SUPPRESSED PERCEPTION OF HEAD-REFERENCED IMAGE MOTION DURING HEAD MOVEMENT

B.D. Adelstein, L. Li[1], J.J. Jerald[2], S.R. Ellis

Human Systems Interface Division, NASA Ames Research Center

[1]San Jose State University Foundation
[2]University of North Carolina at Chapel Hill

A previous investigation in our laboratory showed that virtual environment users rely on image velocity errors rather than image displacement errors to sense head-tracking latency. In that study, the effect of image displacement error may have been suppressed because its peak amplitudes were associated with higher head velocity, which is thought to suppress visual motion sensitivity. To determine whether diminished motion sensitivity could play a role in user discrimination of latency, we investigated whether subjective perception of image motion comparable to that caused by latency is impaired by concurrent horizontal or vertical head movement. Eight participants monocularly viewed a checkerboard pattern in a head-mounted display that was oscillated from side-to-side with half peak-to-peak amplitudes from 0° to 5.64°. Perceptual sensitivity to image motion amplitude was reduced by almost half, during 30° repetitive horizontal head movements at 0.4 Hz. Accordingly, the reduced sensitivity to motion during head movement appears to be a phenomenon that modulates users' tolerance of erroneous image motion caused by virtual environment latency.

INTRODUCTION

Virtual environment (VE) system latency and imperfect predictive compensation for such latencies produce dynamically varying displacement errors in image rendering (Adelstein, Lee, & Ellis, 2003; Azuma & Bishop, 1994). If sufficiently large, such errors will cause the virtual world or objects in that world to appear to move with respect to the external real world.

In a prior study (Adelstein et al., 2003), we noted that in VEs with short to moderate delays (i.e., <150 ms) observer sensitivity to system latency cannot be due solely to direct perception of the elapsed temporal interval. Instead, to discriminate the presence of latency, observers make use of the consequent image "slip" visible in the head-tracked head-mounted display (HMD). We define image slip as the dynamically varying offset of a virtual object or scene from its intended spatially stabilized location. This offset is caused by the difference (i.e., error) between the time-delayed version of the user's head rotation/position that results in the currently viewed VE image and the user's actual current head rotation/position at any particular instant. Referring to Figure 1, slip can be described in terms of displacement, velocity, or even higher-order (not depicted) spatial derivatives. As discussed by Adelstein et al. (2003), the peak magnitude of any of these kinematic measures of image slip grows with the amount of added VE latency.

In a subsequent study (Adelstein, Burns, Ellis, & Hill, 2005), we reported that observers used the velocity of image slip rather than its displacement as a basis for discriminating latency during repetitive sinusoidal head yaw movements. However, because of the sinusoid-like head movement in that experiment, the highest image slip velocities coincided with yaw direction reversals, when head velocity, as depicted schematically in Figure 1, was near zero. Conversely, peak displacement of the image slip occurred near the straight-ahead direction when head movement velocity was at its highest.

Previous research has suggested that head movement can diminish perception of image motion, regardless of whether this motion is world- or head-referenced. Probst, Brandt, and Degner (1986) reported higher thresholds to visually presented object motion in a variety of settings when the head was rotating with respect to the body. They attributed this suppression to a number of head movement sensory systems. In fact, such suppression was also observed when image motion was head-referenced by presentation in an HMD that effectively eliminated the influence of the vestibulo-ocular reflex (VOR). Probst et al., however, measured time-to-onset of velocity detection, rather than changes in perceived motion amplitude. Recently, Durgin, Gigone, and Scott (2005) showed that subjective magnitude estimation of speed from visual flow in an HMD could be reduced both by biomechanical self-motion (regular and treadmill walking) as well as during passive motion while being pushed forward or backward on a chair.

Therefore, the question arises as to whether the subjects' ability to detect image slip displacement in our previous study (Adelstein et al., 2005) could have been impaired by their concurrent head movement. Because of this resulting impairment, subjects may have been instead forced to rely mainly on velocity of the image slip, the dominant signal available when head velocity was near zero during their repetitive yaw movement cycle. The goal of the present study is to examine the potential effect of head movement on perceived concurrent image motion in an HMD, and to make a preliminary assessment of the magnitude of such an effect.

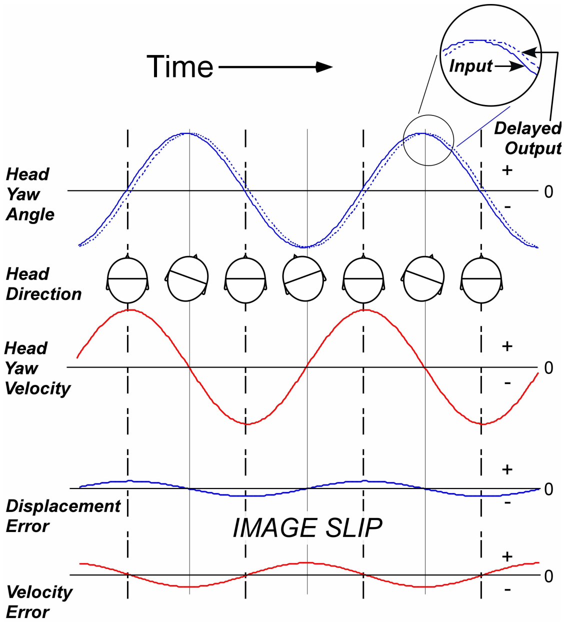

Figure 1. Image slip as a consequence of head tracking delay. The difference between actual and delayed head angle due to VE latency (top trace) produces instantaneous displacement error (2nd trace from bottom) of the VE from ideal spatially constant location. For ideal, single-axis sinusoidal yaw rotations, peak displacement errors coincide approximately with peak head yaw velocities (3rd trace from bottom trace), when the head icon direction is straight-ahead. Conversely, peak velocity errors (bottom trace) occur at the ends of travel, when head velocity passes through zero at direction reversals.

METHODS

Stimulus images were presented to the subjects' right eye via a Kaiser ElectroOptics SR80 HMD. The HMD provided SVGA resolution (1280 X 1024 pixels) with a 64° H X 48° V monocular field of view (FOV). The response time, which is uniform across the entire color-sequential ferroelectric liquid crystal on silicon (FLCOS) display element in this HMD, is ~100 µs per individual color bit-plane. The FLCOS rendering hardware eliminates the line-by-line scanning typical of raster type displays as well as the smearing of moving images endemic to the slow dynamic response of the TFT-LCD picture elements in prior generation HMDs (e.g., Virtual Research V8) and other conventional flat-panel desktop displays. The HMD was covered with an opaque hood and the experiment room was darkened in order to obstruct any external visual stimulus.

The stimulus image was a red checkerboard pattern shown in Figure 2 that covered the HMD's entire right-eye FOV. The bright red squares were set to 8-bit RGB triples of (100,0,0); the dark squares were (50,0,0). Each square subtended a visual angle of ~6.8° on a side. The Michelson luminance contrast between the two square colors was 58% as measured from an external printed checkerboard in our laboratory that was subjectively adjusted by the authors to match the HMD image. A Gaussian window ($\sigma = 12°$) was used to obscure sharp edges of the checkerboard toward the boundaries of the FLCOS image element, which otherwise could have provided differential motion cues.

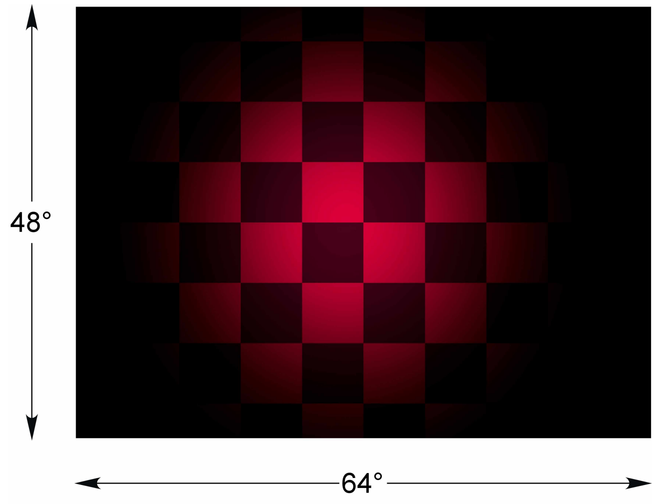

Figure 2. Gaussian-windowed checkerboard stimulus.

The checkerboard pattern was oscillated from side-to-side at 0.4 Hz (full cycle) in each trial at one of seven amplitudes. The half peak-to-peak amplitudes selected were 0°, 0.94°, 1.88°, 2.82°, 3.76°, 4.70°, and 5.64°, nominally equivalent to the image slip expected for 20 ms steps in VE latencies from 0 to and including120 ms (e.g., Adelstein et al., 2005). To eliminate VOR as a factor in subject response, the image was stabilized with respect to the HMD, i.e., the subject's head, not the external world. Moreover, since head tracking is not involved at all in visual stimulus presentation, intrinsic VE system latency has no impact on subject response.

Subjects were seated and, depending on the experiment block, either kept their head still while gazing straight ahead, yawed their head from side to side, or pitched up and down. In all conditions, they made subjective magnitude estimates of the amplitude of the viewed head-referenced motion of the checkerboard. During the yaw (horizontal) and pitch (vertical) head movements, a computer-generated metronome with two beeps/cycle paced the subjects in a 0.5 Hz repetitive motion pattern lasting three full cycles. This motion frequency was selected to match that from our previous studies (Adelstein et al., 2005). Head movements began either from the left (yaw) or at the bottom (pitch). An audible alarm based on the instantaneous reading by a Polhemus FasTrak of HMD rotation reminded participants to limit yaw to approximately ±15° about the straight-ahead direction and pitch to 10° deg above and 20° below the horizontal. The alarm comprised a single-frequency 600 Hz tone, which was audibly distinct from the metronome beeps composed of four concurrent frequencies (0.11, 0.5, 1.6, and 7 KHz). Moreover, unlike the ~180 ms duration of the metronome beeps, the alarm tone continued to sound as long as the subject's head was outside the acceptable yaw or pitch zones.

The checkerboard image was visible for only 5 s of the 6 s comprising each head movement interval. The HMD image was otherwise black. The head movement start was randomized and therefore uncorrelated with respect to the image motion cycle for any trial during which both the head and image moved.

During each head movement block, subjects viewed five repetitions of the seven image motion (0° stationary plus six moving) levels. The resulting 35-trial sequence was presented in the same randomized order for all subjects. The first block was a no-head-motion condition to get a baseline judgment scale. The subjects completed the horizontal and vertical motion blocks, with the order of the two directions balanced across subjects. A final, fourth no-motion block was used to determine whether there had been any shift in subject judgment criteria.

In order to train and anchor their individual response scales at the start of each block, subjects were shown 0°, 2.82°, and 5.64° image motion while they kept their head still. Subjects were instructed to observe these motions and use them as references for 0, 50, and 100% of maximum motion amplitude. The anchor stimuli were presented prior to each block and subjects could view these anchor stimuli as many times as they wished prior to starting. We expected that subjects would make their judgments with respect to the moving head (HMD) reference frame. The 0° image motion (i.e., stationary) condition, inserted repeatedly throughout each block, permitted us to verify whether the subjects maintained this reference frame.

During the experiment, subjects were instructed to rate the magnitude of the checkerboard movement on a continuous scale by verbally announcing their evaluation as a percentage relative to the 0-100% trained range. The experiment monitor then recorded subjects' response and advanced to the next stimulus presentation. Each 35-trial block typically lasted 15 minutes, with the experiment completed in about 1 hour on the same day.

Data were collected from eight participants (3 F, 5 M) aged 28 to 45 years who were recruited from colleagues as well as the paid subject pool at the Ames Research Center. All had normal or corrected-to-normal vision and were naïve to the purpose of study.

RESULTS

First, we computed the medians from the five responses provided by each subject for each image motion level within each head movement condition. Zero image motion levels were analyzed separately from the remainder of the stimuli. All but four of the 32 median responses (four head movement blocks with zero image motion by eight subjects) were reported as 0% amplitude. The remaining four medians ranged between 1 and 4%. A Friedman non-parametric analysis of variance did not show significant differences in judged percentage across the four blocks for zero image motion ($\chi^2 = 2.25$; $p > 0.52$). Separate t-tests did not show the mean judgment at zero image motion for any head movement condition to be significantly different from zero ($p < 0.09$). These observations for the zero image motion condition attest to the subjects maintaining an HMD (i.e., head-based) reference frame when making all of their judgments.

Next, linear regressions of median reported judgment percentage as a function of the six non-zero image motion amplitudes were performed separately for each subject for each of the four head movement conditions. Input stimulus levels were first converted to percentage (i.e., 100% stimulus equaling the maximum 5.64° image motion amplitude) such that a regression with unity slope and zero intercept would correspond to a perfectly accurate subjective response. The slope of the linear regression represents the proportional gain, or sensitivity, in response for a given input stimulus level. The intercept corresponds to a bias or offset in the fitted linear function representing subject response. Figure 3 illustrates median responses and the resultant regression fits for all four head movement conditions from one subject.

The Pearson correlations for all 32 regressions (4 head movement conditions X 8 subjects) performed were statistically significant ($r > 0.707$, $df = 5$, $p < 0.05$). The subsequent analyses and discussion rely on the slopes and intercepts computed from the regressions.

Slope (Sensitivity)

Average regression slopes (mean ± std error) for the "before" no head movement condition (1.032 ± 0.073) and the "after" condition (0.960 ± 0.075) were not significantly different according to a planned pairwise (by subject) t-test ($t_7 = 2.365$, $p_{two\text{-}tail} < 0.28$). Hence, we averaged the "before" and "after" slope for each subject, yielding a single, combined "no head movement" value. The overall mean across subjects for the combined no head movement slopes (0.996 ± 0.067) was not significantly different from the ideal unity slope response to which the subjects were trained ($t_7 = 0.059$, $p_{two\text{-}tail} > 0.95$).

The one-way ANOVA on the regression slopes for the three remaining head motion conditions (vertical, horizontal, none) was significant ($F_{2,14} = 8.882$; $p < 0.003$). Post-hoc Scheffé contrasts of the slopes for the three conditions plotted in Figure 4 show that the suppression for the horizontal versus the no head movement condition was significant (difference $\Delta = 0.458$, $p < 0.005$), as was the difference between the horizontal and vertical conditions ($\Delta = 0.282$, $p < 0.07$). The vertical and no motion conditions did not differ significantly ($\Delta = 0.175$, $p > 0.3$).

Intercept (Bias)

A planned pairwise (by subject) t-test between the regression intercepts for the "before" (-7.9% ± 3.3%, mean ± std error) and "after" (0.3% ± 3.5%) no head movement condition indicates that the change in bias was significant ($t_7 = 4.079$, $p_{two\text{-}tail} < 0.005$). Therefore the intercepts from the initial and final movement blocks could not be combined into a single no-movement condition for subsequent analysis.

The one-way ANOVA on regression intercept plotted in Figure 5 for the four head movement conditions (before no movement, vertical, horizontal, after no movement) was significant ($F_{3,21} = 7.247$; $p < 0.002$). Post-hoc Scheffé contrasts showed that while the "before" no-movement condition was not significantly different from either the horizontal ($\Delta = 0.1\%$, $p > 0.99$) or vertical ($\Delta = 4.2\%$, $p < 0.4$) conditions, the "after" condition was (Horizontal $\Delta = 8.3\%$, $p < 0.05$; Vertical $\Delta = 12.2\%$, $p < 0.005$). The difference between the horizontal and vertical intercepts was not significant ($\Delta = 3.8\%$, $p < 0.6$).

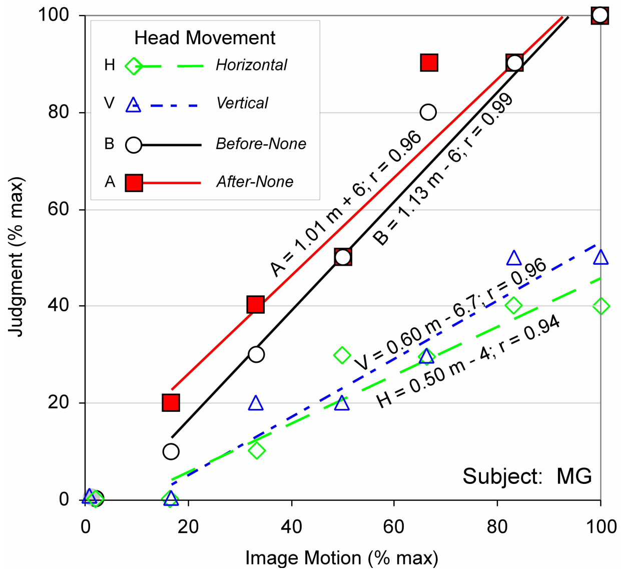

Figure 3. Medians of image motion judgments from one subject at all image motion stimulus levels for each head movement condition (*H* = horizontal; *V*= vertical; *B* = before no-movement; *A* = after no-movement). The regression equation for each line, $J = S\ m + I$, relates reported judgment percentage, *J* (i.e., *H*, *V*, *A*, or *B* for the particular head movement condition), to the input image motion percentage, *m*. *S* and *I*, respectively, are the regression slope and intercept. *r* is the Pearson correlation between input image motion and the six median judgments contributing to each regression. Note that the summary least-squares regression line plotted for each head movement condition does not include zero image motion level.

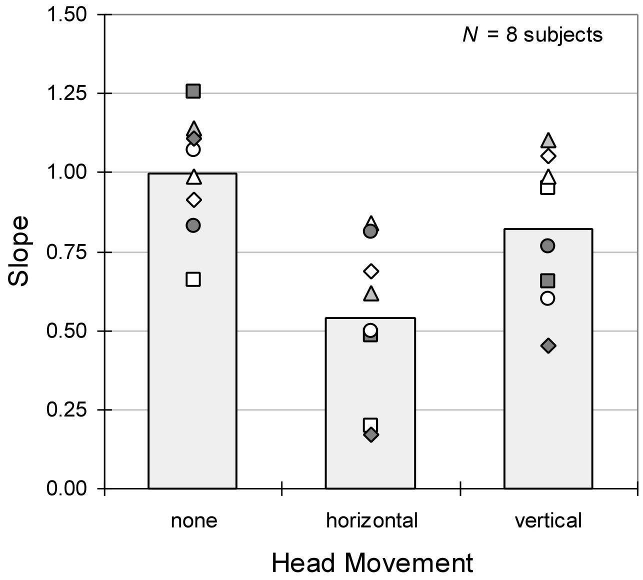

Figure 4. Mean and individual subject slopes for averaged before and after no head movement (i.e., "none"), horizontal and vertical head movement conditions. Slope is expressed in terms of percentage judgment level divided by percentage image motion, and is therefore dimensionless. Unique markers for individual subject data indicate that seven of the eight subjects had minimum slope for the horizontal condition.

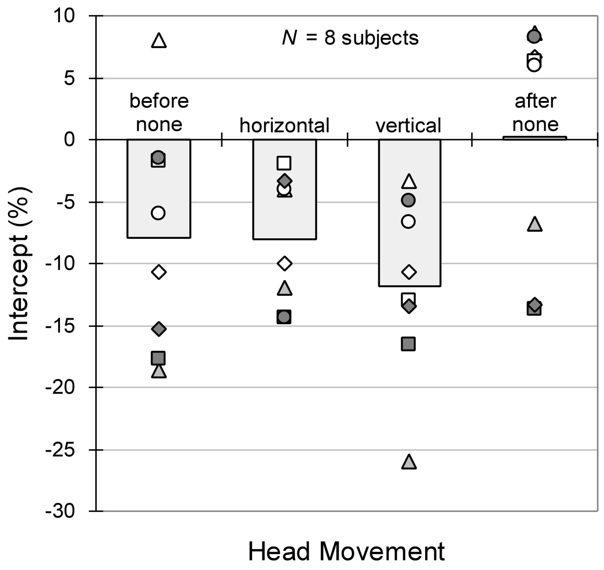

Figure 5. Mean and individual subject intercepts in terms of judgment percent levels for the before and after no-movement and the horizontal and vertical head movement conditions. Intercept is expressed in terms of the judgment percent level.

DISCUSSION

We observed that horizontal head movement reduced the subjects' perceived magnitude of the horizontal image motion by nearly half from the veridical levels reported when the head was still for the range of image motion amplitudes examined. This result, reflected by a decrease in linear regression slope, suggests suppression of "gain," or, equivalently, "sensitivity" while the head was moving. The reduction in sensitivity to horizontal image motion with vertical head movement was not statistically significant. A subsequent study from our lab (Li, Adelstein, Ellis, 2006), in which a range of image frequencies were examined, demonstrated respectively similar reductions in the average perceptual sensitivity to image motion for horizontal and vertical head movement with a different group of participants. However, in this second study, the reduction in sensitivity with vertical head movement did turn out to be significant.

No significant change in bias of the subjective judgments was observed for either horizontal or vertical head movement from the initial no-movement condition. However, there was a significant jump in bias from these three blocks to the final no head movement condition, suggesting a possible criterion shift in subject response. At this point we do not have an explanation for this shift.

Wallach (1987) proposed that the inhibitory effect of self-motion on visual perception of motion reflects a compensation process that stabilizes the world during motor activities. For instance, head movements produce viewer-relative motions similar to those produced when the world is moving. Nevertheless, we normally perceive the world as being stationary. In fact, perceived motion smear has been reported to be lower during VOR-compensated movements than during fixation, particularly for target durations of 100 ms or longer (Bedell & Patel, 2005).

In the present study, image motion was head-referenced by means of presentation in the HMD. The zero magnitude judgments reported for the zero amplitude image motion condition in the present study suggests that the subjects actually employed a reference frame that moved with the head rather than one fixed in the external world. The effect is therefore not attributable to confusion of which frame of reference should be used to estimate the motion amplitude. If eye orientation remained fixed with respect to the head during the experiment, then any image motion on the retina would be governed solely by the movement of the checkerboard pattern in the HMD. However, while such eye fixation relative to the head would essentially suppress the VOR, eye motion was neither measured nor explicitly controlled in the present study. Consequently, an effort was made in our subsequent study (Li et al., 2006) to provide participants with a fixation target (a cross region of dark squared) embedded in the checkerboard pattern. Reduction of image motion perception during head movement if VOR were truly suppressed would therefore support the idea of sensory inhibitory processes based on inner ear (e.g. otoliths) signals or proprioceptive systems.

Barlow (1990) proposed the inhibition theory of sensory correlation as a mechanism for Wallach's (1987) self-motion compensation process. According to this theory, highly correlated events come to mutually specify one another and consequently produce inhibitory interactions between their respective sensory coding. The perceived reduction of image motion amplitude during head movement in this account "serves the functions of de-emphasizing predictable events in favor of detecting deviations from the norm" (Durgin et al., 2005). The findings from our current study support this theory by showing that horizontal head movement, which is more highly correlated with the side-to-side image motion, produces a larger reduction in perceptual sensitivity to image motion than does vertical head movement.

Barlow (1990) did not specify whether the inhibitory process involves a bias shift in response or a gain. Durgin et al. (2005) proposed that the reduction in perceived image motion magnitude in their studies could be modeled by a "subtractive" operation, i.e., a bias shift in response. As noted, suppression of image motion sensitivity with horizontal head movement in the present study indicates a gain reduction, corresponding to what Durgin et al. termed a "divisive" model. Aside from experimental task and stimuli, the difference in our interpretation of the suppression mechanism may stem from the subjective magnitude estimation and data analysis procedures employed. Durgin et al. provided their subjects with only one reference stimulus, termed "100," for comparison, while we tied both ends as well as the middle in an effort to both anchor and linearize our estimation scale. Additionally, Durgin et al.'s use of logarithmically transformed data in their regression analyses potentially obscures simultaneous gain and bias effects.

Head movement triggered suppression of perceived image motion relative to the HMD is potentially advantageous in dealing with artifacts attributable to VE system latency as well as those arising from predictive techniques employed to compensate for such latencies. Both latency and prediction artifacts introduce dynamic errors in image position relative to the real world that appear as slip or jitter in the HMD. Head movement, which is the source of these errors, can therefore also diminish observer sensitivity to such image motion. On the other hand, when the head stops moving the perceptual suppression will also end. However, because image slip based rendering errors will cease as the scene becomes stationary, the suppression would no longer be necessary anyhow.

ACKNOWLEDGEMENT

This work was supported by NASA's Space Human Factors Engineering Program. We thank Prof. Richard So for earlier comments pertinent to the work presented here.

REFERENCES

Adelstein, B.D., Lee, T.G., & Ellis, S.R. (2003). Head Tracking Latency in Virtual Environments: Psychophysics and a Model. *Proceedings, 47th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica CA, pp. 2083-2087.

Adelstein, B.D., Burns, E.M., Ellis, S.R., & Hill, M.I., (2005). Latency discrimination mechanisms in virtual environments: velocity and displacement error factors. *Proceedings, 49th Annual Meeting of the Human Factors and Ergonomics Society*, pp. 2221-2225.

Azuma, R., & Bishop, G. (1994). Improving static and dynamic registration in an optical see-through display. *Proceedings, SIGGRAPH*, Orlando, FL, pp. 197-204.

Barlow, H. B. (1990). A theory about the functional role and synaptic mechanism of visual aftereffects. In C. Blakemore, Ed., *Vision: Coding and Efficiency*. Cambridge UK: Cambridge University Press, pp. 363-375.

Bedell, H.E., & Patel, S.S. (2005). Attenuation of perceived motion smear during the vestibulo-ocular reflex. *Vision Research*, *45*, 2191-2200.

Durgin, F.H., Gigone, K., & Scott, R. (2005). Perception of visual speed while moving. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(2), 339-353.

Li, L., Adelstein, B.D., & Ellis, S.R. (2006). Perception of image motion during head movement. *Proceedings, Applied Perception in Graphics and Visualiztion (APGV06)*. Association for Computing Machinery, Boston MA, July 28-30

Probst, T., Brandt, T., & Degner D. (1986). Object-motion detection affected by concurrent self-motion perception: psychophysics of a new phenomenon. *Behavioural Brain Research*, *22*, 1-11.

Wallach, H. (1987). Perceiving a stable environment when one moves. *Annual Review of Psychology*, *38*, 1-27.