Evaluation of Simulated Visual Impairment

Margarita Vinnikov
Department of Computer Science and Engineering
York University
mvinni@cse.yorku.ca

Robert S. Allison
Department of Computer Science and Engineering
York University
allison@cse.yorku.ca

ABSTRACT

We have developed two novel evaluation techniques for gaze-contingent systems that simulate visual defects. These two techniques can be used to quantify simulated visual defects in visual distortion and visual blur. Experiments demonstrated that such techniques could be useful for quantification of visual field defects to set simulation parameters. They are also useful for quantitative evaluation of simulation fidelity based on measurement of the functional relation between the intended simulated defect and psychophysical results.

Author Keywords

Gaze-contingent displays, foveation, visual simulations, evaluation of visual simulations

ACM Classification Keywords

H.5.2.e Information Interfaces and Representation (HCI): User Interfaces Evaluation/Methodology[Evaluation of gaze-based UI]

INTRODUCTION

Humans as a visual species rely on their sight in numerous everyday situations. One approach to study and modify visually-guided behaviour is through the simulation and control of the visual field. For instance, several studies have looked at visual defects associated with spatial resolution [1, 4, 3]. A few researchers have experimented with other possible visual defects such as visual distortions (metamorphopsia) or glare [11, 9, 2]. Nevertheless, what patients with visual deficits really see remains unknown. In order to close the gap between the simulation and the real experience a quantifying technique to evaluate the amount of simulated visual defect is required.

Visual simulations are most effective when combined with gaze-contingent display (GCD), in which the content and the quality of the rendered displays primarily depends on the user's eye and head position. GCD systems are usually assessed in terms of bandwidth performance, GCD latency or successful task completion. However, there are no systematic approaches that evaluate the GCD system through its primary components - the effectiveness of the system's contingency and how well the visual simulation matches a desired result. Several studies looked at different methods to evaluate the GCD systems and to qualify the property of the displayed image. For example, Loschky at al. [7] have looked at explicit measures such as subjective image quality and implicit measures such as eye movement durations. However, their approach does not quantify the actual acuity represented by their algorithm and does not separate contingent the component from the rendering component. Hence, it is impossible to attribute the experimental results to a particular component. On the other hand, if the visual field is modeled correctly then any noticeable distraction can be attributed to problems with contingency of the system. In our earlier work [10], we described a technique to evaluate the effectiveness of a system's contingency. This technique was based on localizing natural visual field features. In contrast, the present experiments look at the second aspect.

In this paper, we present new quantitative methods to verify or to model the simulated visual field for a GCD system. These techniques psychophysically quantify simulated visual disorders such as visual distortions and visual blur with the goal of validating the simulated behaviour against the expected degradation. The presented evaluation techniques are based on visual testing procedures that are usually used with real patients to examine their visual abilities.

SYSTEM DESCRIPTION

To evaluate our techniques, we have developed a real-time GCD system that simultaneously tracks the user's eye movements and head pose. The system is separated into three primary components - tracking, virtual environment rendering, and image processing. The tracking module is responsible for determining the gaze location from the head and eye coordinates that are received from the tracking devices. More details can be found in [11]. The virtual environment rendering module can render any content from 2D images to 3D scenes. For both experiments the stimuli were rear projected with a resolution of 1024 x 768 pixels at a frame rate of 120 Hz. The image processing module transforms the rendered content to reflect users's visual field and gaze position. The system can also combine several visual defects together (see [11] for more details). All image processing operations were done with Nvidia GeForce 8800 support. A chin-rest was used to stabilize the head and to maintain a

2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction.
February 13, 2011, Palo Alto, California, USA.

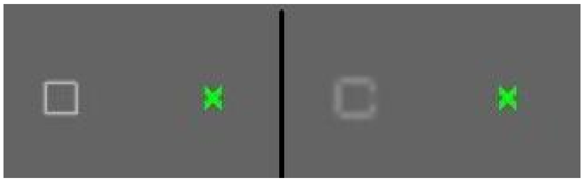

Figure 1. Stimuli examples: highest blur & gap on the right (left); no gap and no blur (right).

viewing distance of 100 cm.

PARTICIPANTS

Three female university students (average age 25) participated in the study. All participants had normal or corrected to normal visual acuity (20/20 or better) and passed a series of visual tests. Participants viewed the display monocularly. One participant tan twice viewing first with right eye and then with the left eye.

EVALUATION OF VISUAL BLUR TECHNIQUE

Visual blur is the most common visual defect or limitation that is simulated in the context of GCDs. Our evaluation technique was developed to quantify the amount of simulated blur and to compare it to visual acuity measurements. This technique capitalizes on a standard procedure that is used to examine visual acuity. We hypothesized that the participants' visual acuity would change as a function of the amount of simulated blur. Errors and imprecisions in GCD could lead to perceptual disruptions such as increase in blur in regions that result in mismatch with user's visual field. On the other hand, if the simulated blur falls below the visual threshold then it will not be perceived. Hence, a more conservative simulation can be still considered to be realistic simulation.

Visual Blur Simulation & Stimuli

The approach for simulating visual blur is a scene-based approach that used graphical hardware support, to achieve blur simulation during a single rendering pass. During the pass the system aligns visual field resolution map and the image of the scene based on the user's gaze position. The degradation level of each pixel is then determined based on an appropriate mipmap level that is encoded in visual field resolution map.

The stimuli used in this experiment was similar to the Landolt C optotype [8]. However, to simplify the stimulus, a square shape was used instead of a ring. Each side of the stimulus subtended 1°. The gap was vertical and was shown either in the middle of the left side or the right side of the square (Figures 1). In addition, three different levels of blur (highest blur (bias 5.4), intermediate blur (bias 2.4), no blur (bias 0)) were applied.

Procedure

During each trial the participant fixated a cross at the center of the screen. Then, a square with a gap on one side was shown for 250 ms at an eccentricity of 5° from the center of

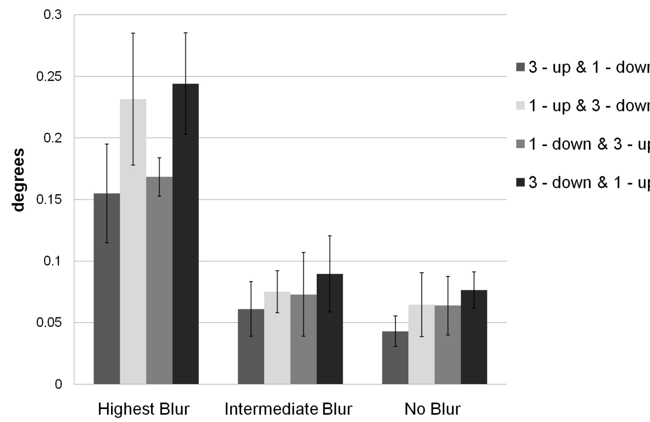

Figure 2. Averaged threshold estimates for all participants. The variability is reflected with the standard deviation bars

fixation. At the end of each trial, the participant had to indicate whether the gap was displayed on the left side of the square or on the right side. The participant's acuity was measured with a transformed Up-Down staircase procedure [6]. Whether the gap changed depended on the staircase rule for sequence (see below). To prevent guessing of the sequence and to further enhance precision, four sequential staircases ran concurrently and randomly intermixed.

Two of the staircase strategies were ascending and started at the lowest gap size (0°), while the other two were descending strategies and started at the maximum gap size (1°). For each case, two strategies were used. For three-up and one-down strategies, the gap was increased after three incorrect responses in succession, and immediately decreased after correct response. The one-up and three-down increased with a single incorrect response, while a decrease required three correct responses in a row. The three-up, one-down and one-up, three-down target the 25% and 75% correct points on the psychometric function respectively and they estimate the upper and lower bounds of the participant's blur sensitivity.

There were two sessions. In the first crude sensitivity thresholds were determined and in the second session more refined thresholds were determined. In order to determine the refined threshold, the threshold was estimated by averaging last ten peaks and valleys of each staircases [6].

Results

Figure 2 shows averaged visual acuity thresholds across participants. From the figure it is apparent that there is a significant decrease in threshold gap size from the highest blur condition to the intermediate blur condition. There was no significant difference between the intermediate and no blur conditions. The simulated ratios of the blur kernel width between different blur levels were 3.0:1.5:1 for high: intermediate: no blur cases. However, the ratio of thresholds between no blur and intermediate blur conditions was about 1, which implies that there was no difference between the results for these two conditions.

Discussion
The experimental hypothesis was that the threshold gap size would grow as a function of amount of simulated blur. Indeed, this hypothesis was supported by the obtained results. Nevertheless, data for the intermediate blur and no blur trials were very similar. However, when directly viewed, the difference in the blur is easily perceived. The explanation for this result is the fact that, since the stimulus was shown 5° away from the center of fixation, the simulated blur was enhanced with the subject's natural acuity decline across the visual field. At 5° eccentricity the subjects natural acuity was the limiting factor. Furthermore, this can also explain why participants could detect smaller gaps on the right rather then on the left. Overall, our technique offers a way to measure perceived blur across the entire visual field and have results comparable with user's visual acuity.

EVALUATION OF SIMULATED VISUAL DISTORTION
Distortion (metamorphopsia) of the visual field is a common symptom of ocular disease such as Age-related Macular Degeneration (AMD). It is typically assessed clinically by sketches or descriptions, such as Amsler grid. However, such assessments do not quantify the degree of distortion. Therefore, our the technique is based on the ability of the participant to discriminate differences in spatial positions of two different segments as a function of imposed distortion. The procedure is similar to a vernier visual acuity test, where one has to line up a point or a segment with a second fixed point or segment as precisely as possible. Shifts in the apparent point of alignment of two objects reflects spatial distortion between them. Our hypothesis was that this technique can also be used to measure perceived visual distortion and to equate a desired distortion with a simulated value.

Visual Distortion Simulation & Stimuli
Visual distortion was achieved by using a modified version of a bump mapping shading techniques. The bump mapping shading technique perturbates each pixel according to the surface normal. The current algorithm deforms the input image at the pixel level by using a normal map as an input.

Two line segments were shown on the screen. Each line segment was 1° long and the two lines were separated with a gap of 3°. The grey background intensity was 3.5 cd/m^2 and the lines were black (0.5 cd/m^2). The lines were either drawn horizontally or vertically. The stationary line was always rendered in the same position on the screen, while the other line segment was randomly placed to the right or to the left of the first segment in the vertical case and above or below the first segment in the horizontal case (Figure 3). For each trial, different levels and direction of distortion were applied to the target area in the visual field. In total, there were 4 levels of horizontal distortions (0.573°, 0.43°, 0.286°, 0.143°), 4 levels of vertical distortions (0.573°, 0.43°, 0.286°, 0.143°) as well as one case where there distortions where applied together. There were 10 conditions x 4 initial offsets x 2 viewing options x 2 repeats. Thus, in total there were 160 trials par subject.

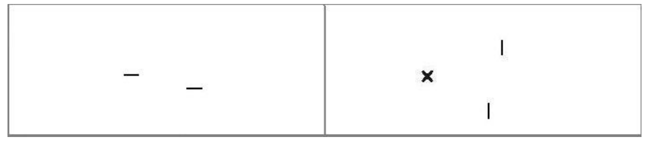

Figure 3. Possible fixation positions and stimuli locations

Procedure
During each trial two line segments were displayed. The participant had to use the arrow keys on the keyboard to align one of the line segments with the other. Once the participant was satisfied with the alignment results, they pressed a key to indicate the end of the trial. There were two types of viewing condition (Figure 3). One set of trials allowed for free viewing, while during the second set of trials, the participant fixated a location 5° away from the two segments. The latter is applicable for GCD evaluation where the segments would be presented in fixed location in the visual field to evaluate distortion of the visual field. We did not present in a GCD manner in order to evaluate the technique without dependence on particulars of a given GCD implementation. However, we conducted an additional experiment, where participant's fixation was tracked. The results were consistent with the main experiment and thus are not presented here.

Results
The alignment settings for both free viewing trials and fixation trials averaged across participants are presented in Figure 4. For free viewing, the alignment adjustment shifted with imposed distortion as predicted when the distortion was in the direction of the alignment task (e.g. horizontal distortion with vertical lines, which need to be horizontally shifted for alignment). On another hand, if the alignment task was perpendicular to the direction of distortion then alignment errors were small, on average 0.04°. This was expected since the distortion should have no effect on the alignment of the line segments (the gap between the lines changes but not their alignment). When the psychophysical results are compared with the predicted results, one can see that the observed values are closer to the predicted results for horizontal alignment and slightly underestimated in the vertical alignment trials. In the fixation tasks, it can be observed that participants had more variability and tended to overshoot or undershoot in their matches. However, these errors were not systematic and on average the data were similar to those in the free viewing tasks.

Discussion
The hypothesis that the bias exhibited in the alignment task would be directly related to the amount of imposed visual distortion was supported by the experimental results from the free viewing trials. A similar pattern was observed in trials with fixations. The task was very precise under free viewing, but it was much more variable with fixation. This

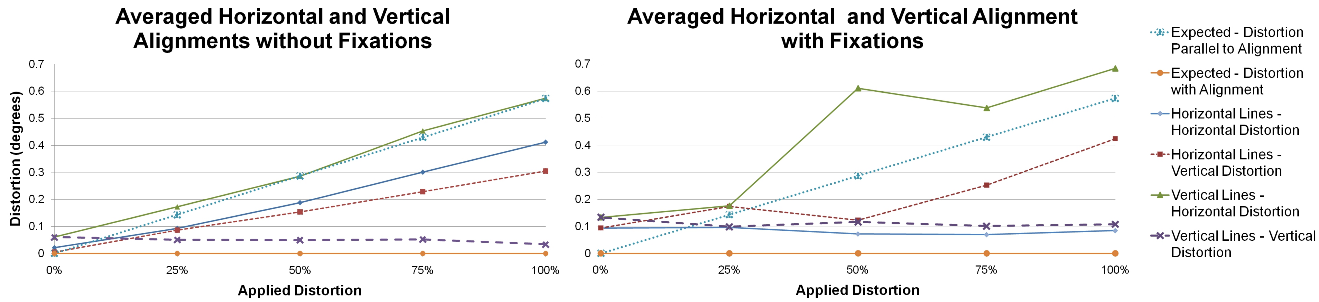

Figure 4. Averaged horizontal and vertical estimations

can be explained by the fact that in addition to the simulated visual defect, the participant's performance was affected by the natural degradation with eccentricity, which is even more pronounced in vernier acuity tasks. Our results are consistent with previous findings [5] that have reported that there is a decline in precision with increasing eccentricity. The ratios between participants' variability in performance between free viewed tasks and fixation tasks was 3.36 for vertical alignment and 3.7 for horizontal alignment. This is an indication that the performance accuracy was less affected by the eccentricity but rather by the stimulus manipulation. Nevertheless, our results show the importance of the area of fixation on the ability to perform a fine detail task, such as alignment. We found a close match between the predicted and obtained alignments. Vernier acuity tasks are precise and repeatable [5]. Thus these tasks are well suited to the measurement of perceived distortion and to the validation of simulated distortion in gaze contingent displays, both in the central visual field and in the near periphery. In the far periphery precision will be an issue and analogous technique based on apparent motion might be preferable.

CONCLUSIONS & FUTURE WORK

The results from the experiments confirmed a strong functional relation between the intended simulation distortion and psychophysical results. These types of evaluation techniques are necessary for ensuring realistic non-subjective clinical simulations when the degree of imposed visual defect is important. As well they provide a methodology to compare the desired visual field to the achieved one that the user is actually perceiving. Furthermore, incorporating these techniques into evaluation practices allows an easy comparison between different systems and algorithms. In the future, we intend to work on similar techniques that can be used to quantify and evaluate other visual defects, so that it will be possible to model a wide range of visual disorders.

REFERENCES

1. Z. Ai, B. K. Gupta, M. Rasmussen, Y. J. Lin, F. Dech, W. Panko, and J. C. Silverstein. Simulation of eye diseases in a virtual environment. In *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, page 5, 2000.
2. A. T. Duchowski and T. D. Eaddy. A gaze-contingent display compensating for scotoma. *EUROGRAPHICS 2009*, 2009.
3. F. C. Fortenbaugh, J. C. Hicks, L. Hao, and K. A. Turano. A technique for simulating visual field losses in virtual environments to study human navigation. *Behavior Research Methods*, 39(3):552, 2007.
4. W. S. Geisler and H. L. Webb. A foveated imaging system to reduced transmission bandwidth of video images from remote camera system. Technical Report 19990025482, AD-A358811, AFRL-SR-BL-TR-98-0858, NASA, Austin TX, 1998.
5. D. M. Levi, S. A. Klein, and A. P. Aitsebaomo. Vernier acuity, crowding and cortical magnification. *Vision Research*, 25(7):963–77, 1985.
6. H. Levitt. Transformed Up-Down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49:467– 447, 1971.
7. L. C. Loschky, G. W. McConkie, J. Yang, and M. E. Miller. Perceptual effects of a gaze-contingent multi-resolution display based on a model of visual sensitivity. In *the ARL Federated Laboratory 5th Annual Symposium-ADID Consortium Proceedings*, page 5358, 2001.
8. J. H. Prince and G. A. Fry. Correction for the guessing bias in the landolt ring test. *Am. J. Ophrhal.*, 46:77–85, 1958.
9. Y. Saito, Y. Hirata, A. Hayashi, T. Fujikado, M. Ohji, and Y. Tano. The visual performance and metamorphopsia of patients with macular holes. *Archives of Ophthalmology*, 118(1):41, 2000.
10. M. Vinnikov and R. S. Allison. Contingency evaluation of gaze-contingent displays for real-time visual field simulations. In *Proceedings of the 2010 symposium on Eye Tracking Research & Applications (ETRA)*, pages 263–266, 2010.
11. M. Vinnikov, R. S. Allison, and D. Swierad. Real-time simulation of visual defects with gaze-contingent display. In *Proceedings of the 2008 symposium on Eye Tracking Research & Applications (ETRA)*, pages 127–130, Savannah, Georgia, 2008. ACM.