

Amplifying Head Movements with Head-Mounted Displays

Abstract

The head-mounted display (HMD) is a popular form of virtual display, due to its ability to immerse users visually in virtual environments (VEs). Unfortunately, the user's virtual experience is compromised by the narrow field of view (FOV) it affords, which is less than half that of normal human vision. This paper explores a solution to some of the problems caused by the narrow FOV, by amplifying the head movement made by the user when wearing an HMD, so that the view direction changes by a greater amount in the virtual world than it does in the real world. Tests conducted on the technique show a significant improvement in performance on a visual search task, and questionnaire data indicates that the altered visual parameters the user receives may be preferable to those in the baseline condition where amplification of movement was not implemented. The tests also show that the user cannot interact normally with the VE if corresponding body movements are not amplified to the same degree as head movements, which may limit the implementation's versatility. Although not suitable for every application, the technique shows promise, and alterations to aspects of the implementation could extend its use in the future.

1 Introduction

This paper describes an idea for overcoming the limited field of view offered by contemporary head mounted displays (HMDs), and reports results of an experiment to test its effectiveness under controlled conditions. The fundamental idea is to amplify rotational head movements, so that, for a given turn of a subject's head, his or her view in a virtual environment rotates by a greater amount. We call this amplifying – or sometimes exaggerating – the head movement. A simple experiment demonstrates how this works. Standing up, it is easy for a human subject to look at their feet: he or she rotates the head forwards by an angle of approximately 45° and then directs the eyes downwards. Attempting the same task while wearing an HMD requires a grossly exaggerated movement, in which the upper body is bent forwards, necessitated by the limited field of view of the HMD. Similar conditions apply to lateral head turning. In the normal world, with a head turn of around 90° , peripheral vision permits one to see directly behind oneself. But in an HMD the view is blinkered; trying to look over one's shoulder is well nigh impossible. By amplifying the head movement we emulate the normal visual range afforded by head movement plus eye movement. Thus, the amplified movement affords an extended range of vision similar to that which we experience in normal vision.

This kind of amplified movement is clearly unlike our normal experience. By choosing a large scaling factor we can make the user's view spin round alarmingly. Similarly, we can exaggerate vertical movements so that tilting the head backwards results in an upside-down, backwards view of the world. The experiment reported here was designed to discover whether human subjects found amplification unnatural or disorientating, and whether they were able to function effectively at a given task with amplification enabled. The task involved users selecting targets, using a hand-held 3D mouse for pointing. This raised another interesting issue, relating to a possible conflict between visual and

proprioceptive cues: was the user's hand displayed where they expected to see it – where they felt it to be? We therefore experimented with amplifying the hand motion to make it consistent with the amplified head movement, and we report on that.

2 Background and related work

For some time the head mounted display has been a popular virtual display device, particularly in the virtual reality research community. It has proven able to provide the user with a much higher perceived level of presence than other devices (Hendrix & Barfield, 1996), and improved performance in tasks such as navigation (Peruch & Mestre, 1999). This is principally because its head-tracking facility enables 'a more intuitive navigational interface' than other types of display (Waller, 1999). Unfortunately, HMDs do have one particular disadvantage. Whilst the human field of view (FOV) is over 200 degrees horizontally and 150 degrees vertically (Arthur, 1996), the great majority of HMDs offer a visual field that is only a fraction of that: around 50 degrees horizontally and 40 degrees vertically (Barfield et al, 1995).

This narrow FOV causes many problems. Peripheral vision is greatly diminished, meaning that considerable head movement is required to see parts of the environment that would normally be brought into view by a short eye movement. This is compounded by the fact that HMDs are generally cumbersome, so any form of head movement requires far more effort than it would in a normal environment. The clipping of the FOV also means that objects pass out of sight more quickly than normal when moving through an environment. As the normal horizontal FOV is greater than 180 degrees in the real world, an object that has passed out of sight would be behind you. However, the much narrower FOV in an HMD frequently leads to the user assuming that objects are further behind him or her than is actually the case. It is not unusual, when navigating a VE, to attempt to turn into a doorway to the side of you, only to find you have not yet reached it! The frequent collisions reported by Witmer, Baily & Knerr (1994), have been attributed by Kline & Witmer (1996) to the limited visual field. As participants pass through a VE, they lose sight of perspective cues that would normally be available to them, such as the contours provided by the intersection of the walls with the floor, ceiling, and other walls. Without these cues it is very difficult to estimate how far away from the wall you are positioned. The reduced vertical FOV also means that it is difficult not only to position oneself in a VE, but also to see areas of the environment that are very close, such as the floor, ceiling and parts of one's own body.

Unfortunately, with current technology, it is not possible to produce a commercially available, reasonably priced, HMD with a wide FOV and a high quality display. The focus in recent developments has been largely on reducing cost and weight. As the LCDs commonly used to construct HMDs are of a finite resolution, a widening of the display and thus an increase in the physical visual field would inevitably lead to a lower resolution of the displayed image. CRTs, though able to provide a higher resolution, are subject to the same trade-off, and are also heavier, making them less practical for use in HMDs. The delay in visual updates to which wide FOV HMDs are subject has also proved problematic, increasing the risk of simulator sickness (Dizio & Lackner, 1997; Arthur, 1996). Developments in HMD design, such as the Virtual Retinal Display (Pryor, Furness & Viirre, 1998), and the Eye Movement Tracking HMD (Iwamoto & Tanie,

1997), potentially seem able to offer a wider FOV by taking a completely different approach to the hardware used. However, these are currently only prototypic technologies, and thus remain in the early stages of development.

Until such hardware becomes widely available, one of the most promising avenues may lie in manipulating the software that supplies the visual parameters to HMDs. The technique used in our study attempts to counteract some of the negative effects of a narrow FOV by taking advantage of the fact that it is surprisingly easy to fool the human sensory system. The validity of the technique is supported by three studies indicating that amplifying the user's movements is both acceptable and useful to the user. A fourth study provides evidence that the amplification of head movements specifically may not only be useful to the user, but may also make his or her virtual experience feel more naturalistic.

Mine et al (1997) have proposed a technique called 'scaled-world grab'. This allows objects that would normally be some distance away from the user to be brought into reach by scaling down the world about the user whilst he or she is reaching towards the object. This allows the user to effectively reach out and 'grab' whichever object is required for manipulation, no matter how far away it is. Users reported having no difficulty using the technique, and often did not notice that scaling had taken place.

In a slightly different vein, Poupyrev et al (2000), have produced a framework for exaggerating the rotation of a virtual object, using a non-isomorphic mapping between the object and a multiple degrees of freedom controller (in this case a Polhemus SpaceBall, 6-degree of freedom magnetic tracker). The technique uses quaternions to represent the orientation of the virtual object, which in turn is determined by the orientation of the controller. To change the orientation of the object, the user simply manipulates the controller. However, rather than an isomorphic (one-to-one) mapping existing between the controller and the object, the controller's rotation is amplified, so that a small change in the orientation of the controller will produce a much bigger, corresponding change in the rotation of the virtual object. As a result, subjects were able to accomplish a rotation task 13% faster than with a one-to-one mapping, without any significant loss in accuracy. Poupyrev et al (1998) have also produced a technique to amplify head rotations in 3D user-interfaces in a desktop system. In a short experiment used to demonstrate the effects, the user's head movements were tracked, and the corresponding, amplified version of these movements was presented on the screen. Poupyrev and his colleagues report that the mappings used felt 'intuitive' to the users, and enhanced rather than degraded their experiences in the virtual environment. As such, their results provide a strong indication that exaggerating head movements in a VE may well be a useful and realistic proposition.

It also seems possible to make large alterations to the visual parameters during navigation (Razzaque, Kohn & Whitton, 2001). Razzaque et al's technique, known as 're-directed walking', gives users the impression that they are zigzagging from side to side down a long room, when in fact they are moving backwards and forwards between two points. This is accomplished by a rotation of the visual scene about the user, whilst he or she is turning round.

The fact that people frequently find the sensory manipulation used in these studies undetectable is not surprising. In the case of Razaque et al (2001), users are completely unaware of the rotation of the scene about their bodies, due to the movement they are making themselves: optic flow occurs in the expected direction, and the vestibular system will not detect any anomalies, as it is not directly stimulated by this kind of movement. As humans carry a notoriously sparse and incomplete internal representation of their surroundings (e.g. Simons & Levin, 1997), they are unlikely to notice any discrepancy between the amount by which the view has changed, and the amount by which it should change.

A recent study has also suggested that not only is the amplification of head movements acceptable to the user, but that it may make the virtual experience feel more natural than the usual isomorphic mapping. Jaekl et al (2002) asked subjects wearing a head mounted display to turn their head and adjust the visual display until it corresponded to the view they felt they should receive. It was found that a large range of movement was tolerated by the subjects and judged to be stable. However, for subjects to feel most comfortable, the display typically needed to be turned by a greater amount than the movement: subjects unwittingly 'amplified' their own head movements.

Our study takes advantage of the malleability of the human perceptual system, and builds on the results of Jaekl et al (2002), by increasing the amount by which the visual field changes in response to a given head movement. It is predicted that 'exaggerating' head movement in such a way will make it easier to survey the environment, thus making the experience of using a HMD more comfortable, and alleviating some of the difficulties caused by the narrow FOV.

3 Method

3.1 Experimental Task and Test Environment

A visual search task, in which users were required to locate and then select targets, was chosen as an effective way of measuring how easy it was to locate objects outside the immediate FOV in a VE. Visual search involves peripheral vision, and is thus seriously constrained by a narrow FOV (Arthur, 1996). Each trial was timed from beginning to end, providing a measure of the participant's performance.

The environment used was a naturalistic and detailed model of the laboratory in which the experiment was conducted (see figure 1). The participant sat in the center of the room. The participant's hand was represented by a purple sphere, 6cm in diameter, which corresponded to the physical location of the 3-D mouse. The targets consisted of ten red spheres, 14cm in diameter, which were located between 0.5m and 3m away from the participant in the horizontal plane, and up to 1 m above or below the participant's eye point. Three 'sets' of target positions, designed to be equivalent to one another, were used in the experiment. This was to control for effects arising from the fact that some targets may be easier to locate and select than others. None of the targets in one set appeared in the same location as any target in the other two sets, aside from the first

target, which was always positioned in the same place: in front of the participant, slightly above eye point and within easy range.

The targets appeared one at a time in the environment, each out of the horizontal and/or vertical field of view of the next. Ten targets were used, to ensure the task was of sufficient difficulty, and that the trials would be long enough to allow any potential differences between the conditions to be detected. In order to make the next target appear, the participant had to 'hit' the current target, by placing the cursor sphere representing their hand over the target sphere, and pressing a button on a 3-D mouse. This ensured that the target had actually been seen, and that each participant was completing the task to the same level of accuracy. To determine whether the target had been hit, ray casting was employed, using a ray from the eye point through the center of the cursor (see Figure 5).

None of the targets was positioned more than 3 meters from the participant, as pilot studies indicated that the intersection task became considerably more difficult when the targets were located beyond this.

3.2 Performing the Amplification

Figure 2 shows the relation of the viewing parameter axes, which constitute the coordinate system of the visual display, to the world axes of the virtual environment. A vector, 'eye', at the origin of the viewing parameter axes is determined by the position of the HMD sensor, and defines the user's eye point in world co-ordinates. Rotational movement of the user's head was amplified around this point, from a position of zero rotation (when the user was facing towards the Polhemus emitter; see Figure 6). The view direction is determined by the z axis, down which the user looks.

A plan view of the way in which lateral turning movements were amplified is depicted in Figure 3. The 'initial position', in which all the participants started when they entered the VE, and which was used as the point from which to amplify movement is shown by (a). An actual head movement taking place, with the corresponding change in the view that the user would normally receive is represented by (b). The change in the FOV that the movement in (b) would equate to when amplification is taking place is shown by (c).

Amplifying the lateral movement of the head is predicted to be a comfortable sensation for the user, as the vestibular system will not be disrupted. Movement in the horizontal plane was thus scaled by a factor of two, as this closely approximates the additional FOV afforded by peripheral vision. In the vertical plane, however, a large amount of amplification could potentially cause a very strange or uncomfortable sensation, particularly if the user is able to see upside-down! For this reason, the amount by which head rotation is amplified vertically decreases the further the user moves his or her head away from its original position (and stops when the user is looking directly up or down). This smaller amount of exaggeration is still able to make looking up and down less effortful, and compensate for the eye movements that make it possible to look at the floor

and ceiling in the real world without having to tilt one's head horizontally backwards or forwards, and which HMDs currently deny the user. Amplification in the vertical plane was achieved by multiplying the y component of the z axis by a factor of two. Rolling motions of the head (side to side tilting) were not amplified.

The method used to calculate the view direction when amplifying the user's head movement can be seen in Figure 4. The z axis taken from the current HMD reading is reflected into the world XZ plane and rotated by the required amount. The z axis is then rotated to the required point in the YZ plane (calculated by multiplying the previous value of the axis in this plane by two). To ensure the user's head remained at the correct orientation, the x axis of the viewing parameters had exactly the same rotation applied to it as the z axis, and the y axis was set as the normal to the x and z axes.

A representation of the user's hand was required for him or her to complete the experimental task; in our experiment this was the purple sphere. Where a large amount of lateral head movement occurred, the situation could arise where the participants felt their hand should be within their visual field, but could not see it. For this reason, hand movements were also amplified in one of the conditions. A vector from the user's head to their hand was computed, and this vector was rotated using amplification, in exactly the same way that the view direction was modified, to obtain the amplified position of the hand.

3.3 Procedure

Each participant sat in front of and facing the Polhemus emitter, so that both his or her head and hand were at zero rotation when pointing directly towards the emitter (see Figure 6).

The participants first completed an untimed practice session, in which they were asked to 'shoot' five targets. This gave them the opportunity to familiarise themselves and adjust the inter-pupillary distance on the HMD. This was followed by nine consecutive test sessions, in which the time to complete each test was recorded. Finally, the participants were asked to answer a short questionnaire, which asked them whether any condition seemed more difficult to complete than the others, and whether they could cite any reasons for this.

3.4 Experimental design

The study used a within-subjects, 3×3 factorial design. The factors represent the three different 'amplification conditions': control (neither hand nor head movements amplified); head-only (only head movements amplified); head-and-hand (both head and hand movements amplified), and the positions in the environment where the targets were presented, which could be in one of three ways. These three differently positioned sets of targets were designed to be equivalent to one another. Participants undertook nine trials altogether: each amplification condition with each set of targets. The position of the targets and the type of amplification always differed from one trial to the next, to avoid participants remembering where the targets were, or completing two trials consecutively in the same condition. The order in which the participants undertook the trials was

calculated according to a Latin square, to ensure each undertook the trials in a different sequence, thus controlling further for practice and order effects.

3.5 Participants

Three female and ten male volunteers aged between 20 and 26 took part in the study. Nine of the participants had some prior experience with VR, and four had used an HMD before.

3.6 Equipment

The VR software used for the study was MAVERIK (Hubbold et al, 2001). The equipment used for the experiment consisted of a head mounted-display, a hand held 3-D mouse and calibrated Polhemus Long Ranger emitter. The visual display was provided by a Virtual Research Systems V8 Head Mount Display. The V8 provides a binocular 49.6 degree (horizontal) x 36.2 degree (vertical) FOV, with 100% overlap and an image resolution of 640 x 480. Images were presented with stereopsis. The average display frame rate was 15Hz. Shared memory was used to store the sensor readings. As two sensors were used in the current experiment, tracker latency in the worst case was 60ms.

4 Results

As the results were skewed away from a Gaussian distribution, the data was transformed using a natural logarithm prior to analysis. Data from each condition was pooled, and a mean calculated, leaving each participant with one score for each of the control, head-only and head-and-hand conditions. A repeated measures ANOVA was conducted, giving a highly significant result ($F_{2,24} = 20.79$, $p < 0.001$). Pair wise comparisons were then conducted to ascertain the extent of the differences between each of the conditions. Results of paired t-tests (quoting 2-tailed significance) were as follows: for the head-only and control conditions, $t = 3.80$, $p < 0.005$; for the control and head-and-hand conditions, $t = 2.40$, $p < 0.05$; for the head-only and head-and-hand conditions, $t = 7.63$, $p < 0.001$ ¹. The average times to hit all ten targets in each condition are shown in Table 1 and Figure 7.

	Control	Head-only	Head-and-hand
Mean	105.0	139.7	80.8
Median	93.0	137.7	95.0
Mode	95	137.7	N/A
Geometric Mean	100.7	128.6	78.9

Table 1: Mean, median, mode and geometric mean time (seconds) to hit all ten targets in each condition

¹ The difference between the *control* and *head-and-hand* scores was more pronounced in subject 10 than the other subjects. The data was reanalysed without these scores, and the results remained significant.

The responses to the questionnaire indicated that exaggerating movements in a VE is indeed a comfortable sensation for the user. Four of the participants said that the head and hand condition felt 'normal', and the control condition felt 'slowed down'. A further two indicated that they preferred the head and hand condition to the control condition. Only one participant claimed to find the condition in which 'a 360 degree turn of the head did not equate to 360 degrees in the virtual environment' more difficult.

Eleven of the participants reported that the condition in which head movements were amplified, but hand movements were not, made the trials much more difficult to complete. When viewed alongside the increased times that were obtained for these trials, there is strong evidence that amplifying head movements without making similar adjustments for the rest of the body produces an uncomfortable and unnatural sensation.

Figure 8 shows the mean time in seconds taken by each subject to complete the three conditions. On average, participants were able to complete those trials in which head and hand movements were amplified 21% faster than the control trials.

None of the participants reported after effects in the form of motion sickness or disorientation having completed the experiment. However, as they were not explicitly asked to discuss this aspect of their experience, and they were only immersed in the environment for an average of 16.3 minutes (5.3 minutes in the control condition, 7 minutes in the head-and-hand condition and 4 minutes in the head-only condition) the possibility of after effects cannot be ruled out. Learning effects were also exhibited to a limited degree, with 69% of subjects completing control and head-only trials progressively faster, and 38% of subjects completing head-and-hand trials progressively faster. However, it is not clear whether this was due to subjects becoming increasingly familiar with the layout of targets in the environment, or increasingly comfortable with the set-up of the equipment.

5 Discussion

The study was designed to answer two questions about the current implementation: does the technique feel comfortable to the users, and if yes, does it actually enhance their interaction with the VE? In both cases, the answer appears to be yes. In every case except one, participants indicated that the amplified condition felt very similar, or preferable, to the control condition, supporting the results of Jaekl et al (2002) and providing clear evidence that the technique does not cause sensory disruption. The fact that participants were able to complete the task significantly faster when their head movements were amplified also indicates that this technique can considerably improve performance in a VE.

The current study was conducted in very specific conditions: the user was sitting down rather than moving around, and had only a very abstract body representation. If amplifying head movements in VEs is to be genuinely useful, it must also be practical to use in situations where the user is navigating through the VE, and when the user has a more concrete body representation, such as an avatar. It remains to be seen whether

amplifying arm movements, which was shown here to be essential if the user wished to interact with the environment whilst head movement was exaggerated, are as practical when they are attached to a body in the virtual world.

However, it is certainly the case that this technique has promise. Despite the limitations to which it may be subject, the inability of most participants to detect when amplified head movements were active, coupled with the improved visual search performance in which it resulted, means that it does seem able to offer some compensation for the negative effects of the narrow FOV in HMDs.

6 References

Arthur, K. (1996). Effects of field of view on task performance with head-mounted displays. CHI '96 Proceedings on Human Factors in Computing Systems, pp 29-30. ACM, New York. ISBN: 0-89791-777-4.

Barfield, W., Hendrix, C., Bjornseth, O., Kaczmarek, K.A. & Lotens, W. (1995). A comparison of human sensory capabilities with technical specifications of virtual environment equipment. *Presence*, 4 (4), pp. 329-356.

DiZio, P. & Lackner, J.R. (1997). Design of Computing Systems: Cognitive Considerations. In Proceedings of the Seventh International Conference on Human-Computer Interaction (HCI International '97), 2, pp. 893-896. Elsevier Science Publishers, San Francisco.

Hendrix, C. & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence*, 5 (3), pp. 274-289.

R. Hubbard, J. Cook, M. Keates, S. Gibson, T. Howard, A. Murta, A. West, & S. Pettifer (2001). GNU/MAVERIK: A micro-kernel for large-scale virtual environments. *Presence*, 10 (1), pp. 22-34.

Iwamoto, I. & Tanie, K. (1997). A head-mounted eye movement tracking display and its image display method. *Systems and Computers in Japan*, 28 (7), pp.879-888.

Jaekl, P.M., Allison, R.S., Harris, L.R., Jasiobedzka, U.T., Jenkin, H.L., Jenkin, M.R., Zacher, J.E. & Zikovitz, D.C. (2002). Perceptual stability during head movement in virtual reality. Proceedings IEEE Virtual Reality 2002, pp. 149-155. IEEE Computer Society Press, Los Alamitos, CA.

Kline, P. & Witmer, B. (1996). Distance perception in virtual environments: Effects of field of view and surface texture at near distances. Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, pp. 1112-1116. The Human Factors and Ergonomics Society, Santa Monica, CA.

Mine, M.R., Brooks Jr., F. P., Sequin, C. H., Moving Objects in Space: Exploiting Proprioception In Virtual-Environment Interaction, Computer Graphics, SIGGRAPH '97 Conference Proceedings, 1997, pp.19-34. ACM, New York. ISBN: 0-89791-896-7.

O'Reagan, J.K. (1992). Solving the real mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, 46 (3), pp. 461-488.

Peruch, P. & Mestre, D. (1999). Between desktop and head immersion: functional visual field during vehicle control and navigation in desktop environments. *Presence*, 8 (1), pp. 54-64.

Poupyrev, I., Weghorst, S. & Fels, S. (2000). Non-isomorphic 3-D rotational techniques. *Proceedings of CHI' 2000*. pp. 540-547. ACM, New York. ISBN: 1-58113-248-4.

Poupyrev, I., Weghorst, S., Otsuka, T. & Ichikawa, T. (1998). Amplifying rotations in 3D interfaces. *Proceedings of CHI'99 Conference Abstracts, Late Breaking Results*. pp.256-257. ACM, New York. ISBN: 1-58113-243-3.

Pryor, H.L., Furness, T.A. & Viirre, E. (1998). The Virtual Retinal Display: a new display technology using scanned laser light. *Proceedings of the Human factors and Ergonomics Society, 42nd Annual Meeting, 2*, pp. 1570-1574. The Human Factors and Ergonomics Society, Santa Monica, CA.

Razzaque, S., Kohn, Z. & Whitton, M. (2001). Redirected walking. Department of Computer Science Technical Report TR01-007, University of North Carolina.

Rolland, J.P., Biocca, F.A., Barlow, T. & Kancherla, A. (1995). Quantification of adaptation to virtual-eye location in see-thru head-mounted displays. *Virtual Reality Annual International Symposium '95*, pp.56-66. IEEE Computer Society Press, Los Alamitos, CA.

Simons, D.J & Levin, D.T. (1997). Change Blindness. *Trends in Cognitive Sciences*, 1 (7), pp. 261-267.

Waller, D. (1999). Factors affecting the perception of inter-object distances in virtual environments. *Presence*, 8 (6), pp. 657-670.

Witmer, B. G., Baily, J. H.. & Knerr, B. W. (1994). Training dismounted soldiers in virtual environments: Route learning and transfer. ARI Technical Report 1022. U. S. Army Research Institute for the Behavioural and Social Sciences.



Figure 1: Virtual environment used in the experiment.

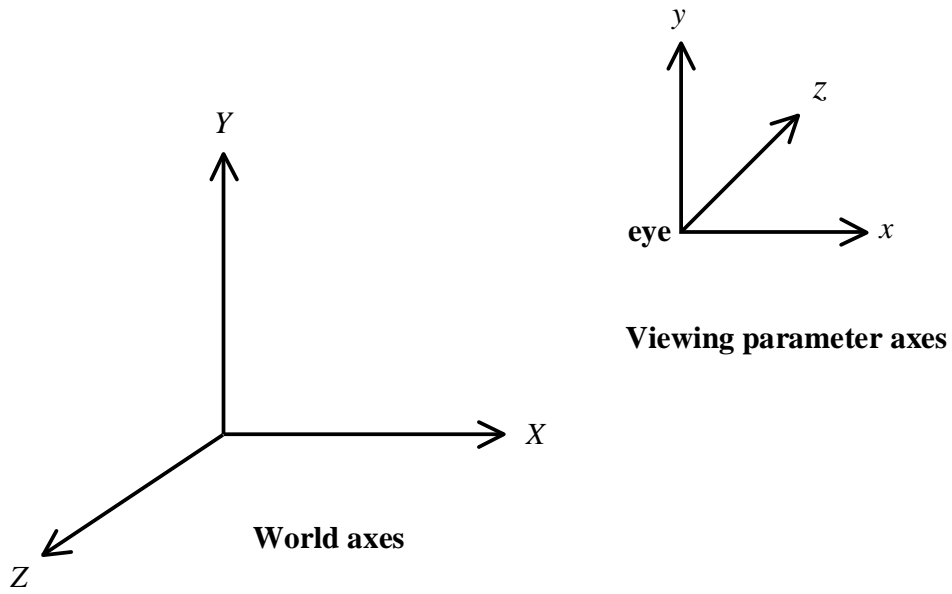


Figure 2: Viewing parameter axes in relation to world axes.

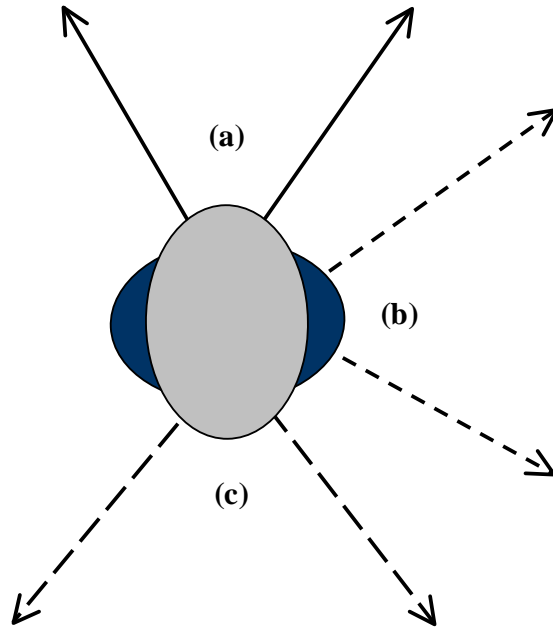
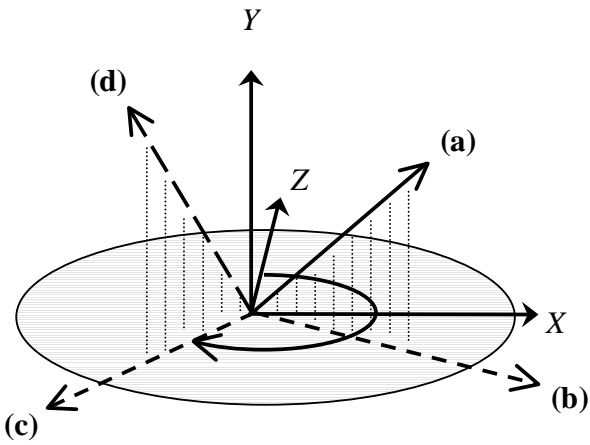


Figure 3: (a) original head direction and FOV, (b) rotated head and FOV, (c) amplified view direction.



X , Y and Z represent world x , y and z axes, where rotation = 0.

(a) Current HMD z axis.

(b) Reflected HMD z axis.

(c) New z axis, after lateral rotation has been amplified.

(d) New z axis, after lateral and vertical movement has been amplified.

Figure 4: Amplifying rotational movement of the head.

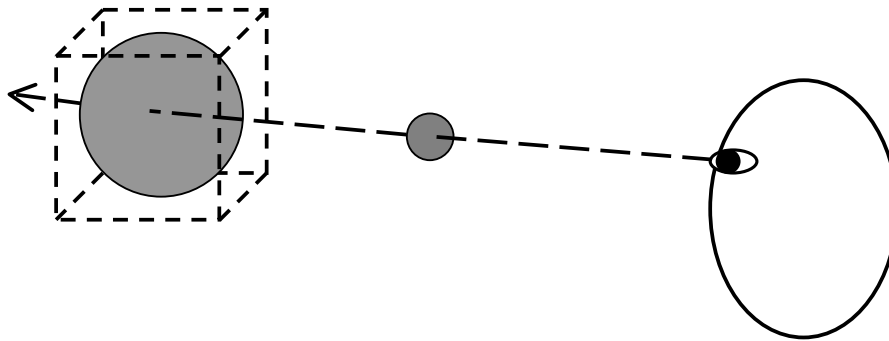


Figure 5: Line of intersection from the participant's eye point in the VE through the cursor, intersecting with the bounding box of the target sphere.

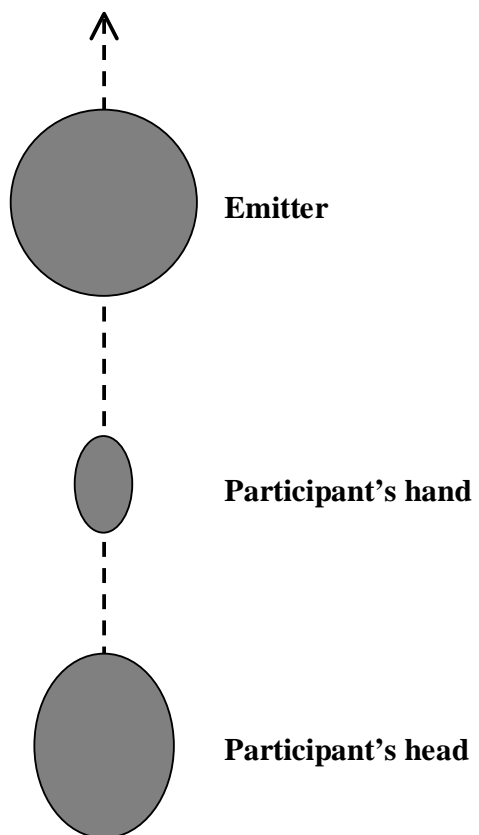


Figure 6: The participant's head and hand at zero rotation.

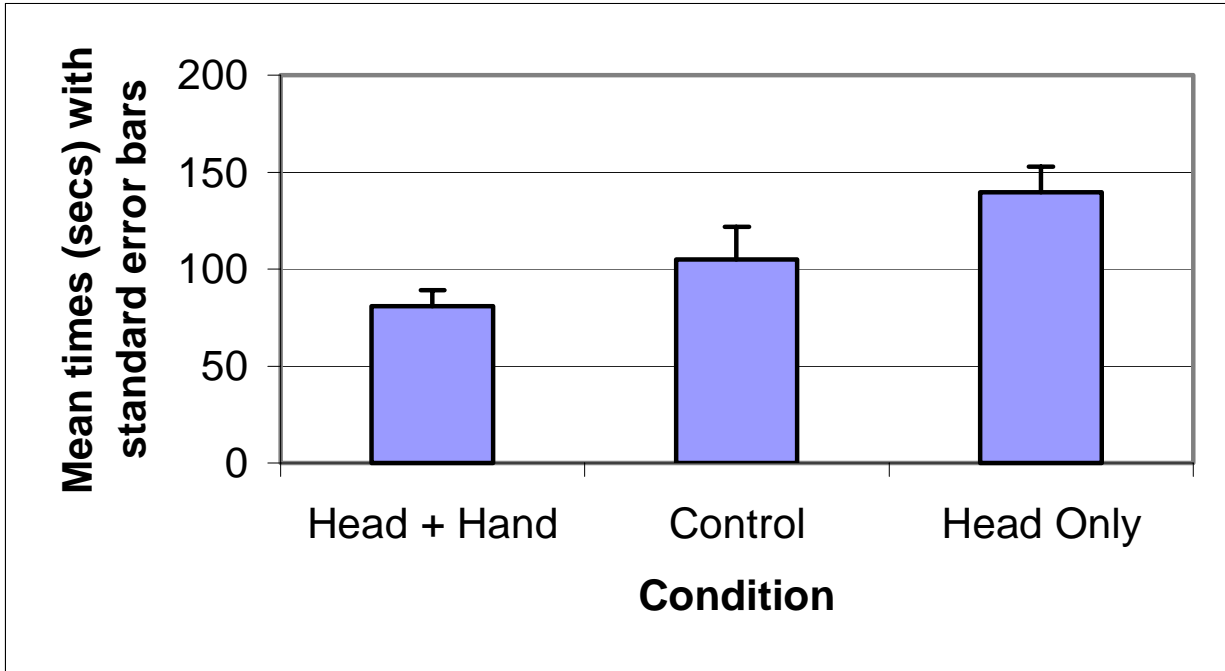


Figure 7: Overall mean time (seconds) to hit all ten targets in each condition.

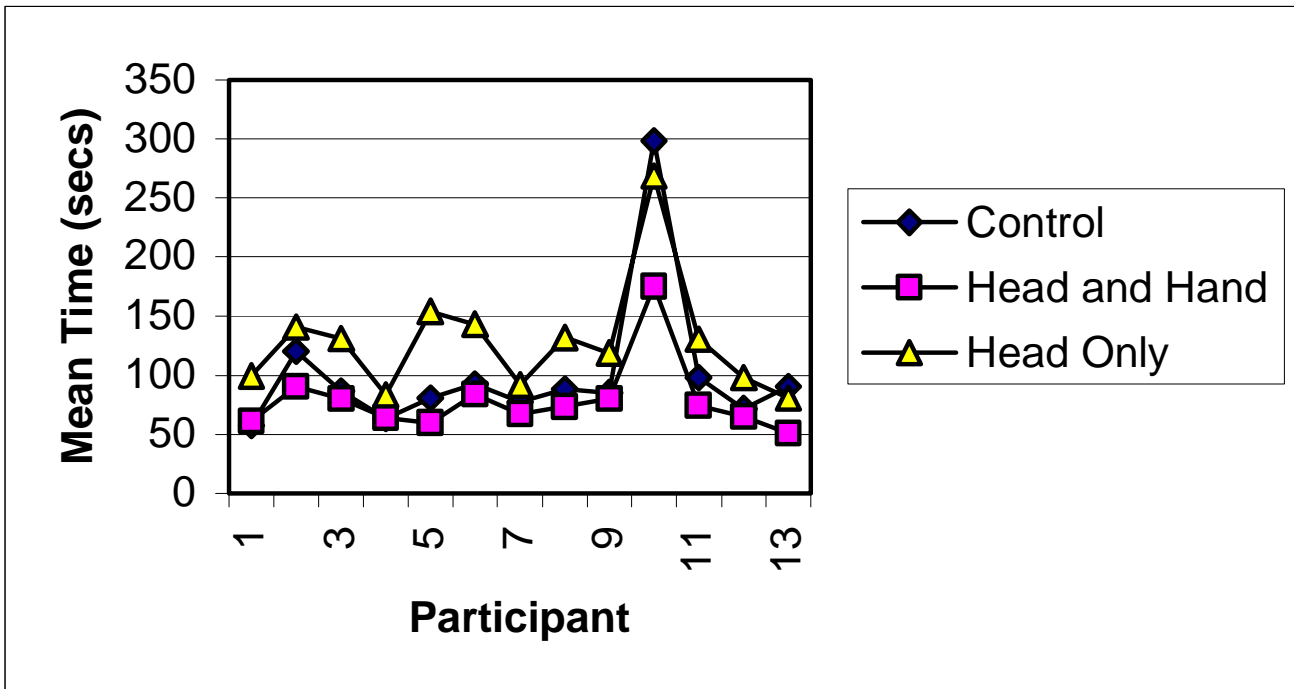


Figure 8: Mean time (seconds) taken by each participant to hit all ten targets in each condition.