Direct Measurement of the System Latency of Gaze-Contingent Displays

Daniel R. Saunders and Russell L. Woods

Schepens Eye Research Institute, Massachusetts Eye and Ear

Author Note

Daniel R. Saunders and Russell L. Woods, Schepens Eye Research Institute, Massachusetts Eye and Ear.

This research was supported by National Eye Institute grant R01EY019100 awarded to Russell L. Woods.

Correspondence concerning this article should be addressed to Daniel R. Saunders, Schepens Eye Research Institute, 20 Staniford Street, Boston, MA 02114. Email: daniel_saunders@meei.harvard.edu

Abstract

Gaze-contingent displays combine a display device with an eyetracking system to rapidly update an image based on the measured eye position. All such systems have a delay, the system latency, between a change in gaze location and the related change in the display. The system latency is the result of delays contributed by the eyetracker, the display computer, and the display, and it is affected by the properties of each component, which may include variability. We present a direct, simple and low-cost method to measure the system latency. The technique uses a device to briefly blind the eyetracker system (e.g. for video-based eyetrackers, a device with infrared LEDs), creating an eyetracker event that triggers a change to the display monitor. The time between these two events, as captured by a relatively low-cost consumer camera with high-speed video capability (1,000 Hz), is an accurate measurement of the system latency. With multiple measurements, the distribution of system latencies can be characterized. The same approach can be used to synchronize the eye position time series and a video recording of the visual stimuli that would be displayed in a particular gaze-contingent experiment. We present system latency assessments for several popular types of display, and discuss what values are acceptable for different applications, as well as how system latency might be improved.

Direct Measurement of the System Latency of Gaze-Contingent Displays

In gaze-contingent visual displays, a computer monitor is rapidly and continuously updated based on information about where an observer is looking. In one use of gaze-contingent displays, the area of the display that corresponds to the fovea is enhanced, degraded, or masked completely. In another, areas of the display that correspond to peripheral vision are modified or masked. Since its introduction in 1973 (Reder), this technique has led to major insights into the respective roles of central and peripheral vision in reading (Rayner, 1998), visual search (Geisler, Perry, & Najemnik, 2006; Loschky & McConkie, 2002), and scene perception (Henderson, McClure, Pierce, & Schrock, 1997). It has also been used in proposals to reduce the bandwidth required for video by encoding parts of the image outside fixation with a lower resolution (Duchowski, Cournia, & Murphy, 2004; Reingold, Loschky, McConkie, & Stampe, 2003). In research on vision disorders, gaze-contingent displays can simulate central or peripheral defects in the visual field, allowing normally-sighted subjects to experience vision similar to that of patients with eye diseases such as macular degeneration or glaucoma (Fine & Rubin, 1999). These simulations make it possible to more efficiently test hypotheses about adaptation to these disorders, as well as strategies for rehabilitation, since it means that patients with the target condition are not required in the early stages of an investigation.

However, the effectiveness of a gaze-contingent display depends on characteristics of the system, in particular its temporal characteristics. In the “simulated scotoma” paradigm, in which loss of central vision is simulated by masking several degrees around the point of fixation, the position of the masking patch will lag behind the eye position by some milliseconds, depending on characteristics of the eyetracker, software, and display device. Figure 1 illustrates a real example of this lag on a 60 Hz iMac monitor during a saccade (creation of this figure is discussed in the Method section, under the heading “Synchronizing eye position data and experimental stimuli recordings”). Sensitivity to visual information is greatly reduced during a saccade (Burr, Morrone, & Ross, 1994; Diamond, Ross, & Morrone, 2000). However at the end of a saccade, visual information begins to have an effect on perception within about 6 ms (McConkie & Loschky, 2002). If the simulated scotoma position is not updated on the screen within that time, then the observer will have a high-resolution glimpse of the new fixation location with central vision before it is masked. We will refer to the delay between an eye movement event and the resulting change to the image on the monitor as the system latency. In the instance depicted in Figure 1, if we define the end of the saccade by the peak of the overshoot (as recommended by Loschky & Wolverton, 2007), and define the arrival of the simulated scotoma by its alignment with the fixation location (rather than merely overlapping it), then the measured system latency was 22 ms. However, there can be considerable variance in system latency from one saccade to another, and so to characterize a particular system it is necessary to take multiple samples to estimate the distribution of system latencies.

How large does system latency need to be before this unmasked glimpse starts to have an effect on perception? Given that many gaze-contingent experiments involve viewing of natural scenes, there are a number of relevant results from the literature on scene gist perception. The typical approach is to present photographs of scenes with varying delays before a full-screen mask appears. While a strong conceptual scene gist can be activated in 40-100 ms (e.g. Castelhano & Henderson, 2008), other types of high-level information about a scene are available at even shorter durations. The presence of animate objects can be detected in 27 ms (Fei-Fei, Iyer, Koch, & Perona, 2007); scenes can be judged as natural or unnatural in 19 ms at the 75% correct level (Greene & Oliva, 2009); photographs of animals can be distinguished from non-animal photographs at above-chance levels in 12 ms (Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005); and 12 ms was sufficient to determine better than chance whether photographs had been correctly matched with one of 10 scene categories (Loschky, et al., 2007). Given these results with high level judgments, more schematic visual stimuli, including basic visual features, might be expected to impact perception with even briefer pre-mask presentations. A study on perception of grating orientation orientation found above-chance performance at 8.4 ms (Bodelón, Fallah, & Reynolds, 2007), while global blur of a natural image was detected at significantly above chance levels in only 6 ms (Loschky & McConkie, 2002). Of direct relevance to simulated scotoma experiments are results from paradigms where a mask appears at the location of fixation at a fixed time delay after the end of a saccade. However many of these studies do not examine exposures below 50 ms (Glaholt, Rayner, & Reingold, 2012), or use only changes in gaze patterns (such as median fixation duration) as dependent measures, rather than acquired information (van Diepen, De Graef, & d'Ydewalle, 1995; van Diepen, Ruelens, & d'Ydewalle, 1999). Critically, the majority of them either do not report measurements of the end-to-end latency of their gaze contingent system, or do not describe the method used to obtain them.

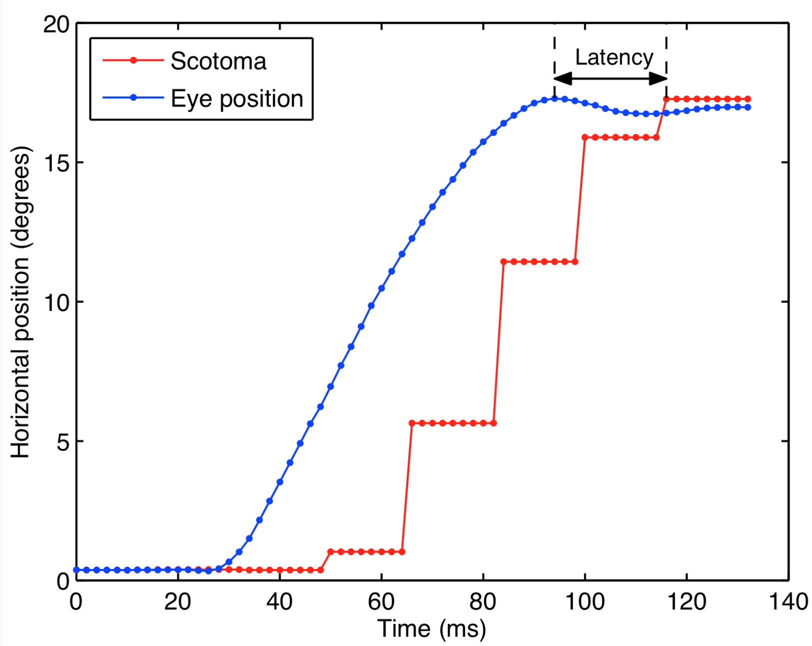

Figure 1. Position of the center of the simulated scotoma relative to the gaze position during one left-to-right saccade. Markers are placed at 2 ms intervals. Recorded on an iMac Intel Core i7 (60 Hz display) connected to an Eyelink 1000, and demonstrating a 22 ms measured system latency.

Accurate estimation of end-to-end system latency is difficult because a number of additive factors may contribute to the delay between measurement of the eye position and display response. At the eyetracker, factors include the sampling rate, the time to compute the gaze position from the camera image, the time for additional processing and filtering, and the time to transmit the information to the display computer. At the display computer, factors include the time to render the updated image to send to the monitor, and possible delays introduced by other aspects of the experiment code, or the operating system. At the monitor, one factor is the time to wait for the next refresh cycle. In cathode ray tube (CRT), liquid-crystal display (LCD) and organic light-emitting diode (OLED) displays, the image in memory is drawn from top to bottom in horizontal scan lines, and therefore there is also the time for the vertical refresh process to reach a particular location on the screen. In digital light processing (DLP) displays, the entire image appears (effectively) simultaneously, but in alternating primary colors; see supplemental videos. Once a location has been signaled to update, pixels of LCD displays take some time to reach their target brightness, which is called the pixel response time, or rise time. Finally, for digital (LCD, DLP and OLED) displays there is the input lag (or display lag, or response lag), which is time added for preprocessing of the input signal. In

LCD displays, this almost always includes buffering at least one frame of the input video to improve pixel response time. Input lag also includes latency contributed by other types of digital preprocessing at the monitor, such as noise reduction or rescaling the image to match the screen resolution.

Estimating the delay added by each component of a gaze-contingent system and then summing them produces an estimate of the system latency, and this is the approach that is taken in the majority of gaze-contingent papers that report system latency (Henderson, et al., 1997; Loschky & McConkie, 2000; McConkie, 1981; Santini, Redner, Iovin, & Rucci, 2007). However, manufacturer specifications for equipment can contain errors, as well omissions of key information, often including the input lag. Some of the earliest published gaze-contingent experiments were contaminated by an undocumented 25 ms delay in the eyetracker circuitry, so that the findings had to be reconfirmed in subsequent work (as recounted in McConkie, 1997, p. 312). Although measuring a monitor's response properties with a photodiode and custom software (Dorr & Bex, 2011; Elze & Tanner, 2012; Wiens, et al., 2004) can improve the estimate of the latency added by the monitor, there are still interactions between monitor and eyetracker timing, as well as interactions involving the software and operating system, that are hard to anticipate. Therefore a more reliable approach is to measure the end-to-end system latency directly.

One approach to direct, end-to-end measurement of latency uses an oscilloscope that receives realtime input from both the saccade detection circuit of an eyetracker, and a photodiode attached to the monitor (Allison, Schumacher, Sadr, & Herpers, 2010; Triesch, Sullivan, Hayhoe, & Ballard, 2002). The distance between the spikes in the oscilloscope produced by a saccade and the consequent gaze-contingent change to the monitor, as captured by a video camera, can be an estimate of the system latency. In another approach, artificial saccades were created using LEDs, and the monitor output was measured with a photodiode, with a centralized controller comparing the timing of the two events (Bernard, Scherlen, & Castet, 2007; Bockisch & Miller, 1999). However, these techniques require a certain degree of electronics knowledge, and do not give an overall sense of the gaze-contingent temporal behaviour of the stimuli at the millisecond level.

The benefit of our novel approach to measuring system latency is that it is direct and intuitive, and economical enough that it can be used routinely by researchers to characterize their gaze-contingent apparatus regardless of the scale of their involvement in eyetracking. The technique uses infrared LEDs to briefly blind the eyetracker camera, which is an eyetracker event that can be used to trigger a change to the display monitor. The time between these two events, as captured by a relatively low-cost consumer camera with high-speed capability (1000 Hz video), is an accurate measurement of the system latency. The same approach can be used to synchronize the eyetracking time series and a video recording of the monitor displaying actual experimental stimuli, such as a video superimposed with a simulated scotoma, to produce graphs such as Figure 1. We present the method, and then discuss the results of measuring the system latency of seven display configurations, representing classes of display that experimenters might choose.

Method

Equipment

This method requires a high speed camera, which has some sensitivity to IR light (as most digital cameras do). We used a Casio Exilim EX-ZR100, which has a full retail price of under $300, and advertises the ability to film at 1,000 frames per second. This sample rate was verified using the fixed timing of monitor refreshes: a 60 Hz monitor was filmed while switching between black and white with each frame (see supplemental video), and the span of 120 screen refreshes in the footage (measured from the first appearance of white at the top of the monitor) was exactly 2000 frames, showing that it was sampling once per millisecond (120 frames / 60 Hz = 2000 ms) on average. The other piece of equipment that is needed is an infrared (IR) illuminator, of the type that is available inexpensively for use with home surveillance cameras, or for mounting on video cameras that have night vision capabilities. The measurements reported here were made using a simple custom circuit connecting IR LEDs to a switch, but we also tested a Polaroid IR Light Bar (PLLED36) and found it to be effective for this purpose. The eyetracker was an Eyelink 1000 (SR Research) capturing at a rate of 1,000 samples per second, with eye position estimated using the Centroid pupil tracking algorithm and the corneal reflection (i.e. not “pupil-only” mode). One level of heuristic filtering was applied (specifically, the Link/Analog Filter setting was “STD”), and the viewer’s head position was fixed using a brow and chinrest. The software was written in MATLAB with the Psychophysics Toolbox (Brainard, 1997), which incorporates the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002). Except in the case of the iMac (Intel Core i7, 2.93GHz), the display software ran on a Mac Pro (Quad-Core Intel Xeon 2.8GHz). The Mac Pro was connected either directly via a DVI connection, or, in the case of the CRT, via a Mini DisplayPort to VGA converter.

Measurement of system latency

In our method, to measure the latency of the system, the experimenter seats a subject in front of the eyetracker and begins a regular eyetracking session, including calibration and verification controlled by a MATLAB

Psychophysics toolbox program. Subsequently, the program monitors the eyetracker input and changes the screen from black to white when tracking is lost. Then the experimenter holds the IR illuminator close to the eyetracking camera lens (Figure 2). The illuminator is turned on, and the glare from the infrared LEDs causes an instantaneous loss of tracking (the LEDs we used had a rise time of 0.8 ms). The highspeed camera is positioned so that it captures both the face of the illuminator and the screen. The entire screen need not be captured, but at least a part of the top must be within the frame for the most accurate measurements. In this configuration, the camera records both events with high temporal precision: the moment of illumination of the LEDs, and the moment when the monitor begins to change from black to white. The LEDs are turned off and then on again for each sample that is desired. In the resulting 1,000 Hz video, the number of frames between the LED illumination and the first appearance of white at

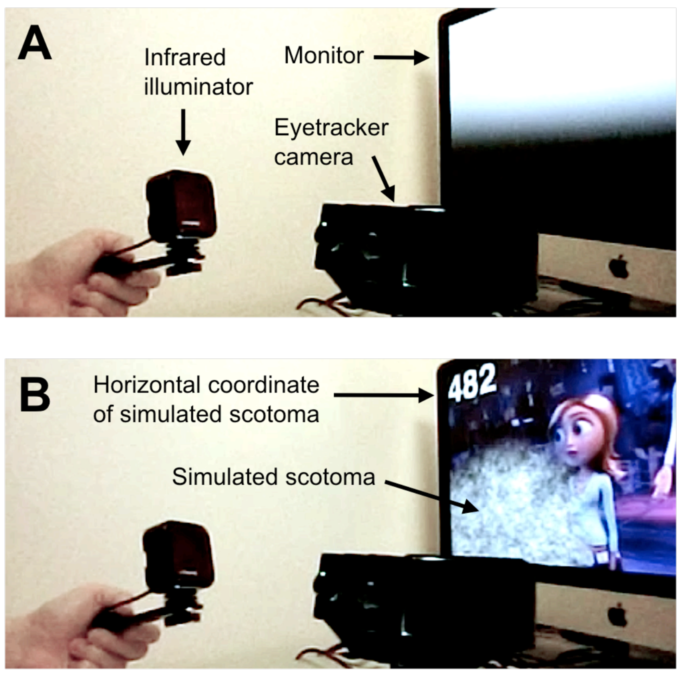

Figure 2. Illustration of the apparatus. The researcher holds an infrared illuminator positioned such that when turned on, the infrared LEDs disrupt the tracking camera, and are visible to the high speed video camera. A) shows the technique for obtaining a measurement of system latency: the monitor changes from black to white when triggered by loss of tracking (monitor is depicted partway through drawing the first white frame) B) shows the technique for synchronizing the time course of the eyetracking data with the displayed simulated scotoma position in an actual gaze-contingent experiment. This technique was used to produce Figure 1.

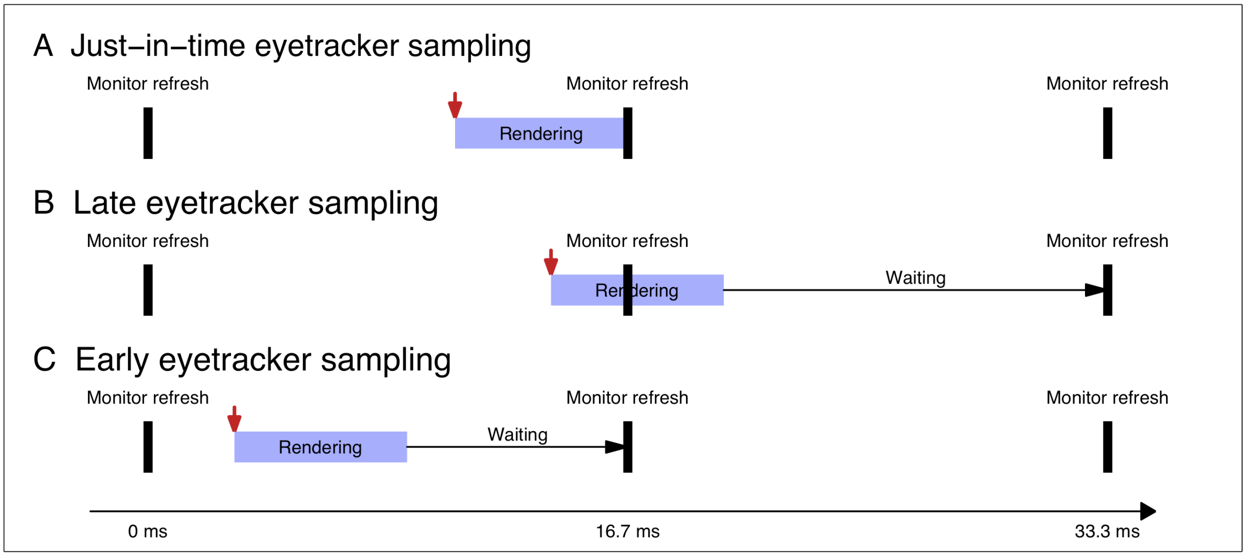

Figure 3. Illustration of the variability that may be contributed by asynchrony with monitor refreshes. Eyetracker samples, taken at the time indicated by the red downwards arrow, will have different delays to being reflected on the screen depending on whether they are taken early or late in the refresh cycle. This example assumes a 60 Hz refresh rate and a 4 ms rendering time (blue bar). The shortest latency will be when the eyetracker is sampled in time for the image to finish rendering just before the refresh deadline (A), and that is 4 ms in this example, while the longest latency will be when this deadline is just missed (B), approximately 20 ms. In other cases (C) the latency will be somewhere in between.

the top of the screen is one latency measurement. This count can be determined easily and unambiguously using a program such as QuickTime Player 7 or FinalCut Pro that allows frame-by-frame inspection of movie files.

The mean of the values obtained this way is an estimate of the average time between any eyetracking event, such as the beginning or the end of a saccade, and the beginning of the refresh reflecting that event. In a real gaze-contingent system, the latency for eye movement events of interest will not be a constant, as demonstrated in Figure 3. A large proportion of this variance is due to the monitor's refresh cycle, especially if the eyetracking sample rate is much higher than the refresh rate. The timing of eye movement events is unlikely to be phase-locked with the refresh cycle. For example, if the end of a saccade occurs just before the new update is about to begin (including enough time to render the new frame), then the latency in that instance will be less than if the end of a saccade occurs well before the update. The range of latencies contributed is a function of the monitor's current refresh rate, and should be uniformly distributed between 0 and $T_{refresh}$, the time between screen refreshes.

The estimates discussed so far represent the latency for a location at the top of the display area, so for objects that are not at the top of the display, additional time needs to be added depending on the vertical distance. The maximum time added will be $T_{refresh}$ for objects at the bottom of the screen. Therefore for a particular apparatus using a 60 Hz monitor, where $T_{refresh}$ = 16.7 ms, the time of drawing initiation and the position on the screen could contribute a combined latency between 0 and 33 ms. With a 120 Hz monitor this range decreases to 0–17 ms, and with a 200 Hz monitor, 0–10 ms.

Synchronizing eye position data and experimental stimuli recordings

Besides helping to obtain individual estimates of the latency, an IR illuminator combined with a highspeed camera can also be used to create a diagram such as Figure 1, showing the relative position of the gaze location and simulated scotoma during each millisecond of a particular saccade. The technique allows the retrospective alignment of two streams of data, one being the measured eye position, and one the position of the scotoma (or other gaze-contingent image modification) on the screen. To closely model the system's behavior during the gaze-contingent experiment, the stimuli from the experiment should be used, with the addition of large numerals near the top of the screen corresponding to the horizontal eye position that was used to draw the overlayed simulated scotoma (Figure 3). The vertical eye position can be displayed as well, if tracking a more complex scanpath than a single horizontal saccade is desired. The time series of the position of the scotoma, as captured by the video camera, is aligned with the eye position time series by activating the IR illuminator at the beginning or end of the recording, temporarily disrupting tracking. Then the frame of the video recording in which the LEDs are seen to turn on represents the same point in time as the first missing sample of eyetracking data, providing synchronization to within 1-2 ms (the variance is due to the eyetracker processing time and the LED rise time). Thereafter the two time series can be compared.

Results

Using the equipment described in the Method section and the first technique, we measured the system latency 50 times for a selection of displays and obtained summary statistics (Table 1). For each system, the time to render the frames was well below the time between two screen refreshes minus the time to retrieve an eye position sample, ensuring that the screen refresh deadline was not missed (which would have increased the latency by the time of one refresh cycle).

We observed that a higher frame rate does not necessarily result in lower latency. The average system latency at the top of the iMac monitor running at 60 Hz was significantly lower than the average system latency at the top of the Acer monitor running at 120 Hz, according to a Welch's *t* test, $t(91.9) = 3.05$, $p < .005$. However, this advantage applied only to the top of the monitor: the time for the refresh process to update items at the midpoint was slightly faster on the Acer monitor (34.6 ms vs 35.9 ms), and notably faster for items near the bottom of the monitor (38.7 ms vs 44.3 ms). A different 120 Hz LCD, the Samsung, had a much lower average system latency at all screen positions. The differences among the displays that had the same refresh rate must be primarily due to differences in input lag, since the software and eyetracker were constant. Therefore, as expected, the best latency (12 ms) was observed using the fast CRT, which has no input lag. In fact our initial estimate overstated the latency, since with the Mac Pro the video signal had to be converted from Mini DisplayPort to VGA before reaching the monitor. Therefore we tested the latency again, using a Windows XP computer which had built-in VGA output. We inferred from the lower latency values that video conversion added approximately 3 ms to the average latency. The DLP display we tested had comparable latency to the best LCD monitor, showing evidence of requiring one frame to be buffered, but no more than one frame.

Table 1

System Latency at the Top of the Monitor

		System latency (ms)				
Monitor	Refresh rate (Hz)	Mean	Median	SD	Min	Max
iMac Intel Core i7 (built-in screen) (LED)	60	27.6	28	5.2	19	37
Acer GD235HZ (LED)	120	30.4	30	4.0	18	37
Samsung 2233RZ (LED)	120	17.5	18	2.0	14	22
Sony PVM-2551MD (OLED)	60	38.8	40	4.4	28	48
InFocus DepthQ (DLP)	120	18.3	18	2.6	14	23
NEC MultiSync FP2141SB (CRT)	120	14.9	15	2.5	10	19
NEC MultiSync FP2141SB (CRT) without video conversion	120	12.0	12	2.4	8	16

Discussion

We have described a method for obtaining measurements of end-to-end system latency in gaze-contingent systems. We showed that there are large discrepancies between the latencies contributed by different displays. The choice of monitor could mean the difference between an average latency at the center of the screen of 16 ms (the CRT) and 35 ms (the iMac).

A limitation of system latency measurements obtained this way is that they only partially account for pixel response time, that is, the time for the screen elements in an LCD monitor to reach the target brightness level after receiving the signal to change. The monitor onset times that are recorded are based on the first change to monitor brightness that can be unambiguously detected in the high speed video camera footage, rather than one of the standard criteria used to assess pixel response time, such as the time to transition between 10% and 90% of maximum luminance. If the precise level of luminance is critical for a gaze-contingent application, then our LED method will underestimate the system latency. However, since the pixel response times of modern displays are usually below 10 ms (Elze & Tanner, 2012), if we assume that the refreshed pixels respond strongly enough to be detected in the recording at half brightness, then latency will be underestimated by no more than 5 ms.

This method requires more time investment to collect samples than the methods using data acquisition devices in combination with photodiodes and artificial saccades (Bernard, et al., 2007; Bockisch & Miller, 1999), since it is necessary to locate the frames in the video where the key events happen. Therefore the photodiode-based methods are viable alternatives to measuring system latency for labs with access to the required technical expertise. However, the ability to synchronize high speed footage of the stimuli with the eyetracking data is an advantage of the present method, facilitating a good understanding of the stimulus as presented to the eye, particularly as it is affected by the monitor properties (see supplemental high speed videos of monitor performance).

What is an acceptable level of system latency?

The maximum acceptable system latency depends on the properties of the visual system, but also on the application. For example, for a multiresolution display, where areas outside of central vision are represented at lower resolution, latencies up to 60 ms are not typically detected (Loschky & Wolverton, 2007). Latencies in this type of application are less critical because only effects on conscious awareness are relevant, not effects outside of conscious awareness. Another factor is whether it is central vision or the periphery that is being masked. When the periphery is masked, the area that is hidden post-saccade will largely overlap with the area that was hidden pre-saccade, and so longer latencies do not result in the viewer obtaining much extra information that could affect results. However, the delay in unmasking the target location should be taken into account when computing reaction times. On the other hand, when central vision is masked, the post-saccade target will not already be masked (unless it is close to the original fixation location), and so there is the danger of a central-vision glimpse of the target when latencies are long. Other experimental paradigms that depend on low system latencies include gaze-contingent investigations into perception during saccades. Most saccadic suppression experiments use eyetracking to retroactively discard trials where the stimulus was not presented at the correct time during the saccade (e.g. Watson & Krekelberg, 2011). However, when the system latency can be well characterized, more precise and efficient experiments are possible, with the presentation of the stimulus contingent on the time course of the saccade.

Broadly speaking, if the goal is to mask central vision, then given the results on rapid perception of scene gist, and the fact that there is an estimated 5-25 ms of visual suppression after a saccade (Loschky & Wolverton, 2007), experimenters should achieve an average system latency at screen midpoint of less than 25 ms. Of the displays we tested, the Samsung LCD, the CRT, and the DLP display offered this latency level in combination with the other hardware and software we used.

Reducing system latency

The choice of monitor is critical for controlling system latency. The monitor should be capable of refreshing at least 120 times per second, and should have a small or non-existent input lag, as with the CRT and the Samsung 2233RZ, the latter of which has been shown to have other desirable temporal properties (Wang & Nikolic, 2011). The presence of a video converter can add to the latency, as we discovered from the extra 3 ms contributed by conversion from a Mini DisplayPort signal to a VGA signal for use with the CRT monitor. The smallest latency we measured, 16 ms average at the center of the monitor, was obtained with a CRT connected to a PC with a video card that had native VGA output. No currently manufactured Apple computer has a native VGA output, and VGA is being phased out of PCs shipped by major manufacturers (Shah, 2012), so just as CRTs are becoming more difficult to obtain, so are computers that can drive analog displays without the delay introduced by conversion from a digital video signal.

What other steps can be taken? Many displays have features that, when active, perform additional processing on the input, increasing the input lag and therefore the system latency. For example, the Adaptive Contrast

Management mode on the Acer GD235HZ increases average system latency by an average of 3 ms, while the Dynamic Contrast mode on the Samsung 2233RZ increases system latency by an average of 8 ms (both based on 10 measurements). In addition to the added latency of these modes, inspection of the high-speed footage shows that they can cause unexpected time-varying changes to the contrast and brightness. Researchers should consult display documentation to find the settings that correspond to the least amount of input processing possible, which may mean deactivating features with names like Adaptive Contrast Management or Dynamic Contrast, or activating features with names like Video Game Mode which are intended to minimize input lag.

The settings of the eyetracker also affect the system latency, although to a lesser degree than the display does. The Eyelink 1000 can sample at 250 Hz, in which case the sampling stage alone would add an average of 2 ms to the latency (4 ms worst case, 0 ms best case). However there appear to be diminishing returns beyond 1000 Hz (the change from 1000 Hz to 2000 Hz only reduces the mean latency by 0.4 ms, SR Research, 2013). At those frequencies, the processing of the eye images is likely the bottleneck at the eyetracker, not the sample rate. Eyelink eyetrackers provide two levels of heuristic filtering (Stampe, 1993), which reduce noise in the eye position data in real time. However, each level of filtering increases the mean latency by 1 ms (1000 Hz sampling rate, fixed head, SR Research, 2013). Since the measurements reported here used one level of filtering, they overestimate the latency by 1 ms relative to what would be possible with no filtering (Link/Analog Filter set to “OFF”). The “Remote Option”, which enables tracking of the participant’s head so that a head-fixing chinrest is not required, also contributes 1.2 ms to the latency when in use (Eyelink 1000 User Manual version 1.5.2, p 9). Therefore to optimize the latency contributed by the Eyelink 1000, experimenters should use at least a 1000 Hz sampling rate, no heuristic filtering, and fixed head mode (with the viewer using a brow and chinrest). With these settings, this model of eyetracker should be responsible for an average of 1.8 ms of the system latency (Eyelink 1000 User Manual version 1.5.2, p 9). For other models of eyetracker, equivalent latency measurements, and settings to optimize latency, should be obtained from the vendor.

Another approach to reducing latency is to change the scheduling of the eyetracker sampling and subsequent rendering. The measurements reported in our results used an algorithm which samples the eye position as soon as possible after the previous refresh, and then immediately renders the new image updated by the gaze location. After that, program execution blocks until the current screen refresh finishes and the next one is ready to begin. During this time, many fresher eye position samples may become available, but are ignored. An alternative to this sample-then-wait scheme is described in Aguilar and Castet (2011): after the monitor refresh is triggered, rather than sampling the eyetracker again immediately, the program waits to sample until there is just enough time left to render the updated image before the refresh. The advantage to a wait-then-sample scheme is that the next monitor refresh is based on a more recent sample of the eye position. How much reduction in latency can be achieved? For the pixels at the top of the monitor, if the sample-then-wait scheme is used, the delay for the sample to reach the screen will always be at least $T_{refresh}$ ms (where $T_{refresh}$ is the time between screen refreshes), whereas if the wait-then-render scheme is used, the delay for the sample to reach the screen can be as little as the rendering time, as in Figure 3 A). On an iMac computer, which has a $T_{refresh}$ = 16.7 ms, rendering an alpha-blended simulated scotoma at a particular location takes approximately 5 ms. So the system latency would be reduced by over 10 ms. However the potential benefit of waiting to sampling later in the refresh cycle will be considerably less when the refresh rate is higher, for example only around 2 ms on the Samsung monitor at 120 Hz. And the time to wait must be based on an estimate of the rendering time, which should be conservative, to avoid missing the monitor refresh deadline. If this occurs, it would add $T_{refresh}$ to the system latency for that frame.

Decreasing the rendering time, for example by preloading frames of a movie into memory, can be a strategy to reduce latency (O'Sullivan, Dingliana, & Howlett, 2002). This applies particularly when the wait-then-sample algorithm is used, where render time is often the limiting factor. In the sample-then-wait algorithm, on the other hand, the duration of rendering does not make a difference, provided it stays within the available time, which is equal to $T_{refresh}$ minus the time to retrieve a sample from eyetracker. Only if this time is regularly exceeded does the rendering time need to be optimized.

Latency can also be reduced slightly by forcing the monitor to begin using the newly generated image immediately, regardless of where it is in the refresh cycle (Loschky & McConkie, 2000; McConkie, Wolverton, & Zola, 1984). Items near the bottom of the screen would be updated more quickly on average using this option, with an average benefit of 0.5 * $T_{refresh}$ ms, but there is the drawback that the new image could be applied partway through drawing an object of interest. If the object is moving then it may appear distorted, for example by a "shearing" effect. These type of artifacts may be acceptable depending on the application.

Finally, in a simulated scotoma paradigm, the effective average latency will decrease as the scotoma size increases, since for shorter saccades the next fixation point may already be covered by the simulated scotoma at its previous location. This applies when it is only the point of fixation that is important to obscure, and not a complete alignment of the visual field modification with the fixation. The expected reduction in latency depends on the distribution of saccade distances relative to the scotoma size.

Conclusion

Gaze-contingent experiments require control of the latency between eye movements and the corresponding changes to the onscreen stimuli. We have described an efficient and inexpensive way to measure the end-to-end latency of a gaze-contingent system directly, and such measurements should be reported in accounts of gaze-contingent research, in place of estimations based on manufacturers' reported device parameters. Given the past research that some information about natural scenes and other complex stimuli can be perceived from very brief presentations, latencies for experiments simulating visual disorders in particular should be as small as possible. High speed CRTs, running at a high refresh rate, should be preferred, or at least high performance LCDs. But for other purposes, such as development of multiple-resolution displays, displays with poorer latencies can be used. However, the latency should still be measured, and the present method is accessible to most researchers.

References

Aguilar, C., & Castet, E. (2011). Gaze-contingent simulation of retinopathy: Some potential pitfalls and remedies. *Vision Research, 51*(9), 997-1012. doi:10.1016/j.visres.2011.02.010

Allison, R., Schumacher, J., Sadr, S., & Herpers, R. (2010). Apparent motion during saccadic suppression periods. *Experimental Brain Research, 202*(1), 155-169. doi:10.1007/s00221-009-2120-y

Bacon-Macé, N., Macé, M. J. M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research, 45*(11), 1459-1469. doi:10.1162/08989290152001880

Bernard, J.-B., Scherlen, A.-C., & Castet, E. (2007). Page mode reading with simulated scotomas: A modest effect of interline spacing on reading speed. *Vision Research, 47*(28), 3447-3459

Bockisch, C. J., & Miller, J. M. (1999). Different motor systems use similar damped extraretinal eye position information. *Vision Research, 39*(5), 1025-1038. doi:10.1016/S0042-6989(98)00205-3

Bodelón, C., Fallah, M., & Reynolds, J. H. (2007). Temporal Resolution for the Perception of Features and Conjunctions. *The Journal of Neuroscience, 27*(4), 725-730. doi:10.1523/jneurosci.3860-06.2007

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436. doi:10.1163/156856897X00357

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature, 371*(6497), 511-513. doi:10.1038/371511a0

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology, 34*(3), 660-675

Cornelissen, F., Peters, E., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34*(4), 613-617. doi:10.3758/bf03195489

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *Journal of Neuroscience, 20*(9), 3449-3455

Dorr, M., & Bex, P. J. (2011). *A gaze-contingent display to study contrast sensitivity under natural viewing conditions.* Paper presented at the SPIE, Human Vision and Electronic Imaging XVI, San Francisco, CA.

Duchowski, A. T., Cournia, N., & Murphy, H. (2004). Gaze-contingent displays: A review. *Cyberpsychology & Behavior, 7*(6), 621-634. doi:10.1089/cpb.2004.7.621

Elze, T., & Tanner, T. G. (2012). Temporal properties of liquid crystal displays: Implications for vision science experiments. *PLoS ONE, 7*(9). doi:10.1371/journal.pone.0044048

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision, 7*(1). doi:10.1167/7.1.10

Fine, E. M., & Rubin, G. S. (1999). Effects of cataract and scotoma on visual acuity: A simulation study. *Optometry & Vision Science, 76*(7), 468-473

Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision, 6*(9), 858-873. doi:10.1167/6.9.1

Glaholt, M. G., Rayner, K., & Reingold, E. M. (2012). The mask-onset delay paradigm and the availability of central and peripheral visual information during scene viewing. *Journal of Vision, 12*(1). doi:10.1167/12.1.9

Henderson, J. M., McClure, K. K., Pierce, S., & Schrock, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Attention, Perception, & Psychophysics, 59*(3), 323-346. doi:10.3758/BF03211901

Loschky, L. C., & McConkie, G. W. (2000). User performance with gaze contingent multiresolutional displays *Proceedings of the 2000 symposium on Eye Tracking Research and Applications* (pp. 97-103). Palm Beach Gardens, Florida, United States: ACM. doi:10.1145/355017.355032

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied, 8*(2), 99-117

Loschky, L. C., Sethi, A., Simons, D. J., Pydimarri, T. N., Ochs, D., & Corbeille, J. L. (2007). The importance of information localization in scene gist recognition. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1431-1450

Loschky, L. C., & Wolverton, G. S. (2007). How late can you update gaze-contingent multiresolutional displays without detection? *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), 3*(4), 1-10. doi:10.1145/1314303.1314310

McConkie, G. W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods & Instrumentation, 13*(2), 97-106. doi:10.1207/s1532799xssr0104_1

McConkie, G. W. (1997). Eye movement contingent display control: Personal reflections and comments. *Scientific Studies of Reading, 1*(4), 303-316

McConkie, G. W., & Loschky, L. C. (2002). Perception onset time during fixations in free viewing. *Behavior research methods, 34*(4), 481-490

McConkie, G. W., Wolverton, G. S., & Zola, D. (1984). Instrumentation considerations in research involving eye-movement contingent stimulus control. *Advances in Psychology, 22*, 39-47. doi:10.1016/S0166-4115(08)61816-6

O'Sullivan, C., Dingliana, J., & Howlett, S. (2002). Eye movements and interactive graphics. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 555-572). Boston: Elsevier

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422

Reder, S. M. (1973). On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behavior Research Methods & Instrumentation, 5*(2), 218-228. doi:10.3758/bf03200168

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 45*(2), 307-328. doi:10.1518/hfes.45.2.307.27235

Santini, F., Redner, G., Iovin, R., & Rucci, M. (2007). EyeRIS: A general-purpose system for eye-movement-contingent display control. *Behavior Research Methods, 39*(3), 350-364. doi:10.3758/BF03193003

Shah, A. (2012). VGA, DVI Display Interfaces to Bow out in Five Years. *PCWorld*. Retrieved from http://www.pcworld.com/article/248421/vga_dvi_display_interfaces_to_bow_out_in_five_years.html

SR Research (2013). Gaze Contingent and Gaze Control Paradigms, from http://eyelinkinfo.com/solutions_gaz_con.html

Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers, 25*(2), 137-142

Triesch, J., Sullivan, B. T., Hayhoe, M. M., & Ballard, D. H. (2002). Saccade contingent updating in virtual reality *Proceedings of the Eye Tracking Research & Applications Symposium 2002* (pp. 95-102). New York: ACM. doi:10.1145/507072.507092

van Diepen, P. M. J., De Graef, P., & d'Ydewalle, G. (1995). Chronometry of foveal information extraction during scene perception *Studies in Visual Information Processing* (Vol. 6, pp. 349-362): North-Holland. doi:http://dx.doi.org/10.1016/S0926-907X(05)80030-3

van Diepen, P. M. J., Ruelens, L., & d'Ydewalle, G. (1999). Brief foveal masking during scene perception. *Acta Psychologica, 101*(1), 91-103. doi:http://dx.doi.org/10.1016/S0001-6918(98)00048-1

Wang, P., & Nikolic, D. (2011). An LCD monitor with sufficiently precise timing for research in vision. *Frontiers in Human Neuroscience, 5*. doi:10.3389/fnhum.2011.00085

Watson, T., & Krekelberg, B. (2011). An equivalent noise investigation of saccadic suppression. *The Journal of Neuroscience, 31*(17), 6535-6541. doi:10.1523/jneurosci.6255-10.2011

Wiens, S., Fransson, P., Dietrich, T., Lohmann, P., Ingvar, M., & Arne, Ö. (2004). Keeping it short: A comparison of methods for brief picture presentation. *Psychological Science, 15*(4), 282. doi:10.1111/j.0956-7976.2004.00667.x