

Scene-Motion- and Latency-Perception Thresholds for Head-Mounted Displays

by
Jason J. Jerald

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2010

Approved by:

Frederick P. Brooks Jr., Advisor

Mary C. Whitton, Reader

Bernard D. Adelstein, Reader

Stephen R. Ellis, Committee Member

Gary Bishop, Committee Member

Anselmo A. Lastra, Committee Member

© 2010
Jason J. Jerald
ALL RIGHTS RESERVED

Abstract

**JASON J. JERALD: Scene-Motion- and Latency-Perception Thresholds
for Head-Mounted Displays.**

(Under the direction of Frederick P. Brooks Jr.)

A fundamental task of an immersive virtual environment (IVE) system is to present images of the virtual world that change appropriately as the user's head moves. Current IVE systems, especially those using head-mounted displays (HMDs), often produce spatially unstable scenes, resulting in simulator sickness, degraded task performance, degraded visual acuity, and breaks in presence. In HMDs, instability resulting from latency is greater than all other causes of instability combined. The primary way users perceive latency in an HMD is by improper motion of scenes that should be stationary in the world. Whereas latency-induced scene motion is well defined mathematically, less is understood about how much scene motion and/or latency can occur without subjects noticing, and how this varies under different conditions.

I built a simulated HMD system with zero effective latency—no scene motion occurs due to latency. I intentionally and artificially inserted scene motion into the virtual environment in order to determine how much scene motion and/or latency can occur without subjects noticing. I measured perceptual thresholds of scene-motion and latency under different conditions across five experiments.

Based on the study of latency, head motion, scene motion, and perceptual thresholds, I developed a mathematical model of latency thresholds as an inverse function of peak head-yaw acceleration. Psychophysics studies showed that measured latency thresholds correlate with this inverse function better than with a linear function.

The work reported here readily enables scientists and engineers to, under their particular conditions, measure latency thresholds as a function of head motion by using an off-the-shelf projector system. Latency requirements can thus be determined before designing HMD systems.

To my parents, Rick and Susan Jerald. None of this work would have been possible without them, in more ways than one.

Acknowledgments

Over the eight years that I have worked towards my Ph.D., many people have helped me. In particular, I thank all my committee members—Fred Brooks, Mary Whitton, Steve Ellis, Dov Adelstein, Anselmo Lastra, and Gary Bishop for their interest in this work, and their inspiration, guidance, patience, and encouragement. There are many others to whom I am also grateful:

- Neeta Nahta for her understanding, support, love, and patience.
- Sharif Razzaque, Eric Burns, Luv Kohli, Tabitha Peck, Dorian Miller, and Frank Steinicke for their work in redirection techniques and countless discussions of how our work relates.
- The entire Effective Virtual Environments (EVE) group during my studies at the UNC-Chapel Hill Computer Science Department, including Chris Oates, Jeff Feasel, Jeremy Wendt, Paul Zimmons, Ben Lok, Jess Martin, Laura Kassler, Angus Antley, Greg Coombe, Drew Chen, and Mark Harris.
- The Department of Computer Science staff including David Harrison, John Thomas, Kurtis Keller, Murray Anderegg, Mike Stone, and Kelli Gaskill for their technical and non-technical support.
- Henry Fuchs for inspiration and creative ideas outside the box.
- Russ Taylor for help with VRPN and other technical and intellectual issues.
- Andrei State for his patience and willingness to work in a dark room so that my experiments could proceed.
- Greg Welch as my first graduate academic advisor.
- The subjects who volunteered for the experiments.
- Sam Krishna for his coaching in helping me to understand that integrity is power, challenging me to be unreasonable, and holding me accountable.
- ACM SIGGRAPH for helping me to understand that computer graphics is about the people as much as it is about computers.

- The anonymous reviewers who reviewed publications resulting from this work.
- Paul Mlyenic and Digital Artforms for providing the final push to finish and motivation to continue pursuing my passions of interactive computer graphics and advanced human-computer interfaces.
- My parents, for their unwavering and unconditional support.

I am also grateful for support and funding from:

- HRL Laboratories
- The National Physical Science Consortium
- The Latané Center for Human Science
- The San Jose State University Foundation
- The Naval Research Laboratory
- NASA Ames Research Center
- The LINK Foundation
- The North Carolina Space Grant
- Digital Artforms

Table of Contents

Abstract	iii
List of Tables	xv
List of Figures	xvi
1 Overview	1
1.1 Introduction	1
1.1.1 Latency	2
1.2 Thesis Statement	3
1.3 Goals	4
1.4 Dissertation Overview	4
1.5 Summary of Research	5
1.5.1 A simulated zero-latency HMD system	5
1.5.2 A Mathematical Model Relating Scene-Motion Thresholds to Latency Thresholds	5
1.5.3 Validating the Model	6
1.6 Summary of Experiments	7
1.6.1 Experiment 1: Quasi-Sinusoidal Head Yaw and Scene-Motion Direction	8
1.6.2 Experiment 2: Single Head Yaw, Luminance, and Scene-Motion Direction	9
1.6.3 Experiment 3: Increasing Head Motion	11
1.6.4 Experiment 4: Validating the Model	13
1.6.5 Experiment 5: Two Case Studies	14
1.7 Summary of Results	18
1.8 Recommendations	18
1.8.1 Measuring Latency Thresholds	18
1.8.2 Latency Guidelines for Head-Mounted Displays	19
1.8.3 Implications for Redirected Walking	19

2	Background	20
2.1	Virtual Environments	20
2.1.1	Head-Mounted Displays versus World-Fixed Displays	21
2.1.2	IVEs Versus Augmented Environments	22
2.2	Visual Perception	22
2.2.1	The Visual System	23
2.2.1.1	The Photoreceptors: Rods and Cones	23
2.2.1.2	Two Visual Pathways	23
2.2.1.3	Central Versus Peripheral Vision	24
2.2.2	The Vestibular System	24
2.2.3	Eye Movements	24
2.2.3.1	Gaze-Shifting Eye Movements	25
2.2.3.2	Fixational Eye Movements	26
2.2.3.3	Gaze-Stabilizing Eye Movements	26
2.2.4	Afference and Efference	28
2.2.5	Intersensory Interactions	29
2.2.5.1	Visual and Vestibular Cues Complement Each Other	29
2.2.5.2	Cue Conflicts	29
2.2.5.3	Other Interactions	30
2.2.6	Top-Down Processing	30
2.2.7	Motion Perception	30
2.2.7.1	Visual Velocity	30
2.2.7.2	Object-Relative versus Subject-Relative Judgments	30
2.2.7.3	Depth Perception Affects Motion Perception	31
2.2.7.4	Motion Perception in Peripheral versus Central Vision	31
2.2.7.5	Motion Perception During Head Movement	32
2.2.7.6	Motion Illusions	33
2.2.8	Perceptual Constancies	35
2.2.8.1	Position Constancy	35
2.2.9	Adaptation	36
2.2.9.1	Sensory and Perceptual Adaptation	37
2.2.9.2	Light and Dark Adaptation	37
2.2.9.3	Position-Constancy Adaptation	37
2.2.9.4	Temporal Adaptation	38
2.3	Psychophysics	39

2.3.1	Psychometric Functions	40
2.3.2	Thresholds	40
2.3.2.1	Point of Subjective Equality	41
2.3.2.2	Just-Noticeable Difference	41
2.3.3	Determining Psychometric Functions and Thresholds	42
2.3.3.1	Method of Adjustment	42
2.3.3.2	Method of Constant Stimuli	42
2.3.3.3	Methods of Limits	42
2.3.3.4	Adaptive Staircases	43
2.3.4	Judgment Tasks	44
2.3.4.1	Yes/No Tasks	44
2.3.4.2	Same/Different Tasks	44
2.3.4.3	Identification Tasks	45
2.4	Latency	46
2.4.1	Human Factors	46
2.4.1.1	Degraded Visual Acuity	46
2.4.1.2	Performance	46
2.4.1.3	Simulator Sickness	47
2.4.1.4	Breaks in Presence	47
2.4.1.5	Negative Training Effects	47
2.4.2	Perception of Latency	48
2.4.2.1	Temporal Perception versus Motion Perception	48
2.4.2.2	Latency PSEs	48
2.4.2.3	Latency JNDs	49
2.4.2.4	Latency Discrimination Mechanisms	49
2.4.2.5	Scene Complexity	49
2.4.3	System Delay	49
2.4.3.1	Tracking Delay	50
2.4.3.2	Application Delay	51
2.4.3.3	Rendering Delay	51
2.4.3.4	Display Delay	51
2.4.3.5	Synchronization Delay	56
2.4.3.6	Timing Analysis	57
2.4.3.7	Measuring Delays	57
2.4.4	Delay Compensation	60

2.4.4.1	Prediction	63
2.4.4.2	Post-Rendering Techniques	63
2.4.4.3	Problems of Delay Compensation	64
2.5	Intentional Scene Motion	64
3	Analysis and Relationships of Latency, Head Motion, Scene Motion, and Perceptual Thresholds	65
3.1	Error due to Latency	65
3.1.1	Scene Displacement	65
3.1.2	Just-in-Time Scanlines	67
3.1.3	Scene Velocity	68
3.2	Real Head Motion	70
3.3	Threshold Assumptions	72
3.4	Relating Scene-Motion Thresholds to Latency Thresholds	73
3.4.1	Model Assumptions	73
3.4.2	The Model	73
4	Overview of Experiments	77
4.1	Materials and Methods	77
4.1.1	The System	79
4.1.2	Head Motion	80
4.1.3	The Scene	81
4.1.3.1	Scene Visibility	82
4.1.3.2	Scene Position	82
4.1.3.3	Scene Motion	83
4.1.3.4	Scene Luminance and Contrast	83
4.1.3.5	The Reference Scene	84
4.1.4	Psychophysics Methods	84
4.1.4.1	Adaptive Staircases Versus the Method of Constant Stimuli	84
4.1.4.2	3-Interval 2-Alternative Forced Choice	85
4.1.4.3	Yes/No Judgments	85
4.1.4.4	Confidence Ratings	86
4.2	Measures	86
4.2.1	Actual Head Motion	86
4.2.1.1	Head-Motion Quantiles	87

4.2.2	Scene Motion	87
4.2.3	Latency	87
4.3	Data Analysis	88
4.3.1	Confidence Ratings	88
4.3.2	Psychometric Functions	88
4.3.3	Thresholds	88
4.3.4	Tracker Jitter	89
4.3.5	Statistics	89
5	Experiment 1: Quasi-Sinusoidal Head Yaw and Scene-Motion Direction	90
5.1	Hypotheses	90
5.2	Experimental Design	91
5.2.1	The Stimulus	91
5.2.2	Controlled Variables	91
5.2.3	Independent Variables	92
5.2.4	Dependent Variables	94
5.3	Methods	94
5.4	Participants	94
5.5	Data Analysis	95
5.6	Summary of Results	97
5.7	Recommendations	98
5.7.1	Applications to Latency Perception	98
6	Experiment 2: Single Head Yaw, Luminance, and Scene-Motion Direction	100
6.1	Hypotheses	101
6.2	Experimental Design	101
6.2.1	The Stimulus	102
6.2.2	Controlled Variables	102
6.2.3	Independent Variables	102
6.2.4	Dependent Variable	103
6.3	Methods	103
6.4	Participants	104
6.5	Data Analysis	104
6.6	An Illusion of Scene Motion	107

6.7	Summary of Results	108
6.8	Recommendations	108
7	Experiment 3: Increasing Head Motion	109
7.1	Hypotheses	109
7.2	Experimental Design	110
7.2.1	The Stimulus	110
7.2.2	Controlled Variables	110
7.2.3	Independent Variables	110
7.2.4	Dependent Variables	111
7.3	Methods	112
7.4	Participants	112
7.5	Data Analysis	112
7.6	Summary of Results	116
7.7	Recommendations	118
8	Experiment 4: Validating the Latency-Thresholds Model	119
8.1	Hypotheses	120
8.2	Experimental Design	120
8.2.1	The Stimulus	121
8.2.2	Controlled Variables	121
8.2.3	Independent Variables	122
8.2.4	Dependent Variables	124
8.3	Methods	124
8.3.1	Head Yaw	125
8.3.2	Scene Motion	125
8.4	Participants	126
8.5	Data Analysis	126
8.5.1	Latency Thresholds	127
8.5.1.1	Average Latency Thresholds	127
8.5.1.2	One Subject's Latency Thresholds for Different Peak Head-Yaw Accelerations	128
8.5.1.3	Fitting the Model to the Measured Latency Thresholds	128
8.5.2	Scene-Motion Thresholds	133
8.5.2.1	Average Scene-Motion Thresholds	133

8.5.2.2	One Subject’s Scene-Motion Thresholds For Different Peak Head-Yaw Accelerations	133
8.5.2.3	Comparison of Scene-Motion Thresholds Between Conditions	133
8.5.3	Estimation of Latency Thresholds with Scene-Motion Thresholds	137
8.6	Summary of Results	138
8.7	Recommendations	139
9	Experiment 5: Two Case Studies	140
9.1	Experimental Design	140
9.1.1	Controlled Variables	141
9.1.2	Independent Variables	141
9.1.3	Dependent Variables	142
9.2	Methods	142
9.3	Participants	143
9.3.1	Subject IDjj	143
9.3.2	Subject ID338	143
9.4	Data Analysis	144
9.4.1	Scene-Motion Thresholds—Correlations	144
9.4.2	Minimum Latency Thresholds	145
9.4.3	Latency Thresholds—Correlations	146
9.4.4	Estimating Model Parameters from Scene-Motion Thresholds . .	146
9.4.5	Scene-Motion Thresholds—Differences Between Conditions . . .	148
9.4.6	Threshold Differences of Yes/No Judgments with Confidence Ratings and Yes/No Judgments Without Confidence Ratings . .	148
9.5	Summary of Results for these Cases	149
9.6	Recommendations	151
10	Discussion	166
10.1	Comparison to Previous Results	166
10.1.1	Scene-Motion Thresholds	166
10.1.2	Latency Thresholds	167
10.1.3	An Illusion of Scene Motion	168
10.1.3.1	Explaining the Illusion	169
10.1.4	Explanation of Differences	171
10.2	Subject Comments	173

10.3	Factors Affecting Thresholds	173
10.3.1	Other Factors for Latency Requirements	175
10.4	Limitations	175
10.4.1	Assumptions	175
10.4.2	Average Thresholds	176
10.4.3	Faster Head Motions	176
10.5	Studying Other Sources of Scene-Motion	176
10.6	Recommendations to IVE developers	176
10.6.1	Application to Reorientation Techniques	176
10.6.1.1	Maximum Scene-Motion as a Function of Physical Head Turns	177
10.6.1.2	Scene Motion Against Versus With Physical Turns	177
10.6.2	Measuring Latency Thresholds	178
10.6.3	Latency Requirements	178
10.6.4	Reducing Latency	179

REFERENCES **181**

List of Tables

2.1	Display rise and fall times for the Virtual Research V8 and Glasstron HMDs.	59
4.1	Factors and Findings for all five experiments.	78
5.1	Experiment 1 conditions.	93
6.1	Experiment 2 conditions.	104
7.1	Experiment 3 conditions.	112
8.1	Experiment 4 conditions.	123
8.2	Pearson correlations of latency thresholds to peak head-yaw accelerations using my model and a line.	128
8.3	95% confidence intervals of the differences between the two methods of determining τ and ψ	138
9.1	Experiment 5 conditions.	141
9.2	Pearson correlations of the linear relationship between scene-motion thresholds and peak head-yaw accelerations.	145
9.3	Pearson correlations of latency thresholds to peak head-yaw accelerations using my inverse model and a line.	146
9.4	Model Parameters τ and ψ	147
9.5	Difference of thresholds obtained from yes/no judgments with confidence ratings and yes/no judgments without confidence ratings.	150

List of Figures

1.1	Theoretical latency thresholds determined from my model.	6
1.2	Overview of the Experiments.	7
1.3	Experiment 1 scene-motion thresholds.	8
1.4	Experiment 2 box plots of scene-motion thresholds.	10
1.5	Experiment 3 scene-motion thresholds for a single subject.	11
1.6	Experiment 3 box plots of Pearson Correlations.	12
1.7	Experiment 4 thresholds for a single subject.	16
1.8	Experiment 5 thresholds for a single subject.	17
2.1	Components of an IVE.	21
2.2	An illustration of the vestibular system.	25
2.3	Efference-copy during rotation of the eye.	29
2.4	A typical psychometric function.	40
2.5	Method of Limits.	43
2.6	A one-up-one-down adaptive staircase.	44
2.7	End-to-end system delay.	50
2.8	Image tearing.	54
2.9	Rise and fall times for a typical LCD.	55
2.10	A timing diagram for a typical VE system.	58
2.11	The latency meter.	59
2.12	Oscilloscope output showing system delay.	60
2.13	Red peak-response time for the NASA Kaiser SR80 HMD.	61
2.14	Red, green, and blue peak-response times for the NASA Kaiser SR80 HMD.	62
3.1	Displacement error.	66
3.2	Peak scene velocity is a linear function of peak head acceleration and latency.	69
3.3	Scene velocity is head acceleration integrated over the time of latency. .	71
3.4	Measured head motion and computed scene motion due to 100 <i>ms</i> of latency.	72

3.5	Peak scene velocity as a function of peak head acceleration and latency, and a hypothetical scene-motion threshold line.	74
3.6	Theoretical latency thresholds.	76
4.1	A subject and scene as the subject yaws her head.	80
4.2	The modified V8 HMD.	81
5.1	Experiment 1 head yaw and scene visibility.	92
5.2	Judgments from a single subject for the With Condition and the resulting psychometric function.	95
5.3	Psychometric functions for all three conditions for a single subject. . .	96
5.4	Thresholds for all nine subjects for all three visibility conditions.	97
5.5	Ratios of scene-velocity thresholds to peak head velocities.	98
6.1	Experiment 2 confidence ratings and the resulting psychometric function for a single subject.	105
6.2	Box plots of all scene-motion thresholds.	106
6.3	Pilot data shows an illusion that the scene appeared to move when the scene was stable and the head was yawing.	107
7.1	Experiment 3 confidence ratings from a single subject for the End Condition and sextile five of the Peak Head-Yaw Accelerations.	113
7.2	Scene-motion thresholds for a single subject for the End Condition where head motion is defined by Peak Head-Yaw Acceleration.	114
7.3	Scene-motion thresholds for a single subject for all four head-yaw phase conditions where head motion is defined by Peak Head-Yaw Acceleration.	115
7.4	Box plots of Pearson Correlations for the eight subjects.	116
8.1	Experiment 4 example scene velocities for the three experimental conditions.	122
8.2	21 confidence ratings versus latency, and the resulting psychometric function, for a range of a subject's peak head-yaw acceleration.	127
8.3	75% thresholds for a single subject.	129
8.4	50% Thresholds for a single subject.	130
8.5	Difference thresholds for a single subject.	131
8.6	Latency thresholds for all six subjects.	132

8.7	Individuals' differences of Latency Condition scene-motion thresholds and Constant Condition scene-motion thresholds.	134
8.8	Individuals' differences of Gaussian Condition scene-motion thresholds and Constant Condition scene-motion thresholds.	135
8.9	Individuals' differences of Latency Condition scene-motion thresholds and Gaussian Condition scene-motion thresholds.	136
9.1	Judgments of various latencies for a range of Subject IDjj's peak head-yaw acceleration and the corresponding psychometric function.	152
9.2	Judgments of various latencies for a range of Subject ID338's peak head-yaw acceleration and the corresponding psychometric function.	153
9.3	75% thresholds for Subject IDjj, computed from yes/no judgments <i>with</i> confidence ratings.	154
9.4	50% thresholds for Subject IDjj, computed from yes/no judgments <i>with</i> confidence ratings.	155
9.5	Difference thresholds for Subject IDjj, computed from yes/no judgments <i>with</i> confidence ratings.	156
9.6	75% thresholds for Subject IDjj, computed from yes/no judgments <i>without</i> confidence ratings.	157
9.7	50% thresholds for Subject IDjj, computed from yes/no judgments <i>without</i> confidence ratings.	158
9.8	Difference thresholds for Subject IDjj, computed from yes/no judgments <i>without</i> confidence ratings.	159
9.9	75% thresholds for Subject ID338, computed from yes/no judgments <i>with</i> confidence ratings.	160
9.10	50% thresholds for Subject ID338, computed from yes/no judgments <i>with</i> confidence ratings.	161
9.11	Difference thresholds for Subject ID338, computed from yes/no judgments <i>with</i> confidence ratings.	162
9.12	75% thresholds for Subject ID338, computed from yes/no judgments <i>without</i> confidence ratings.	163
9.13	50% thresholds for Subject ID338, computed from yes/no judgments <i>without</i> confidence ratings.	164
9.14	Difference thresholds for Subject ID338, computed from yes/no judgments <i>without</i> confidence ratings.	165

Chapter 1

Overview

1.1 Introduction

The real world remains stationary as one rotates one's head, and one perceives the world to be stationary even though the world's image moves on the retina. A computer-generated *immersive virtual environment* (IVE) provides stimuli to a user's senses so that she feels present in a virtual world. A *scene* is a visual representation of the IVE, from a single point of view. *Scene motion* is improper visual motion of the scene that would not occur if the IVE behaved as the real world. However, an IVE scene may not be stationary, as the real world is, due to

- **unintentional scene motion** caused by shortcomings of technology, such as latency, or imprecise calibration (incorrect field of view, incorrect world-to-eye transformations, etc.), and
- **intentional scene motion** injected into the system in order to make the virtual world behave differently than the real world (e.g., redirected walking (Razzaque et al., 2001), a technique that allows users to walk in IVEs larger than the tracked lab space).

Whereas error and scene motion in IVEs are well defined mathematically (Adelstein et al., 2005; Holloway, 1997), users' perception of unnatural scene motion when the head is moving is not as well understood. Investigators do know that noticeable visual instability can degrade an IVE experience by causing simulator sickness (Draper, 1998), reducing task performance (So and Griffin, 1995), lowering visual acuity (Allison et al., 2001), and decreasing the sense of presence (Meehan et al., 2003). However, IVE

researchers know little about the amounts of scene motion required for head turners to notice. I quantify such perceptual tolerances to define system requirements for IVE developers and explore the relationship among visual perception, scene motion, head motion, and latency.

1.1.1 Latency

Latency is the effective time for a system to respond to a user’s actions. Latency in head-mounted display (HMD) systems is a primary factor that detracts from the sense of presence (Meehan et al., 2003). For low latencies, users do not perceive latency directly, but rather the consequences of latency—a static virtual scene becomes unstable in space when users move their heads. Reducing latency provides more believable virtual worlds and is important for natural and precise interaction.

Often engineers build HMD systems with a goal of “low latency” without specifically defining what that “low latency” is. There is little consensus among researchers what latency requirements should be. I propose that HMD systems should ideally have latency requirements such that users are not able to perceive scene motion, with some probability, resulting from that latency. Such requirements will vary greatly depending on tasks and conditions. A relaxing IVE with slow head movements (e.g., intended to induce relaxation) will have quite different requirements than a fighter-pilot training system. I want to know how much scene motion and latency can occur in an IVE without users noticing (i.e., perceptual thresholds), and how this differs for different head motions.

I define a perceptual *threshold* to be the change in intensity of a stimulus that is required for a subject to detect that change with some specified probability or confidence. Scene-motion thresholds are measured in degrees of visual angle per second and latency thresholds are measured in seconds.

Two primary factors contribute to latency thresholds:

- **Scene motion due to latency and head motion** is mathematically defined (Adelstein et al., 2005; Holloway, 1997). With zero latency or no head motion, the virtual scene in an HMD is spatially fixed in the world. As latency is added and the user moves her head, scene motion increases.
- **Scene-motion thresholds during head motion** are subjectively measured through psychophysics experiments. As one moves the head faster, the ability to detect scene motion decreases (as will be shown).

Little is known about the relationship between the two factors above. In this work, I explore the relationships between head motion, latency, scene motion, latency thresholds, and scene-motion thresholds. By understanding and quantifying perception of scene motion and latency, I and others can determine design guidelines for future IVE systems.

The typical way to measure detection thresholds is to present stimuli to subjects at different magnitudes and have them state if they detect the stimulus or not. The stimuli magnitudes range from zero to obviously apparent. Statistics are then performed and thresholds are extracted from the data. A system with zero inherent latency would enable scientists to measure latency thresholds in a straightforward way using standard psychophysics techniques. However, it is not possible to build an HMD system with zero latency.

I build a simulated-by-projector HMD system, where latency is effectively zero, i.e., the scene remains perfectly stable in space independent of system delay and head motion. I then move the scene, as a subject yaws her head, in the same way as would occur in an HMD system with latency. This enables me to measure latency thresholds. I also move the scene in other ways, in order to measure scene-motion thresholds for different types of scene motion.

Measuring scene-motion thresholds and latency thresholds enables me to

- develop and validate an analytical model describing latency thresholds as an inverse function of peak head-yaw acceleration, and
- relate scene-motion thresholds to latency thresholds.

I measure latency thresholds and scene-motion thresholds for various types of head yaw under specific conditions. These thresholds could be different for different conditions (e.g., for a larger field of view). I do not map the whole multidimensional space for these thresholds, but the methods described enable engineers to estimate latency requirements under specific conditions with a basic off-the-shelf projector system.

1.2 Thesis Statement

Human latency perception thresholds for head-mounted displays can be measured using an off-the-shelf projector system requiring neither an HMD nor a system with zero

system delay. Latency thresholds Δt are inversely related to peak head-yaw acceleration ϕ'' such that $\Delta t = \tau + \psi \left(\frac{1}{\phi''} \right)$ where τ is an offset value, in seconds, and ψ is a scale value, in degrees per second. τ and ψ can be estimated from scene-motion thresholds. Directly measured latency thresholds correlate with this inverse function better than with a linear function.

1.3 Goals

I report the following:

- An analysis of scene motion in an HMD due to latency and real head motion.
- A simulated zero-latency HMD system enabling me and others to measure latency thresholds.
- Methods for determining scene-motion thresholds and latency thresholds for HMDs.
- Understanding of how scene-motion thresholds and latency thresholds change for different conditions.
- An analytical model of latency thresholds as a function of peak head-yaw acceleration.
- A mathematical relationship between scene-motion thresholds and latency thresholds.
- Discussion of perceptual tolerances of scene motion and latency that are relevant to HMDs. Establishment of latency guidelines and specifications for the design, implementation, and effective deployment of IVE systems.
- Suggestions for improving re-orientation and redirected walking techniques (Razzaque et al., 2001) for HMDs.

1.4 Dissertation Overview

This dissertation is written for a general Computer Science audience. I expect the reader to have a basic background in statistics and virtual environments.

Chapter 1 provides a synopsis of the dissertation.

Chapter 2 discusses relevant background and literature in the areas of IVEs, visual perception, psychophysics, and latency for HMDs. The areas covered are specific to this work, not general background material. This chapter may be read independently or even skipped.

Chapter 3 analyzes the relationship among latency, head motion, scene motion, and perceptual thresholds. From this analysis, I develop a mathematical model relating scene-motion thresholds to latency thresholds.

Chapter 4 provides an overview of the experiments. Chapters 5-9 describe the experiments in detail, with a chapter devoted to each experiment. These chapters provide not only the results of the experiments, but also examples of experimental designs for measuring perceptual thresholds for IVEs.

Chapter 10 compares my results to previous results, discusses factors affecting thresholds and latency requirements, describes limitations and extensions of my model and methods, and provides recommendations to IVE developers.

1.5 Summary of Research

1.5.1 A simulated zero-latency HMD system

I describe a simulated HMD system that I built. Although the system has 7.2 *ms* of system delay, the system has effectively zero latency—no scene motion occurs due to system delay. Scene motion can then be precisely controlled by injecting arbitrary scene motion into the system, including scene motion like that which latency would cause in an HMD. This system allows me to conduct psychophysics experiments with latency ranging from zero up.

1.5.2 A Mathematical Model Relating Scene-Motion Thresholds to Latency Thresholds

I develop a mathematical model that relates head motion, scene-motion thresholds, and latency thresholds for HMDs. The model is

$$\Delta t = \tau + \psi \left(\frac{1}{\phi^n} \right) \tag{1.1}$$

where Δt is the latency threshold (in seconds) and ϕ'' is peak head-yaw acceleration (in degrees per second squared). τ (in seconds) and ψ (in degrees per second) are parameters that can be estimated from scene-motion thresholds.

Figure 1.1 shows latency thresholds versus peak head-yaw acceleration. The parameters τ and ψ simply shift and scale the curve.

The model shows that latency thresholds decrease as head-yaw acceleration increases: the faster a user moves her head, the more she will notice scene motion due to latency. When there is no head motion, latency thresholds go to infinity; users do not notice scene motion due to latency when not moving the head because no scene motion occurs.

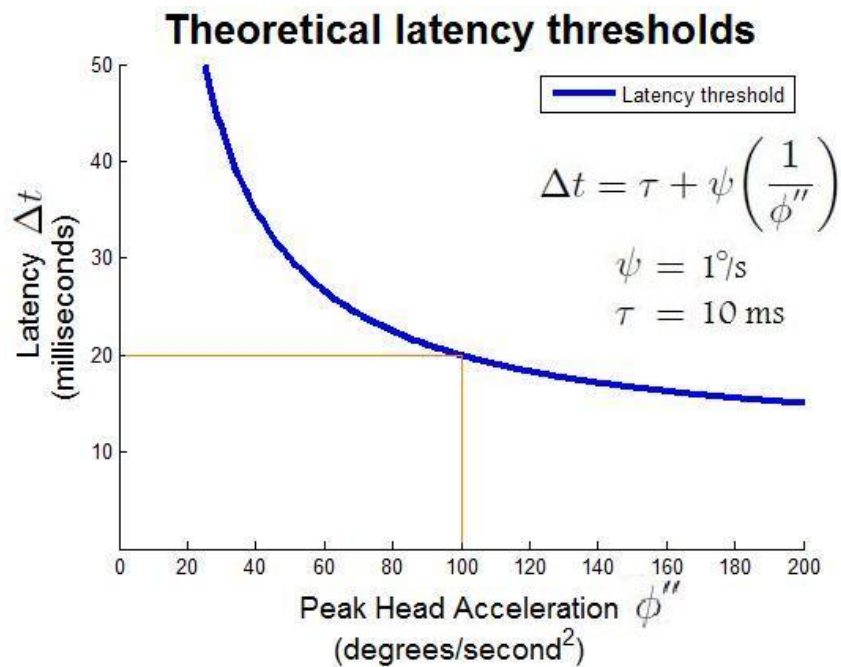


Figure 1.1: Theoretical latency thresholds determined from my model with parameters determined from hypothetical scene-motion thresholds. I hypothesize that measured latency thresholds match this form, and demonstrate, in Experiments 4 and 5, that they do.

1.5.3 Validating the Model

I measure scene-motion thresholds for several subjects under different conditions to verify assumptions of the model. In order to validate the model, I compute correlations of how well measured latency thresholds fit the form of the model. I then compare the correlation values for my model and a linear model. Furthermore, I measure how well scene-motion thresholds estimate parameters of the model.

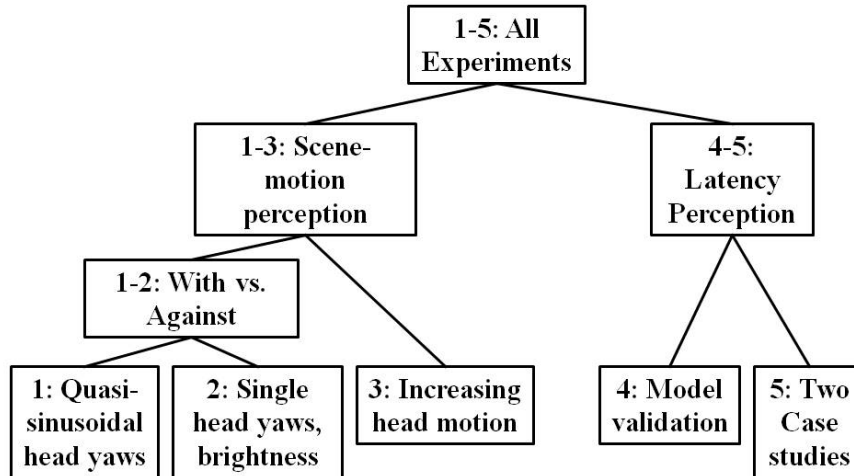


Figure 1.2: Overview of the Experiments. Experiments 1-3 measure scene-motion thresholds. Experiments 4 and 5 measure both scene-motion thresholds and latency thresholds.

1.6 Summary of Experiments

I conduct five experiments that measure perceptual thresholds. The stimulus to be detected for all experiments is scene motion (whether induced by latency or not). Figure 1.2 provides an overview of these experiments. The goals of Experiments 1-3 are to help me understand perception of scene motion in order to develop the mathematical model relating scene-motion thresholds and latency thresholds, to design experiments for determining latency thresholds, and to test assumptions of the latency thresholds experiments (Experiments 4 and 5).

Experiments 1-3 measure and compare scene-motion thresholds under different conditions. Experiments 1 and 2 test if scene-motion thresholds depend upon scene-motion direction relative to head-motion direction. Experiment 3 tests if scene-motion thresholds increase as head motion increases. Experiments 4 and 5 measure both scene-motion thresholds and latency thresholds, and validate my mathematical model relating latency thresholds to scene-motion thresholds.

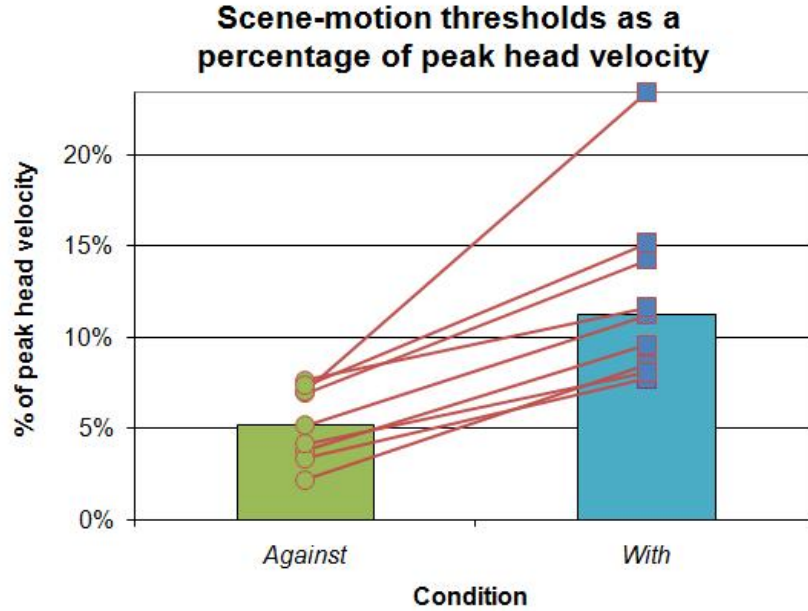


Figure 1.3: Scene-motion thresholds from Experiment 1 for nine subjects when the scene moves *against* head yaw and *with* head yaw. Since I ask different subjects to yaw their heads by different amounts, thresholds are shown as the ratio of scene-motion thresholds (in degrees per second) to the intended peak head-yaw velocities (also in degrees per second). Lines connect results from individual subjects. On average, thresholds for the With Condition are twice that of the Against Condition.

1.6.1 Experiment 1: Quasi-Sinusoidal Head Yaw and Scene-Motion Direction

For Experiment 1, subjects yaw their heads in a quasi-sinusoidal manner. The system presents moving and non-moving scenes for some head-yaw phase (i.e., the scene is visible for only portions of the head yaw and blanked-out otherwise) and subjects select which presentation they believe contains scene motion.

I measure and compare scene-motion thresholds for the following conditions:

- The reversal phase of quasi-sinusoidal head yaw, when the direction of head motion changes (**Reversal Condition**).
- The center phase of quasi-sinusoidal head yaw, when the head moves with approximately constant velocity. Two scene-motion directions were measured for the center phase:
 - The scene moves with the direction of head yaw (**With Condition**).
 - The scene moves against the direction of head yaw (**Against Condition**).

Differences between the With Condition thresholds and the Against Conditions thresholds were statistically significant. On average, the subjects' scene-motion thresholds are twice as high for the With Condition than for the Against Condition. Figure 1.3 shows scene-motion thresholds of the Against and With Conditions for the nine subjects. Since I ask different subjects to yaw their heads by different amounts (in order to generalize results), the thresholds are shown as the ratio of scene-motion velocity thresholds to the intended peak head-yaw velocities, both measured in degrees per second.

Scene motion-thresholds are statistically significantly lower for the Reversal Condition than for the With Condition. Although scene-motion thresholds are greater for the Reversal Condition than for the Against Condition, this finding is not statistically significant.

These results suggest, if developers do not want users to perceive scene motion, the direction of scene motion relative to head yaw direction and the phase of the user's head yaw should be taken into account when inserting visual rotation into IVEs, In particular, there is much more latitude for artificially and imperceptibly rotating a scene, as in Razzaque's redirecting walking technique (Razzaque et al., 2001), in the same direction of head yaw than against the direction of head yaw.

The results also suggest HMD users are less likely to notice latency when beginning a head yaw (when the head accelerates causing the scene to move with the direction of head yaw) than when slowing down a head yaw (when the head decelerates causing the scene to move against the direction of head yaw) or when changing head yaw direction (when head acceleration peaks causing scene velocity to peak).

The results of Experiment 1 are also reported in Jerald et al. (2008).

1.6.2 Experiment 2: Single Head Yaw, Luminance, and Scene-Motion Direction

Experiment 2 further investigates the results from Experiment 1 that scene-motion thresholds are greater when the scene moves with the head than when the scene moves against the head. I test if this result generalizes to other conditions, and is not simply an artifact of the specific conditions of Experiment 1. Specifically, I want to know if this finding holds for conditions similar to those used in Experiments 4 and 5, as the design of those experiments assumes this to be true.

I confirm that scene-motion thresholds are greater when the scene moves with

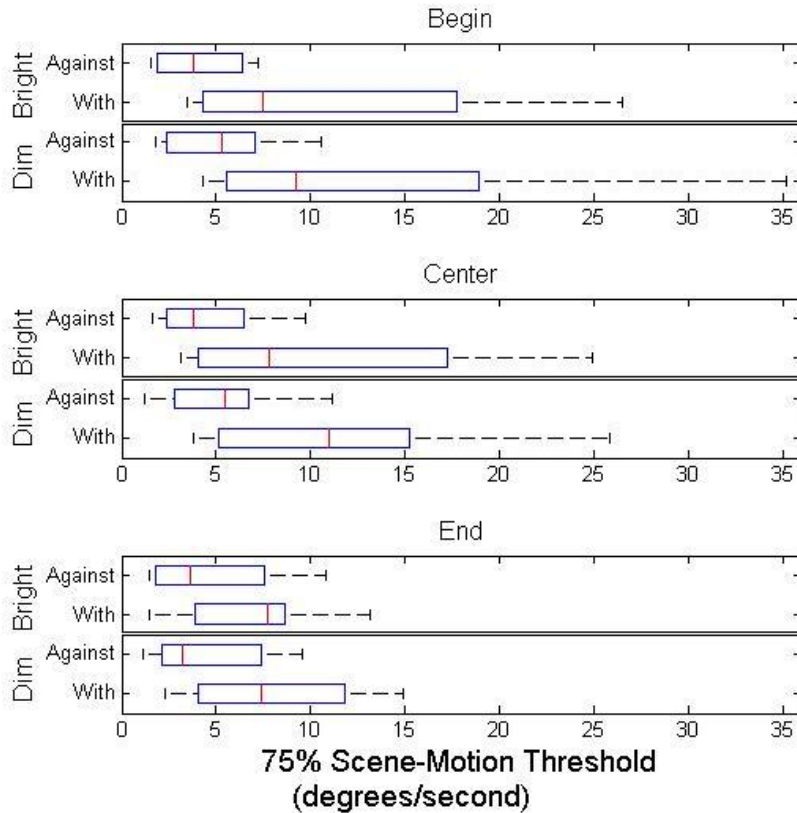


Figure 1.4: Experiment 2 box plots of scene-motion thresholds for the nine tested subjects.

the head (With Condition) than when the scene moves against the head (Against Condition) for each of the following conditions:

- Three phases of single head yaw (**Start, Center, and End Conditions**).
- Two scene-luminance (contrast) conditions (**Dim and Bright Conditions**).

Figure 1.4 shows box plots for all subjects' thresholds for all conditions. With Condition thresholds are greater than Against Condition thresholds for each of the three head-yaw phase conditions and for each of the two scene-luminance conditions (all statistically significant). No statistically significant differences occur in thresholds that are attributable to luminance (contrast).

The median ratio of the With Condition thresholds to the Against Condition thresholds is 1.89. This average ratio confirms the results of Experiment 1 that scene-motion thresholds when the scene moves with head yaw are approximately twice as high as when the scene moves against head yaw. Thus, the results of Experiment 1 are not limited to the specific conditions of that study.

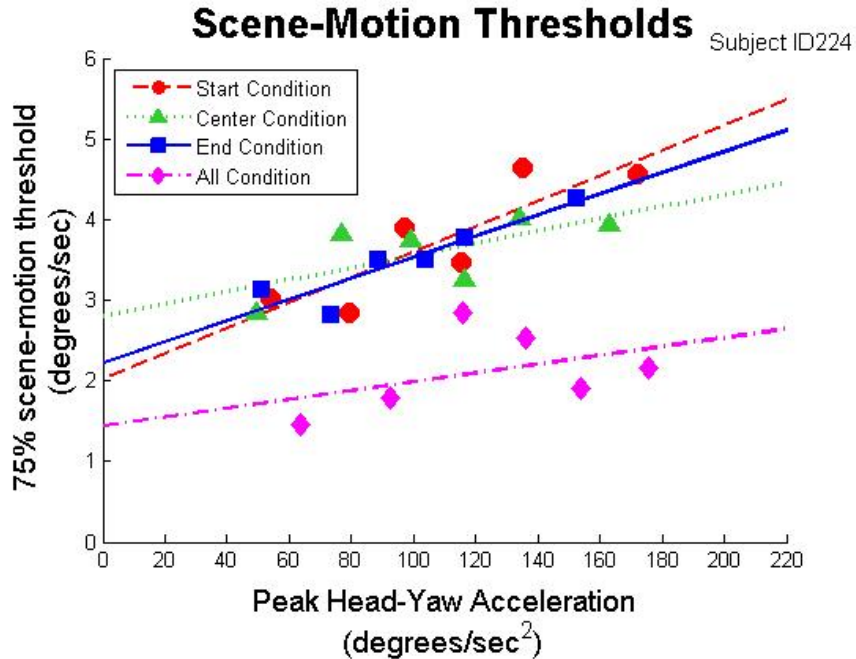


Figure 1.5: Experiment 3 scene-motion thresholds for a single subject where head motion is measured by the Peak Head-Yaw Acceleration measure. Note the lower thresholds for the All condition and the similarities of the other conditions for this particular subject; such patterns are not consistent across all subjects.

Part of Experiment 2 is also reported in Jerald and Steinicke (2009).

1.6.3 Experiment 3: Increasing Head Motion

Experiment 3 tests if scene-motion thresholds increase as head motion increases. Over several hundred trials, as subjects yaw their heads a single time left-to-right or right-to-left, I measure scene-motion thresholds as functions of three head-motion measures:

- Head-Yaw Range, given constant time (i.e., average head-yaw velocity),
- Peak Head-Yaw Velocity, and
- Peak Head-Yaw Acceleration.

Scenes always move against the direction of head yaw, since Experiment 1 shows thresholds to be lower for this direction. Thresholds are determined for four head-yaw phase visibility conditions; i.e., the scene is visible for only a part of the head yaw and blanked-out otherwise. The head-yaw phase conditions are:

- the start of the head yaw (**Start Condition**),

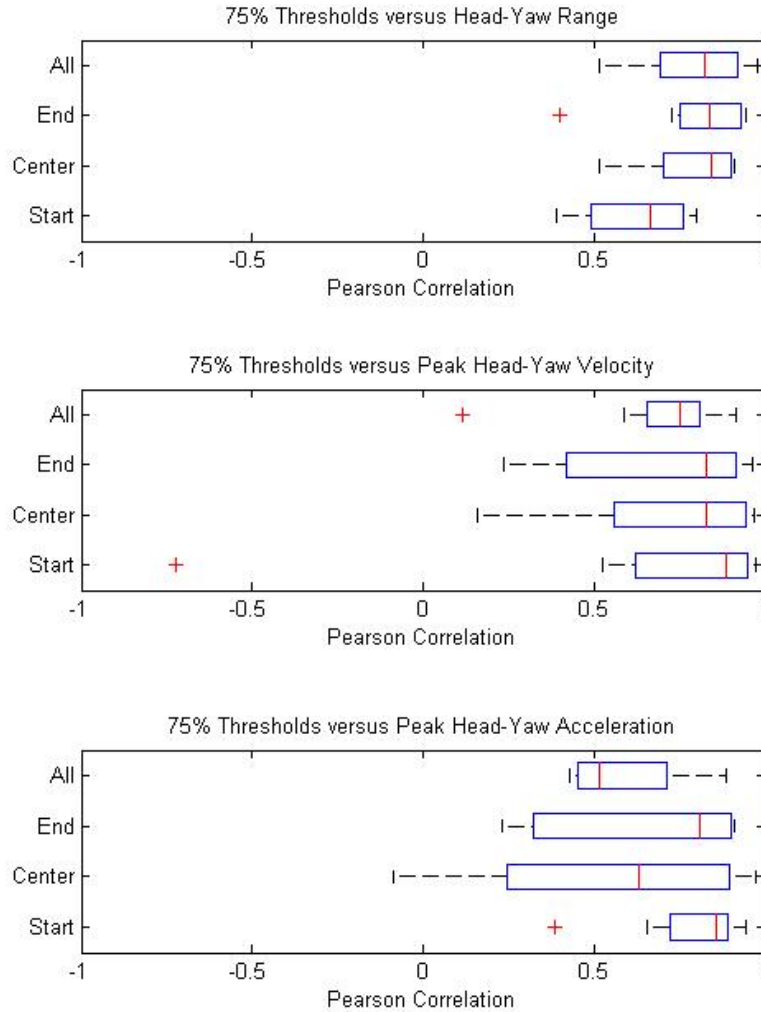


Figure 1.6: Experiment 3 box plots of Pearson Correlations for the eight tested subjects.

- the center of the head yaw (**Center Condition**),
- the end of the head yaw (**End Condition**), and
- all of the head yaw (**All Condition**).

Figure 1.5 shows that scene-motion threshold increases as peak head-yaw accelerations increases for a single subject. Although thresholds for the All Condition are lower than the other head-phase conditions for this head-motion measure and subject, this pattern is not statistically significant across subjects.

Pearson correlations between head motions and scene-motion thresholds are computed for each of the 12 combinations (3 head motion measures \times 4 head-yaw phase conditions) for each of the 8 subjects (a total of 96 correlations). Figure 1.6 shows

box plots of the 8 subjects for all 12 conditions. Although all Pearson correlations are not greater than zero, sign tests across all 8 subjects yield correlations statistically significantly greater than zero for each of the 12 conditions. Thus, I conclude that scene-motion thresholds increase as head motion increases.

Experiment 3 is also reported in Jerald et al. (2009).

1.6.4 Experiment 4: Validating the Model

Experiment 4 tests my model (Equation 1.1) relating scene-motion thresholds to latency thresholds. The goals of this experiment are to:

1. Measure scene-motion thresholds and latency thresholds over a range of peak head-yaw accelerations.
2. Validate my model relating scene-motion thresholds to latency thresholds.
3. Compare thresholds resulting from three types of scene motion that occur during head yaw. The scene moves against the direction of the head yaw and is controlled by one of three scene-motion types:
 - **Constant Scene Velocity (Constant Condition)**—The scene is already moving when it appears and it moves at a constant velocity until it disappears (the same as in Experiments 1-3).
 - **Gaussian-Modeled Scene Motion (Gaussian Condition)**—The scene moves with a Gaussian-velocity profile that models scene motion that typically occurs in an HMD due to head motion and latency.
 - **Latency-Induced Scene Motion (Latency Condition)**—The scene motion is manipulated in real-time based on the subject’s actual head motion and latency. Since latency controls scene motion in this condition, both latency thresholds, in seconds, and scene-motion thresholds, in degrees per second, are determined from the same trials.

The Constant and Gaussian Conditions are independent of head motion, whereas the Latency condition depends upon head motion.

As done in Experiment 3, I measure scene-motion thresholds for ranges of peak head-yaw accelerations for the three conditions. Figure 1.7 (top) shows these results for a single subject.

Across all subjects, the Constant Condition thresholds are statistically significantly less than both the Gaussian Condition thresholds and the Latency Condition thresholds. No statistically significant difference are found between the Gaussian Condition thresholds and the Latency Condition thresholds.

I fit lines to the scene-motion thresholds and transform these lines into latency-thresholds curves using my model. These resulting latency-threshold curves (the three thinner lines) are shown in Figure 1.7 (bottom).

I measure latency thresholds for ranges of peak head-yaw accelerations. For a single subject, Figure 1.7 (bottom) shows measured latency thresholds and the best fit curve (the thickest line) that fits those latency thresholds using my model. This subject had the lowest latency threshold of 3.2 ms

The directly-measured latency thresholds correlate with the form of my model with $\rho > 0.5$ for each subject. Across all subjects, these latency thresholds fit the form of my inverse model better than a linear model.

The results of Experiment 4 are also reported in Jerald and Whitton (2009).

1.6.5 Experiment 5: Two Case Studies

Data from Experiment 4 is quite noisy. For Experiment 5, I choose to collect an extensive amount of data for a smaller number of subjects under conditions similar to those conditions of Experiment 4. This enables me to more precisely measure thresholds from the subjects with the lowest thresholds. The two subjects are myself and SubjectID 338 from Experiment 4.

Experiment 5 determines

- whether scene-motion threshold as a function of peak head-yaw acceleration is approximately linear,
- how well the Gaussian Condition can be used to estimate latency thresholds,
- more precise latency thresholds for the most sensitive subjects, and
- how well thresholds obtained using the experiment task that is specific to my method match thresholds obtained from a standard psychophysics task.

Figure 1.8 shows data collected in Experiment 5 from a single subject (SubjectID 338, of Experiment 4).

Results for both subjects are consistent with Experiment 4. Scene-motion thresholds increase linearly with peak head-yaw acceleration (median $\rho = 0.87$). The minimum measured latency threshold is 4.3 *ms*. The directly-measured latency thresholds correlate with the form of my model for each subject (median $\rho = 0.97$). All directly-measured latency thresholds correlate with the form of my inverse model better than with a linear model.

As can be seen in Figure 1.8, The Gaussian Condition scene-motion thresholds do not estimate the directly-measured latency thresholds as well as the Latency Condition scene-motion thresholds.

Due to the large amount of data collected for Experiment 5, I am able to compare thresholds obtained by my psychophysics method with a standard psychophysics method. In some cases, my method over estimates thresholds obtained from the standard psychophysics method. In such cases, I suggest to first use my method to find conditions that scene motion can most easily be detected and to find subjects who can most easily detect scene motion, and then to use standard psychophysics methods under those conditions and with those subjects to determine conservative thresholds.

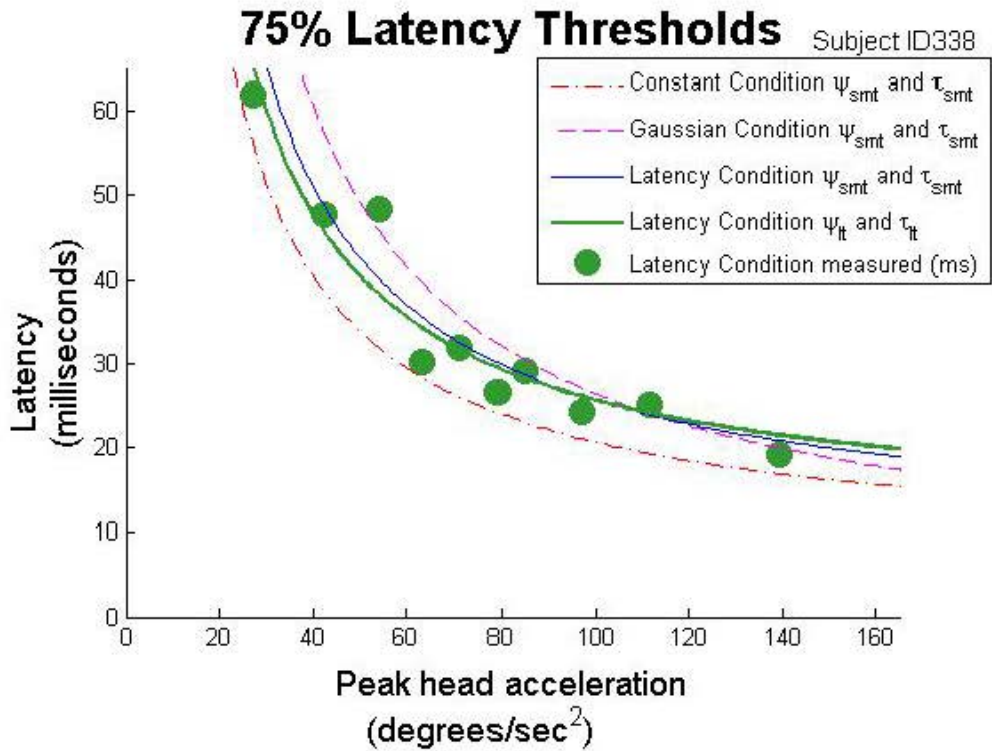
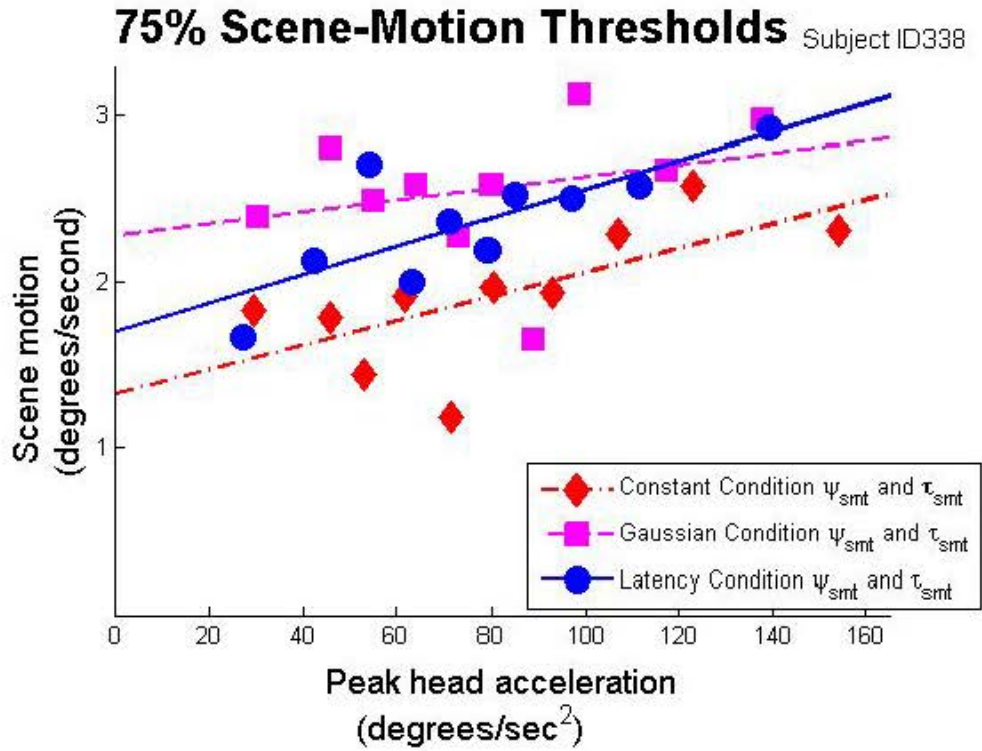


Figure 1.7: Experiment 4 scene-motion thresholds (top) and latency thresholds (bottom) for a single subject.

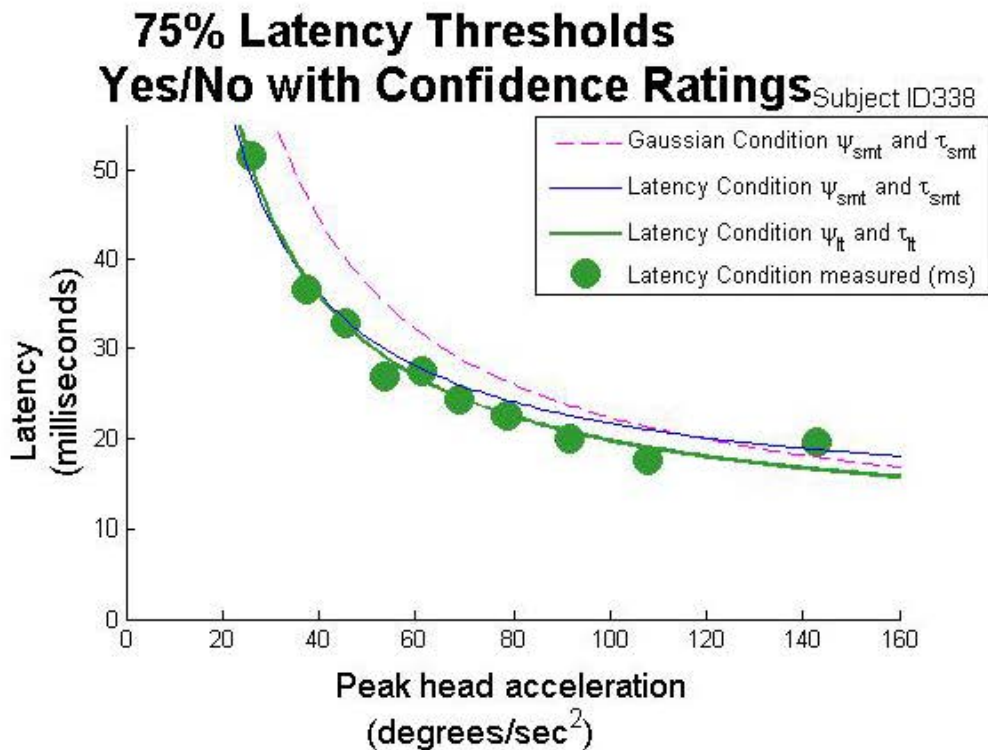
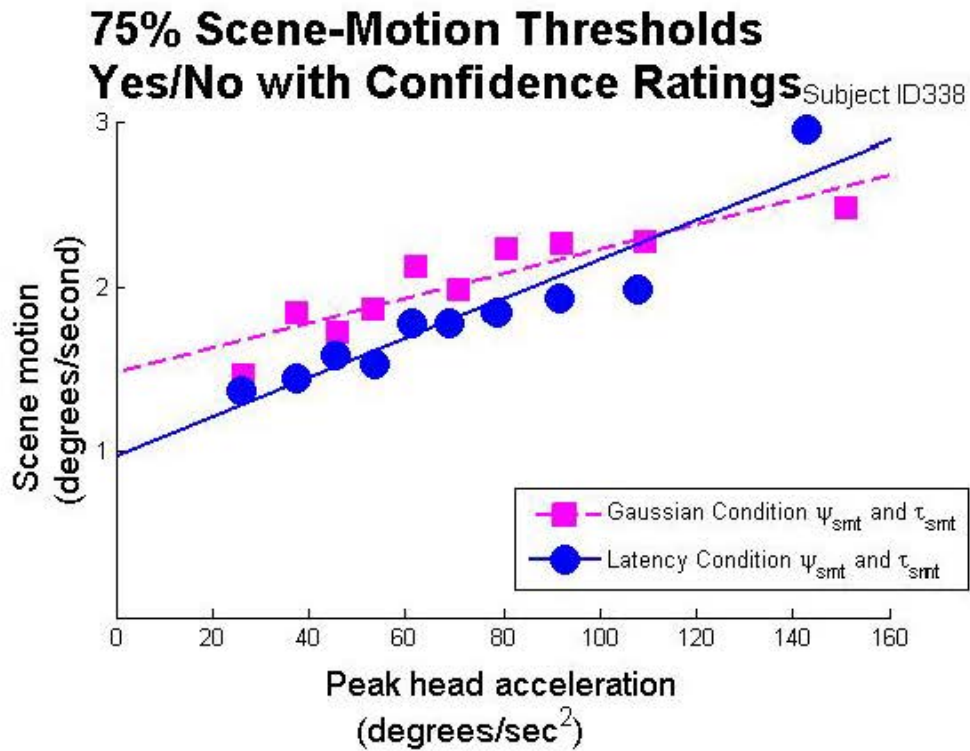


Figure 1.8: Experiment 5 scene-motion thresholds (top) and latency thresholds (bottom) 5 for a single subject.

1.7 Summary of Results

The primary experimental results are:

- **Experiment 1**—For quasi-sinusoidal head-yaw, scene-motion thresholds are greater when the scene moves with the direction of head yaw (as occurs in a lagging HMD when the head accelerates) than when the scene moves against the direction of head yaw (as occurs in a lagging HMD when the head decelerates).
- **Experiment 2**—The results of Experiment 1 hold separately for all phases—start, center, and end—of single side-to-side head yaw and for two scene luminances (contrasts) that differ by two orders of magnitude.
- **Experiment 3**—Scene-motion thresholds increase with head yaw range given constant time (i.e., average velocity), with peak head-yaw velocity, and with peak head-yaw acceleration.
- **Experiments 4**—Directly-measured latency thresholds as a function of peak head-yaw acceleration correlate with the form of my inverse model better than with a linear model. For the most sensitive subjects, the minimum latency-difference threshold is 3.2 *ms*.
- **Experiments 5**—Results are consistent with Experiments 3 and 4, and have stronger correlations.

1.8 Recommendations

1.8.1 Measuring Latency Thresholds

Measuring scene-motion thresholds helped me understand how perception of scene motion relates to perception of latency in HMDs and helped me create the mathematical model. However, measuring scene-motion thresholds are not necessary to measure latency thresholds. I suggest investigators determine the parameters of my latency-thresholds model by measuring latency thresholds directly using a simulated-by-projector HMD system—it is just as easy to measure latency thresholds as it is to measure scene-motion thresholds.

1.8.2 Latency Guidelines for Head-Mounted Displays

Latency thresholds decrease as peak head-yaw acceleration increases. The minimum measured latency threshold is 3.2 *ms*. This suggests that latency requirements should be in the 3 *ms* range, although the exact number may vary depending on the specific conditions of the IVE. This is a challenging number for system designers, but not impossible.

1.8.3 Implications for Redirected Walking

The results of Experiments 1 and 2 suggest that twice as much scene motion can occur without users noticing when the scene moves with the direction of a head turn compared to when the scene moves against the direction of a head turn.

Experiment 3 suggests, if the purpose is for subjects not to notice scene motion, that the maximum amount of allowed scene motion should go up with head velocity and with head acceleration. Experiment 5 provides evidence that for head acceleration, this function is linear.

Chapter 2

Background

2.1 Virtual Environments

Sherman and Craig (2003), define *virtual reality*, also known as a *virtual environment*:

A medium composed of interactive computer simulations that sense the participant’s position and actions and replace or augment the feedback to one or more senses, giving the feeling of being mentally immersed or present in the simulation.

I define an *immersive virtual environment* (IVE) to be defined as above with the additional requirement that only computer-generated visual cues are visible—no real-world cues can be seen.

Figure 2.1 shows a user and an IVE system. I divide IVE systems into their primary components of tracking, application, rendering, and display. *Tracking* calculates the viewpoint of the user. The *application* includes non-rendering aspects of the virtual world including updating dynamic geometry, user interaction other than viewpoint manipulation, physics simulation, etc. *Rendering* is the transformation of a geometric description to pixels. *Display* is the process of physically displaying the computed pixels from a video signal to the user. These components are well understood and the details are not important for this work other than delay-related issues associated with them (Section 2.4.3). This work focuses more on what is less understood—how users perceives IVEs, specifically scene motion and latency.

I define an *object* to be a 3D description of an entity in the IVE, the *virtual world* to be all objects, a *visual* to be a geometric object rendered to the display (relative to that display space) in display space, and the *scene* to be all visuals in real-world space.

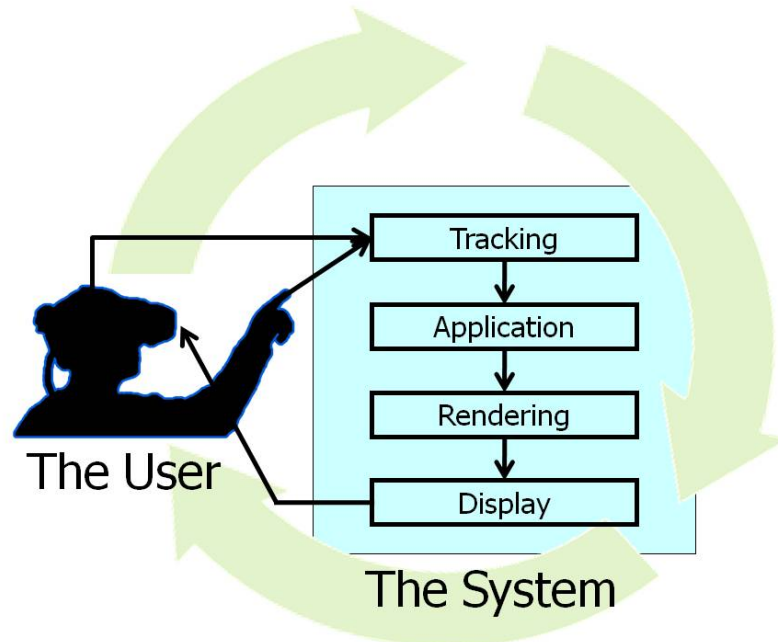


Figure 2.1: An IVE consists of input from a user, four primary components, and output to the user.

2.1.1 Head-Mounted Displays versus World-Fixed Displays

Head-mounted displays (HMDs) are displays semi-rigidly attached to the users head. As the user rotates the head to the right, the visuals on the display should rotate to the left so that the scene appears to be stable in space, as in the real world.

For a world-fixed display (e.g., a desktop monitor or a CAVETM), visuals do not rotate with the head. If a 2D object is collocated with the physical display surface, the rendered visual is independent of viewing position. For objects not collocated with the physical display surface, visuals are dependent on viewing position. As the viewpoint translates, the motion of the visuals on the display surface increases with the distance of the object from the display surface. For objects between the viewpoint and the display surface, the visuals move opposite to viewpoint translation. For objects beyond the display surface, the visuals move in the same direction as the viewpoint translation.

Scene motion can not be precisely controlled in an HMD due to latency and other issues (e.g., precisely calibrating HMDs is difficult). Latency is much less of a problem with world-fixed displays than with HMDs. If objects are collocated with a world-fixed display surface then no scene motion results from latency.

This work focuses upon HMDs. However, world-fixed displays are used as a tool to simulate HMDs. I collocate 2D scenes with the physical display surface such that

latency and/or incorrect viewpoint (or translating viewpoint) does not cause scene motion. The system then injects scene motion in order to conduct experiments under controlled conditions.

2.1.2 IVEs Versus Augmented Environments

An IVE, by my definition, blocks out all real-world cues and provides only computer-generated cues to the user. An *augmented environment* augments reality by mixing computer-generated cues with real world cues. Users are much more sensitive to error for augmented environments, as visual cues can be directly compared with the real world (Azuma, 1997).

This dissertation focuses upon IVEs, specifically IVEs implemented by non-see-through HMDs. The methods resulting from this work can also be applied to specific cases of augmented environments where real-world visual cues are delayed by the same amount as computer-generated visual cues. For example, real-world visual cues in a video-see-through display can be delayed so as to synchronize with computer-generated visuals. The methods will not however apply to optical-see-through HMDs.

World-fixed displays are normally augmented environments since users are normally able to see real-world objects or their own bodies. However, a world-fixed displays can be considered to be an IVE if lighting conditions are carefully controlled so that users can not see their own body or any of the real world. In this case, this work on perception of scene motion applies to world-fixed displays.

2.2 Visual Perception

Visual perception is a broad and deep topic. I will only treat background material relevant to this dissertation—specifically, topics relevant to visual-temporal and visual-motion perception. All information in this section comes from Coren et al. (1999) unless otherwise noted.

The human ability to see motion relies primarily on the visual and vestibular systems. Experiences and expectations also affect how people perceive motion. Although other cues (such as auditory, tactile, and motion of the neck) may contribute to how we see motion in some situations, I do not focus upon these cues in this work.

2.2.1 The Visual System

Light falls onto photoreceptors in the retina of the eye. These photoreceptors transduce photons into electrochemical signals that travel through two primary systems in the brain.

2.2.1.1 The Photoreceptors: Rods and Cones

The *retina* is a structure of photoreceptive cells in the eye and consists of two types of receptors—rods and cones. *Rods* are primarily responsible for vision at low levels of illumination. Rods are extremely sensitive in the dark but cannot resolve fine details. *Cones* are primarily responsible for vision for high levels of illumination, color vision, and detailed vision. The fovea is a small area in the center of the retina that contains only cones packed densely together. The fovea is located on the line of sight, so that when a person looks at something, its image falls on the fovea.

Electrochemical signals from several photoreceptors converge to a single neuron in the retina. The number of converging signals per neuron increases towards the periphery, resulting in higher sensitivity to light, but decreased visual acuity. In the fovea, some cones have a “private line” to the brain (Goldstein, 2007).

2.2.1.2 Two Visual Pathways

The human visual system can be divided into two visual pathways: the parvocellular and the magnocellular. These two pathways extend from cells in the retina to high levels of the cortex.

The *parvocellular pathway* is responsible for spatial analysis and color perception. Although the parvocellular pathway covers all areas of the visual field, a greater portion of parvo neurons are dedicated to the fovea. Parvo neurons have small receptive fields (the area on the retina that influences the firing rate of the neuron) allowing perception of details such as texture, depth, and local shape. These neurons have a slow conduction rate and a sustained response, resulting in poor sensitivity to motion.

The *magnocellular pathway* is responsible for motion detection. Magno neurons have a fast transient response, meaning the neurons briefly fire when a change occurs and then stop responding. Hence, high sensitivity to motion. Magno neurons have large receptive fields but are blind to color. Magno neurons enable observers to perceive large visuals quickly before perceiving small details.

2.2.1.3 Central Versus Peripheral Vision

Central and peripheral vision have different properties, due to not only the retina, but also to processing within the brain. Central vision

- has high visual acuity,
- is optimized for bright daytime conditions, and
- is color-sensitive.

Peripheral vision

- is color insensitive,
- is more sensitive to light than central vision in dark conditions,
- is less sensitive to longer wavelengths (i.e., red),
- has fast response and is more sensitive to fast motion and flicker,
- is less sensitive to slow motions.

Sensitivity to motion in central and peripheral vision is described further in Section 2.2.7.4.

2.2.2 The Vestibular System

The *vestibular system* consists of labyrinths in the inner ears that act as mechanical motion detectors (Figure 2.2). The vestibular system provides input for balance and sensing self-motion. The vestibular organs are composed of the semicircular canals (SCCs) and the otolith organs. Each set (right and left) of the three nearly orthogonal SCCs acts as a three-axis gyroscope, measuring angular velocity to a first order-approximation (for velocities above 0.1 Hz). Each set of two otolith organs (left and right) acts as a three-axis accelerometer, measuring linear motion.

2.2.3 Eye Movements

Six extraocular muscles control rotations of each eye around three axes. Eye movements can be classified in several different ways. I categorize them into gaze-shifting, fixational, and gaze-stabilizing eye movements.

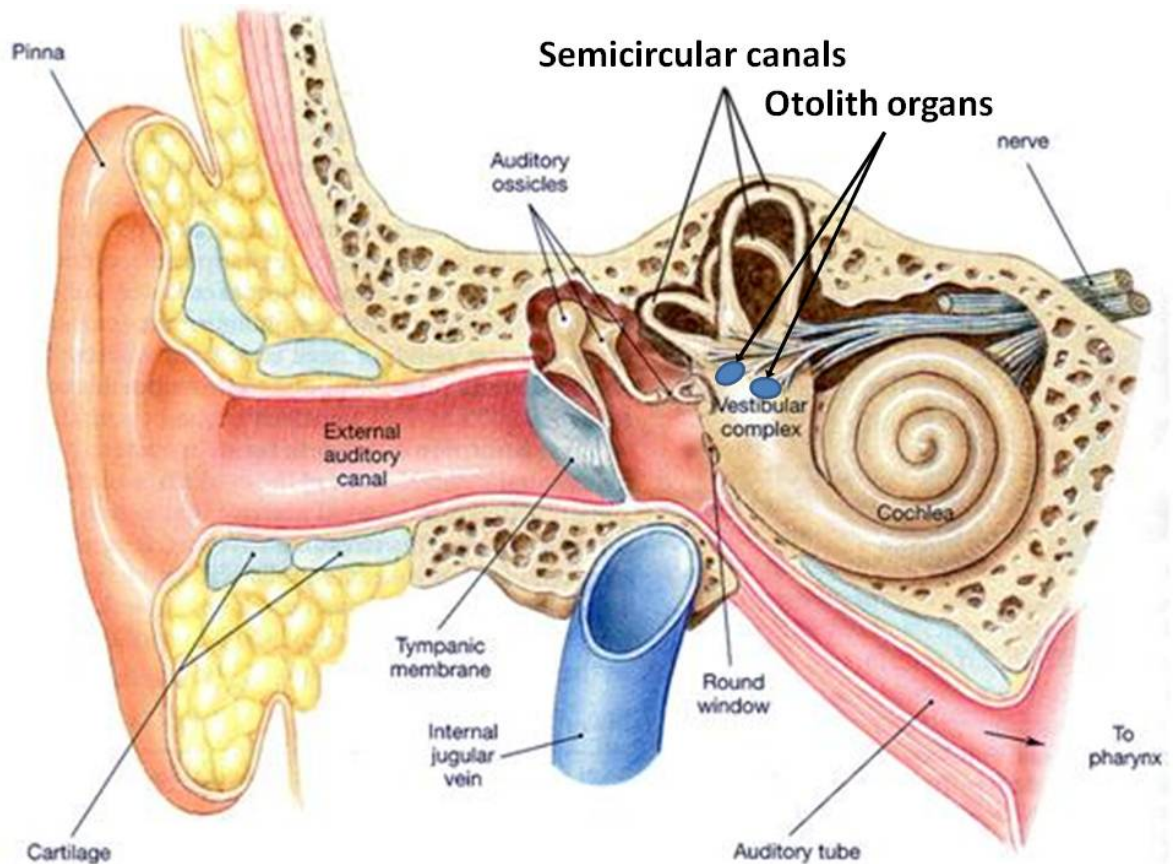


Figure 2.2: A cut-away illustration of the outer, middle, and inner ear, revealing the vestibular system [adapted from Martini (1998)].

2.2.3.1 Gaze-Shifting Eye Movements

Gaze-shifting eye movements enable people to track moving objects or look at different objects.

Pursuit is the voluntary tracking of a visual target. The purpose of pursuit is to stabilize a target on the fovea, in order to provide maximum resolution, and to prevent motion blur, as the brain is too slow to process foveal images moving faster than a few degrees per second. It is important to note that observers still perceive movement even when pursuing a single subject-relative cue (no other objects are visible), due to efference copy (Section 2.2.4).

Saccades are fast voluntary or involuntary movements of the eye that allow different parts of the scene to fall on the fovea. Saccades are the fastest moving external part of the body with speeds up to $1000^{\circ}/s$ (Bridgeman et al., 1994). Most saccades are about 50 ms in duration (Hallett, 1986).

Saccadic suppression greatly reduces vision during and just before saccades, effectively blinding observers. Although observers do not consciously notice this loss of vision, events can occur during these saccades and observers will not notice. A scene can rotate by 8–20% of eye rotation during these saccades without observers noticing (Wallach, 1987). If eye tracking were available in a VE, then the system could perhaps (for the purposes of redirected walking) move the scene during saccades without users perceiving motion.

Vergence is the simultaneous rotation of both eyes in opposite directions in order to obtain or maintain binocular vision for objects at different depths. *Convergence* rotates the eyes towards each other. *Divergence* rotates them away from each other.

2.2.3.2 Fixational Eye Movements

Fixational eye movements enable people to maintain vision when holding the head still and looking in a single direction. These small movements keep rods and cones from becoming bleached. Humans do not consciously notice these small and involuntary eye movements, but without such eye movements the visual scene would fade into nothingness.

Small and quick movements of the eyes can be classified as *microtremors* (less than one minute of arc at 30-100 Hz) and *microsaccades* (about five minutes of arc at variable rates) (Hallett, 1986).

Ocular drift is slow movement of the eye, and the eye may drift as much as a degree without the observer noticing (May and Badcock, 2002). Involuntary drifts during attempts at steady fixation have a median extent of 2.5 arc minutes and have a speed of about 4 arc minutes per second (Hallett, 1986).

In the dark, the drift rate is faster. Ocular drift plays an important part in the autokinetic illusion described in Section 2.2.7.6. As discussed there, this may play an important role in judgments of position constancy.

2.2.3.3 Gaze-Stabilizing Eye Movements

Gaze-stabilizing eye movements enable people to see objects clearly even as their heads move.

Retinal image slip is movement of the retina relative to a visual stimulus being viewed (Stoffregen et al., 2002). The greatest potential source of retinal image slip is due to rotation of the head (Robinson, 1986). Two mechanisms stabilize gaze direction as

the head moves—the *vestibular-ocular reflex* (VOR) and the *optokinetic reflex* (OKR).

The VOR rotates the eyes as a function of vestibular input and occurs even in the dark with no visual stimuli. Eye rotations due to the VOR can reach smooth speeds up to $500^\circ/s$ (Hallett, 1986).

The OKR stabilizes retinal gaze direction as a function of visual input from the entire retina. If uniform motion of the visual scene occurs on the retina, then the eyes reflexively rotate to compensate. Eye rotations due to OKR can reach smooth speeds up to $80^\circ/s$ (Hallett, 1986). VOR is not perfect, and the OKR corrects for residual error.

Gain is the ratio of the velocity of eye rotation divided by the velocity of head rotation (Hallett, 1986; Draper, 1998). Gain due to the VOR alone (i.e., in the dark) is approximately 0.7. If the observer imagines a stable target in the world, gain increases to 0.95. If the observer imagines a target that turns with the head, then gain is suppressed to 0.3 to 0.5. If one adds OKR by providing a stabilized visual target, gain is close to 1 over a wide range of frequencies.

Gain also depends on the distance to the visual target being viewed. For an object at an infinite distance, the eyes are looking straight ahead in parallel. In this case, gain is ideally equal to one so that the target image remains on the fovea. For a closer target, eye rotation must be greater than head rotation for the image to remain on the fovea. These differences in gains are because the axis of the rotation of the head is different than the axis of rotation for the eyes. The differences in gain can quickly be demonstrated by comparing eye movements while looking at a finger held in front of the eyes versus an object further in the distance.

Nystagmus is a rhythmic and involuntary rotation of the eyes (Howard, 1986a). Researchers typically discuss nystagmus caused by rotating the observer continuously at a constant angular velocity. The eyes rotate to stabilize gaze direction. This rotation is called the *slow phase* of nystagmus. As the eyes reach their maximum amount of rotation relative to the head, a saccade snaps the eyes back to looking straight ahead relative to the head. This rotation is called the *fast phase* of nystagmus. This pattern repeats, resulting in a rhythmic movement of the eyes.

Pendular nystagmus occurs when one rotates her head back and forth at a fixed frequency. This results in an always-changing slow phase with no fast phase.

Little is known about how users rotate their eyes while turning their heads when latency-induced scene motion occurs in an HMD. Subjects in one latency experiment (Ellis et al., 2004) claimed to concentrate on a single feature in the moving visual field.

However, this was anecdotal evidence and could not be verified. A user's eye gaze might remain stationary in space (because of the VOR) when looking for scene motion, resulting in retinal image slip, or might follow the scene (because of OKR), resulting in no retinal image slip. I postulate this varies by the amount of head motion, the type of task, the subject, etc. Eye tracking would allow an investigator to study this in detail.

2.2.4 Afference and Efference

Afferent nerve impulses travel from sense organs inwards towards the central nervous system. *Efferent* nerve impulses travel from the central nervous system outwards towards effectors such as muscles. *Efference copy* sends a signal equal to the efference to an area of the brain that predicts afference. This prediction allows the central nervous system to initiate responses before sensory feedback occurs. The brain compares the efference copy with the incoming afferent signals. If the signals match, then this suggests the afference is solely due to the observer's actions. This is known as *re-afference*, since the afference reconfirms that the intended action has taken place. If the efference copy and afference do not match, then the observer perceives the stimulus to have occurred from a change in the external world.

Afference and efference copy result in the world being perceived differently depending upon whether the stimulus is applied actively or passively. For example, when an observer moves her eyes with the eye muscles (active), efference copy compensates for the movement of the scene on the retina (re-afference), resulting in the perception of a stable scene (Figure 2.3). However, if the same observer pushes on her eyeball with her finger (passive), she perceives a moving scene. This perceived shifting of the scene is due to the brain receiving no efference copy from the brain's command to move the eye muscles. The afference (the movement of the scene on the retina) is then greater than the zero efference copy, so scene motion is perceived.

Hence, sensitivity to motion of a visual scene while the head moves depends upon whether the head is actively moved by the observer or passively moved via an external force (Howard, 1986a; Draper, 1998). Thus, in all my experiments, subjects actively rotate their heads, instead of their heads being rotated passively via a controlled rotary chair or other mechanical device.

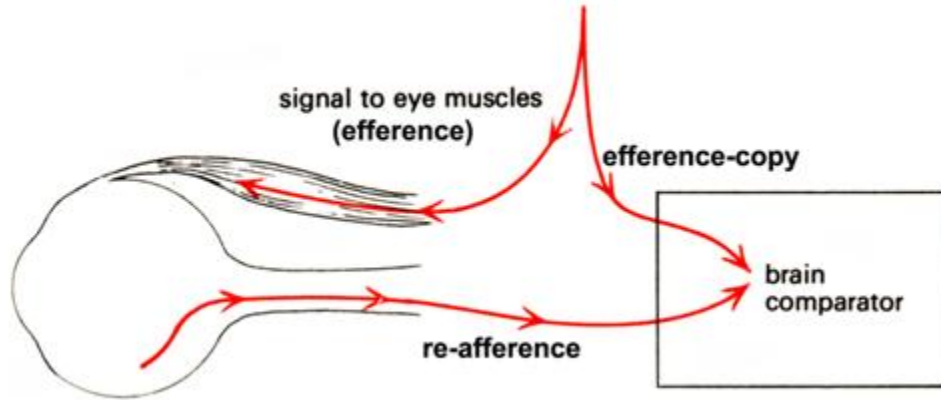


Figure 2.3: Efference-copy during rotation of the eye [Razzaque (2005) after Gregory (1973)]

2.2.5 Intersensory Interactions

Visual-motion perception takes into account sensation from sensory modalities besides just the eye.

2.2.5.1 Visual and Vestibular Cues Complement Each Other

The vestibular system provides only first-order approximations of angular velocity and linear acceleration, and drift occurs over time. Thus, absolute position or orientation cannot be determined from vestibular cues alone, as any airplane pilot has learned. Visual and vestibular cues combine to enable people to disambiguate moving stimuli and self-motion.

The vestibular system is a mechanical system and has a faster response than the slower electrochemical visual response. The VOR is most effective at 1-7 Hz and progressively less effective at lower frequencies, particularly below 0.1 Hz. The OKR is most sensitive at frequencies below 0.1 Hz and has decreasing effectiveness at 0.1-1 Hz. Thus the VOR and OKR complement each other for typical head movements (Draper, 1998).

2.2.5.2 Cue Conflicts

Cue conflicts are mismatches between signals from different sensory modalities. Perceptually unstable worlds and motion sickness (Section 2.4.1.3) can result if visual signals are not consistent with vestibular signals.

2.2.5.3 Other Interactions

Still other sensory organs may affect visual perception. For example, if audio cues are not precisely synchronized with visual cues, then observers may sense the difference. Intersensory interactions may result in surprising results. For example, auditory stimulation can influence the visual critical flicker fusion frequency (the frequency at which subjects start to notice the flashing of a visual stimulus) (Welch and Warren, 1986).

2.2.6 Top-Down Processing

An observer's experiences and expectations influence what she perceives. At any given time, an observer has an internal model of the world. The perception of an external stimulus is heavily biased towards this internal model (Goldstein, 2007; Gregory, 1973).

2.2.7 Motion Perception

2.2.7.1 Visual Velocity

The primate brain has visual receptor systems dedicated to motion detection (Lisberger and Movshon, 1999) which humans use to perceive movement (Nakayama and Tyler, 1981). Whereas visual velocity (i.e., speed and direction of a visual stimulus on the retina) is sensed directly in primates and humans, visual acceleration is not, but is instead inferred through processing of velocity signals (Lisberger and Movshon, 1999). Most visual-perception scientists agree that for motion perception, visual acceleration is not as important as visual velocity (Regan et al., 1986). Thus, I assume scene velocity is a better measure of scene-motion thresholds than scene acceleration.

2.2.7.2 Object-Relative versus Subject-Relative Judgments

Judgment of motion is *object-relative* when an object is judged to move relative to another object. Object-relative judgments depend solely upon stimuli on the retina and do not take into account extra-retinal information such as eye or head motion. Which object is moving is ambiguous.

Judgment of motion is *subject-relative* when an object is judged to move relative to the observer. This egocentric frame of reference is provided by extra-retinal information such as vestibular input. This egocentric frame may be the body (body-centric), the head (head-centric), or the eyes (ocular-centric).

Subject-relative cues always exist. Even when the head is held still, the brain still receives input from the vestibular system that the head is not moving. Thus object-relative cues cannot occur in isolation—when object-relative cues are present then subject-relative cues are also present. A subject-relative cue without any object-relative cues occurs only when a single visual stimulus is visible or all visuals in the scene have the same motion.

People use both object-relative and subject-relative cues to judge motion of the external world. Humans are much more sensitive to object-relative cues than subject-relative cues (Mack, 1986). Incorrectly moving visuals are more easily noticed in optical-see-through HMDs (i.e., augmented reality) than non-see-through HMDs (Azuma, 1997), since users can directly compare rendered visuals relative to the real world, and hence see spurious object-relative motion. This work focuses upon non-see-through HMDs, and thus I explore only subject-relative cues (See Section 4.1.1).

2.2.7.3 Depth Perception Affects Motion Perception

The pivot hypothesis (Gogel, 1990) states that a point stimulus at a distance will appear to move as the head moves if it's perceived distance differs from its actual distance. A related effect is demonstrated by focusing on a finger held in front of the eyes and noticing that the background further in the distance seems to move with the head. Likewise if one focuses on the background then the finger seems to move against the direction of the head.

Objects in HMDs appear to be closer to users than their intended distance (Loomis and Knapp, 2003). If a user is looking at an object as if it is close, but it moves in the HMD as if it is further away (when turning the head), then, according to the pivot hypothesis, the object will appear to move with the head. In this case, the scene would have to move against the direction of the head turn to appear stable in space.

2.2.7.4 Motion Perception in Peripheral versus Central Vision

The literature conflict as to whether sensitivity to motion increases or decreases with eye eccentricity (the distance from the stimulus on the retina to the center of the fovea). These conflicting claims are likely due to differences of experimental conditions as well as interpretations.

Anstis (1986) states that it is sometimes mistakenly claimed that peripheral vision is more sensitive to motion than central vision. In fact, the ability to detect slow moving

stimuli actually decreases steadily with eye eccentricity. However, since sensitivity to static detail decreases even faster, peripheral vision is relatively better at detecting motion than form. A moving object seen in the periphery is perceived as something moving, but it is more difficult to see what that something is.

Coren et al. (1999) states that the detection of movement depends both on the speed of the moving stimulus and eye eccentricity. A person's ability to detect slow-moving stimuli (less than $1.5^\circ/s$) decreases with eye eccentricity, which is consistent with Anstis. For faster-moving stimuli, however, the ability to detect moving stimuli increases with eye eccentricity. These differences are due to the dominance in the periphery of the magnocellular pathway, which consist of transient-response cells. Transient-response cells respond best to fast-changing stimuli. Scene motion in the peripheral visual field is also important in sensing self-motion.

2.2.7.5 Motion Perception During Head Movement

Loose and Probst (2001) found increasing angular velocity of the head significantly suppresses the ability to detect visual motion when the visual motion moves relative to the head. They found no statistically significant effect of head angular acceleration on the ability to detect visual motion. They controlled head velocity to be greater than zero and approximately constant for different amounts of head acceleration, and vice versa. They did not measure the ability to detect visual motion with near-zero head velocity and some head acceleration (i.e., at the start, end, or reversal of head turns). Head acceleration may be an important factor when head velocity is near zero.

Loose and Probst also found subjects could better detect visual motion when the visual motion moved by a lesser amount than head motion compared to when the visual motion moved by a greater amount than head motion.

The moving visual stimuli in their experiment were presented in head-centric coordinates, visual motion was judged object-relative to a head-stabilized target, and the subjects were passively rotated via a rotary chair. These conditions contrast to those conditions in an IVE, where scenes are judged to be moving in world coordinates, judgments are subject-relative, and subjects actively rotate their own heads.

Adelstein et al. (2006) and Li et al. (2006) also showed head motion suppresses perception of visual motion. They used an HMD without head tracking, resulting in the image moving in head-centric coordinates; subjects judged motion relative to the head. In my experiments, I investigate if head motion suppresses perception of motion of images in world-centric coordinates; subjects judge visual motion relative to the

world.

2.2.7.6 Motion Illusions

Motion illusions occur in specific situations of real life and often cause people to mischaracterize the environment. Knowledge of these illusions can help investigators improve understanding of human perception, better design perceptual experiments, interpret experimental results, and better design IVEs.

The Autokinetic Effect

The *autokinetic effect* is the apparent movement of a single stable point-light in a homogeneous surround (no object-relative cues are present). This effect occurs even when the head is held still. This is because no efference copy occurs for small and slow eye movements. The brain has no relative cues to judge motion and the movement is ambiguous, i.e., did the movement occur due to the eye moving or due to the point-light source moving? An observer does not know the answer, and the amount and direction of perceived motion varies. The autokinetic effect decreases as the size of the target increases, due to the brain's assumption that objects taking up a large field of view are stable.

The autokinetic effect suggests that large scenes should be used when measuring the ability to detect subject-relative scene motion and I do so in my experiments. Otherwise, subjects may claim to perceive motion even when no motion occurs.

The Aubert-Fleischl and Filhene Illusions

The *Aubert-Fleischl illusion* causes a moving object to appear to be slower when one pursues the object with the eyes than when the eyes and head are stable and the image of the object moves across the retina. This illusion occurs as the object moves in front of a stationary background (i.e., the effect only appears when object-relative cues are present).

The *Filhene Illusion* causes a stable background to appear to move against the direction of eye motion when one tracks a moving object with the eyes.

I designed my experiments to include only subject-relative cues by blocking out the stationary real world. Thus, the Aubert-Fleischl and Filhene Illusions should not be a factor in this work.

Motion aftereffects

Motion aftereffects are illusions that occur after one views stimuli that move in the same direction for at least 30 seconds. After this time, the observer may perceive the motion to slow down or stop completely due to fatigue of motion-detection neurons. When the motion stops or the observer looks away from the moving stimuli to non-moving stimuli, she may perceive motion in the opposite direction as the previously moving stimuli.

I therefore had to be careful of motion aftereffects when measuring scene-motion

thresholds. In order to prevent motion aftereffects, I

- do not present the moving scenes for more than 1.2 seconds,
- randomly move the scenes from left-to-right or right-to-left, and
- present stable reference scenes before and after the moving scenes.

Induced Motion

Induced motion occurs when motion of one object induces the perception of motion in another object. The moon-cloud illusion is an excellent example of induced motion—one may perceive clouds to be stable and the moon to be moving, when in fact the moon is stable and the clouds are moving. This illusion occurs because object-relative cues tell only how objects move relative to each other, not relative to the world. In such circumstances, the mind assumes smaller objects are more likely to move than larger surround objects.

My experiments include subject-relative cues only, by blocking out the stationary real world. Thus induced motion should not be a factor in this work.

Vection

Vection is an illusion similar to induced motion, but instead causes a perception of self-motion. If an observer is presented with a steadily moving visual pattern, the observer may feel as if she is moving. If the pattern is moved steadily to the side, she may feel that she is moving or leaning to the opposite side and will compensate by leaning into the direction of the visual pattern. Vection is more likely to occur when a large stimulus is moving and when that stimulus is seen in peripheral vision (rather the stimulus is seen or not in the foveal area of vision). Normally, the observer correctly perceives herself to be stable and the stimulus to be moving before the onset of vection. This delay of vection usually lasts several seconds. However, at stimulus accelerations of less than about $5^\circ/s^2$, vection is not preceded by a period of perceived stimulus motion, as it is at higher rates of acceleration (Howard, 1986b).

Vection can occur when one is seated in a car and an adjacent stopped car pulls away. One experiences the car one is seated in, which is actually stationary, to move in the opposite direction of the moving car. Vection can occur in virtual environments when the entire scene moves independent of user movement.

Latency in HMDs may also cause vection, due to the visual scene moving in a way that does not correspond to head movements (a cue conflict). The observer may incorrectly attribute the movement of the scene to his own motion.

Vestibular stimulation suppresses vection (Lackner and Teixeira, 1977). Since scene motion due to latency occurs only when one moves the head (stimulating the vestibular system), I suspect vection due to latency is relatively rare.

Delayed Perception as a Function of Luminance

In dark environments, the eye trades its acuity in space and time for increased light sensitivity. The retina integrates over a longer period of time in the dark, delaying perception of stimuli. The differences in delay between a light and dark environment can be up to 100 *ms* (Anstis, 1986). This delay produces a lengthening of reaction-time for automobile drivers in dim light (Gregory, 1973). Given that visual delay varies for different amounts of dark adaptation or stimulus intensity, then why do people not perceive the world to be unstable or experience motion sickness when they experience greater delays in the dark or with sunglasses as they do in delayed HMDs? I can think of two possible answers to this question:

- The brain might recognize stimuli to be darker and calibrate for the delay appropriately. If this is true, this suggests users can adapt to latency in HMDs. Furthermore, once the brain understands the relationship between wearing an HMD and delay, position constancy could exist for both the delayed HMD and the real world; the brain would know to expect more delay when an HMD is being worn. However, the brain has had a lifetime of correlating dark adaptation and/or stimulus intensity to delay. The brain may take years to correlate the wearing of an HMD to delay.
- The delay inherent in dark adaptation and/or low intensities is not a single precise single delay but the average delay of the integrated signals from a stimulus. This imprecise delay results in motion smear (Coren et al., 1999) and makes precise localization of moving objects more difficult. Perhaps observers are biased to perceive objects to be more stable in such dark situations, but not for the case of brighter HMDs. If this is true, perhaps darkening the scene or adding motion blur to the scene would cause it to appear to be more stable and reduce simulator sickness (although computing motion blur typically adds latency unless prediction is used).

2.2.8 Perceptual Constancies

A *perceptual constancy* is the apparent stability of objects even though the actual stimuli change due to changes in lighting, viewing position, etc. Perceptual constancies occur partly due to the observer's understanding that objects tend to remain constant in the world.

All information in this section comes from Wallach (1987) unless otherwise noted.

2.2.8.1 Position Constancy

Position constancy is the perception that an object appears to be stationary in the world even as the eyes and the head move (Mack and Herman, 1972). A perceptual matching process between motion on the retina and extra-retinal cues cause this head-centric motion of the visual field to be discounted.

The real world remains stable as one rotates the head. Turning the head 10° to the right causes the visual field to rotate to the left by 10° relative to the head. The *displacement ratio* is the ratio of the angle of environmental displacement, in world coordinates, to the angle of head rotation. The real-world is stable no matter what head movement is and thus has a displacement ratio of zero. If the environment rotates with the head then the displacement ratio is positive. Likewise, if the environment rotates against head direction the displacement ratio is negative.

Minifying glasses, i.e., glasses with concave lenses, increase the displacement ratio and cause the object to appear to move with the head. Magnifying glasses decrease the displacement ratio below zero and cause the object to move in the opposite direction as head rotation.

The *range of immobility* is the range of displacement ratio where position constancy is perceived. Wallach determined the range of immobility to be 0.04 to 0.06 displacement ratios wide. If an observer rotates her head 100° to the right, then the environment can move by up to $\pm 2^\circ$ to 3° with or against the direction of her head turn without her noticing that movement.

In an untracked HMD, the scene moves with the user's head and has a displacement ratio of one. In a tracked HMD with latency, the scene first moves *with* the direction of the head turn (a positive displacement ratio), then the world moves back to its correct position, *against* the direction of the head turn (a negative displacement ratio), after the system has caught up to the user.

In my experiments, I measure how much scene motion can occur while subjects perceive position constancy. These scene motions are measured in terms of angular scene velocity divided by the angular range of head turns, angular scene velocity divided by angular head velocity, and angular scene velocity divided by angular head acceleration.

2.2.9 Adaptation

Humans must be able to adapt to new situations in order to survive. People must not only adapt to an ever-changing environment, but also adapt to internal changes such as neural processing times and spacing between the eyes, both of which change over a lifetime. When studying perception in virtual environments, investigators must be aware of adaptation, as it may confound measurements.

Negative aftereffects are changes in perception of the original stimulus after the adapting stimulus has been removed. These negative aftereffects provide the most common measure of adaptation (Cunningham et al., 2001).

2.2.9.1 Sensory and Perceptual Adaptation

Adaptation can be divided into two categories: sensory adaptation and perceptual adaptation (Wallach, 1987).

Sensory adaptation alters a person's sensitivity to detect a stimulus. Sensitivity increases or decreases over time—one stops detecting or starts detecting the original stimulus after a period of constant stimulus intensity. Dark adaptation is an example of sensory adaptation.

Perceptual adaptation alters a person's perceptual processes. Welch (1986) defines *perceptual adaptation* to be “a semipermanent change of perception or perceptual-motor coordination that serves to reduce or eliminate a registered discrepancy between or within sensory modalities or the errors in behavior induced by this discrepancy”. VOR adaptation is an example of perceptual adaptation.

2.2.9.2 Light and Dark Adaptation

The sensitivity of the eye changes by as much as six orders of magnitude, depending on the lighting of the environment. The cones reach maximum dark adaptation in approximately 10 minutes after initiation of dark adaptation. The rods reach maximum dark adaptation in approximately 30 minutes after initiation of dark adaptation. Complete light adaptation occurs within five minutes after initiation of light adaptation.

Rods are relatively insensitive to red light. Thus, if red lighting is used, rods will dark adapt whereas cones will maintain high acuity.

Dark adaptation causes a person's perception of stimuli to be delayed. By presenting a relatively large and bright scene between trials, I keep subjects light-adapted. Alternately only red stimuli could be used to maintain dark adaptation.

2.2.9.3 Position-Constancy Adaptation

The compensation process that keeps the environment stable during head rotation can be altered by perceptual adaptation. Wallach (1987) calls this adaptation to be “adaptation to constancy of visual direction”, and I call it *position-constancy adaptation*. This adaptation corrects for the effect of eyeglasses that cause the stationary environment to move optically during head movements.

Wallach and Kravitz (1965) had subjects wear minifying glasses that caused the world to seemingly move with the direction of head turns. They found that over time subjects' perceived motion of the environment during head turns subsided, resulting in apparent position constancy. After adaptation and removal of the optical device, subjects reported a negative aftereffect such that the real world seemed to move against the direction of their

head turns. Draper (1998) found similar adaptation for subjects in HMDs when the rendered field of view was intentionally modified to be different from the true field of view of the HMD.

People can learn to perceive position constancy for different displacement ratios if there is a cue (e.g., glasses on the head or scuba-diver face masks (Welch, 1986)).

Position-constancy adaptation could be due to vestibular adaptation, eye-movement adaptation, or visual-field adaptation.

If vestibular adaptation is a cause, then sounds should be perceived to move in the same way the visual scene appears to move after position-constancy adaptation. Wallach and Kravitz (1968) tested if this was the case, but sounds remained spatially fixed after position-constancy adaptation. Thus, vestibular adaptation is not at work here.

If eye-movement adaptation is a cause, then the observer should adapt when tracking a small target, with a displacement ratio different than zero, even with a surrounding stationary pattern (with a displacement ratio equal to zero). Wallach and Canal (1976) indeed found this to be the case.

If visual-field adaptation is a cause, then the observer should adapt when when keeping their gaze stabilized on a stable target (with a displacement ratio equal to zero), even when when a surrounding visual pattern moves with some displacement ratio. Wallach and Canal (1976) indeed found this to be the case.

Thus, the position-constancy adaptation process is due to both eye-movement adaptation, or visual-field adaptation, but not due to vestibular adaptation.

2.2.9.4 Temporal Adaptation

Cunningham et al. (2001) provided behavioral evidence that humans can adapt to a new intersensory temporal relationship caused by delayed visual feedback. A virtual airplane was displayed moving downward with constant velocity on a standard monitor. Subjects attempted to navigate through an obstacle field by moving a mouse that controlled only left/right movement. Subjects first performed the task in a pre-test with a visual latency of 35 *ms*. Subjects were then trained to perform the same task with 200 *ms* of additional latency introduced into the system. Finally, subjects performed the task in a post-test with the original minimum latency of 35 *ms*. The subjects performed much worse in the post-test than in the pre-test.

Toward the end of training, with 235 *ms* of visual latency, several subjects spontaneously reported that visual and haptic feedback seemed simultaneous. All subjects showed very strong negative aftereffects. In fact, when the latency was removed, some subjects reported the visual stimulus seemed to move before the hand that controlled the visual stimulus, i.e., a reverse of causality occurred, where effect seemed to occur before cause!

Previous studies had not been able to show adaptation to latency. The authors reasoned

that sensorimotor adaptation to latency requires exposure to the consequences of the discrepancy. Subjects in previous studies were able to reduce discrepancies by slowing down their movements when latency was present, whereas in this study subjects were not allowed to slow the constant downward velocity of the airplane.

These results suggest HMD users may adapt to latency, thereby changing latency thresholds over time. If this is the case, the adaptation could cause latency thresholds to vary over time. I present reference scenes with zero effective latency between test scenes containing some latency to prevent latency adaptation from occurring. I also randomly vary latency between trials so that subjects do not have a constant latency to adapt to.

2.3 Psychophysics

Psychophysics is the study of relationships between physical stimuli and sensations (Fechner, 1860).

Ferwerda (2008) provides motivation for using psychophysics to help understand perception of computer graphics:

One of the hallmarks of a science is the development of reliable techniques for measuring the phenomena of interest. For example, in physics there are standard techniques for measuring an object’s mass or velocity. In chemistry there are techniques for measuring the strength of bonds or the energy given off by a reaction. In computer graphics, we would like to be able to measure what people perceive when they look at our images, so we can create more effective and compelling images, and create them more efficiently. However, while physical quantities can be measured with standard methods, our visual experiences are subjective and cannot be measured directly.

If computer graphics is to become a “science of visualization”, then researchers need to be able to conduct experiments that quantify the relationships between the parameters of our algorithms, the images they produce, and the visual experiences the images create. Psychophysics provides a sound set of methodologies for conducting and analyzing these experiments.

Ideally, one would like to minimize variance, bias, and confounding factors, so that one measures only systematic effects that are due to a change in the dependent variable(s). Psychophysics experiments often limit experimental conditions to very specific conditions (i.e., control as many variables as possible) in order to minimize these affects. This comes at the cost of external validity, i.e., the results may not apply in situations other than those tested. I do not expect all results obtained from my work to hold across conditions different

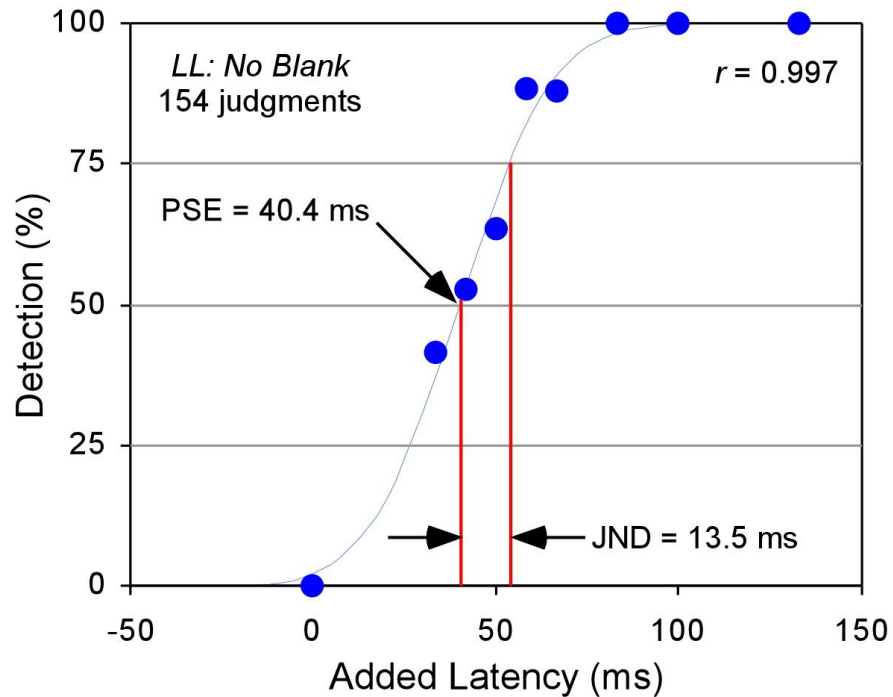


Figure 2.4: A typical psychometric function. The greater the stimulus (added latency in this case), the greater the detection rate [adapted from Adelstein et al. (2005)].

from my narrowly defined conditions. My goal is to develop a method that will allow scientists to determine latency requirements for their specific conditions.

All information in this section is derived from Gescheider (1997) unless otherwise noted.

2.3.1 Psychometric Functions

A *psychometric function* relates stimulus intensity to probabilities of detecting that stimulus (Macmillan and Creelman, 2005).

Figure 2.4 shows the percentages of times that a subject detects different levels of stimulus intensities (added latency in this case). The psychometric function in this case is the cumulative Gaussian that best fits the proportions.

2.3.2 Thresholds

A perceptual *threshold* is the change in intensity of a stimulus that is required for a subject to detect it with some probability. Fechner (Fechner, 1860) defined a threshold as a stimulus intensity that “lifted the sensation... over the threshold of consciousness.”

A 50% threshold is the level of stimulus at which the stimulus is correctly detected 50% of the time. 50% thresholds (typically the point of subjective equality), 75% thresholds, and

just-noticeable differences (typically defined to be the difference between the 75% threshold and the 50% threshold) are commonly reported.

2.3.2.1 Point of Subjective Equality

The *point of subjective equality* (PSE) is the stimulus intensity at which a subject perceives two or more options (e.g., less than or greater than, present or not present, same or different, selecting the different stimulus from multiple stimuli) to be equally likely. The PSE requires a perceptual comparison of a test stimulus with a reference, or at least one other stimulus in order for the subject to judge the test stimulus to be “subjectively equal”. In this sense, one can say that the stimuli are perceptually equivalent on the dimension of interest. If the subject has two options to choose from, then the PSE will be the 50% threshold.

The PSE reflects subject bias (i.e., the tendency to say yes or no) and is influenced heavily by a subject’s expectations and whether the subject is a liberal or conservative responder. If the subject is a liberal responder, i.e. for a variety of different possible reasons she is more inclined to state “signal present”, then the PSE will be low. Liberal responders result in higher hit rates (stating “signal present” when a signal exists), but at the cost of higher false alarm rates (stating “signal present” when no signal exists). On the other hand, if the subject is a conservative responder, i.e., she is more inclined to state “signal not present”, then the PSE will be high. Conservative responders result in a lower false alarm rate, but at the cost of a higher miss rate.

Discriminability refers to how easy it is to detect that a signal is present. Signal detection theory can be used to differentiate between a subject’s discriminability and her bias. However, signal detection theory usually requires an enormous collection of data. Bias can be largely removed by using an alternative forced-choice identification task (Section 2.3.4.3).

2.3.2.2 Just-Noticeable Difference

The *just-noticeable difference* (JND) is the change in a stimulus required for an observer to perceive that change. The JND is, by convention (although arbitrarily chosen), defined to be the change in stimulus intensity that causes a change in correct detection probability from 50% to 75% (Ellis et al., 2004). The JND is proportional to the standard deviation of a subject’s observations—the greater the JND then the greater the dispersion of the observations. The greater the JND, then the smaller the slope of the psychometric function at the point of inflection. The JND reflects observer sensitivity to changes in the physical stimulus. The greater the JND then the less sensitive is the subject.

For most stimuli, changes in the intensity of the stimuli that are just noticeably different are a constant proportion of the stimulus intensity. This is known as *Weber’s Law* and is

written as

$$\text{JND} = kI \quad (2.1)$$

where k is a constant dependent on the stimulus and I is stimulus intensity.

2.3.3 Determining Psychometric Functions and Thresholds

2.3.3.1 Method of Adjustment

The *method of adjustment* approximates 50% detection thresholds by having a subject directly adjust (e.g., via a knob) a stimulus intensity until she can just start to perceive the stimulus, or just start to stop perceiving the stimulus. The experimenter can then take the average of multiple thresholds for better precision and accuracy of the actual threshold.

JNDs can also be approximated by the method of adjustment. In this case, the subject is instructed to adjust a comparison stimulus until it seems equal to a standard stimulus. The subject is presented with the stimulus many times. The subject will sometimes underestimate and sometimes overestimate the standard by a considerable amount, but most of the matches tend to cluster around the value of the standard stimulus. This will often result in a normal distribution. The mean of this distribution is equivalent to a 50% PSE. The 25% JND is equivalent to .68 standard deviations (i.e., from 50% to 75% or from 50% to 25%).

2.3.3.2 Method of Constant Stimuli

The *method of constant stimuli* is a procedure of presenting a predetermined ordering of stimuli for the entire experiment. The responses of the subject do not affect future stimuli levels. The higher stimulus values should be high enough so that they are almost always detected. Likewise, lower stimulus values should be low enough that they are almost never detected. These values can be estimated, before the experiment is conducted, by a method of adjustment as described above. The proportion of hits for each stimulus level are then graphed and fit to a cumulative Gaussian curve (or some other curve), resulting in a psychometric function.

2.3.3.3 Methods of Limits

The *method of limits* begins by presenting a stimulus intensity well above or well below threshold. When starting with a low stimulus, an ascending series ensues—the stimulus is increased every trial until the subject detects the stimulus. The series then terminates and a new series begins. When starting with a high stimulus, a descending series ensues—the stimulus is decreased until the subject can no longer detect the stimulus. Averaging multiple termination points give the 50% threshold. Figure 2.5 shows an ascending and descending

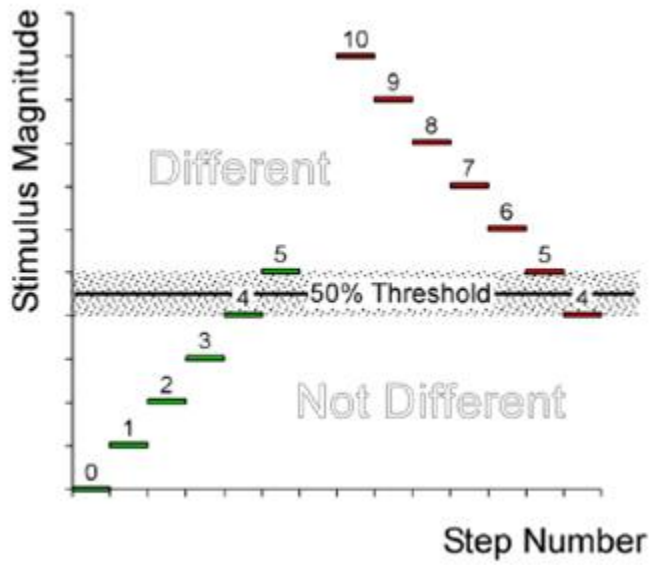


Figure 2.5: Method of Limits [adapted from Swan et al. (2007)].

series with the 50% threshold estimate. Proportions of detections at each stimulus level can also be analyzed to determine the entire psychometric function (Swan et al., 2007).

2.3.3.4 Adaptive Staircases

An *adaptive staircase* is a psychophysics procedure that results in stimulus levels concentrating near a point of interest. A starting stimulus level is chosen and the following stimulus levels are determined by preceding stimulus levels and subject responses (Levitt, 1970). For a one-up-one-down staircase, the stimulus level decreases after the subject successfully detects the stimulus and increases after the subject fails to detect the stimulus. Figure 2.6 shows a one-up-one-down adaptive staircase, it converges to the 50% detection threshold. Other adaptive staircases converge to different threshold levels, e.g., a one-up-two-down staircase converges to the 70.7% detection threshold. Values near the 70.7% detection threshold are helpful for exploring the region between the 50% PSE and the 75% PSE+JND.

The step size (the change in stimulus level) may start high and then become lower each time a staircase reversal occurs. This allows the threshold to converge faster than when using a constant step size.

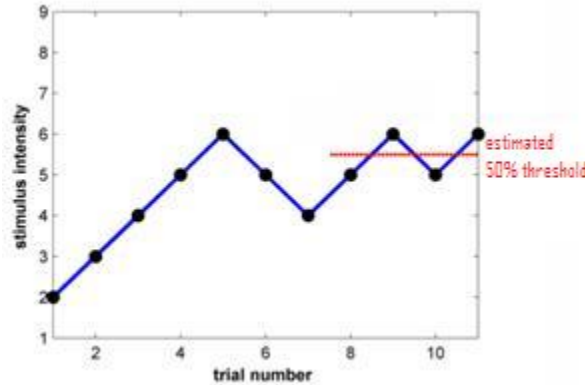


Figure 2.6: A one-up-one-down adaptive staircase. The stimulus intensities converge to a 50% detection threshold. [adapted from Razzaque (2005)].

2.3.4 Judgment Tasks

Detection rates, psychometric functions, and thresholds vary depending upon the experimental task. For example, experiment designs differ in how many presentations the user sees before making a judgment as to whether the stimulus is present or not.

2.3.4.1 Yes/No Tasks

In a *yes/no task*, the experimenter provides a presentation and the subject states if she believes the presentation contained the stimulus or not. Yes/no tasks can be performed quickly, since they require only a single presentation to be judged.

Yes/no tasks assume the subject knows what the “signal present” looks like. By providing a reference presentation with no stimulus before the test presentation, the subject is better able to judge what the “signal present” looks like.

The disadvantage of a yes/no task is that there is a large amount of bias because the resulting threshold largely depends upon the subject’s belief of “how much” perceived signal constitutes a “yes” response. A liberal responder may almost always say “yes” resulting in a low threshold whereas a conservative responder may almost always say “no” resulting in very different thresholds. Yes/no tasks are useful for determining single subjects’ thresholds (assuming their criterion remains the same throughout the experiment) but yes/no tasks are not very useful for computing statistics across subjects due to the large variance between subjects.

2.3.4.2 Same/Different Tasks

In a *same/different task*, the subject states if she believes two presentations are the same or if there is a difference between the presentations. The objective of this design is not to test

specific cues; the subject is free to discriminate on any basis she chooses. This is useful when a change in stimulus level causes several effects that are not well defined, or when “greater” and “less” are not well defined. The subject only states if the presentations different, without stating in what way they are different or if one is “greater” or “less” than the other.

The disadvantage of a same/different task is, like a yes/no task, there is a large amount of bias; the amount of difference required for the subject to state “different” is a subjective judgment. A subject may almost always choose “different” due to a criterion where even the slightest possibility of differences between the presentations results in a “different” judgment. Or a subject may almost always choose “same” unless the presentations are completely different.

2.3.4.3 Identification Tasks

In an *n*-alternative forced-choice (*nAFC*) *identification task* multiple presentations are provided, with one presentation containing the stimulus. The subject selects which presentation she believes contains the stimulus. When the stimulus is too weak to be detected, the PSE will take on a range of values with a detection rate of $1/n$ where n is the number of alternatives. In this case the threshold is often defined to be the stimulus level with a detection rate halfway between guessing and perfect judgment: $1/n + ((n - 1)/n)/2$. (e.g., the 75% threshold for a 2AFC task).

Bias is smaller for an nAFC identification task than for a yes/no task and a same/different task because for an nAFC identification task, the subject is forced to choose between alternative presentations. Assuming the presentations are randomized, subjects are forced to choose the “best” answer and their choice is largely independent of their tendency to be a liberal or a conservative responder.

2AFC identification tasks assume that the subject knows what “signal present” looks like. If the subject perceives two presentations to be different, but she does not understand what a stimulus looks like then she may consistently choose the wrong presentation (the presentation with no stimulus instead of the presentation with the stimulus).

A *three-interval two-alternative forced-choice (3I-2AFC) identification task* is similar to a 2AFC identification task, except that the first presentation contains a reference stimulus and the subject chooses from the following two presentations. This has the advantages of both the same/different task (since the subject chooses which of the two is different from the reference) and the 2AFC identification task (since bias is minimized). The 3I-2AFC identification task also has an advantage over a 3AFC identification task because the subject knows the first presentation is a reference presentation with no stimulus. The disadvantage of the 3I-2AFC identification task (and the 3AFC identification task) compared to all previously described tasks is that trials take a longer period of time when presentations are presented sequentially

in time.

2.4 Latency

I define *system delay* in an HMD system to be the true time from the start of head movement to the time a pixel resulting from that movement responds to the viewpoint change without any compensation (e.g., prediction). I consider pixel response to be a change of intensity of 50% from starting intensity to intended intensity, unless otherwise noted.

I define *latency* to be the *effective* time from the start of head movement to the time a pixel resulting from that movement responds to the viewpoint change. If no pixel motion occurs due to system delay, then I consider latency to be zero. Compensation techniques can reduce latency but cannot reduce system delay. Section 2.4.4 provides an overview of some common compensation techniques to reduce system delay and latency.

2.4.1 Human Factors

Latency in an HMD-based system causes visual cues to lag behind other perceptual cues, creating sensory conflict. With latency and some head motion, the visual scene presented to a subject in an HMD moves incorrectly. This scene motion due to latency is known as “swimming” and has serious usability consequences. In this section, information is derived from Allison et al. (2001) unless otherwise referenced.

2.4.1.1 Degraded Visual Acuity

Latency can cause degraded vision. Given some latency, as an HMD user moves her head then stops, the image is still moving when the head has stopped moving. If the image velocity on the retina is greater than $2-3^\circ/s$, then motion blur and degraded visual acuity result. Typical head motions and latencies result in image motion greater than $3^\circ/s$ in world coordinates. For example, a sinusoidal head motion of 0.5 Hz and $\pm 20^\circ$ and 133 ms of latency results in a peak scene motion of $\pm 8.5^\circ/s$ (Adelstein et al., 2005). It is not known if users’ eyes tend to follow a lagging scene, resulting in no retinal image slip, or if their eyes tend to stay stabilized in space, resulting in retinal image slip.

2.4.1.2 Performance

The level of latency necessary to negatively impact performance may be different from the level of latency that can be perceived. So and Griffin (1995) studied the relationship between latency and operator learning in an HMD. The task consisted of tracking a target with the

head. Training did not improve performance when latency was ≥ 120 ms. Thus, the subjects were unable to learn to compensate for these latencies in this task.

2.4.1.3 Simulator Sickness

Symptoms of motion sickness include nausea, dizziness, blurred vision, disorientation, and vertigo (Kennedy and Lilienthal, 1995; Kolasinski, 1995). The *sensory conflict theory of motion sickness* is the most widely accepted explanation for the initiation of motion sickness symptoms (Harm, 2002). The theory states that motion sickness may result when the environment is altered in such a way that information from the senses is not compatible and does not match previously stored neural patterns.

Latency in an HMD causes scene motion in the HMD that conflicts with other sensory cues (e.g., visual cues get out of phase with vestibular cues) and is different from what users expect of the real world. *Simulator sickness* is sickness similar to motion sickness that results from a simulator but not in the real situation that is being simulated (Pausch et al., 1993).

2.4.1.4 Breaks in Presence

Presence is a subjective illusion of “being there” in the virtual world. Latency is a factor that detracts from the sense of presence in HMDs (Meehan et al., 2003). A *break-in-presence* is a moment when the illusion generated by a virtual environment breaks down and the user finds himself where he truly is—in a laboratory wearing an HMD (Slater and Steed, 2000). Latency combined with head movement causes the scene to move in a way not consistent with the real world. This incorrect scene motion can distract the user, who might otherwise feel present in the virtual environment, and cause her to realize the illusion is only a simulation.

2.4.1.5 Negative Training Effects

I have not found any specific research documenting negative-training effects due to latency in HMDs. However, latency has been shown to result in negative training effects with desktop displays (Cunningham et al., 2001) and driving simulators utilizing large screens (Cunningham et al., 2001). For non-HMD vehicle simulators, Denne (2004) claims no motion cues are better than bad motion cues, as the motion cues are worse than useless because they teach the wrong intuition and can cause negative training effects. I suspect similar effects can occur for HMDs.

2.4.2 Perception of Latency

2.4.2.1 Temporal Perception versus Motion Perception

A *perceptual moment* is the smallest psychological unit of time that an observer can sense (Coren et al., 1999). Based on several research findings, (Stroud, 1986) estimated that perceptual moments are about 100 *ms* in duration. Stimuli presented within the same perceptual moment would be perceived as occurring simultaneously or would not be distinguishable from each other. The length of these perceptual moments depends upon the sensory modality, stimulus, task, etc.

Efron (1967) concluded that human visual-perceptual moments are 60 to 70 *ms*. He had subjects compare a flash of light that lasted 1 *ms* to longer flashes of light. The subjects were only able to differentiate between the 1 *ms* flash of light and a longer flash of light when the longer flash of light lasted a minimum of 60 to 70 *ms*.

In a similar study, subjects were able to discriminate temporal differences in some cases as low as ~ 10 *ms* Nilsson (1969). However, subjects reported they made judgments, at least in part, based upon perceived brightness, instead of temporal cues alone.

Evidence suggests perceptual moments become unstable when one judges with more than one sensory modality, such as when judging the order of presentation of a sound and a light (Ulrich, 1987) Directly sensing time differences between visual stimuli and head turns might result in temporal thresholds greater than Efron's perceptual moments of 60 to 70 *ms*.

Latency thresholds for HMDs are typically less than perceptual moments of 60 to 70 *ms* (see Sections 2.4.2.2 and 2.4.2.3). Adelstein et al. (2003) found subjects, at least in part, judge latency based indirectly upon scene motion resulting from latency.

In order for observers to detect the visual-temporal aspect of latencies directly, they must have some visual cue to base their temporal judgments on. In the case of detecting latency in HMD-based systems, the only thing that changes visually over time is the motion of the scene as the observer moves her head (assuming a static scene). If subjects do not perceive a scene-motion cue then there is nothing to compare temporal differences to. Observers must first detect scene-motion before making visual-temporal judgments when detecting the latency. Thus, I assume subjects detect latency based upon the visual instability of the scene rather than latency directly.

2.4.2.2 Latency PSEs

Various experiments conducted at NASA Ames Research Center (Adelstein et al., 2003; Ellis et al., 2004; Mania et al., 2004) reported latency thresholds during quasi-sinusoidal head yaw for a same/different task where the probe latency was the stimulus added to a scene with some base latency. They found PSEs (the 50% threshold in their case) to vary considerably

(ranging from -0.5 ms to $\sim 85\text{ ms}$) due to bias, type of head movement, individual differences, differences of experimental conditions, and other known and unknown sources of variance.

2.4.2.3 Latency JNDs

The NASA experiments (Ellis et al., 1999; Adelstein et al., 2003; Ellis et al., 2004; Mania et al., 2004) found JNDs of latency to be more consistent than PSEs. This is consistent with the fact that JNDs are less sensitive to bias than PSEs (Section 2.3.2.2). They found JNDs to range from $\sim 4\text{ ms}$ to $\sim 40\text{ ms}$.

Ellis et al. (1999) showed that users are just as sensitive to changes in latency with a low base latency as those with a higher base latency. Thus, latency discrimination does not follow Weber’s law (see Section 2.3.2.2). Consistent latency is important even when average latency is high.

2.4.2.4 Latency Discrimination Mechanisms

Adelstein et al. (2005) found subjects wearing a non-see-through HMD to be more sensitive to latency at the reversal of head turns, when the head changes direction and latency-induced scene velocity peaks, than the center of head turns, when the head moves with near-constant velocity and latency-induced scene displacement peaks. These results suggest users are more sensitive to incorrect velocity of a scene than to incorrect position of a scene. Since my work also focuses upon non-see-through HMDs, their result has led me to focus on scene velocity rather than scene displacement.

2.4.2.5 Scene Complexity

The NASA measurements showed no differences in latency thresholds for different scene complexities, ranging from single, simple objects (Ellis et al., 2004) to detailed photorealistic rendered environments (Mania et al., 2004). For this work, I use a simple geometric pattern for the scene and expect the results to generalize to more complex scenes.

2.4.3 System Delay

Miné (1993) and Olano et al. (1995) characterize system delays in IVEs and discuss various methods of reducing latency. System delay is the sum of delays from tracking, application, rendering, display, and synchronization among components. Figure 2.7 shows how these delays contribute to total system delay. Note that system delay can be greater than the inverse of the update rate, i.e., a pipelined system can have a frame rate of 60 Hz but have a delay of several frames.

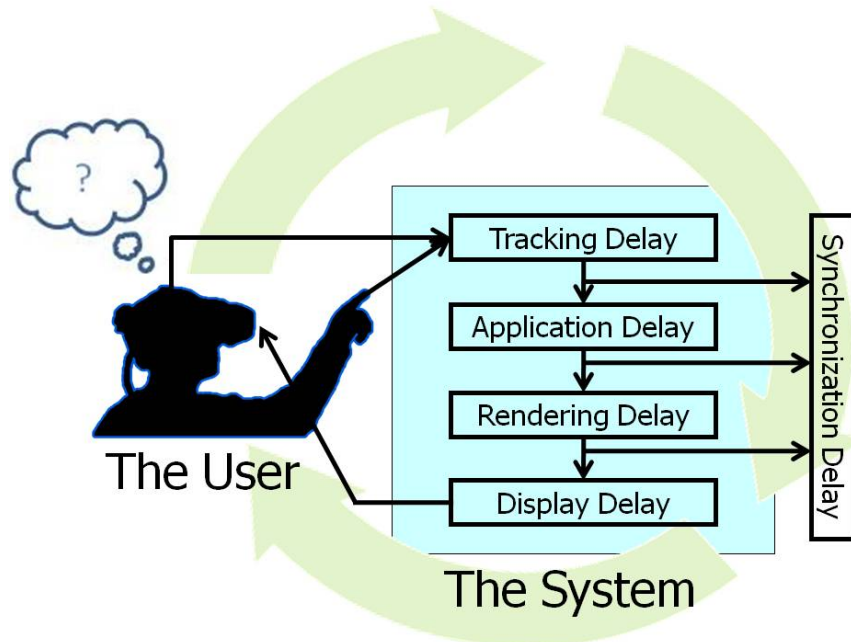


Figure 2.7: End-to-end system delay comes from the delay of the individual systems components and from the synchronization of those components.

2.4.3.1 Tracking Delay

Tracking delay is the time from when the tracked part of the body moves until movement information from the tracker’s sensors resulting from that movement is input into the application or rendering component of the VE system. Tracking products can include techniques that complicate delay analysis. For example, many tracking systems incorporate filtering to smooth jitter. If filters are used, the resulting output pose is only partially determined by the most recent tracker reading, so that precise delay is not well defined. Some trackers use different filtering models for different situations—delay during some movements may differ from that during other movements. For example, the 3rdTech HiBall™ tracking system (3rdTech, 2002) allows the option of using multi-modal filtering. A low-pass filter is used to reduce jitter if there is little movement, whereas a different model is used for larger velocities.

Commercial tracker vendors report the delay of their systems—this is typically the minimum tracker delay, and actual delay varies depending on the update rate of the trackers (Section 2.4.3.5).

Tracking is often processed on a different computer from the computer that executes the IVE application and rendering. In that case tracking delay includes the network delay.

2.4.3.2 Application Delay

Application delay is time due to computation other than tracking or rendering. This includes updating the world model, user interaction, physics simulation, etc. This application delay can vary greatly depending on the complexity of the task and the virtual world. Application processing can often be executed asynchronously from the rest of the system (Bryson and Johan, 1996). For example, a weather simulation with input from remote sources could be delayed by several seconds and computed at a slow update rate, whereas rendering needs to be tightly coupled to the latest head pose and occur at a fast frame rate.

2.4.3.3 Rendering Delay

Rendering delay is the time from when new data enters the graphics pipeline to the time an image resulting from that data is completely drawn into a buffer. Rendering delay depends on the complexity of the virtual world, the desired quality of the resulting image, and the performance of the graphics software/hardware. The *rendering rate* is the number of times the system can render the entire scene per second. The *rendering time* is the inverse of the rendering rate, and in non-pipelined rendering systems is equivalent to rendering delay.

Rendering computation is normally separate from the application, typically on dedicated graphics hardware. Current graphics cards can achieve rendering rates of several thousand Hertz for simple scenes. Although rendering is not the dominant factor in most systems, it still contributes to system delay.

2.4.3.4 Display Delay

I define *display delay* to be the time from when a pixel is output from the graphics card to the time that pixel changes its displayed intensity by 50%, from its previous intensity to its new intensity. Display delay depends on the technology of the display hardware.

Response Time

I define the *rise time* of a display to be of a display to be the time it takes for a pixel to change from zero intensity to 50% intensity, unless otherwise noted. I define the *fall time* to be the time it takes for a pixel to change from full intensity to 50% intensity, unless otherwise noted. I define *display response time* to be the longer of the rise or fall time.

Raster Displays

A *raster display* sweeps pixels out to the display, scanning out left-to-right in a series of horizontal scanlines from top to bottom (Whitton, 1984). This pattern is called a *raster*. Timings are precisely controlled to draw pixels from memory to the correct locations on the screen.

Pixels on raster displays have inconsistent intra-frame delay (i.e., different parts of the frame have different delay), because pixels were rendered from a single time-sampled viewpoint but are presented over a period of time—the bottom pixels are displayed at a later time than the top pixels.

A *frame* is the full-resolution image that is scanned out to the display hardware. The *frame rate* is the number of frames scanned out per second (Hz). Note this can be different from the rendering rate described in Section 2.4.3.3. The *frame time* is the inverse of the frame rate (seconds per frame). A display with a frame rate of 60 Hz has a frame time of 16.7 *ms*. Typical computer displays have frame rates from 30-120 *Hz*.

Double Buffering

In order to avoid memory access issues, the system should not read the frame at the same time it is being written to by the renderer. The frame should not scan out pixels until the rendering is complete, otherwise visuals may not be occluded properly. Furthermore, the rendering time and rate varies, depending on scene complexity and hardware capability, whereas the frame rate is set solely by the hardware, which is designed to update at a rate such that users perceive smooth transitions between frames.

A solution to this problem of dual access to the frame can be solved by using a double-buffer scheme. The display processor renders to one buffer while the refresh controller feeds data to the display from an alternate buffer. The *vertical sync* signal occurs between frames. Most commonly, the system waits for this vertical sync to swap buffers. The newly rendered frame is then scanned out to the display while a yet-newer frame is rendered.

Waiting for vertical sync causes additional delay, since rendering must wait up to 16.7 *ms* (for a 60 *Hz* display) before starting to render a new frame. Waiting for vertical sync also results in higher intra-frame delay, because raster displays present pixels at different times—the bottom-most pixels are always presented nearly a full-frame time after the top-most pixels are presented.

Tearing

If the system does not wait for vertical sync to swap buffers, then the buffer swap occurs while the frame is being scanned out to the display hardware. In this case, *tearing* occurs during viewpoint motion and appears as a spatially discontinuous image, due to two or more frames (each rendered from a different sampled viewpoint) contributing to the same displayed image. When the system waits for vertical sync to swap buffers, no tearing is evident, because there is a single rendering for the entire frame.

Figure 2.8 shows a simulated image that would occur with a system that does not wait for vertical sync superimposed over a simulated image that would occur with a system that does wait for vertical sync. The figure shows what a static virtual block would look like when a user is turning her head from right to left. The tearing is obvious when the swap

does not wait for vertical sync, due to four renderings of the images with four different head poses. Thus, most current VR systems avoid tearing at the cost of additional and variable intra-frame delay.

Just-In Time Pixels

Tearing decreases with decreasing differences of head pose. As the sampling rate of tracking increases and the rendering rate increases, pose coherence increases and the tearing becomes less evident. If the system were to render each pixel with the correct up-to-date viewpoint, then the tearing would occur between pixels. The tearing would be small compared to the pixel sizes, resulting in a smooth image—without perceptual tearing. Miné and Bishop (1993) discuss this *just-in-time pixels*.

Rendering could conceivably occur at a rate fast enough that buffers would be swapped for every pixel. Although the entire image would be rendered, only a single pixel would be displayed for each rendered image. However, a 640x480 image would require a rendering rate of over 18 MHz—clearly impossible for the foreseeable future using standard rendering algorithms and commodity hardware. If a new image were rendered for every scanline, then a 640x480 image would require nearly 29 kHz. In practice, today’s systems can render at rates up to 20 kHz for very simple scenes, which make it possible to show a new image every couple of scanlines.

In current HMD systems, delay caused by waiting for vertical sync can be the largest source of system delay. Ignoring vertical sync greatly reduces overall delay at the cost of image tearing.

Field-sequential displays

A *field-sequential display* sequentially displays fields of different colors and/or intensities.

Some field-sequential displays present successive colors in raster patterns (e.g., older color-sequential HMDs that update at 180 Hz with each of the red, green, and blue components updating at 60 Hz).

Other field-sequential displays flash bursts of light for each field such that all pixels in each field are displayed in parallel (e.g., a 24 bit-plane DLP flashes 8 intensities and/or durations for all pixels in parallel for each of the red, green, and blue bits of a frame). One perceives each frame as single analog image due to the persistence of vision.

Whereas these displays have some nice properties (typically fast response times for individual fields), they have several disadvantages:

- An additional frame buffer is often used to store fields before display. This adds an additional 16.7 *ms* of delay for a 60 *Hz* display.
- Field separation can occur in an HMD with fast head motion because of the different latencies for the different fields. By the time a new field is displayed, the user may be

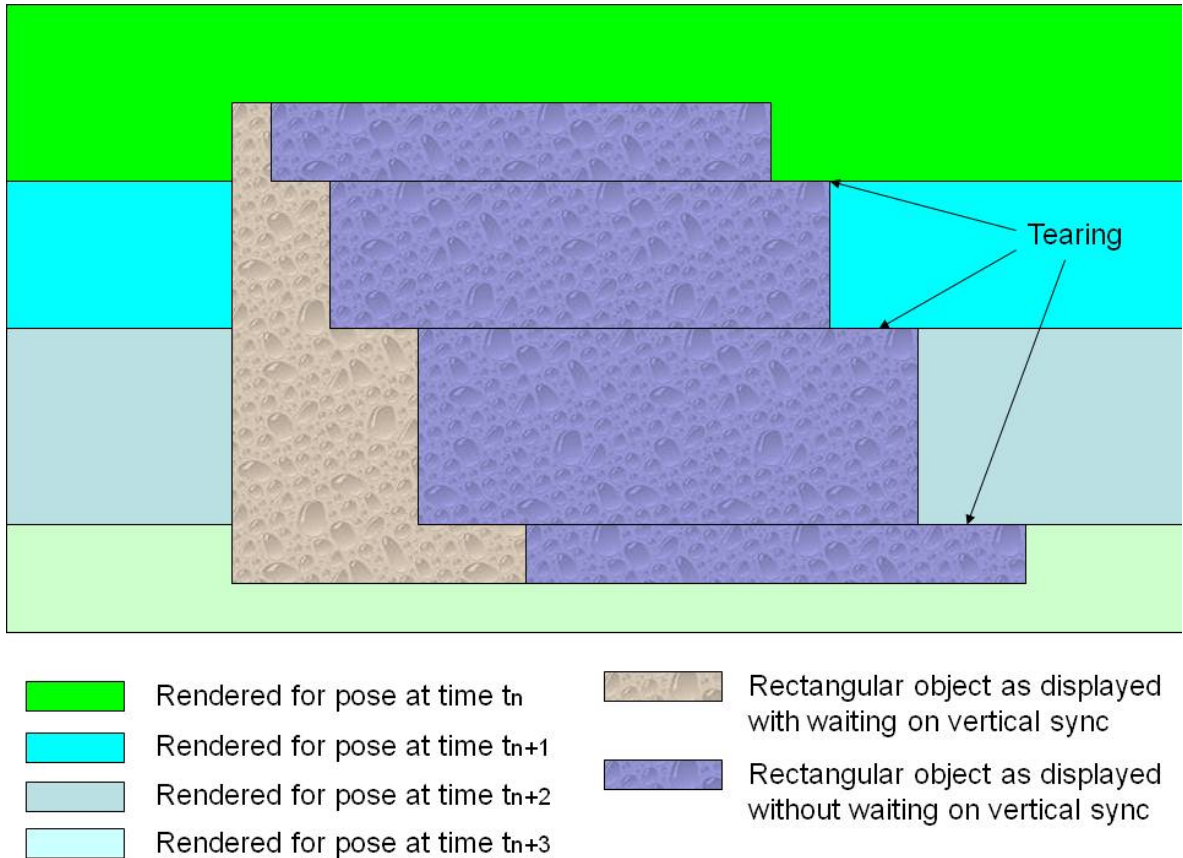


Figure 2.8: A frame showing a rectangular object as the user is looking from right to left with and without waiting for vertical sync to swap buffers. The result of not waiting for vertical sync to swap buffers is perceived as image tearing.

looking in a direction different from where she was looking when the previous field was displayed.

- Updating every field with a new rendering from a new pose would require tracking and rendering updates at $60 \text{ Hz} \times 24 \text{ bit planes} = 1440 \text{ Hz}$. Commodity tracking and rendering hardware can achieve such rates. However, I know of no commercial field-sequential displays that allow direct access to the display hardware without buffering the entire frame.

Liquid Crystals

Various forms of liquid crystals are used in head-mounted displays. These displays may be raster displays or field-sequential displays.

Active-matrix liquid-crystal displays (AMLCDs) have become the display of choice for HMDs (Rolland and Cakmakci, 2005). AMLCDs are typically implemented as raster displays. Unfortunately the response times for AMLCDs are large and variable due to the time it takes

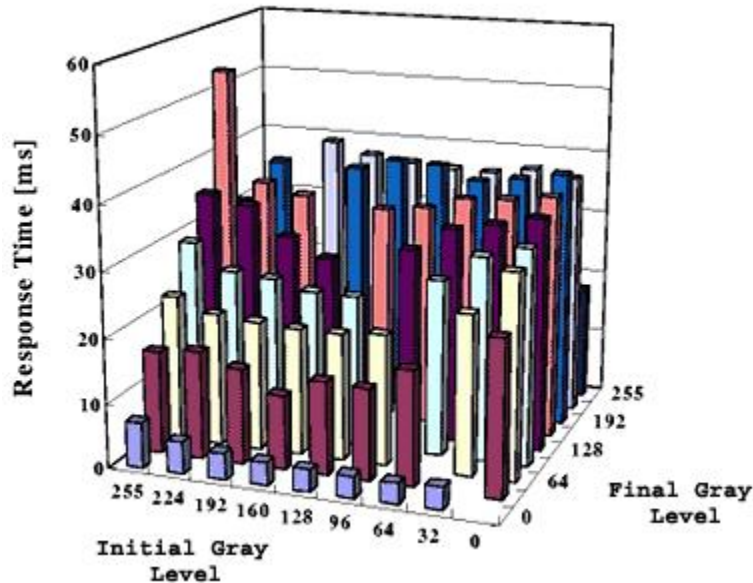


Figure 2.9: Rise and fall times as a function of initial and intended intensities for a typical liquid crystal display [adapted from Nakanishi et al. (2001)]. Response times of such displays vary greatly depending on starting and ending intensities.

the liquid crystals to turn on or off. Figure 2.9 shows response times for a typical AMLCD. In Section 2.4.3.7, I report response times for two popular HMDs that use AMLCDs.

Ferroelectric displays (e.g., as used in the Kaiser SR80 and NVision nVisorSX HMD) use liquid crystals with extremely fast response times of 1-100 μ s. Ferroelectric displays are field-sequential displays.

Organic Light Emitting Diode

Organic Light Emitting Diodes (OLEDs) use polymers that emit light when an electrical current is passed through and have response times less than 1 ms (Rolland and Cakmakci, 2005). OLED displays are typically Raster Displays.

Digital Light Processing

Digital Light Processing (DLP) projectors are field-sequential displays that use micro-mechanical mirrors, one for each pixel, that are turned in one of two orientations so that light is reflected or not reflected onto a display surface. DLPs projectors have become a popular alternative to CRT or LCD projectors. Response times from the time the signal to turn the mirror to the time the pixel is displayed is well below one millisecond. However, all DLP projectors I know of buffer the entire frame before starting to display, such that there is more than an additional frame of delay (the time for the buffered frame plus the time from vertical sync to the time that the individual field is displayed).

Virtual Retinal Display

The Virtual Retinal Display (VRD) scans a modulated light directly onto the retina of the eye, eliminating the need for screens and heavy, expensive imaging optics (Kollin, 1993). There is zero response time as the modulated light falls directly on the retina the same way that light from the real world does.

Cathode Ray Tubes

A *Cathode Ray Tube* (CRT) emits a stream of electrons from a heated cathode to a phosphor-coated screen (Foley et al., 1990). The phosphor emits light when excited by the electrons. The phosphor determines the response time. Modern CRTs have a response time of less one millisecond and scanout in a raster pattern.

The advantage of CRTs are

- black is black (a zero signal produces no light),
- response times (rise and fall times) are below one millisecond (for modern monitors and projectors),
- no frame buffers exist between video input and pixel output, and
- display times for individual pixels can be precisely predicted to compensate for system delay (Section 2.4.4.1).

Because of these advantages, I use a CRT projector for all experiments.

2.4.3.5 Synchronization Delay

Total system delay is not simply a sum of component delays. I define *synchronization delay* to be the delay that occurs due to integration of pipelined components. Synchronization delay is equal to the sum of component delays subtracted from total system delay. Synchronization delay can be due to components waiting for a signal to start new computations and/or can be due to asynchrony among components.

Pipelined components depend upon data from the previous component. When a component starts a new computation and the previous component has not updated data, then old data must be used. Alternatively, the component can in some cases wait for a signal or wait for the input component to finish.

Trackers provide a good example of a synchronization problem. Commercial tracker vendors report their delays as the tracker response time—the minimum delay incurred until the tracker outputs pose information. If the tracker is not synchronized with the application or rendering component, then the tracking update rate is also a crucial factor and affects both average delay and delay consistency. If a tracker reports n times per second then the average

tracking delay, in milliseconds, is

$$delay_{average} = response + (1/2) \times (1/n) = (response + 1/2n) \quad (2.2)$$

The delay range is

$$delay_{range} = delay_{average} \pm (1/2n) = (response + 1/2n \pm 1/2n) \quad (2.3)$$

For a tracker with 5 *ms* response time and an update rate of 50 *Hz*, tracking delay varies between 5 and 25 *ms*. A faster update rate and/or synchronizing components produces more consistent and less delay.

2.4.3.6 Timing Analysis

This section presents an example timing analysis and discusses complexities encountered when analyzing system delay. Figure 2.10 shows a timing diagram for a typical VE system. In this example, the scanout to the display begins just after the time of vertical sync. The display stage is shown for discrete frames, even though typical displays present individual pixels at different times. The image delays are the time from the beginning of tracking until the start of scanout to the display. Response time of the display is not shown in this diagram.

As can be seen in the figure, asynchronous computations often give repeated results when no new input is available. For example, the rendering component cannot start computing a new result until the application component provides new information. If new application data is not yet available, then the rendering stage repeats the same computation.

In the figure, the display component displays frame *n* with the results from the most up-to-date rendering. All the component timings happen to line up fairly well for frame *n*, and image *i* delay is not much more than the sum of the individual component delays. Frame *n* + 1 repeats the display of an entire frame because no new data is available when starting to display that frame. Frame *n* + 1 has a delay of an additional frame time more than the image *i* delay. Frame *n* + 4 is delayed even further due to similar reasons. No duplicate data is computed for frame *n* + 5, but image *i* + 2 delay is quite high because the rendering and application components must complete their previous computations before starting new computations.

2.4.3.7 Measuring Delays

To better understand system delay, one can measure timings for not only the end-to-end system delay but also for sub-components of the system. Means and standard deviations can be derived from several such measurements.

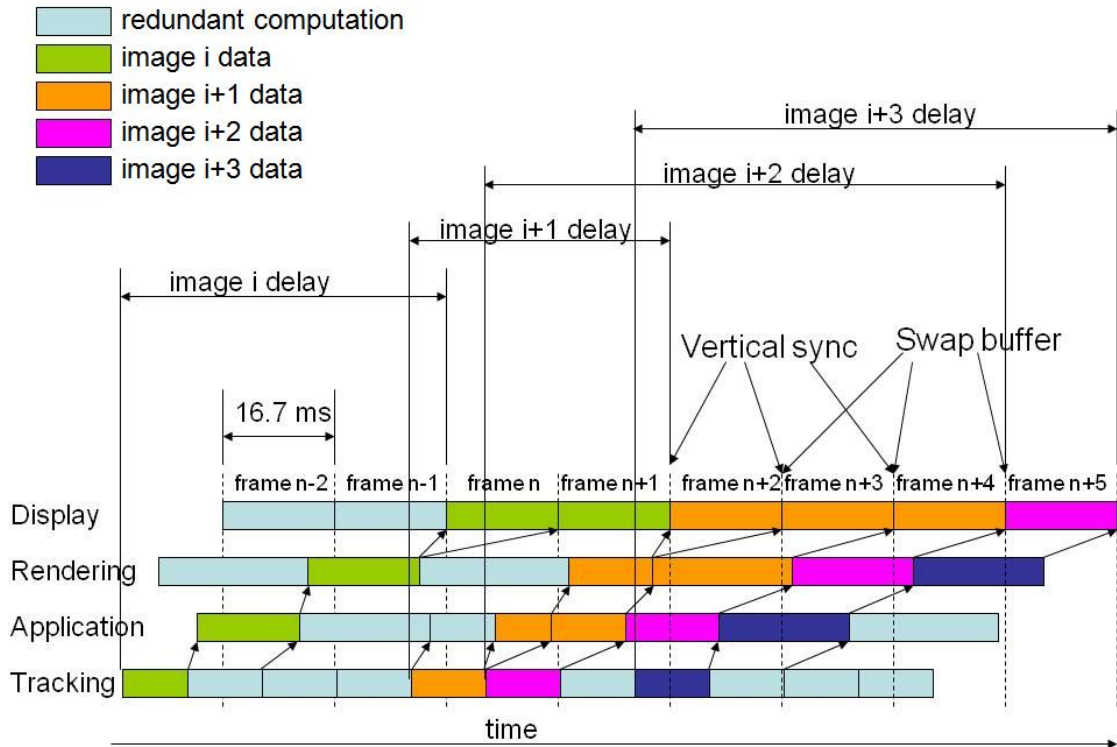


Figure 2.10: A timing diagram for a typical VE system. In this non-optimal example, the pipelined components execute asynchronously. Components are not able to compute new results until preceding components compute new results themselves.

The *latency meter* (Miné and Bishop, 1993; Miller and Bishop, 2002) is a device that measures system delay. The device is shown in Figure 2.11. This device sends a signal to the oscilloscope as the arm of the latency meter crosses the vertical (where the vertical is defined as the low point of the arc of a pendulum’s motion). The latency meter application then renders an alternating black/white screen when the tracker reports its sensor crossing the vertical. A photodiode attached to the display then sends another signal to the oscilloscope when it senses the change of black to/from white. The difference of the times between the two signals, as measured on the oscilloscope, is the system delay.

The video signal can also be sent directly to the oscilloscope. One can then view the video signal to see when the change occurs. In this case the delay measured is total system delay minus the display delay.

Figure 2.12 shows an image from an oscilloscope with input from the latency meter and photodiode measuring the display intensity of a Virtual Research V8 HMD. The system takes ~ 100 ms to go from the time of motion to 50% intensity (the screen goes from white to black). Also note the liquid crystals do not turn instantaneously—the fall time from white to 50%

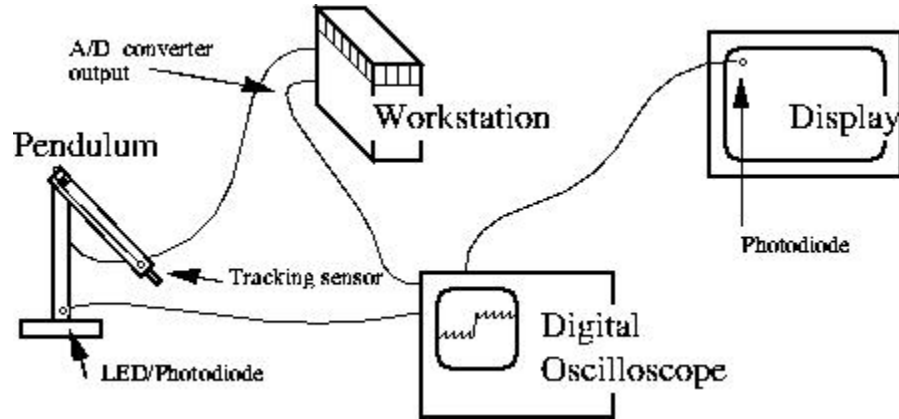


Figure 2.11: The latency meter [adapted from (Miné and Bishop, 1993)].

Table 2.1: Mean display rise and fall times for the Virtual Research V8 and Glasstron HMDs. Rise and fall times are from the time of vertical sync of the VGA output to 16.7%, 50%, and 83.3% of the intended maximum intensity.

	display response		
	16.70%	50%	83.30%
Virtual Research V8 HMD			
rise time (msec)	16.45	31.6	73.3
fall time (msec)	14.5	33.85	113.25
Sony Glasstron			
rise time (msec)	18.76	40.56	101.68
fall time (msec)	19.76	43.64	123.24

intensity is approximately 20 *ms*.

Timings can further be analyzed by sampling signals at various stages of the pipeline and measuring the time differences. The parallel port can be used to output timing signals from both the tracking PC and the application PC. These signals are precise in time since there is no additional delay due to a protocol stack; writing to the parallel port is equivalent to writing to memory;

I measured response times for the UNC Graphics Lab’s Virtual Research V8 and Sony Glasstron HMDs, and for the NASA Ames Kaiser SR80 HMD.

Mean response times for the Virtual Research V8 and Sony Glasstron are shown in Table 2.1. Note the rise time (black to white) is less than than the fall time (white to black) for all but one case. This is because liquid crystals take a longer time to return to their natural “black” state then they take to rotate to their electrically charged “white” state.

I measured the end-to-end latency of the HMD system used by Adelstein et al. (2006) and Li et al. (2006) to be 27.9 *ms*. Figure 2.13 shows the intensity of a red input signal for their

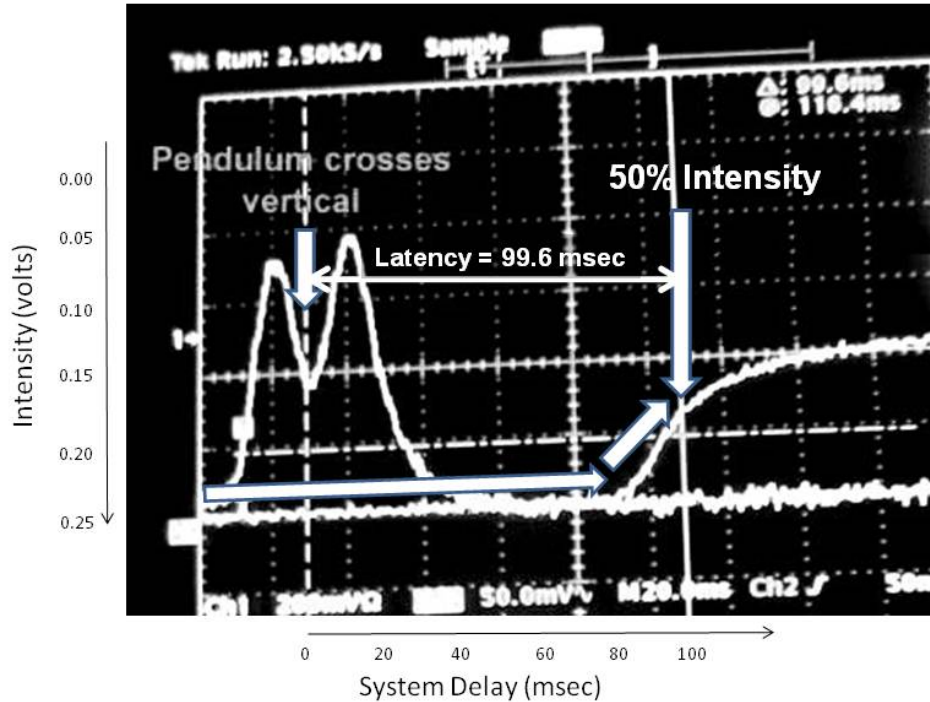


Figure 2.12: This oscilloscope output shows there is about 100 *ms* of system delay from the time of tracking to 50% display response [adapted from Razzaque (2004)]. Note white is at the bottom and black is at the top such that the display changes from white to black.

Kaiser SR80 HMD as a function of time. This display is a 24-sequential-bit-plane display. The local minima are the times that the intensity peaks for several bit-planes; the global minimum is for the most-significant bit's plane. Figure 2.14 shows intensity for the same display with maximum red, green, and blue input. Note the blue intensity peaks approximately 4 *ms* after the red intensity peaks.

Synchronization delays between two adjacent components of the pipeline can also be measured indirectly. If the delays of individual components are known, then the sum of two adjacent components can be compared with the measured delay across both components. The difference is the synchronization delay between the two components.

2.4.4 Delay Compensation

Because computation is not instantaneous, virtual environments will always have delay. Delay-compensation techniques can reduce the deleterious effects of system delay, effectively

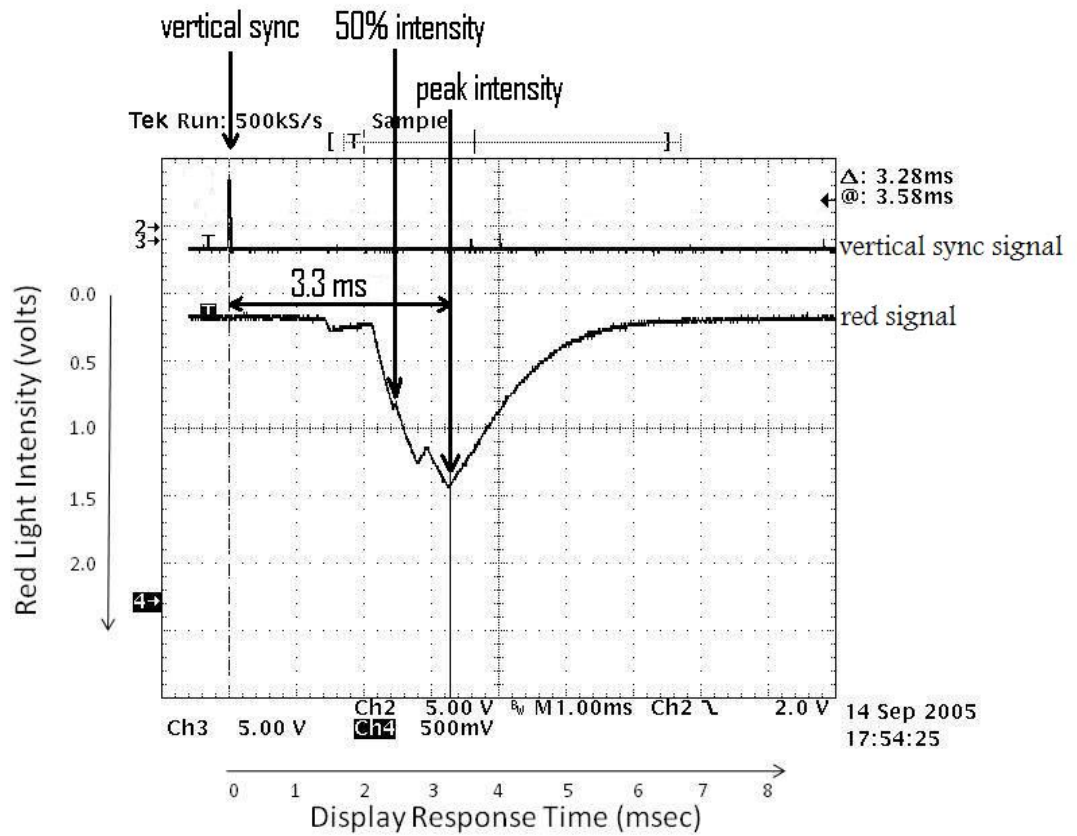


Figure 2.13: Red peak-response time for the NASA Kaiser SR80 HMD (that uses a ferroelectric display). There is 3.3 *ms* of delay from vertical sync to maximum intensity and 2.4 *ms* of delay for 50% intensity. However, there is an additional frame buffer in the hardware to collect all pixels from the VGA input signal, causing an additional 16.7 *ms* of delay.

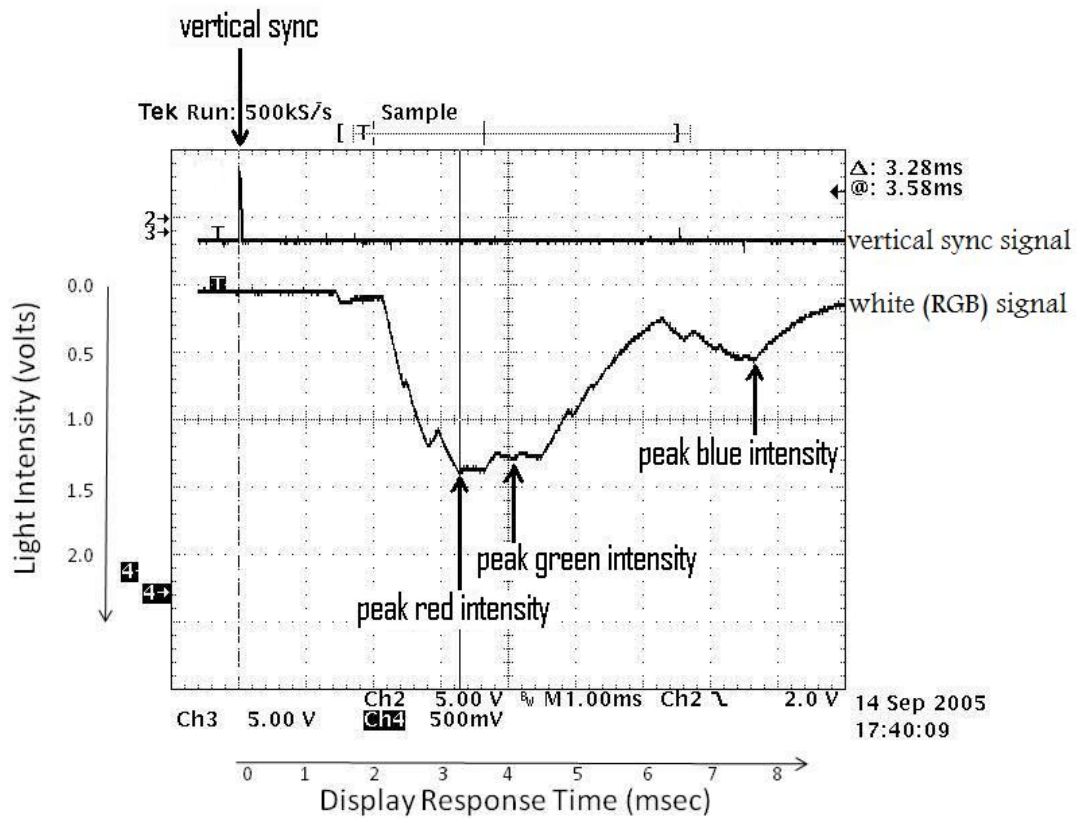


Figure 2.14: Red, green, and blue peak-response times for the NASA Kaiser SR80 HMD. Note the blue intensity peaks approximately 4 ms after the red intensity peaks.

reducing latency.

2.4.4.1 Prediction

Head-motion prediction is a commonly used delay-compensation technique for HMD systems. Prediction produces reasonable results for low system delays or slow head movements. However, prediction increases motion overshoot and amplifies sensor noise (Azuma and Bishop, 1995). Displacement error increases with the square of the prediction interval and the square of the angular head frequency. Furthermore, prediction is incapable of instantaneous response to rapid transients.

2.4.4.2 Post-Rendering Techniques

Post-rendering techniques first render an image larger than the final display and then, late in the display process, select the appropriate subset to be presented to the user.

The simplest post-rendering technique renders the scene to a single image plane. The pixels to be displayed are then selected from the larger image plane to reduce yaw and pitch error (Breglia et al., 1981; Burbidge and Murray, 1989; Murray et al., 1984; Kano et al., 2004; Kijima et al., 2001; Kijima and Ojika, 2002; Mazuryk and Gervautz, 1995; So and Griffin, 1992; Yanagida et al., 1998; Jerald et al., 2007). This technique results in little error if the user is looking in the same general direction as the original projected image plane, the scene is static, and objects are at a distance from the user.

A cylindrical panorama, as in Quicktime VR (Chen, 1995), can be used, instead of a single image plane, to completely remove error due to yaw rotation. However, rendering to a panorama is computationally expensive, because standard graphics pipelines are not optimized to render to such a projection surface.

A compromise between a single image plane and a panorama is a cubic environment map. This technique extends the image plane technique by rendering the scene onto the six sides of a large cube (Greene, 1986). Head rotation simply alters what part of display memory is accessed—no other computation is required. Cubic environment maps do not correct for motion parallax, and thus are not appropriate for geometry close to the viewpoint.

Regan and Pose (1994) took environment mapping further by projecting geometry onto concentric cubes surrounding the viewpoint. Larger cubes that contain projected geometry far from the viewpoint do not require re-rendering as often as smaller cubes that are close to the viewpoint. In order to minimize dynamic registration error due to large translations, a full 3D warp (Mark et al., 1997) or a pre-computed light field (Regan et al., 1999) is required.

2.4.4.3 Problems of Delay Compensation

None of the delay compensation techniques described above are perfect. For example, the single image plane technique minimizes error at the center of the display but error increases towards the periphery in the direction of the head turn (Mazuryk and Gervautz, 1995; Kijima et al., 2001; Kijima and Ojika, 2002; Jerald et al., 2007).

In addition, these compensation algorithms usually focus on reducing displacement error and rarely address velocity error. Velocity error is more important than displacement error with respect to latency perception in HMDs (see Section 2.4.2.3). In fact, some latency-compensation techniques may even degrade perceptual stability of the world due to ignoring velocity error—when a new head pose is determined, systems usually update pixel position as soon as possible, reducing displacement error, but at the cost of the visuals “jumping” with instantaneous velocity.

Compensation techniques can reduce error in real systems but visual artifacts resulting from such techniques can interfere with judgments of latency. Thus, I do not use delay compensation techniques for my systems or my experiments.

2.5 Intentional Scene Motion

People often do not know if they are facing north but they would immediately know if the world were quickly rotated 90 degrees. However, if a virtual world rotated at a slow enough rate such that the motion was imperceptible, it would be even more difficult for someone to know, even after a few moments, which way was true north. Razzaque uses this fact to create *redirected walking* (Razzaque, 2005), a technique that allows users to walk in IVEs larger than the tracked lab space by rotating the scene at velocities below human perception. All 11 subjects in one of his experiments (Razzaque et al., 2001) were surprised to find at the conclusion of the experiment that they were walking back and forth between the ends of the lab—they did not notice the rotational offset of the virtual world.

I define a *reorientation technique* to be any method that causes a user to change his perception of his orientation in an IVE. Redirected walking is a type of reorientation technique. The scene-motion thresholds, and the experimental procedures used to determine scene-motion thresholds, described in this work, can be used as guidelines of how much scene motion can be injected into IVEs without subjects noticing.

Chapter 3

Analysis and Relationships of Latency, Head Motion, Scene Motion, and Perceptual Thresholds

3.1 Error due to Latency

In order to simplify analysis, all error analysis assumes a monocular HMD, ignoring depth error due to disparity. I assume a static virtual world—the world, and objects in that world, are intended to be stable in space. All discussion and analysis are in world space and in degrees (or degree derivatives with respect to time) unless otherwise noted.

3.1.1 Scene Displacement

Displacement error (equivalent to registration error as defined by Holloway (1995) and Azuma and Bishop (1994)) is the difference, in degrees from the viewpoint, of where a point-object is displayed and where it is intended to be displayed. Figure 3.1 shows an example of displacement error.

Displacement error can be classified as static or dynamic (Azuma and Bishop, 1994). *Static error* is error that is independent of head motion—this error occurs even when an observer keeps her head completely still. Factors contributing to static error include optical distortion, incorrect viewing parameters, tracker noise (jitter and drift), and imprecise calibration of equipment. *Dynamic error* is error that occurs when the head moves. As latency or head motion increase, dynamic error increases.

Even for moderate head velocities, latency causes more displacement error than all other sources of displacement errors combined (Holloway, 1995). Displacement error due to latency

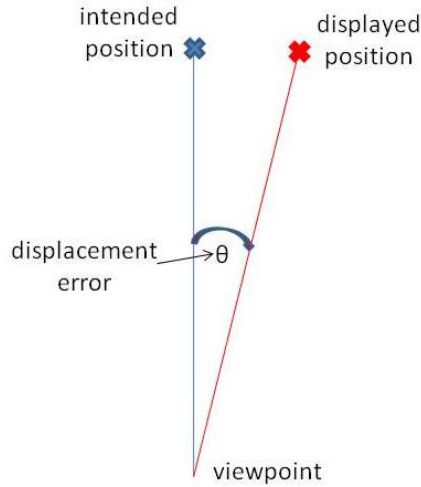


Figure 3.1: Displacement error is the difference, in degrees from the viewpoint, of where a point-object is displayed and where it is intended to be displayed.

and head rotation is greater than displacement error due to latency and head translation for all but the closest objects. For a system with 30 *ms* of latency, a moderate head rotation of $50^\circ/s$, and a fast head translation of 0.5 *meters/s*, displacement error due to rotation is 1.5° , and error due to translation is 0.86° for objects at 1.0 meter. Translation error is a relatively insignificant 0.086° for objects at 10 meters. I assume objects are at a far enough distance such that with error due to translation is small compared to error due to rotation, and hence argue an analysis performed using only head rotation is acceptable. In my experiments, I place a single object at a distance of approximately four meters and subjects yaw their head with little translational movement so that error due to head translation is small.

Scene displacement is the average displacement error over the entire display. Since my analysis considers only pure rotation, scene displacement equals displacement error of any individual point.

Scene displacement θ , in degrees, at time t_0 is

$$\theta_{t_0} = \Delta\phi_{t_0} = \phi_{t_0} - \phi_{t_0-\Delta t} \quad (3.1)$$

where ϕ is head displacement, in degrees, and Δt is latency, in seconds.

Head displacement ϕ at time t_0 can be written as a function of head velocity ϕ' :

$$\phi_{t_0} = \int_{-\infty}^{t_0} \phi'(t) dt \quad (3.2)$$

so that Equation 3.1 can be written as

$$\theta_{t_0} = \int_{-\infty}^{t_0} \phi'(t)dt - \int_{-\infty}^{t_0-\Delta t} \phi'(t)dt \quad (3.3)$$

By the addition property of equality, scene displacement θ simplifies to integrating head velocity ϕ' over the latency interval:

$$\theta_{t_0} = \int_{t_0-\Delta t}^{t_0} \phi'(t)dt \quad (3.4)$$

Given average head velocity $\phi'_{average}$ over the latency interval, the equation simplifies to

$$\theta_{t_0} = \Delta t \phi'_{average} \quad (3.5)$$

For small latency intervals Δt and/or near-constant head velocity over the time of integration, scene displacement θ can be approximated by

$$\theta_{t_0} \approx \Delta t \phi'_{t_0} \quad (3.6)$$

3.1.2 Just-in-Time Scanlines

Given a display width of 640 pixels and a horizontal field-of-view of 32° , there are 0.05° per pixel. Assuming a moderate head yaw of $50^\circ/s$, the maximum amount of latency allowed for single-pixel error can be determined by placing these values into equation 3.6:

$$0.05^\circ = \Delta t 50^\circ/sec \quad (3.7)$$

and solving for latency:

$$\Delta t = \frac{0.05^\circ}{50^\circ/sec} = 0.001 \text{ sec} \quad (3.8)$$

In this case, latency must be reduced to one millisecond to obtain single-pixel accuracy during head yaw. Faster head yaw cause more pixel error (i.e., scene displacement) for any given latency.

As discussed in Section 2.4.3.4, typical raster displays render at 60 Hz, switching to a new image each time the vertical sync signal is detected. This introduces an additional $\sim 15 \text{ ms}$ latency difference between the top scanline and bottom scanline of a frame (depending on the specific video timings). To reduce displacement error to near zero across the entire frame, traditional frame rendering at 60 Hz will not suffice for a raster display. Instead, pixels should be updated at a faster rate. This can be done by using a just-in-time scanlines algorithm, most easily achieved by not waiting to swap buffers on vertical sync (Section 2.4.3.4).

3.1.3 Scene Velocity

As a subject turns her head (assuming objects are at a distance) the scene should move with equal and opposite angular velocity in the HMD in order for the scene to have zero scene velocity (i.e., so that the image appears stable in world coordinates). For example, if the head turns to the right at $50^\circ/s$ in world coordinates, the scene should rotate to the left by $50^\circ/s$ in head-centric coordinates, with a new part of the scene becoming visible on the right side of the HMD screen.

Velocity error is displacement error differentiated with respect to time (Adelstein et al., 2005).

Scene velocity is the average velocity error in world space over the entire display. Since I assume a static virtual world, all scene velocity is considered error. Note, scene velocity can occur due to latency or scene velocity can be intentionally injected into the system for purposes of determining perceptual thresholds using psychophysics methods. As stated, I place a single 2D object at a distance in my experiments such that error due to head motion is nearly the same for all pixels. For purposes of this work, scene velocity is equal to velocity error for any single point.

Scene velocity θ' , in degrees per second, at time t_0 is

$$\theta'_{t_0} = \Delta\phi'_{t_0} = \phi'_{t_0} - \phi'_{t_0-\Delta t} \quad (3.9)$$

Note that scene displacement θ due to latency (Equation 3.1) can be greater than zero while scene velocity θ' due to latency (Equation 3.9) 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 is a constant offset distance from where it should be in the world.

Head velocity ϕ' at time t_0 can be written as a function of head acceleration ϕ'' :

$$\phi'_{t_0} = \int_{-\infty}^{t_0} \phi''(t)dt \quad (3.10)$$

so that Equation 3.9 can be written as

$$\theta'_{t_0} = \int_{-\infty}^{t_0} \phi''(t)dt - \int_{-\infty}^{t_0-\Delta t} \phi''(t)dt \quad (3.11)$$

By the addition property of equality, scene velocity θ' simplifies to integrating head acceleration ϕ'' over the latency interval:

$$\theta'_{t_0} = \int_{t_0-\Delta t}^{t_0} \phi''(t)dt \quad (3.12)$$

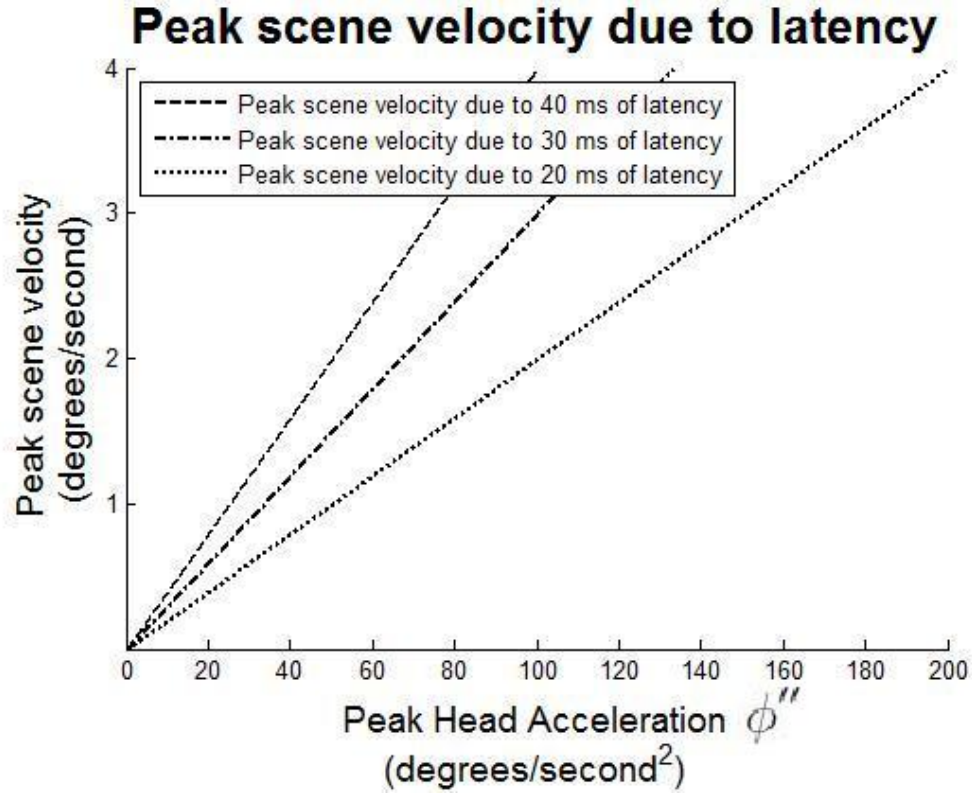


Figure 3.2: Peak scene velocity as a function of peak head acceleration is linear with slope equal to latency (Equation 3.15). The y-intercept equal to zero shows velocity error is zero regardless of latency when head acceleration is zero.

Given average head acceleration $\phi''_{average}$ over the latency interval, the equation simplifies to

$$\theta'_{t_0} = \Delta t \phi''_{average} \quad (3.13)$$

For small latency intervals Δt and/or near-constant head acceleration over the time of integration, scene velocity θ' can be approximated by

$$\theta'_{t_0} \approx \Delta t \phi''_{t_0} \quad (3.14)$$

Head acceleration is not normally constant. However, this formula is useful because it shows that scene velocity can be estimated by head acceleration multiplied by latency. By measuring peak head acceleration ϕ''_{peak} during a head turn, I can, for a small known latency, estimate (with some overestimation) the peak scene velocity θ'_{peak} that occurs during that head turn:

$$\theta'_{peak} \approx \Delta t \phi''_{peak} \quad (3.15)$$

Using equation 3.15, Figure 3.2 plots peak scene velocity θ'_{peak} as a function of peak head acceleration ϕ''_{peak} for three different latencies.

Peak scene velocity θ'_{peak} does not occur exactly in phase with peak head acceleration ϕ''_{peak} . Figure 3.3 shows Equation 3.12 geometrically. The areas underneath the head acceleration curves represent peak scene velocities. Shifting the integration intervals left or right decreases the area underneath the curve. Given increasing head acceleration then sudden decreased head acceleration, peak scene velocity θ'_{peak} occurs at time $t_{\phi''_{peak}}$ (Figure 3.3(a)). Given sudden head acceleration then decreasing head acceleration, peak scene velocity θ'_{peak} occurs at time $t_{\phi''_{peak}} + \Delta t$ (Figure 3.3(b)). For a peak head acceleration that increases then decreases symmetrically (a “bump” function), peak scene velocity θ'_{peak} occurs at time $t_{\phi''_{peak}} + \frac{1}{2}\Delta t$ (Figure 3.3(c)).

For small integration times Δt , Peak head-acceleration curves are typically shaped like Figure 3.3(c). On average, scene velocity peaks $\frac{1}{2}\Delta t$ after head acceleration peaks, i.e.,

$$t_{\theta'_{peak}} \approx t_{\phi''_{peak}} + \frac{1}{2}\Delta t \quad (3.16)$$

Rarely does scene velocity peak later than Δt after head acceleration peaks.

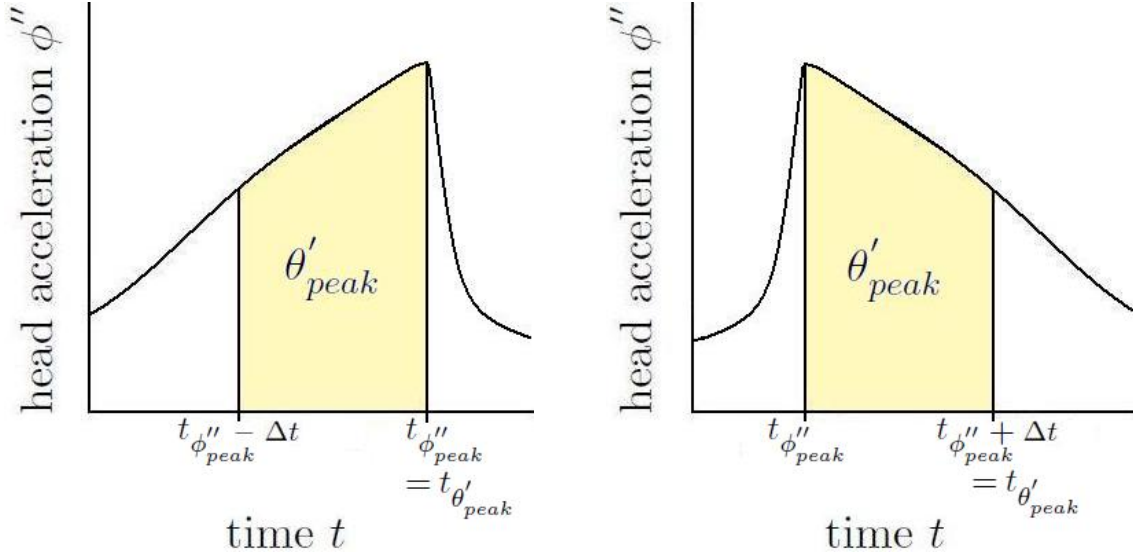
3.2 Real Head Motion

Figure 3.4 shows an actual head yaw of a subject and the same head yaw delayed by 100 milliseconds to demonstrate the effect of 100 milliseconds of latency.

The subject starts yawing her head from the left shortly before 2.0 seconds and stops yawing at approximately 3.0 seconds. The difference between the head-yaw angle and the delayed head-yaw angle is the scene displacement (Equation 3.1). Likewise, the difference between the head velocity and the delayed head velocity is the scene velocity (Equation 3.9).

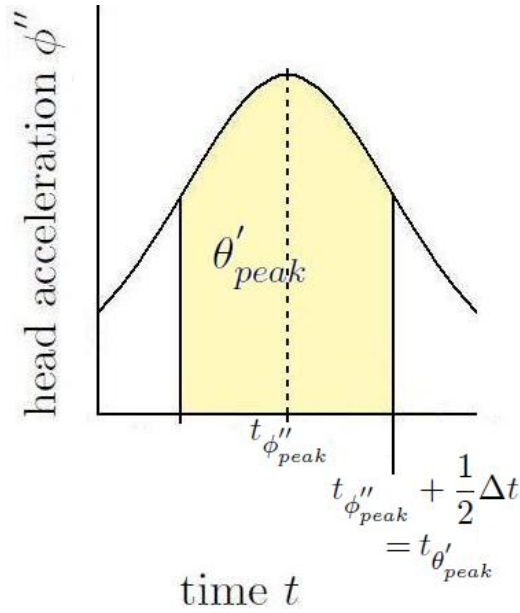
Notice that head velocity peaks during the center of the head yaw and that head acceleration peaks near the beginning and end of the head yaw. Scene displacement is consistent with Equation 3.6—scene displacement is a fraction of head velocity and both peak near the center of the head yaw. Scene velocity is consistent with Equations 3.15 and 3.16—scene velocity peaks approximately 50 *ms* after head acceleration peaks at the beginning and end of the head yaw.

Also notice that the scene moves in the same direction of the head yaw as the head accelerates. The scene displacement then nearly 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 yaw, towards its zero-error position.



(a) Given increasing head acceleration then sudden decreased head acceleration, maximum scene velocity θ'_{peak} occurs at time $t_{\phi''_{peak}}$.

(b) Given sudden head acceleration then decreasing head acceleration, maximum scene velocity θ'_{peak} occurs at time $t_{\phi''_{peak}} + \Delta t$.



(c) Given a symmetric head acceleration “bump” function, maximum scene velocity θ'_{peak} occurs at time $t_{\phi''_{peak}} + \frac{1}{2}\Delta t$

Figure 3.3: The shaded areas represent the maximum scene velocity θ'_{peak} that occurs over the integration time Δt (i.e., latency) and the given head acceleration. Shifting the integration interval left or right results in a smaller shaded area (i.e., less scene velocity). For typical head acceleration peaks, scene velocity peaks some time between 0 and Δt after head acceleration peaks.

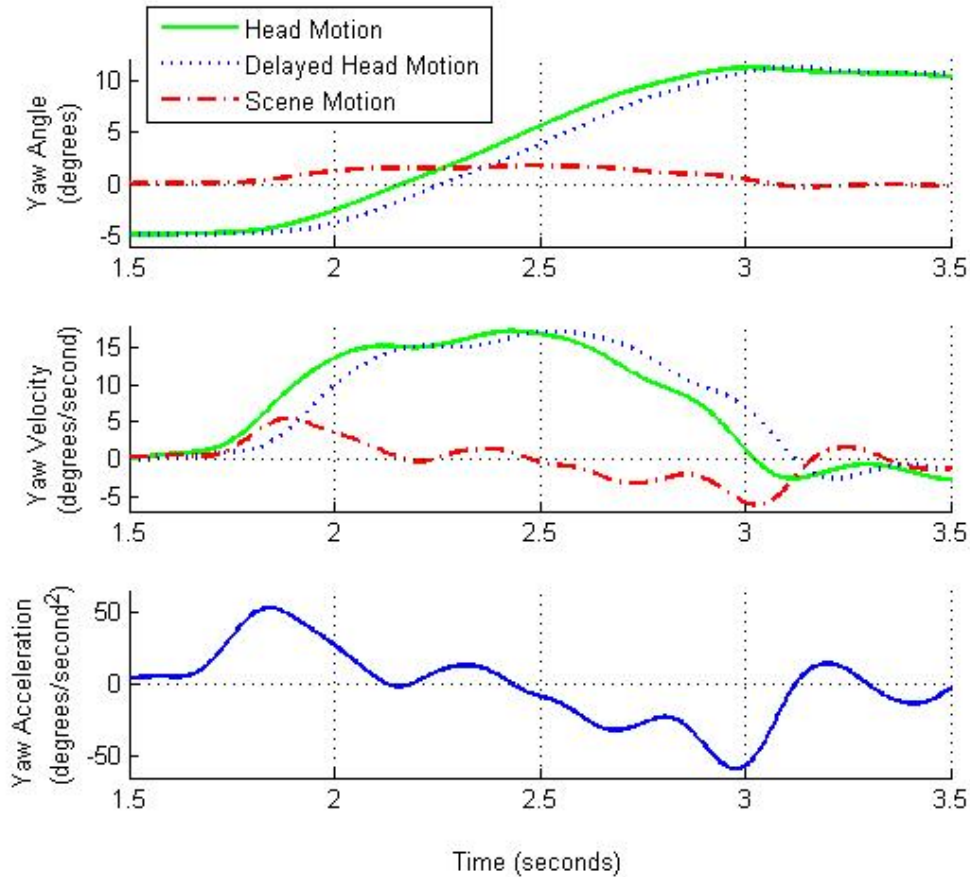


Figure 3.4: Measured head motion and computed scene motion due to 100 *ms* of latency. Top—displacement; middle—velocity; and bottom—acceleration.

Faster head turns often result in harder stops. This would result in a larger peak acceleration, and thus larger scene velocity, at the end of the head turn.

3.3 Threshold Assumptions

I made certain assumptions when designing my experiments. I claim my measured thresholds to be valid for HMDs given the following assumptions:

- The virtual world is predominately static. Users expect the scene to be fixed in world space.
- Objects are at a distance so that error due to head translation is insignificant (Sections 3.1.1, and 3.1.3).
- Latency thresholds for a simple scene are equivalent to latency thresholds for a more complicated scene (Section 2.4.2.5).

- Scene motion delayed by ~ 7.4 ms (the delay of my simulated HMD system, Section 4.1.1) is below human perception (Section 2.4.2.1) and does not affect my measured thresholds.

3.4 Relating Scene-Motion Thresholds to Latency Thresholds

The goal of this section is to determine a mathematical relationship between scene-motion thresholds and latency thresholds, both plotted against peak head acceleration.

3.4.1 Model Assumptions

I make the following assumptions for the mathematical model described in Section 3.4.2:

- Humans perceive low latencies by scene motion induced by latency, rather than latency directly (Section 2.4.2.1).
- Scene velocity is what the human-visual system detects when it detects scene motion (Section 2.2.7.1) and a perceptual threshold is the lowest amount of stimulus an observer will reliably detect. Therefore, I assume peak scene velocity is what lifts the sensation of scene motion over the threshold of consciousness and I measure scene-motion thresholds in terms of peak scene velocity.
- Equation 3.15 is a workable approximation relating peak scene velocity to latency and peak head-yaw acceleration. Furthermore, peak scene velocity is almost always nearly simultaneous (within the time period of the latency, Section 3.1.3 and Equation 3.16) with peak head-yaw acceleration.
- The head-yaw motion shown in Figure 3.4 is typical of head yaw that starts from a stopped position and ends in a stopped position—head acceleration peaks at the beginning and end of head yaw.
- Scene-motion thresholds, in degrees per second, are a linear function of peak head-yaw acceleration (Experiment 5).

3.4.2 The Model

Figure 3.5 repeats Figure 3.2, showing plots of latency-induced peak scene velocities as functions of peak head acceleration for three latency values (Equation 3.15). In addition, the thicker solid line is a plot of a hypothetical scene-motion threshold line. This line shows a

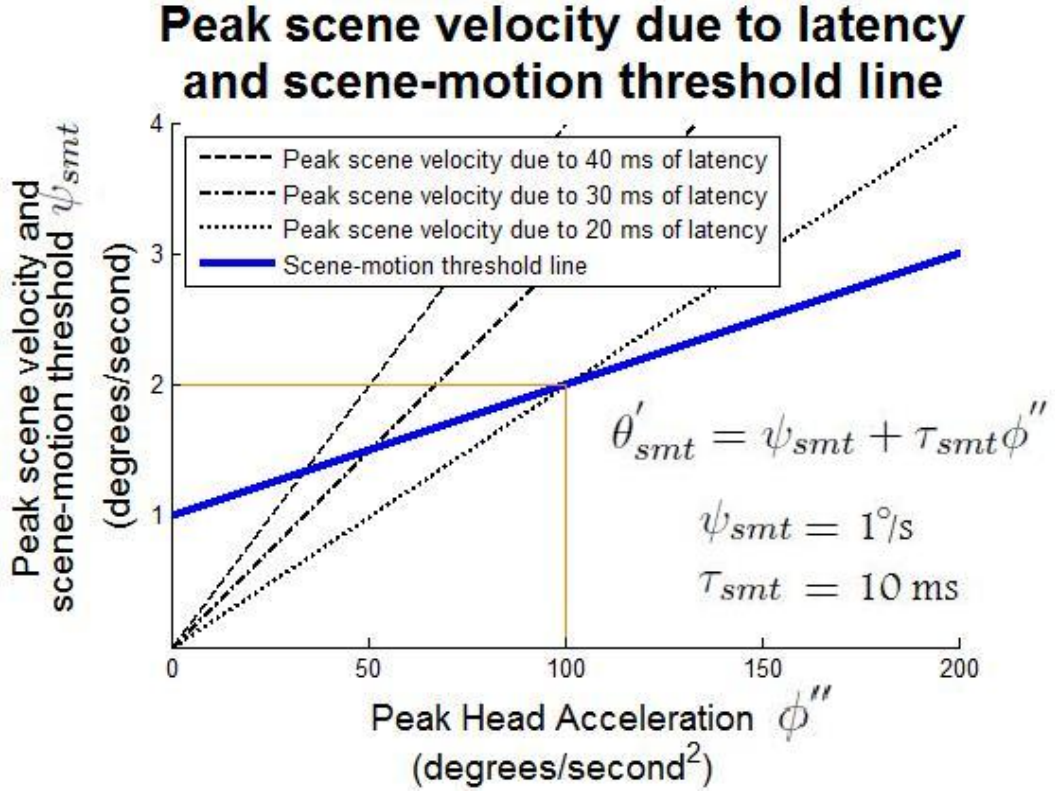


Figure 3.5: Peak scene velocities as a function of peak head acceleration and latency, and a hypothetical scene-motion threshold line similar to results found in Experiments 3-5. For the given hypothetical scene-velocity threshold line, the subject will theoretically detect 20 ms of latency when she rotates her head with a peak acceleration of $100^\circ/s^2$ or more.

linear relationship between scene-motion thresholds and peak head accelerations. For the first attempt at creating the model, I assumed a linear relationship and experimentation showed this to be a reasonable assumption (Experiments 3-5).

The linear equation for the scene-motion thresholds, θ'_{smt} , can be written as

$$\theta'_{smt} = \psi_{smt} + \tau_{smt} \phi'' \quad (3.17)$$

where ψ_{smt} is the y -intercept of the scene-motion threshold line, in degrees per second, τ_{smt} is the slope of the scene-motion threshold line, in seconds, and ϕ'' is peak head acceleration, in degrees per second squared.

Projecting the point at the intersection of the scene-motion threshold line and one of the latency-induced scene-velocity lines onto the x -axis gives the value of peak head acceleration at which a user would start to perceive latency-induced scene motion for that latency. The slope of the latency-induced scene-velocity line that intersects the scene-motion threshold line at a given peak head acceleration is the theoretical latency threshold for that peak head

acceleration. The task is then to find the slope of the latency-induced scene-velocity line (equivalent to latency) that intersects the scene-motion threshold line at a given peak head acceleration.

To determine the intersections of the lines and the latencies required to detect scene motion, the equation of the latency-induced scene-velocity lines (Equation 3.15) is set equal to the equation of the scene-motion thresholds line (Equation 3.17) and solved for Δt :

$$\theta'_{peak} = \theta'_{smt} \quad (3.18)$$

$$\Delta t \phi''_{peak} = \psi_{smt} + \tau_{smt} \phi'' \quad (3.19)$$

$$\Delta t = \tau_{smt} + \psi_{smt} \left(\frac{1}{\phi''_{peak}} \right) \quad (3.20)$$

According to this model, latency thresholds are inversely related to peak head acceleration where τ_{smt} is an offset parameter and ψ_{smt} is a scale parameter. Figure 3.6 shows the latency threshold function for the scene-motion threshold line shown in Figure 3.5. 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 low amounts of latency will be detected only with large head accelerations. As head acceleration goes to infinity, the latency thresholds approach an asymptote—the slope of the scene motion thresholds line, τ_{smt} . For the given hypothetical scene-motion threshold line, a subject with head acceleration of $100^\circ/sec^2$ will start to detect latency at 20 *ms*.

I conduct experiments to validate this model. If measured latency thresholds match the model well, then this provides evidence for the mathematical assumptions and/or the robustness of the mathematical assumptions stated in Section 3.4.1.

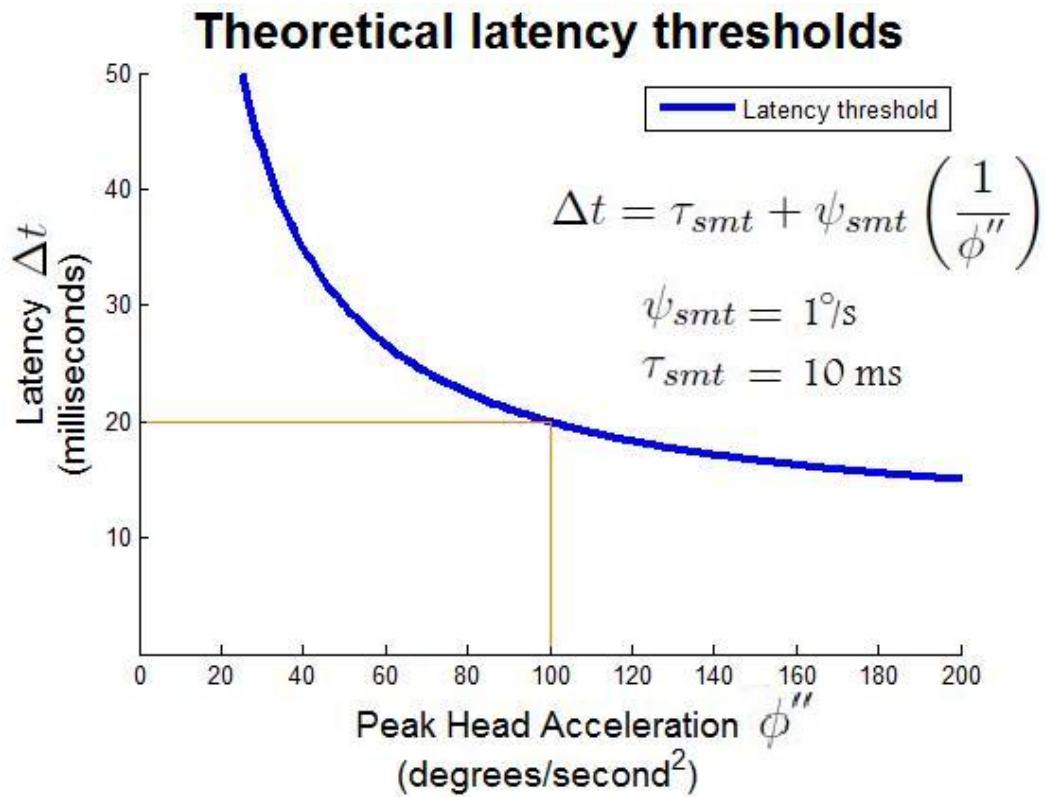


Figure 3.6: Theoretical latency thresholds determined from equation 3.20 for the scene-motion threshold line plotted in Figure 3.5. If the subject rotates her head with a peak head acceleration of $100^\circ/s^2$ then she will theoretically perceive scene motion due to a latency of 20 *ms* or more.

Chapter 4

Overview of Experiments

I conducted five experiments to better understand perception of scene motion and latency in IVEs. Experiments 1-3 focus on scene-motion thresholds whereas Experiments 4 and 5 focus on latency thresholds and their relationship to scene-motion thresholds. Experiments 1-3 are used to validate assumptions employed in Experiments 4 and 5.

I report experiments in logical order rather than temporal order. Experiments were conducted in the order of 1, 3, 4, 2, 5. Experiment 2 was conducted post hoc after Experiment 4 in order to further validate my assumptions for Experiment 4 and 5.

4.1 Materials and Methods

Each subject experienced hundreds of trials. Subjects were encouraged to take breaks at any time.

For each trial, subjects yawed their heads while a moving or non-moving scene was presented. At the conclusion of each trial, subjects made judgments about scene motion via a three-button mouse. No communication between the experimenter and the subject occurred during trials.

Experiment 1 was a mixed design (within and between subjects) where different subjects yawed their heads with different frequencies. Because Experiment 1 found that thresholds varied considerably among subjects, I thereafter used a repeated-measures design (all subjects experienced all conditions).

The specifics of the head motions, scene presentation times, scene motions, scene luminances, psychophysics methods, and analysis varied across experiments. Table 4.1 shows factors and findings of each experiment.

Table 4.1: Factors and Findings for all five experiments.

		Experiment					
		1 Quasi-Sinusoidal Head Yaw and Scene- Motion Direction	2 Single Head Yaw, Luminance, and Scene- Motion Direction	3 Increasing Head Motion	4 Validating the Latency- Thresholds Model	5 Two Case Studies	
Factor	# of subjects	9	9	8	8	2	
	Head Motion	head turn type	Quasi-Sinusoidal	Single	Single	Single	Single
		head-yaw frequency	0.35 Hz, 0.5 Hz, 0.65 Hz	0.5 Hz	0.5 Hz	0.5 Hz	0.5 Hz
		head-yaw range	±11°	±15°	±5°, ±10°, ±15°, ±20°	±0.5° to ±22.5°	±0.5° to ±22.5°
		head-motion measure	Theoretical peak velocity (deg/s)		Range (deg/s) Peak Velocity (deg/s) Peak Acceleration (deg/s ²)	Peak Acceleration (deg/s ²)	Peak Acceleration (deg/s ²)
	Scene Motion Visibility	head-yaw phase (time visible)	Center (0.5 s) Edge (0.5 s)	Center (0.6 s) Start (0.6 s) End (0.6 s)	Center (0.6 s) Start (0.6 s) End (0.6 s) All (1.2 s)	End (0.7 s)	End (0.7 s)
		scene-motion type	Constant	Constant	Constant	Constant Gaussian Latency	Constant Gaussian Latency
		scene-motion direction	With Against	With Against	Against	Against	Against
		scene-motion measure	Constant Velocity (deg/s)	Constant Velocity (deg/s)	Constant Velocity (deg/s)	Peak Velocity (deg/s) Latency (s)	Peak Velocity (deg/s) Latency (s)
	Brightness	foreground luminance	Between Bright and Dim Conditions of Exp 2	Bright: 10.0 cd/m ² Dim: 0.11 cd/m ²	same as Exp 1	same as Exp 1	1.2 cd/m ²
		background luminance	Between Bright and Dim Conditions of Exp 2	Bright: 0.06 cd/m ² Dim: 0.003 cd/m ²	same as Exp 1	same as Exp 1	0.008 cd/m ²
		contrast		Bright: 167 Dim: 37			150
	Methods	Stimulus Selection	Adaptive staircase	Adaptive staircase	Method of constant stimuli	Method of constant stimuli	Method of constant stimuli
		Judgment Task	3I-2AFC	Yes/no with confidence ratings	Yes/no with confidence ratings	Yes/no with confidence ratings	Yes/no with confidence ratings
		Threshold	75%	75%	75%	75% 50% Difference	75% 50% Difference
	Findings	Hypothesis 1	Scene-motion thresholds are greatest during the reversal of quasi-sinusoidal head yaw	Scene-motion thresholds are greater for the With Condition than the Against Condition for each head-yaw phase and for each brightness.	Scene-motion thresholds increase as head-yaw range, given constant time, (i.e., average head-yaw velocity) increases.	Directly-measured latency thresholds correlate with the modeled inverse function better than with a linear function.	
Hypothesis 2		Scene-motion thresholds are greater when the scene moves with head yaw than when the scene moves against head yaw.	Scene-motion thresholds are lower than those for the Dim Condition for all head-yaw phases.	Scene-motion thresholds increase as peak head-yaw velocity increases.	The Constant Condition, Gaussian Condition, and Latency Condition result in different scene-motion thresholds.		
Hypothesis 3				Scene-motion thresholds increase as peak head acceleration increases.			
Results		Hypothesis 2 confirmed	Hypothesis 1 confirmed	Hypothesis 1, 2, and 3 confirmed	Hypothesis 1 confirmed, Hypothesis 2 partially confirmed	Consistent with previous experiments	

4.1.1 The System

I used a minimum virtual environment that consisted of only head tracking, an HMD simulated by a projector (1024×768 pixels), and a virtual scene consisting of a simple 2D object 4 meters in front of the observer. Using such a minimal system, as is often done in psychophysical experiments, enabled me to control variables precisely.

A 3rdTech HiBallTM 3000 provided head pose data at ~ 1500 Hz. Head pose was used to control auditory cues, to check for acceptable head rotations, and to record motion for post-analysis. Head pose did not affect the position or motion of the scene for Experiments 1-3. Head pose did not affect the position or motion of the scene for all but one condition of Experiments 4 and 5; head pose did affect the one condition.

Experiments 1-3 waited to swap buffers on vertical sync (Section 2.4.3.4) whereas Experiments 4 and 5 did not, resulting in a just-in-time scanline implementation (Section 2.4.3.4) at 1500 Hz (such that head pose was updated every few scanlines).

I measured the end-to-end system delay (from the time of tracking to the time the visuals were displayed) of Experiments 4 and 5 using the technique described by Steed (2008) I measured and analyzed 10 samples of system delay resulting in a mean system delay of 7.4 *ms*.

Even though some system delay existed in the system, I designed all experiments to emulate a zero-latency HMD using a world-fixed display; no scene motion occurred due to system delay. This was possible because scene motion is independent of head motion for a world-fixed display. Scene motion and/or simulated latency was then artificially injected in order to measure scene-motion thresholds and latency thresholds.

A BARCO CRT projector displayed a simple scene onto a world-fixed planar surface 4 meters in front of the seated subject (Figure 4.1). The CRT projector was chosen for its fast phosphor response (so that display time was minimized) and decay times (such that no ghosting occurred), and because 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 (Figure 4.2). The cutouts limited the user's field-of-view to $48^\circ \times 36^\circ$, as is the case for an unmodified V8 HMD. Total weight of the modified HMD and tracker was ~ 0.6 kg.

All object-relative cues were removed by darkening the room and providing a uniform visual field. Since only the computer-generated scene was visible, subjects could make only subject-relative judgments. Subjects confirmed they could see no visual elements other than the computer-generated scene presented by the projector.

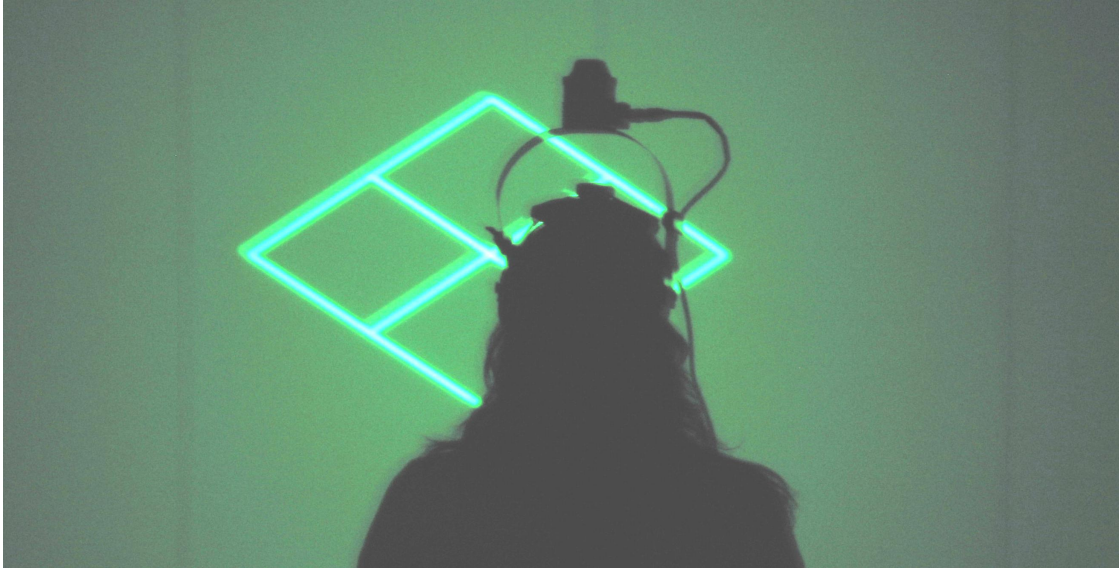


Figure 4.1: A subject and scene as the subject yaws her head. Note the motion blur occurs from the exposure time of the camera and the image has been brightened substantially.

4.1.2 Head Motion

I asked subjects to yaw their heads to the pacing of metronome beeps. I trained subjects using both visual and auditory cues; the experiment room was otherwise kept quiet. A tone sounded when head yaw exceeded a minimum head-yaw amplitude. Subjects were trained to move their heads far enough to hear this tone at the reversal of head yaw. If subjects yawned their heads beyond a maximum head-yaw amplitude, a buzzing sound occurred.

Subjects optionally practiced head yaw with visual cues before each trial, and then pressed a button to start the trial. The visual cues then disappeared, and the subject yawned her head in time to the auditory cues as had been practiced. If head motion deviated from the prescribed minimum and maximum range, the trial was canceled and repeated at a later time.

For Experiment 1, subjects yawned their heads with quasi-sinusoidal motions over an intended head-yaw amplitude of $\pm 11^\circ$ from straight ahead for one of three different head-yaw frequencies. For Experiments 2-5, subjects yawned their heads a single time left-to-right or right-to-left for each trial over a period of one second. For Experiment 2, the intended head-yaw amplitude was $\pm 15^\circ$ from straight ahead. For Experiment 3, the intended head-yaw amplitude was randomly selected to be $\pm 5^\circ$, $\pm 10^\circ$, $\pm 15^\circ$, or $\pm 20^\circ$ from straight ahead. For Experiment 4 and 5, the intended head-yaw amplitude was randomly selected on each trial to be between $\pm 0.5^\circ$ to $\pm 22.5^\circ$ from straight ahead.



Figure 4.2: The modified V8 HMD ,with display elements removed so that subjects could see through the casing to the world-fixed display. Note the Hiball tracker is not shown in this picture but was mounted on top of the device during experimentation.

4.1.3 The Scene

A simple 2D visual scene (a rotated green monochrome square with diagonals and a 20° horizontal span, as shown in Figure 4.1) was chosen for the following reasons:

- A 2D scene colocated with the display surface was chosen to eliminate scene motion due to incorrect viewing position (e.g., incorrect head-tracking, incorrect interocular distance for stereo display, misregistration of 3D geometry, etc.) that could confound the results (Section 2.1.1).
- A simple scene minimizes the rendering time component of the unavoidable end-to-end system delay.
- Prior work has not found statistically significant differences in latency thresholds across scene complexities (Section 2.4.2.5).
- Tearing effects, due to the system not waiting to swap buffers on vertical sync (Section 2.4.3.4), are more noticeable in vertical lines than diagonal lines. Experiments 1-3 waited on vertical sync to swap buffers, whereas Experiments 4 and 5 did not.

The scene was presented during some phase of the head yaw and either moved or did not move.

4.1.3.1 Scene Visibility

The scene was presented for different phases of head yaw. When the scene was not presented, no visual cues were visible.

For Experiment 1, the scene was presented during the center of head yaw (during the time that head-yaw velocity peaks) or during the reversal of head yaw (during the time that head yaw direction changes and head acceleration peaks). For each trial the scene was presented three times for a period of 0.5 seconds for each presentation.

For Experiment 2, the scene was presented a single time, over a period of 0.6 second, during the start of the head yaw, the center of the head yaw, or the end of the head yaw.

For Experiment 3, the scene was presented a single time during the start of the head yaw (a maximum of 0.6 second), the center of the head yaw (a maximum of 0.6 second), the end of the head yaw (a maximum of 0.6 second), or for the entire head yaw (a maximum of 1.2 seconds). The system turned off the scene in this experiment when head-yaw velocity fell below a threshold value so that subjects would not judge motion when their heads were not moving

For Experiments 4 and 5, the scene was presented only during the end of head yaw, over a period of 0.7 seconds. This was done because the results of Experiments 1 and 2 suggested that latency thresholds would be the lowest during the end of head yaw. The scene was visible for the entire period of 0.7 seconds, even if the head was not moving.

4.1.3.2 Scene Position

I varied only the horizontal position of the scene; vertical position was constant. Although the scene was presented on a planar surface, I performed all analysis in degrees.

The angular offset of the scene from the center of the display is

$$\theta = \arctan\left(\frac{x}{d}\right) \quad (4.1)$$

where x , in meters, is the position of the center of the scene and d , in meters, is the distance from the subject's head to the display.

In analyzing Experiments 1-3, I approximated the distance d to be four meters. d was required to be constant for Experiments 1 and 2, since the adaptive staircase design used in those Experiments (4.1.4.1) requires that scene-motion levels be set to constant values. In hindsight, d could have been measured at the start of each trial in Experiment 3.

For Experiments 4 and 5, I measured d at the start of every trial. Since subjects did not always position their head at exactly 4 meters from the display, by up to ± 0.4 meters in some cases, measuring d improved the accuracy of θ calculations by up to $\sim 1^\circ$ and scene-motion by up to $\sim 1^\circ/s$ (for fast scene motions of $10^\circ/s$, which is greater than measured scene-motion

thresholds for all subjects).

For Experiments 1 and 3, the system randomly set the starting position of the scene to be $\pm 3.2^\circ$ of center.

For Experiments 2, 4, and 5, the system set the starting position of the scene to be half-way between 0° and the end of the suggested head yaw (when the scene was visible for the second half of the head yaw), so that the scene would be presented approximately in the center of the field of view.

4.1.3.3 Scene Motion

The stimulus to be detected for all experiments was scene motion (whether induced by latency or not). Some scenes moved and some scenes did not move. For each trial, scene velocity or latency varied between trials.

For Experiments 1-3, scene motion was independent of head motion and scene velocity was constant within trials. Scene velocity varied between trials.

For Experiment 1, when the scene was presented during the center of head yaw, the scene moved with or against the direction of head yaw. When the scene was presented during the reversal of head yaw, the scene moved both with and against the direction of head yaw since scene motion did not change direction but head motion did change direction.

For Experiment 2, the scene moved with or against the direction of head yaw.

Experiments 1 and 2 demonstrated thresholds to be lowest when the scene moved against the direction of head yaw. For Experiment 3, the scene moved against the direction of head yaw.

For Experiments 4 and 5, the scene moved in one of three ways within trials: (1) constant scene velocity against the direction of head yaw, (2) a scene motion modeled after that which typically occurs due to latency (against the direction of head yaw), and (3) a latency-induced scene motion that moved as a function of actual head motion. For the latency-induced scene motion, the scene mostly moved against the direction of head yaw but occasionally moved with the direction of head yaw for brief periods of time depending on the specific motion of the subject's head.

4.1.3.4 Scene Luminance and Contrast

Experiment 2 contained scenes with two luminance (contrast) conditions. The scene foreground (the rotated green monochrome square with diagonals) of the Bright Condition had a luminance of $10.0\text{cd}/\text{m}^2$ whereas the scene background had a luminance of $0.06\text{cd}/\text{m}^2$. The scene foreground of the Dim Condition had a luminance of $0.11\text{cd}/\text{m}^2$ whereas the scene background had a luminance of $0.003\text{cd}/\text{m}^2$. These luminances resulted in a contrast of 167

for the Bright Condition and a contrast of 37 for the Dim Condition.

Experiment 5 had a foreground luminance of $1.2cd/m^2$ and a background luminance of $0.008cd/m^2$. The contrast was 150.

The scene luminances for Experiments 1, 3, and 4 were somewhere between the luminances of the Dim and Bright Conditions of Experiment 2. The luminances were not measured for these experiments.

4.1.3.5 The Reference Scene

For all experiments, a green reference scene was always visible between trials to prevent dark adaptation, so that brightness sensitivity would be consistent across trials. The reference scene also served to provide subjects with a non-moving scene that allowed them to better judge scene motions presented during the trials.

For Experiments 1, 3, and 4, the between-trial non-moving reference scene consisted of a background and instructions. In hindsight, I should have included the 2D object between trials in these experiments so that subjects could have a stable reference scene that more closely resembled the trial scene in order to better judge scene-motion. For Experiments 2 and 5, the reference scene contained the same 2D object presented during trials.

4.1.4 Psychophysics Methods

At the conclusion of each trial, subjects made judgments about scene motion via a three-button mouse. No communication between the experimenter and the subject occurred during trials.

4.1.4.1 Adaptive Staircases Versus the Method of Constant Stimuli

An adaptive staircases algorithm allows stimulus levels to quickly converge to a point of interest along a single dimension (Section 2.3.3.4). For example, the experimenter may want to collect most of the data near a subject's 75% scene-motion threshold.

Adaptive staircases (Section 2.3.3.4) require all variables to be known before starting a trial in order to select the stimulus level. Since all variables could be precisely controlled for Experiments 1 and 2, I used an adaptive staircase for selecting scene velocities for individual trials.

Since head motion was an uncontrolled and measured variable for Experiments 3-5, I was not able to use adaptive staircases for those experiments. I used the *method of constant stimuli* (stimulus levels are randomly selected independent of previous results, Section 2.3.3.2) for selecting the stimulus level (the amount of scene motion or latency) for individual trials. The range of stimulus levels for each participant was chosen based on judgments and confidence

ratings given during training. The stimulus range was adjusted until users gave approximately the same number of “yes” responses as “no” responses. As the experiment progressed, I increased or decreased the maximum scene-velocity value for individual subjects as necessary to generate the spread of data needed for analysis. Thus, in some respect, Experiments 4 and 5 used an adaptive “algorithm” controlled by myself rather than by the system.

4.1.4.2 3-Interval 2-Alternative Forced Choice

The trial task in Experiment 1 was a three-interval two-alternative forced-choice (3I-2AFC) identification task (Section 2.3.4.3). Three presentations were provided and subjects were forced to choose which of the latter two presentations contained scene motion. The first presentation was a reference scene containing no scene motion so that subjects knew what a stable scene looks like. Some scene motion was randomly assigned to either the second or third presentation. The subject then selected which of the latter two presentations she believed was different—i.e., the presentation that contained scene motion.

Experiment 1 had correct answers—only one of the three presentations contained scene motion and subjects attempted to choose the presentation with scene motion. After each response the system informed the subject whether she was correct. I rewarded subjects \$0.05 for every correct response to encourage good performance.

4.1.4.3 Yes/No Judgments

Time per trial in Experiment 1 was more than three times longer than a single presentation due to three presentations per trial and many trials being canceled for incorrect head motions. The head motions were difficult to make consistently and subjects required extensive training to achieve three valid head motions in a row.

To collect data more efficiently for Experiments 2-5, I used a yes/no judgment task (Section 2.3.4.1). A trial consisted of a single presentation of the scene, and a judgment whether the scene contained motion or not. Such a design contains more bias than an AFC task, but bias was kept as constant as possible within subjects by randomly intermixing conditions across trials. A reference scene was shown between trials so that subjects had a stable scene to compare the trial scene to.

The wording of the yes/no 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*

Yes/no judgment tasks do not have “correct” responses. If I had rewarded subjects for every correct response then they could have optimized their earnings by always selecting “yes”. Penalties could have been imposed for incorrect positive judgments. However, balancing rewards and penalties changes the subjective decision process of subjects (analyzed by signal detection theory (Macmillan and Creelman, 2005)), which greatly complicates the problem. Hence, I paid subjects a constant rate for Experiments 2-5. I paid \$7 per hour for Experiments 2-4. I paid \$10 per hour for Experiment 5, which required multiple sessions.

4.1.4.4 Confidence Ratings

For Experiments 2-5, I collected confidence ratings in addition to yes/no judgments. After subjects judged whether the scene seemed to move or not move, they rated their 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 subjects selected the “scene seemed to not move” with the word “not” inserted into the rating options.

Experiment 5 explored how using confidence ratings affected thresholds.

4.2 Measures

4.2.1 Actual Head Motion

I analyzed head motion for Experiments 3-5. Head motion was a measured variable (via the 3rdTech HiBall™ tracking system) since it could not be precisely controlled.

For Experiment 3, I measured head motion during the time that the scene was visible for each trial using the following measures:

- the range of head yaw (degrees),

- the peak head-yaw velocity (degrees per second), and
- the peak head-yaw acceleration (degrees per second squared).

For Experiments 4 and 5, I used peak head-yaw acceleration as the measure of head motion.

4.2.1.1 Head-Motion Quantiles

I divided each head-motion measures into bins with each bin containing a range of head motions. I set each bin to have an equal number of samples (quantiles). I chose to use quantiles instead of a uniform equal range of head motion for each bin in order to allow for the calculation of psychometric functions. If I used a uniform range for each bin, then some bins would only have a few samples—not enough samples to compute a psychometric function.

For Experiment 3, I divided each measures of head motion into sextiles (six bins). For Experiments 4 and 5, I divided the measured peak head-yaw accelerations into deciles (ten bins). I chose the quantile size (i.e., the number of bins) to ensure enough samples to calculate reasonable psychometric functions.

4.2.2 Scene Motion

For all experiments, I used peak scene velocity (because my model demands it (Section 3.4.1 and the human-visual system directly detects scene velocity (Section (Section 2.2.7.1)), in degrees per second, as a measure of scene motion.

For Experiments 1-3, I was able to precisely control scene motion and I used predetermined constant scene-velocities (such that constant scene velocity was equivalent to peak scene velocity) within trials. I conducted conducted multiple trials for each level of scene velocity. This allowed me to compute average responses for different levels of scene velocity.

For Experiments 4 and 5, scene motions could not be precisely controlled and varied on a continuum. Although, averages of multiple responses could not be computed for individual levels of peak scene velocities, psychometric functions could still be fit to responses on the peak scene-velocity continuum.

4.2.3 Latency

For one of the conditions of Experiments 4 and 5, I randomly selected latency (constant within trials), in seconds, on a continuum. I chose to select latency on a continuum instead of from discrete levels of latency, so that the analysis for latency would be similar to analysis for scene motion due to latency and head motion (which could not be precisely controlled).

4.3 Data Analysis

Regardless of the specifics of scene motion and whether confidence ratings were used, the analysis method used to determine psychometric functions was identical for Experiments 2-5; only the input changed. Experiment 1 used slightly different analysis methods, due to a different experiment design.

4.3.1 Confidence Ratings

For Experiments 2-5, I constructed a six-level rating scale in order to have a larger amount of data for analysis than the binary yes/no judgments. Three ratings for “yes” responses and three ratings for “no” responses resulted in six possible levels. I quantified these responses as 0% (no, with full certainty), 20% (no, with some certainty), 40% (no, guessed), 60% (yes, guessed), 80% (yes, with some certainty), and 100% (yes, with full certainty) and I considered the quantifications to be ratio values.

4.3.2 Psychometric Functions

For all experiments, I fit cumulative-Gaussian psychometric functions to the data collected from each of the subjects, conditions, head-motion measures, and/or head-motion quantiles.

A psychometric function (Section 2.3.1) normally relates detection probabilities, determined from binary responses, to stimulus intensities. However, the “psychometric functions” in Experiments 2-5 are based on subjects’ confidence ratings in addition to the binary yes/no judgments. Experiment 5 explored how ratings affected the psychometric functions.

4.3.3 Thresholds

For all experiments, I extracted 75% thresholds from each of the psychometric functions. For Experiments 4 and 5, I broke the 75% thresholds into 50% thresholds and difference thresholds (75% thresholds - 50% thresholds) in order to compare my results with the results of previous experiments performed at NASA Ames Research Center (Ellis et al., 1999; Adelstein et al., 2003; Ellis et al., 2004; Mania et al., 2004). This comparison is discussed in Section 10.1.2.

I chose to use the terms “50% threshold” instead of the point of subjective equality (PSE, Section 2.3.2.1) and “difference threshold” instead of the just-noticeable difference (JND, Section 2.3.2.2) because my thresholds were obtained from yes no judgments and confidence ratings. Since I collected a large amount of data for Experiment 5, I was also able to compute thresholds with yes/no judgments alone. For yes/no judgments alone, my 50% threshold are equivalent to PSEs and my difference thresholds are equivalent to JNDs.

4.3.4 Tracker Jitter

For Experiments 4 and 5, I removed trials where subjects reported that the scene appeared to jitter due to no tracker problems. This jitter was due to lack of tracker updates and occasionally appeared as instantaneous movement of the scene one or more times during a single trial of the Latency Condition. The problems were only apparent to subjects for the Latency Condition, where scene motion depended upon head motion. In addition, I also removed trials (for all conditions) where the tracker did not update for $10ms$ when the scene was visible and the head yawed more than 0.25° during the dead interval(s).

4.3.5 Statistics

Parametric tests assume normally distributed samples and homogeneity of variance, or large sample sizes. Since I could not assume this was the case for my data and most of the samples sizes were small, I used nonparametric tests unless otherwise noted.

I used two-tailed tests unless otherwise noted. I used one-tailed tests only when I expected one of the conditions to be greater than the condition to which it was being compared.

I set $\alpha = 0.05$. When performing multiple tests, I use Bonferroni Correction when one or more of the tests resulted in $p > 0.05$. The intention of correcting for multiple tests is for the case when many tests result in $p > 0.05$. No correction is required if the largest p-value is less than the uncorrected α (Hochberg, 1988).

Chapter 5

Experiment 1: Quasi-Sinusoidal Head Yaw and Scene-Motion Direction

The goal of this experiment was to find if scene-motion thresholds depend upon head-motion phase and scene-motion direction relative to head motion direction. Subjects yawed their heads in a quasi-sinusoidal manner for four cycles while two non-moving moving scenes and one moving scene was presented to them. They then chose which of the presented scenes they thought moved. From these judgments, I measured and compared scene-motion thresholds for the following cases:

- The reversal phase of quasi-sinusoidal head yaw, when the direction of head motion changes.
- The center phase of quasi-sinusoidal head yaw, when head velocity peaks and head velocity is approximately constant. Two scene-motion directions were measured for the center phase:
 - The scene moves with the direction of head yaw.
 - The scene moves against the direction of head yaw.

5.1 Hypotheses

My initial hypothesis was:

- **Hypothesis 1:** Scene-motion thresholds are greatest during the reversal of quasi-sinusoidal head yaw, when head-yaw direction changes.

During pilot studies, I noticed a trend that for the center-phase of head yaw (when head velocity was near constant), scene-motion thresholds seemed to depend upon the direction the scene was moving relative to the head. Due to this trend, I added the following hypothesis:

- **Hypothesis 2:** Scene-motion thresholds are greater when the scene moves with the direction of head yaw than when the scene moves against the direction of head yaw.

5.2 Experimental Design

I measured scene-motion thresholds for different conditions. The experiment was a mixed (within-subject and between-subject) adaptive staircase design. Three head-yaw frequencies were controlled between subjects, two head-yaw phases varied within subjects, and two scene-motion directions varied within subjects. Each trial consisted of a three-interval two-alternative forced-choice (3I-2AFC) identification task (Section 2.3.4.3). Three scenes were presented as subjects yawed their heads in a quasi-sinusoidal manner. One of the scenes moved, and at the end of the trial subjects selected which scene they believe moved.

I asked subjects to keep their eyes centered at the center of the stimulus figure so that the scene was seen mostly in the foveal region of the eye. However, the scene appeared randomly within $\pm 3.2^\circ$ of center so that subjects could not precisely predict where the scene would appear and I did not measure eye movements. Thus, I do not know how well subjects kept their eyes centered on the scene.

5.2.1 The Stimulus

The stimulus to be detected was scene motion. For each trial, scene velocity was constant for the one presentations that contained scene motion. Scene velocity varied between trials.

5.2.2 Controlled Variables

Head-Yaw Frequency

To generalize results across head motions, I controlled head-yaw frequency between subjects with three subjects per head-yaw frequency (with the data analyzed within subjects for all nine subjects). Subjects yawed their heads at 0.35 Hz, 0.5 Hz, or 0.65 Hz corresponding to side-to-side head swings in 1.43 seconds, 1.0 seconds, or 0.77 seconds. The top element of Figure 5.1 shows a specified head yaw at 0.5 Hz and a single subjects' actual head yaw over time for several trials. If head motion deviated from that prescribed, the trial was discarded and repeated later.

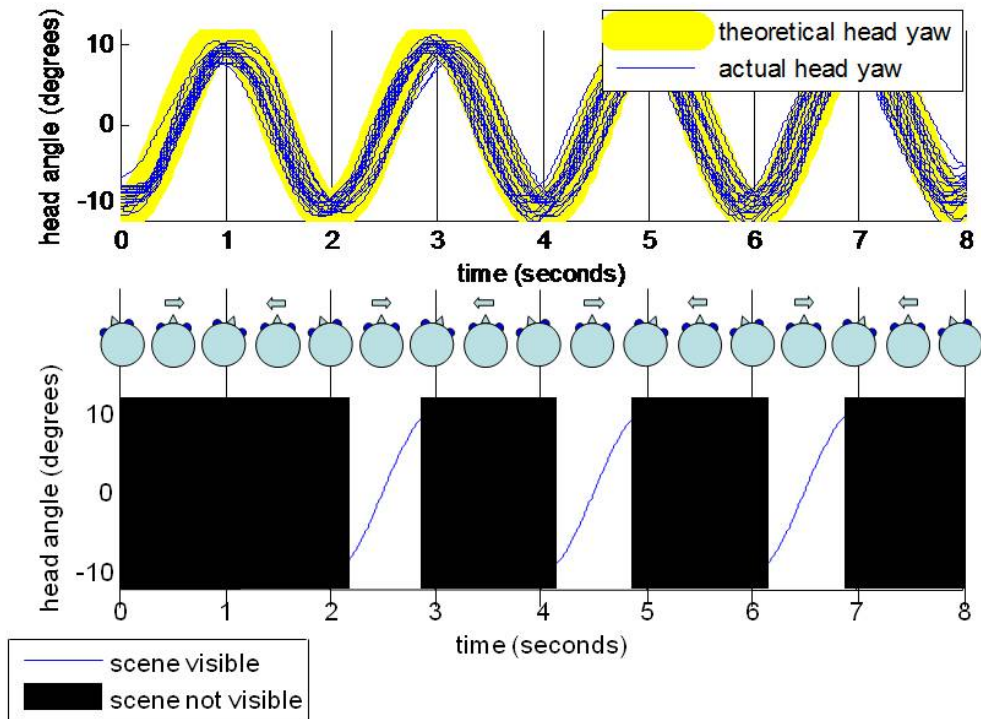


Figure 5.1: Head yaw and scene visibility. The top part of the diagram shows specified and actual head yaw with a head frequency of 0.5 Hz and a head-yaw amplitude of $\pm 11^\circ$ from straight ahead. The head icons show head aim and the direction the head is yawing. The bottom part of the diagram shows that the scene is visible during the center-phase of head yaw. The scene is presented three times for 0.5 seconds each as the head yaws from the left to the right. The scene is not visible for the black area of the diagram.

Head-Yaw Amplitude

Visual cues before the trials, and auditory cues before and during the trials suggested subjects yaw their heads by $\pm 11^\circ$ from straight ahead.

5.2.3 Independent Variables

Two variables were manipulated—head-yaw phase and scene-motion direction—resulting in three conditions.

Head-Yaw Phase

The scene was visible during the *reversal*-phase of head-yaw (when head motion changes direction) or during the *center*-phase of head yaw (when head velocity peaks and is nearly

Table 5.1: Experiment 1 conditions.

		Condition (Within Subjects)		
		With	Against	Reversal
Head Frequency (Between Subjects)	0.35 Hz	Subjects 1-3	Subjects 1-3	Subjects 1-3
	0.50 Hz	Subjects 4-6	Subjects 4-6	Subjects 4-6
	0.65 Hz	Subjects 7-9	Subjects 7-9	Subjects 7-9

constant). The scene was visible for 0.5 seconds for each of the three presentations per trial. The bottom element of Figure 5.1 shows a scene visible for the center-phase of head yaw.

Scene Motion

For the presentation that moved in each trial, the scene moved at a constant velocity *with* or *against* the direction of the head yaw for the center head phase. The scene velocity was determined by an adaptive staircase design at the start of each trial.

Conditions

The conditions were:

- **With Condition**—The scene moved in the same direction of the head yaw and was presented only during the center phase of quasi-sinusoidal head yaw.
- **Against Condition**—The scene moved against the direction of head yaw and was presented only during the center phase of quasi-sinusoidal head yaw.
- **Reversal Condition**—The scene moved left or right and was presented only during the extremes of quasi-sinusoidal head yaw, during the time that head direction changed, i.e., the scene moved part of the time with the direction of the head yaw and part of the time against the direction of the head yaw.

The bottom part of Figure 5.1 shows an example trial of the With Condition (assuming the scene moves from the left to the right). After a stationary first presentation, the scene moved to the right (with the direction of the head yaw) for the second or third presentation, but not for both presentations. At the end of the trial, the subject selected the second or third presentation that she believed contained scene motion.

5.2.4 Dependent Variables

Scene-Motion Threshold

I defined the scene-motion threshold to be the scene velocity (degrees per second) at which the subject correctly chose the presentation that contained scene motion 75% of the time. The 75% threshold is the halfway point between random guessing at 50% correct detection (guessing which of the two presentations contained scene motion) and 100% detection. Assuming subjects would choose correctly 50% of the time if all presentations contained the same stimulus, then the 75% threshold is also the just-noticeable difference (JND).

I measured scene-motion thresholds for the three conditions and the three head-yaw frequencies shown in Table 5.1.

5.3 Methods

Trial procedures used a one-up-two-down adaptive staircases algorithm (Section 2.3.3.4) to determine perceptual thresholds.

Each subject experienced six sessions over multiple days. Each session consisted of three randomly interleaved staircases, with one staircase for each condition. This interleaving of staircases minimized order effects, and made it difficult to differentiate among the conditions. Each staircases started with a scene motion of $7.1^\circ/s$ and terminated after eight staircase reversals, resulting in each subject judging a total of 148 to 219 trials for each of the three conditions. The staircase step size started at $3.6^\circ/s$ and was halved at every reversal until a minimum step size of $0.22^\circ/s$ was reached.

Each trial consisted of a three-interval two-alternative forced choice (3I-2AFC) identification task (Section 2.3.4.3). Three presentations were provided with one presentation per head-yaw cycle. The first presentation was a reference scene containing no scene motion so that subjects knew what a stable scene looked like. Some scene motion was randomly assigned to either the second or third presentation. Subjects then selected which of the latter two presentations they believed contained scene motion (i.e., which of the latter two presentations was different from the first presentation?). After each response the system informed subjects whether they were correct. To encourage good performance, I rewarded subjects \$0.05 for every correct response.

5.4 Participants

Nine subjects (age 18-44, 7 male and 2 female) participated. I served as one of the subjects; all other subjects were naive to the experimental conditions. Subjects were encouraged to

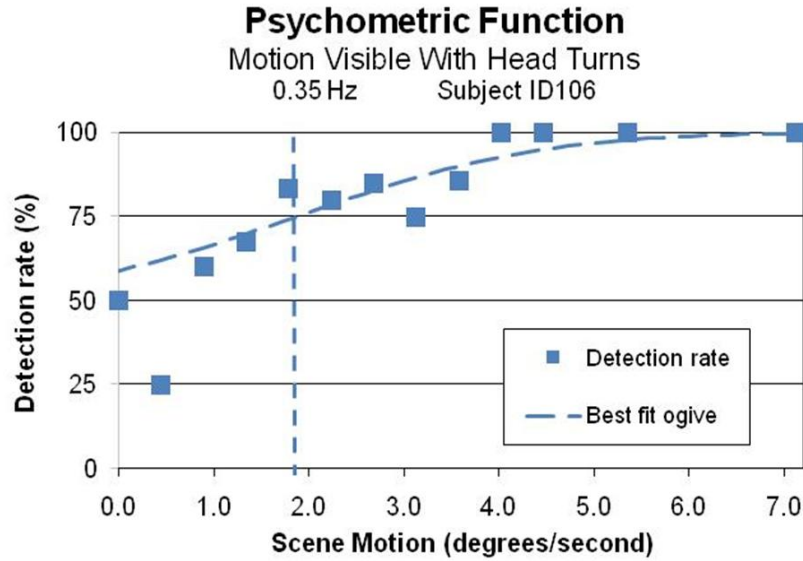


Figure 5.2: Judgments from a single subject for the With Condition. The curve is a psychometric function determined by the best-fit cumulative Gaussian function. The curve fits the data with a correlation of $r = 0.95$. The 75% scene-motion threshold is $1.9^\circ/s$.

take breaks at any time. For each subject, six sessions were conducted over one or more days. Total time per subject, including consent form, instructions, training, experiment sessions, breaks, and debriefing, took three to six hours, with an average of approximately five hours.

5.5 Data Analysis

For each subject and condition, I computed proportions of correct responses for each scene velocity presented. A cumulative Gaussian psychometric function was fit to these proportions (weighted by the number of samples in each proportion). I required a minimum of two judgments for a scene-motion judgment proportion to contribute to the fit. In addition, I added a theoretical proportion of 0.5 correct responses for zero scene motion (since the “stimulus” was identical to the null stimulus, subjects had a 50% chance of selecting the “correct” presentation). The cumulative Gaussian psychometric function ranged between 50% (random guessing) to 100%. The Gaussian distribution’s mean at 75% yields the scene-motion threshold.

Figure 5.2 shows a single subject’s proportion of correct responses for the With Condition scene motions along with the psychometric function fit to those proportions. This subject had a With Condition scene-motion threshold of $1.9^\circ/s$ or 7.7% of the intended peak head-yaw velocity ($24.2^\circ/s$).

Figure 5.3 shows percentage of correct responses and psychometric functions for all three

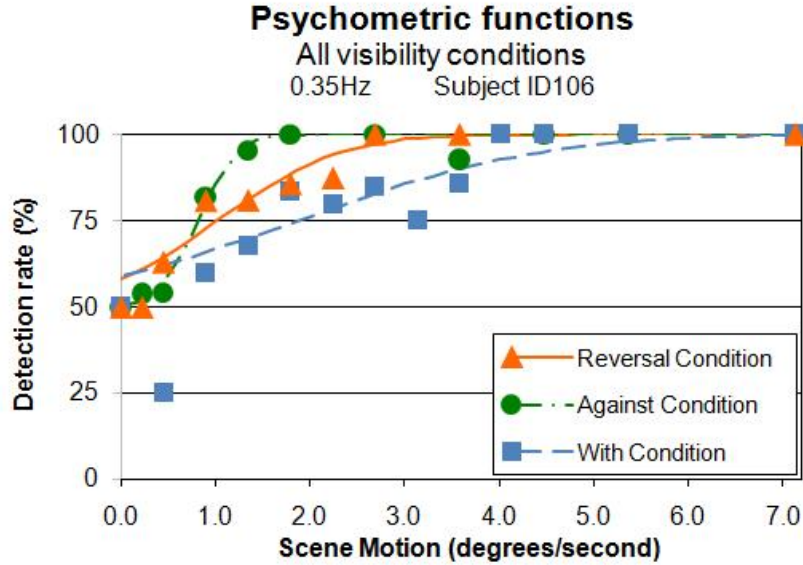


Figure 5.3: Psychometric functions for all three conditions for a single subject.

conditions for a single subject at a head frequency of 0.35 Hz.

I fit 27 psychometric functions (9 subjects \times 3 conditions) to a cumulative Gaussian function. Each of the 27 sets of data were statistically significant (each $p < 0.05$) with the single lowest goodness-of-fit Pearson correlation of $r = 0.56$, the next eight with $0.71 < r < 0.87$, and the remaining 18 at $r > 0.90$.

Figure 5.4 shows scene-motion thresholds for the three conditions from all nine subjects. It is visually evident that scene-motion thresholds are greater for the With Condition. The slopes of the scene-motion psychometric functions increase (not shown) as scene-motion thresholds increase, implying that the equal variance requirements for parametric tests do not hold. Thus I use non-parametric tests.

Friedman analyses of variance (ANOVA) shows that scene-motion thresholds were significantly affected by the conditions, ($Q_2 = 16.22, p < 0.001$). For comparing differences between conditions, I set $\alpha = 0.05/3 = 0.017$ due to Bonferroni Correction. With Condition thresholds were statistically significantly greater than both the Against Condition thresholds and Reversal Condition thresholds (both Wilcoxon matched-pairs signed-rank tests: $S_9 = 0, p_{two-tail} < 0.01$). The Against Condition thresholds were less than the Reversal Condition thresholds. However, this finding was not statistically significant (Wilcoxon matched-pairs signed-rank test: $S_9 = 6, p_{two-tail} < 0.055$). The median of the subject's ratios of Reversal Condition thresholds to Against Condition thresholds was 1.2.

Since there was no statistically significant difference between the Against and Reversal Condition thresholds, I computed the 95% confidence interval of the differences and

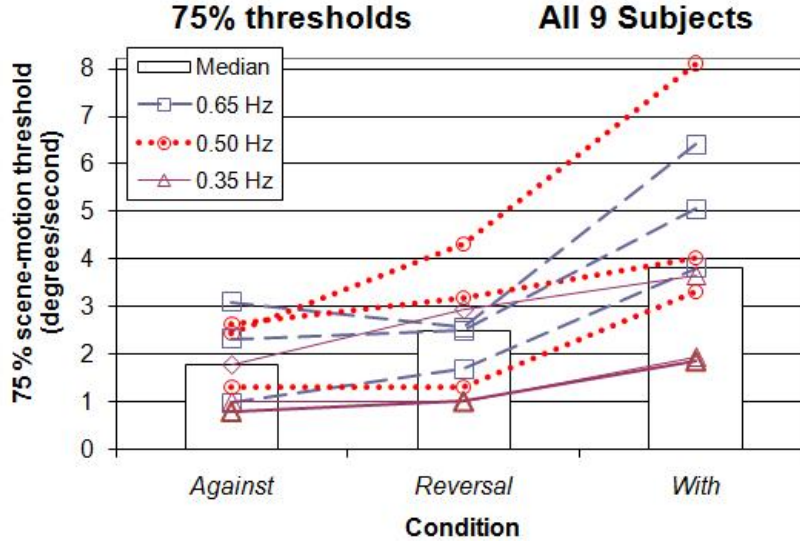


Figure 5.4: Thresholds for all nine subjects for all three visibility conditions. Bars indicate medians for each condition. The lines connect results of individual subjects.

the corresponding range of statistical equivalence. The 95% confidence interval of the differences (Against Condition thresholds subtracted from Reversal Condition thresholds) was $(-0.01^\circ/s, 1.05^\circ/s)$. Thus, the differences were statistically equivalent within $\pm 1.05^\circ/s$ ($p < 0.05$). $1.05^\circ/s$ is 39% of the average of all computed thresholds ($2.7^\circ/s$).

In order to summarize across head frequencies, I computed the ratio of scene-motion thresholds to intended peak head-yaw velocities for the Against and With Conditions. Figure 5.5 shows these ratios. The ratios ranged from 2.2% to 7.7% (median of 5.2%) for the Against Condition and 7.7% to 23.5% (median of 11.2%) for the With Condition. Thus, on average across subjects and trials, subjects did not notice scene motion that was 5.2% to 11.2% of intended peak head-yaw velocity. Note these percentages are provided only as guidelines since they do not take into account actual peak head-yaw velocity.

The median of the ratios of the With Condition to the Against Condition was 2.2; twice as much motion can occur without users noticing when the scene moves with the direction of a head yaw than when the scene moves against the direction of a head yaw.

5.6 Summary of Results

I found the following:

- Scene-motion thresholds were greatest when the scene moved in the same direction as head yaw, compared to when the scene moved against the direction of head yaw and, separately, compared to when the scene moved while head yaw changed direction.

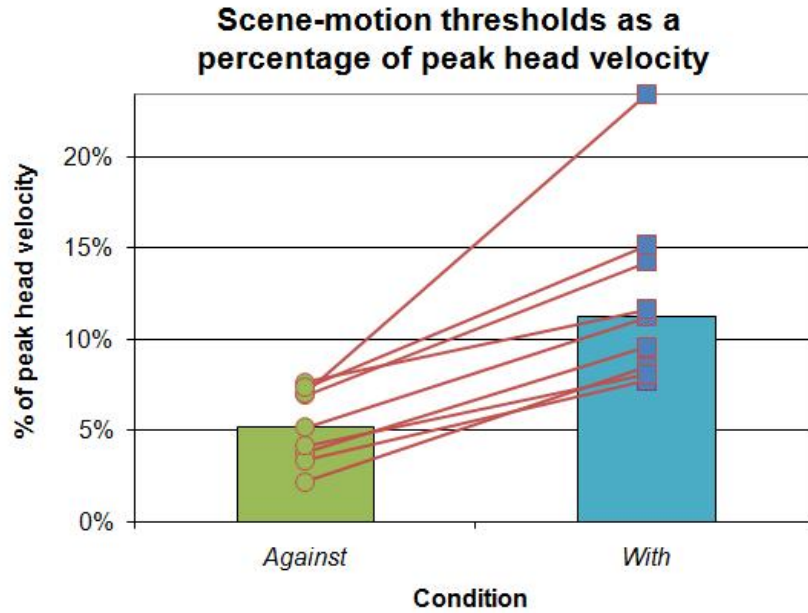


Figure 5.5: Ratios of scene-velocity thresholds to peak head velocities.

Twice as much scene motion occurred without users noticing when the scene moved with the direction of head yaw than when the scene moved against the direction of head yaw.

- Scene-motion thresholds were lower when the scene moved against the direction of head yaw than when the scene moved while head yaw changed direction. However, this finding was not statistically significant. Scene-motion thresholds were equivalent within $1.05^\circ/s$ when the scene moved against the direction of head yaw and when the scene moved while head yaw changed direction.

5.7 Recommendations

The results suggest that in order to determine lower bounds of scene-motion thresholds during quasi-sinusoidal head yaw, the scene should move against the direction of head yaw. I further explore this in Experiment 2.

5.7.1 Applications to Latency Perception

Latency in an HMD system causes the scene to move with the user's head until the system "catches up". The scene moves in the direction the head is turning until shortly after head acceleration goes to zero (constant head velocity) or less (the head decelerates). With constant

head velocity (zero head acceleration), the scene appears to be stable in space (i.e., no scene motion) with a constant offset from where it would correctly appear with no head motion or zero latency. As the head decelerates, the scene starts to move back to where it should be in space, moving against the head turn. Maximum latency-induced scene velocity occurs near the time that head-motion direction reverses or starts/ends from/at a stopped position (Section 3.2), when head acceleration peaks. These facts, combined with the results of this experiment (that the With Condition thresholds are greater than the Against Condition and the Reversal Condition thresholds), suggest users are less likely to notice latency in an HMD when beginning a head turn (when the scene moves with the head turn) than when slowing down a head turn (when the scene moves against the head turn) or when reversing head-motion direction (when latency-induced scene velocity peaks).

The results also suggest that the reason subjects have the lowest latency thresholds at the reversal of head yaw (Adelstein et al., 2005) is because latency-induced scene velocity is maximized at the reversal of head yaw, not because scene-motion thresholds are highest when the direction of head motion changes.

Chapter 6

Experiment 2: Single Head Yaw, Luminance, and Scene-Motion Direction

Experiment 1 found that scene-motion thresholds were greater when the scene moved with the direction of head yaw (the With Condition) than when the scene moved against the direction of head yaw (the Against Condition). However, this finding was only for very specific conditions—for the center of head yaw (approximately constant head velocity). I did not define start or end head-yaw phase conditions in that experiment because head yaw was quasi-sinusoidal—the “end” of a head yaw was also the “start” of a head yaw; the head did not accelerate from a fully stopped position or decelerate to a fully stopped position. I wanted to provide a stronger case that scene-motion thresholds are greater for “with” conditions than “against” conditions, and that the results were not simply an artifact of the head motions and other specifics of Experiment 1. Therefore, for Experiment 2, I compared thresholds for “with” and “against” conditions during the time that head yaw started and ended.

I also wanted to know if scene luminance and contrast affected the results of Experiment 1. Graham (1965) described experiments that found subjects are more sensitive to motion of brighter stimuli. Although the measurements were taken when the head was held still, I suspected the findings would be similar when the head moves.

It is also possible, for Experiment 1, that the scene luminance and contrast could have biased scene-motion thresholds to be greater for the With Condition than for the Against Condition of Experiment 1. Such a dependence on luminance and contrast is plausible because of previous findings:

- Brightness is a depth cue that can cause brighter objects to appear to be in front of darker objects (Coren et al., 1999). The pivot hypothesis (Section 2.2.7.3) states that

incorrectly perceived distance can cause a stimulus to seem to move as the head moves when it is fact stable on the display surface. If a non-moving bright scene appears to be in front of the display surface, then the scene could be perceived to move with the direction of head turns and the scene would have to move against the direction of head turns to appear stable. In this case, higher scene-motion thresholds would result when the scene moves against the direction of head turns and lower thresholds would result when the scene moves with the direction of head turns (or the opposite result if a darker scene was perceived to be behind the display surface). I wanted to know if a brighter scene could nullify the results found in Experiment 1, i.e., would “with” condition thresholds still be greater than “against” condition thresholds for bright scenes?

- Contrast can affect the direction a scene seems to move. Freeman and Banks (1998) found that low-contrast scenes can reverse the direction of perceived movement in the Aubert-Flieschl and Filehne illusions (Section 2.2.7.6). For low-contrast scenes, an object can appear to move faster (instead of slower) when pursued with the eyes compared to when the eyes are stable, and a stable background can appear to move with (instead of against) eye motion when a moving object is pursued with the eyes.
- Dimmer scenes take longer to perceive than brighter scenes (Section 2.2.7.6). In situations that lack real-world cues, delayed perception of dimmer scenes may result in perceived scene motion, similar to that which occurs with a lagging HMD system (Sections 2.4 and 3.1).

6.1 Hypotheses

I tested the following hypotheses:

- **Hypothesis 1:** Scene-motion thresholds are greater when the scene moves with the direction of head yaw than when the scene moves against the direction of head yaw, regardless of head-yaw phase and regardless of scene luminance (contrast).
- **Hypothesis 2:** Scene-motion thresholds are lower for a bright scene than for a dim scene.

6.2 Experimental Design

I measured scene-motion thresholds for different conditions. The experiment was a repeated-measures adaptive-staircase design. Each subject experienced all conditions.

I asked subjects to yaw their heads starting and ending at stopped positions (i.e., one time from left-to-right or right-to-left) over a period of one second. A moving or non-moving scene was presented to them during head yaw. Subjects then stated if they believe the scene moved or not. They then rated their confidence in their judgments.

I measured scene-motion thresholds for three head-yaw phases, two scene-motion directions, and two scene luminances (contrasts).

6.2.1 The Stimulus

The stimulus to be detected was scene motion. Scene velocity was constant within trials and varied between trials.

6.2.2 Controlled Variables

Head-Yaw Period

Subjects yawed their heads over a period of one second.

Head-Yaw Amplitude

Visual cues before the trials, and auditory cues before and during the trials suggested subjects yaw their heads by $\pm 15^\circ$ from straight ahead.

6.2.3 Independent Variables

Three variables were manipulated—three head-yaw phases, two scene-motion directions, and two scene luminances—resulting in 12 conditions.

Head-Yaw Phase

The system presented a scene for a single phase of a head yaw from a start position to a stopped position. The scene appeared at some point during the head yaw and moved with constant velocity or did not move at all. I trained subjects to start yawing their heads at two seconds of a metronome beep and stop yawing their heads at three seconds. The head-yaw phase conditions were:

- **Start Condition**—The scene was presented just before the intended start of the head yaw (1.9 seconds) to half of the intended head yaw (2.5 seconds).
- **Center Condition**—The scene was presented during the central part of the head yaw (2.2 to 2.8 seconds).

- **End Condition**—The scene scene was presented from half of the intended head yaw (2.5 seconds) to just after the end of the intended head yaw (3.1 seconds).

Scene-Motion Direction

For each trial, the scene did not move or moved with some constant velocity in a direction relative to the head-yaw direction:

- **With Condition**—The scene moved in the same direction that the head yawed.
- **Against Condition**—The scene moved in the opposite direction that the head yawed.

Scene luminance and contrast

Two scene luminances were presented to subjects:

- **Dim Condition**—The scene foreground was 10.0 cd/m^2 and the scene background was 0.067 cd/m^2 resulting in a contrast of 167.
- **Bright Condition**—The scene foreground was 0.11 cd/m^2 and the scene background was 0.003 cd/m^2 resulting in a contrast of 37.

The Bright Condition foreground emitted two orders of magnitude more light than the Dim Condition foreground. The Bright Condition contrast was 4.5 times greater than the Dim Condition contrast.

6.2.4 Dependent Variable

Scene-motion thresholds

I defined the scene-motion threshold, determined from yes/no judgments with confidence ratings (Sections 4.1.4.4 and 4.3.1), to be the scene motion, in degrees per second, at which a subject was 75% confident that the scene moved.

For each subject, I measured scene-motion thresholds for the four head-yaw phases and the three head-yaw measures shown in Table 6.1.

6.3 Methods

Each condition consisted of two adaptive staircases, one staircase starting with zero scene motion and one staircase starting with a scene motion of $10^\circ/s$, with the staircase terminated on the eighth reversal. The staircase step size started at $5^\circ/s$ and was halved at every reversal until a minimum step size of $0.625^\circ/s$ was reached. Staircases were randomly interleaved in order to minimize bias.

Table 6.1: Experiment 2 conditions. Each subject experienced all conditions.

luminance	Scene Direction	Head-Yaw Phase		
		Start	Center	End
Dim	With			
	Against			
Bright	With			
	Against			

I used a yes/no rating task; subjects judged if the scene appeared to move or not and then stated their confidence in that judgment. Three levels of confidence resulted in six possible responses per trial (Sections 4.1.4.4 and 4.3.1).

6.4 Participants

Nine subjects (7 male and 2 female, age 18–34) participated. I served as one of the subjects; all other subjects were naive to the experimental conditions.

6.5 Data Analysis

For each of the 9 subjects and 12 conditions, I fit a cumulative Gaussian psychometric function to data, derived from yes/no judgments with confidence ratings, and then extracted 75% thresholds. I then computed differences of conditions between subjects and compared these differences across subjects.

The average false alarm rate (the average confidence rating that there was scene motion when there was no scene motion) across all conditions and all subjects was 24.8% with one subject having the largest false alarm rate of 56.4%. These high alarm rates suggest the subjects were largely guessing when there was no scene motion.

Figure 6.1 shows data collected from a single subject for the Center-Dim-Against Condition. Individual trial ratings are shown (some Xs represent multiple ratings) along with the mean rating per scene-motion level, a psychometric function fit to the data, and the 75% scene-motion threshold.

Figure 6.2 shows box plots for all subjects’ thresholds for all conditions.

My primary goal for this experiment was to determine whether the With Condition thresholds are greater than the Against Condition thresholds (as found in Experiment 1) for different scene luminances (contrasts) and different phases of head yaw. Thus, I conducted Wilcoxon matched-pairs signed-rank tests comparing the With Condition thresholds to the

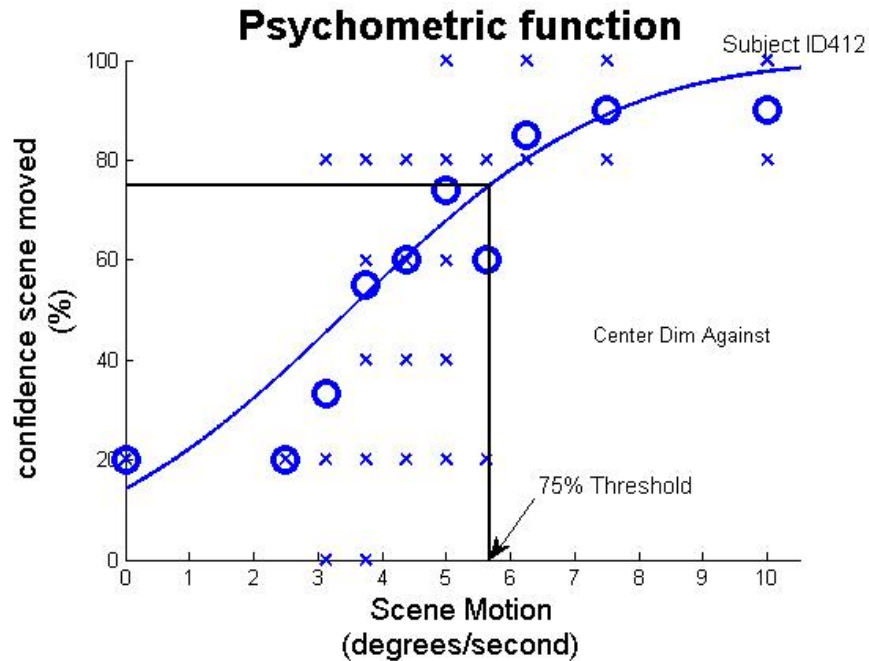


Figure 6.1: Confidence ratings and the resulting psychometric function and scene-motion threshold for a single subject for the Center-Dim-Against Condition. The small Xs represent ratings for one or more trials (some Xs represent multiple ratings), the large Os are the mean rating for that scene-motion level, and the curve is the psychometric function (a cumulative Gaussian function fit to the data). For this condition, the subject rated the scene to move with 75% confidence when the scene moved at $5.7^\circ/s$.

Against Condition thresholds for each of the six tested Conditions (3 head-yaw phases \times 2 scene luminances). I used a one-tail test since I expected scene-motion thresholds to be greater for the With Condition than the Against Condition. The With Condition thresholds were found to be statistically significantly greater (each $S_9 \leq 7$, each $p_{one-tail} < 0.05$) than the Against Condition thresholds. No correction was required for $\alpha = 0.05$ since the largest p-value was less than 0.05.

The median ratio of With Condition thresholds to Against Condition thresholds was 1.89 and the mean ratio was 1.98 with a standard deviation of 0.93 (mean ratios were calculated in addition to median ratios because of the large number of 54 ratios (2 luminances \times 3 head-yaw phases \times 9 subjects)). These average ratios agree with the results of Experiment 1 that scene-motion thresholds when the scene moves with the direction of head yaw are approximately twice as large as when the scene moves against the direction of head yaw.

I also thought scene luminance (contrast) might affect thresholds. I conducted Wilcoxon matched-pairs signed-rank tests comparing the Dim Condition thresholds to the Bright Condition thresholds for each of the six tested Conditions (3 head-yaw phases \times 2 scene-motion directions). None of the differences were statistically significant with Bonferroni

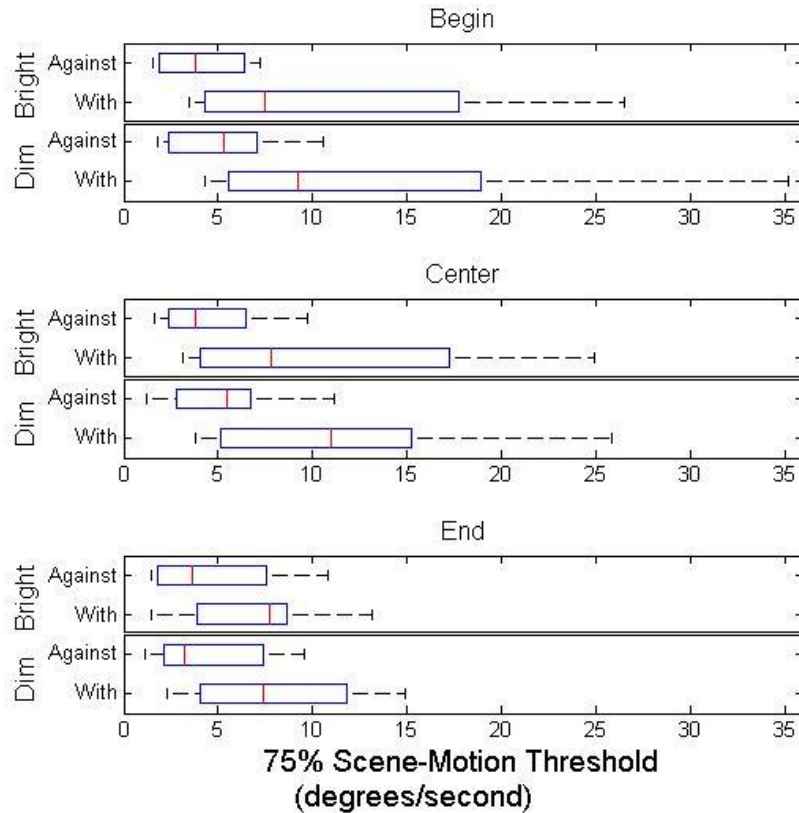


Figure 6.2: Box plots of all scene-motion thresholds. With Condition thresholds were statistically significantly greater than Against Condition thresholds for each of the three head-yaw phases for each of the two scene-luminance conditions. No statistically significant differences were found between the two scene-luminance conditions.

Correction ($\alpha = 0.05/6 = 0.008$) although there was a marginally statistically significant difference ($p_{two-tail} < 0.039$) between the Dim Condition thresholds and the Bright Condition for the Start-Against Condition (with the Dim Condition thresholds being greater than the Bright Condition thresholds).

I suspect the tests would find differences for the tested conditions with more statistical power. However, the lack of statistically significant differences implies scene luminance and contrast plays less of a role than the statistically significant differences found between the With and Against Conditions. Larger differences in scene luminance and/or contrast would also likely result in threshold differences.

I also compared thresholds across head-phase conditions. I conducted Wilcoxon matched-pairs signed-rank tests comparing the Head-Phase Conditions. Because no differences were found between the two scene-luminance conditions, I collapsed the Dim and Bright Conditions together. Thus, there were 6 comparisons total (3 head-yaw phases \times 2 scene-motion directions). I set $\alpha = 0.05/6 = 0.008$ due to Bonferroni Correction. For the With

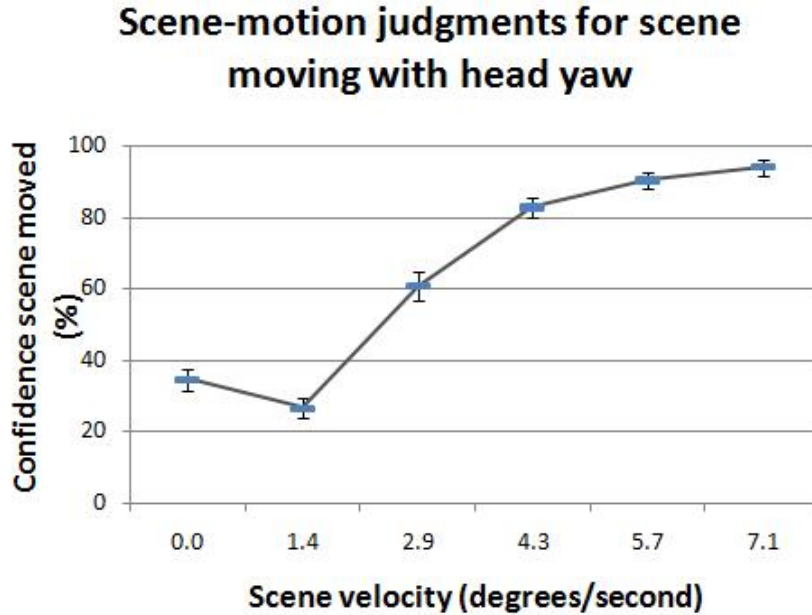


Figure 6.3: Pilot data (288 yes/no judgments with confidence ratings). shows an illusion that the scene appeared to move when the scene was stable and the head was yawing. Confidence that the scene moved are plotted against six scene motions. Confidence that the scene moved is greater when the scene did not move than when the scene moved at $1.4^\circ/s$ with the direction of head yaw.

Condition, the End Condition thresholds were less than both the Center Condition thresholds ($S_{18} = 15, p_{two-tail} < 0.002$) and the Start Condition thresholds ($S_{18} = 8, p_{two-tail} < 0.001$). I do not claim this to generalize across general head yaw since some subjects stopped yawing earlier than I intended and the scene kept moving, which could have caused the End Condition to result in lower thresholds.

6.6 An Illusion of Scene Motion

It may be, as other investigators have reported (Wallach and Kravitz, 1965; Steinicke et al., 2008, 2010), that scenes appear most stable when there is a small amount of scene motion with the direction of head turns. This could explain why my With Condition scene-motion thresholds are greater than my Against Condition scene-motion thresholds.

In a pilot study on myself, I measured scene-motion thresholds as the scene moved with a single left-to-right or right-to-left head yaw for six levels of scene-motion. Results are shown in Figure 6.3. I rated the scene to be more stable when the scene moved at $1.4 \text{ deg}/s$ than when the scene did not move, although this result was not statistically significant ($z_{94} = 1.92, p_{two-tail} < 0.058$).

This trend warrants further investigation using subjects other than myself. Possible explanations of why such an illusion may occur is discussed in Section 10.1.3.

6.7 Summary of Results

Experiment 2 confirmed that scene-motion thresholds are greater when the scene moves with the direction of head yaw than when the scene moves against the direction of head yaw for the following conditions:

- Three phases of single head yaw (Start, Center, and End Conditions).
- Two scene-luminance (and contrast) levels (Dim and Bright Conditions).

The median ratio of With Condition thresholds to Against Condition thresholds was 1.89.

These results suggest that the finding from Experiment 1 that the With Condition thresholds are approximately twice as high as Against Condition thresholds holds across a broad range of head motions and scene-luminance (contrast) conditions. The results of Experiment 1 are not an artifact of those specific conditions.

I did not find statistically significant differences between scene luminance differing by two orders of magnitude (and contrasts differing by a factor of 4.5).

6.8 Recommendations

No differences of scene-motion thresholds were found between the two scene luminances, which differed by two orders of magnitude. This suggests that differences of thresholds due to scene luminance or contrast, if any, are minor compared to the differences found when the scene moves with the direction of head yaw versus when the scene moves against the direction of head yaw.

The results strengthen the claims from Experiment 1 that subjects can more easily detect

- scene motion when the scene moves against the direction of head yaw than when the scene moves with the direction of head yaw, and
- latency at the end of head yaw, during the time the head is decelerating and thus latency-induced scene motion moves against the direction of head yaw.

For experiments 3-5, I only measure thresholds for scenes moving against the direction of head yaw such that most conservative thresholds (i.e., lower thresholds) are measured.

Chapter 7

Experiment 3: Increasing Head Motion

The goal of Experiment 3 was to find whether scene-motion thresholds increase as head motion increases.

I asked subjects to yaw their heads starting and ending at stopped positions (i.e., one time from left-to-right or right-to-left), with different suggested head-yaw amplitudes, over a period of one second. A moving or non-moving scene was presented to them during head yaw. Subjects then stated if they believed the scene moved or not. They then rated their confidence in that judgment.

I measured scene-motion thresholds for different phases and measures of head motion. Because scene-motion thresholds are lower when the scene moves against the direction of head yaw (Experiments 1 and 2), I only measured scene-motion thresholds for such “against” motion.

7.1 Hypotheses

I tested the following hypotheses for four phases of head yaw:

- **Hypothesis 1:** Scene-motion thresholds increase as head angle range, given constant time, (i.e., average head-yaw velocity) increases.
- **Hypothesis 2:** Scene-motion thresholds increase as peak head-yaw velocity increases.
- **Hypothesis 3:** Scene-motion thresholds increase as peak head-yaw acceleration increases.

7.2 Experimental Design

The experiment was a repeated-measures (all subjects experienced all conditions) constant-stimuli (stimulus levels were randomly selected independent of previous results) design. Each subject experienced all conditions. For each trial, scenes were presented as subjects yawed their heads a single time from left-to-the-right or from right-to-the-left. The system presented either a moving or non-moving scene. Subjects then stated if they believed the scene moved or not and then stated their confidence in that judgment.

7.2.1 The Stimulus

The stimulus to be detected was scene motion. Scene velocity was constant within trials and varied between trials.

7.2.2 Controlled Variables

Head-Yaw Period

Subjects yawed their heads, a single time from left-to-right or right-to-left, over a period of one second.

Scene-Motion Direction

Based on the results of Experiment 1, scenes moved against against the direction of the head yaw, when thresholds are lowest.

7.2.3 Independent Variables

Two variables were manipulated—head-yaw phase and intended head-yaw amplitude.

Head-Yaw Phase

with some random and constant velocity or did not move at all. I trained subjects to start yawing their heads two seconds after an audio beep and stop yawing their heads three seconds after the audio beep. The head-yaw phase conditions were:

- **Start Condition**—The scene was presented just before the intended start of the head yaw (1.9 seconds) to half of the intended head yaw (2.5 seconds).
- **Center Condition**—The scene was presented during the central phase of the head yaw (2.2 to 2.8 seconds).
- **End Condition**—The scene scene was presented from half of the intended head yaw (2.5 seconds) to just after the end of the intended head yaw (3.1 seconds).

- **All Condition**—The scene was presented for the entire duration of the intended head yaw (1.9 to 3.1 seconds).

Unlike in Experiment 2, the system turned off the scene when head-yaw velocity fell below a threshold value, so that subjects would not judge motion when their heads were not moving.

Intended Head-Yaw Amplitude

Visual cues before the trials, and auditory cues before and during the trials suggested subjects yaw their heads by $\pm 5^\circ$, $\pm 10^\circ$, $\pm 15^\circ$, or $\pm 20^\circ$ from straight ahead.

7.2.4 Dependent Variables

Head-Motion Measure

From the tracker log, I analyzed three measures of head motion:

- **Head-Yaw Range Measure**—The range of the head yaw, in degrees, during the time of the head-yaw phase condition (i.e., the the scene was visible), given constant time (i.e., average head-yaw velocity).
- **Peak Head-Yaw Velocity Measure**—The absolute value of the peak head-yaw velocity (i.e., left or right peak velocity), in degrees per second, during the time that the scene was visible.
- **Peak Head-Yaw Acceleration Measure**—The absolute value of the peak head-yaw acceleration (i.e., peak deceleration or acceleration), in degrees per second squared, during the time that the scene was visible.

The suggested head-yaw amplitude, and checks for head yaw to be reasonably close to the suggested head yaw amplitude over a period of one second, ensured head head motion varied over a wide range of head motion.

Scene-Motion Threshold

As in Experiments 1 and 2, I defined the scene-motion threshold to be the scene velocity, in degrees per second, at which a subject was 75% confident that the scene moved.

For each subject, I measured scene-motion thresholds for the four head-yaw-phases and three head-yaw measures shown in Table 7.1.

Table 7.1: Experiment 3 conditions. Each subject experienced all conditions.

		Head Phase			
		Start	Center	End	All
Measure	Head-Yaw Range (<i>deg</i>)				
	Peak Head-Yaw Velocity (<i>deg/s</i>)				
	Peak Head-Yaw Acceleration (<i>deg/sec²</i>)				

7.3 Methods

The method of constant stimuli (Section 2.3.3.2) determined random ordering of trials. As in Experiment 2, judgments consisted of a yes/no rating task with three levels of confidence resulting in six possible responses per trial (Sections 4.1.4.4 and 4.3.1).

7.4 Participants

Eight subjects (5 male and 3 female, age 18–27) participated. I informed subjects that the scene motion, if any, would occur only in the opposite direction of their head yaw—scene motion would never occur in the same direction as their head yaw. Otherwise, all subjects were naive to the experimental conditions. I encouraged subjects to take breaks at will. Total time per subject, including consent form, instructions, training, experiment sessions, breaks, and debriefing, was four hours or less. Seven subjects judged 576 trials—144 judgments for each phase condition. One subject judged only 535 trials due to time limitations.

7.5 Data Analysis

False alarm rates (average confidence rating that there was scene motion when there was no scene motion) across all conditions were 12.9% or less for each of the six subjects.

Since head motion could not be precisely controlled, I divided head motions for each of the three head-motion measures and for each of the four head-yaw phase conditions into sextiles (i.e., six bins of head motion from slow to fast with each bin having an equal number of samples per bin). This resulted in 72 scene-motion thresholds per subject.

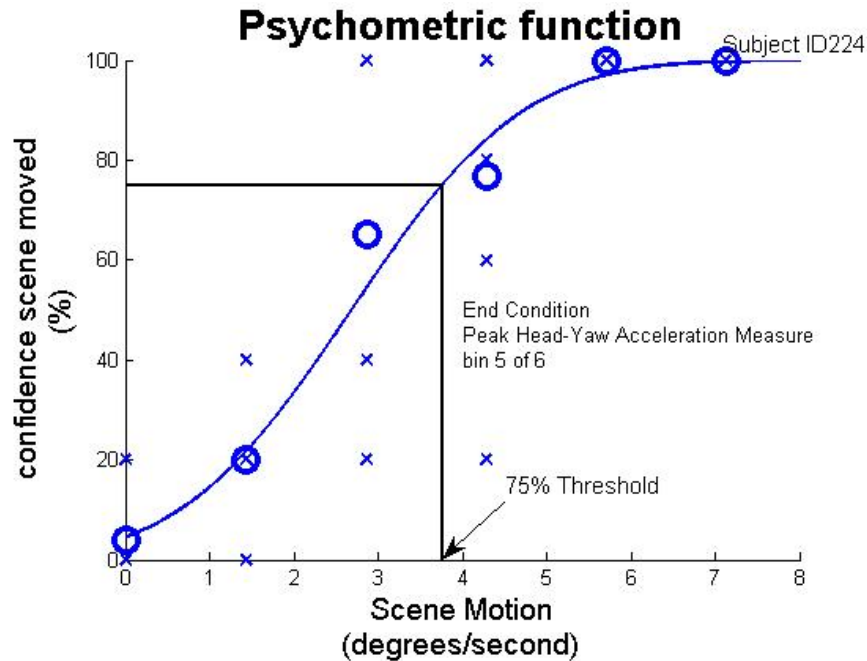


Figure 7.1: Confidence ratings from a single subject for the End Condition and sextile five of the Peak Head-Yaw Acceleration Measure. The small Xs represent ratings for one or more trials (some Xs represent multiple ratings), the large Os are the mean rating for that scene-motion level, and the curve is the psychometric function (a cumulative Gaussian function fit to the data). For this sextile, the subject rated the scene to move with 75% confidence when the scene moved at $3.8^{\circ}/s$.

For all subjects and all conditions from Experiment 3, the mean of scene-motion thresholds (in degrees per second) as the percentage of the corresponding Peak Head-Yaw Velocity Measure (192 ratios) was 14.7% (median = 11.5%), with a standard deviation of 11.2%.

I fit a cumulative Gaussian function to the data for each sextile. 75% scene-motion thresholds were then extracted for each sextile and plotted against the sextile's head-motion mean.

Figure 7.1 shows data collected from a subject for sextile five (head motion bin five of six) for the End Condition and the Peak Head-Yaw Acceleration Measure. Individual trial ratings are shown along with the mean rating per scene-motion level (some Xs represent multiple ratings), a psychometric function fit to the rating data, and the 75% scene-motion threshold.

Figure 7.2 shows the same subjects' 75% scene-motion thresholds versus the Peak Head-Yaw Acceleration Measure for the End Condition. The vertical lines represent the sextile boundaries.

Figure 7.3 shows the same subjects' 75% scene-motion thresholds versus the Peak Head-Yaw Acceleration Measure for all four head-yaw phase conditions.

Pearson correlations of scene-motion thresholds to head-motion measures were computed for the 12 conditions/measures (3 head motion measures \times 4 phases of head yaw) for each of

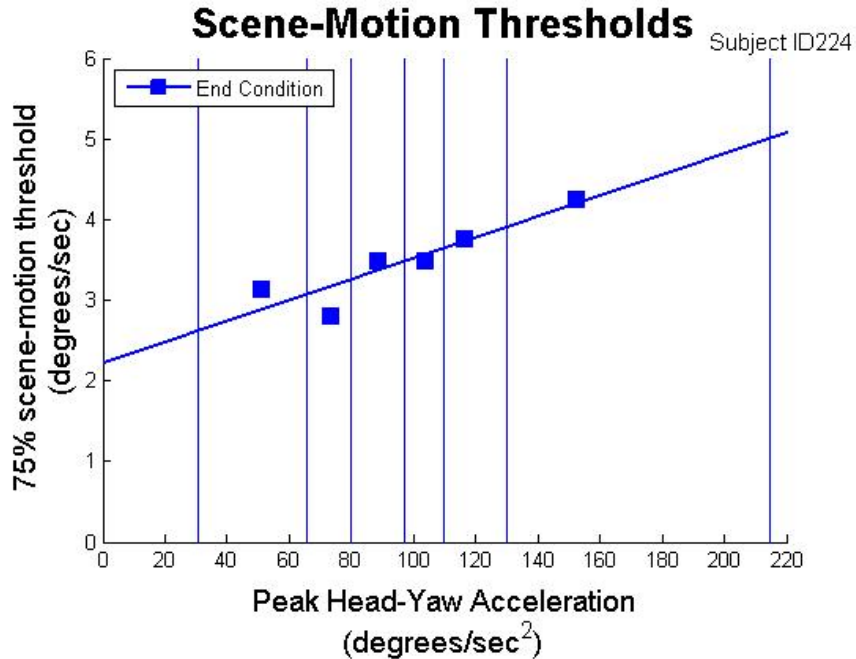


Figure 7.2: Scene-motion thresholds for a single subject for the End Condition where head motion is defined by the Peak Head-Yaw Acceleration Measure. The vertical lines represent boundaries of the sextiles.

the 8 subjects. Figure 7.4 shows box plots of all 8 subjects for all 12 conditions. Although all Pearson correlations were not statistically significantly greater than zero, sign tests across all 8 subjects yielded correlations greater than zero for each of the 12 conditions/measures. I used a one-tailed test since I expected positive correlations. 8 of 8 subjects had positive correlations for 10 of the 12 conditions/measures ($p_{one-tail} < 0.004$). 7 of 8 subjects had positive correlations for the remaining two conditions/measures ($p_{one-tail} < 0.035$): the Peak Head-Yaw Velocity Measure for the Start Condition and the Peak Head-Yaw Acceleration Measure for the Center Condition. No correction was required for $\alpha = 0.05$ since the largest p-value was less than 0.05.

The results show that scene-motion thresholds increase as head motion increases, independently of the head-motion measure used and the phase of the head yaw for which the scene was presented.

I conducted three two-way ANOVA tests (one test for each head-motion measure) to determine if there were differences between head-yaw phase conditions. I matched the eight subjects and sextiles and tested across conditions (each test consisted of 48 thresholds for each of the four head-yaw phase conditions). Two-way ANOVA found no statistically significant differences between the head-yaw phase conditions for the Head-yaw Range Measure ($F_{3,47} = 0.88$) or for the Peak Head-Yaw Acceleration Measure ($F_{3,47} = 0.67$). There were statistically

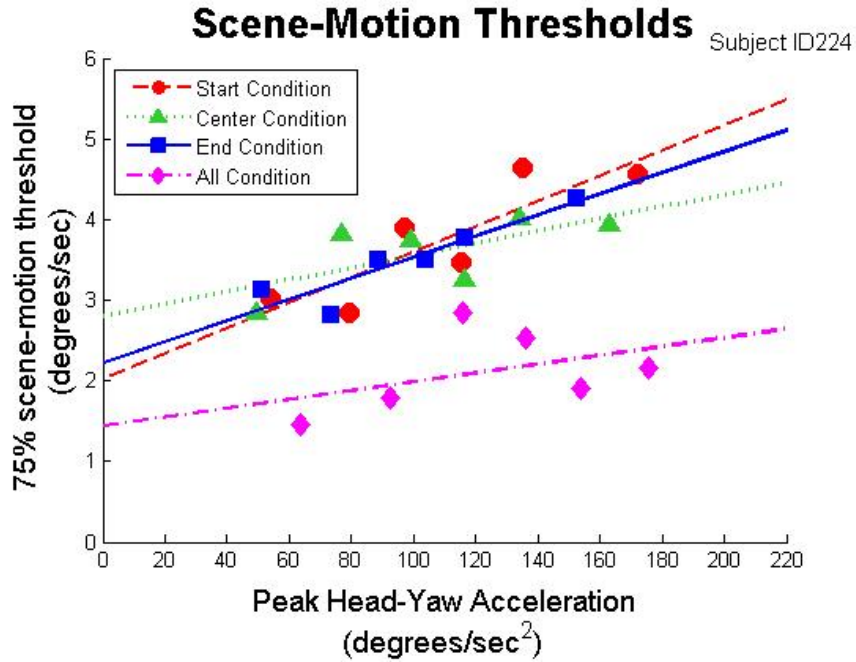


Figure 7.3: Scene-motion thresholds for a single subject for all four head-yaw phase conditions where head motion is defined by the Peak Head-Yaw Acceleration Measure. Note the lower thresholds for the All Condition and the similarities of the other conditions—such trends were not consistent for all subjects.

significant differences between conditions for the Peak Head-Yaw Velocity Measure ($F_{3,46} = 4.12, p < 0.008$) (note one row of data was not used for the peak head-yaw velocity due to not enough data to compute a psychometric function for one of the data points).

Since there were statistically significant differences between conditions for the Peak Head-Yaw Velocity Measure, I conducted z-tests to find the differences. I set $\alpha = 0.5/6 = 0.0083$ due to Bonferroni correction. Start Condition thresholds were greater than All Condition thresholds ($z_{46} = 3.67, p_{two-tail} < 0.001$) and End Condition thresholds ($z_{46} = 3.0, p_{two-tail} < 0.004$).

Lower thresholds for the All Condition is no surprise because the scene was visible for twice as long as for the other conditions.

I cannot conclude too much about the End Condition thresholds being less than the Start Condition thresholds because of the following:

- Head yaw may not have been symmetric resulting in biased results. Although I attempted to have all subjects make a symmetric head-yaw motion, it was impossible to force users to yaw their heads in a precise way. For example, a subject may have suddenly started turning her head at 2.0 seconds and more slowly stopped moving her head early at 2.9 seconds, such that head motion was slower and/or stopped for a longer period of time for the end of the head turn than for the start of the head turn.

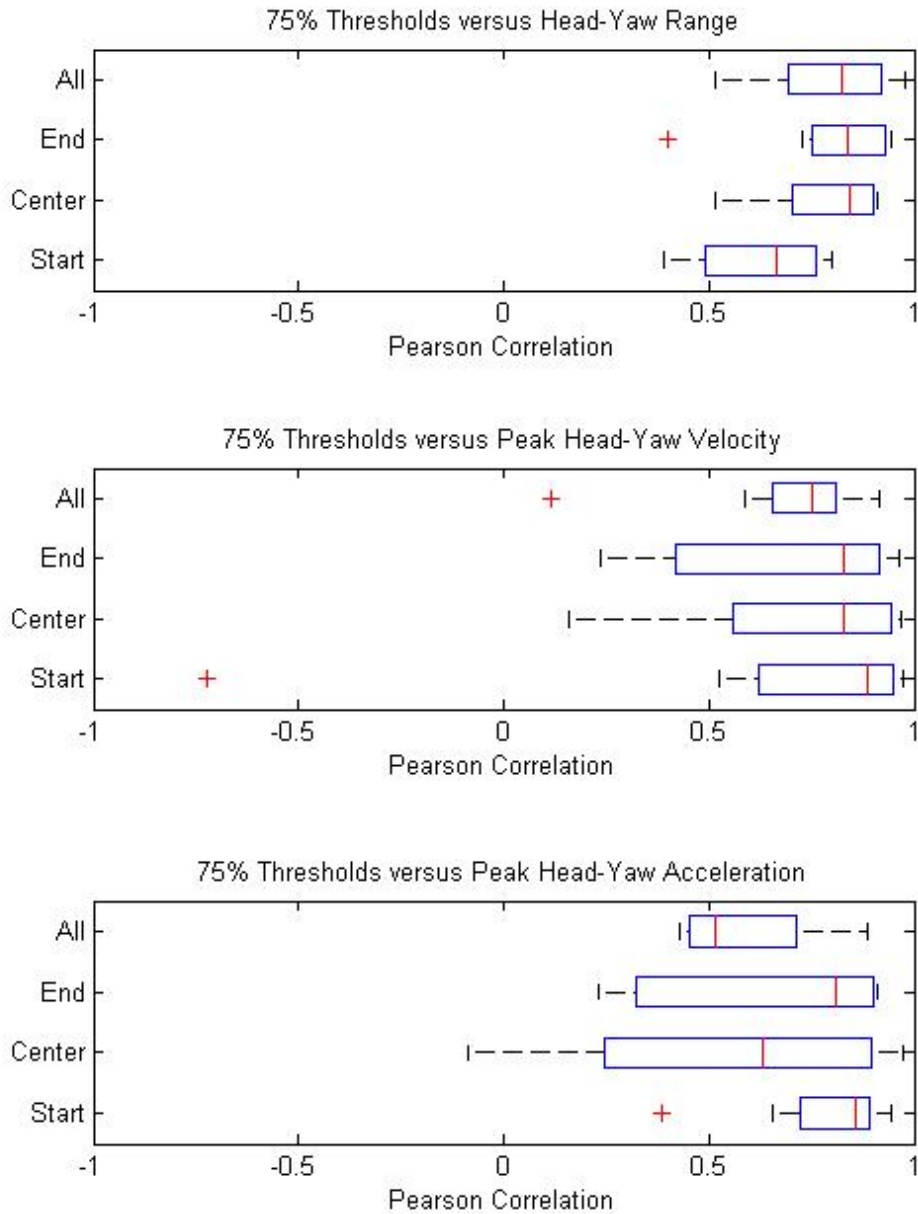


Figure 7.4: Box plots of Pearson Correlations for the eight subjects.

- The Start Condition may have been presented for shorter durations than the End Condition (or vice versa), due to the scene being turned off when head-yaw velocity was below a threshold.

7.6 Summary of Results

Scene-motion thresholds positively correlated with three measures of head motion for four phases of head yaw.

I found that scene-motion thresholds increased as

- the head-yaw range increased,
- the peak head-yaw velocity increased, and
- the peak head-yaw acceleration increased.

I found that scene-motion thresholds increased regardless of

- the start-phase of head yaw,
- the center-phase of head yaw,
- the end-phase of head yaw, and
- the entire head yaw.

7.7 Recommendations

The results suggest that perception of scene motion decreases as head motion increases, independent of the head-motion measure used and independent of the phase of the head yaw, i.e., scene motion is more imperceptible in IVEs for faster head motions. Injected scene motion that is intended to be below conscious perception should not be limited by a fixed maximum value, but should be limited according to a function of the current amount of head motion.

Chapter 8

Experiment 4: Validating the Latency-Thresholds Model

Experiment 4 investigated my mathematical model relating latency thresholds to peak head-yaw accelerations (Section 3.4.2). The form of the model is

$$\Delta t = \tau + \psi \left(\frac{1}{\phi''} \right) \quad (3.20)$$

where Δt is the latency threshold, in seconds, ϕ'' is peak head acceleration, in degrees per second squared, τ is an offset parameter, in seconds, and ψ is a scale parameter, in degrees per second.

The goals of this experiment were to:

1. Measure scene-motion thresholds and latency thresholds over a range of peak head-yaw accelerations.
2. Validate my model relating latency thresholds to
 - peak head-yaw accelerations and
 - scene-motion thresholds.
3. Compare thresholds resulting from three types of scene motion that occur during the ends of head yaw. The three types of scene motion were:
 - **Constant scene velocity (Constant Condition)**—The scene moves with a constant velocity against the direction of the head yaw, in the same way as in Experiments 1-3.

- **Gaussian-modeled scene motion (Gaussian Condition)**—Head acceleration for a head motion that comes to a sudden stop can be modeled with a Gaussian-temporal profile (head acceleration as a function of time). Scene velocity due to latency scales with head acceleration (Section 3.1.3). I postulate that scene velocity due to latency can be approximated with a Gaussian-temporal profile (head velocity as a function of time) during the time that head motion comes to a stop. For this Gaussian Condition, the scene velocity is modeled with Gaussian-temporal profile such that scene velocity peaks at the same time as scene velocity typically peaks in an HMD, due to head motion and latency.
- **Latency-induced scene motion (Latency Condition)**—The scene motion is manipulated in real-time based on the subjects’s actual head yaw and injected latency. Since latency controls scene motion in this condition, both latency thresholds, in seconds, and scene-motion thresholds, in degrees per second, were determined from the same trials.

The Constant and Gaussian Conditions are independent of the subject’s head motion, whereas the Latency Condition is dependent upon the subject’s head motion.

8.1 Hypotheses

I tested the following hypotheses:

- **Hypothesis 1:** Constant scene velocity, Gaussian-modeled scene velocity, and latency-induced scene velocity derived from actual head motion result in different scene-motion thresholds.
- **Hypothesis 2:** Directly-measured latency thresholds correlate with the inverse function of the form $\Delta t = \tau + \psi \left(\frac{1}{\phi''} \right)$, where ϕ'' is peak head-yaw acceleration, better than with a linear function of the form $\Delta t = b + m\phi''$.

8.2 Experimental Design

The experiment was a repeated-measures constant-stimuli design. Each subject experienced all conditions.

For three different types of scene motions and increasing head motions, subjects judged whether they believed the scene moved and rated their confidence in that judgment. I fit psychometric functions to their responses and extracted thresholds. I then compared and related the thresholds among the three types of scene motions.

8.2.1 The Stimulus

The stimulus to be detected was scene motion (controlled by the system and the subject's head motion). Scene motion was one of three types. Peak scene velocity or latency was constant within trials and varied between trials.

8.2.2 Controlled Variables

Head-Yaw Motion

As in Experiments 2 and 3, I trained subjects to start yawing their heads at two seconds after the start of the trial and stop yawing their heads at three seconds.

Head-Yaw Phase

The scene was visible only for the end of the head yaw, over the interval from 2.5 to 3.2 seconds (similar to the End Condition of Experiments 2 and 3). I presented the test scene at the end of the head yaw because:

- Head acceleration and scene velocity typically peak near the end of head yaw (at about 3 seconds for this experiment) (Sections 3.1.3 and 3.2).
- Subjects are most sensitive to latency when head acceleration peaks (Adelstein et al., 2005).
- Latency causes the scene to move with the head at the beginning of a head yaw and to move against the head at the end of a head yaw (Section 3.2). Subjects are more sensitive to scene motion when the scene moves against the direction of head yaw (Experiments 1 and 2).

Scene-Motion Direction

I set the scenes to move against the direction of head yaw, when thresholds are lowest (Experiments 1 and 2). For the case of latency-induced scene motion, the scene mostly moved against the direction of head yaw but occasionally moved with the direction of head yaw for brief periods of time depending on the specific motion of the subject's head.

Scene Position

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

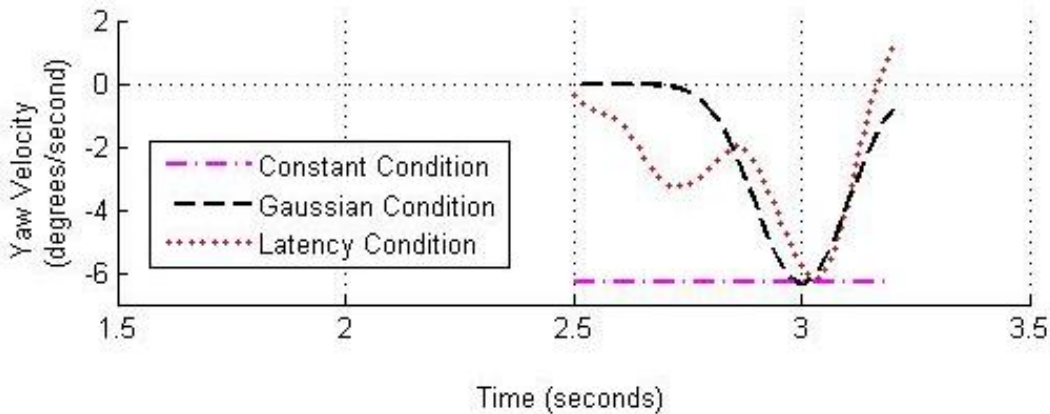


Figure 8.1: Example scene velocities for the three experimental conditions. In this example, each condition has the same peak scene velocity. The scene is visible only from 2.5 seconds to 3.2 seconds during the end of the head yaw.

8.2.3 Independent Variables

Scene Motion

The scene motion was controlled by one of three scene-motion types to simulate motion due to latency. Three conditions with the same peak scene velocity are graphed in Figure 8.1.

- **Constant scene velocity (Constant Condition)**—The scene is already moving when it appears and it moves at a constant velocity until it disappears. The experiment varied scene velocity between trials.
- **Gaussian-modeled scene velocity (Gaussian Condition)**—The scene moves with a Gaussian-velocity profile ($\sigma = 0.1 s$) centered at the end of the intended head yaw (at 3 seconds) such that scene velocity peaks near the time that scene velocity typically peaks in an HMD (because peak head acceleration typically occurs at the time head motion stops and scene velocity peaks near the time of peak head acceleration (Section 3.1.3)), due to head motion and latency. The experiment varied peak scene velocity between trials.
- **Latency-induced scene motion (Latency Condition)**—The scene motion is manipulated in real-time based on the subject’s measured head yaw and latency. The scene displacement θ_{t_0} is determined by Equation 3.1:

$$\theta_{t_0} = \Delta\phi_{t_0} = \phi_{t_0} - \phi_{t_0-\Delta t} \quad (3.1)$$

where ϕ is head displacement, in degrees, and Δt is latency, in seconds. The current

Table 8.1: Experiment 4 conditions. Each subject experienced all conditions. Three conditions resulted in four psychometric functions for ten ranges of head motion. For the Latency Condition, two psychometric functions (one in degrees per second and one in seconds) were computed from the same trials. 75% thresholds, 50% thresholds, and difference thresholds were extracted from the psychometric functions. The ranges of head motion for each decile varied per subject and, to a lesser extent, per condition.

		Peak Head-Yaw Acceleration (deg/s^2)										
		Decile										
		Measure	1	2	3	4	5	6	7	8	9	10
Condition	Constant (deg/s)	75% threshold										
		50% threshold										
		difference threshold										
	Gaussian (deg/s)	75% threshold										
		50% threshold										
		difference threshold										
	Latency (deg/s)	75% threshold										
		50% threshold										
		difference threshold										
	Latency (ms)	75% threshold										
		50% threshold										
		difference threshold										

scene velocity θ'_{t_0} is determined by Equation 3.9:

$$\theta'_{t_0} = \Delta\phi'_{t_0} = \phi'_{t_0} - \phi'_{t_0-\Delta t} \quad (3.9)$$

where ϕ' is head velocity, in degrees per second.

Since latency controls scene motion in this condition, both latency thresholds, in seconds, and scene-motion thresholds, in degrees per second, were determined from the same trials. The experiment varied latency between trials.

The left-most column of Table 8.1 shows the three conditions, including the two measures (in degrees per second and seconds) for the Latency Condition.

Intended Head-Yaw Amplitude

Visual cues before the trials, and auditory cues before and during the trials suggested subjects yaw their heads by a random amplitude between $\pm 0.5^\circ$ and $\pm 22.5^\circ$ from straight ahead..

8.2.4 Dependent Variables

Both head motion and scene motion due to latency vary continuously during head yaw. However, a single value for each is needed to generate the data points that establish the psychometric function, from which thresholds are extracted. Based on the mathematical analysis described in Section 3.1, I chose to use peak head-yaw acceleration and peak scene-velocity as measures of head motion and scene motion.

Peak Head-Yaw Acceleration

For each trial, I measured peak head-yaw acceleration for the time that the scene was visible (2.5 to 3.2 seconds). I divided peak head-yaw accelerations into deciles (ten bins with each bin containing an equal number of samples). I threw out the highest peak head-yaw accelerations in order to have an equal number of samples per decile and because my model predicts that latency thresholds vary the least for such trials.

Peak Scene Velocity

For each trial, I measured peak scene velocity for the time that the scene was visible (2.5 to 3.2 seconds).

Thresholds

For each subject, I computed psychometric functions for each of the deciles, conditions, and measures (a total of 40 psychometric functions per subject) from yes/no judgments with confidence ratings (Sections 4.1.4.4 and 4.3.1). In addition to extracting 75% thresholds from each psychometric function, I also extracted 50% and difference thresholds in order to compare my results with the results of previous experiments performed at NASA Ames Research Center (Ellis et al., 1999; Adelstein et al., 2003; Ellis et al., 2004; Mania et al., 2004). This comparison is discussed in Section 10.1.2.

8.3 Methods

As in Experiment 3, the method of constant stimuli determined random ordering of trials. As in Experiments 2 and 3, I used a yes/no confidence rating task; subjects judged if the scene appeared to move or not and then provided confidences in those judgments. Each subject completed 216 trials per condition for a total of 648 trials.

8.3.1 Head Yaw

Trials with a single head yaw from left-to-right or from right-to-left were randomly intermixed. The beginning of a trial was signaled by a tone; metronome beats occurred every second so that the subjects knew when to start (at 2 seconds) and stop (at 3 seconds) yawing their heads.

Visual and auditory cues suggested head yaw to range in amplitude between $\pm 0.5^\circ$ and $\pm 22.5^\circ$ from straight ahead. This resulted in moderate head motions with measured peak head-yaw accelerations less than $361^\circ/s^2$. Head accelerations in this study were not near the peak head accelerations possible for humans (fighter pilots can rotate their heads in the range of $2000^\circ/s^2$ with peaks of $6000^\circ/s^2$ (List, 1983)). My mathematical model as presented in Section 3.4.2 and Figure 3.6, predicts that latency thresholds vary most over the range of slow to moderate head rotations, so I measured thresholds for that range of peak head accelerations.

8.3.2 Scene Motion

For the reference scene presented between trials and catch trials, the scene remained stable in space. By using a world-fixed projector, I completely eliminated unwanted latency effects that would occur in an HMD (Section 4.1.1). Scene-motion was artificially injected into the system as described by the conditions.

For the Latency Condition, peak scene-velocities could not be determined in advance because scene motion depends on head motion. Thus, for the Latency Condition, I set latency for each trial and peak-scene velocity was a measured variable. Because peak velocity could not be controlled for this condition, I could not use predefined discrete levels of scene motion as used in Experiments 1-3. In order to maintain consistency between the conditions, I randomly set peak scene velocities on a continuous scale for the Constant and Gaussian Conditions.

The range of latencies and scene velocities for each participant was based on judgments and confidence ratings given during training: the stimulus range was adjusted until subjects gave responses distributed between 0% and 100% correct. As the experiment progressed, I increased or decreased the maximum latency and maximum scene-velocity values for individual subjects if necessary to generate the spread of data needed to generate the psychometric functions (thus not a true method of constant stimuli).

8.4 Participants

Eight subjects (six male; two female) participated. All subjects were naive to the experimental conditions. Three subjects had participated in one of the previous experiments. All subjects made a minimum of 216 judgments for each condition. Total time per subject, including consent form, instructions, training, experiment sessions, breaks, and debriefing, was less than six hours. Data for two male subjects were too noisy for useful analysis—their responses were largely independent of the amount of scene motion. Thus, data for six subjects (four male; two female) was analyzed.

8.5 Data Analysis

I removed trials with tracker problems, as described in Section 4.3.4.

For each trial, head-pose data logged from the tracker system yielded the peak head-yaw acceleration occurring during the time that the scene was visible.

Binning the measured peak head-yaw accelerations into ten deciles gave 21 valid trials for each decile. I chose to disregard the trials with the greatest peak head-yaw accelerations in order to have an equal number of samples per decile and because my model predicts that latency thresholds vary the least for such trials. The 21 data points for each decile were used to generate psychometric functions. Figure 8.2 shows data from a subject’s first peak head-yaw acceleration decile for the Latency Condition. The psychometric function and thresholds are also shown.

For each subject, I fit psychometric functions to the data from

- each decile for the peak scene-velocity measure (in degrees per second) of the Constant Condition,
- each decile for the peak scene-velocity measure (in degrees per second) of the Gaussian Condition,
- each decile for the peak scene-velocity measure (in degrees per second) of the Latency Condition, and
- each decile for the the latency measure (in seconds) of the Latency Condition.

I extracted 75% thresholds, 50% thresholds, and difference thresholds from each of these psychometric functions

False alarm rates (average confidence rating that there was scene motion when there was no scene motion) across all conditions were 11.5% or less for each of the six subjects.

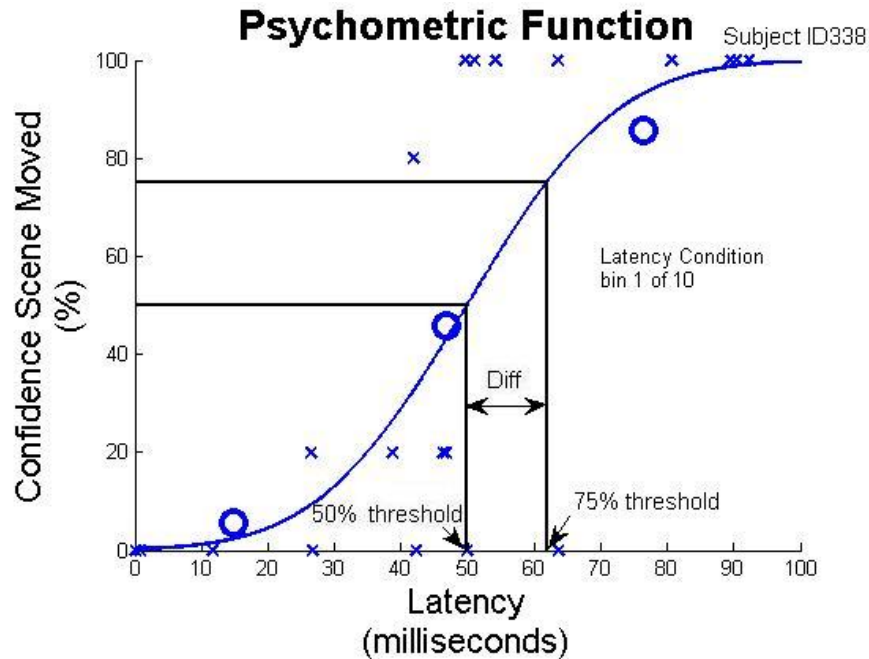


Figure 8.2: 21 confidence ratings versus latency, and the resulting psychometric function, for a range of a subject’s peak head-yaw acceleration (the eighth decile). This subject was fairly confident in her judgments in this case—she did not guess in her judgments; there were no confidences of 40% or 60%. The Os represent the means over seven adjacent samples. The 75% threshold, 50% threshold, and difference threshold are extracted from the psychometric function.

8.5.1 Latency Thresholds

8.5.1.1 Average Latency Thresholds

The mean latency 75% threshold across all subjects and all six deciles (60 75% thresholds) was 55.6 *ms* with a standard deviation of 23.4 *ms*. The lowest 75% threshold was 19.2 *ms* and the maximum 75% threshold was 154.1 *ms*.

The mean latency 50% thresholds across all subjects and all six deciles was 38.7 *ms* with a standard deviation of 18.4 *ms*. The lowest 50% threshold was -0.2 *ms* (a negative 50% threshold implies that the subject rated a stable scene to be moving with over 50% confidence) and the highest 50% threshold was 103.2 *ms*.

The mean latency difference threshold across all subjects and all six deciles was 16.9 *ms* with a standard deviation of 10.4 *ms*. The lowest difference threshold was 3.2 *ms* and the highest difference threshold was 60.5 *ms*. The minimum latency difference threshold of 3.2 *ms* over 60 difference thresholds suggest that end-to-end system latency in the 3 *ms* range is sufficiently low to be imperceptible in HMDs.

Table 8.2: Pearson correlations of latency thresholds to peak head-yaw accelerations using my model and a line. Bold values indicate the larger of the model and line correlations (absolute values).

		Subject						
		Correlation	ID331	ID332	ID333	ID336	ID337	ID338
Threshold	75% threshold	model	0.89	0.87	0.89	0.98	0.78	0.95
		line	-0.79	-0.74	-0.89	-0.86	-0.57	-0.89
	50% threshold	model	0.91	0.90	0.89	0.93	0.58	0.98
		line	-0.85	-0.72	-0.90	-0.84	-0.37	-0.88
	difference threshold	model	0.52	0.80	0.79	0.50	0.65	0.68
		line	-0.33	-0.76	-0.74	-0.40	-0.55	-0.71

8.5.1.2 One Subject’s Latency Thresholds for Different Peak Head-Yaw Accelerations

For a single subject, Figures 8.3(a), 8.4(a), and 8.5(a) show measured latency thresholds for 75% thresholds, 50% thresholds, and difference thresholds, respectively. The vertical lines represent the boundaries of peak head-yaw acceleration deciles.

8.5.1.3 Fitting the Model to the Measured Latency Thresholds

I fit curves with the form of my model to the measured latency thresholds. Figures 8.3(a), 8.4(a), and 8.5(a) show these best-fit curves (the thickest curve in each figure) for a single subject. Figure 8.6 shows latency threshold curves for all six subjects. Table 8.2 shows Pearson correlations of how well the threshold data fits the model for each individual subject; Pearson correlations were all greater than or equal to 0.50 (sign tests across subjects for each threshold type: $p_{one-tail} < 0.016$). No correction was required for $\alpha = 0.05$ since the largest p-value was less than 0.05. Table 8.2 also shows linear correlations of thresholds to peak head-yaw accelerations.

In order to determine if the thresholds fit the form of my model better than a line, I compared the model correlations to the absolute value of the linear correlations using Wilcoxon matched-pairs signed-rank tests. I used one-tail tests since I expected the data to fit my model well. The thresholds fit my data better than a line for 75% thresholds ($S_6 = 0, p_{one-tail} < 0.016$), 50% thresholds, ($S_6 = 1, p_{one-tail} < 0.031$), and difference thresholds ($S_6 = 1, p_{one-tail} < 0.031$). No correction was required for $\alpha = 0.05$ since the largest p-value was less than 0.05.

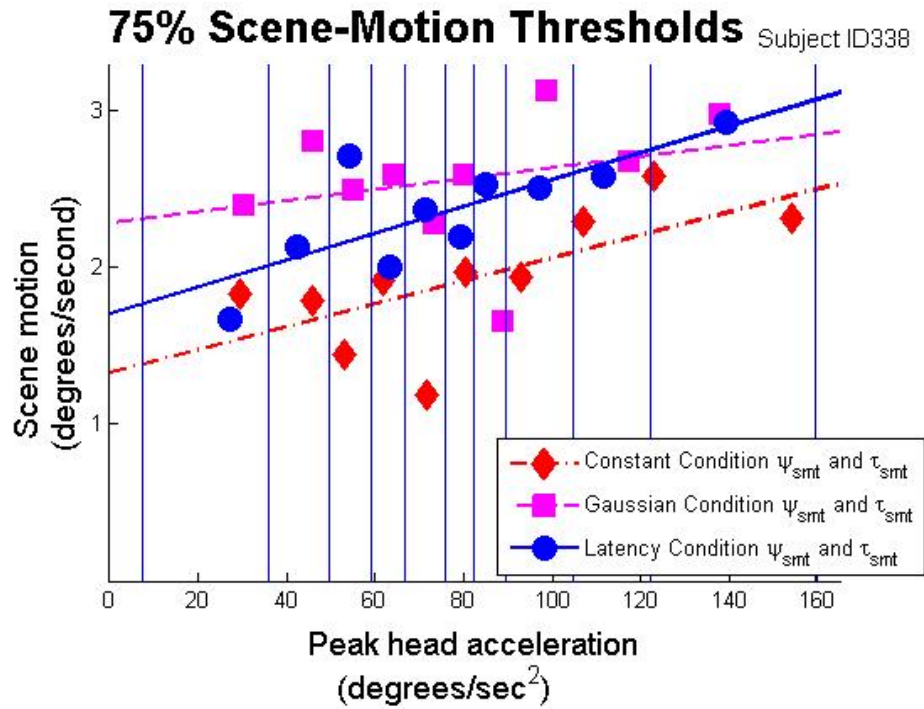
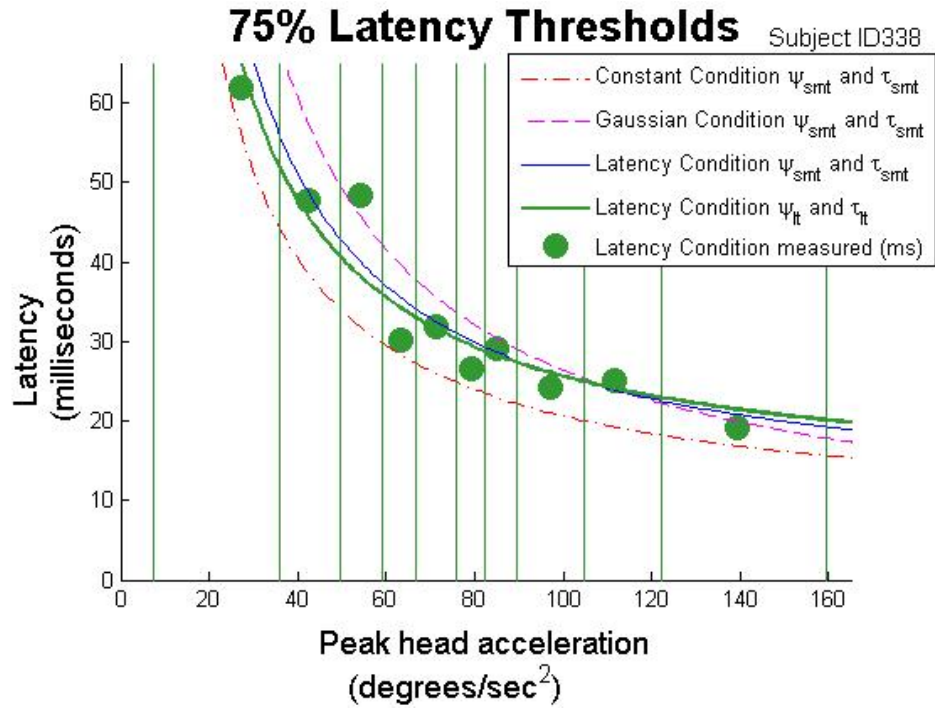


Figure 8.3: 75% thresholds for a single subject. The vertical lines represent the boundaries of the Latency Condition deciles.

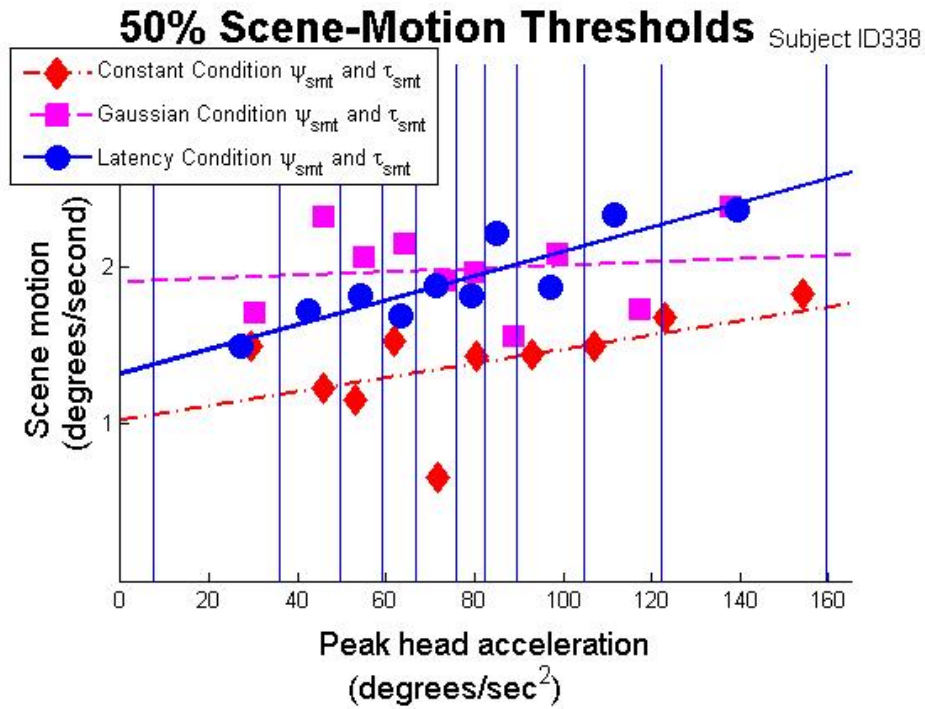
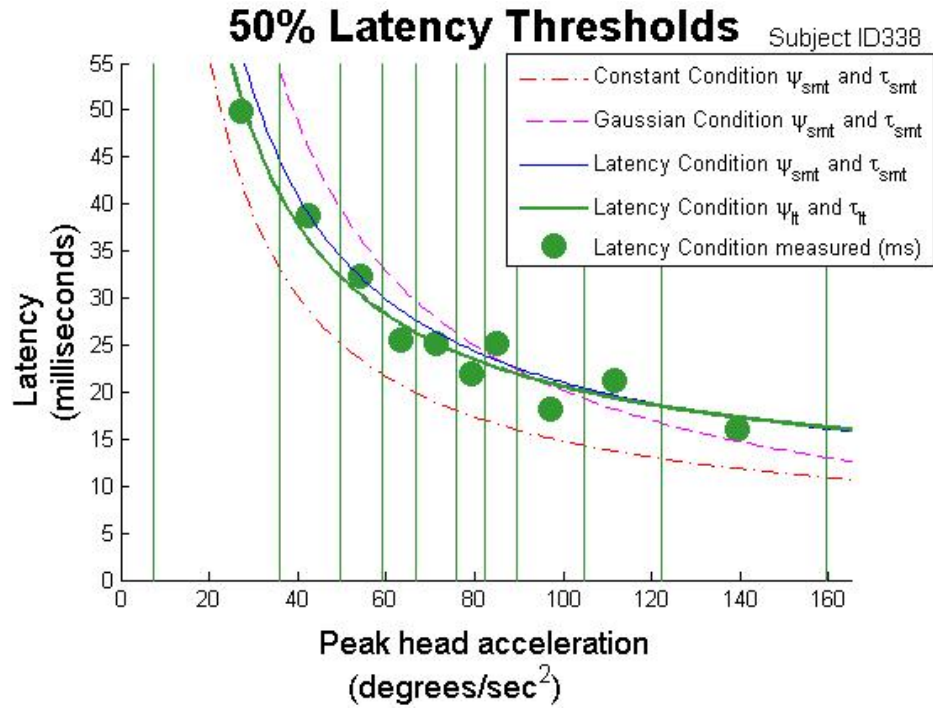


Figure 8.4: 50% Thresholds for a single subject. The vertical lines represent the boundaries of the Latency Condition deciles.

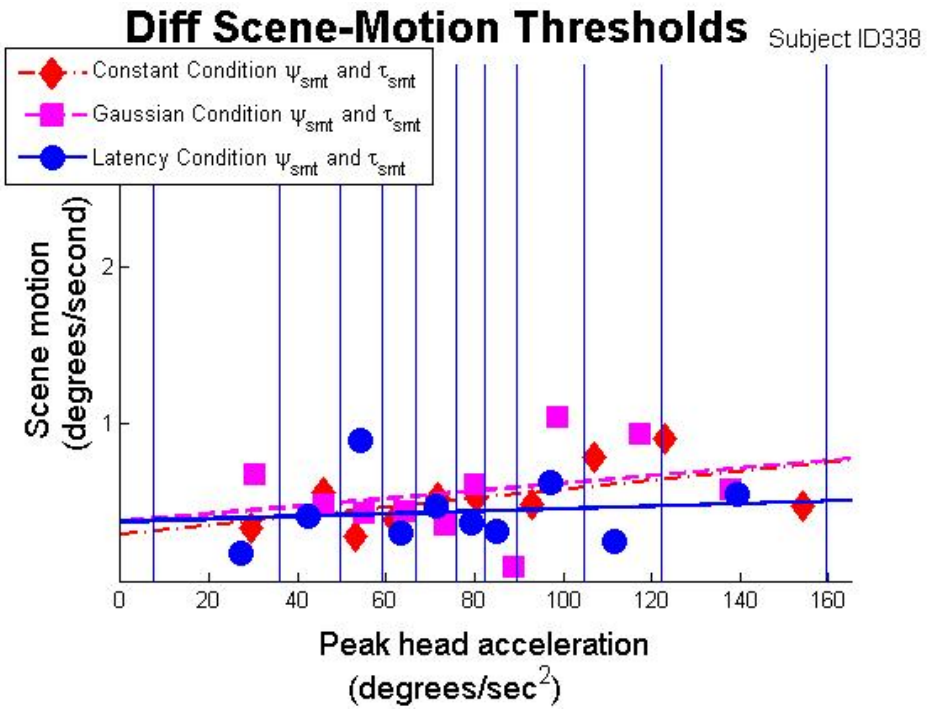
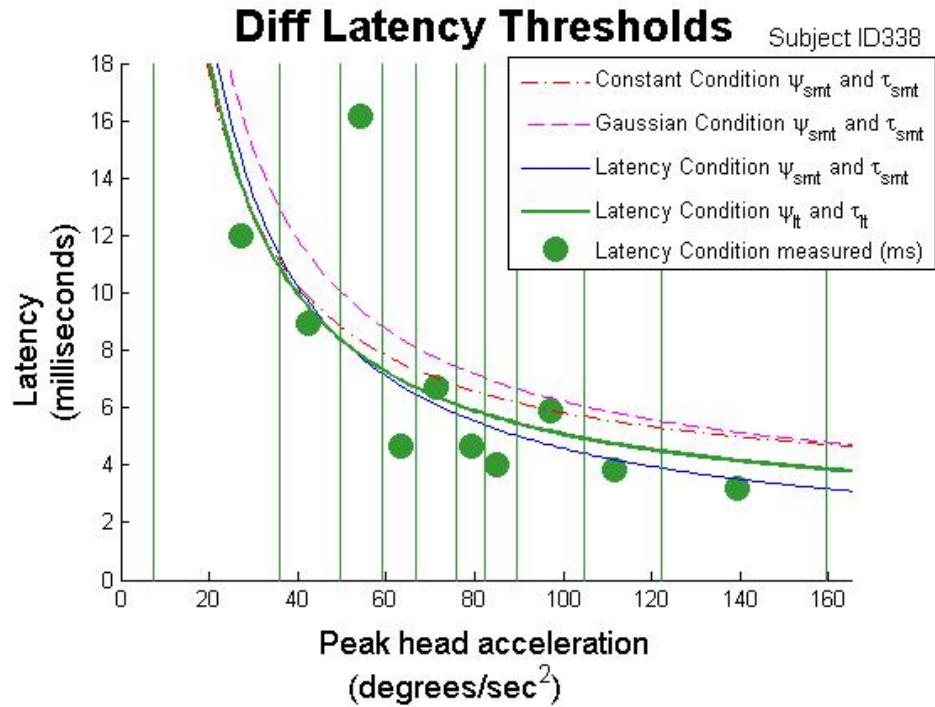
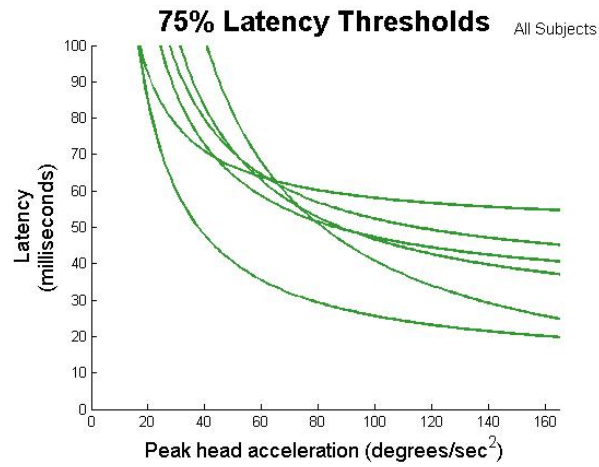
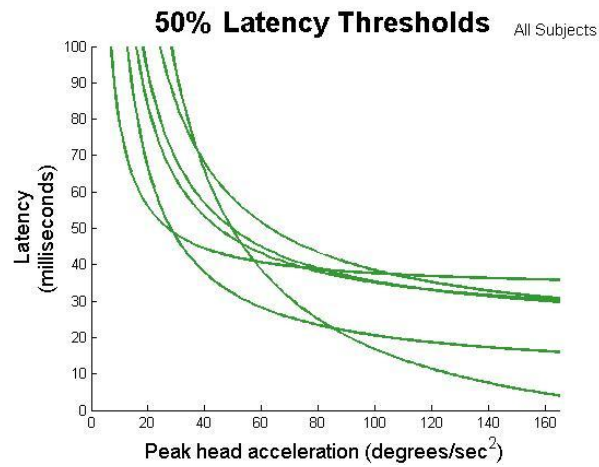


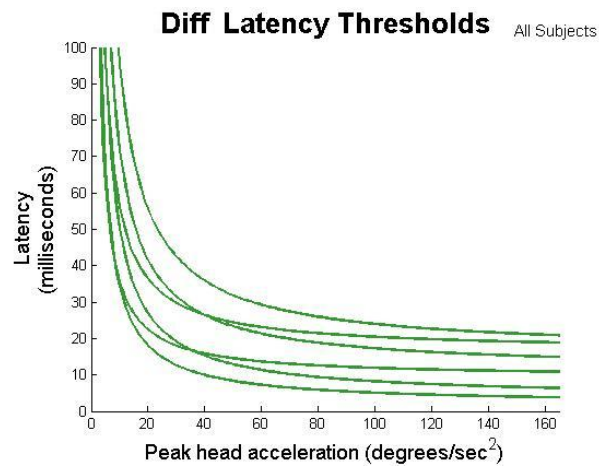
Figure 8.5: Difference thresholds for a single subject. The vertical lines represent the boundaries of the Latency Condition deciles.



(a) Latency 75% thresholds.



(b) Latency 50% thresholds.



(c) Latency difference thresholds.

Figure 8.6: Latency thresholds for all six subjects. The curves were determined by the best fit of my model to measured latency thresholds.

8.5.2 Scene-Motion Thresholds

For each of the subjects, 30 psychometric functions (10 deciles \times 3 condition) were computed. From these psychometric functions, 75% thresholds, 50% thresholds, and difference thresholds extracted.

8.5.2.1 Average Scene-Motion Thresholds

The mean scene-motion 75% threshold across all six subjects, all deciles, and all three conditions (180 75% thresholds) was $4.2^\circ/s$ with a standard deviation of $1.5^\circ/s$. The mean scene-motion 50% threshold was $2.9^\circ/s$ with a standard deviation of $1.2^\circ/s$. The mean scene-motion difference threshold was $1.3^\circ/s$ with a standard deviation of $0.8^\circ/s$.

8.5.2.2 One Subject's Scene-Motion Thresholds For Different Peak Head-Yaw Accelerations

Figures 8.3(b), 8.4(b), and 8.5(b) show scene-motion thresholds for a single subject for the three conditions described in section 8.2.3. Linear regressions of the scene-motion thresholds are also shown.

8.5.2.3 Comparison of Scene-Motion Thresholds Between Conditions

Figure 8.7 shows all six individuals' Constant Condition scene-motion thresholds subtracted from their Latency Condition scene-motion thresholds. Likewise, Figure 8.8 shows Constant Condition scene-motion thresholds subtracted from Gaussian Condition scene-motion thresholds and Figure 8.9 shows Gaussian Condition scene-motion thresholds subtracted from Latency Condition scene-motion thresholds. There is a pattern that the subjects' Latency Condition thresholds and Gaussian Condition thresholds are greater than their Constant Condition thresholds. There is no evident pattern of differences between the Latency Condition and the Gaussian Condition.

Due to the large number of samples (60 psychometric functions per condition), I used parametric tests to compare between conditions. A two-way ANOVA showed that scene-motion thresholds were statistically significantly affected by peak head-yaw acceleration decile (each of 75%, 50%, and difference thresholds: $F_{9,150} \geq 2.88, p < 0.01$) and by condition (each of 75%, 50%, and difference thresholds: $F_{2,150} \geq 4.91, p < 0.01$). No interaction effects were found.

To further investigate the data, I compared the differences between conditions by matching scene-motion thresholds across subjects and decile (6 subjects \times 10 deciles for a total of 60 samples per condition). I then performed three pair-wise t-tests for each of the three threshold types with $\alpha = 0.05/9 = 0.0056$ (Bonferroni correction). The Latency Condition thresholds

Latency Condition - Constant Condition Scene Motion Thresholds All Subjects

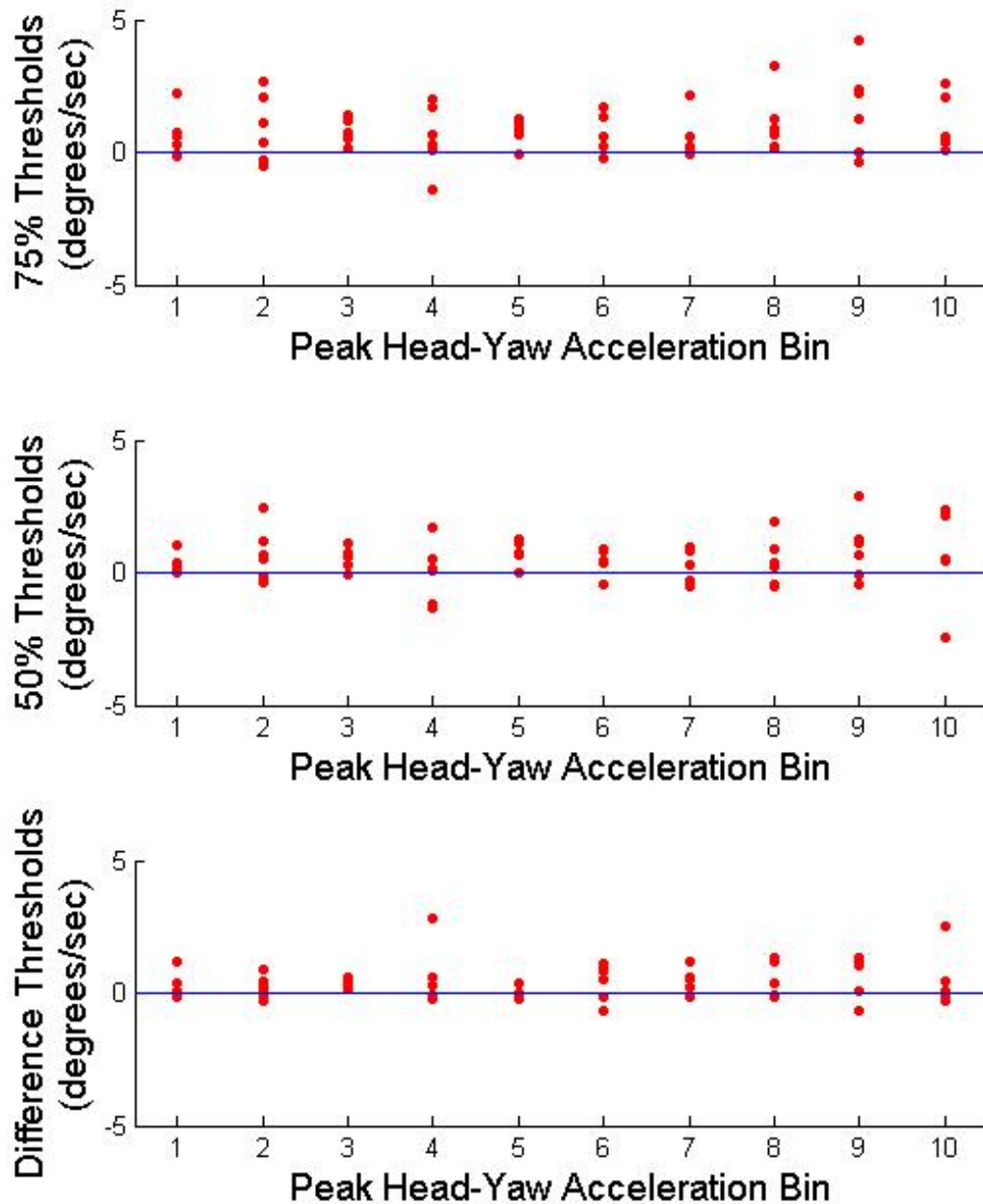


Figure 8.7: Individuals' differences of Latency Condition scene-motion thresholds and Constant Condition scene-motion thresholds. There is a pattern for the Latency Condition thresholds to be greater than the Constant Condition thresholds.

Gaussian Condition - Constant Condition Scene Motion Thresholds All Subjects

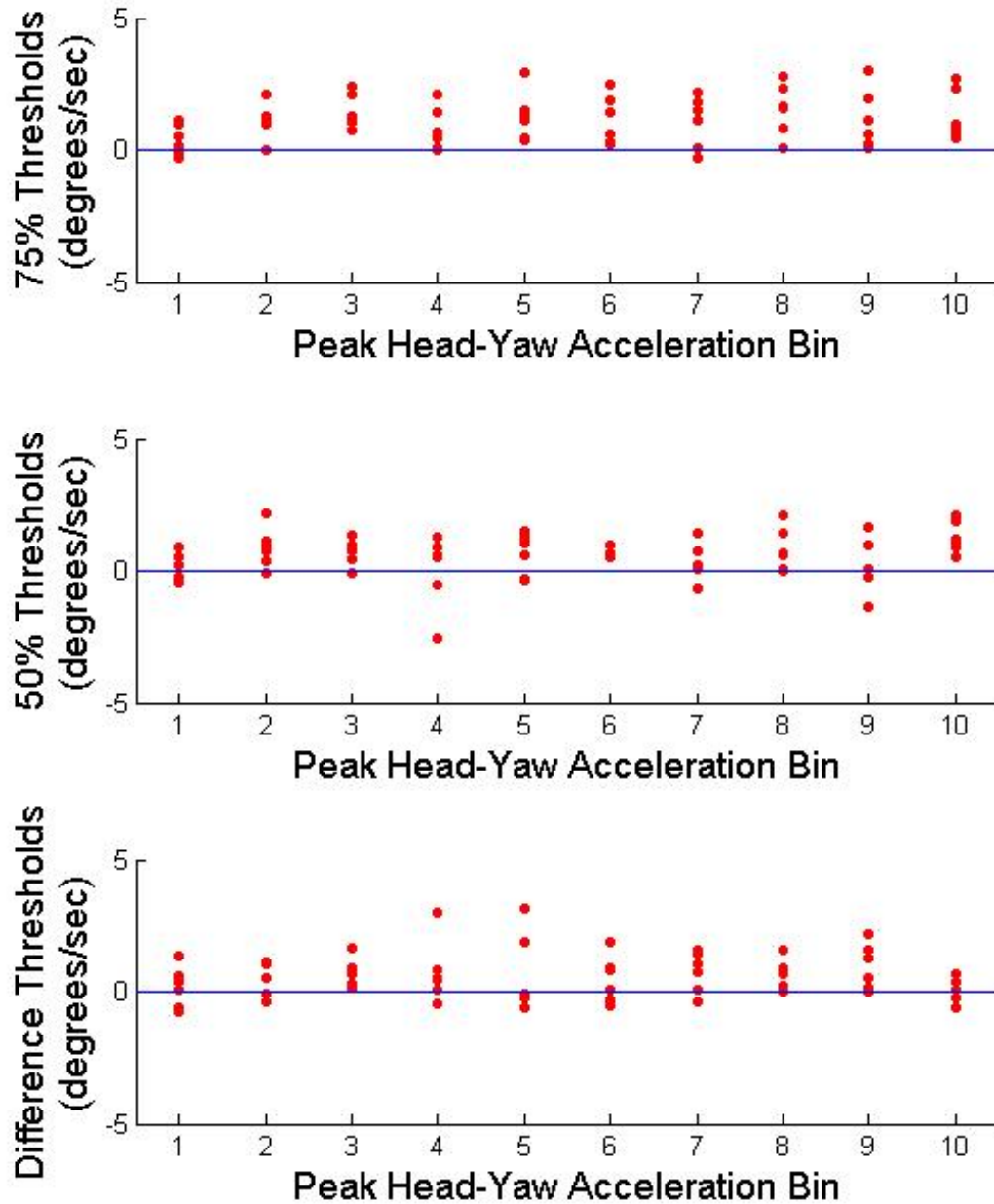


Figure 8.8: Individuals' differences of Gaussian Condition scene-motion thresholds and Constant Condition scene-motion thresholds. There is a pattern for the Gaussian Condition thresholds to be greater than the Constant Condition thresholds.

Latency Condition - Gaussian Condition Scene Motion Thresholds All Subjects

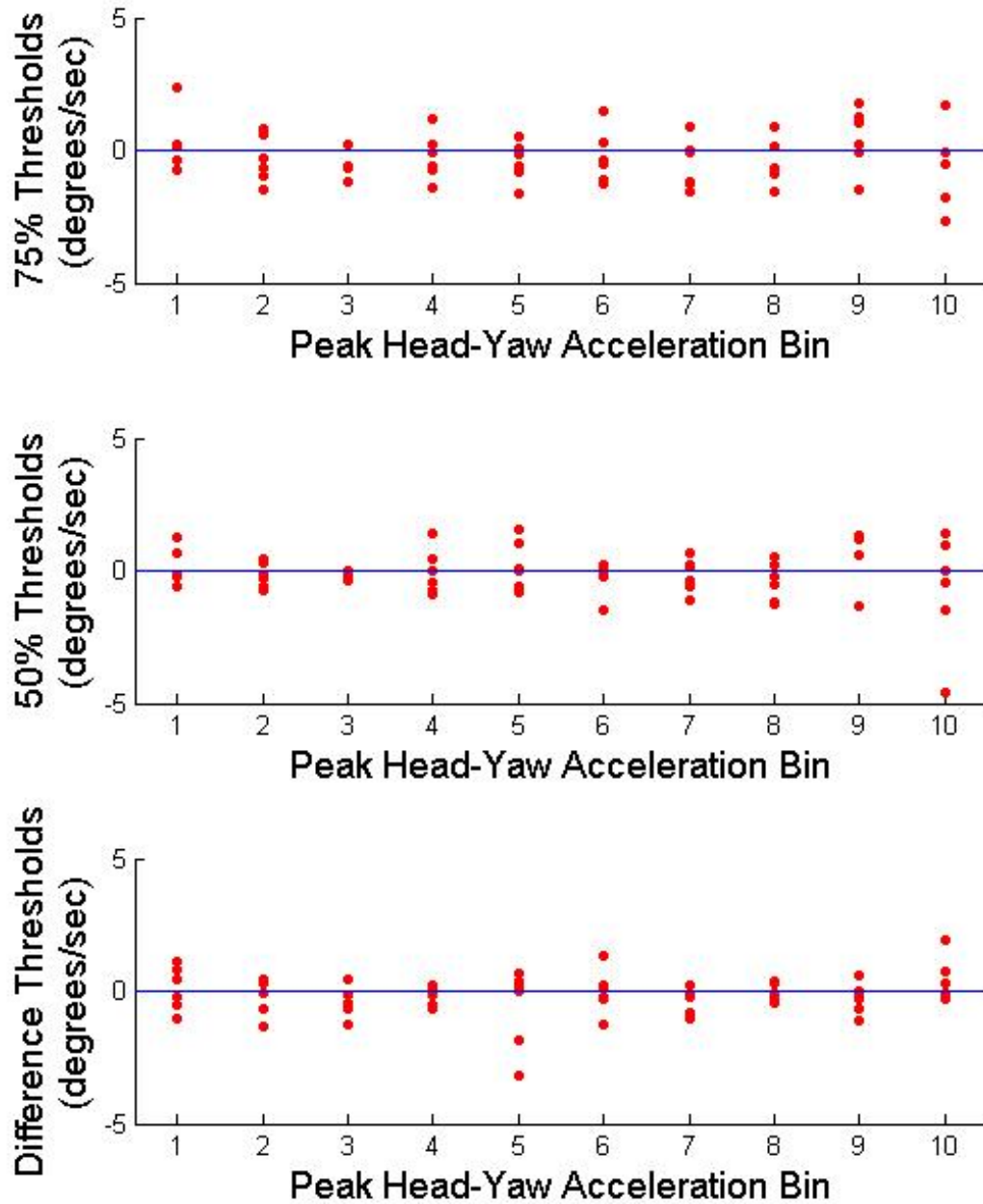


Figure 8.9: Individuals' differences of Latency Condition scene-motion thresholds and Gaussian Condition scene-motion thresholds. There is no consistent pattern that one condition has greater thresholds than the other condition.

were greater than (75% thresholds: $0.9^\circ/s$, 50% thresholds: $0.5^\circ/s$, difference thresholds: $0.4^\circ/s$) the Constant Condition thresholds (each of 75%, 50%, and difference thresholds: $t_{59} \geq 4.2$, $p_{two-tail} < 0.001$). The Gaussian Condition thresholds were also greater than (75% thresholds: $1.1^\circ/s$, 50% thresholds: $0.6^\circ/s$, difference thresholds: $0.6^\circ/s$) than the Constant Condition thresholds (each of 75%, 50%, and difference thresholds: $t_{59} \geq 5.1$, $p_{two-tail} < 0.001$). I suspect the Constant Condition resulted in lower thresholds because peak scene velocity occurs over the entire display interval instead of for a shorter time as in the other conditions.

Since I found no statistically significant differences of scene-motion thresholds between the Gaussian and Latency Conditions, I computed confidence intervals and the corresponding ranges of statistical equivalence. The 95% confidence intervals of the differences (Latency Condition thresholds subtracted from the Gaussian Condition thresholds) were ($-0.01^\circ/s$, $0.51^\circ/s$) for 75% thresholds, ($-0.17^\circ/s$, $0.33^\circ/s$) for 50% thresholds, and ($-0.04^\circ/s$, $0.36^\circ/s$) for difference thresholds. Confidence intervals in terms of percentage of average scene-motion thresholds (from Section 8.5.2.1) were (-0.2% 12%) for 75% thresholds, (-6% , 11%) for 50% thresholds, and (-3% , 28%) for difference thresholds. Thus, the Gaussian Condition and Latency Conditions thresholds were statistically equivalent within $\pm 0.51^\circ/s$ and within $\pm 28\%$ of the average scene-motion thresholds (each $p < 0.05$). This small range suggests that, on average, the Gaussian Condition scene-motion thresholds estimate the Latency Condition scene-motion thresholds.

8.5.3 Estimation of Latency Thresholds with Scene-Motion Thresholds

For each subject and condition, lines were fit to the scene-motion thresholds as a function of peak head-yaw acceleration using standard linear regression. The y-intercept ψ_{smt} , in degrees per second, and slope τ_{smt} , in seconds, of the lines were used as input for my mathematical model. The form of the model is

$$\Delta t = \tau + \psi \left(\frac{1}{\phi''} \right) \quad (3.20)$$

where Δt is the latency threshold, in seconds, ϕ'' is peak head acceleration, in degrees per second, τ is an offset parameter, in seconds, and ψ is a scale parameter, in degrees per second.

Figures 8.3, 8.4, and 8.5 show these scene-motion threshold lines and corresponding transformed latency threshold curves.

The input parameters can also be obtained by fitting the model to directly measured

Table 8.3: 95% confidence intervals of the differences between the two methods of determining τ and ψ . Confidence intervals are given in milliseconds and degrees per second as well as percentage of the average thresholds over all subjects (from Sections 8.5.1.1 and 8.5.2.1).

Parameter	Threshold		
	75%	50%	Diff
$\tau_{lt} - \tau_{smt}$	(-10.6 <i>ms</i> , 3.0 <i>ms</i>) (-19%, 5%)	(-8.5 <i>ms</i> , 6.1 <i>ms</i>) (22%, 16%)	(-3.0 <i>ms</i> , 4.6 <i>ms</i>) (18%, 27%)
$\psi_{lt} - \psi_{smt}$	(-0.05°/s, 0.94°/s) (-1%, 22%)	(-0.30°/s, 0.76°/s) (10%, 26%)	(-0.22°/s, 0.38°/s) (17%, 29%)

latency thresholds. I call the parameters obtained in this manner τ_{lt} and ψ_{lt} .

For the Latency Condition, I wanted to know how well τ_{smt} matched τ_{lt} and how well ψ_{smt} matched ψ_{lt} .

I performed Wilcoxon matched-pairs signed-rank tests to compare differences of parameter values τ_{smt} and τ_{lt} , and differences of ψ_{smt} and ψ_{lt} . I found no statistically significant differences.

Since I found no statistically significant differences between the two methods of determining the parameter values τ and ψ , I computed 95% confidence intervals of the differences and the corresponding ranges of statistical equivalence. Table 8.3 shows the confidence intervals. The parameter values τ_{smt} and τ_{lt} were statistically equivalent within ± 10.6 *ms* (and within $\pm 27\%$ of the average latency thresholds) and the parameter values ψ_{smt} and ψ_{lt} were statistically equivalent within $\pm 0.94^\circ/s$ (and within $\pm 29\%$ of the average scene-motion thresholds) (each $p < 0.05$). These results suggest that latency thresholds can be estimated from scene-motion thresholds. I suspect these ranges of equivalence would be smaller if more data was collected.

8.6 Summary of Results

I found the following:

- Latency thresholds as a function of peak head-yaw acceleration correlate (each $\rho \geq 0.5$) with the form of my model (Equation 3.20). (Section 8.5.1.3)
- Directly-measured latency thresholds correlate with the form of my inverse model better than with a line. (Section 8.5.1.3)
- Scene-motion thresholds for the Latency Condition can be used to estimate parameters for my model of latency thresholds. Scene-motion thresholds estimate the offset parameter τ within 27% of the average latency thresholds. Scene-motion thresholds

estimate the scale parameter ψ within 29% of the average scene-motion thresholds. (Section 8.5.3)

- Scene-motion thresholds for the Constant Condition are lower than scene-motion thresholds for the Latency Condition. Scene-motion thresholds for the Gaussian Condition are equivalent to scene-motion thresholds for the Latency Condition within 28% the average scene-motion thresholds. (Section 8.5.2.3)

8.7 Recommendations

With slow head accelerations, high amounts of latency can occur without users perceiving scene motion. As head acceleration increases, users are more likely to notice scene motion resulting from latency.

Latency requirements can have larger tolerances for applications in which users make only slow head motions. As head acceleration increases, latency requirements level off to near constant value for head accelerations in the $100^\circ/s^2$ range. My data suggests latency requirements for the most-sensitive users should be in the 3 *ms* range.

The variability of scene-motion thresholds and latency thresholds among subjects makes it difficult to determine conservative thresholds (i.e., minimum thresholds) when conducting and analyzing experiments across multiple subjects. Requirements should be determined for subjects with the lowest thresholds, rather than averaging across all subjects, so that thresholds will be below perception of all users.

Although latency thresholds obtained indirectly by measuring scene-motion thresholds estimate latency thresholds reasonably well, I recommend measuring latency thresholds directly—it is just as easy using a simulated-latency system and is more accurate.

Chapter 9

Experiment 5: Two Case Studies

Data from Experiment 4 were quite noisy. I wanted to more precisely measure scene-motion thresholds and latency thresholds from the subjects with the lowest thresholds.

In Experiment 5, I collected an extensive amount of data from two individual subjects under conditions very similar to those of Experiment 4. Moving or non-moving scenes were presented to subjects as they yawed their heads. Subjects then judged if they believed the scene moved or not, and rated their confidence in those judgments.

The primary purpose of this experiment was to explore further the latency-thresholds model

$$\Delta t = \tau + \psi \left(\frac{1}{\phi''} \right) \quad (3.20)$$

where Δt is the latency threshold, in seconds, ϕ'' is peak head-yaw acceleration, in degrees per second squared, τ is an offset parameter, in seconds, and ψ is a scale parameter, in degrees per second.

I also compared thresholds derived from yes/no judgments with confidence ratings to yes/no judgments without confidence ratings.

The two subjects were myself and SubjectID 338 from Experiment 4.

9.1 Experimental Design

The experiment was nearly identical to Experiment 4, except that there were only two subjects, more data was collected, and scene luminance was different. The experiment was a repeated-measures constant-stimuli design.

Scene-motion thresholds, in degrees per second, and latency thresholds, in seconds, were measured for different types of scene motion. Measured thresholds were 75% thresholds, 50% thresholds, and difference thresholds. Table 9.1 shows the experimental design.

Table 9.1: Experiment 5 conditions for the two subjects. Three conditions resulted in 40 psychometric functions for 30 ranges of head motion. For the Latency condition, 20 psychometric functions (10 with the dependent variable being degrees per second and 10 being seconds) were computed. 75% thresholds, 50% thresholds, and difference thresholds were extracted from the psychometric functions. The ranges of head motion for each decile varied per subject and per condition.

		Peak Head-Yaw Acceleration (deg/sec^2)										
		Decile										
		Measure	1	2	3	4	5	6	7	8	9	10
Condition	Constant ¹ (deg/s)	75% threshold										
		50% threshold										
		difference threshold										
	Gaussian (deg/s)	75% threshold										
		50% threshold										
		difference threshold										
	Latency (deg/s)	75% threshold										
		50% threshold										
	Latency (ms)	difference threshold										
		75%										
		50% threshold										
		difference threshold										

9.1.1 Controlled Variables

The scene foreground luminance was $1.2\text{ cd}/m^2$ and the scene background luminance was $0.008\text{ cd}/m^2$ resulting in a contrast of 150. Otherwise, conditions were similar to Experiment 4. Subjects yawed their heads to the timing of metronome beats; they started to yaw their heads at two seconds and stopped at three seconds. The scene was visible for the end of the head yaw, over the interval from 2.5 to 3.2 seconds.

9.1.2 Independent Variables

Independent variables in Experiment 5 were the same as in Experiment 4, as described in Section 8.2.3. In summary:

- The scene moved against the direction of the head yaw and was controlled by one of three conditions to simulate motion due to latency:

- **Constant scene velocity (Constant Condition)** (subject IDjj only)—The

¹Subject IDjj only.

scene moved with a constant velocity against the direction of the head yaw. The experiment varied scene velocity between trials.

- **Gaussian scene velocity (Gaussian Condition)**—The scene moved with a Gaussian velocity profile such that scene velocity peaks near the time that head deceleration typically peaks (at three seconds, when the head stops) in an HMD due to head motion and latency. The experiment varied peak scene velocity between trials.
- **Latency-induced scene velocity (Latency Condition)**—The scene motion was manipulated in real-time based on the user’s measured head motion and injected latency. Since latency controls scene motion in this condition, both latency thresholds, measured in seconds, and scene-motion thresholds, measured in degrees/second, were determined from the same trials. The experiment varied latency between trials.

9.1.3 Dependent Variables

The dependent variables in Experiment 5 were identical to Experiment 4, as described in Section 8.2.4. In summary:

- The stimulus to be detected was scene motion. The particular type of scene motion was defined by one of the three conditions.
- Thresholds were measured for deciles of peak head-yaw accelerations (10 peak head-yaw acceleration bins with an equal number of trials per bin).
- Scene-motion thresholds, in degrees per second, were obtained for each of the three conditions.
- Latency thresholds, in seconds, were also obtained for the Latency Condition.
- Measured thresholds were 75% thresholds, 50% thresholds, and difference thresholds.

9.2 Methods

The experimental task was identical to Experiment 4 (Section 8.3). In summary:

- Subjects started to yaw their heads at 2.0 seconds after a metronome signaled the start of the trial and stopped yawing their heads at 3.0 seconds.
- Suggested head yaw ranged in amplitude between $\pm 0.5^\circ$ and $\pm 22.5^\circ$ from straight ahead over a period of one second.

- The scene was presented at the end of the head yaw, from 2.5 to 3.2 seconds
- The type of scene motion was controlled according to the conditions.
- Peak scene velocities and latencies were randomly chosen on a continuous scale.
- The range of peak scene velocities and latencies for each participant was chosen based on judgments and confidences given during training. The stimulus range was adjusted until users gave responses distributed between 0% and 100% correct. As the experiment progressed, I occasionally increased or decreased the maximum scene-velocity value for individual subjects if necessary to generate the spread of data needed to generate the psychometric functions.
- The horizontal starting position of the scene for each presentation was set to be half-way between 0° (straight ahead) and the end of the suggested head yaw, so that the scene would be presented approximately in the center of the field of view.

9.3 Participants

Two subjects participated.

9.3.1 Subject IDjj

Finding quality subjects willing to volunteer 15 to 20 hours proved to be difficult. I required that subjects could efficiently and consistently perform the task of yawing their heads. I also wanted the most-motivated and most-sensitive subjects. For these reasons, investigators often use themselves as subjects for psychophysics studies (as done in the NASA Ames latency Experiments (Adelstein et al., 2003; Mania et al., 2004; Adelstein et al., 2005; Li et al., 2006)). I used myself as a subject. I analyzed 3556 trials (1203 trials for the Constant Condition, 1195 for the Gaussian Condition and 1158 trials for the Latency Condition).

9.3.2 Subject ID338

I wanted to collect more data from at least one sensitive subject, in addition to myself. I asked top performers from the previous experiments to return. Subject ID338 agreed to return for 15-20 hours of data collection. Subject ID338 was female, age 20. She had music experience since age six. Her only experience with IVEs was her participation in the previous experiments. She played a small amount of video games (a maximum of 2-5 hours per week over the past two years). She had some moving-target-shooting practice as a child, otherwise she had no experience in tracking objects or judging scene motion.

I analyzed 3746 trials (1868 for the Gaussian Condition and 1878 trials for the Latency Condition) from Subject ID338. Since the Constant Condition did not estimate the Latency Condition well in Experiment 4 (Section 8.5.2.3), I chose to not collect data for the Constant Condition, allowing more trials for the Gaussian and Latency Conditions.

9.4 Data Analysis

I removed trials with tracker problems, as done in Experiment 4 and described in Section 4.3.4.

As in Experiment 4, conditions and stimulus values were randomly selected to prevent bias among conditions and to minimize ordering effects. I sorted the data into peak head-yaw acceleration deciles. Trials with the highest peak head-yaw accelerations were removed to yield an equal number of trials per decile. This resulted in a maximum removal of 9 trials per condition.

Psychophysics does not normally use confidence ratings. In Experiments 2-5, I used confidence ratings (Sections 4.1.4.4 and 4.3.1) in addition to yes/no judgments in order to obtain more data for analysis. Since I collected a large amount of data for Experiment 5, I was also able to compute thresholds from yes/no judgments alone.

Figures 9.1(a) and 9.2(a) (all figures for this chapter are presented at the end of this chapter) show the two subjects' data for decile four from the Latency Condition with the psychometric function fit to the yes/no judgments *with* confidence ratings. Figures 9.1(b) and 9.2(b) show data from the same trials with the psychometric function fit to the yes/no judgments *without* confidence ratings.

For Subject IDjj, Figures 9.3, 9.4, and 9.5 show thresholds obtained from yes/no judgments *with* confidence ratings, and 9.6, 9.7, and 9.8 show thresholds obtained from yes/no judgments *without* confidence ratings. Subject IDjj's false alarm rate (average confidence rating there was scene motion when there was no scene motion) across all conditions was 15.7%.

For Subject ID338, Figure 9.9, 9.10, and 9.11 show thresholds obtained from yes/no judgments *with* confidence ratings, and 9.12, 9.13, and 9.14 show thresholds obtained from yes/no judgments *without* confidence ratings. Subject ID338's false alarm rate across all conditions was 7.0%.

9.4.1 Scene-Motion Thresholds—Correlations

Scene-motion thresholds appear to be linear as a function of peak head-yaw acceleration; no pattern is apparent that matches the data better than a line. For both subjects, all Pearson correlations are better than 0.75 for thresholds obtained using yes/no judgments

Table 9.2: Pearson correlations of the linear relationship between scene-motion thresholds and peak head-yaw accelerations.

(a) Subject IDjj, yes/no judgments *with* confidence ratings.

Condition	Threshold		
	75%	50%	Diff
Constant	0.90	0.82	0.92
Gaussian	0.88	0.76	0.90
Latency	0.96	0.84	0.91

(b) Subject IDjj, yes/no judgments *without* confidence ratings.

Condition	Threshold		
	75%	50%	Diff
Constant	0.81	0.54	0.86
Gaussian	0.87	0.83	0.23
Latency	0.86	0.83	0.82

(c) Subject ID338, yes/no judgments *with* confidence ratings.

Condition	Threshold		
	75%	50%	Diff
Gaussian	0.91	0.86	0.81
Latency	0.94	0.92	0.90

(d) Subject ID338, yes/no judgments *without* confidence ratings.

Condition	Threshold		
	75%	50%	Diff
Gaussian	0.96	0.87	0.76
Latency	0.91	0.90	0.81

with confidence ratings (Table 9.2 (a and c)). Pearson correlations are 0.23 or better with thresholds obtained using yes/no judgments alone (Table 9.2 (b and d)). The smaller correlation variability for the case of yes/no judgments with confidence ratings is presumably due to the additional information taken into account, resulting in less noise.

9.4.2 Minimum Latency Thresholds

Latency difference thresholds for the most sensitive subjects can be used for conservative latency requirements.

For Subject IDjj, the minimum-measured latency difference threshold was 10.9 *ms* when using yes/no judgments with confidence ratings and 5.8 *ms* when using yes/no judgments alone. For Subject ID338, the minimum-measured latency difference threshold was 5.8 *ms* when using yes/no judgments with confidence ratings and 4.3 *ms* when using yes/no judgments alone. These values were greater than the minimum-measured latency difference threshold of 3.2 *ms* for Subject ID338 obtained from Experiment 4 (and all other subjects, who had higher minimum thresholds), which used yes/no judgments with confidence ratings. Whereas the lower minimum-measured latency threshold from Experiment 4 may have occurred by chance due to the smaller sample size, the higher minimum-measured latency threshold from Experiment 5 may have been because responses varied over several days due to training effects and/or other unknown day-to-day variance (difference thresholds are a measure of variance). Thus, I suggest that latency requirements be the more conservative

Table 9.3: Pearson correlations of latency thresholds to peak head-yaw accelerations using my inverse model and a line. In each case, the measured latency thresholds correlate with my model better than they correlate with a linear model (larger absolute values of Pearson correlations).

(a) Subject IDjj, yes/no judgments *with* confidence ratings.

		Threshold		
		75%	50%	Diff
Type	model	0.99	0.94	0.97
	line	-0.88	-0.84	-0.87

(b) Subject IDjj, yes/no judgments *without* confidence ratings.

		Threshold		
		75%	50%	Diff
Type	model	0.97	0.91	0.82
	line	-0.90	-0.83	-0.79

(c) Subject ID338, yes/no judgments *with* confidence ratings.

		Threshold		
		75%	50%	Diff
Type	model	0.99	0.99	0.96
	line	-0.81	-0.81	-0.79

(d) Subject ID338, yes/no judgments *without* confidence ratings.

		Threshold		
		75%	50%	Diff
Type	model	0.98	0.98	0.92
	line	-0.80	-0.82	-0.70

value of 3.2 *ms*.

9.4.3 Latency Thresholds—Correlations

I fit curves with the form of my model, $\Delta t = \tau + \psi \left(\frac{1}{\phi^\pi} \right)$, to the measured latency thresholds. The bottom of Figures 9.3-9.14, shows these best-fit curves (the thickest curve in each figure). Table 9.3 shows Pearson correlations of how well the latency threshold data fits the model; Pearson correlations were 0.94 or better for all latency thresholds obtained using yes/no judgments with confidence ratings and 0.82 or better for all latency thresholds obtained using yes/no judgments alone. For each threshold type, model correlations were better when confidence ratings were taken into account (each ρ difference ≥ 0.03).

Table 9.3 also shows linear correlations of thresholds to peak head-yaw accelerations. All directly-measured latency thresholds correlate with my inverse model better than with a line (larger absolute values of Pearson correlations).

9.4.4 Estimating Model Parameters from Scene-Motion Thresholds

As can be seen in the bottom of Figures 9.3 - 9.14, the latency threshold curves resulting from input parameters τ_{smt} (the slope of the line fit to the Latency Condition scene-motion

Table 9.4: Model Parameters τ and ψ . τ is in milliseconds and ψ is in degrees per second.

(a) Subject IDjj, yes/no judgments *with* confidence ratings.

Parameter	Threshold		
	75%	50%	Diff
τ_{lt}	12.8	7.5	5.3
τ_{smt}	14.7	8.0	6.7
$\tau_{lt} - \tau_{smt}$	-1.9	-0.5	-1.4
ψ_{lt}	1.62	1.00	0.61
ψ_{smt}	1.54	1.01	0.53
$\psi_{lt} - \psi_{smt}$	0.07	-0.01	0.08

(b) Subject IDjj, yes/no judgments *without* confidence ratings.

Parameter	Threshold		
	75%	50%	Diff
τ_{lt}	13.4	9.1	4.3
τ_{smt}	12.2	8.4	3.8
$\tau_{lt} - \tau_{smt}$	1.2	0.6	0.5
ψ_{lt}	1.16	0.91	0.24
ψ_{smt}	1.29	0.99	0.30
$\psi_{lt} - \psi_{smt}$	-0.14	-0.08	-0.05

(c) Subject ID338, yes/no judgments *with* confidence ratings.

Parameter	Threshold		
	75%	50%	Diff
τ_{lt}	9.0	5.8	3.1
τ_{smt}	12.0	7.8	4.2
$\tau_{lt} - \tau_{smt}$	-3.0	-2.00	-1.0
ψ_{lt}	1.08	0.73	0.35
ψ_{smt}	0.97	0.66	0.31
$\psi_{lt} - \psi_{smt}$	0.11	0.07	0.04

(d) Subject ID338, yes/no judgments *without* confidence ratings.

Parameter	Threshold		
	75%	50%	Diff
τ_{lt}	7.8	6.2	1.6
τ_{smt}	11.2	7.8	3.4
$\tau_{lt} - \tau_{smt}$	-3.4	-1.6	-1.8
ψ_{lt}	1.04	0.71	0.33
ψ_{smt}	0.90	0.66	0.24
$\psi_{lt} - \psi_{smt}$	0.14	0.05	0.09

thresholds) and ψ_{smt} (the y-intercept of the line fit to the Latency Condition scene-motion thresholds) match closely to the latency thresholds curve resulting from the fit parameters τ_{lt} and ψ_{lt} . The difference between these values are shown in Table 9.4. The greatest differences between τ_{lt} and τ_{smt} was 3.4 *ms* and the greatest differences between ψ_{lt} and ψ_{smt} was $0.14^\circ/s$. These small differences suggest that τ_{smt} and ψ_{smt} are equivalent to τ_{lt} and ψ_{lt} as my model predicts and are consistent with Experiment 4 (Section 8.5.3).

9.4.5 Scene-Motion Thresholds—Differences Between Conditions

I did not compare differences of scene-motion thresholds between the different conditions due to the low value of $\alpha = 0.05/24 = 0.002$ (Bonferroni correction) and the low number of samples ($n = 10$).

For Subject IDjj, a pattern is evident in Figures 9.3-9.8 that the Constant Condition thresholds are less than both the Gaussian Condition thresholds and the Latency Condition thresholds. There is no obvious difference between the Latency Condition thresholds and the Gaussian Condition thresholds. These results for Subject IDjj are consistent with the results of Experiment 4 (Section 8.5.2.3).

For Subject ID338, a pattern is evident in Figures 9.9-9.14 that the Latency Condition thresholds and the Gaussian Condition thresholds are different. Because of the difference between Conditions for Subject ID338, I suggest that investigators do not estimate latency thresholds with Gaussian-modeled scene-motion thresholds.

9.4.6 Threshold Differences of Yes/No Judgments with Confidence Ratings and Yes/No Judgments Without Confidence Ratings

I computed thresholds in two ways: using yes/no judgments with confidence ratings and using yes/no judgments alone. I compared differences between the resulting thresholds with Wilcoxon matched-pairs signed-rank tests, where thresholds were matched by peak head-yaw acceleration decile. I found the differences to be statistically significant for 75% thresholds and difference thresholds (each $p_{two-tail} < 0.05/12 = 0.004$), but not for 50% thresholds. Table 9.5 shows the results of these tests.

For Subject IDjj, 75% thresholds obtained from yes/no judgments with confidence ratings were, on average, 7.6 *ms* and $0.38^\circ/s$ greater than 75% thresholds obtained from yes/no judgments alone. Difference thresholds obtained from yes/no judgments with confidence ratings were, on average, 7.4 *ms* and $0.39^\circ/s$ greater than difference thresholds obtained from

yes/no judgments alone.

For Subject ID338, 75% thresholds obtained from yes/no judgments with confidence ratings were, on average, 1.8 *ms* and 0.12°/s greater than 75% thresholds obtained from yes/no judgments alone. Difference thresholds obtained from yes/no judgments with confidence ratings were, on average, 1.8 *ms* and 0.12°/s greater than difference thresholds obtained from yes/no judgments alone.

The fact that the resulting differences for the 75% thresholds and difference thresholds were nearly identical to each other suggests 50% thresholds are the same independent of the analysis method. This is consistent with the fact that the difference of 50% thresholds were not statistically significant.

Since no statistically significant differences were found for 50% thresholds, I computed confidence intervals of the differences and the corresponding ranges of statistical equivalence. The 95% confidence intervals of the differences (50% thresholds determined from judgments without confidence ratings subtracted from 50% thresholds determined from judgments with confidence ratings) for Subject IDjj were (−1.2 *ms*, 1.7 *ms*) and (−0.05°/s, 0.04°/s), and for Subject IDjj were (−0.4 *ms*, 0.5 *ms*) and (−0.02°/s, 0.02°/s). For Subject IDjj, confidence intervals of the differences in terms of percentage of his average thresholds were (−5%, 7%) for latency thresholds and (−3%, 3%) for scene-motion thresholds. For Subject ID338, confidence intervals of the differences in terms of percentage of her average scene-motion thresholds were (−2%, 3%) for latency thresholds and (−1%, 1%) for scene-motion thresholds.

Thus, for these subjects, 50% thresholds determined from yes/no judgments with confidence ratings were statistically equivalent to 50% thresholds determined from yes/no judgments without confidence ratings within ± 1.7 *ms*, $\pm 0.05^\circ/s$, and $\pm 7\%$ (each $p < 0.05$).

9.5 Summary of Results for these Cases

- Scene-motion thresholds, in degrees per second, correlate linearly (each $\rho \geq 0.76$ for yes/no judgments with scene-motion thresholds and each $\rho \geq 0.23$ for yes/no judgments alone) to peak head-yaw acceleration. (Section 9.4.1)
- For one subject, there is an apparent difference between the Gaussian Condition scene-motion thresholds and the Latency Condition scene-motion thresholds.
- Directly-measured latency thresholds correlate with the form of my inverse model (each $\rho \geq 0.94$ for yes/no judgments with scene-motion thresholds and each $\rho \geq 0.82$ for yes/no judgments alone) better than with a line (each ρ difference ≥ 0.03) (Section 9.4.3).
- The minimum directly-measured latency difference threshold is 4.3 *ms* (Section 9.4.2).

Table 9.5: Difference of thresholds obtained from yes/no judgments with confidence ratings and yes/no judgments without confidence ratings. A \checkmark indicates the differences were statistically significant. The + indicates the yes/no judgments with confidence ratings thresholds were greater than the yes/no judgments without confidence ratings thresholds. An = indicate thresholds were statistically equivalent within $0.05^\circ/s$ or $1.7 ms$.

(a) Subject IDjj

		Threshold		
		75%	50%	Diff
Scene-Motion Thresholds	With Rating - Without Rating	$\checkmark +$	=	$\checkmark +$
Latency Thresholds	With Rating - Without Rating	$\checkmark +$	=	$\checkmark +$

(b) Subject ID338

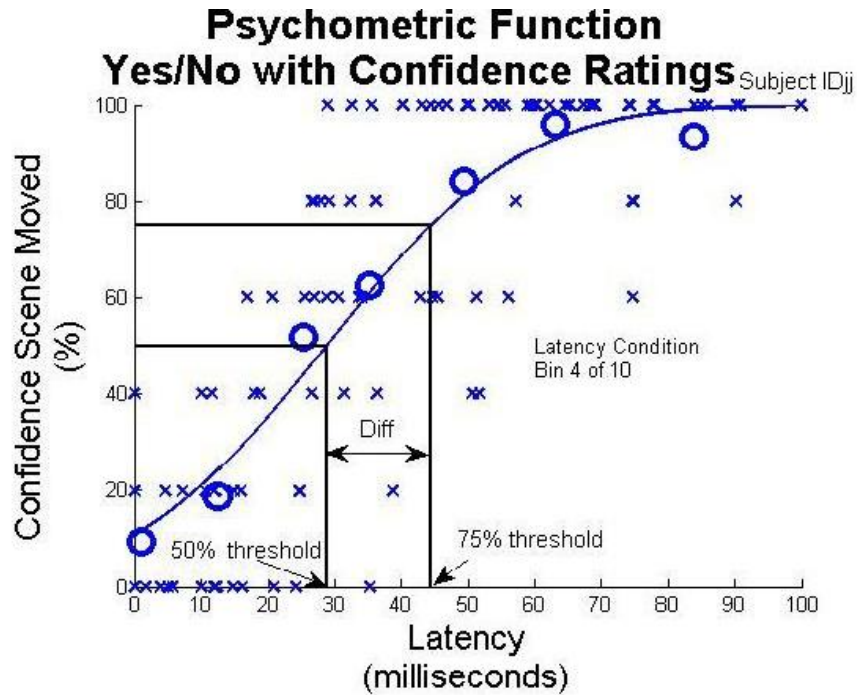
		Threshold		
		75%	50%	Diff
Scene-Motion Thresholds	With Rating - Without Rating	$\checkmark +$	=	$\checkmark +$
Latency Thresholds	With Rating - Without Rating	$\checkmark +$	=	$\checkmark +$

- The slope, τ_{smt} , and y-intercept, ψ_{smt} , of scene-motion threshold regressions can be used as input parameters, τ and ψ , for my latency model. The maximum difference between the fit parameter τ_{lt} and τ_{smt} is 3.4 *ms*. The maximum difference between the fit parameter ψ_{lt} and ψ_{smt} is 0.14°/s. These small differences suggest that τ_{smt} and ψ_{smt} are equivalent to τ_{lt} and ψ_{lt} , as my model predicts (Section 9.4.4) and are consistent with Experiment 4.
- 75% thresholds and difference thresholds obtained using yes/no judgments with confidence rating are greater than 75% thresholds and difference thresholds obtained using yes/no judgments alone. 50% thresholds using yes/no judgements with confidence ratings and 50% thresholds using yes/no judgments without confidence ratings are equivalent within 7% of average 50% thresholds. (Section 9.4.6)
- Correlations of thresholds to peak head-yaw accelerations vary less when using yes/no judgments with confidence ratings than when using yes/no judgments alone (Sections 9.4.1 and 9.4.3).

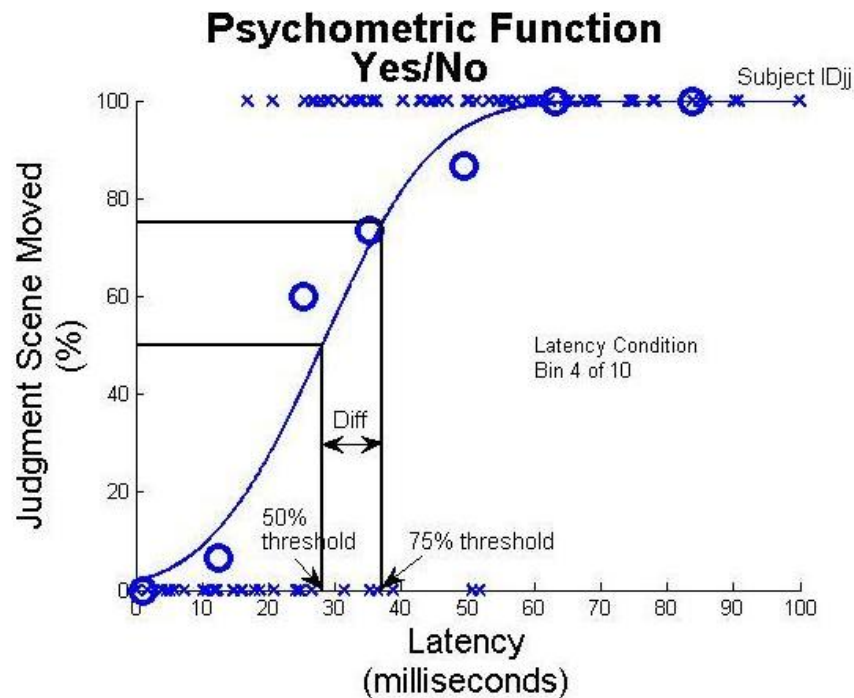
9.6 Recommendations

The results suggest the following recommendations:

- Although scene-motion thresholds may estimate latency thresholds reasonably well, it is just as easy to measure latency thresholds directly using my simulated-latency projector system. These directly-measured latency thresholds can be fit to the mathematical model (Equation 3.20), providing an analytical function of latency thresholds as a function of peak head-yaw acceleration.
- If an investigator wants to determine 50% thresholds, then using yes/no judgments with confidence ratings is sufficient for measuring scene-motion thresholds and latency thresholds.
- An investigator wanting to determine 75% thresholds and/or difference thresholds for the most sensitive subjects should first collect yes/no judgments with confidence ratings for several subjects. Then he can collect an extensive amount of data from the most sensitive subjects using yes/no judgments without confidence ratings, to determine more conservative (i.e., lower) thresholds.



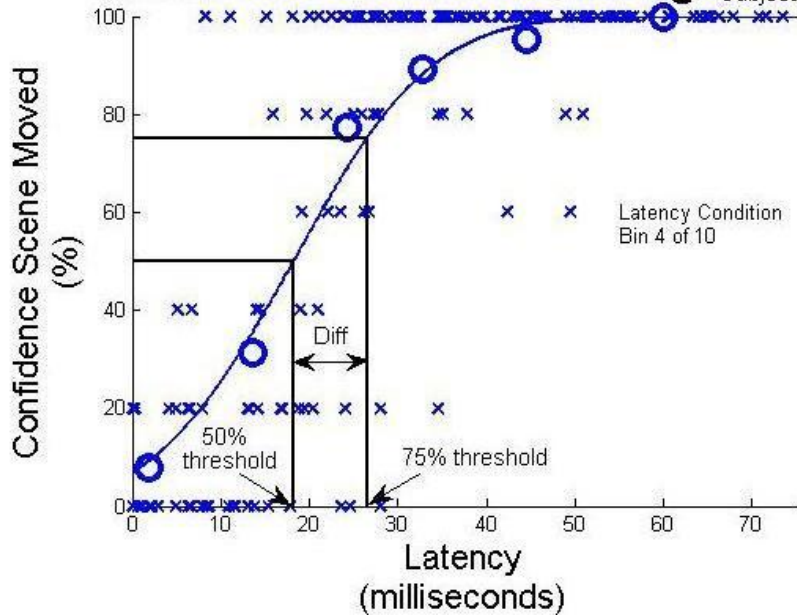
(a) Yes/no judgments *with* confidence ratings.



(b) Yes/no judgments *without* confidence ratings.

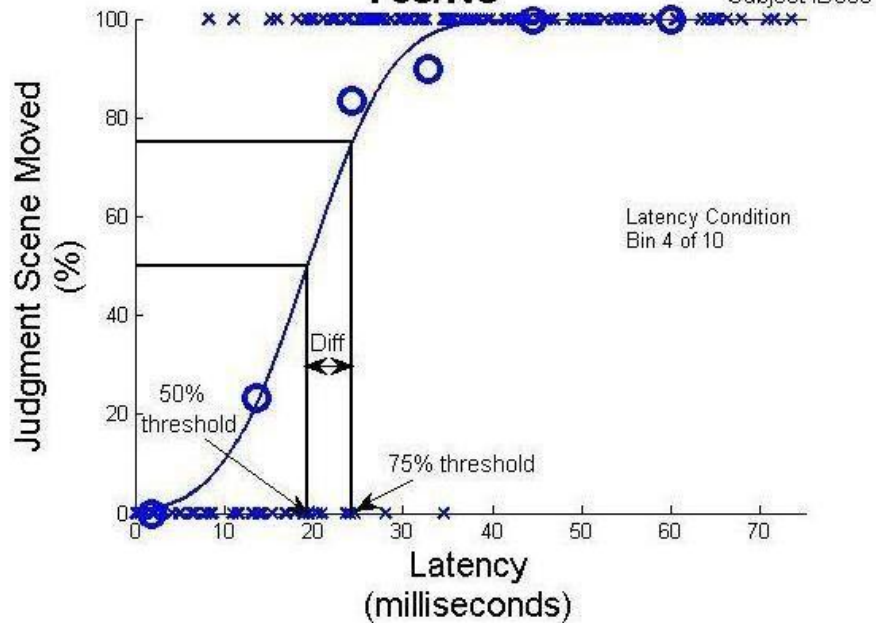
Figure 9.1: 115 judgments of various latencies for a range of Subject IDjj's peak head-yaw acceleration (decile four). A cumulative Gaussian function fit to this data is the psychometric function. The 50% threshold and difference threshold are extracted from the psychometric function. The Os represent the mean of every fifteen samples.

Psychometric Function Yes/No with Confidence Ratings



(a) Yes/no judgments *with* confidence ratings.

Psychometric Function Yes/No



(b) Yes/no judgments *without* confidence ratings.

Figure 9.2: 186 judgments of various latencies for a range of Subject ID338's peak head-yaw acceleration (decile four). A cumulative Gaussian function fit to this data is the psychometric function. The 50% threshold and difference threshold are extracted from the psychometric function. The Os represent the mean of every thirty samples.

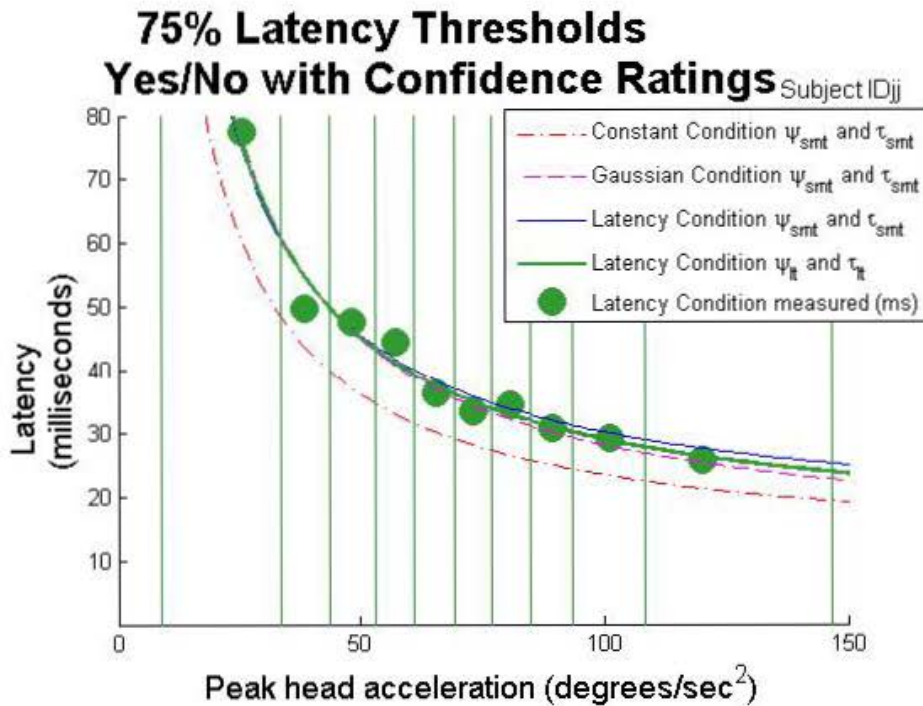
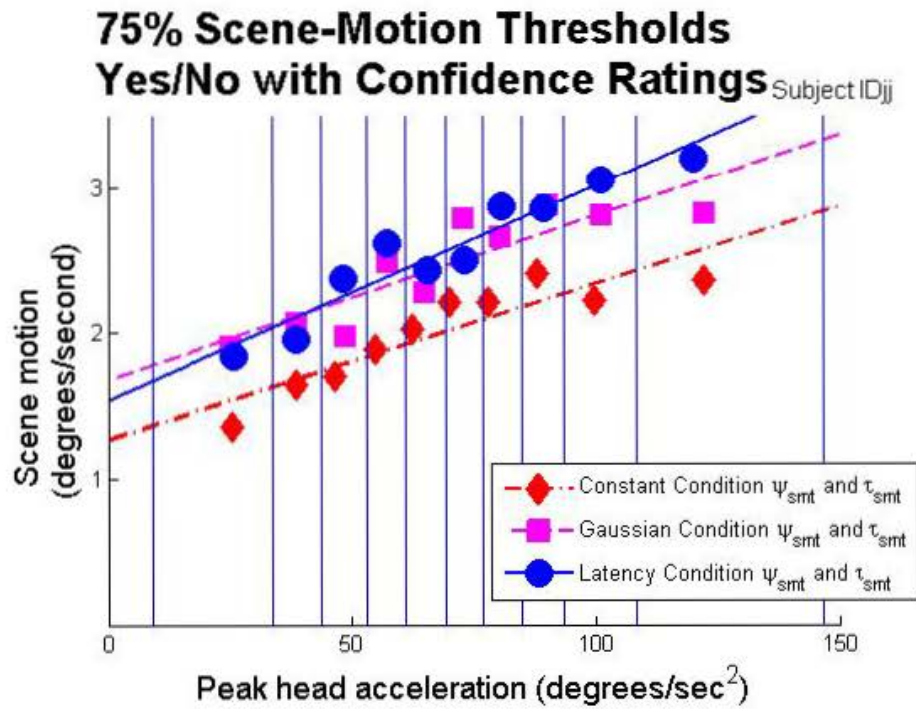


Figure 9.3: Scene-motion 75% thresholds (top) and latency 75% thresholds (bottom) for Subject IDjj, computed from yes/no judgments *with* confidence ratings.

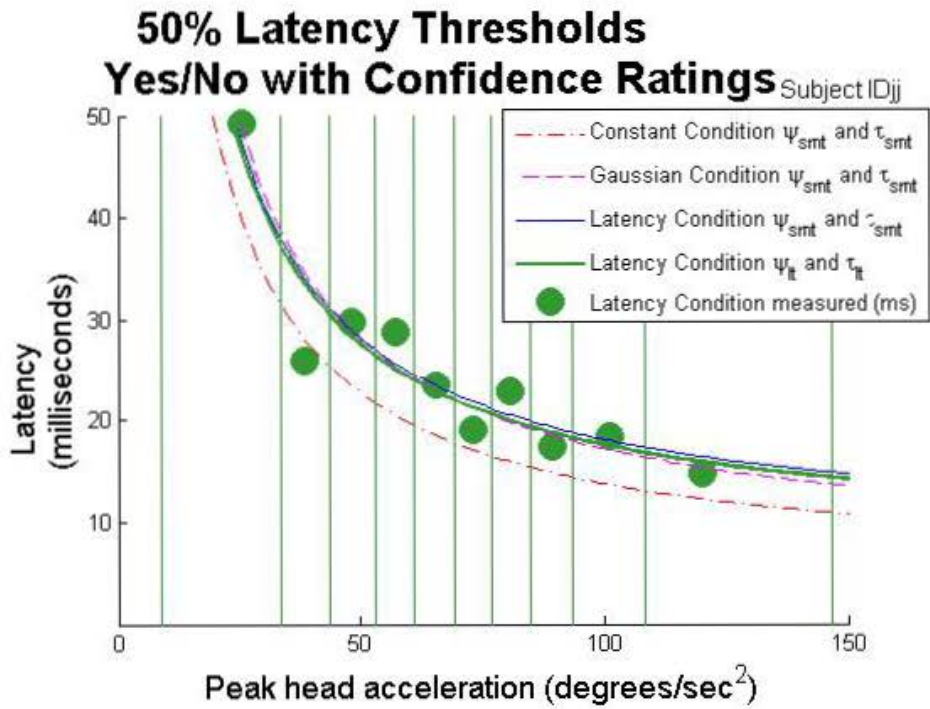
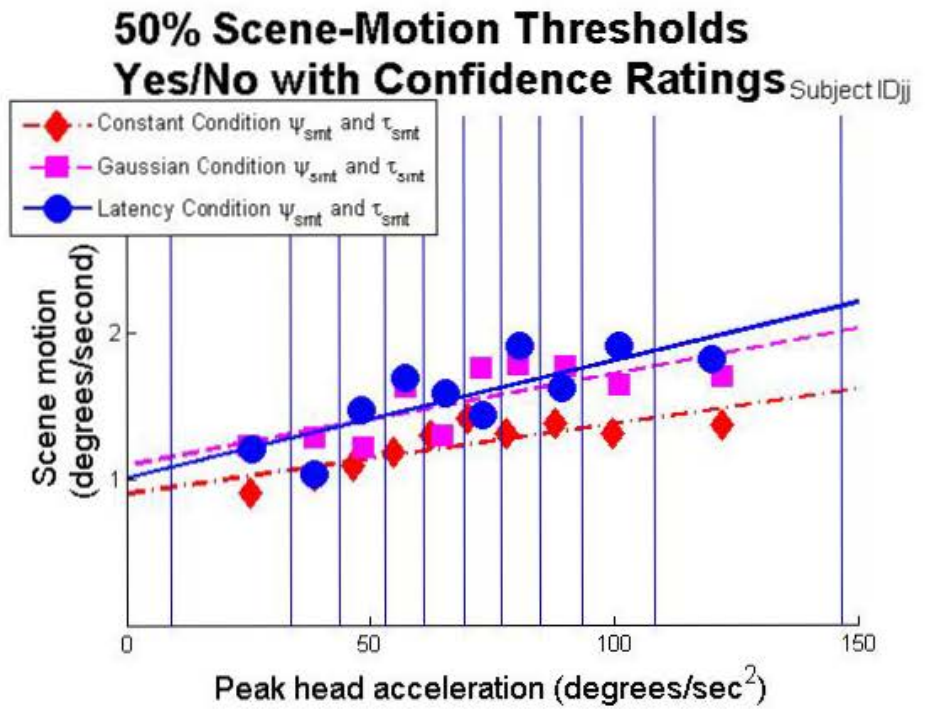


Figure 9.4: Scene-motion 50% thresholds (top) and latency 50% thresholds (bottom) for Subject IDjj, computed from yes/no judgments *with* confidence ratings.

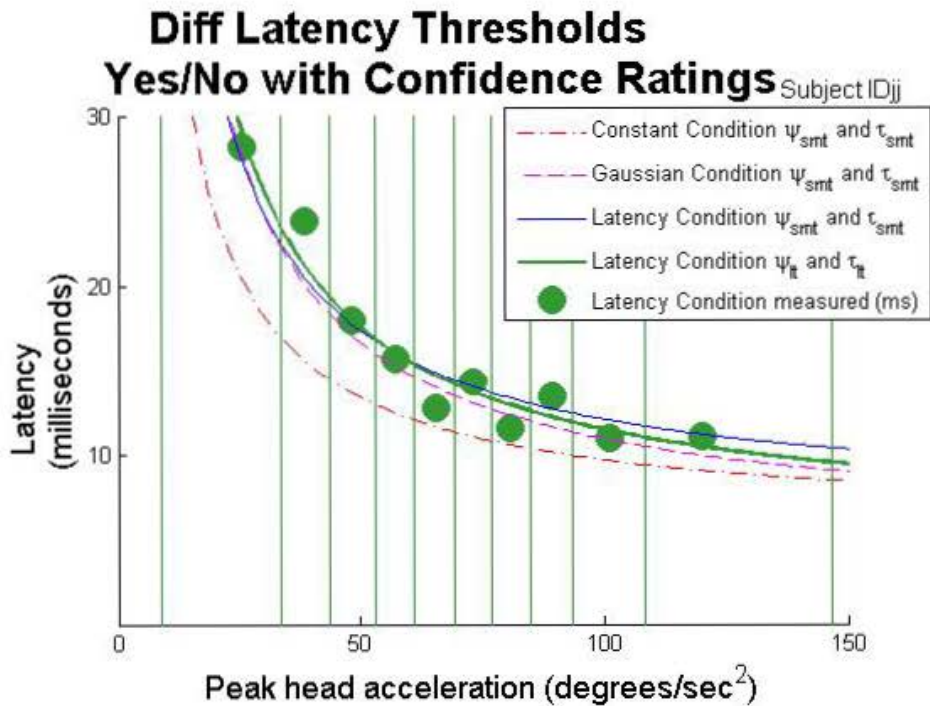
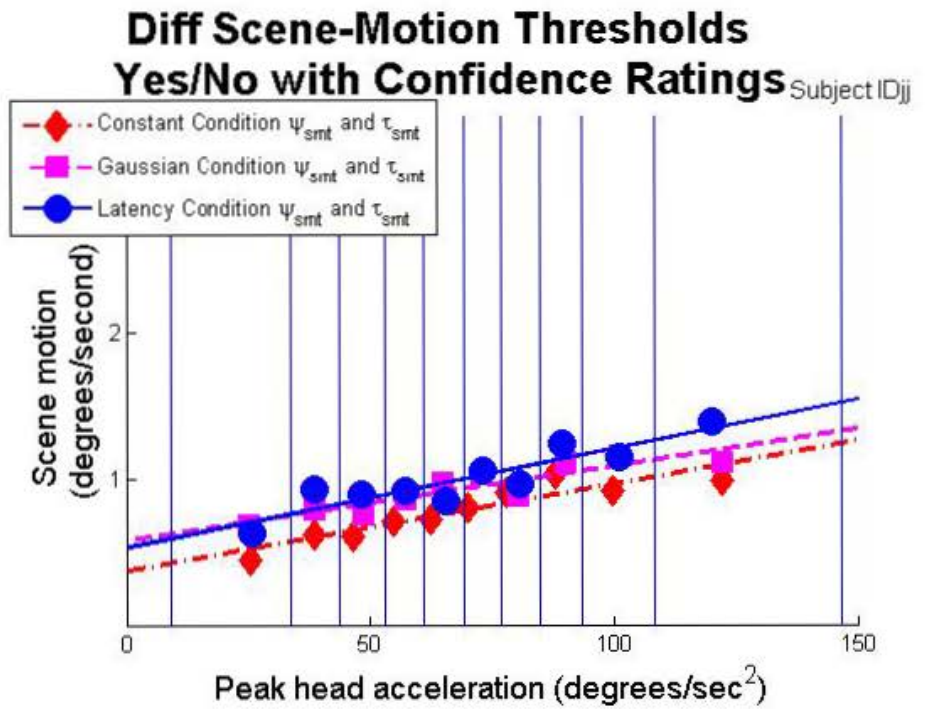


Figure 9.5: scene-motion difference thresholds (top) and latency difference thresholds (bottom) for Subject IDjj, computed from yes/no judgments *with* confidence ratings.

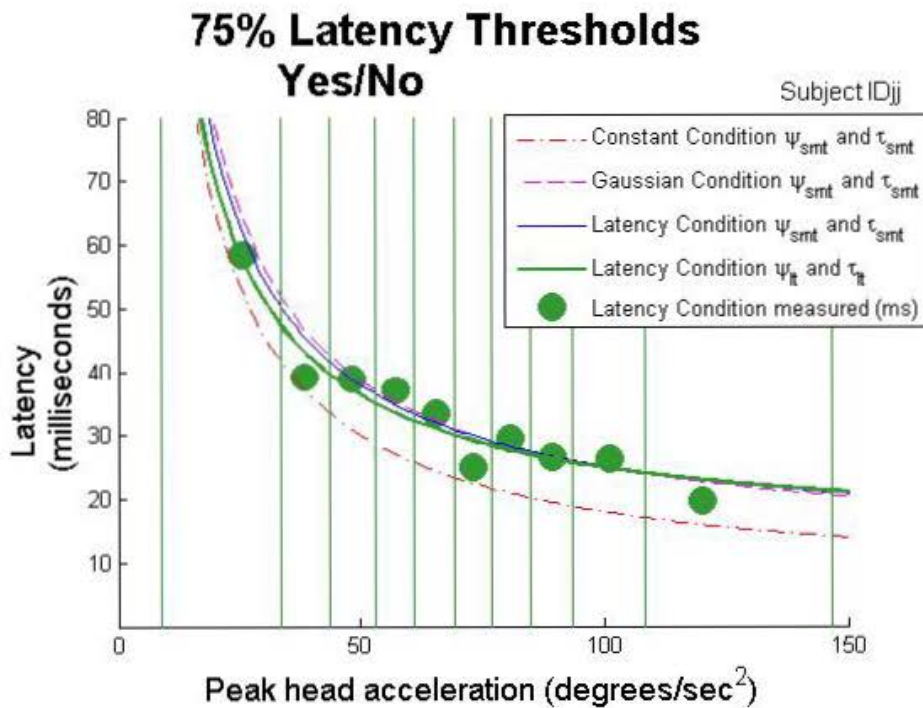
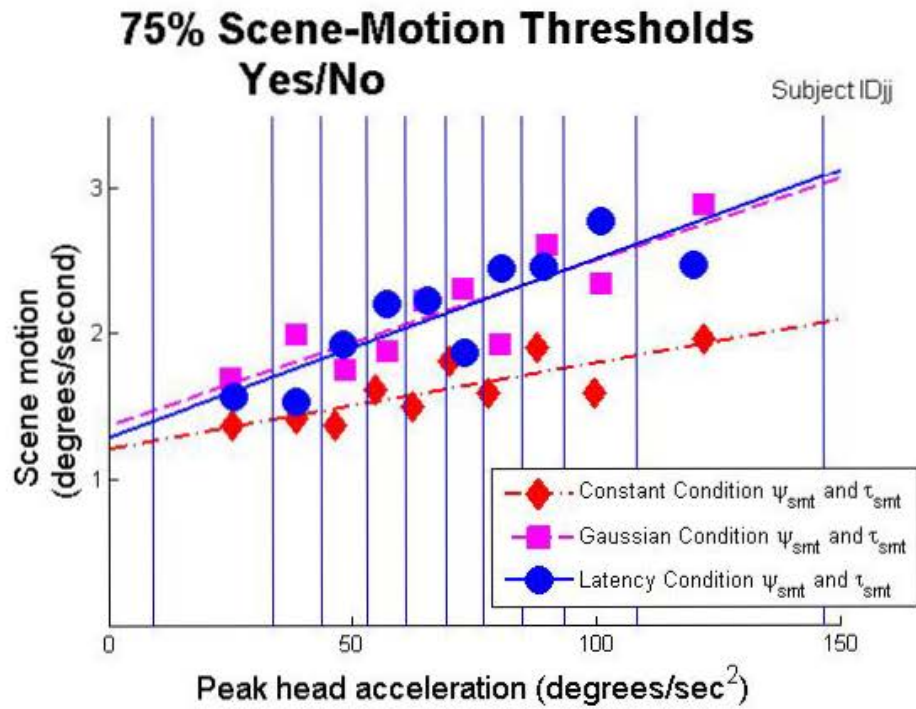


Figure 9.6: Scene-motion 75% thresholds (top) and latency 75% thresholds (bottom) for Subject IDjj, computed from yes/no judgments *without* confidence ratings.

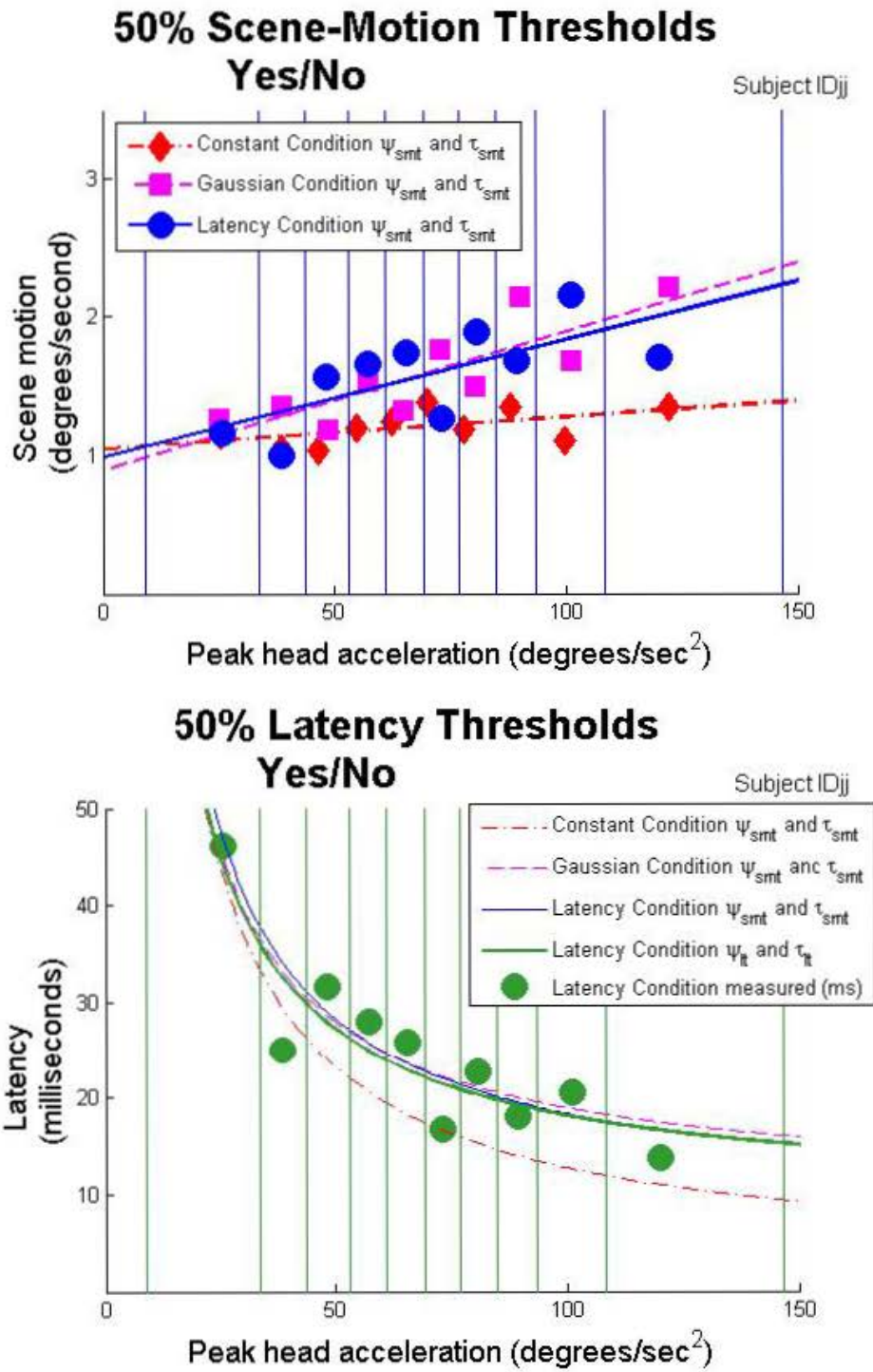


Figure 9.7: Scene-motion 50% thresholds (top) and latency 50% thresholds (bottom) for Subject IDjj, computed from yes/no judgments *without* confidence ratings.

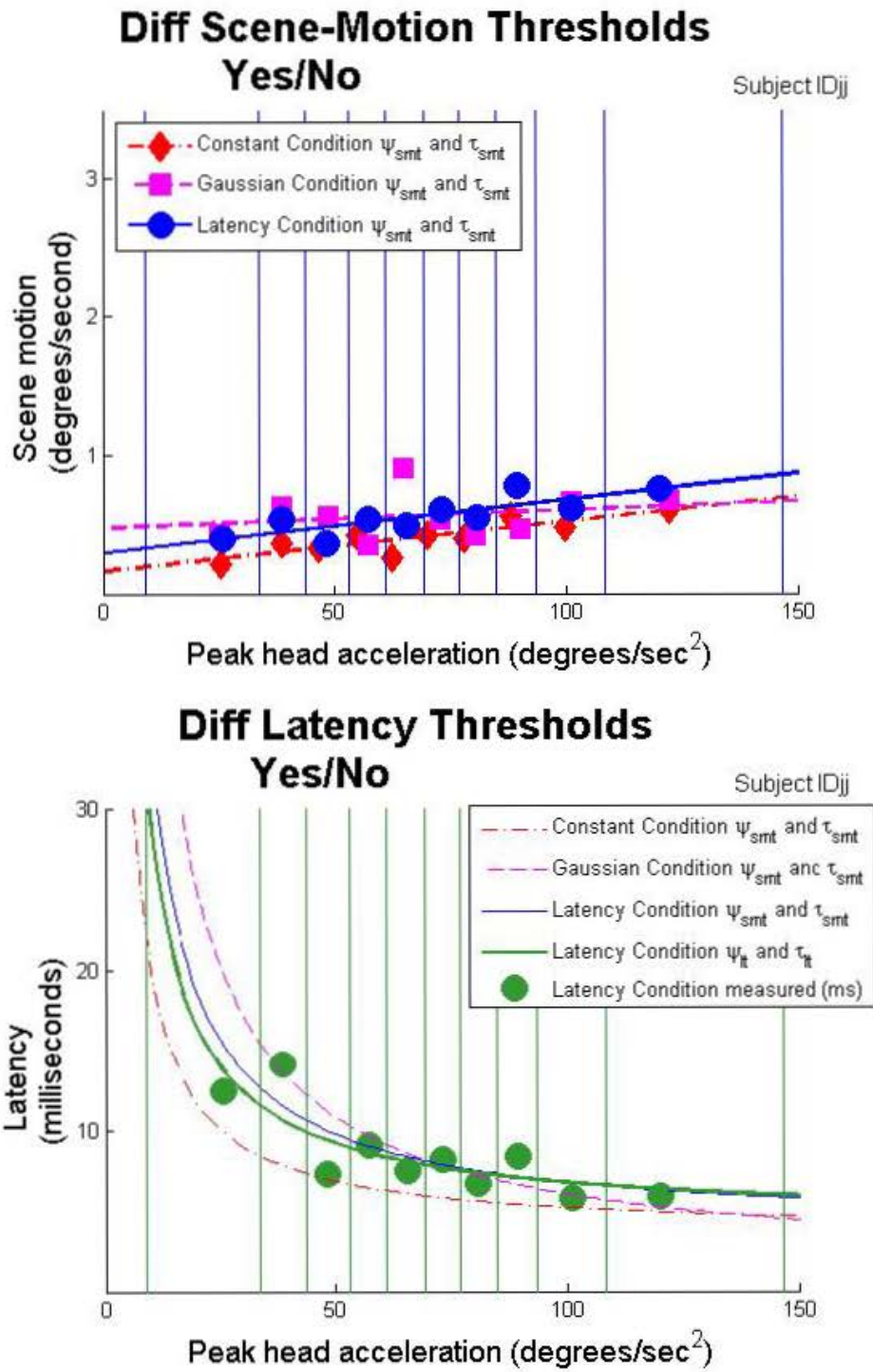


Figure 9.8: Scene-motion difference thresholds (top) and latency difference thresholds (bottom) for Subject IDjj, computed from yes/no judgments *without* confidence ratings.

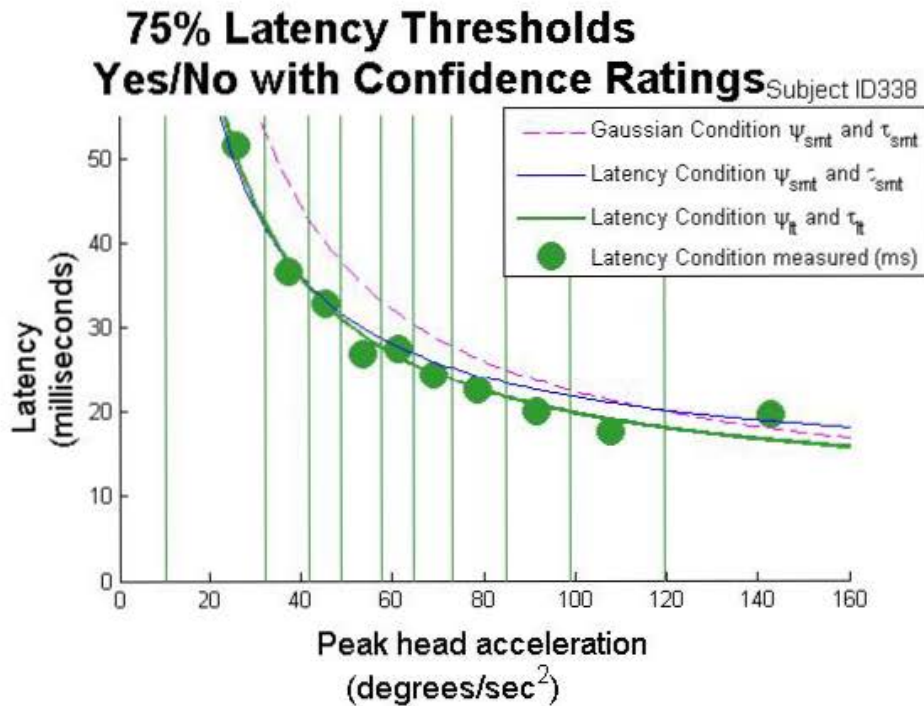
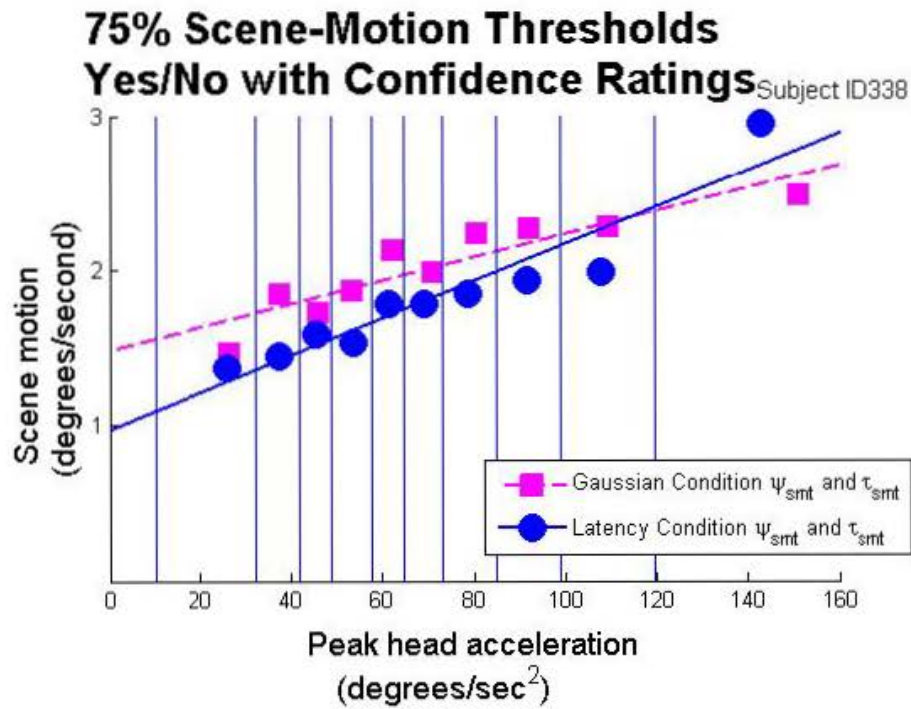
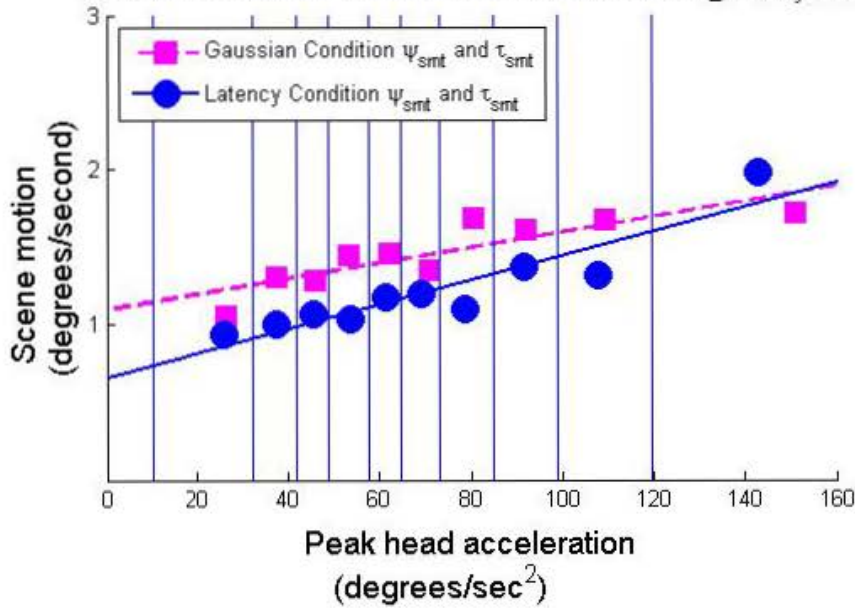


Figure 9.9: Scene-motion 75% thresholds (top) and latency 75% thresholds (bottom) for Subject ID338, computed from yes/no judgments *with* confidence ratings.

50% Scene-Motion Thresholds Yes/No with Confidence Ratings Subject ID338



50% Latency Thresholds Yes/No with Confidence Ratings Subject ID338

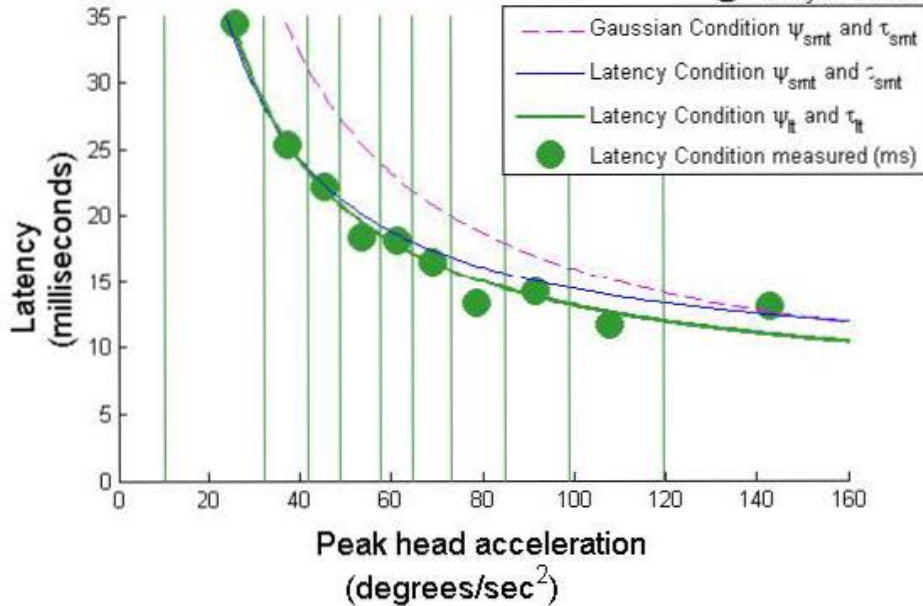
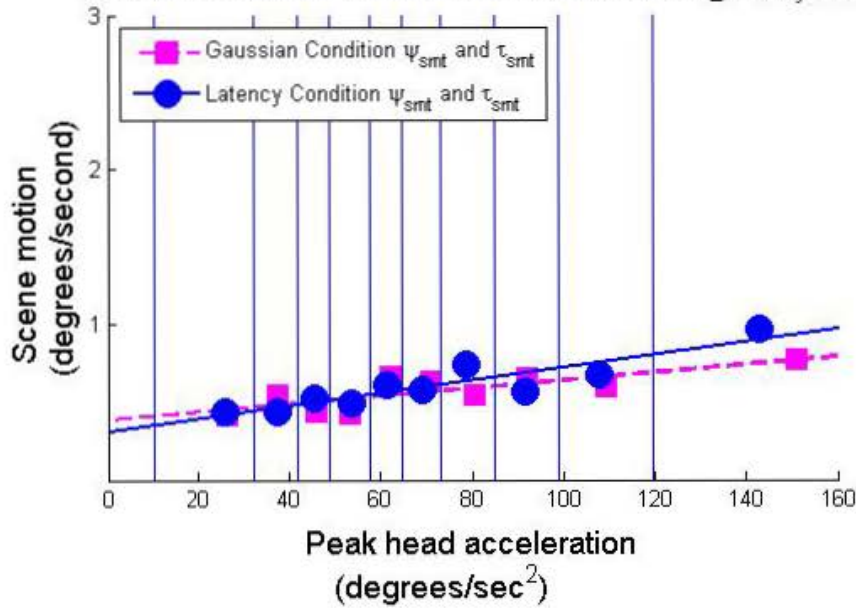


Figure 9.10: Scene-motion 50% thresholds (top) and latency 50% threshold (bottom) for Subject ID338, computed from yes/no judgments *with* confidence ratings.

Diff Scene-Motion Thresholds Yes/No with Confidence Ratings Subject ID338



Diff Latency Thresholds Yes/No with Confidence Ratings Subject ID338

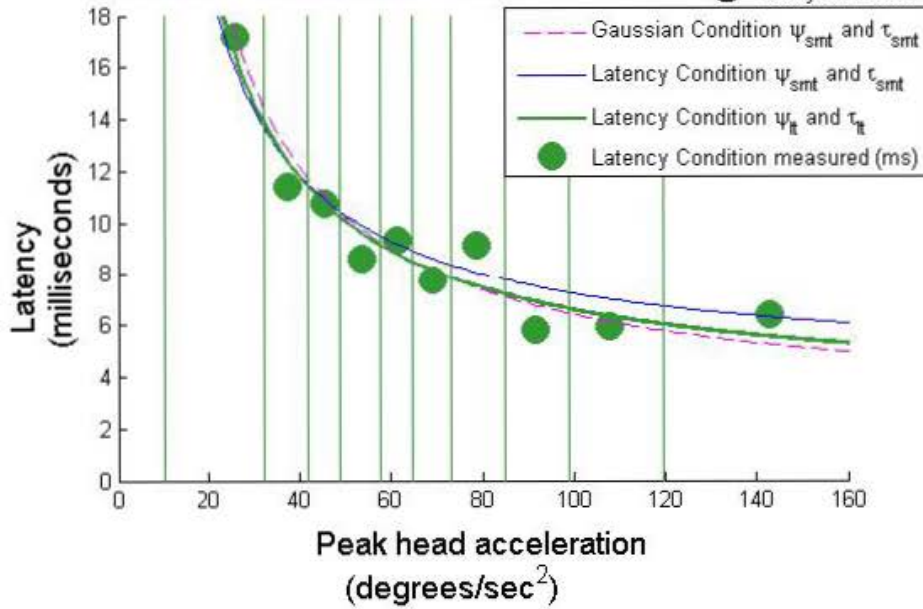


Figure 9.11: Scene-motion difference thresholds (top) and latency difference thresholds (bottom) for Subject ID338, computed from yes/no judgments *with* confidence ratings.

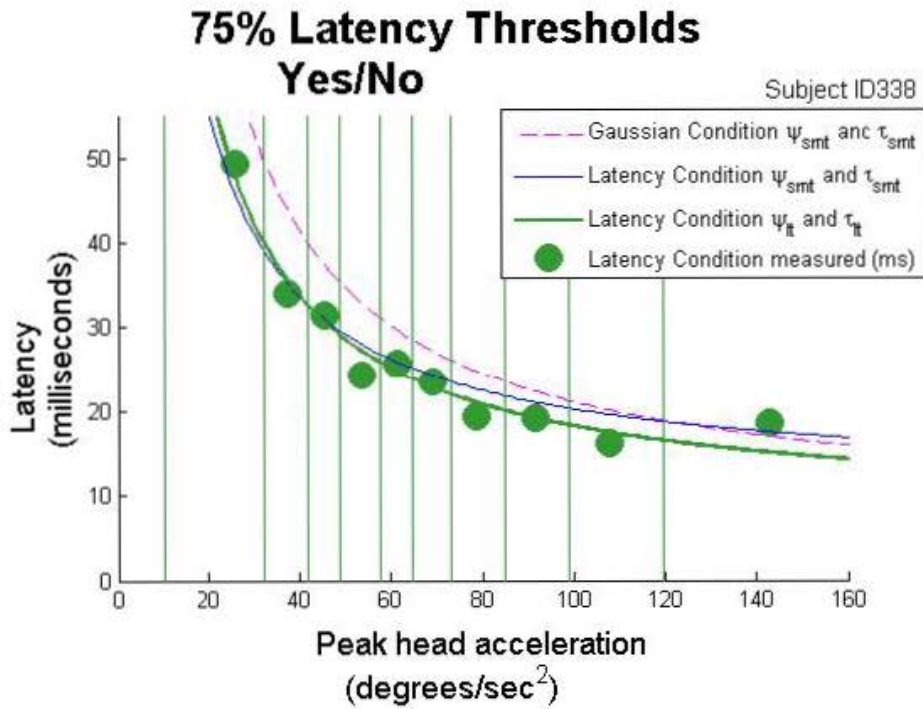
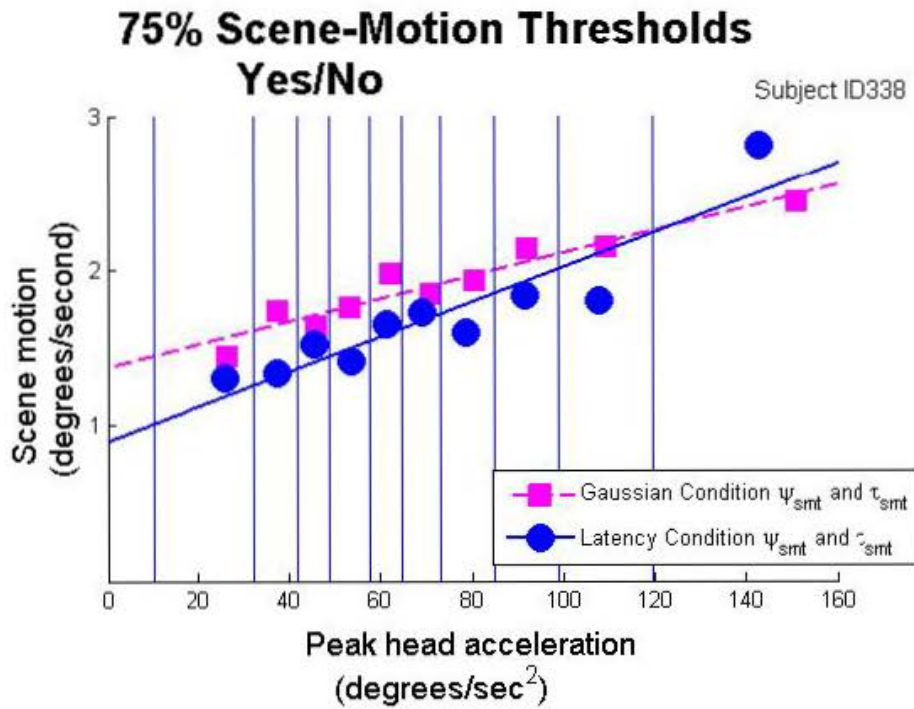


Figure 9.12: Scene-motion 75% thresholds (top) and latency 75% thresholds (bottom) for Subject ID338, computed from yes/no judgments *without* confidence ratings.

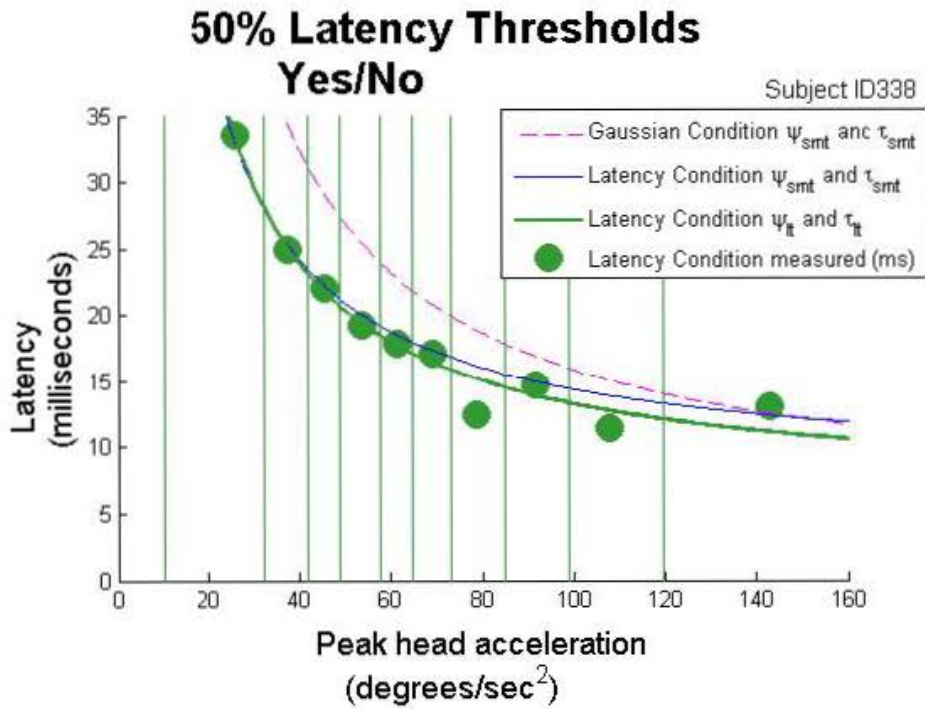
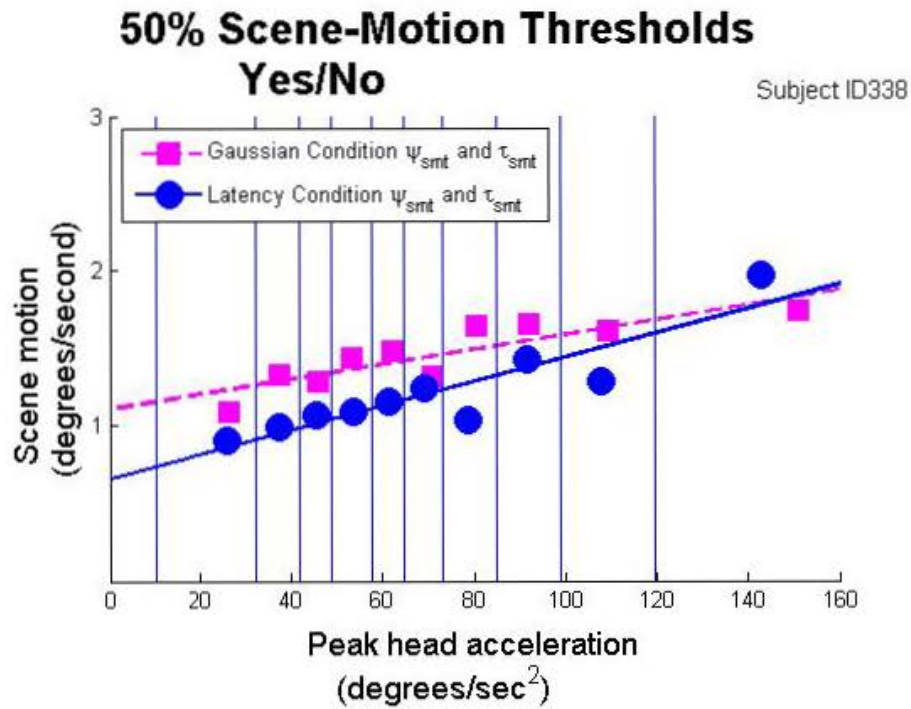


Figure 9.13: Scene-motion 50% thresholds (top) and latency 50% thresholds (bottom) for Subject ID338, computed from yes/no judgments *without* confidence ratings.

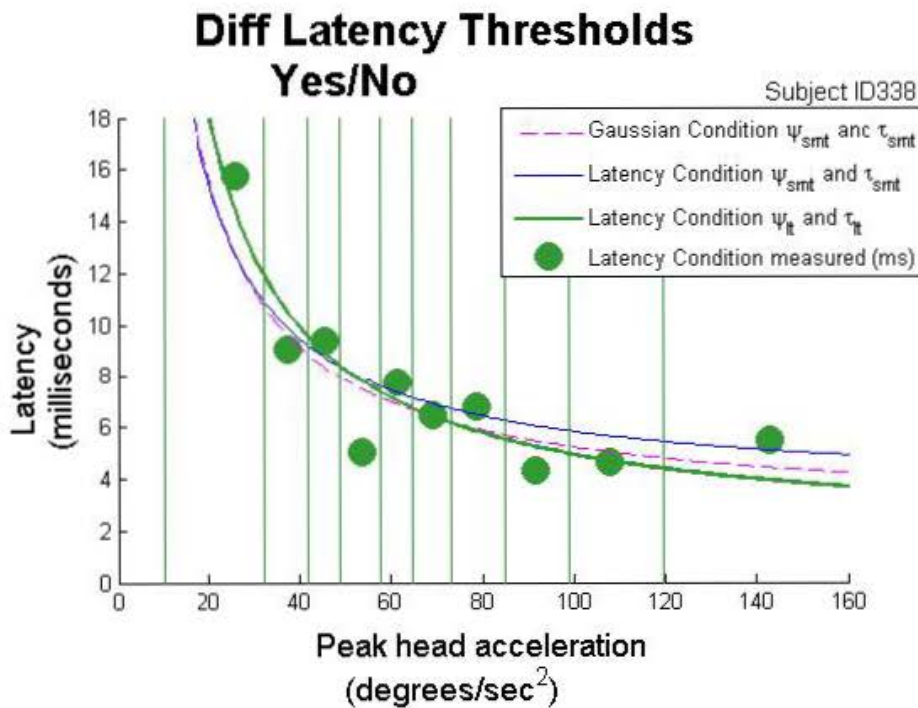
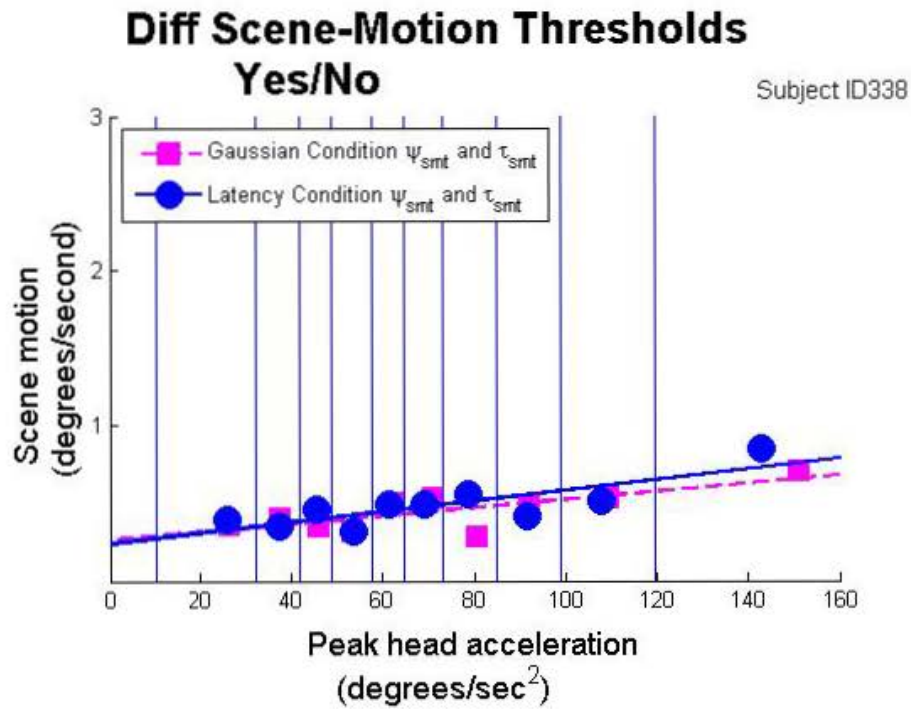


Figure 9.14: Scene-motion difference thresholds (top) and latency difference thresholds (bottom) for Subject ID338, computed from yes/no judgments *without* confidence ratings.

Chapter 10

Discussion

In this chapter, I compare my results to previous results, discuss factors affecting thresholds and latency requirements, describe limitations and extensions of my model and methods, and provide recommendations to IVE developers.

10.1 Comparison to Previous Results

In this section, I compare my results to the results of other investigators and discuss differences.

10.1.1 Scene-Motion Thresholds

The mean scene-motion threshold of Experiment 3 was 14.7% of peak head-yaw velocity, with a standard deviation of 11.2%. For Experiment 3, scene motion moved against the direction of head yaw.

Wallach

Wallach (1987) reported that healthy subjects perceived visual scenes to be unstable when the scene velocity is more than 2-3% of head velocity. He did not report the types of head motion in his summary or in earlier work (Wallach and Kravitz, 1965). His results are consistent with my results from Experiments 1 and 2 where scene-motion thresholds are less when the scene moves against the direction of head yaw than when the scene moves with the direction of head yaw. However, his scene-motion thresholds are much lower than the scene-motion thresholds I found in Experiment 3. The lack of details in his short paper make it difficult to speculate why.

Steinicke

Steinicke et al. (2008, 2010) investigated perception of scene-motion when walking in

general, without focusing on head rotations. Their system injected scene motion moving with or against the direction of physical body/head turns. They found subjects could be turned $\sim 49\%$ more than the theoretically correct (i.e., a stable scene in the world) visual rotation in the HMD (similar to my With Condition) and $\sim 24\%$ less than the theoretically correct visual rotation in the HMD (similar to my Against Condition) (Steinicke et al., 2010). Their results are consistent with my results of Experiment 1 and 2 in that twice as much scene motion can occur when the scene moves with the direction of head turns then when the scene moves against the direction of head turns without subjects noticing.

Loose and Probst

In contrast to Loose and Probst (2001), Experiment 3 found thresholds increased as head acceleration increased. These differences may be due to experiment differences:

- My subjects' head velocities varied as head acceleration varied, whereas Loose and Probst carefully controlled head velocity to be approximately constant for different amounts of head acceleration.
- My stimuli moved in world coordinates, whereas the stimuli of Loose and Probst moved in head coordinates.
- My subjects experienced only subject-relative cues, whereas the subjects of Loose and Probst experienced both subject-relative cues and a stable object-relative cue.
- My subjects actively rotated their own heads, whereas the subjects of Loose and Probst were actively rotated via a rotary chair.

10.1.2 Latency Thresholds

Difference thresholds from Experiments 4 and 5, across all subjects and all peak head-yaw acceleration deciles, ranged from 3.2 ms to 60.5 ms . 50% thresholds ranged from -0.2 ms to 103.2 ms .

NASA Experiments

Experiments conducted at NASA Ames Research Center (Adelstein et al., 2003; Ellis et al., 2004; Mania et al., 2004) reported latency thresholds during quasi-sinusoidal head yaw for a same/different task. The NASA-measured PSEs and JNDs are in the same general range as the 50% and difference thresholds I measured. Their measured PSEs ranged from -0.5 ms to $\sim 85\text{ ms}$ (compared to my 50% thresholds of -0.2 ms to 103.2 ms). Their measured JNDs ranged from $\sim 4\text{ ms}$ to $\sim 40\text{ ms}$ (compared to my difference thresholds of 3.2 ms to 60.5 ms).

My subjects had some higher thresholds than the NASA results, due to high thresholds for the slow head motions.

Allison

Allison et al. (2001) reported the proportion of times subjects reported a scene to appear to be unstable in an HMD for four latencies (0 *ms*, 50 *ms*, 100 *ms*, and 200 *ms*) and three frequencies of single left-to-right head yaw (45° over a period of 0.5 *s*, 1 *s*, or 2 *s*). Subjects increased their proportion of judgments that the scene appeared unstable as head motion increased and as latency increased. This is consistent with my results from Experiment 4 and 5 that latency thresholds decrease as peak angular head acceleration increases.

Allison et al. reported the average of 50% latency thresholds to be approximately 60 *ms* for the fast head turns and approximately 200 *ms* for the slow head turns (compared to my average 50% thresholds of -0.2 *ms* to 103.2 *ms* for a wide range of head turns). However, their latencies were in addition to their baseline latency of approximately 120 *ms*. Thus, their 50% latency thresholds were actually 180 – 320 *ms*, much higher than my measured 50% latency thresholds.

10.1.3 An Illusion of Scene Motion

I propose a perceptual illusion—when no object-relative cues are visible, scenes appear to be most stable when the entire scene moves slightly in the same direction as head turns and scenes appear to move slightly against the direction of head turns when in fact the scene is not moving.

Wallach and Kravitz (1965), Steinicke et al. (2008, 2010), and I (Section 6.6) found that scenes appear most stable when there is a small amount of scene motion with the direction of physical turns (although only the results of Wallach and Kravitz were statistically significant). This could explain why my With Condition scene-motion thresholds are greater than my Against Condition scene-motion thresholds (Experiments 1 and 2). This could also explain why the false alarm rates for Experiment 2 were so high in comparison to Experiments 3-5. For Experiment 2, it is possible that subjects perceived scenes moving slightly with the direction of head yaw to be most stable, resulting in higher confidence ratings that the scene moved when there was no scene motion. For Experiment 1, the scene always moved since I assumed subjects would guess correctly 50% of the time when there was no scene motion (due to the 3I-2AFC identification task). For Experiments 3-5, the scene did not move or moved against the direction of head yaw—no scenes moved with the direction of head yaw that could have biased false alarm rates.

Wallach

In Wallach's review of his own work (Wallach, 1987), he reported that scene-motion thresholds were the same whether the scene moved with the head or against the direction of head turns. However, in his original work (Wallach and Kravitz, 1965), he stated subjects judged scenes to be the most stable when scene rotation was 1.5% of head turns (with the scene moving in the the same direction as the head turns). For example, if subjects turned their head to the right at $100^\circ/\text{sec}$ then they would perceive the scene to be the most stable when the scene moved to the right at $1.5^\circ/\text{sec}$. Wallach stated this 1.5% to be negligible, even though the result was statistically different from zero. This "negligible" result is consistent with my findings.

Steinicke

Steinicke et al. (2008, 2010) report similar findings as Wallach and Kravitz (1965)—scenes appear to be most stable when the scene moves slightly with the head—that are consistent with my findings. Steinicke et al. did not perform tests to determine if their finding was statistically significant.

Jaekl

Jaekl et al. (2002) had subjects adjust scene motion until the scene appeared most stable. Their subjects found scenes moving 26% against the direction of head motion to appear to be the most stable. This contradicts my findings from Experiments 1 and 2.

10.1.3.1 Explaining the Illusion

I suggest that understanding how and when such an illusion of scene motion occurs would be best studied by investigating the following, in approximate ordering of decreasing likelihood of explanation.

1. **Pivot hypothesis.** The pivot hypothesis (Gogel, 1990) states that a point stimulus at a distance will appear to move as the head moves if it's perceived distance is different from its actual distance. A related effect is demonstrated by focusing on a finger held in front of the eyes and noticing that the background further in the distance seems to move with the head. Likewise if one focuses on the background then the finger seems to move against the direction of the head turn.

For the case that objects appear to be closer to users than they should, as is the case for HMDs (Loomis and Knapp, 2003), the scene should appear to move with the head. In this case, the scene would have to move against the direction of the head turn to appear stable in space. This is consistent with the findings of (Jaekl et al., 2002) with

their HMD, but in the opposite direction of my measurements with the simulated HMD (where perception of distance to the scene is more likely to be correct).

2. **Latency.** Scene motions injected into an HMD system counteract latency-induced scene motion for some phase of a head turn. For example, latency-induced scene motion moves with the direction of the head turn as the head accelerates and moves against direction of the head turn as the head decelerates. Adding scene motion in the direction of head turns would counteract latency-induced scene motion moving against the direction of head turns at the end of head turns. Likewise, adding scene motion against the direction of head turns would counteract latency-induced scene motion moving with the direction of head turns at the beginning of head turns. Latency could have affected scene-motion thresholds for the experiments of (Jaekl et al., 2002) and Steinicke et al. (2008, 2010), but not in my experiments or Wallach’s experiments, as the scene was presented by a world-fixed display so that latency caused no scene motion.
3. **Edge of HMD.** The edge of an HMD might cause subjects to judge scene motion relative to the head instead of relative to the world. Subjects might be biased towards greater thresholds when the scene moves with the direction of head turns due to the edge of the HMD serving as an object-relative cue (i.e., the scene moves by a smaller amount relative to the edge of the HMD when moving with the direction of head turns than when moving against the direction of head turns).
4. **Scene size.** The scene in my experiments had a 20° horizontal span. I expect larger scenes to appear to be more stable and to reduce the illusion since humans assume large objects do not move (Section 2.2.7.6) (althoughvection is more common for large objects seen in the periphery).
5. **Scene motion and eye motion relative to head motion.**

The difference between head motion and scene motion is greater when the scene moves against the direction of the head turn than when the scene moves with the direction of the head turn. Likewise, if the eyes follow the scene as the head turns, the difference between head motion and eye motion is greater when the scene moves against the direction of the head turn than when the scene moves with the direction of the head turn. Observers may be biased towards the smaller difference versus the larger difference.
6. **Eye stabilizing motions.** The vestibular-ocular reflex and optokinetic reflex cause the eyes to rotate in the opposite direction the head turns, so that the visual representation of an object remains stabilized on the retinas (Section 2.2.3.3). For a scene at an infinite distance the eyes rotate in an equal and opposite direction of the head (a gain of 1.0). For closer objects, the eyes must rotate by a greater amount (a gain greater than 1.0)

for the visual representation of the object to remain stabilized on the retina. Gains different than 1.0 might cause a bias resulting in apparent motion when no motion exists.

7. **Spatial frequency, brightness, and/or contrast.** It is possible that my results could be a function of spatial frequency, brightness, and/or contrast. Experiment 2 did not find any threshold differences for the two brightness conditions. This may have been due to not having enough difference in brightness and/or not enough statistical power. Freeman and Banks (1998) found that the Filehne illusion (when tracking a moving object with the eyes, the background appears to move against the direction of eye motion) and the Aubert-Fleischl illusion (the false impression that objects seem to move slower when they are pursued with the eyes as compared to when the eyes are kept stationary) could be reversed by manipulating spatial frequency. Manipulating spatial frequency, brightness, and/or contrast might reverse the illusion I found.

10.1.4 Explanation of Differences

I believe the results of Steinicke (higher scene-motion thresholds than my measured thresholds), Allison (much higher latency thresholds than my measured thresholds), and Jaekl (scenes appear more stable when the scene moves against the direction of head motion) are different than the results of the NASA experiments and my experiments due to the following:

- **No reference scene.** None of the subjects of Steinicke, Allison, and Jaekl were provided with within-trial or between-trial reference scenes which to compare the test scenes. The NASA subjects were provided with within-trial reference scenes with some baseline latency. My subjects were provided with stable (zero effective latency) within-trial reference scenes (Experiment 1) or between-trial reference scenes (Experiments 2-5).
- **Training.** The subjects of Steinicke, Allison, and Jaekl might have been naive to latency and its consequences of scene motion. They were provided with little training or explanation of what to look for. Steinicke provided subjects with two extreme examples of both conditions (scene motion with and against the direction of body/head turns) at the beginning of the experiment (Steinicke, 2009). Allison's subjects were asked if the scene appeared to be stable or not. Before the start of the studied trials, the NASA subjects and my subjects were provided with extensive training of what scene motion and latency looked like.
- **Baseline Latency.** The systems of Steinicke, Allison, and Jaekl all contained baseline latencies (the minimum latency their system could achieve) causing some scene motion

that was likely perceived by all their subjects. Since, in these systems, there was no reference scene for subjects to know what “zero” latency looked like, subjects’ judgments of “stable” might have been arbitrarily chosen to be “good enough” instead of “no perceived scene motion”. Steinicke and Jaekl did not report baseline latency; Allison’s baseline latency was 122 *ms*. The NASA system also had some baseline latency (reported to be as low as 14.5 *ms* (Adelstein et al., 2005) not including a display latency of 32-34 ms that I measured for a V8 HMD (Section 2.4.3.7), like what they used), but that latency was taken into account by having the reference scene contain the baseline latency. My baseline latency was effectively zero, i.e., no scene motion occurred due to system delay (Section 4.1.1).

Allison’s baseline latency was 122 *ms*, a higher latency than my highest measured 50% latency threshold of 103 *ms* (Allison also measured 50% latency thresholds). One of Allison’s subjects judged that the scene appeared to be unstable for all conditions. Allison considered this single subject’s judgments to be outliers and did not include that subject’s data in his analysis. I believe that subject might have been the only subject to report “perceived scene motion” instead of “good enough”.

- **Type of physical motion.** Steinicke investigated perception of scene-motion when walking in general, without focusing on head rotations. The physical movement of the entire body might have reduced the ability to perceive scene motion.
- **Other sources of error.** Head-mounted displays can be very difficult to calibrate precisely. Similar to baseline latency, other sources of error cause additional scene motion. Incorrect depth cues, incorrect rendered field of view, and other incorrect parameter values, cause unwanted and unnecessary scene motion in HMDs as the head moves. For example, similar to minifying glasses (Section 2.2.9.3), a scene displayed with a smaller field of view than it’s rendered field of view causes the scene to move with the direction of head turns—in this case the virtual scene would have to move against the direction of head turns to remain stable in world space.
- **Psychophysics methods.** Jaekl’s subjects adjusted scene motion until the scene appeared most stable (some scene motion presumably occurred due to some unreported latency). Steinicke’s subjects were asked, while walking, if the virtual movement was less than or greater than physical movement. Allison’s subjects were asked if the scene appeared to be stable or not. The NASA subjects were asked if two presentations (one reference presentation with baseline latency and one probe presentation with some additional latency) were the same or different. My subjects were asked which of two presentations were different from a reference presentation (Experiment 1) or if the scene seemed to move (compared to a reference scene presented between trials).

10.2 Subject Comments

I expected subjects to judge scene motion directly from the visual motion, or lack thereof, on the display. However, some subjects reported making judgments about the smaller motions based on avection-like sensation (i.e., self-motion). Some subjects commented on this orally and some reported this in written form for the exit questionnaire. Two subjects from Experiment 1 wrote in response to “Explain in your own words how you went about making the scene-motion judgments.”:

- “Most of the times I detected the scene with motion by noticing a slight dizziness sensation. On the very slight motion scenes there was an eerie dwell or “suspension” of the scene; it would be still but had a floating quality.” (Subject ID103)
- “At the beginning of each trial [sic: session—when scene velocities were greatest because of the adaptive staircases], I used visual judgments of motion; further along [later in sessions] I relied on the feeling of motion in the pit of my stomach when it became difficult to discern the motion visually.” (Subject ID109)

The finding that scene-motion thresholds increase as head motion increases is consistent with some subjects’ statements about their perception of scene motion. For example, one subject stated the following for Experiment 2:

- “It sometimes feels still up to the end, then it moves when you stop your head.” (Subject ID415).

This statement implies the subject perceived variable scene velocity as head motion changed, when in fact the scene moved with constant velocity.

These comments suggest perception of scene motion is not based solely upon image motion on the retina. Other cues, such as efference copy (Section 2.2.4), vestibular cues (Section 2.2.2), andvection (Section 2.2.7.6), play a role in the perception of scene motion.

10.3 Factors Affecting Thresholds

I expect latency thresholds to vary under different experimental and application settings. The thresholds I measured provide guidelines for latency, but I do not claim that these thresholds will necessarily hold across all conditions. I encourage scientists and engineers to use the methods described to establish latency requirements for their specific conditions.

I believe the following factors affect scene-motion and latency thresholds, in approximate declining order of importance.

1. **Object-relative cues.** This work deals with only subject-relative cues—all real world cues are removed so as to simulate a non-see-through HMD. Motion thresholds with object-relative cues (as seen in an optical optical-see-through HMD) are much lower (Section 2.2.7.2).
2. **Individual differences.** My observations and data suggest that latency thresholds differ significantly for different people. Other than object-relative cues, I believe differences between individual subjects may be the the most important factor of scene-motion thresholds. Some subjects had more difficulty judging scene motion than others. The data of two subjects from Experiment 4 could not be used due to the noise in their judgments; it appeared these two subjects were largely guessing. I 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.

I suggest that, when possible, engineers should, in order to help ensure users will not perceive scene motion or latency, design systems with latency below the lowest measured latency thresholds.

3. **Attention.** The latency thresholds presented in this work are conservative values (i.e., low thresholds) since subjects focused exclusively on scene motion. I am reasonably sure that more motion could occur without users noticing if they were performing some other cognitive task such that their attention were not solely devoted to judging scene motion.
4. **Training.** Training with feedback of correct or incorrect responses likely lowers thresholds substantially. Even without feedback, subjects' thresholds probably decrease as they experience more trials due to improving their understanding of scene motion (in contrast to fatigue).
5. **Fatigue.** Although I encouraged subjects to take breaks at any time, I suspect fatigue may have have led to inattention and inaccurate responses. I noticed that judgments did not seem to consistently relate to the session number across subjects: some subjects performed better during the early sessions, some performed better during the middle sessions, and some performed better during later sessions. The interaction of attention, training, and fatigue make it difficult to analyze any one factor by itself.
6. **Field of view and/or scene size.** The human ability to detect slow-moving visual stimuli decreases with eye eccentricity (Section 2.2.7.4). However, vection occurs more strongly for large scenes that are seen in peripheral vision (Section 2.2.7.6). The two

visual pathways (Section 2.2.1.2) imply two separate processes occur when judging scene motion. Scene motion may occur differently to users for these two processes. I informally noticed that I intuitively “feel” scene motion when it occurs in the periphery versus whereas I more logically “think” about scene motion when it occurs only in central vision.

7. **Experimental Methods and Analysis.** The specifics of the psychophysics methods and analysis result in different threshold values. For example, a 3AFC task will result in different 75% thresholds than a 2AFC task. Another example is how thresholds from Experiment 5 were different using yes/no judgments with confidence ratings than using yes/no judgments alone.
8. **Type of head turns.** I measured thresholds only for head yaw. Thresholds might be different for head pitch, although I expect the differences to be small. Head turns other than quasi-sinusoidal head turns (Experiment 1) and single left-to-right and right-to-left head turns (Experiments 2-5) might also result in different thresholds.

10.3.1 Other Factors for Latency Requirements

This work focused upon perception of scene motion resulting from latency. It may be the case that more scene motion / latency can occur without affecting performance and/or without causing simulator sickness, even though subjects may consciously notice the scene motion / latency. Or, it may be the case that imperceptible scene motion / latency can negatively affect performance and/or cause simulator sickness. I suspect scene-motion / latency affect on performance to be highly task dependent.

10.4 Limitations

10.4.1 Assumptions

My experiment assumptions (Section 3.3) may not hold in all cases. For example, thresholds may be different for scenes that are close to the viewpoint of subjects. I suspect minor violations of the stated assumptions have relatively little effect compared to the factors described in Section 10.3.

My mathematical assumptions (Section 3.4.1) are not completely veracious. For example, Equation 3.15 slightly overestimates peak scene velocity. The results that measured latency thresholds correlate with my mathematical model provide evidence for the robustness of the mathematical assumptions.

10.4.2 Average Thresholds

The thresholds reported are average responses over many trials; users might occasionally notice lower scene-motions and latencies than the threshold values reported.

10.4.3 Faster Head Motions

I measured thresholds only for low to moderate head yaw accelerations and my model may not hold for faster head yaw. If users are expected to make faster head turns (e.g., fighter-pilot training), then thresholds should be measured for those faster head turns.

10.5 Studying Other Sources of Scene-Motion

Scene motion can occur in IVE systems for reasons other than artificially inserting scene motion or latency. Other forms of dynamic error (i.e., scene motion) can occur as the viewpoint changes, and perception of this error could be studied in a manner similar to the methods I have created in this work. Examples of sources of dynamic error are:

- **Incorrect field of view**—the display field of view does not match the geometric field of view set in the application.
- **Inaccurate calibration**—e.g., an incorrect tracker-to-viewpoint transformation matrix.
- **Incorrect depth cues**—e.g., Stereo cues and eye accommodation.

10.6 Recommendations to IVE developers

10.6.1 Application to Reorientation Techniques

Reorientation techniques enable users to walk in larger-than-lab sized IVEs by imperceptibly rotating the scene around the user. Razzaque's maximum rotation values (Razzaque et al., 2001; Razzaque, 2005) were determined by informal evaluation of imperceptibility for a small group of users. The thresholds I measured, or thresholds obtained through similar psychophysical measurements, could be used to establish the maximum amount of allowed scene motion.

10.6.1.1 Maximum Scene-Motion as a Function of Physical Head Turns

Redirected walking systems typically limit the maximum scene-velocity θ'_{max} to be a proportion of the current angular head velocity $m_{vel}\phi'$:

$$\theta'_{max} = m_{vel}\phi' \quad (10.1)$$

where m_{vel} is a unitless number.

Experiment 3 suggests a better function of maximum scene-velocity θ'_{max} to be a linear function of the current angular head velocity ϕ' with the form

$$\theta'_{max} = b_{vel} + m_{vel}\phi' \quad (10.2)$$

where b_{vel} is in units of degrees per second.

Experiments 3-5 suggest developers could also limit scene-velocity θ'_{max} to be a linear function of peak head acceleration ϕ'' with the form

$$\theta'_{max} = b_{acc} + m_{acc}\phi'' \quad (10.3)$$

where b_{acc} is in units of degrees per second and m_{acc} is in units of seconds. Experiment 5 suggests this form may be the best model, since scene-motion thresholds (in degrees per second) have a high linear correlation with peak head-yaw acceleration.

10.6.1.2 Scene Motion Against Versus With Physical Turns

The results of Experiments 1 and 2 suggest different maximum scene velocities should be used depending if the scene moves with the direction of head turns or against the direction of head turns. Twice as much scene motion can occur imperceptibly if the scene moves with the direction of head turns than against the direction of head turns.

If the system wants to reorient the user to the right by 10° , there are two options:

- **Rotate the scene with the direction of the physical rotation.** For example, rotate the scene to the right by 10° (in world space) as the user physically turns 100° to the right. The visual scene updates as if the user had turned 90° to the right ($100^\circ - 10^\circ$). The user will visually perceive that she rotated 90° to the right when she physically turned 100° to the right.
- **Rotate the scene against the direction of the physical rotation.** For example, rotate the scene to the right by 10° (in world space) as the user physically rotates 100° to the left. The visual scene updates as if the user had rotated 110° to the left. The

user will visually perceive that she turned 110° to the left when she physically turned 100° to the left.

My results imply that users are less likely to perceive scene motion if the first option is implemented.

Distractors, as first implemented by Miller (2002) and later investigated by Peck et al. (2008), can be used for reorientation purposes by occasioning the user to turn his viewpoint so that the scene can be rotated without the user noticing. My results suggest these distractors could be used most effectively by having a distractor trigger a head turn in the same direction that the system rotates the world, such that users would turn in the same way that the world turns.

I suspect scene-motion thresholds to be higher when subjects pay attention to a distractor or some distracting task, similar to change blindness (Goldstein, 2007). Further experimentation would be required to confirm if and how distractors affect scene-motion thresholds.

10.6.2 Measuring Latency Thresholds

Measuring scene-motion thresholds helped me understand how perception of scene motion relates to perception of latency in HMDs and helped me create the mathematical model. However, measuring scene-motion thresholds are not necessary to measure latency thresholds. I suggest investigators determine the parameters of my latency-thresholds model by measuring latency thresholds directly using a simulated-by-projector HMD system—it is just as easy to measure latency thresholds as it is to measure scene-motion thresholds.

10.6.3 Latency Requirements

Assuming a goal is for users not to perceive scene motion due to latency, developers of HMD systems should first measure latency thresholds for their specific conditions in order to determine latency requirements. If resources are not available to do so, I suggest as a guideline that HMD systems should not have more than 3 *ms* of latency. Although achieving this level of latency is quite challenging, it is not impossible. The simulated-HMD system I built with off-the-shelf components contained 7.4 *ms* of system delay. If a developer had access to an HMD with fast-response optics and a fast head tracker, then an IVE with 3 *ms* of system delay could be built.

10.6.4 Reducing Latency

In order to develop minimal-latency HMD systems, I recommend the following be done, in order of declining importance.

1. **Use displays with minimum response time.** Display response time is currently the greatest contributor to latency in today's systems (See Table 2.1 for example response times of two popular HMDs). Developers should not waste their efforts improving the speed of their rendering algorithms by shaving off a few milliseconds if a slow response HMD with 100 *ms* of latency is used.
2. **Use displays without internal frame buffers.** Buffering images for an entire frame time within the display hardware adds 16.7 *ms* of delay for a display with a 60 Hz update rate. Unfortunately many modern displays use a frame buffer so that all pixels are stored before displaying (Field-Sequential Displays of Section 2.4.3.4). Using a display that draws pixels as they are received from the image generator in scanline order is ideal (Raster Displays of Section 2.4.3.4).
3. **Use fast tracker systems.** Use a low-latency tracker with a fast and consistent update rate. An example of such a tracker is the 3rdTech Hi-BallTM tracking system that has approximately 1 *ms* of latency and an update rate of ~ 1500 *Hz*.
4. **Do not wait on vertical sync.** Not waiting to swap on vertical sync can approximate a just-in-time scanlines approximation (Raster Displays of Section 2.4.3.4). This comes at the cost of image tearing, but if the update rate is fast enough (e.g., ~ 1000 *Hz*) tearing can be reduced to imperceptible levels.
5. **Use minimal tracker filtering.** Smoothing tracker data comes at the cost of latency.
6. **Use fast Rendering algorithms.** Many developers focus on optimizing the rendering stage of IVE systems when in fact this stage offers minor speed up compared to other stages (when using modern graphics hardware). However, for complex scenes, optimizing this stage can offer significant speed up. For especially complex scenes, use post-rendering techniques to shift pixels at the last possible moment before display (Section 2.4.4.2).
7. **Reduce synchronization delay.** Asynchrony among components can add significant and variable delay (Sections 2.4.3.5 and 2.4.3.6). However, if all components have a fast update rate, then synchronization delay will be low.
8. **Prediction.** Prediction of head movement can be used to reduce latency-induced scene motion (Section 2.4.4.1). However, it is impossible to perfectly predict human motion,

and scene motion artifacts occur if prediction is not perfect. These artifacts sometimes result in more scene motion than would occur with no prediction. These scene motion artifacts are roughly proportional to the square of the angular head frequency and the square of the prediction interval (Azuma and Bishop, 1995). Minimizing latency before doing prediction is essential to minimizing scene motion. Also, consistent delay is important in order to know what time to predict for.

REFERENCES

- 3rdTech (2002). Hiball-3000 wide-area tracker. Technical Manual.
- Adelstein, B. D., E. M. Burns, S. R. Ellis, and M. I. Hill (2005). Latency discrimination mechanisms in virtual environments: Velocity and displacement error factors. In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, pp. 2221–2225.
- Adelstein, B. D., T. Lee, and S. R. Ellis (2003). Head tracking latency in virtual environments: Psychophysics and a model. In *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*, pp. 2083–2987.
- Adelstein, B. D., L. Li, J. Jerald, and S. R. Ellis (2006). Suppression of head-referenced image motion during head movement. In *Proceedings of the 50th Annual Meeting of Human Factors and Ergonomics Society*, pp. 2678–2682.
- Allison, R. S., L. R. Harris, M. Jenkin, U. Jasiobedzka, and J. E. Zacher (2001). Tolerance of temporal delay in virtual environments. In *Proceedings of IEEE Virtual Reality*, pp. 247–253.
- Anstis, S. (1986). Motion perception in the frontal plane: Sensory aspects. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 16. New York: Wiley-Interscience.
- Azuma, R. T. (1997). A survey of augmented reality. *PRESENCE: Teleoperators and Virtual Environments* 6(4), 355–385.
- Azuma, R. T. and G. Bishop (1994). Improving static and dynamic registration in an optical see-through HMD. In *Proceedings of ACM SIGGRAPH 94*, pp. 197–204.
- Azuma, R. T. and G. Bishop (1995). A frequency-domain analysis of head-motion prediction. In *Proceedings of ACM SIGGRAPH 95*, pp. 401–408.
- Breglia, D., M. Spooner, and D. Lobb (1981). Helmet mounted laser projector. In *Proceedings of the Image Generation/Display Conference II*, pp. 241–258.
- Bridgeman, B., A. V. d. Heijden, and B. M. Velichkovsky (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences* 17(2), 247–292.

- Bryson, S. and S. Johan (1996). Time management, simultaneity and time-critical computation in interactive unsteady visualization environments. In *Proceedings of IEEE Visualization '96*, pp. 255–262.
- Burbidge, D. and P. Murray (1989). Hardware improvements to the helmet mounted projector on the visual display research tool (vdrt) at the naval training systems center. In *Proceedings of SPIE Head-Mounted Displays 1989*.
- Chen, S. E. (1995). Quicktime VR—an image-based approach to virtual environment navigation. In *Proceedings of ACM SIGGRAPH 95*, pp. 29–38.
- Coren, S., L. M. Ward, and J. T. Enns (1999). *Sensation and Perception* (5th ed.). Harcourt Brace College Publishers.
- Cunningham, D., V. Billock, and B. Tsou (2001). Sensorimotor adaptation to violations in temporal contiguity. *Psychological Science* 12(6), 532–535.
- Cunningham, D., A. Chatziastros, M. V. D. Heyde, and H. Bulthoff (2001). Driving in the future: Temporal visuomotor adaptation and generalization. *Journal of Vision* 1(2), 88–98.
- Denne, P. (2004, September 2004). Motion platforms or motion seats? Technical report, Guilden Ltd.
- Draper, M. H. (1998). *The Adaptive Effects of Virtual Interfaces: Vestibulo-Ocular Reflex and Simulator Sickness*. Dissertation, University of Washington.
- Efron, R. (1967). The duration of the present. *Annals of the New York Academy of Science* 138, 713–729.
- Ellis, S. R., B. D. Adelstein, S. Baumeler, G. J. Jense, and R. H. Jacoby (1999). Sensor spatial distortion, visual latency, and update rate effects on 3d tracking in virtual environments. In *Proceedings of IEEE Virtual Reality*, pp. 218–221.
- Ellis, S. R., K. Mania, B. D. Adelstein, and M. I. Hill (2004). Generalizeability of latency detection in a variety of virtual environments. In *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society*, pp. 2083–2087.
- Ellis, S. R., M. J. Young, B. D. Adelstein, and S. M. Ehrlich (1999). Discrimination of changes in latency during head movement. In *Proceedings of Human-Computer Interaction 99*, pp. 1129–1133.
- Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Hrtel.

- Ferwerda, J. A. (2008). Psychophysics 101: How to run perception experiments in computer graphics. In *ACM SIGGRAPH 2008 Courses*, pp. 1–60.
- Foley, J., A. Dam, S. Feiner, and J. Hughes (1990). *Computer Graphics: Principles and Practice*. Reading, Massachusetts: Addison-Wesley Publishing.
- Freeman, T. and M. Banks (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research* 38(7), 941–945.
- Gescheider, G. A. (1997). *Psychophysics: The Fundamentals* (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Gogel, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception and Psychophysics* 48(2), 105–123.
- Goldstein, E. B. (2007). *Sensation and Perception* (7th ed.). Belmont, CA: Wadsworth Publishing.
- Graham, C. (1965). Perception of movement. In C. Graham (Ed.), *Vision and Visual Perception*, Chapter 20. New York: John Wiley and Sons, Inc.
- Greene, N. (1986). Environment mapping and other applications of world projections. *IEEE Computer Graphics and Applications* 6(11), 21–29.
- Gregory, R. L. (1973). *Eye and Brain: The Psychology of Seeing* (2nd ed.). London: Weidenfeld and Nicolson.
- Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 10. New York: Wiley-Interactive.
- Harm, D. L. (2002). Motion sickness neurophysiology, physiological correlates, and treatment. In K. M. Stanney (Ed.), *Handbook of Virtual Environments, Human Factors and Ergonomics*, pp. 637–661. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika* 75(4), 800–802.
- Holloway, R. (1997). Registration error analysis for augmented reality. *Presence: Teleoperators and Virtual Environments* 6(4), 413–432.
- Holloway, R. L. (1995). *Registration Errors in Augmented Reality Systems*. Ph.d. dissertation, Department of Computer Science, University of North Carolina at Chapel Hill.

- Howard, I. P. (1986a). The perception of posture, self motion, and the visual vertical. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 18. New York: Wiley-Interscience.
- Howard, I. P. (1986b). The vestibular system. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 11. New York: Wiley-Interscience.
- Jaekl, P. M., R. S. Allison, L. R. Harris, U. T. Jasiobedzka, H. L. Jenkin, M. R. Jenkin, J. E. Zacher, and D. C. Zikovitz (2002). Perceptual stability during head movement in virtual reality. In *Proceedings of IEEE Virtual Reality*, pp. 149–155.
- Jerald, J., A. Fuller, A. Lastra, M. Whitton, L. Kohli, and F. Brooks (2007). Latency Compensation by Horizontal Scanline Selection for Head-Mounted Displays. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, Volume 6490.
- Jerald, J., T. Peck, F. Steinicke, and M. Whitton (2008). Sensitivity to scene motion for phases of head yaws. In *Proceedings of the Symposium on Applied Perception in Graphics and Visualization*, pp. 155–162.
- Jerald, J. and F. Steinicke (2009). Scene instability during head turns. In *Proceedings of IEEE Virtual Reality Workshop on Perceptual Illusions in Virtual Environments (PIVE)*, pp. 4–6.
- Jerald, J., F. Steinicke, and M. Whitton (2009). Scene-motion thresholds correlate with angular head motions for immersive virtual environments. In *Proceedings of Advances in Computer-Human Interaction*, pp. 69–75.
- Jerald, J. and M. Whitton (2009). Relating scene-motion thresholds to latency thresholds for head-mounted displays. In *Proceedings of IEEE Virtual Reality*, pp. 211–218.
- Kano, K., K. Kitabayash, and R. Kijima (2004). Reflex head mounted display: Head mounted display for virtual reality with time lag compensation. In *Proceedings of Virtual Systems and Multimedia*, pp. 119–127.
- Kennedy, R. S. and M. G. Lilienthal (1995). Implications of balance disturbances following exposure to virtual reality systems. In *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS)*, pp. 35–39.
- Kijima, R. and T. Ojika (2002). Reflex HMD to compensate lag and correction of derivative deformation. In *Proceedings of IEEE Virtual Reality*, pp. 172–179.

- Kijima, R., E. Yamada, and T. Ojika (2001). A development of reflex HMD—HMD with time delay compensation capability. In *Proceedings of the 2nd International Symposium on Mixed Reality (ISMR2001)*, pp. 172–179.
- Kolasinski, E. (1995). Simulator sickness in virtual environments. Technical report, U.S. Army Research Institute for the Behavioral and Social Sciences.
- Kollin, J. (1993). A retinal display for virtual-environment applications. In *Proceedings of the Society for Information Display*, Volume XXIV.
- Lackner, J. R. and R. Teixeira (1977). Visual-vestibular interaction: Vestibular stimulation suppresses the visual induction of illusory self-rotation. *Aviation, Space and Environmental Medicine* 48, 248–253.
- Levitt, H. (1970). Transformed up-down methods in psycho-acoustics. *Acoustical Society of America* 49(2), 467–477.
- Li, L., B. D. Adelstein, and S. R. Ellis (2006). Perception of image motion during head movement. In *Proceedings of the ACM Symposium on Applied Perception in Graphics and Visualization*, pp. 45–50.
- Lisberger, S. G. and J. A. A. Movshon (1999). Visual motion analysis for pursuit eye movements in area mt of macaque monkeys. *The Journal of Neuroscience* 19(6), 2224–2246.
- List, U. H. (1983). Nonlinear prediction of head movements for helmet-mounted displays. Technical report, Operations Training Division, Williams Air Force Base, Arizona.
- Loomis, J. M. and J. M. Knapp (2003). Visual perception of egocentric distance in real and virtual environments. In L. J. Hettlinger and M. W. Haas (Eds.), *Virtual and adaptive environments*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Loose, R. and T. Probst (2001). Velocity not acceleration of self-motion mediates vestibular-visual interaction. *Perception* 30(4), 511–518.
- Mack, A. (1986). Perceptual aspects of motion in the frontal plane. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 17. New York: Wiley-Interscience.
- Mack, A. and E. Herman (1972). A new illusion: The underestimation of distance during pursuit eye movements. *Perception and Psychophysics* 12, 471–473.
- Macmillan, N. A. and C. D. Creelman (2005). *Detection Theory: A User's Guide* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

- Mania, K., B. D. Adelstein, S. R. Ellis, and M. I. Hill (2004). Perceptual sensitivity to head tracking in virtual environments with varying degrees of scene complexity. In *Proceedings of the ACM Symposium on Applied Perception in Graphics and Visualization*, pp. 39–47.
- Mark, W. R., L. McMillan, and G. Bishop (1997). Post-rendering 3d warping. In *Proceedings of Interactive 3D Graphics*, pp. 7–16.
- Martini, F. (1998). *Fundamentals of Anatomy and Physiology*. Upper Saddle River, Prentice Hall.
- May, J. G. and D. R. Badcock (2002). Vision and virtual environments. In K. M. Stanney (Ed.), *Handbook of Virtual Environments, Human Factors and Ergonomics*, pp. 29–63. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Mazuryk, T. and M. Gervautz (1995). Two-step prediction and image deflection for exact head tracking in virtual environments. In *Proceedings of EUROGRAPHICS 95*, Volume 14 (3), pp. 30–41.
- Meehan, M., F. Brooks, S. Razzaque, and M. Whitton (2003). Effects of latency on presence in stressful virtual environments. In *Proceedings of IEEE Virtual Reality*, pp. 141–148.
- Miller, D. (2002). Personal Communication.
- Miller, D. and G. Bishop (2002). Latency meter: A device for easily monitoring VE delay. In *Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems IX*, Volume 4660.
- Miné, M. and G. Bishop (1993). Just-in-time pixels. Technical Report TR93-005, Department of Computer Science, University of North Carolina at Chapel Hill.
- Miné, M. R. (1993, March 1993). Characterization of end-to-end delays in head-mounted display systems. Technical report TR93-001, Department of Computer Science, University of North Carolina at Chapel Hill.
- Murray, P. M., G. J. Olive, M. Roberts, and O. Wynn (1984). Air combat visual simulation using a head slaved projector. In *Proceedings of the Image III Conference*, pp. 295–311.
- Nakanishi, K., S. Takahashi, H. Oura, T. Matsumura, S. Miyake, K. Kobayashi, K. Oda, S. Tahata, A. Yuuki, J. Someya, M. Yamakawa, and E. Gofuku (2001). A fast-response 15-in. xga tft-lcd with feed-forward driving (ffd) technology for multimedia application. In *Proceedings of the Society for Information Display*, pp. 488–491.
- Nakayama, K. and C. Tyler (1981). Psychological isolation of movement sensitivity by removal of familiar position cues. *Vision Research* 21, 427–433.

- Nillson, T. H. (1969). Two-pulse-interval vision thresholds. *Journal of the Optical Society of America* 59(6), 753–756.
- Olano, M., J. Cohen, M. Miné, and G. Bishop (1995). Combatting rendering latency. In *Proceedings of the ACM Symposium on Interactive 3D Graphics*, pp. 19–24.
- Pausch, R., T. Crea, and M. J. Conway (1993). A literature survey for virtual environments: Military flight simulator visual systems and simulator sickness. *PRESENCE: Teleoperators and Virtual Environments* 1(3).
- Peck, T., M. Whitton, and H. Fuchs (2008). Evaluation of reorientation techniques for walking in large virtual environments. In *Proceedings of IEEE Virtual Reality*, pp. 121–127.
- Razzaque, S. (2004). Personal Communication.
- Razzaque, S. (2005). *Redirected Walking*. Dissertation, Department of Computer Science, University of North Carolina at Chapel Hill.
- Razzaque, S., Z. Kohn, and M. Whitton (2001). Redirected walking. In *Eurographics*, pp. 289–294.
- Regan, D. M., L. Kaufman, and J. Lincoln (1986). Motion in depth and visual acceleration. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 19. New York: Wiley-Interscience.
- Regan, M., G. Miller, S. Rubin, and C. Kogelnik (1999). A real-time low-latency hardware light-field renderer. In *Proceedings of ACM SIGGRAPH 99*, pp. 287–290.
- Regan, M. and R. Pose (1994). Priority rendering with a virtual reality address recalculation pipeline. In *Proceedings of ACM SIGGRAPH 94*, pp. 155–162.
- Robinson, D. (1986). Control of eye movements. In V. Brooks (Ed.), *The Handbook of Physiology*, Volume II, Part 2, The Nervous System, pp. 1275–1320. Baltimore: Williams and Wilkens.
- Rolland, J. and O. Cakmakci (2005). The past, present, and future of head mounted display designs. In *SPIE International Society of Optical Engineers*, Volume 5638, pp. 368–377.
- Sherman, W. R. and A. B. Craig (2003). *Understanding Virtual Reality*. The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling. Morgan Kaufmann Publishers.
- Slater, M. and A. Steed (2000). A virtual presence counter. *Presence: Teleoperators and Virtual Environments* 9(5), 413–434.

- So, R. and M. Griffin (1995). Effects of lags on human operator transfer functions with head-coupled systems. *Aviation, Space and Environmental Medicine* 66(6), 550–556.
- So, R. H. Y. and M. J. Griffin (1992). Compensating lags in head-coupled displays using head position prediction and image deflection. *Journal of Aircraft* 29(6), 1064–1068.
- Steed, A. (2008). A simple method for estimating the latency of interactive real-time graphics simulations. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST)*, pp. 123–129.
- Steinicke, F. (2009). Personal Communication.
- Steinicke, F., G. Bruder, J. Jerald, H. Frenz, and M. Lappe (2008). Analyses of human sensitivity to redirected walking. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST)*, pp. 149–156.
- Steinicke, F., G. Bruder, J. Jerald, H. Frenz, and M. Lappe (2010). Psychophysiological estimation of detection thresholds for redirected walking techniques. *Transactions on Visualization and Computer Graphics*, to appear.
- Stoffregen, T. A., M. H. Draper, R. S. Kennedy, and D. Compton (2002). Vestibular adaptation and aftereffects. In K. M. Stanney (Ed.), *Handbook of Virtual Environments, Human Factors and Ergonomics*, pp. 773–790. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Stroud, J. (1986). The fine structure of psychological time. In H. Quastler (Ed.), *Information Theory in Psychology: Problems and Methods*, pp. 174–207. Glencoe, IL: Free Press.
- Swan, J. E., S. R. Ellis, and B. D. Adelstein (2007). Tutorial proceedings: Conducting human-subject experiments with virtual and augmented reality. In *Proceedings of IEEE Virtual Reality*.
- Ulrich, R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception and Psychophysics* 42, 224–239.
- Wallach, H. (1987). Perceiving a stable environment when one moves. *Annual Review of Psychology* 38, 1–27.
- Wallach, H. and T. Canal (1976). Two kinds of adaptation in the constancy of visual direction. *Perception and Psychophysics* 19, 445–449.
- Wallach, H. and J. H. Kravitz (1965). The measurement of the constancy of visual direction and of its adaptation. *Psychonomic Science* 2, 217–218.

- Wallach, H. and J. H. Kravitz (1968). Adaptation in the constancy of visual direction tested by measuring the constancy of auditory direction. *Perception and Psychophysics* 4(5), 299–303.
- Welch, R. B. (1986). Adaptation of space perception. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 24. New York: Wiley-Interscience.
- Welch, R. B. and D. Warren (1986). Intersensory interactions. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Volume 1, Chapter 25. New York: Wiley-Interscience.
- Whitton, M. C. (1984). Memory design for raster graphics displays. *IEEE Computer Graphics and Applications* 4(3), 48–65.
- Yanagida, Y., M. Inami, and S. Tachi (1998). Improvement of temporal quality of HMD for rotational motion. In *Proceedings of the 7th IEEE International Workshop on Robot and Human Communication (RO-MAN)*, pp. 121–126.