Published in final edited form as:

Proc IEEE Virtual Real Conf. 2009 ; : 211–218. doi:10.1109/VR.2009.4811025.

Relating Scene-Motion Thresholds to Latency Thresholds for Head-Mounted Displays

Jason Jerald* and Mary Whitton†

Department of Computer Science, The University of North Carolina at Chapel Hill

Abstract

As users of head-tracked head-mounted display systems move their heads, latency causes unnatural scene motion. We 1) analyzed scene motion due to latency and head motion, 2) developed a mathematical model relating latency, head motion, scene motion, and perception thresholds, 3) developed procedures to determine perceptual thresholds of scene-velocity and latency without the need for a head-mounted display or a low-latency system, and 4), for six subjects under a specific set of conditions, we measured scene-velocity and latency thresholds, and compared the relationship between these thresholds. Resulting PSEs and JNDs of latency thresholds are in the same range of Ellis and Adelstein. The results are a step toward enabling scientists and engineers to determine latency requirements before building immersive virtual environments using head-mounted display systems.

Index Terms

H.1.2 [Models and Principles]: User/Machine Systems—Human Factors; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Ergonomics, Evaluation/Methodology, Theory and Methods, and User-Centered Design; I.3.6 [Computer Graphics]: Methodology and Techniques—Ergonomics and Interaction Techniques; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality; I.3.m [Computer Graphics]: Miscellaneous—Perception; J.4 [Social and Behavioral Sciences]: Psychology

1 Introduction

The real world remains stationary as we rotate our heads, and we perceive the world to be stable even when the image of the world moves on our retinas. Immersive virtual environments (IVE) should have this same property: the virtual scene should appear spatially fixed as users turn their heads. However, an IVE, in particular one viewed in a head-mounted display (HMD), may not appear stationary due to imprecise calibration (incorrect field-of-view, poor head tracking, etc.) or, the dominant cause [9], system latency.

Latency, the time required for a system to respond to a user's actions, lowers users' sense of presence in IVEs using HMDs [16]. When users move their heads, latency causes a mismatch between the user's current head pose and the delayed pose used to render the currently displayed frame. The result is that the virtual scene, which should be fixed in the world, is visually unstable. We call this instability *scene motion*.

*jjerald@cs.unc.edu

†whitton@cs.unc.edu

Although scene motion due to latency is well understood mathematically [9,2], perceptual sensitivity to this motion is not. If there is zero latency or no head motion, the virtual scene is spatially fixed. As users move their heads faster, scene motion due to latency increases, but sensitivity to scene motion decreases [3,12,11]. Little is known about the relationship between sensitivity to scene motion and latency-induced scene motion in HMDs.

In this work we look specifically at latency, scene motion caused by latency, and perception of that scene motion. We present methods that enable researchers to *quantify perception of scene motion and latency as a function of peak angular acceleration of the user's head*. Our methods can be used to determine latency requirements for IVE systems using HMDs. For our conditions, our results suggest an acceptable maximum end-to-end system latency in the range of 5 ms.

2 Background

Interspersed throughout this paper is information that supports simplifying assumptions made for our experimental design and analysis. The assumptions are called out in bold.

2.1 Latency Thresholds

Allison et al. [4] reported 50% latency thresholds for HMDs to be in the 180-320 ms range depending on head velocity. In their study, subjects performed single interval judgments—i.e., they did not compare each presentation to a reference stimulus, but relied on their own internal notion of when the scene appeared unstable.

Various experiments conducted at NASA Ames Research Center [8,2,7,15] reported latency thresholds for a judgment of whether scenes presented in HMDs are the same or different from a minimal-latency reference scene. Subjects viewed the scene while rotating their heads back and forth. These studies found that individual subjects' *point of subjective equality* (PSE—the amount of scene motion at which the subject is equally likely to judge a stimulus to be different from one or more reference stimuli) vary considerably (in the 0 to 80 ms range) depending on factors such as different experimental conditions, bias, type of head movement, and individual differences. They found just-noticeable differences (JND—the stimulus required to increase or decrease the detection rate by 25% from a PSE with a detection rate of 50%) to be in the 5 to 20 ms range. They also found that subjects are more sensitive to latency during the phase of head rotation when direction of head rotation reverses (when scene velocity due to latency peaks) than the middle of head turns (when scene velocity due to latency is smaller) [1].

2.2 Perceptual Moments

A question for our work is whether users perceive latency directly as a temporal quantity or perceive it indirectly by observing its visual consequence—unnatural scene motion. A *perceptual moment* is the smallest unit of time that an observer can sense [5]. The length of perceptual moments depends upon the sensory modality, stimulus, task, etc.

Efron [6] concluded that human visual-perceptual moments are 60 to 70 ms: In his study two flashes of light had to differ in length by 60 ms in order for users to judge one longer than the other. In a similar study, subjects were able to discriminate temporal differences in some cases as low as 10 ms [18]. However, subjects reported they made judgments, at least in part, based upon perceived brightness, instead of temporal cues alone.

There is also some suggestion that the perceptual moment becomes unstable when judgments involve more than one sensory modality [21]. Directly sensing time differences between visual stimuli and head turns might result in larger temporal thresholds. Adelstein

et al. [2] reported that subjects, at least in part, judge latency based indirectly upon scene motion resulting from latency.

We believe latency thresholds for HMDs to be lower than perceptual moments. For small latencies, we do not expect users to detect latency directly, but rather the consequences of latency—the visual instability of the virtual world as users move their heads. We assume subjects make judgments of latency based upon scene motion resulting from that latency rather than latency directly.

2.3 Sensitivity to Motion of Visual Stimuli

Velocity of visual stimuli is sensed directly by humans, whereas acceleration of visual stimuli is inferred through processing of velocity signals [13]. For perception of visual stimuli motion, most visual perception scientists assume acceleration is not as important a cue as velocity [19]. We assume scene velocity is a better measure of scene-motion thresholds than scene acceleration.

3 Mathematical Foundations: Latency, Head motion, and Scene Motion

Like any motion, scene motion can be quantified by displacement and derivatives of displacement. The magnitude of scene motion is dependent on both head motion and latency. In the equations in the next sections, θ is scene motion in degrees, φ is head motion in degrees, and Δt is latency in seconds.

We assume static worlds: No scene motion is caused by objects moving in the virtual world.

3.1 Displacement

Displacement error—Is the difference between where a displayed pixel appears to be in world space at a specific point in time and where the pixel would appear in world space at that same time if it were a static physical object in the real world.

For even moderate head motion, latency causes more displacement error than all other sources of displacement errors combined [9]. To simplify the analysis, we consider only latency-induced displacement error. Latency-induced displacement error from head rotation is larger than from head translation for all but the closest objects. We assume the scene is at a distance so that an analysis of head rotation, ignoring head translation, is sufficient.

Scene displacement—Is the average displacement error over all pixels. Since our analysis considers only pure rotation, scene displacement equals the displacement error of any individual pixel.

Scene displacement θ , in degrees, at current time t_0 is

$$\theta_{t_0} = \Delta\phi_{t_0} = \phi_{t_0} - \phi_{t_0 - \Delta t} \quad (1)$$

where φ is head rotation in degrees (where 0° is looking straight ahead) and Δt is latency in seconds.

Scene displacement θ can also be computed by integrating head velocity φ' over the latency interval:

$$\theta_{t_0} = \int_{t_0 - \Delta t}^{t_0} \phi' dt \quad (2)$$

For small latency Δt and/or constant head velocity ϕ' over the time of integration, scene displacement θ at any one instance in time t_0 is

$$\theta_{t_0} = \Delta t \phi'_{t_0}. \quad (3)$$

3.2 Velocity

Velocity error—Is displacement error differentiated with respect to time [1]. We measure velocity error in degrees/second. For a static world, velocity error is the velocity of the scene motion.

Scene velocity—Is the average velocity error in world space over all pixels in a frame.

For a static virtual world, any and all scene velocity is considered error. Note that error can be due to latency or, as is the case in our work, error can be intentionally injected into the system to test subject sensitivity to scene motion. We assume velocity error, like displacement error, is identical across all pixels in our scenes, so scene velocity is equal to the velocity error for any single pixel.

Scene velocity θ' due to latency is

$$\theta'_{t_0} = \Delta \phi'_{t_0} = \phi'_{t_0} - \phi'_{t_0 - \Delta t}. \quad (4)$$

Note that scene displacement θ due to latency (Equation 1) can be greater than zero while scene velocity θ' due to latency is zero. This occurs when a user moves her head with constant velocity such that $\phi'_{t_0} = \phi'_{t_0 - \Delta t}$. In such a case the scene appears to be at a constant offset from where it should appear in the world.

Scene velocity θ' can also be computed by integrating head acceleration ϕ'' over the latency interval Δt such that

$$\theta'_{t_0} = \int_{t_0 - \Delta t}^{t_0} \phi'' dt. \quad (5)$$

For small latency Δt and/or constant head acceleration ϕ'' over the time of integration, scene velocity θ' at any one instance in time t_0 is

$$\theta'_{t_0} = \Delta t \phi''. \quad (6)$$

Head acceleration is not normally constant. However, this formula is useful because it shows that for a constant latency Δt , scene velocity is proportional to head acceleration weighted by that latency. Identifying and measuring peak head acceleration during a head turn means

we can, for a known latency, estimate the peak scene velocity occurring during that turn, and vice-versa.

Based on Equation 6, Figure 1 shows peak scene velocity plotted as a function of peak head acceleration for three values of latency Δt . Peak scene velocity increases linearly with peak head acceleration. Latency Δt is simply the slope of the line.

3.3 Real Head Motion

Figure 2 (top) shows yaw head motion for a typical subject. The subject starts yawing her head from the left shortly before two seconds and stops turning at approximately three seconds. We assume such head turns are typical of head yaws.

To demonstrate the effect of 100 ms of latency, the graph also shows the same head motion delayed by 100 ms. The difference between the head angle and delayed head angle is the scene displacement (Equation 1). Likewise (middle), the difference between the head velocity and delayed head velocity is the scene velocity (Equation 4).

Notice that head velocity peaks during the center of the head turn, head acceleration (bottom) peaks near the beginning of the head turn, and head deceleration peaks near the end of the head turn. As the head accelerates, the scene moves in the same direction as the head (both positive values). The scene displacement then peaks and remains relatively stable while head acceleration is near zero. Once the head starts to decelerate, scene velocity drops below zero, and the scene starts to move against the direction of the head turn, towards its correct position in space.

Since scene velocity is approximated by head acceleration scaled by latency (Equation 6), peak scene velocity occurs near the beginning and end of head turns.

With latency, the scene moves *with* the head turn as the head accelerates and the scene moves *against* the direction of the head turn as the head decelerates. Previous research has shown that subjects are more sensitive to scene motion when the scene moves against the direction of the head turn than when the scene moves with the direction of the head turn [10,20]. Thus we assume users are more sensitive to latency at the end of head turns than the beginning of head turns. We measure sensitivity to scene motion at the end of head turns when both scene motion due to latency and sensitivity to scene motion peak.

3.4 Relating Scene-Motion Thresholds to Latency Thresholds

This section derives a mathematical relationship between scene-motion detection thresholds and latency detection thresholds, both as functions of peak head acceleration.

1 shows a hypothetical linear relationship between scene-velocity threshold and peak head acceleration that is based on results of a previous study [11]. Projecting the point of intersection of the scene-velocity threshold line and one of the latency-induced scene velocity lines onto the x-axis gives the value of peak head acceleration that would produce noticeable scene motion in a system with that amount of latency. The analysis task is then to find the slope of the scene velocity line (equivalent to latency) that intersects the threshold line at a given peak head acceleration.

The linear equation for the scene-velocity threshold line is

$$\theta'_{threshold} = m_{threshold} \phi'' + b_{threshold} \quad (7)$$

where $m_{threshold}$ is the slope of the scene-velocity threshold curve, ϕ'' is peak head acceleration, and $b_{threshold}$ is the y intercept of the scene-velocity threshold line.

To determine the latency required to detect motion, Equations 6 and 7 are set equal to each other and solved for latency:

$$m_{threshold}\phi'' + b_{threshold} = \Delta t \phi'' \quad (8)$$

$$\Delta t = m_{threshold} + b_{threshold} \left(\frac{1}{\phi''} \right) \quad (9)$$

Figure 3 shows the latency threshold function for the hypothetical scene-velocity threshold line shown in Figure 1. The latency threshold curve shows that large amounts of latency with little head acceleration will not be detected (because scene velocity is near zero), whereas small amounts of latency will be detected only with large head accelerations. For the given hypothetical scene-velocity threshold line, a subject with head acceleration of $100^\circ/sec^2$ will detect latency of 20 ms or more.

4 Experiment

We designed an experiment to investigate the mathematical model proposed in Equation 9. The goals of the study were to validate our model and compare thresholds resulting from three scene-motion conditions. In a repeated-measures study, we experimentally determined *psychometric functions* (relationships of detection probabilities to stimulus intensities) for three models of scene motion over ten ranges of head acceleration.

4.1 Scene-Motion Conditions

In our test trials we inserted scene-motion to simulate the visual effects of latency. The beginning of a trial was signaled by a tone and metronome beats occurred every second so that the subjects knew to start turning their heads at two seconds and stop turning their heads at three seconds. The scene was visible for the interval 2.5 to 3.2 seconds.

The scene moved against the direction of the head turn and was controlled by one of three scene velocity conditions (Three conditions with the same peak scene velocity are graphed in Figure 4):

- **Constant Scene Velocity (Constant)** - The scene is already moving when it appears and it moves at a constant velocity until it disappears.
- **Gaussian Scene Velocity (Gaussian)** - The speed of the scene ramps up and down after it appears. The velocity ramp is defined by a truncated Gaussian curve ($\sigma = 0.1$ seconds) centered at the end of the intended head turn.
- **Latency-Induced Scene Velocity (Latency)**, the scene-velocity is manipulated in real-time based on the user's measured head motion and Equations 1 and 4. Since latency controls scene motion in this condition, both latency thresholds measured in milliseconds and scene-velocity thresholds measured in degrees/second were determined from the same trials.

The Constant and Gaussian conditions are independent of head motion whereas the Latency condition depends upon head motion. The left column of Table 1 shows the three scene-velocity conditions, including the two measures from the Latency condition.

4.2 Variables and Experimental Data

In our experiment, both scene motion and head motion vary continuously during head turns. However, a single value for each is needed to generate the data points that establish the psychometric function. Based on our mathematical analysis described in Section 3, we choose to use peak scene-velocity and peak head-acceleration as measures of scene motion and head motion.

The stimulus to be detected in our study was scene motion: Test trials either contained a random amount of scene motion or were *catch trials* containing no scene motion. The scene motion in a test trial was defined by one of the three conditions described above. Trials for the three conditions were randomly intermixed to reduce bias. Head turns from left to right and from right to left were also randomly intermixed. To simplify our analysis, we assume an analysis using only head yaw, ignoring pitch and roll, is sufficient for analysis.

We chose to use measures of PSE and JND determined from psychometric functions as done in the NASA experiments [8,2,7,15]. We used visual cues during the per-trial instructions to suggest a head turn speed and amplitude. This insured, over the 210 test trials per scene velocity condition (630 test trials for each participant), that users made head turns with a wide range of head-acceleration values.

Starting with head-pose data logged from the tracker system, we computed the peak-head acceleration occurring during the head turn in each test trial. We then divided the measured peak head-acceleration data into ten bins. Each bin contained data for 21 judgments of scene motion, and those 21 data points were used to generate a psychometric function (see Figure 6).

We derived PSE and JND threshold values from the psychometric functions as shown in Figure 7. Linear regressions of scene-velocity thresholds versus peak head accelerations were then computed. Applying Equation 9 transforms these regressions into latency threshold curves relating peak head acceleration and latency thresholds.

Table 1 shows experiment conditions, measures, and peak head-acceleration bins for a single subject.

4.3 Materials

Materials were similar to those used in previous experiments to determine scene-velocity thresholds for different types of head turns [10,11].

4.3.1 The Display Environment—We designed a system to emulate a zero-latency HMD. A BARCO CRT projector displayed images of approximately one lux onto a world-fixed planar surface four meters in front of the seated subject (Figure 5). The CRT projector was chosen for its fast phosphor response and decay times. With the CRT, there was no ghosting, and no light was projected for black pixels, which is not the case for LCD and DLP projectors. A Virtual Research V8 HMD was modified by replacing the display elements with cardboard cutouts so that subjects could see through the casing to the world-fixed display. The cutouts limited the user's field-of-view to 48° by 36°, as in an unmodified V8 HMD. Total weight of the modified HMD with attached tracker sensor was 0.6 kg.

All object-relative cues were removed by darkening the room and providing a uniform visual field. Material identical to the display surface was curved around the subject in a quasi-cylindrical shape. The floor was similarly covered, so that the subject saw a uniform visual field. Subjects confirmed they could not see any visual elements other than the computer-generated scene presented by the projector. Since only the computer-generated scene was visible, subjects could make only subject-relative judgments. A ten-lux green screen was shown between trials to prevent dark adaptation.

4.3.2 The Visual Stimulus—A simple 2D visual scene (a green monochrome rhombus with bisectors and a 20° horizontal span) was chosen because:

- A simple scene minimizes the rendering time component of the unavoidable end-to-end system latency.
- Prior work has not been able to find significant differences in scene-motion thresholds across scene complexities [7,15].
- A 2D scene minimize depth issues (stereo cues, motion parallax, incorrect head tracking, etc.) that could confound the results.
- Tearing effects, if any, are more noticeable in vertical lines than diagonal lines. We eliminated vertical lines so that participants would respond to real scene motion and not to visual artifacts caused when the system does not wait to swap display buffers during the vertical retrace interval (our system swapped buffers every 50 scanlines approximating a just-in-time pixels [17] implementation in order to remove shearing artifacts).

The horizontal starting position of the scene for each presentation was set to be half-way between 0° and the extent of the suggested head turn so that the scene would be presented approximately in the center of the field of view.

4.3.3 Scene Motion—Scene motion was controlled by the conditions described in Section 4.1. For our reference scene and catch trials (i.e. trials with no injected stimulus), the scene remained stable in space—no scene motion occurred due to latency (as it would have in an HMD) because we used a world-fixed projector to display our scenes. For the Constant condition and the Gaussian condition, head position/orientation did not affect the position or velocity of the visual scene. For the Latency condition, scene motion was controlled by Equations 1 and 4.

The range of scene velocities for each participant was chosen based on judgment responses given during training. The range was adjusted until users gave responses between zero and near 100% correct; this range is necessary for generating the psychometric functions. As the experiment progressed, the maximum scene-velocity value for a subject was sometimes increased or decreased to get more useful data.

4.3.4 Head Tracking—Head orientation was determined by a 3rdTech HiBall 3000 tracking system. The tracking data were used to check for acceptable head rotations, to record motion for post analysis, and as a parameter used to set scene motion velocity in the Latency condition.

4.4 Methods

Thresholds may vary substantially between subjects, so we chose a repeated measures design (all subjects experienced all conditions).

4.4.1 Head Turns—We asked subjects to yaw their heads over a period of one second starting and ending at designated positions. During the display of the reference/instructions scene, visual and audio cues suggested a target head turn amplitude and speed. The subjects optionally practiced head turns before each trial and then pressed a button to start the trial. The reference/instructions scene then disappeared and the subject rotated her head in time to auditory cues as had been practiced. If head motion got out of sync with the auditory cues, the trial was immediately canceled, and the reference/instruction scene for a new trial appeared.

The suggested head turns ranged in amplitude between 1 and 45 degrees. This resulted in moderate head motions with head accelerations less than $361^{\circ}/sec^2$. Head accelerations in this study were not near peak head accelerations possible for humans (fighter pilots can rotate their heads in the range of $2000^{\circ}/sec^2$ with peaks of $6000^{\circ}/sec^2$ [14]). Our mathematical model presented in section 3.4 is most interesting for small to moderate head rotations. We assumed that smaller head rotations are representative of head motion during everyday activities.

We trained subjects to start turning their heads at two seconds after the start of the trial and stop moving their heads at three seconds. Since our variables head acceleration and scene velocity peak near the end of head turns (at about 3 seconds), we presented the test scene from 2.5 to 3.2 seconds.

4.4.2 Data Collection—A method of constant stimuli (random stimulus levels are presented independent of previous judgments) determined the random ordering of trials. Judgments consisted of a yes/no task (Did the scene seem to move relative to the stable reference scene presented between trials?) along with a confidence rating (1 to 3). This resulted in six possible judgment levels per trial (0%, 20%, 40%, 60%, 80%, 100%). Such a design contains bias, but we attempted to keep bias constant across trials for each individual by randomly interleaving trials from all three conditions.

The wording of the judgment question was:

- Did the scene seem to move or did it seem not to move?
 - The scene seemed to not move
 - * Left button
 - The scene seemed to move
 - * Right button

After pushing the left or right mouse button, the subject rated her confidence in that judgment:

- How confident are you that the scene seemed to move on a scale of 1 to 3?
 - 1. I guessed that the scene seemed to move
 - * Left Button
 - 2. The scene seemed to move, but I could be wrong
 - * Middle Button
 - 3. The scene certainly seemed to move
 - * Right Button

A similar question was asked if the subject selected the “scene seemed to not move” with the word “not” inserted into the rating options.

4.5 Participants

Eight subjects (six male and two female) participated. Data for two of the male subjects were too noisy for successful analysis. All subjects were naive to the experimental conditions. Three subjects had participated in similar motion perception experiments before. All subjects judged 630 trials—210 judgments for each condition. Total time per subject, including consent form, instructions, training, experiment sessions, breaks, and debriefing, took less than six hours.

4.6 Results

False alarm rates (judgment of scene motion when no scene motion was present) were less than 12% for all six subjects.

For each subject, we fit 40 psychometric functions to 30 bins (10 bins $\times$ 3 conditions). We extracted 40 PSE and 40 JND values from these psychometric functions (240 psychometric functions total). Figure 6 shows one of a subject's psychometric function.

For scene-velocity thresholds, 30 PSEs and 30 JNDs (10 per condition) were computed. Row 1 of Figure 7 shows these thresholds for a single subject for the three scene-motion conditions described in section 4.1.

For the Latency condition, we also determined 10 PSEs and 10 JNDs by fitting by fitting psychometric functions to latency instead of peak scene velocity. Latency thresholds are plotted in Row 2 of Figure 7.

The mean latency JND across all subjects and all bins (60 JNDs) was 16.6 ms with a standard deviation of 9.7 ms. The data was positively skewed (median = 16.3) with the smallest JND of 3.2 ms and the maximum JND of 51.0 ms.

The mean PSE across all subjects and all bins (60 PSEs) was 38.9 ms with a standard deviation of 18.1 ms. The data was less skewed than the JNDs (median = 35.5 ms) with the smallest PSE of 9.9 ms and the maximum PSE of 103.2 ms.

The ten scene-velocity threshold values for each condition were then fit to a line using standard linear regression. The scene-velocity threshold lines were then converted to latency threshold curves using Equation 9. These latency threshold curves are plotted in Row 2 of Figure 7.

For all subjects, the directly-measured latency-threshold values were all “close” to the curves determined from the latency-induced scene-velocity threshold lines. While not a statistical measure, this “closeness” leads us to believe that further study of (Equation 9) and use of peak scene-velocity as a measure of latency thresholds will be worthwhile.

Figure 8 shows mean thresholds across all subjects. There is an evident pattern that the Constant condition thresholds are lower than the other two conditions.

JNDs were consistently positively skewed whereas PSEs were not. We performed a log transform on the JND values to make the distribution more normal so that we could perform parametric statistical tests. A two-way ANOVA showed that thresholds were significantly affected by peak head acceleration bin (PSE: $p < 0.001$, JND: $p < 0.001$) and by condition (PSE: $p < 0.01$, JND: $p = 0.013$).

To further investigate the data we matched thresholds across subjects and bin rank for 60 samples per condition (6 subjects $\times$ 10 bins). We then performed six pair-wise t-tests with $\alpha = 0.05/6 = 0.0083$ (Bonferroni correction). The Constant condition had lower PSEs and JNDs than both the Latency condition (PSE: $p < 0.001$, JND: $p < 0.001$) and the Gaussian Condition (PSE: $p = 0.003$, JND: $p < 0.001$). The Constant condition results in the lowest scene-velocity thresholds.

5 Discussion

5.1 Latency Thresholds

The latency JND mean of 16.6 ms and minimum of 3.2 ms out of 60 JND values suggest that end-to-end system latency in the 5 ms range is sufficiently low to be imperceptible in HMDs.

The range of our latency thresholds match better with the latency thresholds from the NASA experiments [8,2,7,15] than with the measurements of Allison et al. [4]. We believe the differences between our results and those of Allison et al. may be because their experiment differed from ours in that, among other things, they provided no stable reference scene against which to judge scene motion.

5.2 Individual Differences

Some subjects had more difficulty judging scene motion than others. The data of two subjects could not be used due to the noise in their judgments; it appeared these two subjects were largely guessing. Inconsistency in judgments manifests itself in large JNDs and as poor fits between the data and the linear regressions of scene-velocity thresholds. We do not believe that judgment inconsistency was necessarily due to lack of motivation, as one of the subjects whose data could not be analyzed went through more training than the others—at the beginning of training, he sincerely thought there was scene motion when there was none.

We selected the particular subject whose data is shown in rows 1 and 2 of Figure 7 because:

- The subject was one of the more sensitive subjects. We are interested in determining thresholds for the most sensitive of the population so that latency requirements for HMDs can be determined that will be appropriate for a large part of the population of users.
- The subject had low variance in terms of scene-motion judgments and had head turns that matched the suggested turns well.

The diversity in sensitivity to scene motion makes it difficult to conduct and analyze experiments between subjects, and it means that permissible latency values must be established using subjects who most sensitive to scene motion.

5.3 Differences between Conditions

We suspect the Constant Condition resulted in smaller thresholds because peak scene velocity occurs over the entire display interval instead of for a shorter time as in the other conditions. The lower thresholds may also be because humans can track visual targets with constant velocity more easily than they can track targets exhibiting less predictable motions.

The fact that no statistical differences were found between the Gaussian and Latency conditions does not prove that the Gaussian condition may be used to predict latency thresholds. Further investigation is required to determine if the correspondence of data from these two conditions is high enough for this use.

5.4 Limitations

Our assumptions may not hold in all cases. E.g., Users may move their heads in a different manner than our subjects did. We suspect minor violations of the stated assumptions have relatively little effect compared to differences in sensitivity across subjects.

Other factors could be added to our model, such as scene/head displacement, scene acceleration, and temporal differences such as when the head is stopped and the image keeps moving.

The thresholds reported are average responses over many trials; fatigue may have led to inattention and inaccurate responses. It may be that rested users can notice smaller scene motions or latencies than fatigued ones. The current work did not include any analysis of judgment accuracy over the course of the 630 trials per subject.

We expect latency thresholds to vary under different experimental and application settings. The thresholds suggested here provide guidelines for latency, but we do not claim that these thresholds will necessarily hold for higher-contrast scenes, wide field-of-view HMDs, different levels of experience in HMDs and VEs, etc.

5.5 Conclusions

This paper presents a new method for measuring scene-motion thresholds and latency thresholds for HMDs without the need for a low latency HMD.

Our observations and data suggest that latency thresholds differ significantly for different people. Given the diversity of scene-motion thresholds between subjects, we suggest that, when possible, engineers design systems with latency below that perceived by the most sensitive subjects.

The work reported here adds to our understanding of the relationship between perception of latency-induced scene motion and head-motion. We encourage scientists and engineers to use the methods described in this paper to establish latency requirements for their specific conditions.

Acknowledgments

Support for this work was provided by a LINK Foundation Fellowship, an NC Space Grant Fellowship, the Office of Naval Research, and the NIH National Institute for Biomedical Imaging and Bio-engineering. We thank Fred Brooks, Steve Ellis, Bernard Adelstein, Dennis Proffitt, Jonathan Bakdash, Chris Wiesen, David Harrison, and John Thomas for valuable discussions and help with experimental design, analysis, and review. We thank the anonymous reviewers for their extensive feedback and the UNC Effective Virtual Environments (EVE) research group for their support.

References

1. Adelstein, BD.; Burns, EM.; Ellis, SR.; Hill, MI. Latency discrimination mechanisms in virtual environments: Velocity and displacement error factors. The 49th Annual Meeting of the Human Factors and Ergonomics Society; 2005. p. 2221-2225.
2. Adelstein, BD.; Lee, T.; Ellis, SR. Head tracking latency in virtual environments: Psychophysics and a model. The 47th Annual Meeting of the Human Factors and Ergonomics Society; 2003. p. 2083-2987.
3. Adelstein, BD.; Li, L.; Jerald, J.; Ellis, SR. Suppression of head-referenced image motion during head movement. The 50th Annual Meeting of Human Factors and Ergonomics Society; 2006. p. 2678-2682.
4. Allison RS, Harris LR, Jenkin M, Jasiobedzka U, Zacher JE. Tolerance of temporal delay in virtual environments. *IEEE Virtual Reality*. 2001:247–253.

5. Coren, S.; Ward, LM.; Enns, JT. Sensation and Perception. 5th. Harcourt Brace College Publishers; 1999.
6. Efron R. The duration of the present. *Annals of the New York Academy of Science*. 1967; 138:713–729.
7. Ellis, SR.; Mania, K.; Adelstein, BD.; Hill, MI. Generalizeability of latency detection in a variety of virtual environments. The 48th Annual Meeting of the Human Factors and Ergonomics Society; 2004. p. 2083-2087.
8. Ellis SR, Young MJ, Adelstein BD, Ehrlich SM. Discrimination of changes in latency during head movement. *Computer Human Interaction*. 1999:1129–1133.
9. Holloway R. Registration error analysis for augmented reality. *Presence: Teleoperators and Virtual Environments*. 1997; 6(4):413–432.
10. Jerald, J.; Peck, T.; Steinicke, F.; Whitton, M. Sensitivity to scene motion for phases of head yaws. APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization; New York, NY, USA: ACM; 2008. p. 155-162.
11. Jerald J, Steinicke F, Whitton M. Scene-motion thresholds correlate with angular head motions for immersive virtual environments. *Advances in Computer-Human Interaction*. 2009 page to appear.
12. Li, L.; Adelstein, BD.; Ellis, SR. Applied Perception in Graphics and Visualization. Boston, MA: ACM Press; 2006. Perception of image motion during head movement; p. 45-50.
13. Lisberger SG, Movshon JAA. Visual motion analysis for pursuit eye movements in area mt of macaque monkeys. *The Journal of Neuroscience*. 1999; 19(6):2224–2246. [PubMed: 10066275]
14. List, UH. Final technical paper, Operations Training Division. Williams Air Force Base; Arizona: December. 1983 Nonlinear prediction of head movements for helmet-mounted displays. At top: AFHRL-TP-83-45
15. Mania K, Adelstein BD, Ellis SR, Hill MI. Perceptual sensitivity to head tracking in virtual environments with varying degrees of scene complexity. *ACM Applied Perception in Graphics and Visualization*. 2004:39–47.
16. Meehan, M.; Brooks, F.; Razzaque, S.; Whitton, M. IEEE Virtual Reality. Los Angeles, CA: 2003. Effects of latency on presence in stressful virtual environments; p. 141-148.
17. Miné M, Bishop G. Just-in-time pixels. Technical Report TR93-005, the University of North Carolina at Chapel Hill. Department of Computer Science. 1993
18. Nillson TH. Visual temporal discriminations of brief pulse intervals. *Vision Research*. 1979; 19(41):41–48. [PubMed: 419700]
19. Regan, DM.; Kaufman, L.; Lincoln, J. Ch. 19, motion in depth and visual acceleration. In: Boff, KR.; Kaufman, L.; Thomas, JP., editors. *Handbook of Perception and Human Performance*. Wiley-Interscience; New York: 1986.
20. Steinicke, F.; Bruder, G.; Jerald, J.; Frenz, H.; Lappe, M. Analyses of human sensitivity to redirected walking. ACM Symposium on Virtual Reality Software and Technology (VRST); ACM Press; 2008. p. 149-156.
21. Ulrich R. Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception and Psychophysics*. 1987; 42:224–239. [PubMed: 3671048]

Figure 1.

Peak scene velocity as a function of peak head acceleration is linear with slope equal to latency. A hypothetical scene-velocity threshold line, based on results of Jerald et al. [11], is also shown.

Figure 2. Measured head motion and computed scene motion due to 100 ms of latency for a single subject. Top—displacement; middle—velocity; and bottom—acceleration.

Figure 3. Latency thresholds determined from Equation 9 and the hypothetical scene-velocity threshold line plotted in Figure 1.

Figure 4. Scene motion for the three experimental conditions. The scene is only visible from 2.5 seconds to 3.2 seconds when the subject is decelerating her head as shown in Figure 2.

Figure 5.
A subject and scene as the subject yaws his head.

Figure 6.

Confidence ratings versus peak scene velocity for a range of a subject's peak head acceleration (bin three of ten). This subject was fairly confident in her judgments in this case—she did not guess in her judgments resulting in no confidences of 40% or 60%. A cumulative Gaussian function fit to this data is the psychometric function. The PSE and JND are extracted from the psychometric function.

Figure 7. Scene-velocity thresholds (Row 1) and latency thresholds (Row 2) for a single subject. The vertical lines represent the boundaries of the bins used for analysis.

Figure 8.
Mean scene-velocity thresholds (Row 1) and latency thresholds (Row 2) for all subjects.

Experiment conditions for each subject. Three conditions resulted in four psychometric functions for ten ranges of head motion. The ranges of head motion varied per subject and per condition since an equal number of judgments were fit to each bin. For the Latency condition, two psychometric functions (one in degrees/second and one in milliseconds) were computed. PSEs and JNDs were extracted from the psychometric functions.

Table 1

Condition		Measure	Peak Head Acceleration (deg/sec^2) Bin									
			1	2	3	4	5	6	7	8	9	10
	Constant (deg/sec)	PSE										
		JND										
	Gaussian (deg/sec)	PSE										
		JND										
	Latency (deg/sec)	PSE										
		JND										
	Latency ($msec$)	PSE										
		JND										