problem of how to provide posture instructions by rotating the head beyond its limit movable angle. For the solution of 3), when the operator removes the gazing point from the cursor and target object in the posture indication state, the target object and cursor can be improved by continuing to display them in front of the operator in AR. However, cutting out an object in real space and displaying it in the AR is quite heavy.

In order to display AR in real time, it is necessary to devise a way to reduce the processing, such as detecting the mesh of objects in real space and displaying them.

4. DISCUSSION ON THE PRACTICAL APPLICATION OF 3D HEAD POINTER

In this section, we discuss the practical application of our proposed method.

The advantages of the 3D Head Pointer are clarified by comparing it with other manipulation methods. After that, we will discuss the concerns of using this interface in real life.

4.1. Comparison with Other Similar Methods

Based on the results of the previous section, the proposed method was compared with other similar methods.

a. *Physical controller*
   Some SRLs, such as those made by Vernonia [1], use a physical controller which is like gamepad with analog stick and some buttons as the method of operation. The advantage of the 3D head pointer is that its operation is more intuitive and easier to understand than that of a physical controller, and it can be operated hands-free.

b. *SRL manipulation method by foot*
   The proposed method can operate the SRL in any standing or seated position., compared to operating by legs [5]. However, manipulation of the legs can simultaneously indicate the position and attitude of the SRL. A short operation time is the main advantage of the leg operation.

c. *Head joystick and nodding to switch between the vertical and horizontal planes*
   Because the 3D head pointer uses compensatory motion of the head, it has less operational burden than methods that use the head as a joystick [11,12]. In contrast, the nodding method [12] allows for digital input from the head alone and may be used in conjunction with the 3D head pointer.

4.2. Limitation

In this study, we used tone recognition to give command instructions such as switching instructions, but voice recognition had the disadvantage of not being able to operate in a noisy environment or while the operator is having a conversation. Some prior examples of command-type instructions use gaze to provide command instructions [26,27]. The combination of pointing instructions with the head and gaze command instructions could provide a more flexible environment for SRL indications.

If there is a need to use SRL for complex or long movements in daily life, we have to register the movements and play them back. Registering and replaying behaviors require many commands, but the number of command-type instructions that can be intuitively memorized and selected is as few as six [28]. When building a system with more than seven commands, it is necessary to devise a way to remember commands, such as displaying a menu screen in the HMD.

5. CONCLUSIONS

In this study, we proposed a spatial position and posture indication interface for SRLs to improve work efficiency in daily life chore execution. We described the required functions for indicating spatial position and posture, and proposed a position indication method, 3D head pointer, which combines head bobbing type depth indication for spatial position and polar direction indication by face orientation. In a VR environment, evaluation tests of the 3D head pointer and IDF were conducted. The results showed that the 3D head pointer had sufficient accuracy without requiring the operator to interrupt his/her actions.

In addition, to provide not only position but also postural guidance by using a 3D head pointer, we proposed a postural guidance method using head rotation as an isometric input, and two types of switching guidance methods using voice recognition and head gestures. We also conducted a comparative study of two switching instruction methods using an optical see-through HMD and a test to evaluate the usefulness of the 3D head pointer as a whole. The results showed that the switching method based on voice recognition was effective in using the assumed SRL, and the 3D head pointer was confirmed accurate enough to be useful in operating the actual robot arms using an optical see-through HMD. These results provide useful knowledge for improving the SRL interface.

In the future, we intend to develop an intuitive posture instruction method that is not affected by compensatory head movements and incorporate a command instruction method that replaces voice recognition.

In the future, we further consider using SRL as an interface to naturally use it as a third arm in situations, such as banquets and construction sites, where an individual's hands are not sufficient.

ACKNOWLEDGEMENT

This research is supported by Waseda University GlobalRobot Academia Institute, Waseda University Green Comput-ing Systems Research Organization and ERATO Inami JIZAIBody Project.

REFERENCES

[1] C. Veronneau, J. Denis, L. Lebel, M. Denninger, V. Blanchard, A. Girard, J. Plante, Multifunctional 3-DOF Wearable Supernumerary Robotic Arm Based on Magnetorheological Clutches, IEEE Robotics and Automation Letters, Vol. 5, Issue 2, pp. 2546-2553, 2020.

[2] C. D. Parietti, H. H. Asada, Design and Biomechanical Analysis of Pernumerary Robotic Limbs, Proc. IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2012.

[3] H. H. Asada, F. Parietti, Supernumerary Robotic Limbs for Aircraft Fuselage Assembly: Body Stabilization and Guidance by Bracing, Proc. IEEE/ICRA, pp. 119-125, 2014.

[4] Y. Iwasaki and H. Iwata, Research on the Third Arm: proposal of a face vector interface for voluntary and intuitive control of a wearable robot arm, IEEE/RSJ Int. Conf. Intell. Robot. Syst., 2017.

[5] T. Sasaki, M. Saraiji, K. Minamizawa, M. Inami, MetaArmS: Body remapping using feet-controlled artificial arms, UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, pp. 140-142, 2018.

[6] S. G. Terashima, J. Sakai, T. Ohira, H. Murakami, E. Satho, C. Matsuzawa, S. Sasaki, K. Ueki, Development of a Tongue Operative Joystick - For Proposal of Development of an Integrated Tongue Operation Assistive System (I-to-AS) for Seriously Disabled People, The Society of Life Support Engineering, Vol. 24, Issue 4, 2012.

[7] R. Barea, L. Boquete, M. Mazo, E. Lopez, System for assisted mobility using eye movements based on electrooculography, IEEE Trans Neural Syst Rehabil Eng, Vol. 10, Issue 20, pp. 209-218, 2002.

[8] R. C. Simpson, S. P. Levine, Voice control of a powered wheelchair, IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 10, Issue 2, pp. 122-125, 2002.

[9] S. Nishio, C. I. Penaloza, BMI control of a third arm for multitasking," Science Robotics Vol. 3, Issue 20, 2018.

[10] T. Miura, "Behavioral and visual attention, Japan, CA: Kazama Shobo, 1996.

[11] R. Hasegawa, Device For Input Via Head Motions, Patents WO 2010/110411 Al, 2010.

[12] A. Jackowski, M. Gebhard, A. Gräser, A novel head gesture based interface for hands-free control of a robot, Proc. 2016IEEE Int. Symp. Med.Meas. Appl. MeMeA, 2016.

[13] M. Kouchi, M. Mochimaru, AIST Anthropometric Database, National Institute of Advanced Industrial Science and Technology, H16PRO 287, 2005.

[14] L. Drohne, K. Nakabayashi, Y. Iwasaki, H. Iwata, Design Consideration for Arm Mechanics and Attachment Positions of a Wearable Robot Arm, Proc.IEEE / SICE International Symposium on System Integration (SII), pp. 645-650, 2019.

[15] Windows Dev Center - Hardware, Pointer Ballistics for Windows XP, Available: http://archive.is/20120907165307/msdn.microsoft.com/en-us/windows/hardware/gg463319.aspx#selection-165.0-165.33, 2002.

[16] K. Amano, Y. Iwasaki, K. Nakabayashi, H. Iwata, Development of a Three-Fingered Jamming Gripper for Corresponding to the Position Error and Shape Difference, RoboSoft 2019, pp. 137-142, 2019.

[17] HTC VIVE, Available: https://www.vive.com/eu/product/vive/, 2011

[18] C. Zaiontz, Wilcoxon Signed-Ranks Table, Available: http://www.real-statistics.com/statistics-tables/wilcoxon-signed-ranks-table/, 2020.

[19] © Unity Technologies Japan/UCL, Unity-chan!, Available: https://unity-chan.com/, 2014.

[20] Committee on Physical Disability, Japanese Orthopaedic Association, Joint range of motion display and measurement methods, The Japanese Journal of Rehabilitation Medicine, Vol. 11, Issue 2, pp. 127-132, 1974.

[21] S. A. Douglas, A. K. Mithal, The ergonomics of computer pointing devices, Germany, CA: Springer, 1997.

[22] J. P. Rolland, R. L. Holloway, H. Fuchs, Comparison of optical and video see-through, head-mounted displays, Proc. of SPIE, 1994.

[23] Hololens2, Available: https://www.microsoft.com/en-us/hololens/buy

[24] Mixed Reality Toolkit, Available: https://hololabinc.github.io/MixedRealityToolkit-Unity/README.html

[25] A. Yasuoka, M. Okura, Binocular Depth and Size Perception in the Peripheral Field, the Journal of the Vision Society of Japan, Vol. 23, issue 2, pp. 103-114, 2011.

[26] M. Yamato, A. Monden, Y. Takada, K. Matsumoto, K. Tori, Scrolling the Text Windows by Looking, Transactions of Information Processing Society of Japan, Vol. 40, Issue 2, pp. 613-622, 1999.

[27] T. Ohno, Quick Menu Selection Task with Eye Mark, Transactions of Information Processing Society of Japan, Vol. 40, Issue 2, pp. 602-612, 1999.

[28] Y. Iwasaki, H. Iwata, Research on a Third Arm: Analysis of the cognitive load required to match the on-board movement functions, LIFE2018, Session. 2-4-1-2, 2018