Gaze Contingent Displays: Analysis of Saccadic Plasticity in Visual Search

Melchi M. Michel and Wilson S. Geisler
Center for Perceptual Systems, University of Texas at Austin, Austin, Texas, USA

Abstract

Gaze contingent displays were used to simulate visual field loss during performance of visual search tasks. Comparison of human saccadic strategies with those of a Bayesian ideal searcher show that humans rapidly adapt their saccadic strategies, but that this learning is limited and depends on the spatial pattern of vision loss.

1. Introduction

In contrast to the typical digital camera, whose resolution and sensitivity are relatively uniform across the sensor array, the human visual system implements a dramatically inhomogeneous design in which spatial resolution decreases rapidly as a function of angular distance from the center of gaze. This design is part of an elegant solution for a visual system attempting to maximize field of view and spatial resolution while minimizing the required neural resources. Humans encode a large field of view using the low-resolution peripheral retina and use high-speed eye movements (saccades), when necessary, to direct the high-resolution central part of the retina (the fovea) toward behaviorally relevant regions of the scene. Unfortunately, this combination of a variable-resolution retina with frequent and rapid eye movements can make it difficult for vision researchers to precisely control the retinal stimulus and, consequently, to evaluate hypotheses about the way observers use visual information in complex visual tasks that involve many eye movements, such as visual search and reading. Gaze contingent displays, which update the displayed stimulus continuously based on the observer's gaze direction, provide a powerful method for controlling the retinal stimulus during such complex tasks.

As a pertinent example, consider the proposed use of gaze-contingent displays for bandwidth reduction. A number of researchers have suggested that gaze-contingent displays might be useful for reducing the bandwidth necessary to transmit image information by coarsely representing information in the periphery, where the degradation might be imperceptible [1,2]. However, such schemes should be pursued with caution. A recent study [3] investigating visual search with gaze-contingent displays found that even imperceptible changes in the peripheral degradation of visual information can significantly impact search performance. This result underlines the importance of studying complex, ecologically relevant tasks to determine how human observers use visual information, and of using gaze-contingent displays to precisely control the presentation of that information.

In this paper, we report the results of a study that exploited the flexibility of one of these display systems to examine how human observers adapt their saccadic strategies in a visual search task following simulated visual field loss.

2. System Configuration

2.1. Hardware and Experimental Setup

The stimuli were 8-bit images displayed on a calibrated monochrome Image Systems monitor (M2IL) with white phosphor at a frame rate of 60 and a resolution of 800 × 600 pixels located 1 meter from the observer. To provide fine control over target contrast, we combined the 24-bit RGB output signals from the graphics card to drive the electron gun voltage [4]. Eye position was measured using a Fourward-Technologies SRI Mark VI dual Purkinje eye tracker. Head position was maintained using a bite bar and headrest, and eye position signals were sampled from the eye tracker at 500 Hz. An 18-point calibration routine was used to establish a transformation between the output voltages of the eye tracker and the position of the observer's gaze on the computer display.

2.2. Software

Gaze-contingent displays were generated using the Space Variant Imaging System (SVIS) library developed by Perry & Geisler [5]. Though a variety of gaze-contingent systems are now available, we chose the SVIS library because it is fast, lightweight, hardware-independent and freely available (from http://www.svi.cps.utexas.edu). The software takes as input an arbitrary video sequence, a gaze location provided by the output of the eye tracker, and an arbitrary 8-bit integral valued two-dimensional map that specifies the desired display resolution at each eccentricity and direction from the current gaze location and returns the appropriate gaze-contingent variable-resolution video sequence in real time (i.e., about 100Hz for the stimuli used in the current study). The average latency between a gaze measurement and its corresponding display update was 18ms or about 1 frame.

3. Experimental Motivation

When an observer develops a retinal pathology, the pattern of sensitivity across the observer's visual field can change in such a way that the patterns of eye movements, or saccadic strategies, appropriate for certain tasks before the injury become inadequate or suboptimal for performing those same tasks following the injury. A normal observer asked to identify a small target presented in a peripheral part of the visual field for example, typically executes a saccade to bring the image of that target onto the fovea. For a normal observer, this is optimal because the detectability (signal-to-noise ratio) for such a target is greatest in the fovea. However, this may not be the best strategy for an observer suffering, for example, the effects of macular degeneration, which causes loss of foveal vision. Such an observer would gain little or no information by placing the image of the target onto the fovea.

In the current study, we were interested in determining the ability of observers to adapt their saccadic strategies to two different types of simulated changes to their visual fields. The first type, the "shifted fovea," is motivated by the observation that, while researchers have repeatedly demonstrated that humans and monkeys adapt rapidly to the effects of real or simulated occulomotor hypermetria (overshooting) or hypometria (undershooting) by adjusting the gain of their saccades [6], clinical studies in patients with macular disorders [7] and monkeys with macular lesions [8] have indicated that they do not similarly adapt to changes in the preferred retinal location (PRL) following loss of foveal vision. However, such observers do not merely experience a shift in foveal location. Rather, they typically have a complex retinal sensitivity landscape qualitatively different from that of an intact retina. The “shifted fovea” experiment was designed to determine whether observers can adapt to a simulated shift in the location of the fovea.

The second type of simulated change, the “central scotoma”, more closely represents the pattern of loss experienced by patients with central visual field deficits. Studies of these patients report, unsurprisingly, that their performance on a variety of visual tasks is reduced compared to that of normal observers. However, it is unclear whether this reduced performance simply reflects the loss of visual information from the fovea or whether the performance also reflects the use of suboptimal saccadic strategies. Resolving this question requires both a well-defined visual task and a standard of optimal performance for that task, which we provide in the form of the ideal visual searcher [9]. The “central scotoma” experiment was designed to determine how well observers can adapt to a simulated loss of central vision.

4. Experimental Methods

4.1. Human Observers

Two observers participated in each of the experiments described below. One (MMM) was an author of the study while the other (TUB) was naïve to the purpose of the study. Both were experienced psychophysical observers.

4.2. Visibility Maps

The performance of the ideal searcher depends critically on the retinotopic pattern of the detectability of the target in the noise background. For example, a searcher with a uniform pattern of sensitivity to the target across the visual field would gain no benefit from eye movements. Accordingly, our characterization of ideal observers for the search tasks required mapping out these sensitivity patterns (visibility maps) for each of our human observers.

To characterize these visibility maps, detection accuracy for the 6 cpd sine wave target was measured as a function of blur level, target contrast and target location, and detection trials were blocked by blur level, target contrast, and target location. The detection experiment was similar to that reported in Geisler & Najemnik [9]. In each trial, two 250 ms (roughly matching the median duration of individual fixations) stimulus displays were presented in random order, separated by a 500 ms interval. One of these displays contained only a 1/*f* noise background, while the other also contained a target. The observer’s task was to maintain fixation in the center of the display and indicate which of the two intervals contained the target. Psychometric functions [3, Appendix B] were fit to these data in each of 8 directions, along evenly spaced radial ‘spokes’ extending from the center of the display.

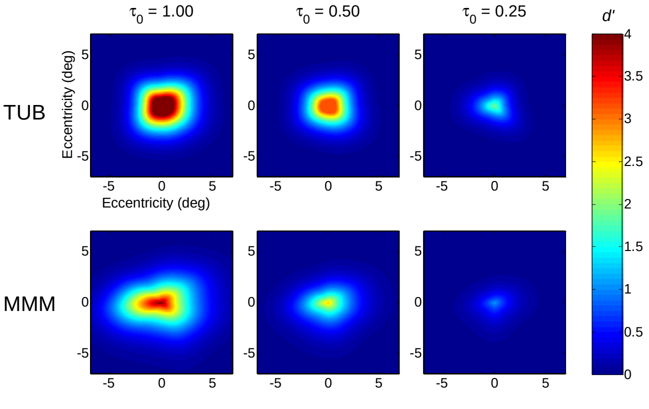

Figure 1. Visibility maps for human observers

Figure 1 illustrates the measured falloff in the visibility of a fixed target (foveal *d’* = 4.0) as a function of retinal eccentricity and blur for each of our two human observers. We indicate the level of blur using the percent of contrast transfer (τ_0) at the target’s peak frequency. Three key features of the resulting visibility maps are that they differ significantly between observers; they are anisotropic, with a generally slower falloff along the horizontal meridian than along the vertical meridian; and the pattern of falloff changes with the amount of stimulus blur. An important consequence of these features is that the search strategy and performance for an ideal searcher differs significantly across human observers and blur levels. For example, switching the ideal searcher’s visibility map from TUB’s to MMM’s cuts the search time by more than half, for a 2% error rate and the same d’ values in the fovea.

4.3. Visual Search Task

In our search task, the target was a small 6 cpd sine wave pattern randomly located at one of 85 target locations densely covering a circular 1/*f* noise background region 13.5 deg in diameter. The observer began each trial by fixating the center of the display and pressing a button. After a random interval of 500-1000ms, the search display appeared. The observer was instructed to find the target as quickly as possible without making any errors. As soon as the target’s location was detected, the observer pressed a button to mark the search time. The observer then fixated the target location and pressed the button again to indicate the target location. To be counted as a correct response, the observer’s gaze direction had to be closer to the actual target location than to any of the other 84 possible target locations. Thus, the probability of being correct by chance was approximately 1.2%. After each trial, an auditory signal indicated whether the observer’s indicated location was correct or incorrect.

In each of the search experiments, human search performance was measured for different target contrast levels set such that they corresponded to 4 different levels of foveal target visibility (*d’= 4.0, 5.0, 6.0, 7.0*). Search trials were blocked by foveal target visibility, with 50 trials per block and 10 blocks per *d’* condition.

4.4. Simulated Searchers

To quantitatively assess the search strategies of the human observers, we compared their search performance in each search experiment to that of three different simulated searchers:

1. An ideal searcher that selects fixation locations optimally.
2. A random searcher that selects its fixations randomly from a uniform distribution across possible fixation locations.
3. A ‘naïve’ searcher that selects its fixations to be optimal under the human observer’s unaltered visual sensitivity map (i.e., as if the display were not gaze-contingent).

Note that in Experiment 1, the display is not gaze-contingent and the ideal and 'naïve' searchers are thus equivalent.

All three of the simulated searchers were optimal in the sense that they used Bayesian updating to evaluate the posterior distribution over target locations. That is, given the same history of fixation locations and observations, all three simulated searchers would compute identical posterior probability distributions over the location of the target. The searchers differed only in how they selected fixations.

The ideal searcher used in our simulations was based on the Bayesian ideal searcher model described elsewhere [3,8]. Briefly, we assumed that the ideal searcher's goal is to find the target as quickly as possible, with the constraint that the average target localization accuracy exceeds a particular criterion value c that was fit to match the localization accuracy of the human observer.

The ideal searcher begins in the center of the display, assuming equal prior probabilities over each of the 85 possible target locations. At each time step, the ideal searcher collects matched-template energy responses in parallel across all target locations and uses the responses encoded during this fixation to compute the posterior probability for the target at each location following the current fixation. If one of these probabilities exceeds the criterion (c), the search stops, otherwise the observer calculates the fixation location that maximizes the prospective probability of correctly identifying the target location, and fixates that location, continuing the search.

The ideal searcher is used to provide an upper limit on the performance of our human searchers. A searcher whose performance matches that of the ideal is performing optimally. The other two simulated observers provide upper bounds on human performance given certain plausible suboptimal search strategies. In particular, the random searcher provides an upper bound on performance for an observer that selects its fixations randomly, while the 'naïve' searcher provides an upper bound on performance for an observer that cannot learn a new strategy following visual loss. These observers can be useful in ruling out certain human search strategies. For example, if human observers outperform the random searcher, then we can conclude that humans do not select their fixations randomly. Likewise, if human observers outperform the 'naïve' searcher, then we can conclude that humans indeed learn to use new strategies following simulated visual loss.

4.5. Visual Search Experiments

4.5.1. *Experiment 1: Unfoveated Search*

If human observers were inefficient in a standard visual search task, then it would not be surprising to find that they are also inefficient when performing visual search with a distorted sensitivity map. Based on previous results [3,8] however, we expected observers to achieve near-ideal performance in the standard (unfoveated) search task. To confirm that our observers are indeed efficient in the unfoveated search task, we first ran the observers in an experiment replicating the task used in Najemnik and Geisler [9]. A key difference in our version of the task, however, is that the stimuli were uniformly blurred to approximate an 'unshifted' version of Experiment 2, described below.

Figure 2 shows the performance of our human observers in Experiment 1, along with the performance of the random and ideal searchers (Note that the grey bars represent error rates for both the human and simulated searchers). In line with previous results [9], human searchers outperformed the random searcher with performance near that of the ideal searcher.

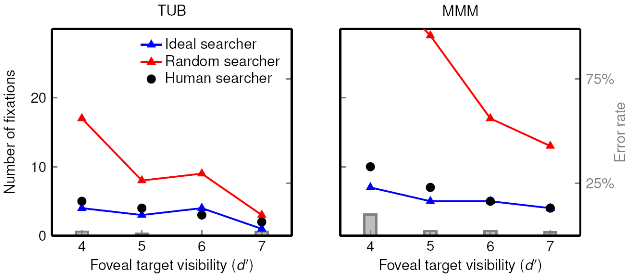

Fig. 2. Visual search performance in Exp. 1 (unfoveated search).

4.5.2. *Experiment 2: 'Shifted Fovea' Search*

In the 'shifted fovea' experiment, we sought to simulate the effect of physically translating the location of the observer's highest resolution region in the visual field upward by 2.5 degrees, displacing the pattern of falloff in visibility measured in the detection experiment. This transformation requires reducing resolution in some regions of the visual field while increasing resolution in other regions. Of course, we cannot directly increase the resolution of an observer's retinae. Starting with a blurred stimulus, however allowed us to arbitrarily change the pattern of *relative* resolution using the following algorithm:

1. Choose an initial blur level τ_0 for the display.
2. Compute the desired d' for the current target at all locations in the 'shifted' visibility map.
3. Calculate the amount of blurring or unblurring (contrast transfer) required to achieve the desired d' at each location in the current visibility map.
4. If the maximum amount of transfer required in step (3) $\leq 1/\tau_0$, stop. Otherwise, reduce τ_0 and return to step (2).

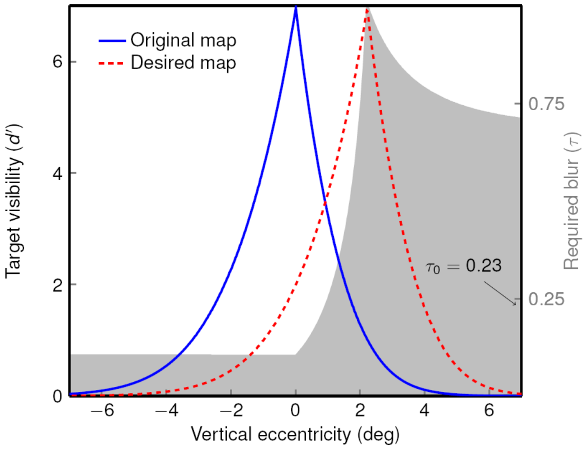

Fig. 3. Shifting the observer's visibility map

Figure 3 illustrates a vertical cross-section through such a shift for observer MMM. The blur level τ_0 (which varied across observers and d' levels) was used both as the baseline blur for Experiment 2, and as the uniform blur applied to the corresponding stimulus in Experiment 1. In Figure 3, this means that blur levels above τ_0 require unblurring with respect to the baseline blur, while blur levels below τ_0 require additional blurring.

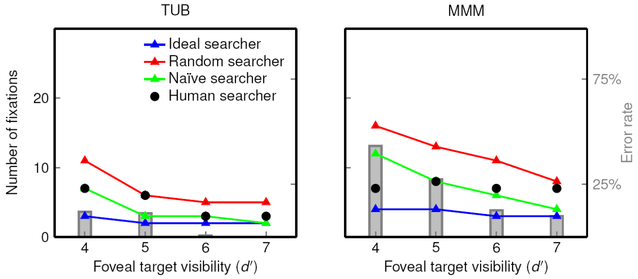

Fig. 4. Visual search performance in Exp. 2 (shifted fovea).

Figure 4 shows the search performance of the human, ideal, 'naïve', and random searchers in the 'shifted fovea' search task. Comparing these results to those of Experiment 1, it is immediately evident that error rates have risen substantially and that the median number of fixations required by the simulated searchers has fallen. Recall that the 'quitting' criterion *c* for the simulated searchers is controlled by the human observer's detection accuracy. The lowered accuracy in this experiment led to lower *c* values and shorter searches for the simulated searchers.

It is also clear that humans now fall quite short of optimal performance, with performance nearer to that of the 'naïve' searcher.

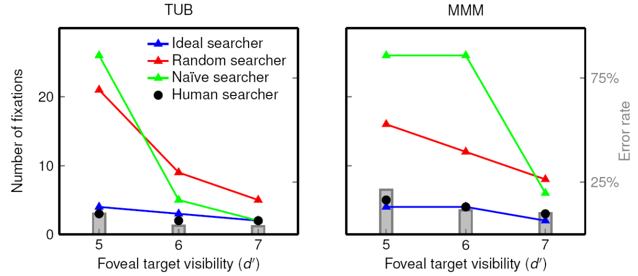

Fig. 5. Visual search performance in Exp. 3 (central scotoma).

4.5.3. Experiment 3: 'Central Scotoma' Search

In the 'central scotoma' experiment, we sought to simulate the effect of foveal vision loss by blurring the central 4 degrees of the observer's visual field so that it carried virtually no information (τ = 0.10) in the target's frequency band. The stimulus was unblurred (i.e., τ = 1.0) in the periphery, and the indicated foveal target visibilities (d'= 4, 5, 6) are labeled according to their unblurred detectability.

Figure 5 shows the search performance of the human, ideal, 'naïve', and random searchers in this task. Though errors remain somewhat higher than in Experiment 1, human performance in this task matched that of the ideal observer and far exceeded the performance of the random and 'naïve' searchers.

5. Conclusions

We used a gaze contingent display along with an ideal observer model of visual search to investigate whether human observers can appropriately modify their saccadic strategies following two types of changes in their visibility maps. When we simulated shifting the fovea, observers' performance moved away from that of the ideal observer, and nearer to that of a naïve observer using an unmodified search strategy. This failure to adapt saccadic strategies for a shifted visibility map supports the suggestion [8] that humans cannot quickly learn to center saccadic endpoints on a new retinal location (i.e., a new oculomotor locus). However, when we simulated a central scotoma, humans were able to match the performance of the ideal observer, greatly outperforming the simulated searcher using an unmodified search strategy. This suggests that following central vision loss humans may be able to quickly adapt their eye movement strategies for certain common visual tasks, and that residual performance deficits may largely be due to visual information loss.

6. Acknowledgements

This work was supported by NIH Grant EY02688.

7. References

[1] E. M. Reingold, L. C. Loschky, G. W. McConkie, D. M. Stampe, "Gaze-contingent multiresolution displays: An integrative review," Human Factors 45(2), 307-328, (2003).

[2] A. T. Duchowski, and A. Çöltekin, "Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging," ACM Trans. Multimedia Comput. Comm. Appl. 3(4), Article 24 (2007).

[3] W. S. Geisler, J. S. Perry, and J. Najemnik, "Visual Search: The role of peripheral information measured using gaze-contingent displays," Journal of Vision 6(9):1, 858–873, (2006).

[4] D. G. Pelli and L. Zhang, "Accurate control of contrast on microcomputer displays." Vision Research 31, 1337-1350, (1991).

[5] J. S. Perry and W. S. Geisler, "Gaze-contingent real-time simulation of arbitrary visual fileds." In: B. Rogowitz and T. Pappas (Eds.), Human Vision and Electronic Imaging, SPIE Proceedings, (2002).

[6] J. J. Hopp, A. F. Fuchs, "The characteristics and neuronal substrate of saccadic eye movement plasticity," Progress in Neurobiology 72, 27-53, (2004).

[7] S. G. Whittaker, R. W. Cummings, L. R. Swieson, "Saccade control without a fovea," Vision Research 31, 2209-2218, (1991).

[8] S. J. Heinen and A. A. Skavenski, "Adaptation of saccades and fixation to bilateral foveal lesions in adult monkey," Vision Research 32, 365-373, (1992).

[9] J. Najemnik and W. S. Geisler, "Optimal eye movement strategies in visual search," Nature 434, 387-391 (2005)