N95-15984

A VIRTUAL WORK SPACE FOR BOTH HANDS MANIPULATION WITH COHERENCY BETWEEN KINESTHETIC AND VISUAL SENSATION

Masahiro Ishii, P.Sukanya, Makoto Sato

Precision and Intelligence Laboratory
Tokyo Institute of Technology
Nagatsuda 4259, Midoriku, Yokohama 227, Japan
Tel: 81-45-924-5050, Fax: 81-45-921-0898
E-mail: {mishii, sukanya, msato}@pi.titech.ac.jp

ABSTRACT

This paper is about constructing a virtual work space for performing any tasks by both hands manipulation. We intend to provide a virtual environment that can encourage users to accomplish any tasks as they usually act in the real environment. Our approach is using a three dimensional spatial interface device that allows the user to handle virtual objects directly by free hands and be feel-able some physical properties of the virtual objects such as contact, weight, etc. We have investigated the suitable conditions for constructing our virtual work space by simulating some basic assembly work, a Face-and-Fit task. Then select the conditions that the subjects feel most comfortable in performing this task to set up our virtual work space. Finally, we have verified the possibility to perform more complex tasks in this virtual work space by providing some simple virtual models then let subjects create new models by assembling these component models together. The subjects can naturally perform assembly operations and accomplish the task. Our evaluation shows that this virtual work space has potential to be used for performing any tasks that need hands manipulation or cooperation between both hands in natural manner.

KEYWORDS: Virtual reality, 3D modeling, cooperation between both hands, multi-modalities

INTRODUCTION

Recently, many three dimensional (3D) spatial interface devices have been proposed. However, each is appropriate for each kind of work[1][2][6][7]. Now, we still lack of the interface device that can immerse the user into the virtual work space, then allows him/her to perform any tasks as desire. For example, to create a new 3D model, it allows the user to grip, rotate or twist virtual models at any orientations arbitrarily by hands manipulation directly. To construct such the virtual work space, it is indispensable to consider the effective interaction communication between human and machine[8][9]. The multi-modal system is a concept which we adopt to provide information to the user in multi-sensory channels simultaneously as we usually get information in the real environment. Primarily, sight and touch are the senses that we have utilized. We use SPIDAR (SPace Interface Device for Artificial Reality) as the 3D spatial interface device to construct such the virtual work space. SPIDAR has been previously proposed by M.Sato et al.[3][4].

On the other hand, let us mention methods for forming a 3D model manually as we practise in the real environment. There are two basic methods:- Extraction method and Combination method[5].

- *Extraction method*
  It is the method that a new model is created by deforming the original model

- *Combination method*
  It is the method that a new model is created by combining one model with another model in an arbitrary orientation

Actually, to form a new model both by extraction method and by combination method, we are familiar with manipulation by cooperative between both hands. So the interface device that has a capacity of handle by both hands is needed to per-

C-2

form this task in the virtual work space.

Originally, SPIDAR has been developed for single hand manipulation. Since some kinds of work are performed by cooperative between both hands more effective than by single hand, here we have enhanced it for both hands manipulation.

This paper aims to construct a virtual work space primarily for performing a task such as forming a 3D model manually such mentioned above. Finally, we have evaluated it by simulating an environment for forming a 3D model by combination method. Then let subjects perform this task.

OVERVIEW OF SPIDAR

SPIDAR is a 3D spatial interface device previously proposed by M.Sato et al. Originally, it was developed for single hand manipulation. The general structure of it is shown in Figure 1. To perform a task with this interface, the user needs only to put his/her thumb and index finger into the provided caps. Each cap is held by four strings that are wound around pulleys attached with electrical motors at each corner of the cubic frame. The user can move both the thumb and the index finger arbitrarily. The motion of each finger that is the motion of each cap also is detected by the rotary encoders attached with the motors, so the position of each finger can be caculated[4]. By controlling the tension of the strings, we can provide force sensations to the user via the caps. Moreover, we can vary both the magnitude and the direction of forces arbitrarily in the range from 0N to 4N with a step of 0.016 N [4].

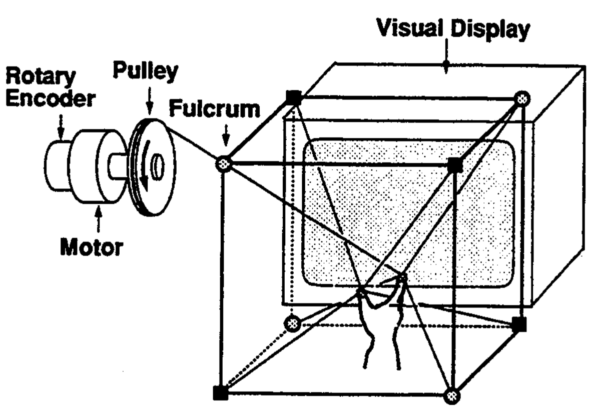


Figure 1 SPIDAR for single hand manipulation

Figure 2 shows the range of motion of the thumb and the index finger that applying force-feedback sensation is effective[3]. Each space for each finger is the tetrahedron whose vertices are the four fulcrums.

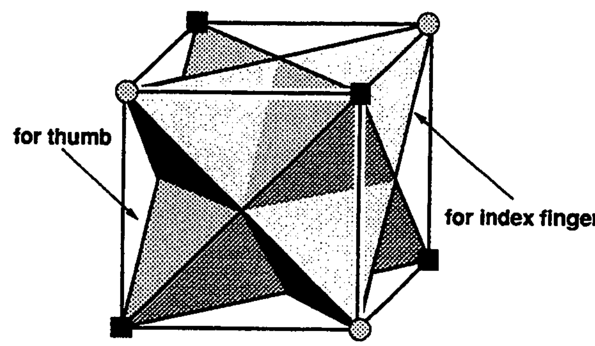


Figure 2 Range of motion of the thumb and the index finger

SPIDAR is a 3D interface device that can track fingers' position and generate force-feedback sensation to the user during manipulating virtual objects. It allows the user to touch and handle virtual objects directly by free hand.

SUITABLE VIRTUAL ENVIRONMENT INVESTIGATION

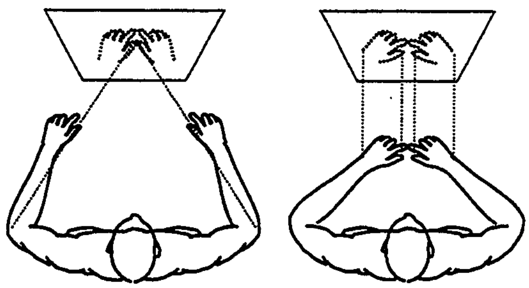
Coherency Between Kinesthetic and Visual Sensation

In our virtual work space, we use the interface device, SPIDAR, which the user has to communicate with the machine by hand movements controlling away from the display screen surface, so we must consider a natural act of eye-hand coordination to make information coincide. Here, our approach is reflecting the images of the thumbs and the index fingers on the display screen then apply force-feedback sensation and change the poses of the virtual objects corresponding to the situations of the images of fingers relative to the virtual objects. We have considered 2 methods for reflecting the images of fingers on the display screen as follows:

- *Lengthened arms method*
  This method reflects the images of fingers by

assuming that the user's arms are lengthened from the current position to the display screen (see Figure 3a). We consider particularly how far the distance between the left and the right hands should be to make the images of fingers of both hands be reflected as if they just touch each other. Let us call *latent distance.* We search for the suitable latent distance by an experiment detailed in Experimental Study section.

- *Projected hands method*
  By this method, the images of fingers are reflected parallel to the actual position (See Figure 3b),so the latent distance is zero.



(a) lengthened arms (b) projected hands

Figure 3 Methods for reflecting the images of fingers on the display screen

We determine which method is suitable for coherency between kinesthetic and visual sensation by an experimental study detailed later.

For providing visual information, we use computer graphic display system for generating stereoscopic images that the user needs to wear stereoscopic glasses to perceive 3D perspective view of the images. We currently do not plan to use head-mounted displays (HMDs) since we feel the current HMD technology is too encumbering and of too limited resolution for viewing complex data.

SPIDAR for Both Hands Manipulation

As mentioned in the previous section, originally, SPIDAR has been developed for single hand manipulation. To enhance it for both hands manipulation, we have considered various styles of setting strings. Some examples of new structures are shown in Figure 4, where the circles are the rough boundaries of each hand motion.

Figure 4a is the structure that is constructed by connecting SPIDAR for single hand manipulation 2 sets together. In this case, the problem of the interference of strings rarely occurs. However, the positions of the left and the right hands are separated rather far, so it is difficult to perform some operations that need the cooperation between both hands in the near distance in the real environment and it can not use with reflecting the images of fingers by projected hands method. On the other hand, the structures shown in Figure 4b and 4c can be used with reflecting the images of fingers both by lengthened arms method and by projected hands method, but the interference of strings occurs more often than the structure shown in Figure 4a.

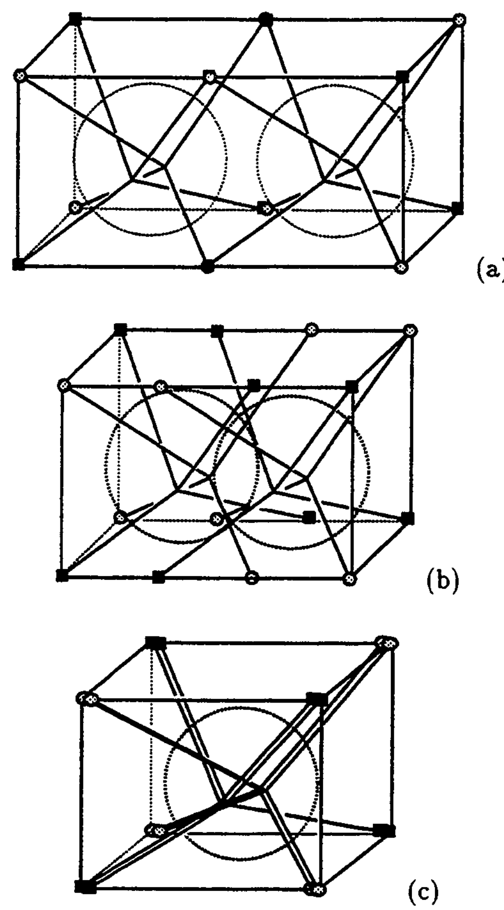


Figure 4 Examples of setting strings of SPIDAR for both hands manipulation

Experimental Study

From the conditions that we have considered above, we determine which ones are the suitable conditions by experimental study. In the experiment, we simulate an environment for performing a Face-and-Fit task (pick up two objects from a number of provided objects, then turn them until the desired sides face each other and unite them). The virtual target objects are two cubic cubes (size 40x40x40 mm^3, weight 50 g) and the initial distance between them is 120 mm. We use 9 styles of SPIDAR structure with reflecting the images of fingers by lengthened arms method varying 8 values of the latent distance: 10, 20, 30, 40, 50, 60, 70, 80 cm. For projected hands method ,which the latent distance is zero, we experiment with the SPIDAR structure shown in Figure 4b. The computer graphic display system is used for generating a real time image with screen updated rate 10 times per second. Force-feedback sensation is generated with force-feedback updated rate 30 times per second. The subject has to wear the provided stereoscopic glasses to see virtual objects as 3D objects. The distance from the subject's eyes to the virtual objects on the display screen is 75 cm. Before the subjects do this experiment, they have been trained until they have enough skills to use this interface device. After each subject finishes the experiment, we have interviewed him/her to collect information about conditions that he/she satisfies in performing the Face-and-Fit task in this environment. Figure 5 is a scene of the experiment.

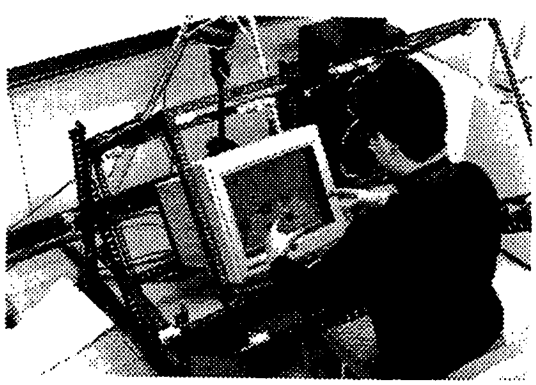
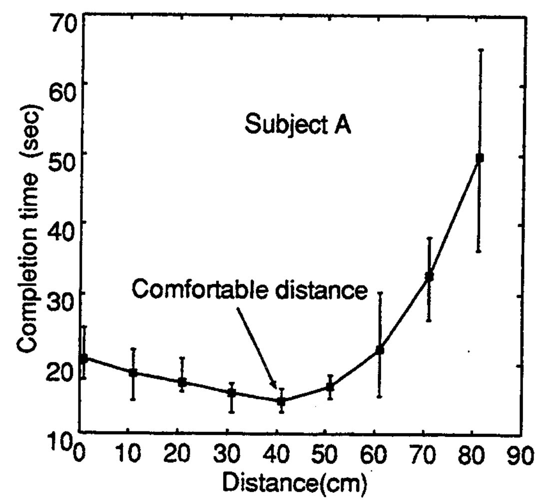
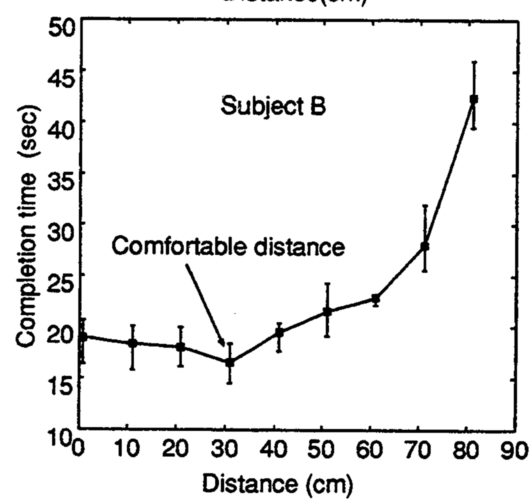


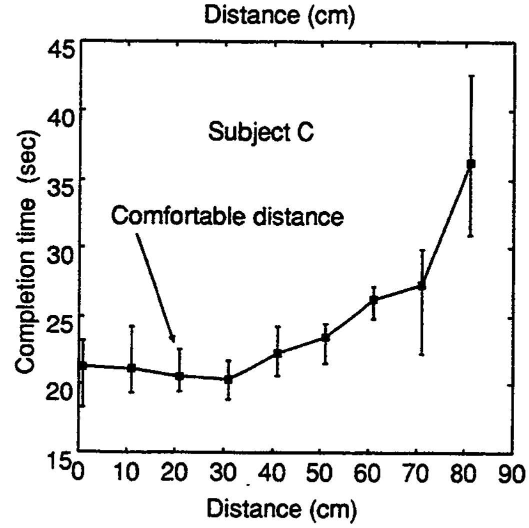
Figure 5 A scene of the experiment

Results and Considerations

Figure 6 is the result that shows the relation between the latent distance and the task completion time of four subjects. The point that is pointed by the arrow is the latent distance that the subject feels most comfortable in performing the task.







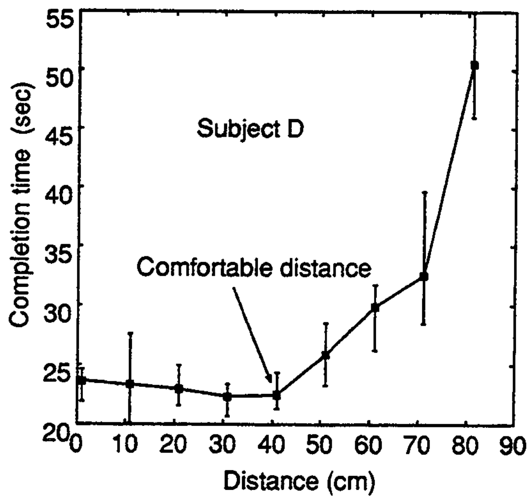


Figure 6 The relation between the latent distance and the task completion time

Subject A can accomplish the task fastest when the latent distance is 40 cm and it is the latent distance that subject A feels most comfortable in performing the task. Subject B can accomplish the task fastest when the latent distance is 30 cm and same as subject A it is the latent distance that subject B feels most comfortable in performing the task. Subject C can accomplish the task fastest when the latent distance is 30 cm, and the latent distance that subject C feels most comfortable in performing the task is 20 cm. Subject D can accomplish the task fastest when the latent distance is 30 cm, and the latent distance that subject D feels most comfortable in performing the task is 40 cm. Although the latent distance that subject C and subject D accomplish the task fastest, and the latent distance that they feel most comfortable in performing the task are different. They are not the discrepant results because when the latent distances are 20 cm and 30 cm for subject C, and 30 cm and 40 cm for subject D, they take time for accomplishing the task in the vicinity.

The most suitable latent distance is 30-40 cm that both makes the users comfortable in performing the task and helps the users to accomplish the task most effective.

Let us consider the latent distance again. Refer to Figure 7, we can establish the formula for calculating the latent distance (FlFr) as follows:

$$\mathrm{FlFr} : \mathrm{ElEr} = \mathrm{VcFr} : \mathrm{VcEr}$$

$$\begin{aligned} \mathrm{FlFr} &= \frac{\mathrm{VcFr} \times \mathrm{ElEr}}{\mathrm{VcEr}} \\ &= \frac{(\mathrm{VcEr} - \mathrm{ErFr}) \times \mathrm{ElEr}}{\mathrm{VcEr}} \\ &= \mathrm{ElEr} - \frac{\mathrm{ErFr} \times \mathrm{ElEr}}{\sqrt{\mathrm{VcH}^2 + (\mathrm{ElEr}/2)^2}} \end{aligned} \quad (1)$$

In the experiment, the distance from the subject's eyes to the display screen is 75 cm (VcH = 75 cm). The average distance from the elbow to the thumb of four subjects is about 40 cm (ErFr = ElFl = 40 cm) and the average distance from the left elbow to the right elbow is 60 cm (ElEr = 60 cm). By substituting these values in equation (1), we obtain the latent distance value that is about 30 cm corresponding to the result from our experiment.

From these results, it indicates that lengthened arms method is better than projected hands method.

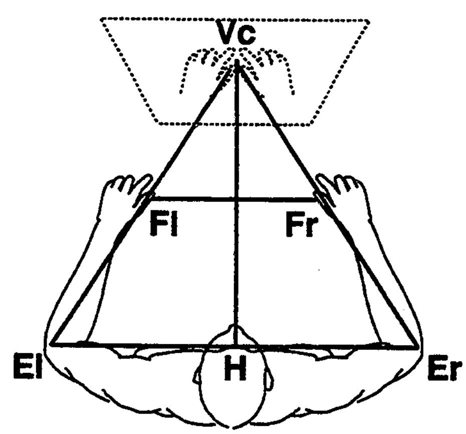
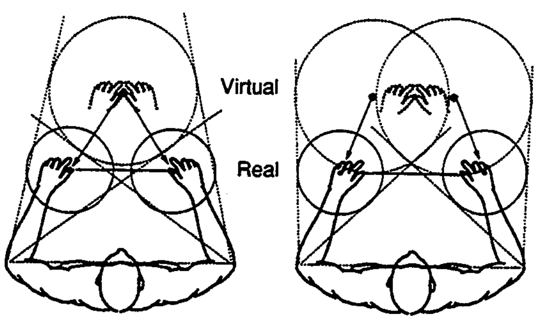


Figure 7 Geometric relation of perceptual information

VIRTUAL WORK SPACE CONSTRUCTION

Setting up the Virtual Work Space

We adopt the suitable conditions that are considered in the previous section to construct our virtual work space. Although it is suggested that the suitable latent distance is 30-40 cm, we have to consider additionally about the boundary of the images of fingers that should be reflected on the display screen. Figure 8 shows two types for defining the boundaries of work spaces of the left and the right hands on the display screen. Type-(a), the boundaries of work spaces of the left and the right hands are located in the same area. Type-(b) the boundaries of work spaces of the left and the right hands are separated but some parts join each other at the center of the total frame. The total boundary of type-(b) is larger than that of type-(a) but type-(b) can not support some operations such as moving the right hand to the leftmost of the total boundary or moving the left hand to the rightmost of the total boundary. However, in performing any tasks, we hardly operate by crossing hands and the larger boundary can present more information to the user, so we select type-(b) for defining the boundaries of work spaces on the display screen.



(a) same work space (b) separate work spaces

Figure 8 Boundaries of work spaces of both hands on the display screen

From the selected conditions, we have constructed a virtual work space for both hands manipulation which has the rough structure as shown in Figure 9.

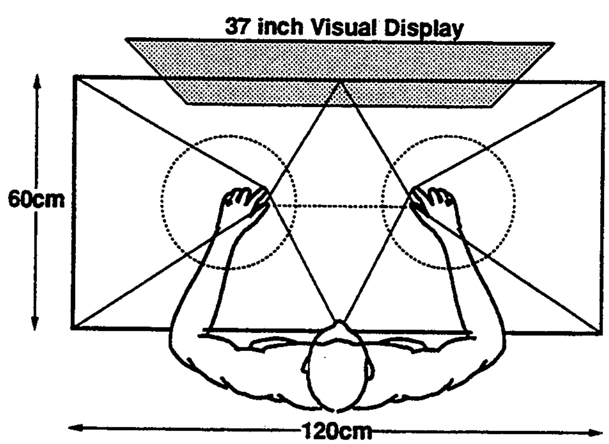


Figure 9 Rough structure of a virtual work space for both hands manipulation

Initial Application

We have evaluated our virtual work space by simulating an environment for forming a 3D model manually by combination method. We provide some simple 3D virtual models such as a sphere, a rod, etc, then let the users create new models by assembling these component models together. The users can perform naturally assembly operations and accomplish this task. Figure 10 is an example that shows some situations of the virtual models during a user is performing this task.

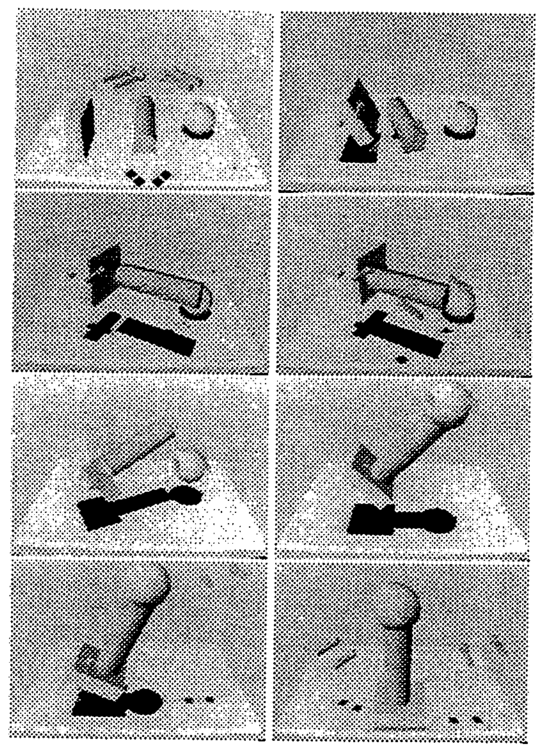


Figure 10 Example of some situations of the virtual models during a user forming a 3D model by combination method

CONCLUSION

In this paper, we have considered the suitable conditions for setting up a comfortable virtual work space for both hands manipulation. Since we use the interface device that communicates with the machine by hand movements controlling away from the display screen surface, the congruity of hand movements and visual information must be considered. In addition to force-feedback sensation, we have proposed reflecting the images of fingers on the display screen to help users to perceive the situations of fingers relative to virtual objects more clearly. Lengthened arms and projected hands are two methods that we have considered and compared by an experiment. The result shows that lengthened arms method is better. We have constructed our virtual work space according to this result then let subjects perform some simulated assembly work. Our evalution shows that this virtual work space has sufficient conditions for supporting users in performing any tasks that need hands manipulation directly or cooperation between both hands in natural manner such as creating a 3D model by assembling the provided component models together. We intend to enhance it for performing more delicate work and plan to utilize auditory sensory to provide supplemental information to the user in the future.

REFERENCES

1. Brooks,F.P., Ouh-Young,M., Batter,J.J. and Kilpatrick,P.J.: "project GROPE - Haptic display for scientific visualization", Computer Graphics, 24,4,1990,pp.177-185.

2. Hinckley,K., Pausch,R., Gobel,J., Kassell, N.:"Passive Real-World Interface Props for Neurosurgical Visualization", ACM CHI'94 Conference on Human Factors in Computing Systems, 1994, pp.452-458.

3. Hirata,Y. and Sato,M.: "3 dimensional interface device for virtual work space", Proceedings of the 1992 IEEE/RSJ International Conference on IROS, 2,1992,pp.889-896.

4. Ishii,M. and Sato,M.: "A 3D Spatial Interface Device Using Tensed Strings", PRESENCE, 3,1,1994,pp.81-86.

5. Isshiki Hachiro:"The world of hand", Education Publication,1980 (In Japanese).

6. Iwata,H.: "Artificial Reality with Force-Feedback: Development of Desktop Virtual Space with Compact Master Manipulator", Computer Graphics, 24,4,1990,pp.165-170.

7. Kabbash,P.,Buxton,W.,Sellen,A.:"Two-Handed Input in a Compound Task", ACM CHI'94 Conference on Human Factors in Computing Systems,1994,pp.417-423.

8. Minsky,M., Ouh-Young,M., Steele,O., Brooks,F.P., and Behensky,M.: "Feeling and seeing: Issues in force display", Computer Graphics, 24,2,1990,pp. 235-244.

9. Yanagida,Y. and Tachi,S.: "Virtual Reality System with Coherent Kinesthetic and Visual Sensation of Presence", Proceedings of 1993 JSME International Conference on Advanced Mechatronics (ICAM),1993,pp.98-103.