Contingency Evaluation of Gaze-Contingent Displays for Real-Time Visual Field Simulations

Margarita Vinnikov * Robert S. Allison †
York University York University

Abstract

The visual field is the area of space that can be seen when an observer fixates a given point. Many visual capabilities vary with position in the visual field and many diseases result in changes in the visual field. With current technology, it is possible to build very complex real-time visual field simulations that employ gaze-contingent displays. Nevertheless, there are still no established techniques to evaluate such systems. We have developed a method to evaluate a system's contingency by employing visual blind spot localization as well as foveal fixation. During the experiment, gaze-contingent and static conditions were compared. There was a strong correlation between predicted results and gaze-contingent trials. This evaluation method can also be used with patient populations and for the evaluation of gaze-contingent display systems, when there is need to evaluate a visual field outside of the foveal region.

CR Categories: Computing Methodologies [COMPUTER GRAPHICS]: Methodology and Techniques —Standards; Computing Methodologies [IMAGE PROCESSING AND COMPUTER VISION]: Miscellaneous—Image processing software

Keywords: Head-Eye tracking System, Gaze Contingent, Display, Contingency Evaluation

1 Introduction

Gaze Contingent Displays (GCD) applications are interactive real-time graphics applications that modify the content of a graphical display based on spatial or temporal characteristics of a user's eye movements. GCD are often used to study the effects of visual field characteristics and visual mechanisms [Rayner 1998]. Many early GCD applications were intended to improve effective display resolution under limited video bandwidth, rendering or other constraints. Traditionally, such GCD systems have been evaluated in terms of bandwidth and CPU processing speed improvements. Standard evaluation also includes latency measures such as gaze measurement time and impact on refresh rate time. Such measures have the advantage of being easily specified and validated but in many cases are not sufficient to describe the impact of the system on perception or task performance. Furthermore, it is very important to asses a GCD in terms of the contingency of the display. Many of the existing GCD evaluation techniques rely on a participant's ability to fixate at predefined area on a screen. In essence, these techniques compare how well the GCD can map the center of the simulated fovea to the point of regard. However, this might be problematic in case the participant can not fixate well or at all. For example, this type of task can be nearly impossible when the image in a central visual field is missing or is distorted. Also, such techniques rely on estimates of the impact of the GCD on what can be seen (often foveally or parafoveally where visual acuity and sensitivity often exceed the capabilities of the display device). Overall, although, the GCD methodology has been practiced for several decades, there are still no established procedures for their evaluation and objective comparison.

In this paper we will present a new technique to evaluate the contingency of a GCD system. The technique relies on the assessment of the impact of the GCD system on what is normally unseen. Utility of the technique was examined in experiments on an exemplar GCD system.

2 Visual Field-Based Contingent Display Evaluation

When building a gaze-contingent system one has to overcome hardware limitations and ensure that the inaccuracy and latency of the system do not effect the contingent experience. Furthermore, in complex simulations where the entire visual field is modeled, it is very important to check that the contingency is effective not only in the central regions of the visual field but also in the periphery.

We have developed a visual-field-based contingent display evaluation (VFB-GCE) technique to evaluate the ability of a GCD system to present an arbitrary visual field to a user. The basis for this assessment method is Goldman's perimetry test. Goldman's perimetry test is usually used to detect and determine the location and the size of visual scotomas.

The technique relies on the fact that a GCD application with a properly modeled visual field can produce comparable perimetry measurements to those acquired from participants under controlled and stable fixation as in clinical tests. In other words, VFB-GCE compares participants' psychophysical detection of targets presented in a gaze-contingent visual field simulation (where the head and eye position is not fixed) and a static perimetry session (where the gaze locations are carefully controlled and monitored). This technique can be applied with visually impaired patients as well as with visually healthy individuals. With healthy individuals the technique relies on the localization of the visual blind spot, which is a natural scotoma common to all humans. Furthermore, the blind spot's position is known or can be easily determined and any stimuli shown in this area will not be perceived by an observer. Simple targets can be used removing dependencies on display capabilities (relative to human acuities) and adding minimal processing load or latency to to the GCD system under test.

3 System Description

We have developed a multi-threaded, real-time GCD system that simultaneously tracks the user's eye movements and head pose. The multi-threaded structure of the system dissociates the eye and head

---

*e-mail: mvinni@cse.yorku.ca
†e-mail:allison@cse.yorku.ca

ETRA 2010, Austin, TX, March 22 – 24, 2010.

trackers from the graphical components of the system thus allowing multi-tasking. This minimizes the delay in data collection from trackers. The system is separated into three primary components: tracking, virtual environment rendering, and image processing. The relationships and data flow between the primary components are illustrated in Figure 1. The tracking module is responsible for calculating the line of sight and point of regard from head and eye coordinates that are received from the tracking devices. This gaze information is then used to transform a three-dimensional (spherical) visual field into a two-dimensional map in the display space. The two-dimensional map is essential for creating a realistic visual defect simulation. The current system can simulate visual defects and characteristics such as visual blur, visual distortions, change in color values of the image, and glare. In order to ensure fast image processing, all of these operations are performed at a hardware level. The system also supports simulation of arbitrary visual defects for experimental and demonstration purposes by allowing processing a combination of several basic visual defects. More details can be found in Vinnikov et al. [Vinnikov et al. 2008; Vinnikov 2009].

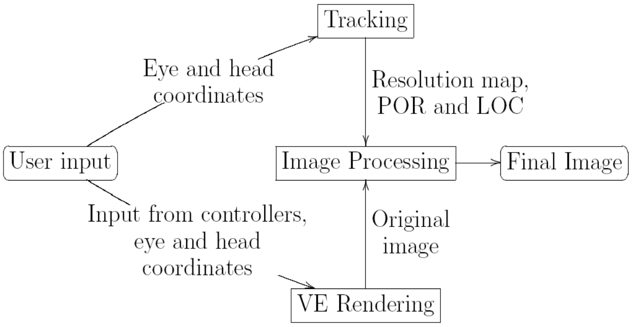

Figure 1: *System structure*

The current head-eye tracking system is based on a video eye tracking subsystem [EL-MAR inc. 2002] and an acoustic-inertial head tracking subsystem [InterSense 2002]. The head tracker is rigidly mounted on the eye tracker's head gear. Both the head tracker and eye tracker work at a synchronized frequency of 120Hz and data are sent to the workstation over an RS232 serial ports at a baud-rate of 38400. The system concurrently collects data from both trackers and performs a series of homogeneous transformations in order to determine the user's real-time gaze direction and position. For more details see Huang [2004].

The system was developed for the Linux operating system on a PC with a 3.2 GHz Pentium 4 processor and 4 GB of RAM. As well all image processing operations were done with Nvidia GeForce 8800 support.

4 Experimental Setup

We have setup an experiment to verify the VFB-GCE technique and to test our own system's performance. The independent variable for this experiment was the display type (static and GCD). During the static display trials, the participant's gaze was not tracked except to monitor fixation as in standard perimetry. During GCD trials, eye and head tracking were used and the participant's gaze direction was estimated in real time. Our experimental hypothesis was that the detection rate of a stimuli will be the same in both gaze contingent and static condition, if the the GCD has no perceptual impact. Furthermore, the clinically determined visual blind spot should be similar in size and location for both the static and the gaze-contingent cases.

4.1 Procedure

The participants were seated in front of a rear-projection screen and donned the head and eye trackers. After calibration, the participants were asked to fixate the center of the screen before stimulus onset. The experimental procedure consisted of a number of steps. First the participant was required to fixate a cross at the center of the screen. Then, the fixation cross was shifted to one of five different locations on the screen and the participant was required to direct her/his gaze to the new fixation location (Figure 2). Then during each trial, a spot of light appeared randomly on the screen in the vicinity of the blind spot. A second spot of light appeared at the end of the trial in the area of fixation. After each trial the participant had to specify how many spots were detected and in the case that two light spots were detected, the participants had to compare the intensity of the two lights. In addition to the psychophysical data collection, the head and eye positions of the participants were recorded.

4.2 Stimuli

The stimuli presented to the participants were a series of light spots with intensity of 60 cd/m^2 on a gray background with average intensity of 3.5 cd/m^2. The size of each light spot was of Goldman standard size III, which under these experimental settings was $0.43°$. There were 15 test stimuli: 5 placed inside the previously measured blind spot, 5 on the boundary and 5 outside of the blind spot (Figure 2). Each stimulus was presented relative to one of two types of reference point. In the static case, each stimulus was presented relative to the fixation point (a green X). In total, there were 5 fixation point locations: one in the center of the screen and the other four were distributed $\pm\, 5°$ horizontally or vertically (Figure 2). Consequently, there were 5 different sets of stimuli for the static case. In the contingent case, there were also 5 fixation points. However, there was only one set of stimuli that was presented dynamically relative to the participant's measured gaze position. Each condition was repeated twice.

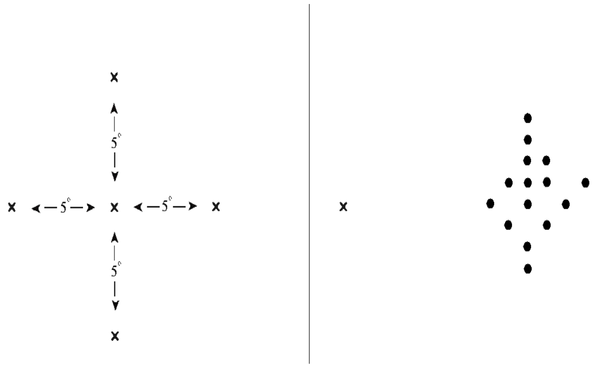

Figure 2: *Possible fixation and stimuli locations*

4.3 Apparatus

The images were projected on rear projection screens with a Barco 808 projector with 1024x768x120Hz resolution. The maximum luminance was 265 cd/m^2 and minimum luminance was $0.1cd/m^2$. The participant sat 100 cm away from the screen. A chinrest was used to stabilize the head and to maintain viewing distance with

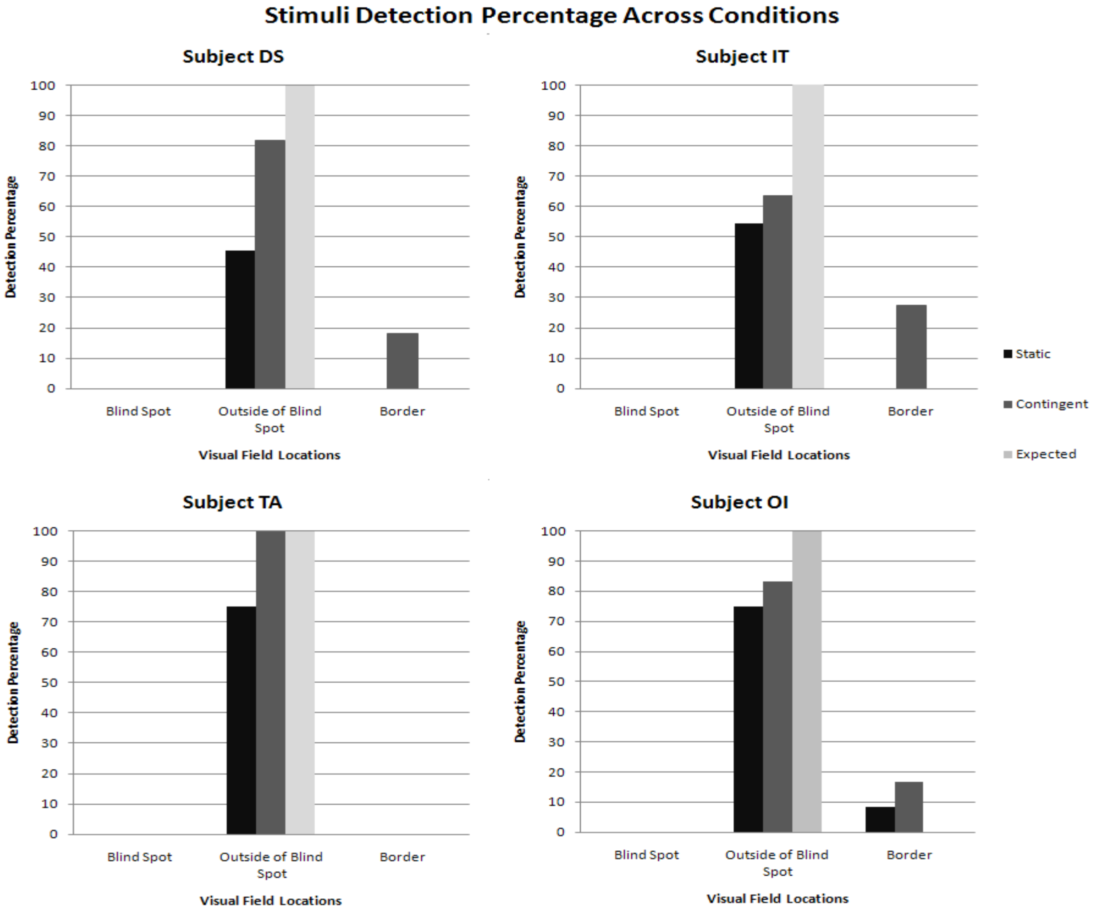

Figure 3: *Results for all four participants*

the right or left eye, as appropriate, centered on and at the level of the center of the screen. Observer's eye movements were recorded using the tracking subsystem. The experiment was conducted in a darkened room.

4.4 Participants

Four university students (3 females and 1 male, ranging in age from 20 to 30, average age 25) participated in the study. All participants had uncorrected distance visual acuity of 20/30 or better and could see clearly at the viewing distance without glasses. Prior to the experiment, a series of simple visual tests (visual acuity, color discrimination, and perimetry) were conducted to ensure that all participants had healthy visual fields. Written informed consent was obtained from all participants in accordance with a protocol approved by the York University Ethics Board.

5 Results

Figure 3 shows the results that were obtained the four subjects. The figure shows the percentage of detected stimuli as a function of location either away from the blind spot (expected to be detected), within the blind spot (expected be never detected), or on the border of the blind spot (where detection rates were expected to be intermediate). In all cases stimuli presented within a participants' blind spots were not detected. However there were significant differences between static and contingent conditions ($\chi^2(2) = 8.44$, p = .015), particularly in the detection rates outside of the blind spot (see Figure 3). In the static case, participants often failed to detect stimuli that were intended to be placed outside of their blind spot. The pattern suggests that the blind spot in the static condition was effectively wider in diameter than for the contingent conditions and for the blind spot determined by Goldman perimetry. For all four participants the number of detected stimuli outside the blind spot was closer to the expected 100% in contingent cases and smaller for the static case. The correlation between expected performance and contingent data is significantly higher than for static case, as can be observed from table 1. A Wilcoxon Signed-ranks test indicated that there was no significant difference between the contingent and expected conditions ($Z = -1.65$, $p = .098$), yet there were significant differences between static and expected ($Z = -2.97$, $p = .003$) and

Table 1: *Correlation between expected performance (EXP) and experimental results for static (ST) and contingent (CON) conditions. Bold face indicates significant correlations (p<0.05)*

Part.	EXP - ST	EXP - CON	ST - CON
TA	**0.52**	**0.81**	**0.67**
OI	**0.70**	**0.76**	**0.86**
DS	0.43	**0.78**	**0.60**
IT	0.46	**0.63**	0.29

static and contingent conditions (Z = -3.12, p = .002).

6 Discussion

The highest correlation between the predicted and actual results was for the gaze-contingent conditions. In other words, when compared with perimetry measurements, the blind spot mapping was more accurate in contingent trials than in static trials. Further, it appears that more stimuli went undetected outside of the mapped blind spot in the static condition than in the contingent condition. Thus, the blind spot is effectively smaller in diameter and better matched to predicted size for the contingent trials than for the static trials.

Another experimental hypothesis was that the gaze contingent stimuli should be detected at the same rate as in the static condition. The difference between static detection and contingent detection could potentially be explained by two factors - errors in tracking estimation and imprecision in fixation. Errors in tracking included the system's inaccuracy and latency. There are several contributing factors to latency and the end-to-end latency using the Barco projector used is 19.22 - 27.22 ms [Vinnikov 2009]. However, the participants had to fixate and thus the eye movements during stimulus presentation should be very small and errors due to latency should be negligible. In particular, the variability of fixation on average was 0.59 degree across participants. On the other hand, a potentially more important factor is the quality of the gaze tracking calibration and eye tracker estimation errors. It is important to note that errors in tracking would predict that the blind spot estimation would be worse under contingent compared to static conditions, which is opposite to the results.

Since effective blind spot size and the correlation with predictions were better in the contingency case than in the static case, we concluded that errors in subject's fixation were a bigger source of variability than GCD tracking errors. In other words, difference between the static and contingent performance appeared to be mainly due to variability in subjects' fixations, which was compensated for in the GCD case but not in the static case. Once again, we found that average variation in fixation was 0.59 degrees. In the static case, this would cause misregistration of the presented stimuli with the expected visual field map and produce errors in blind spot localization, particularly around the boundary of the blind spot.

It can be concluded that VFB-GCE technique is a useful method to assess the contingency of the system. Specifically, this methodology can be useful when peripheral contingency is important, since it not only relies on foveal fixations but also relies on visual properties of other areas of the visual field such as the blind spot. This can be very important for visual disease simulations as well as for experiments with peripheral vision, since the method ensures that a much greater area of the visual field is properly mapped versus only the foveated region in the traditional calibration. VFB-GCE relies on measurement of detection ability in this objectively known and characterized visual field landmark. The mismatch between the GCD and clinical maps of such a well-defined and clear landmark reflects errors in contingency. The technique is simple and relies on straightforward perceptual judgements that do not depend on the fidelity of the display or require measurement of performance decrements on tasks such as visual search. In tasks such as search, subjects may have spare capacity and negative effects of errors in contingency may affect task difficulty without affecting performance measures.

7 Future Work

It will be interesting to compare the methods of GCD evaluation of visual defect simulation described in this paper with previously used methods. We evaluated the technique described in this paper during stable fixation. However, the technique can also be extended to evaluate dynamic contingency. In other words, the system should be tested for conditions when the participants are performing different types of eye movements such as saccades. Errors due to latency for instance would be expected to give increased localization discrepancies during pursuit eye movements or immediately following saccades. Furthermore, we intend to conduct experiments with clinical populations and use different scotomas as reference for contingency validation outside of the fovea.

8 Summary

In this paper we presented a perimetry-based detection task to evaluate the effectiveness of a system's contingency. The method employs visual blind spot localization as well as foveal fixation. During the experimental procedure gaze contingent and static conditions were compared. The experimental results showed that the stimulus detection rate was more closely correlated with the expected results for contingent conditions compared to static conditions. The technique relies on localization of natural visual field features and could be extended to localize pathological scotomas or even full visual field maps. This allows considerable flexibility for use with patient populations particularly those that cannot foveate targets. Furthermore, it can be useful for the evaluation of GCD systems, when the designers are interested in evaluating the visual field outside of the foveal region.

References

EL-MAR INC. 2002. *Video Eye-Tracking System*, 5.0 ed. EL-MAR VISION2000, June.

HUANG, H. 2004. *A Calibrated Combined Head-Eye Tracking System*. Master's thesis, York University.

INTERSENSE. 2002. *IS-900 Theory of Operation and Specification*. InterSense's IS-900 VET & VWT.

LOSCHKY, L. C., MCCONKIE, G. W., YANG, J., AND MILLER, M. 2005. The limits of visual resolution in natural scene viewing. *Visual Cognition*, 1057–1092.

PARKHURST, D., CULURCIELLO, E., AND NIEBUR, E. 2000. Evaluating variable resolution displays with visual search: task performance and eye movements. 105–109.

RAYNER, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin 124*, 3, 372–422.

VINNIKOV, M., ALLISON, R., AND SWIERAD, D. 2008. Real-time simulation of visual defects with gaze-contingent display. *ETRA*, 127–130.

VINNIKOV, M. 2009. *Gaze-Contingent Real-Time Simulation of Impaired Vision in Naturalistic Displays*. Master's thesis, York University.

WATSON, B., WALKER, N., HODGES, L. F., AND WORDEN, A. 1996. Effectiveness of peripheral level of detail degradation when used with head-mounted displays. Tech. rep., GVU Center, Georgia Institute of Technology.

WATSON, B., WALKER, N., HODGES, L. F., AND WORDEN, A. 1997. Managing level of detail through peripheral degradation: effects on search performance with a head-mounted display. *ACM Transactions on ComputerHuman Interaction 4*, 4, 323–346.