

depth cue to accommodation variable. In fact, the depth cue to accommodation is fixed to one meter on our displays. This depth is the most effective to maintain binocular fusion between the infinite far point and the area for hand-eye operation. In this case, the depth of the nearest point is 0.2 meter. The sense of reality for the operator is not broken under this condition, a proof of which was given by an experiment of space perception [2].

Therefore, convergence angle is the only cue of the absolute depth on this type of binocular display. To keep this cue physically correct, it is needed to match the following parameters to physical ones:

- 1) The viewpoint of eyes. This requires matching the width between the center of the displays and the width between the center of the cameras to the width between the operator's eyes physically.
- 2) The visual angle, i.e. retinal image size. This requires matching the view angle of the displays to the view angle of the cameras physically.
- 3) The direction of eyes. This requires matching the direction of the displays and the direction of the cameras physically.

The binocular parallax is an effective cue of relative depth among objects. This cue has such a high resolution to depth that it is always one of the essential cues on human binocular space perception. The performance is degraded, when the horizontal visual angle is smaller than 45 degree.

The motion parallax is an effective depth cue, when the viewpoint of the operator is moving. Therefore, this is the most effective cue to give the operator a sense of locomotion. The performance is degraded when the horizontal and vertical visual angles are smaller than 60 x 45 degree.

3. Development of head-mounted displays without a link mechanism to support their weight.

This type of display has no link mechanism to support its weight. Therefore, head movements of an operator using it are not restricted, and it is easy to make the system for the display smaller than that with a link mechanism. However, the operator must support all of its weight by his neck. Therefore, almost all of the technological problems to develop such displays are concerned with down sizing and lightening its weight. Moreover, operator's head movements are usually measured with some remote sensing method in this type display, therefore such sensing methods often do not give enough accuracy of measurement and not enough frequency on sampling to construct high speed control system. This type of display is

designed to teleoperate a mobile robot. In this case, they must be very light weight, and have wide angle of view to give the operator effective motion parallax. For the purpose, such displays need to have a very wide view angle (H60 x V40 degrees minimum) and moderate resolution (0.1 on equivalent visual acuity minimum).

Prototype display Mark 1 (here after HMD-1) is shown in Figure 1. It is 1.7 kg in weight, and has 4 inch size full-color liquid crystal displays for each eye to display a 3-dimensional visual image to the operator. Each display has H320 x V220 pixels resolution. The view angles of the displays are H72 x V54 degrees. The head position and movements of the operator are measured with an electromagnetic sensor (3SPACE TRACKER) in real time (60Hz). The block diagram of this display system is shown in Figure 2.

To reduce the weight, eye lenses are worn as glasses, separated from the displays. With the lenses, accommodations of the operator are fixed to the depth of 1 meter, when he is watching the displays. The most serious problem on this display is that the width between displays is larger than the width between eyes (Fig.3(a)). In this case, the cameras need to have similar optical system to produce correct 3-dimensional visual image.

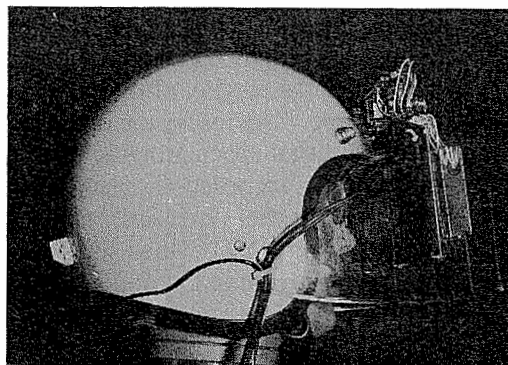


Fig.1 Head-mounted display HMD-1

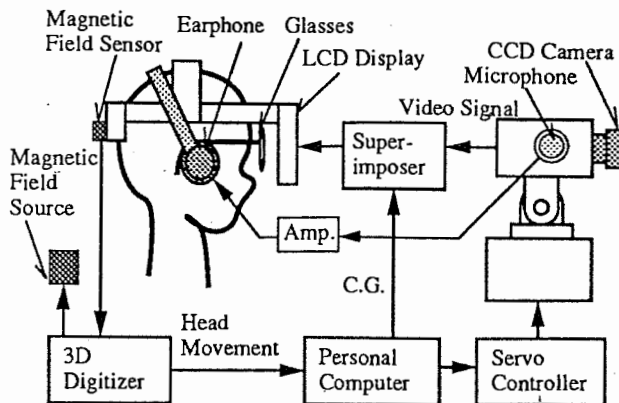
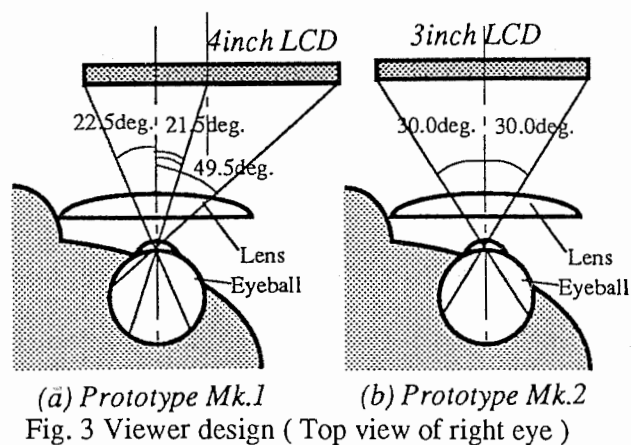


Fig. 2 Blockdiagram of head-mounted display



To check the sense of reality, computer graphical images of a virtual world and camera images of a real environment were shown to an operator with this display system. The operator reported problems as follows :

- 1) The resolution of represented images was too rough.
- 2) The border between binocular fusion area and the other area in sight was too prominently displayed.
- 3) The optical distortion disturbed binocular fusion at the side of binocular fusion area, when real images are displayed.

The reason for 1) is that the equivalent visual acuity of this display was only 0.07. Improvement of the acuity was needed.

The reason for 2) is the presence of not corresponding area on binocular vision. The whole area of the view represented in this display has an angle of 100 degrees. On the other hand, the binocular corresponding area has a view angle of only 45 degrees. The binocular not corresponding area is more than half of the whole image. It has too much area to produce natural 3-dimensional visual image.

The reason for 3) is the distortion of wide conversion lenses of the cameras. In this case, the lenses of the cameras needed to have a view angle of 100 degrees. Such lenses have more optical distortion at the side area than at the center area. When the operator watches the side area of his image, one eye sees a very distorted image, and the other does not. As a result, binocular fusion is disturbed.

To improve problems 2) and 3), the width between displays needs to coincide with the interpupillary distance. It decreases the difference between the images of right eye and left eye for both the not corresponding area and optical distortion.

On the other hand, with separate eye lenses, it is easier to decrease the distance between lens and eye than with "view window " to peep into lenses established to the display. It is also easier to prevent the lens shifting from the center of the sight. When the eye lenses have large degree (20 diopters in this case), it becomes important to prevent lens shifting.

In consideration of these problems, the prototype display Mark 2 (here after HMD-2) was developed. It is shown in Figure.4. The measurement of head position and the method to set up eye lenses are same as on HMD-1. It is 0.6 kg in weight, and has 3 inch size full-color liquid crystal displays for each eye to display 3-dimensional visual image to an operator. Those displays have H372 x V240 pixels. Also, by using smaller LCDs, the construction aligns the center of the line of sight with the center of the display screen, solving the problem of the Mark 1. The field of vision is 30° both inside and outside (Figure 3(b)) and the equivalent visual acuity on this display was 0.1. Except for the inadequate resolution, the binocular fusion region is more than 45° horizontally, the monocular field of vision angle is more than 60° horizontally and more than 45° vertically, which practically satisfied the requirements for the field of vision which were mentioned previously. The transmission of color images using NTSC standard composite signals under similar conditions gives a maximum resolution of 0.13 when converted to the equivalent visual acuity, and this is the upper limit.

By separating the circuits to drive LCDs and locating them behind the head, the weight distribution does not incline to the front. The use of lighter and smaller LCDs and distributing the center of gravity of the equipment closer to the center of gravity of the head has resulted in a head fixture which is simpler than the Mark 1 device. The major reduction in the weight of the equipment has had a considerable effect.

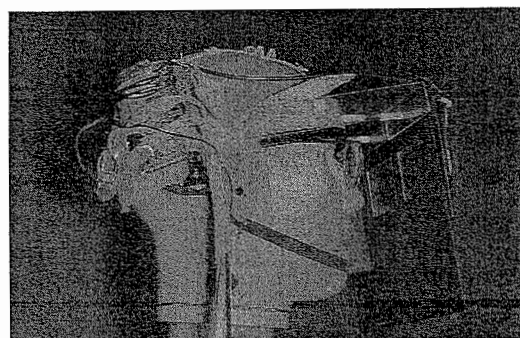


Fig.4 Head-mounted display HMD-2

4. Development of head-mounted displays with a link mechanism to support their weight.

This type of display has a link mechanism to support their weight. In result, the operator need not to support all of its weight by his neck, and it is easy to measure the movements of head with the link mechanism. However, to prevent head movements of an operator using it from being restricted, the link mechanism to support the display needs 6 degrees of freedom which is too complicated to make the display system for the display small. Therefore, almost all of technological problems to develop it are concerned with simplifying the link mechanism and lightning its inertia.

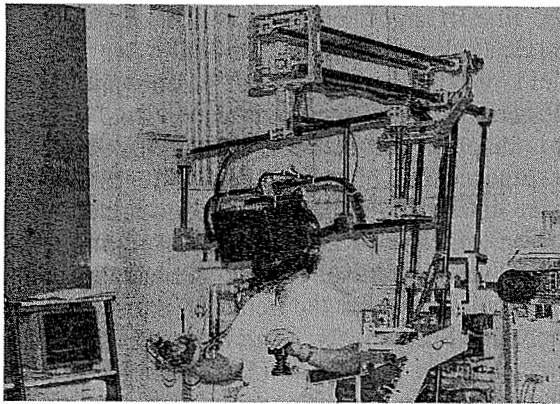


Fig.5 6-degrees of freedom link mechanism to support a head-mounted display

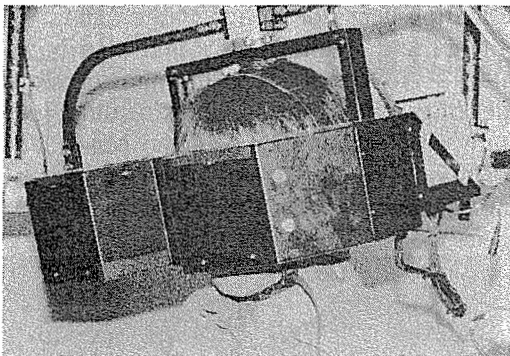


Fig.6 Head-mounted display HLD-1.

This type of display is designed to teleoperate an anthropomorphic robot for fine manipulation tasks. In this case, such displays need to have moderately wide view angle and fine resolution. Because the acuity of the position control for the tip on the manipulator of this robot is 1 mm on

the depth of 1 m, it is necessary for the display to be able to represent the small visual difference. For the needs, the equivalent visual acuity of the display is needed 0.291 at least. Also, free space in front of an operator's face is needed to place his hand there, because he has to watch very closely objects gripped in the manipulator hand.

We developed a link mechanism to support a display which had 6 degrees of freedom. The mechanism is shown in Figure 5. This mechanism was designed always to place the center of the weight including the link mechanism, the display, and counter weights, at a fulcrum on the basement of the system in spite of any movements of the operator's head. Therefore, when the operator is moving his head to use this system, he doesn't feel the weight of them but only their moments of inertia. In this system, the movements of the operator's head are measured by 6 rotary encoders attached to the link mechanism, enabling to separate measurement of the posture and position of the head.

Prototype display Mark 1 (here after HLD-1) is shown in Figure 6. It is 5.5 kg in weight, and has a 4 inch size full-color CRT for each eye to display 3-dimensional visual image to an operator. The CRTs are placed around the head of the operator to reduce the moment of inertia of the display through an optical system shown in Figure 6 in which each visual image is reflected twice by mirrors. Also, the system keeps free the space in front of the face of the operator needed to manipulate objects. Accommodations of the operator are fixed to the depth of 1 meter when he is watching the displays by the eye lenses, as same as on HMD-1 and HMD-2. In this display, the width between eyes is 65 mm, the view angles are $H33 \times V24$ degrees, and the equivalent visual acuity is 0.3.

Generally, useless mirror images are seen in such an optical system using reflection. Because such mirror images prevent the operator from having a natural visual image, they have to be removed. We removed such mirror images with polarizers instead of blinds in order to prevent only slight shifting viewpoints of the operator from decreasing the visible area.

To check the sense of reality, computer graphical images of virtual world and camera images of real environment were shown to an operator with this display system, and he teleoperated manipulation tasks with virtual and real master-slave systems. The operator reported problems as follows.

- 1) It was sometimes inconvenient that the visual angle was a little small.

2) The moment of inertia of display was too large for an operator to move his head rapidly without support by his hand.

3) Eye lenses clouded up in long continuous use.

The reason for 1) is that the view angles are very small, $H33 \times V24$ degrees, in order for the equivalent visual acuity on this display to be more than 0.291. This factor makes it especially difficult to track a moving object in addition to the factor of 2).

Problem 1 should be alleviated by using a high resolution display and a wide visual angle. Reducing the moment of inertia and changing to a closer head fitting may improve Problem 2. About Problem 3, the eyepieces are a fixed to the equipment and are designed to fit close to the face to avoid any deviation from the head position. The cause of the problem is that the lenses are fitted closer to the eyes and ventilation is poor. Since feedback for the whole system depends on sight, it must be cleared of any element which would interfere with its continuous use as a working presentation system.

With due regard to these problems, we proceeded to construct Prototype display Mark 2 (here after HLD-2). An external view is shown in Figure 7. A six-inch LCD is used to reduce weight and provide higher resolution and the close fitting helmet enables rapid movement without help from the hand. The total weight is 5.5 kg (including 1.5 kg for the helmet), which is not lighter than Mark 1 but, by constructing it so that each component is located close to the head, the moment of inertia has been kept small.

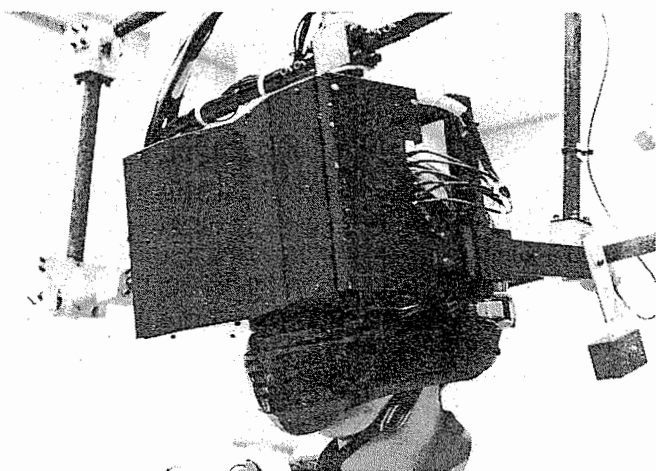


Fig.7 Head-mounted display HLD-2.

The LCD employs a six-inch ($H720 \times V240$ dot) element for RGB signal input. We constructed a compact optical system by placing the LCD

upright with two mirrors to provide a three-dimensional optical path. This enables the space in front of the face to be reserved, which is essential when manipulating objects.

As with HLD-1, the eyepieces are set in order for the virtual image on the display screen to appear 1 m away and for useless reflections to be removed from view by the polarizer. The interpupillary distance is 65 mm, the monocular visual angle is 40° and the equivalent visual acuity is 0.3. A fan has been installed to prevent clouding of the lenses. This avoids the view being impaired during long continuous use.

5. Tracking Experiment

The reality which is obtained in the HMD is not simply static but also includes dynamic factors. In order to compare these characteristics, in this paper we used an experimental tracking simulator to measure the operating motion characteristics of HMDs. This experiment was done with a single eye display image, in order to evaluate not three-dimensional sensory perception but operating characteristics as a visual presentation system for tracking head movements. The assumptions for this tracking experiment are shown in Figure 8.

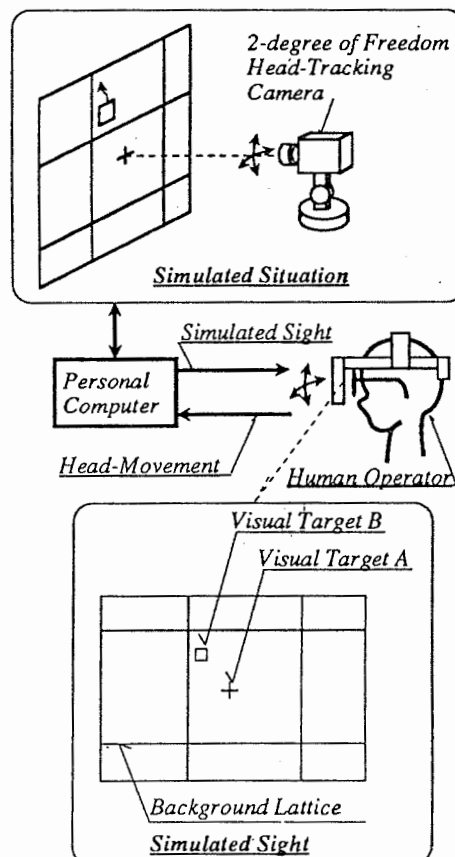


Fig.8 Model of tracking simulation

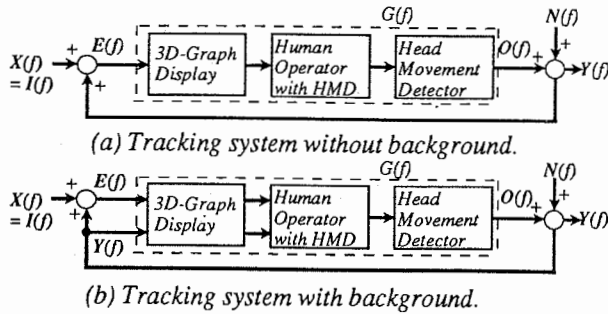


Fig.9 Block diagram of tracking simulation system

The slave system for tracking head movements is a camera with two degrees of freedom, pan and tilt. The human operator has to move his head in order for Target A which is fixed in the center of the camera's sight to track Target B which moves around on the plane facing the camera. Then, a lattice is also shown as background on the plane. If there is no background, the visual feedback to the operator is only the movement of Target B relative to the head movement, that is to say, relative to Target A. But the presence of a stationary background effectively provides the operator with visual feedback on the absolute movement of both the head and Target B.

So, whether or not there is a background makes the differences in the control mechanism for the experimental system, including the operator, as shown in Figure 9, and an experiment conducted by us has confirmed that it changes the tracking characteristics [3]. In the actual experiment, a computer graphics image was generated to simulate this environment (camera, lattice, Targets A and B) with a personal computer which measured the head movements. In this way, the experiment removes the influence of the camera motion characteristics.

Target B movement, $i(t)=x(t)$, used quasi random signals from the combining of sine waves such as Equation (1). The frequencies of the sine waves were arranged at regular intervals on the logarithmic scale and amplitude and frequency were inversely proportional. This standardizes the energy component and speed amplitude of each frequency.

$$x(t) = \sum_{k=0}^n a_0 p^{-k} \sin(2\pi f_{\min} p^k t + \phi_k) \quad \dots(1)$$

($p=1.25, n=17, f_{\min}=0.0222\text{Hz}, \phi_k=\text{random}$)

Also, double buffering was used for the presentation to remove flicker and scatter. The presentation and sampling cycles were 44 ms. The tracking result corresponding to the operator's actual command value, $o(t)$, and including noise, $n(t)$, is measured. It has the form $y(t)=o(t)+n(t)$.

The open loop transfer function, $G(f)$, of the presentation and operating system, which includes the operator, must be evaluated here. It is calculated from the closed loop transfer function, $T(f)$, using Equation (2).

$$G(f) = \frac{T(f)}{1 - T(f)} \quad \dots(2)$$

The closed loop characteristic, $T(f)$, is defined by Equation (3) using the Fourier transforms $I(f)$ and $O(f)$ of the input and output $i(t)$ and $o(t)$.

$$T(f) = \frac{O(f)}{I(f)} = \frac{\phi_{IO}}{\phi_{II}} \quad \dots(3)$$

From the Fourier transform of the results of sampling 1024 points, we get $X(f) = I(f)$ and $Y(f) = O(f) + N(f)$ and set the mean of the ensemble to ϕ_{xy}, ϕ_{xx} . Now, for the evaluate of $G(f)$, here we used the McRuer's crossover model, which is a model of human operator characteristics, for the analysis and assessed the characteristics according to the crossover frequency, f_c , and the equivalent delay, t_d .

$$G'(f) = \frac{f_c}{jf} \exp(-j2\pi f t_d) \quad \dots(4)$$

Based on this experimental method, we took up HLD-1 and HLD-2 (hereafter called L1 and L2) and HMD-2 (hereafter called M2) and compared the motion characteristics. HMD-1 was not included in the experiment because of the many differences in its presentation condition with the other displays.

6. Experimental Results

Figure 10 shows an example of the frequency characteristics of least squares fits with the crossover model which were obtained from the results of the experiment. Results which compare the crossover frequency, f_c , and the equivalent delay, t_d , obtained in this way for each display are shown in Table 1. We can see a trend for f_c to increase from L1 to L2 to M2 and for t_d to decrease from L1 to L2 to M2. The moments of inertia decrease and the visual angles increase for each item of equipment in the order L1 to L2 to M2. Therefore, we can see that these are the primary causes which improve motion characteristics by the increase in f_c and the decrease in t_d . Accordingly, in the HMD systems with link supports which have no difference in the weight of the equipment, compared with L1, the characteristics of L2 are markedly improved, showing motion characteristics close to the much lighter M2. This shows that, with gravity compensated link supports, instead of reducing the weight of the equipment, it is more effective to

improve the motion characteristics by decreasing the moment of inertia of the HMD. Also, both f_c and t_d show a tendency to increase with the background display but, according to reference [4], this is a trend which we have observed in other experiments.

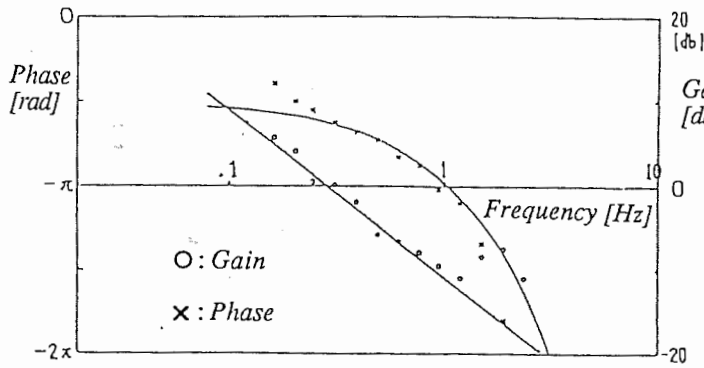


Fig.10 Frequency response

Device	Back-Ground	f_c [Hz]	t_d [s]
L1	on	0.3090	0.2802
L1	off	0.3042	0.2627
L2	on	0.4033	0.2673
L2	off	0.3613	0.2420
M2	on	0.4272	0.2567
M2	off	0.3762	0.2067

Table.1 Dynamic responses

7. Discussion

According to the McRuer's crossover model, f_c and t_d are essential elements in determining the $G(f)$ gain and the equivalent delay, respectively. It is the model of the human operator's motion characteristics, then we could consider that the equivalent delay element t_d , depends on the visual perception, and the gain f_c , depends on the operational mobility.

Following this reasoning, we can group the elements which are thought to influence the frequencies of each item of equipment. An analysis of f_c and t_d according to these elements, leads to the following:

$$f_c = f_0 K_p K_r, \quad t_d = t_0 + t_q + t_s \quad \dots(5)$$

Here f_0 and t_0 indicate gain and equivalent delay due to the basic characteristics of the operator in conditions of direct observation in the tracking experiment with no background. In this regard, K_p represents the proportional change in gain due to the moment of inertia of the HMD and K_r represents the proportional change in gain due to the presence of the background. On the other hand,

t_q represents the delay component due to the limitation of the visual angle and t_s represents the delay component due to the increased sensory perception due to the background display.

By applying parameter analysis according to each of these elements to the crossover model of Equation (4).

Modelling is possible with Equation (6) for the tracking experiment with the background,

$$G(f) = \frac{f_0 K_p K_r}{f} \exp(-j2\pi((t_0 + t_q + t_s)f + 0.25)) \dots(6)$$

and Equation (7) for the tracking experiment without the background, where it is taken that $K_r = 1.0$ and $t_s = 0.0$.

$$G(f) = \frac{f_0 K_p}{f} \exp(-j2\pi((t_0 + t_q)f + 0.25)) \dots(7)$$

The experiment on this paper is not intended to compare conditions of direct observation with the use of the HMD but it is an experiment to compare the characteristics of various HMDs. So it is not necessary to obtain f_0 and t_0 with a direct observation experiment. The tracking characteristics for the M2 with no background is taken as the standard characteristics corresponding to direct observation, because it has the smallest operating load and the largest visual angle among our HMDs. Under these conditions it is accepted that $K_p K_r = 1.0$ and $t_q + t_s = 0.0$. The resulting gain and equivalent delay, f_0 and t_0 , are determined to be standard. The constants obtained by this model are shown in Table 2.

Device	K_p	t_q [s]	K_r	t_s [s]
L1	0.821	0.0560	1.02	0.0175
L2	0.961	0.0353	1.12	0.0253
M2	1.00	0.00	1.14	0.0500

$$(t_0 = 0.207 \text{ [s]}, f_0 = 0.376 \text{ [Hz]})$$

Table.2 Comparison among tracking schemes.

We can consider that, since K_p and t_q are characteristic differentials between items of equipment, they reflect the differences in operating loads, such as moments of inertia and visual angles of the presentation region. And, since K_r and t_s are the amounts of change due to the background display, they reflect mainly differences in visual angle of the visual stimulus. L1 has horizontal visual angle of 33°, L2 has 40°, and M2 has 60°.

K_p is the component which can be considered to be due mainly to changes in the moment of inertia, as the moment of inertia decreases from L1 to L2 to M2, K_p is increasing. Since this is probably pure inertia of the HMD working in accordance

with head movements, the moment of inertia and the gain are in a reciprocal proportion. Moreover, K_r is the component due to the background display but for all the equipment it is greater than 1.0 and because there is feedback of the absolute movement of the scene and head due to the background display, we can see that the gain is improved. Again, K_r increases from L1 to L2 to M2 and we can see that it is increasing in accordance with the widening of the visual angle of the background display region. From this it follows that the rate of increase of the gain becomes larger in accordance with the expansion of the visual angle.

On the other hand, t_q , the component of changes in the effective visual field of the display, decreases from L1 to L2 to M2, and the equivalent delay due to the increase of the visual field of the presentation region is decreased. This is probably because the expansion of the visual angle causes the speed of the scene to increase, causing a speed stimulus which shortens the visual delay. By contrast, the background presentation component, t_s , increases from L1 to L2 to M2 and, as the visual angle of the background display increases, we can see the amount of increase of the equivalent delay also increases. This is probably due to the fact that, as the presentation area of the visible stimulus on the retina increases, the time required to process the visual information is also increasing monotonously.

This discussion points up the following guidelines in the design of the HMD concerning weight distribution and the visual angle of the presentation in order to improve its operational mobility. From the point of view of weight distribution, the gain with L2 is a marked improvement over L1 and, since it has a value to that of M2, to the extent that link supports are used, reducing the moment of inertia is more important than the total weight for improving operational mobility. Also, from the point of view of the visual angle of the presentation, since gain increases and an equivalent delay occurs when the visual angle is expanded, we can say that increasing the visual angle is an effective way to improve operational mobility.

Again, the increase in gain and equivalent delay due to the presence of the background both tend to grow with the expansion of the visual angle. This is because, if this relationship is used when the working environment is being designed, it is possible to work a trade off with the presence or absence of a background for operational mobility between gain and equivalent delay. In this case, a large change will be obtained with a wide visual angle.

8. Conclusion

We have designed and made an experimental model of a head movement tracking type three-dimensional sight presentation equipment for tele-existence systems for movement and work. Also, we have used this equipment for tracking experiments and investigated its motion characteristics. From the results of the experiments, we have also compared the differences between equipment items and studied the primary factors which control the motion characteristics.

With the spread of artificial reality, sensory technology and tele-existence technology in the future, the HMD is a device which will have a wide application as a human interface, giving a high degree of reality. The essential design elements and standards for the evaluation proposed in this paper are fundamental to the development of this sort of human interface for visual presentation technology.

ACKNOWLEDGMENT

This research was performed at Mechanical Engineering Laboratory / MITI. We would like to express our appreciation of the great support we have received from the Director of the Robot Engineering Department of the Mechanical Engineering Laboratory, Dr. Naotake Ohyama, and the Director of the Bio-Robotics Division of the Mechanical Engineering Laboratory, Dr. Kazuo Tanie, in promoting this research.

REFERENCES

- [1] J.D.Foley, "Interfaces for Advanced Computing", Scientific American, pp.83-90, Dec.1987.
- [2] S.Tachi, K.Komoriya and M.Kaneko, "Tele-existence(I) - Design and evaluation of a visual display with sensation of presence - ", Proceedings of 5th Symposium on Theory and Practice of Robots and Manipulators (RoManSy84), pp.245-254, CISM-IFTOMM, Udine, Italy, June 1984.
- [3] H.Arai and S.Tachi, "Study on tele-existence (12) - Tracking characteristics of camera for tele-operation -", Proceedings of RSJ 5th Conference, pp.97-98, Tsukuba, Japan, Nov. 1987 (in Japanese).
- [4] S.Tachi, R.W.Mann and D.Rowell, "Quantitative comparison of alternative sensory displays for mobility aids for the blind", IEEE Transactions on Biomedical Engineering, Vol.BME-30, No.9, pp.571-577, 1983.