Psychophysical Evaluation of Persistence- and Frequency-Limited Displays for Virtual and Augmented Reality

T. Scott Murdison, Christopher McIntosh*, James Hillis*, Kevin J. MacKenzie**

***Facebook Reality Labs, Redmond, WA**

Abstract

Little is known about user sensitivity to flicker and eye movement-induced ghosting at refresh rates and persistence levels relevant for head-mounted displays (HMDs). In this report, we describe a pair of psychophysical experiments that we performed to comprehensively quantify these artifacts and make general recommendations for HMD design.

Author Keywords

eye movements; stroboscopic effect; phantom array; flicker; temporal light artifacts

1. Introduction

When we view a naturally illuminated scene, photons are continuously projected onto the retina. Conventional display technology, however, subsamples image information in both space (pixel rasterization) and time (frame buffers). In order to provide the impression of motion continuity, multiple discrete images are presented in succession at a given rate, and each frame persists for a given length of time (sometimes represented as a proportion of the frame time called the duty cycle). Thus, a display not designed with an understanding of how the visual system integrates information can lead to perceptual artifacts, including detectable flickering during fixation, strobing when the display moves relative to the eye, or ghosting (or 'phantom arrays') when the eye moves relative to the display[1–4]. These artifacts are governed by several display parameters, including: refresh rate, persistence, temporal waveform and between-frame light modulation depth[2].

As an example, we illustrate ghost artifacts that can arise during eye-movements in Fig. 1. Here, we illustrate three different potential light delivery conditions: (1) continuous light from a 'real' object (*top*), (2) discontinuous light from a 60 Hz, 200 us display (*middle*) and (3) discontinuous light from a 120 Hz, 500 us display (*bottom*). Ignoring the influence of modulation depth, Fig. 1 demonstrates how one can recover a percept nearly-identical to a percept aligning with natural expectations (*top*) by increasing the refresh rate and persistence levels (*bottom*).

Head-mounted displays (HMDs), such as those for virtual (VR) or augmented reality (AR), pose novel challenges to minimizing spatiotemporal artifacts because they are updated for head movements. This added constraint means that: (1) HMDs are frequency- and persistence-limited; meaning that updates to content must be made at high refresh rates with short frame times (i.e., persistence) to minimize visible blur during head movements, (2) for additive displays, visible real-world objects can potentially make the user more sensitive to[5,6], or alternatively might mask[7], ghosting[8–10], and (3) for blocked-light displays, high contrast levels may exist between pixels, also making the user more sensitive[6] to ghosting during eye movements than they would be in more conventional display-viewing contexts. As a result, artifacts may be unavoidable under certain HMD architectures. Therefore, having an understanding of how spatiotemporal light delivery contributes to the perception of, and sensitivity to, flicker and ghosting artifacts is an important step for the design of VR and AR HMDs.

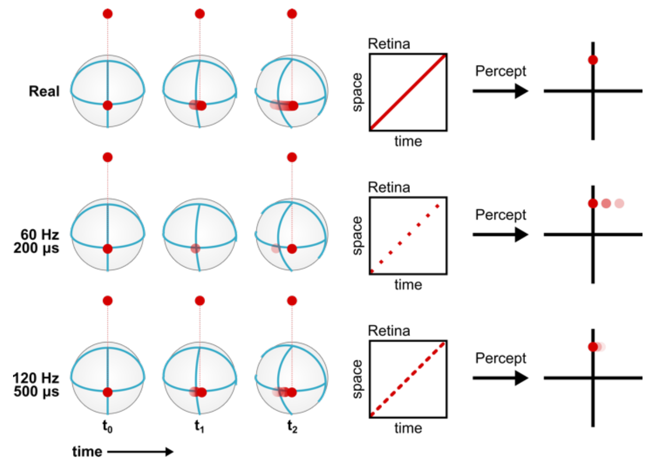

Figure 1. Illustration of ghosting on retina and perceptual artifact during eye movements for a continuously illuminated 'real' object (top) and a 'virtual' object displayed at two different refresh rates and persistence levels (middle and bottom).

Previous efforts to understand the perceptual limits of frequency- and persistence-limited displays have helped define basic thresholds for spatiotemporal artifact perceptibility but are limited for HMD development. For example, simplified models of human spatiotemporal contrast sensitivity such as the Window[11] and Pyramid of Visibility[12] are capable of predicting the perceptual outcomes for many typical displays, however, these models' predictive power are limited by their underlying assumptions. The predictions of these models are based on static detectability limits during fixation, i.e., without a dynamic observer. As such, it does not account for all possible components of visual sensitivity such as saccadic suppression[13], contrast masking[14], chromaticity[15], orientation[16] and changes in sensitivity with eccentricity[17]. Even if these models were extended to include these components, however, their utility would remain hampered by the current technical limits of computing power for high-spatiotemporal resolutions. For example, to simulate persistence levels in the micro- and nanosecond ranges, sufficiently capturing the temporal frequency components at full spatial resolution becomes technically infeasible due to computational memory demands.

Findings from empirical psychophysical investigations on conventional (non-HMD) displays are not necessarily generalizable to HMDs. Because conventional displays are not actively updated during head movements, the temporal parameter requirements are more clearly defined for minimizing temporal light artifacts[4]. Furthermore, relaxing these requirements (e.g., allowing for longer persistence) has allowed display engineers to implement strategies to minimize the appearance of visible flicker and ghosting artifacts. For example, they can do so by adding a constant backlight to minimize the between-frame light modulation while rasterizing frame updates in a global fashion (e.g., VSYNC[18]). Additionally, studies of temporal artifacts for conventional displays or lighting often investigate refresh rate while ignoring persistence[3,19–21] though the appearance of these artifacts depend on both components. Finally, in any study with a visible background, there exists a potential confound in the masking of ghosting contrast energy by the surrounding visual scene[7,14]. Such a scenario is especially relevant for additive AR displays. Thus, a more complete analysis of spatiotemporal artifact sensitivity as a function of basic temporal display parameters under controlled contrast conditions is necessary to inform HMD development.

We designed a pair of psychophysical experiments to comprehensively assess spatiotemporal artifact detectability and sensitivity across a range of refresh rates (up to 960 Hz) and persistence levels (as low as 400 nanoseconds). We carried out this experiment under sensitive viewing conditions (e.g., 7:1 contrast, wide display field of view, high spatial frequency content). This allowed us to derive spatiotemporal artifact detectability and sensitivity thresholds. We use this data to make general recommendations for the design of any display to minimize flicker and eye movement-induced ghosting artifacts.

2. Methods

We asked participants to take part in two two-interval, forced-choice (2IFC) tasks in which they were specifically asked to discriminate between two stimuli. In the following sections we describe the apparatus and procedure we used to do so.

2.1. Apparatus

We built a 50° x 30° field of view array of 3 x 3 (9 total) white LED lightboxes (3500 K color temperature), each approximately 2° x 0.7° across (width x height) and positioned at +/- 25° horizontally and +/- 15° vertically (in a grid) across a high-spatial frequency fabric background, as shown in Fig. 2. LEDs used were Citizen CLU048 series. Each row of LEDs was illuminated sequentially and the timing was controlled over a universal asynchronous receiver-transmitter (UART) controller via a custom Matlab API (The Mathworks, Inc., Natick, MA). The controller would take experimenter input for refresh rates and persistence levels and calculate the timing parameters needed for each row. The controller then updated each integrated PWM driver. The row timing is shown in Fig. 2 (*bottom*) for a single frame. Using this design, we simulated a rolling-style display akin to conventional CRT displays.

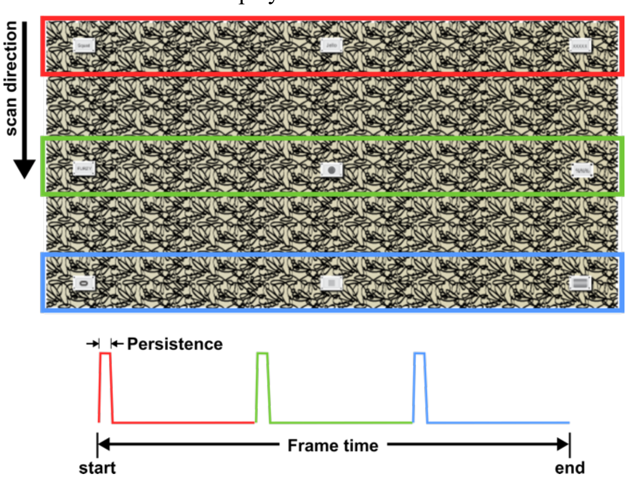

Figure 2. Participant view of the physical apparatus (top) and row-wise temporal light delivery profile (bottom). Persistence and frame time (1/refresh rate) are demarcated, and color represents row illuminations.

We produced refresh rates between 60 Hz and 1 kHz and persistence levels as low as 400 ns at minimum (and 13 ms at maximum). Although the electronics could drive the persistence timing much lower, the LEDs had a phosphorus coating, resulting in a trailing light discharge rate. To account for this, we limited our investigations to a minimum persistence value of 400 ns. We confirmed that we were indeed delivering photons at these speeds using a high-speed photodiode attached through an oscilloscope at every desired combination of refresh rate and persistence. The photodiode response allowed direct measurements for the rise, fall, and on time durations.

To match output luminance at different duty cycles, we changed the LED current by adjusting the power supply voltage (ranging between 48V and 58V) between trial intervals. We created a voltage lookup table using a Konica Minolta CS-160 luminance meter (Konica Minolta Sensing Americas, Inc., Ramsey, NJ) to modulate the current to the LEDs, ensuring identical luminance output at different duty cycles.

2.2. Procedure

We recruited a total of 68 participants (28 female, 40 male, aged 18-40 years) for two psychophysical experiments. Each had normal or corrected-to-normal visual acuity and reported no known neurological disorders. For most-sensitive testing conditions, participants were briefed on the flicker and ghosting artifacts they might observe and given the chance to practice inducing the artifacts prior to commencing the study.

In the first experiment, we asked participants (N = 50, 18 female, 32 male, aged 18-40 years) to judge the relative subjective quality of two stimuli on each trial in a 2IFC design. Each stimulus had a luminance of 300 cd/m^2 at a 7:1 contrast ratio (corresponding to 42 cd/m^2 background luminance). In one interval, we presented an ‘artifact-free’ standard reference stimulus at 1 kHz and 90% duty cycle (900 us persistence). The test stimulus was presented at one of 30 different combinations of refresh rate (60 Hz, 90 Hz, 120 Hz, 240 Hz or 480 Hz) and duty cycle (0.25%, 10%, 20%, 40%, 60% and 80%), which represents the percentage of frame time for which the LED was illuminated. This design resulted in persistence values that ranged from 5 microseconds up to 13 milliseconds, depending on the refresh rate. The order of the stimuli was randomized, and participants performed 20 repetitions of each 2-interval trial, for a total of 600 trials, split across 3 sessions. The order of trials was randomized from person to person to allow population-level statistical analyses of the data.

To fully characterize the sensitivity of viewers to artifacts resulting from persistence changes down to the hundreds of nanoseconds, we ran a second experiment with a separate group of participants (N = 18, 10 female, 8 male, aged 19-39 years) using a 2IFC QUEST adaptive staircase procedure[22]. In this experiment, we again asked participants to discriminate between two stimuli, this time at the same refresh rate but different persistence values. Also, because of the high voltage levels required to achieve photopic conditions with ultra-low persistence levels, we instead used a foreground luminance of 50 cd/m^2 with a background luminance of 7 cd/m^2 (achieving 7:1 contrast). Participants began staircases at any combination of refresh rates at 90 Hz, 120 Hz, 240 Hz, 480 Hz and 960 Hz and persistence levels at 400 ns, 1 us, 5 us, and 150 us, giving a total of 20 possible staircase conditions (each performed twice, 15 trials per staircase). The staircase algorithm converged to the persistence level at which the participants could discriminate between the two stimuli.

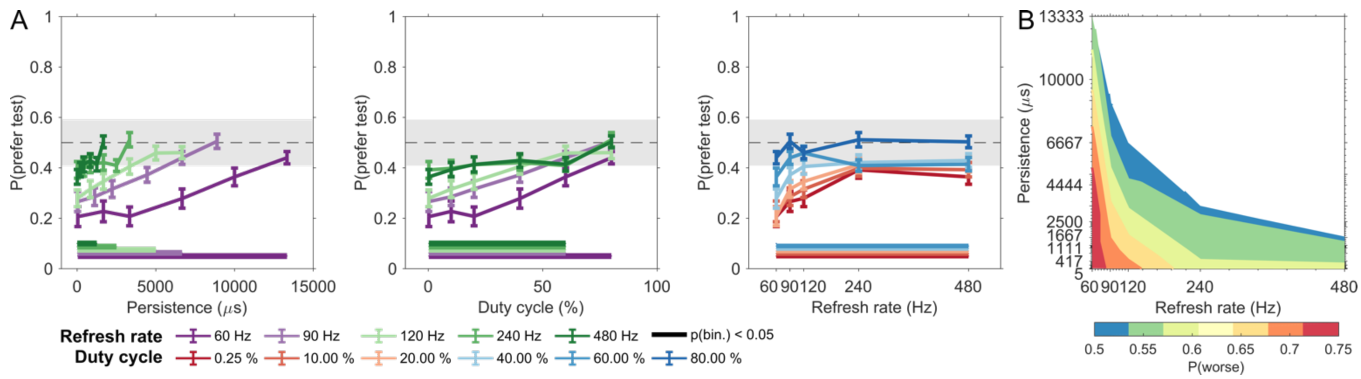

Figure 3. Experiment one results. **A** Spatiotemporal artifact detectability varies with refresh rate and duty cycle. Color-coded horizontal bars near the bottom of each panel represent statistically significant detections, per a one-sided binomial test (p<0.05). Error bars represent standard error. **B** Functional psychometric surface across refresh rate and persistence. At refresh rates <90 Hz, refresh rate has the dominant influence on artifact detectability. At refresh rates >90 Hz, persistence is the dominant factor.

3. Results

We conducted two separate experiments to determine (1) the detectability limits of spatiotemporal flicker and ghosting artifacts and (2) the sensitivity of observers to changes in persistence at ultra-low levels relevant for AR and VR HMDs. We describe our findings in the following two sections.

3.1. Experiment one

To understand the effects of refresh rate and persistence on the detectability of spatiotemporal eye movement artifacts we asked participants to repeatedly discriminate between a test and reference stimulus. We present these probabilities (computed across participants) in Fig. 3A, as a function of persistence and corresponding duty cycle (collapsed across refresh rate) in the left and middle panels, and as a function of frequency (collapsed across duty cycle) in the rightmost panel. Also depicted in these plots are significance bars, representing significant deviation from chance probability based on a one-sided binomial test for significance (color-matched to refresh rate/duty cycle as shown in the legends, p<0.05). All significant deviations from chance corresponded to a preference for the standard stimulus over the degraded test stimulus, or significant artifact detectability.

As expected, artifact detection was reduced with both increases in refresh rate and persistence. Participants' ability to discriminate between the artifact-free and test stimuli disappeared systematically, with a persistence threshold corresponding to 60% duty cycle across all refresh rates except 60 Hz. This effect at 60 Hz was likely due to visible flicker, which was consistently verbally reported by participants. We found significant artifact detectability regardless of frequency for all persistence levels except those corresponding to duty cycles of 80%. In isolation, refresh rate appears to modulate the overall appearance of artifacts, but only when absolute persistence falls below a threshold corresponding to 80% duty cycle – regardless of refresh rate.

One of the main goals of this study was to understand the combined contribution of refresh rate and persistence to the visibility of spatiotemporal artifacts. We present this detectability as a function of both refresh rate and persistence in Fig. 3B. This surface describes anisotropic effects of persistence and refresh rate. At low refresh rates (<90 Hz), the gradient of the probability surface varies primarily along the refresh rate axis, indicating that refresh rate itself drives artifact sensitivity at low refresh rates. At higher refresh rates (>90 Hz), the gradient changes direction, instead lying primarily along the persistence axis, indicating that persistence drives artifact sensitivity at high refresh rates.

3.2. Experiment two

To fully investigate the relationship between spatiotemporal artifact detectability and temporal display parameters, we carried out a second experiment, at the engineering limits of our setup again under 7:1 contrast conditions, down to 400 ns persistence levels. This experiment provided us with estimates of the difference in persistence required for a noticeable change in artifact strength, represented in Fig. 4. These difference thresholds reveal an absolute persistence level between 694 us and 794 us (across reference persistence levels), or on average 721 us (+/- 33 us SD), at which spatiotemporal artifacts become noticeably worse, regardless of reference persistence level. For any persistence values lower than this threshold, including those in the nanosecond range, we found that participants could not distinguish between stimuli.

4. Discussion

We designed two experiments to determine how sensitive people are to spatiotemporal flicker and ghosting artifacts across a range of refresh rates and persistence levels, with direct applicability for AR and VR HMD design. Under sensitive viewing circumstances, we asked participants to perform psychophysical tasks in which they discriminated between different test and reference stimuli.

In the first experiment, we measured the probability of artifact detection across refresh rate and persistence levels. At refresh rates

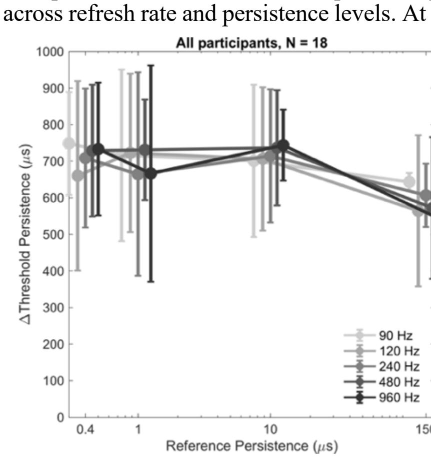

Figure 4. Threshold persistence difference (+/- SD) as a function of reference persistence, grouped by refresh rate (gray levels), offset for visibility.

below 90 Hz, participants were most sensitive to changes in refresh rate. Verbal participant reports indicated that artifacts were most likely comprised of flicker in this range. At refresh rates above 90 Hz, participants were most sensitive to changes in persistence. Verbal participant reports indicated that artifacts were most likely comprised of ghosting in this range. We also carried out a second experiment in which we determined how sensitive participants were to changes in ultra-low persistence levels potentially relevant for high speed display development. We found that users were insensitive to differences in artifact magnitude for persistence levels below 721 us, on average, regardless of refresh rate.

These results are limited in at least two ways that could be important for HMD development. First, our study ignores the influence of concurrent geometrical distortions due to rasterization. When summed with the velocity of the eyeball, this pattern's velocity can cause visible distortions of the displayed image. Our extremely sparse display was likely not sensitive to these kinds of distortions. The second effect likely to be present during HMD use that our study did not explicitly account for is blur. Blur can occur for a variety of reasons in an HMD, ranging from delays in frame updates during head movements (i.e., judder) to blurring of real content relative to virtual content during eye movements with additive AR displays. In each of these situations, it is unclear whether blur would act as a mitigating factor for spatiotemporal artifact detection via contrast masking[8], or if blur would make such artifacts more apparent by serving as a visible reference. Further investigations into geometrical distortions, head movement-induced blur and display fill factors are necessary for a more complete understanding of spatiotemporal artifact prevalence and prevention in AR and VR HMDs.

5. Impacts

- Using two experiments, we provide a comprehensive assessment of spatiotemporal artifact detectability across an HMD-relevant range of refresh rates and persistence.
- Based on the outcomes of these experiments, we recommend designing conventional non-head-mounted displays with a refresh rate above greater than 90 Hz and persistence corresponding to a duty cycle of at least 60% (7 ms at 90 Hz) to minimize flicker and ghosting artifacts.
- Similarly, HMDs should have a refresh rate of at least 90 Hz to avoid visible flicker, but persistence will depend on the exact implementation of the display (e.g., AR vs. VR, see-through vs. light-blocked, etc.).

6. Acknowledgments

The authors would like to acknowledge the many research assistants who collected over 40,000 trials for this study, especially Mark Opaliski, Annelise Smith, Brian Jackson and Emily Higgins who led these efforts. We would also like to thank Keith Hoffman, who created the mechanical design for our LED lightboxes.

7. References

[1] Hershberger WA, Jordan JS. The phantom array: A perisaccadic illusion of visual direction. Psychol Rec 1998;48(1):21–32.

[2] Lehman B, Wilkins AJ. Designing to Mitigate the Effects of Flicker in LED lighting. IEEE Power Electron Mag 2014;(September):18–26.

[3] Thomson WD, Saunders J. The Perception of Flicker on Raster-Scanned Displays. Hum Factors J Hum Factors Ergon Soc 1997;39(1):48–66.

[4] International Commisson On Illumination. Visual Aspects of Time-Modulated Lighting Systems – Definitions and Measurement Models - CIE TN 006:2016. 2016.

[5] Geisler WS. Physical limits of acuity and hyperacuity. J Opt Soc Am A 1984;1(7):775–82.

[6] Westheimer G, Brincat S, Wehrhahn C. Contrast dependency of foveal spatial functions: orientation, vernier, separation, blur and displacement discrimination and the tilt and Poggendorff illusions. Vision Res 1999;39(9):1631–9.

[7] Tolhurst DJ, Tadmor Y. Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra. Vision Res 1997;37(23):3203–15.

[8] Eckstein MP, Ahumada AJ, Watson AB. Visual signal detection in structured backgrounds II Effects of contrast gain control, background variations, and white noise. J Opt Soc Am A 1997;14(9):2406.

[9] Westheimer G, McKee SP. Visual acuity in the presence of retinal-image motion*. J Opt Soc Am 1975;65(7):847.

[10] Demer JL, Amjadi F. Dynamic visual acuity of normal subjects during vertical optotype and head motion. Investig Ophthalmol Vis Sci 1993;34(6):1894–906.

[11] Watson AB, Ahumada AJ, Farrell JE. Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. J Opt Soc Am 1986;3(3):300.

[12] Watson AB, Ahumada AJ. The pyramid of visibility. Electr Imaging 2016;2016(July):1–6.

[13] Burr DC, Morrone MC, Ross J. Selective suppression of the magnocellular visual pathway during saccadic eye movements. Nature 1994;371(6497):511–3.

[14] Legge GE, Foley JM. Contrast masking in human vision. J Opt Soc Am 1980;70(12):1458–71.

[15] Kelly DH. Spatio-temporal frequency characteristics of color-vision mechanisms. J Opt Soc Am 1974;64(7):983–90.

[16] Adelson EH, Bergen JR. Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 1985;2(2):284.

[17] Koenderink JJ, Bouman MA, Bueno de Mesquita E, Slappendel S. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. I. The near peripheral visual field (eccentricity 0°-8°). J Opt Soc Am 1978;68(6):845–9.

[18] Zhang H, Yeo C, Ramchandran K. VSYNC: A novel video file synchronization protocol. ACM Multimed 2008 Short Pap Netw Track 2008;(3):757–60.

[19] Perz M, Vogels IMLC, Sekulovski D, Wang L, Tu Y, Heynderickx IEJ. Modeling the visibility of the stroboscopic effect occurring in temporally modulated light systems. Light Res Technol 2015;47(3):281–300.

[20] Perz M, Sekulovski D, Vogels I, Heynderickx I. Stroboscopic effect : contrast threshold function and dependence on illumination level. J Opt Soc Am A 2018;35(2):309–19.

[21] Boschman MC, Roufs JAJ. Reading and screen flicker. Nature 1994;372(6502):137.

[22] Watson AB, Pelli DG. Quest: A Bayesian adaptive psychometric method. Percept Psychophys 1983;33(2):113–20.