

Visual search: The role of peripheral information measured using gaze-contingent displays

Wilson S. Geisler

Center for Perceptual Systems and Department of Psychology,
University of Texas at Austin, Austin, TX, USA



Jeffrey S. Perry

Center for Perceptual Systems and Department of Psychology,
University of Texas at Austin, Austin, TX, USA



Jiri Najemnik

Center for Perceptual Systems and Department of Psychology,
University of Texas at Austin, Austin, TX, USA



Two of the factors limiting progress in understanding the mechanisms of visual search are the difficulty of controlling and manipulating the retinal stimulus when the eyes are free to move and the lack of an ideal observer theory for fixation selection during search. Recently, we developed a method to precisely control retinal stimulation with gaze-contingent displays (J. S. Perry & W. S. Geisler, 2002), and we derived a theory of optimal eye movements in visual search (J. Najemnik & W. S. Geisler, 2005). Here, we report a parametric study of visual search for sine-wave targets added to spatial noise backgrounds that have spectral characteristics similar to natural images (the amplitude spectrum of the noise falls inversely with spatial frequency). Search time, search accuracy, and eye fixations were measured as a function of target spatial frequency, $1/f$ noise contrast, and the resolution falloff of the display from the point of fixation. The results are systematic and similar for the two observers. We find that many aspects of search performance and eye movement pattern are similar to those of an ideal searcher that has the same falloff in resolution with retinal eccentricity as the human visual system.

Keywords: visual search, eye movements, contrast, noise, ideal observer, gaze-contingent displays, peripheral vision

Introduction

For humans and other primates, there are few perceptual tasks more fundamental than visual search, which is performed almost continuously during waking hours and is essential for locating and interacting with objects, places, and other organisms. To perform visual search, the visual system implements an elegant compromise between the competing goals of maximizing field of view, maximizing spatial resolution, and minimizing neural resources: It encodes a large field of view with a retina having variable spatial resolution and then uses high-speed eye movements to direct the highest resolution region of the retina (the fovea) at potential target locations in the visual scene. In human visual search, eye fixations have a duration averaging 200–300 ms. During each of these fixations, detection and identification processes are applied across the visual field, and eye movements to subsequent fixation locations are planned and programmed.

Although the generic form of visual search in the normal environment involves multiple fixations, most research has focused on response-time tasks designed so that the number of fixations is minimal (e.g., see Wolfe, 1998) or on single-fixation tasks, where stimuli are presented briefly so that only a single fixation is possible

(e.g., see Palmer, Verghese, & Pavel, 2000; Verghese, 2001). The single-fixation task has the advantages of allowing relatively easy control over eccentricity effects, of reducing the need to measure eye movements, and of facilitating the study of covert selection. However, the single-fixation task is relatively unnatural and provides little information on some important aspects of visual search: planning and programming of eye movements (fixation selection), and integration of information across successive fixations.

Fixation eye movements in visual search tasks are complex and depend on the information collected across the visual field during search, as well as on the observer's prior knowledge of the task and stimuli (for a review, see Findlay & Gilchrist, 2003). For example, (1) If a target is highly visible, then observers tend to make a saccade directly toward the target (Eckstein, Beutter, & Stone, 2001; Findlay, 1997), but under more difficult conditions, observers may fixate some average location within a group of possible target locations (Findlay, 1997; He & Kowler, 1989; Zelinsky, Rao, Hayhoe, & Ballard, 1997); (2) The duration of fixations during visual search tends to increase as the discriminability of the target from background decreases (Hooge & Erkelens, 1999; Jacobs & O'Regan, 1987); (3) The "classification image" technique (Beard & Ahumada, 1998) applied to visual search for targets in noise shows that the eye is attracted (at least

some of the time) to features in the noise that match features of the target (Rajashekar, Cormack, & Bovik, 2002). The clear implication of these and other studies (e.g., Engel, 1977; Geisler & Chou, 1995; Motter & Belky, 1998; Zelinsky, 1996) is that if we are to understand multiple-fixation visual search, then we need to understand what information is being extracted from the periphery during search and how it is being used to guide the sequence of eye movements.

The role of peripheral information in multiple-fixation search is not easy to study because of the difficulty of controlling the stimulus on the retina during a sequence of eye movements. A powerful technique to solve this problem is to update the visual display frame-by-frame, contingent on the observers' eye position, as measured with an eye tracker. The earliest use of such "gaze-contingent displays" was to stabilize images on the retina (Riggs, Ratliff, Cornsweet, & Cornsweet, 1953; Yarbus, 1967). The most productive use of gaze-contingent displays has been in the study of reading (for a review, see Rayner, 1998). For example, McConkie & Rayner (1975), and similar subsequent studies, found that normal reading is achieved with about 14–15 letter spaces to the right of fixation and about 3–4 letters to the left of fixation.

Gaze-contingent display techniques can provide similar information for other complex tasks that involve eye movements, such as visual search. For example, Loschky & McConkie (2002) investigated visual search in natural scene images using a gaze-contingent display with two levels of resolution and found that shrinking the window of high resolution produced longer search times, more fixations, shorter saccade lengths, and longer fixation durations. Similar results were obtained by Kortum and Geisler (1996a, 1996b) using a display where pixel size increases with distance from the point of fixation. A limitation of these studies is that display resolution jumps in discrete steps rather than smoothly dropping off, a common problem in early gaze-contingent display systems (Juday & Fisher, 1989; Warner, Serfoss, & Hubbard, 1993; Weiman, 1990).

Recently, there has been considerable effort to improve gaze-contingent displays (for reviews, see Duchowski, Cournia, & Murphy, 2004; Parkhurst & Niebur, 2002; Reingold, Loschky, McConkie, & Stampe, 2003). As part of this effort, we developed software that runs on standard PCs and that is able to produce artifact-free, gaze-contingent displays at update rates of 40–60 Hz (Geisler & Perry, 1998; Perry & Geisler, 2002).

Here, we describe experiments where this software was used to measure how peripheral information is used during visual search for sine-wave targets (Gabor patches) embedded in broadband noise whose amplitude spectrum falls off inversely with frequency ($1/f$ noise). We varied the spatial frequency of the target, the contrast of the noise background, and the rate of falloff in display resolution from the point of gaze. Our aims were to obtain a broad picture of search performance under naturalistic condi-

tions and to obtain an estimate of how much information can be removed from the periphery without affecting eye movement patterns or search time. Najemnik and Geisler (2005) showed that modest differences in peripheral detection sensitivity across stimulus conditions can substantially affect search time; thus, our expectation was that mild reductions in peripheral resolution would significantly increase search time.

There are several reasons for measuring search performance in $1/f$ noise. First, the Fourier amplitude spectra of natural images fall off approximately as $1/f$ (Burton & Moorehead, 1987; Field, 1987), and thus, searching for targets in $1/f$ noise should be representative, in at least some ways, of search in the natural environment. Second, there is substantial literature concerning the detection and identification of targets in broadband noise; this literature provides a solid foundation for understanding search performance in broadband noise (Burgess, Wagner, Jennings, & Barlow, 1981; Lu & Doshier, 1999; Pelli & Farell, 1999). Third, it is possible to derive an ideal observer theory of visual search for targets in broadband noise (Najemnik & Geisler, 2005); this ideal searcher provides the appropriate benchmark against which to compare real performance and a useful starting point for proposing realistic (suboptimal) models of visual search.

Methods

Search performance and fixation patterns were measured for a sine-wave target randomly located in a $1/f$ noise background, as a function of target spatial frequency, noise contrast, and the rate of falloff in display resolution from the fixation location. There were two observers with normal vision; one was an author while the other was naive to the aims of the study.

Stimuli

Eight-bit gray-scale images were displayed on a calibrated monitor (Phillips Brilliance 21A) that was set to a resolution of 640×480 pixels at 60 Hz noninterlaced and placed at a distance of 120 cm from the eyes. The target stimuli were one-octave Gabor patterns (in sine phase) tilted 45 deg to the left. Four target spatial frequencies were tested: 1, 2, 4, and 6 cpd. The root-mean-square (rms) contrast of the target was fixed at 0.35.

The background was a circular region 13 deg (400 pixels) in diameter, filled with noise having an amplitude spectrum that declined inversely with spatial frequency (i.e., $1/f$ noise); the remaining display pixels were set to the mean luminance (20 cd/m²). The $1/f$ noise was created by filtering white noise, truncating the waveform at $\pm 2 SD$, scaling to obtain the desired rms amplitude and then

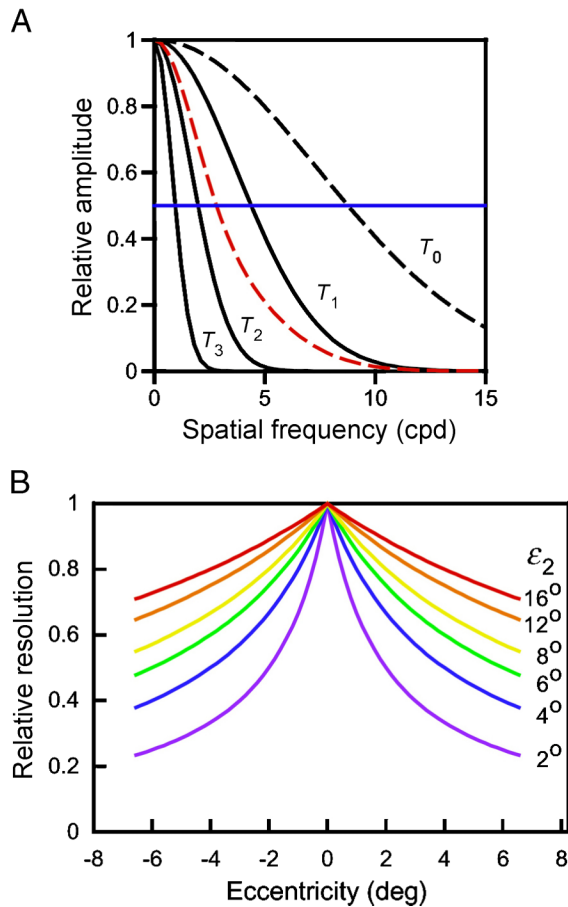


Figure 1. Creation of gaze-contingent displays. (A) Transfer functions for the first three levels of the multiple-resolution Gaussian pyramid. The dashed red curve shows the transfer function associated with a particular interpolation between Pyramid Levels 1 and 2. T_0 is a hypothetical transfer function used for interpolation between the original image and the first level of the pyramid. The horizontal blue line intersects the transfer functions at the half-height resolution, $r(\epsilon)$. (B) Relative resolution maps, $r(\epsilon)/r_0$, used in the present experiment; the parameter ϵ_2 is the eccentricity where display resolution reaches one half of the maximum value.

adding a constant to obtain the mean luminance. Four noise contrast levels were tested: 0.25, 0.125, 0.0625, and 0.03125 rms.

On each search trial, the target was placed at a different random location within the circular noise background; however, the target center was not allowed to fall within 40 pixels of the edge of the background. A different random sample of $1/f$ noise was displayed on each search trial.

Gaze-contingent displays were generated using the method described in Perry & Geisler (2002). The algorithm takes as input an arbitrary video sequence, a gaze location provided by the output of an eye tracker, and an arbitrary real-valued two-dimensional map that specifies the desired display resolution at each eccentricity and direction from the current gaze location.

To create the gaze-contingent displays, the following operations were performed in real time, on each video frame: (1) Obtain an input image (in this case the same image for each video frame of the search trial). (2) Compute a multiresolution Gaussian pyramid representation of the input image (typically six to seven levels deep). The solid curves in Figure 1A show the transfer functions associated with the first three levels of the Gaussian pyramid. (3) Obtain the current gaze direction from the eye tracker. (4) Shift the resolution map to align with the current gaze direction. (5) Up-sample and interpolate the multiresolution pyramid using the shifted resolution map. (6) Display the output image. (7) Go to Step 1. The software that performs these operations runs on a standard PC and is available in either C++ or MatLab, from the authors or from the web site <http://www.svi.cps.utexas.edu/>.

The transfer function associated with the j th level of the Gaussian pyramid is given by

$$T_j(f) = \exp\left(-0.5\left(\frac{2^j f}{\sigma_0}\right)^2\right), \quad (1)$$

where f is spatial frequency in cycles per degree, $\sigma_0 = 0.248w_{\text{pix}}/w_{\text{deg}}$, w_{pix} is the width of the display in pixels, and w_{deg} is the width of the display in degrees. The half-height resolution of the j th level of the Gaussian pyramid is $r_j = \sigma_0\sqrt{2\ln 2}/2^j$. The horizontal blue line in Figure 1A intersects the pyramid transfer functions at their half-height resolution (thus, half-height resolution is also in units of cycles per degree). The dashed black curve in Figure 1A is the hypothetical transfer function obtained when $j = 0$ in Equation 1. We use this transfer function for interpolation between the original image and the first level of the pyramid. In effect, we are assuming that the original image is the first level in a Gaussian pyramid for an image with twice the resolution of the original image. This trick keeps the interpolation procedure between the original image (Level 0) and Level 1 of the pyramid consistent with the interpolation procedure between other neighboring levels of the pyramid (e.g., between Levels 1 and 2).

The local transfer function at an arbitrary interpolated resolution, r , is given by

$$T(f, r) = \begin{cases} \frac{[0.5 - T_{j+1}(r)]T_j(f) - [0.5 - T_j(r)]T_{j+1}(f)}{T_j(r) - T_{j+1}(r)}, & r_{j+1} < r \leq r_j \\ \frac{[0.5 - T_1(r)] - [0.5 - T_0(r)]T_1(f)}{T_0(r) - T_1(r)}, & r_1 < r \leq r_0 \end{cases} \quad (2)$$

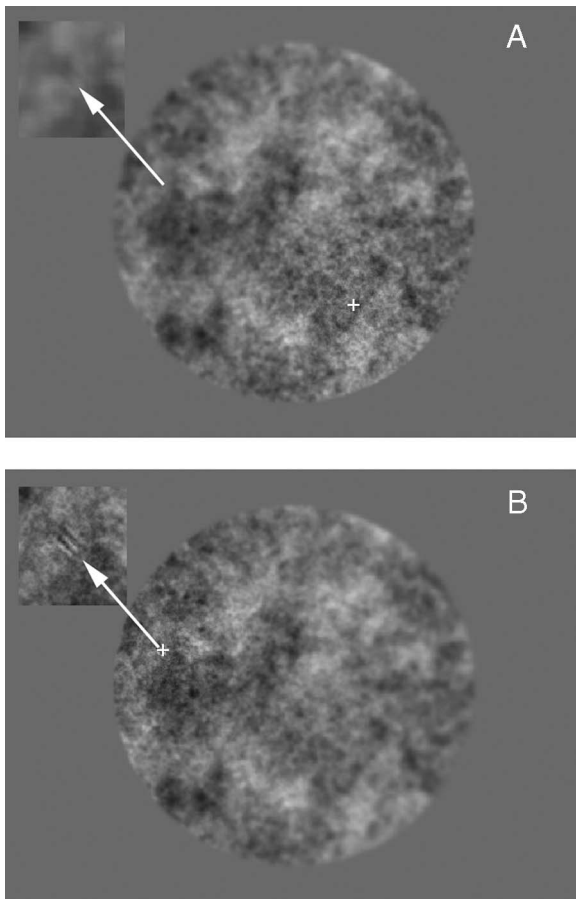


Figure 2. Gaze-contingent display during search task. The stimulus consisted of a Gabor patch target added to a 13 deg background of $1/f$ noise. (A) Gaze-contingent display when fixation is in lower right (white plus sign). (B) Gaze-contingent display when fixation is on the target in the upper left. Insets show enlargements of the region containing the target.

For example, see the dashed red curve in Figure 1A. In other words, this equation precisely specifies the local transfer function that is applied to the original image at each pixel, for any display resolution r desired at that pixel location.

Figure 1B shows cross sections of the radially symmetric resolution maps used in this study. The relative display resolution was highest at the point of fixation and declined smoothly away from that point according to the function,

$$\frac{r(\varepsilon)}{r_0} = \frac{\varepsilon_2}{\varepsilon_2 + \varepsilon}, \quad (3)$$

where ε_2 is the eccentricity (in degrees) at which the display resolution drops to one half of its value at the fixation point. The value of ε_2 controls the rate of falloff in display resolution; the smaller the value of ε_2 , the faster

the rate of falloff. Six values of ε_2 were tested: 2, 4, 6, 8, 12, and 16 deg.

The falloff in display resolution created using Equation 2 is similar in shape to the falloff in resolution of the human visual system (see, e.g., Geisler & Perry, 1998; Wilson, Levi, Maffei, Rovamo, & DeValois, 1990). The half-resolution constant of the human visual system (e_2) is in the range of 2.0–2.5 deg. Thus, the most rapid falloff of display resolution in our experiment is only slightly more rapid than the falloff in resolution of the visual system. However, it is important to keep in mind that the display resolution combines multiplicatively with visual resolution. For example, if ε_2 is set equal to e_2 , then at an eccentricity of 2–2.5 deg, the total effective resolution is reduced by a factor of 4 rather than a factor of 2.

Figure 2A illustrates the appearance of a display ($\varepsilon_2 = 4$ deg, background contrast = 0.25 rms, target frequency = 6 cpd) when fixation is to the right and below the center of the display. Figure 2B illustrates the appearance of the same display when the fixation is on the target. The insets show enlargements of the region containing the target.

In almost all conditions, eye positions were sampled and the display was updated at a rate of 60 times per second; however, a few conditions (the largest values of ε_2) required greater image processing, and thus, the update rate was only 45 times per second. However, these differences had no noticeable effect on the display because the display frame rate always remained at 60 Hz non-interlaced. For more details, see Perry and Geisler (2002). This subjective impression is consistent with a recent gaze-contingent, blur-detection experiment (Loschky & McConkie, 2005).

Gaze direction was measured with an SRI version-6 dual Purkinje eye tracker. Head position was maintained with a bite bar and heavy-duty headrest. The algorithm that was used to compute fixation points from eye positions was a modified version of one in the Applied Science Laboratories Series 5000 data-analysis software (see Appendix A). Observers were allowed to blink between search trials. Loss of tracking occurred very infrequently, and there were no noticeable eye-tracking artifacts during the search trials.

Procedure

Two 30-trial blocks of search trials were run for each of the 96 stimulus conditions (4 target spatial frequencies \times 4 background noise contrasts \times 6 ε_2 values). One block was run for each of the 96 stimulus conditions, then, after all 96 conditions were completed, the second block was run for each of the 96 conditions, but the order of conditions was reversed. The observer knew the target spatial frequency, background noise contrast, and value of ε_2 in each block.

Prior to each 30-trial block, the observer was required to execute a 9-fixation-point calibration procedure. In

addition, after each trial, a gaze-location marker appeared. If this marker was not within close tolerance of the central fixation dot (which the observer was trying to fixate), the observer was given the opportunity to run the calibration procedure again (this rarely occurred).

Each search trial began with the observer keeping fixation on a central dot in a uniform field at mean luminance. When the observer was ready, he or she pushed a response key, which caused the fixation dot and gaze direction marker to disappear. After a random time delay of 500–1,500 ms, the search display appeared; this was done to reduce any temptation to make anticipatory responses in the easy blocks. The instructions were for the observer to find the target as quickly as possible without making errors. As soon as the observer identified the target location, he or she pressed the response key, which determined the search time. The observer then fixated the identified target location and pressed the response key again. To be counted as a correct response, the gaze direction at the time of this second press had to be within 1 deg of the actual target location (thus, the probability of being correct by chance was approximately 2.5%). After the observer responded, the true location of the target was indicated with a small dot. Finally, the uniform field with the central fixation dot and gaze direction marker reappeared.

Ideal observer analysis

To facilitate interpretation of the data, we compared the measured search performances and fixation patterns with those of a Bayesian ideal searcher, which is described in detail elsewhere (see Najemnik & Geisler, 2005 and the supplement to that publication). Briefly, we derived the ideal searcher for a task where a known target is located randomly in a field of spatial noise. We assumed that there are n possible nonoverlapping target locations and that the searcher's goal is to find the target as quickly as possible, with the constraint that the average target localization accuracy exceeds some particular criterion value.

The ideal search strategy is as follows. The searcher begins with fixation in the center of the display, and it assumes that all target locations are equally probable (which they were in the experiment). These initial probabilities across the possible target locations are, in Bayesian terminology, the “prior probabilities.” At the onset of the search display, the searcher collects matched-template responses in parallel from all possible target locations while fixating the center of the search area. The searcher is given the same visibility map (detection sensitivity across the visual field) as the human observer, and thus, the template responses in the fovea are generally more informative than those in the periphery. The searcher uses the responses encoded during this first fixation to compute the “posterior

probability” of the target being at each of the possible target locations. If the maximum of these posterior probabilities exceeds a criterion (e.g., 95%), then the search is stopped and the location where the posterior probability exceeded the criterion is reported. If the criterion is not exceeded, the searcher computes (using the current posterior probabilities and knowledge of its own visibility map) the fixation location that will maximally increase the likelihood of correctly identifying the target location after the eye movement is made. The optimal next fixation location is not necessarily the location with the highest posterior probability because it is often possible to gain more information by fixating elsewhere (e.g., the centroid of several locations with high posterior probabilities). After making the optimal eye movement, the ideal searcher again collects responses in parallel, updates the posterior probabilities, compares them to the criterion, and so on. The cycle repeats until the stopping criterion is exceeded.

The performance of the ideal searcher depends critically upon the retinotopic map of the detectability of the target in the background (visibility map), which we express in terms of a signal-to-noise ratio (d'). Najemnik & Geisler (2005) measured the visibility maps of two observers, for a 6-cpd sine-wave target, as a function of target contrast, eccentricity, and $1/f$ noise contrast. They also found that modest changes in the maps had substantial effects on ideal search performance. Unfortunately, it was not practical to directly measure the visibility maps for the large number of conditions in this study.

However, it was possible to carry out an approximate ideal observer analysis for the 6-cpd conditions using the previously measured maps. This is justified because W.S.G. was an observer in both studies, because the maps were very similar for the two observers in the previous study, and because both observers in this study performed similarly. The predictions for the other conditions are likely to be less accurate. Here, we also generate predictions for the other conditions by extrapolating from the 6-cpd maps, using published facts in the psychophysical literature; however, these predictions must be viewed as more approximate (see Appendix B). Our aim here is not to obtain rigorous measures of efficiency but to get a semiquantitative picture of the search behavior expected from a rational search strategy (see Discussion).

It is important to note that although template matching is the ideal detection mechanism for targets in Gaussian noise, the formal assumption of template matching has little effect on the ideal search performance reported here. The reason is simply that we separate the model of target detection from the search model by estimating the visibility maps for each target and noise contrast from empirical measurements. In other words, the search predictions are largely the same no matter what neural mechanisms give rise to the visibility maps. The only way

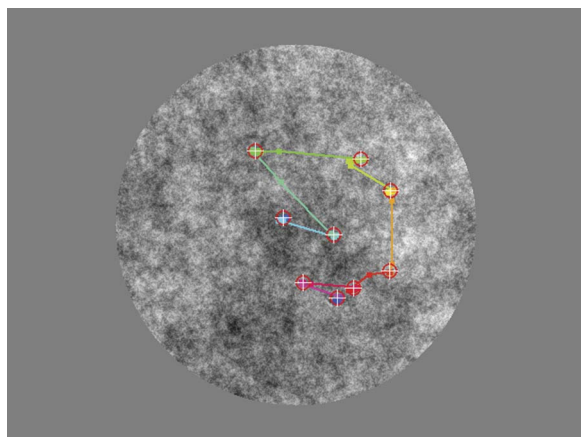


Figure 3. Example fixation sequence in the visual search task. Each plus sign represents a fixation. Time is coded by the color of the scan path (blue = beginning, red = end).

that ideal template matching enters the predictions is in estimating the ratio of external to internal noise level, which has a modest effect on the predictions (see [Appendix B](#)).

Results

Figure 3 shows a typical sequence of eye movements during a search trial. The plus signs show the fixation points, and time is coded by color (blue = beginning, red = end). Note that the scan path is drawn on the unfoveated image; during the actual trial, the display was foveated in a gaze-contingent fashion.

Figure 4 shows the mean search times for correct responses measured from the two observers, where each mean is based on 60 trials. Each panel plots search time as a function of the rate of falloff in display resolution with retinal eccentricity, for one target spatial frequency, at each of the four levels of background contrast. Recall that the rms contrast of the one-octave Gabor target was fixed at 0.35. As can be seen, the data are very systematic; search time increases with target spatial frequency, with background contrast, and with the rate of falloff in display resolution from the point of gaze (lower values of ε_2 correspond to higher rates of falloff in display resolution; see [Figure 1](#)). [Figure 5](#) plots the error rates, which were generally low except for the hardest search conditions (high target spatial frequency, high background contrast, and small display half-resolution).

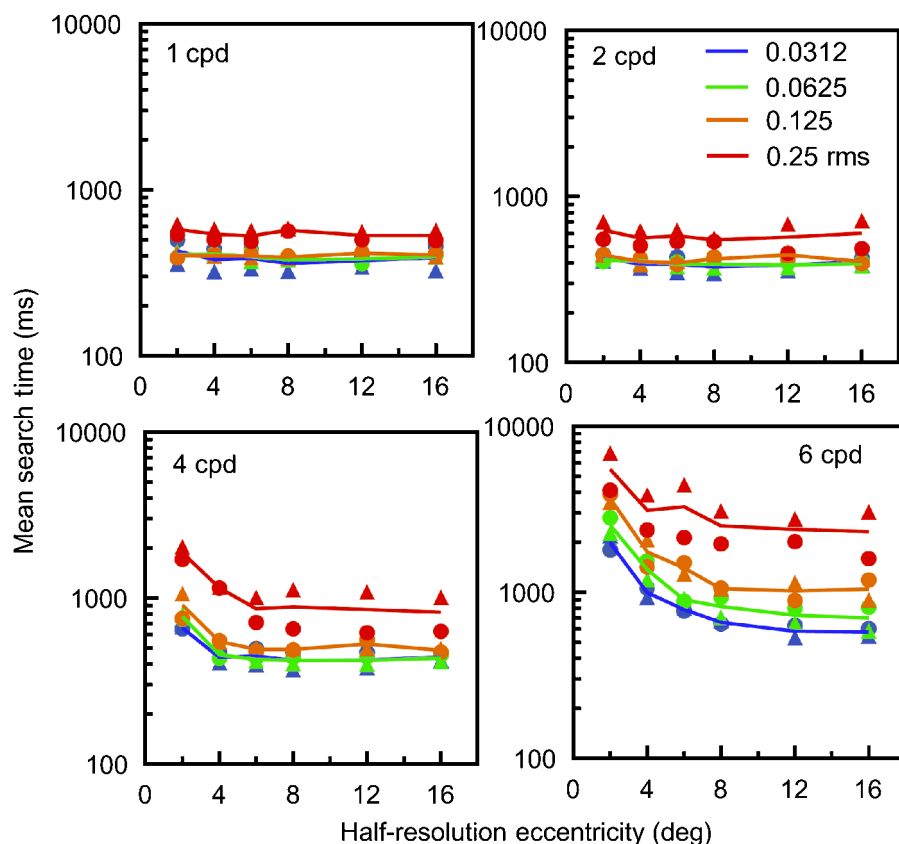


Figure 4. Mean search time for correct responses as a function of the falloff in display resolution, contrast of noise background, and the spatial frequency of the Gabor target. Each point is based upon 60 search trials. Circles, M.E.W.; triangles, W.S.G.; curves, average of the two observers.

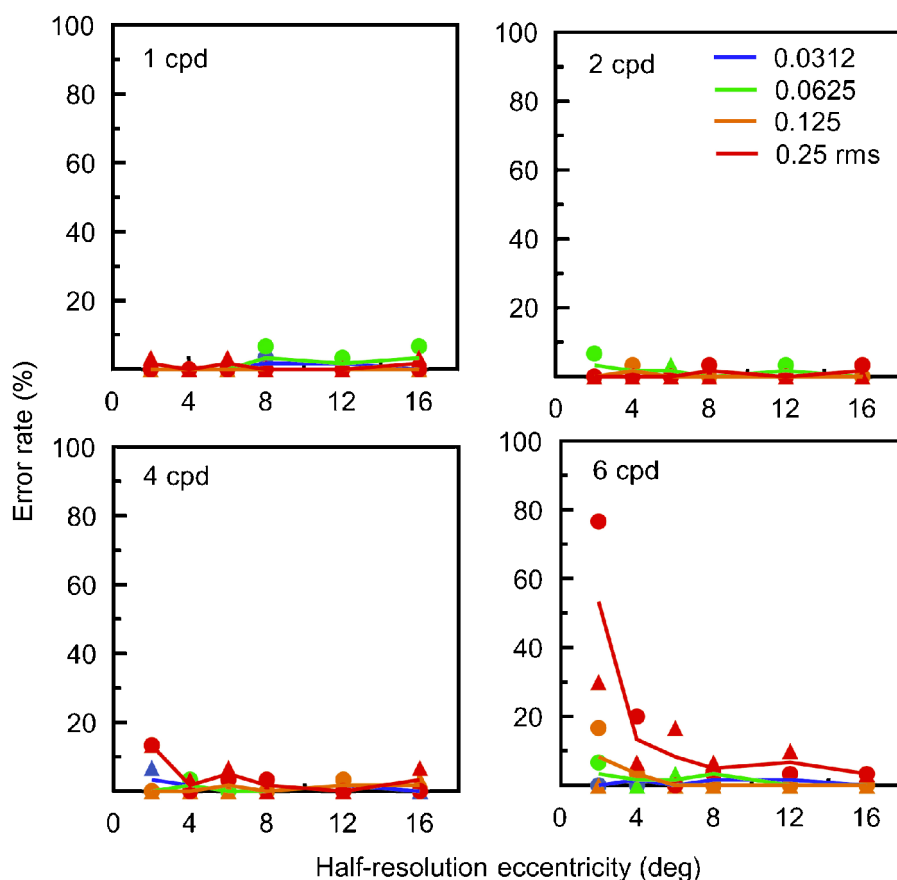


Figure 5. Error rates in search experiment. Circles, M.E.W.; triangles, W.S.G.; curves, average of the two observers. Chance performance is approximately 2.5% correct.

Interestingly, the shape of the function describing search time as a function of the falloff in display resolution is approximately invariant with background contrast. This can be seen more easily in Figure 6, where the curves in each panel of Figure 4 have been scaled to the mean. This was done by scaling each curve (i.e., translating on the log axis) to the mean search time for that spatial frequency.

Examination of the 6-cpd data in Figure 6 shows that half-resolution eccentricities of 6–8 deg produce a reliable increase in search time. These are quite mild attenuations of display resolution and are often subjectively unnoticed when performing the task because of the human visual system's reduced resolution in the periphery. The fact that there is an effect at all suggests a rather subtle use of peripheral information in search.

Search time can vary either because of changes in the number of fixations, changes in the duration of the fixations, or both. Figure 7 plots the number of fixations corresponding to the search times in Figure 4. (In counting fixations, we assumed that button presses occurring less than 125 ms following the onset of the last saccade were initiated before the last saccade, and hence, in these cases, the final fixation was not counted.) The median numbers of fixations vary from 1 to more than 10, and the overall

pattern mirrors that of average search time quite closely. In fact, there is an almost-perfect linear relationship between the average number of fixations and the average search time ($r = .996$ for M.E.W.; $r = .998$ for W.S.G.).

There is also a clear, but weaker, relationship between fixation duration and search time—the greater the search time, the greater the fixation duration ($r = .91$). Thus, search time varies across the various stimulus conditions both because of the number of fixations and because of the duration of the fixations. However, the dominant factor tends to be the number of fixations. This can be seen in Figure 8A, which plots average fixation duration as a function of average number of fixations to find the target, for all 96 stimulus conditions, for both observers. The fixation durations vary by a factor of approximately 1.5 (from about 200 ms to a little more than 300 ms), whereas the number of fixations varies by a factor of approximately 10.

There are several other strong relationships in the eye movement statistics. Figure 8B plots the average distance of the fixation points from the center of the display as a function of the average number of fixations. Given that fixation begins at the center of the display, it is not surprising that the average distance of the fixations from the center increases with number of fixations, but the

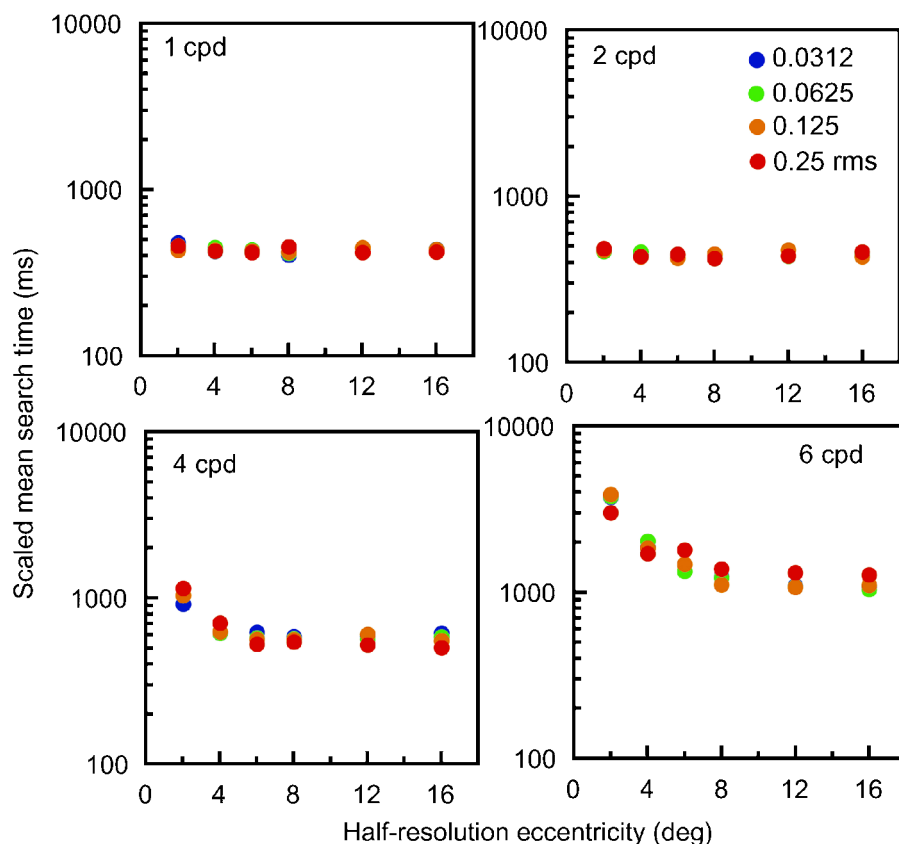


Figure 6. Average search times for correct responses from Figure 4, scaled to compare shapes. Each point is the combined data from the two observers. Each curve within a panel was scaled separately to best fit the average curve (the average curve is not shown).

relationship is very systematic. Figure 8C plots the variance in the number of fixations as a function of the average number of fixations. The variance increases strongly with the number of fixations; the relationship is described approximately by a power function with an exponent of 2–3. Figure 8D plots average saccade length as a function of the average number of fixations. As the number of fixations increases (i.e., as the task gets more difficult), the average saccade length tends to decrease; however, the correlation is relatively weak.

Discussion

There were two major aims of this study. The first was to obtain a parametric picture of human performance in a naturalistic visual search task by recording eye movements and target localization responses while observers searched for Gabor targets in $1/f$ noise (which has the amplitude spectrum typical of natural images). The second aim was to examine the role of the peripheral visual field in visual search by systematically manipulating peripheral spatial information using gaze-contingent display technology.

In general, we obtained very systematic results for the two observers, and although there were some individual differences in performance and in eye movement statistics, one is struck more by the similarities than the differences. This fact and the similarity in search performance for the two observers in Najemnik & Geisler (2005) suggest that the findings reported here are robust.

The finding that search time and number of fixations both increase with spatial frequency, noise contrast, and the falloff in display resolution with eccentricity agrees with common sense expectation and with the existing search literature. However, there are some quantitative properties that are not so intuitive. One clear property is that the shape of the search time functions approximately scales with the power of the $1/f$ noise (i.e., they are approximately parallel on a logarithmic axis; see Figure 6). This result also holds for number of fixations, which is not surprising given the very high correlation (.997) between search time and number of fixations in this study. A second property is that search times (and number of fixations) tend to jump up rather more sharply when the contrast increases from 0.125 to 0.25 than for the other steps in contrast (see Figures 4 and 7). A third property is that for the 1- and 2-cpd targets, increasing the background contrast causes an increase in the number of

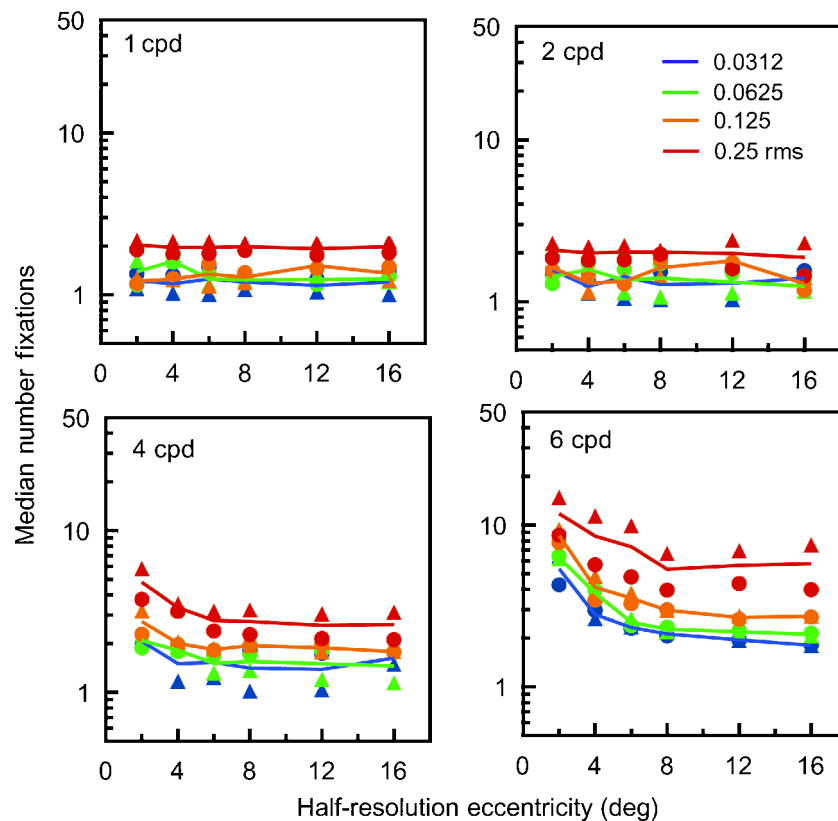


Figure 7. Median number of fixations to find the target for correct trials. Circles, M.E.W.; triangles, W.S.G.; curves, average of the two observers.

fixations, yet varying the rate of falloff in display resolution does not (i.e., search time is flat as a function of ε_2 , yet increases with background noise contrast).

One way to gain insight into these quantitative properties is to consider what would be expected of an ideal searcher that optimally processes the display in parallel on each fixation and that optimally selects successive fixation locations. The solid curves in Figure 9 show the simulated (parameter free) performance of an ideal searcher that is constrained with approximately the same visibility map (detection sensitivity across the retina) as the human observers. The symbols are the average data for the two observers. Like the human search functions, the ideal search functions are roughly parallel on a logarithmic axis and search time jumps up more substantially for the contrast step from 0.125 to 0.25.

Najemnik & Geisler (2005) measured search performance for 6-cpd targets in $1/f$ noise as a function of target and noise contrast, and they found that human performance approaches that of the ideal searcher. As can be seen in Figure 9, the absolute performance level of the human observers in the present experiment also approaches optimal for the conditions with the 6-cpd target. Thus, this study extends the previous finding of a close match between human and ideal search performance to a wider range of conditions.

For the conditions with the 4-cpd target, human performance is somewhat poorer than ideal (although human and ideal roughly parallel each other). Similarly, for the conditions with 1- and 2-cpd targets, human performance is poorer than ideal. We do not show these latter predictions because under these circumstances, the ideal searcher rarely makes more than one fixation. Figure 7 shows that the number of fixations for the human observers increases from approximately 1.0 at the lowest noise contrast (0.03) to 2.0 for the highest noise contrast (0.25). One possible explanation for this is that humans are less efficient at searching for lower frequency targets. Another possibility is that our estimates of the visibility maps for the low-frequency targets are less accurate. As described in the Methods section and in Appendix B, the visibility maps are likely to be fairly accurate for the conditions where the target is 6 cpd because we (Najemnik & Geisler, 2005) directly measured the visibility maps in a very similar experiment for observer W.S.G. The maps for the lower target frequencies are likely to be less accurate because they were extrapolated from the 6-cpd maps. This is a plausible explanation because we have found that relatively small changes in the visibility maps can have substantial effects on search performance (Najemnik & Geisler, 2005). We made some follow-up measurements in the fovea to test this

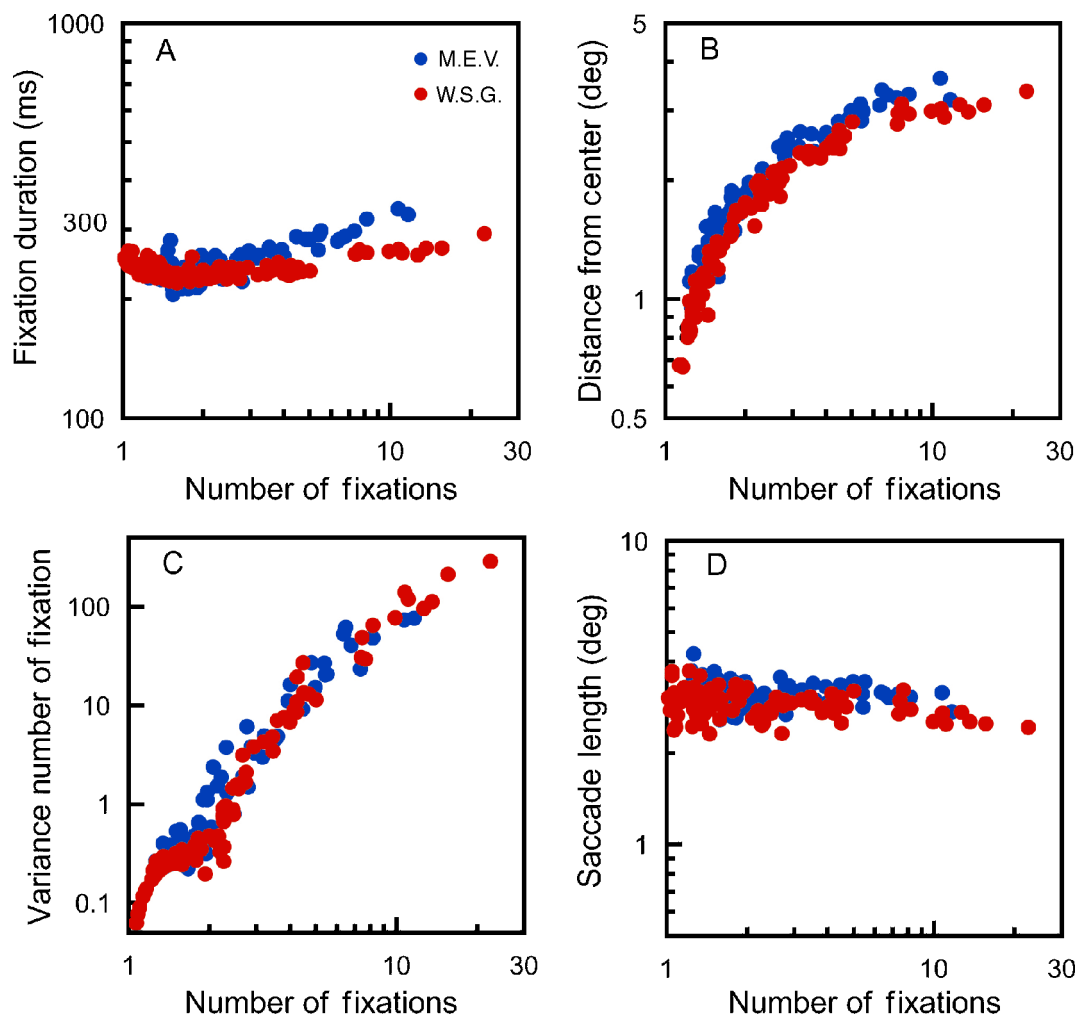


Figure 8. Eye movement statistics for the two observers. (A) Relationship between average fixation duration and average number of fixations to find the target (average correlation = .91). (B) Relationship between average fixation distance from the center of the display and average number of fixations (average correlation = .93). (C) Relationship between the variance and the average number of fixations (average correlation = .96). (D) Relationship between average saccade length and average number of fixations (average correlation = -.21). (Note that all the axes are logarithmic and that the correlations are for log values.)

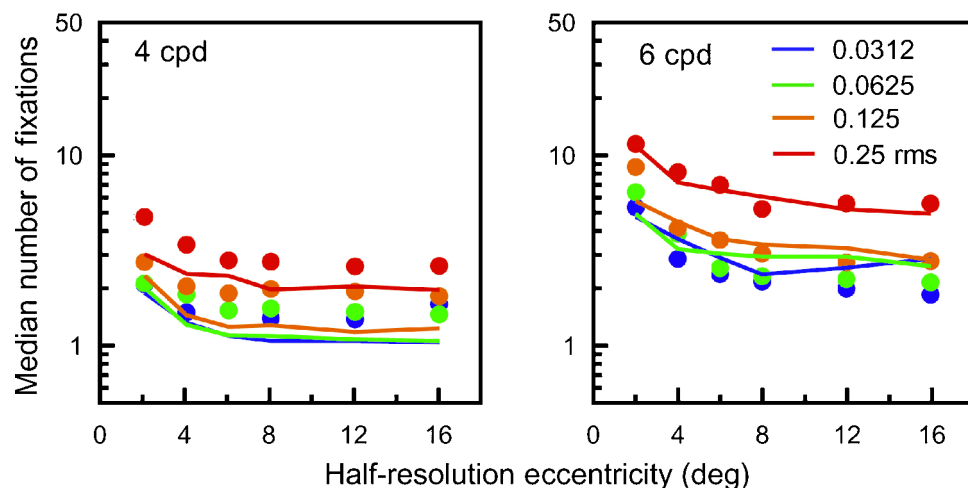


Figure 9. Comparison of human and ideal search performance. Circles, W.S.G. and M.E.W. combined. Curves, ideal searcher.

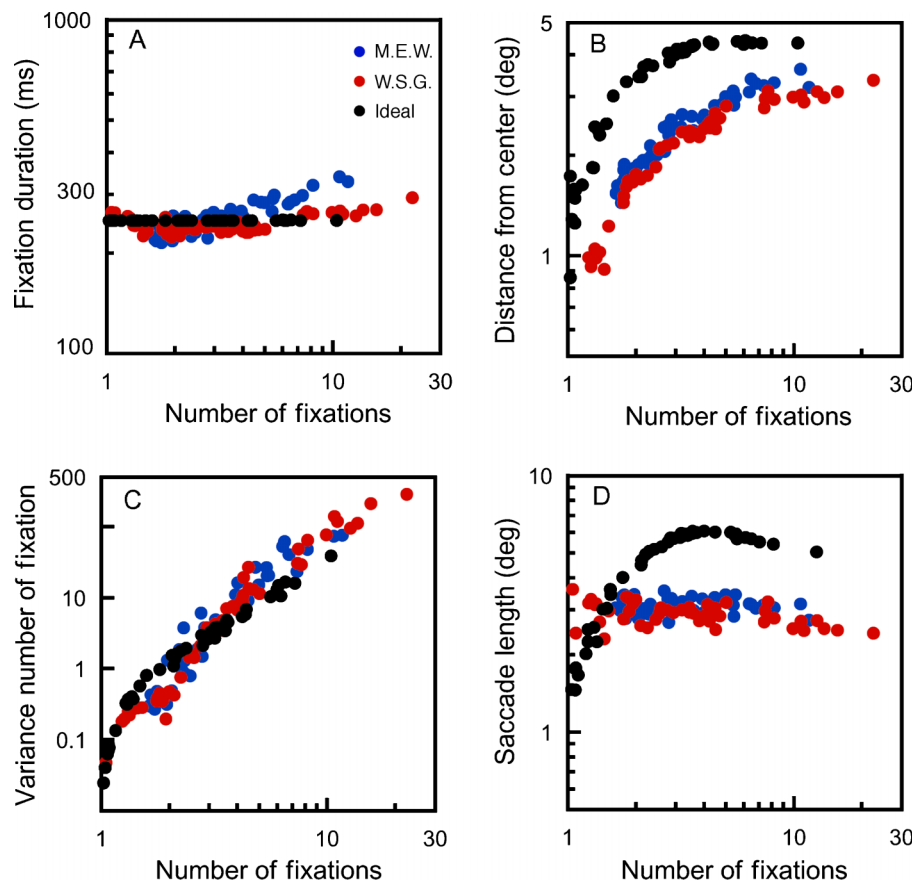


Figure 10. Comparison of human and ideal eye movement statistics for target spatial frequencies of 4 and 6 cpd.

hypothesis and found that, indeed, our extrapolated maps for the lower frequencies underestimate the effect of noise masking (Najemnik & Geisler, unpublished observations).

Figure 10 compares the eye movement statistics of the ideal searcher with those of the human observers. As expected, the ideal searcher does not display an increase in fixation length with increasing task difficulty, as indexed by the average number of fixations to find the target (see Figure 10A). The only mechanism the ideal searcher has for increasing fixation duration is to hold fixation for more than one fixation interval (the fixation interval is set to the average human fixation duration); it does this only very rarely because there is generally more information to gain by moving the eyes to a new location. Figure 10B shows that the average fixation distance from the center of the display increases in a similar fashion for human and ideal searchers, as a function of the average number of fixations. For all conditions, humans tend to fixate on average closer to the center of the display than optimal, although the modal distance for human and ideal is approximately the same (Najemnik & Geisler, 2005). The variance in the number of fixations increases similarly for human and ideal searchers (see Figure 10C). Note that random serial search with replacement predicts similar

behavior (a geometric distribution for number of fixations), although the random searcher's performance is much poorer than human or ideal. For the ideal searcher, mean saccade length increases rapidly and then decreases gradually as a function of number of fixations, whereas for human searchers, it decreases gradually (Figure 10D). The rapid rise in saccade length for the ideal is presumably due to two factors: (1) the first saccade is from the middle of the display and, hence, is more restricted in length, and (2) at least for the first saccade, the larger visibility maps in the easier conditions tend to push the fixation location where the maximum information is gained toward the center of the circular search region.

Figures 9 and 10 show that human searchers are similar in many ways to an ideal searcher. They are less efficient than ideal for some of the easiest search conditions, but overall, they perform very well. To perform so well, human searchers must perform efficient parallel processing across the search area on each fixation, they must select fixation locations with high efficiency, and they must have inhibition of return (Najemnik & Geisler, 2005).

The average distance of fixations from the center of the display and the average saccade length of human searchers differ somewhat from those of the ideal

searcher. However, the fact that human and ideal search times are similar suggests that these non-ideal eye movement behaviors are not critical for efficient search.

At first thought, it is rather surprising that humans perform so well in extended visual search tasks, given that they have relatively poor memory for image details and poor ability to integrate image information across fixations (Hayhoe, Bensinger, & Ballard, 1998; Irwin, 1991; Rensink, 2002). However, by analyzing suboptimal searchers, we have shown that memory for image details and ability to integrate across fixations add relatively little to search performance. More important is having a memory system sufficient to support inhibition of return. In other words, it is not necessary to remember “what I’ve seen,” but it is essential to remember “where I’ve been.”

Figure 6 shows that when the target was 6 cpd, even a very mild level of display foveation (e.g., $\varepsilon_2 = 6$ deg) was sufficient to cause a reliable increase in search time. This level of display foveation was subjectively invisible for medium and low noise contrasts. This is not surprising given that the gaze-contingent displays were free of artifacts and that the half-resolution eccentricity for the human visual system is approximately 2.3 deg. The fact that accentuating the human falloff in resolution by a small amount causes a significant drop in performance confirms the conclusion from the ideal-observer analysis that peripheral information is being used efficiently in guiding eye movements. This is quite different from reading tasks (McConkie & Rayner, 1975), where peripheral information plays little role, but is qualitatively consistent with other analyses of peripheral information use in search (Eckstein et al., 2001; Rajashekar et al., 2002).

Our subjects’ impression that a gaze-contingent display ε_2 of 6 deg often produced an undetectable level of blur is consistent with the recent blur-detection experiments of Loschky, McConkie, Yang, & Miller (2005), who report that blur is undetectable when $\varepsilon_2 = 6$ deg and only detectable 5% of the time when $\varepsilon_2 = 3$ deg. However, their study used a divided attention task, which may have underestimated sensitivity to blur. Nonetheless, their results in conjunction with ours suggest that even when blur goes unnoticed in the periphery, it can affect sensitivity for detecting peripheral targets and, hence, affect search performance.

We have demonstrated that it is possible to generate clean gaze-contingent displays on conventional PCs using conventional software (C++ or MatLab) and that this technology holds considerable promise for rigorously analyzing the role of peripheral information in complex extended tasks that involve eye movements.

The rather surprising finding of this study and of Najemnik & Geisler (2005) is that humans are very efficient at visual search in complex naturalistic backgrounds and that many human eye movement statistics parallel those of the ideal searcher. But how general are

these results and how representative are they of natural search in the natural environment?

In our paradigm, the target was always present somewhere in the display. Thus, the task is similar to searching the ground for a dropped object or searching a stand of foliage for an animal whose presence is known from, say, a sound. However, in many natural search tasks, there is uncertainty about whether the sought-after object is present at all. This adds a level of complexity not addressed in the current task, namely, deciding when to give up the search and conclude the target is absent. Ideal search predictions for this case can be derived within our theory simply by setting the visibility (d') of an imaginary target location to zero (or near zero) and giving this location a prior probability corresponding to the probability of a target absent trial. (Note that the posterior probability of the “target absent” location climbs during the search as other locations are ruled out.) It remains to be seen how efficient humans are in this task; however, because of the greater memory demands in target absent trials, humans may be less efficient, especially under conditions where the target is relatively difficult to detect.

Our search paradigm and ideal searcher theory are restricted in other ways. For example, the search region consists of just one kind of texture (albeit one that is similar in complexity and amplitude spectrum to regions of natural images) and there is only one search target of one type. We are currently working on extending the ideal searcher theory to allow multiple texture regions and possible targets. These extensions together with gaze-contingent display technology may lead to a better understanding of the processing requirements in more complex and natural search tasks and may help to identify the search strategies/mechanisms humans use under these circumstances.

Appendix A

The algorithm used to compute fixation points from eye positions was a modified version of one used by the Applied Science Laboratories Series 5000 data-analysis software. The input to the algorithm is a list of eye position samples, $P[1], \dots, P[N]$, and the output is a list of fixation points, F . Initially, list F and a working list T are set to be empty, and an index i is set to 0. Also, we note that P' is a temporary list that accumulates all the eye positions corresponding to a given single fixation. The algorithm proceeds as follows:

1. If i exceeds N , the algorithm ends. Otherwise, compute the centroid C of eye positions $P[i], P[i + 1], \dots, P[i + k]$, such that $P[i + k]$ is the last eye position that occurs less than 75 ms after $P[i]$. If there is not 75 ms worth of eye positions following $P[i]$, the algorithm ends.

2. If the standard deviation of the distances of $P[i]$, $P[i + 1]$, ..., $P[i + k]$ from C is greater than a degrees, $P[i]$ is not the beginning of a fixation; thus, ignore this eye position by setting $i = i + 1$; continue at 1.
3. If the standard deviation of the distances of $P[i]$, $P[i + 1]$, ..., $P[i + k]$ from C is less than a degrees, $P[i]$ is the beginning of a fixation; hence, save these $k + 1$ eye positions in a list P' , and set $i = i + k + 1$.
4. If the distance from $P[i]$ to C is less than b degrees, add $P[i]$ to P' . Set $i = i + 1$. If i does not exceed N , continue at 4; otherwise, compute the mean of the eye positions in P' and add it to a list of fixations F ; the algorithm ends.
5. If the distance from $P[i]$ to C is greater than c degrees, compute the mean of the eye positions in P' and add it to a list of fixations F ; continue at 1.
6. If the distance from $P[i]$ to C is less than c degrees and greater than b degrees, clear the list T of potential fixation eye positions.
7. If i exceeds N , compute the mean of the eye positions in P' and add it to a list of fixations F ; the algorithm ends. Otherwise, add $P[i]$ to the list T . Set $i = i + 1$. If the time difference between the first and last eye positions in T is less than 50 ms, continue at 7.
8. Compute the mean of the eye positions in T . If this mean is less than b degrees from C , add this mean eye position to P' and clear the list T ; continue at 4. Otherwise, compute the mean of the eye positions in P' and add it to the list of fixations F ; continue at 1.

In this study, we set $a = 0.1$ deg, $b = 0.2$ deg, and $c = 0.3$ deg.

Appendix B

In a simple detection task with no uncertainty and a sinusoidal (Gabor) target, an ideal observer computes the cross correlation of a matched Gabor template with the stimulus at the known target location and responds that the target is present, if the cross correlation exceeds a criterion. Under fairly general conditions, the detection performance of this ideal observer, when limited by external and internal noise, can be described by the signal to noise ratio (d'):

$$d'(c, e_n, \varepsilon)^2 = \frac{c^2}{\alpha e_n + \beta(c, e_n, \varepsilon)}, \quad (\text{B1})$$

where c is the rms contrast of the target (i.e., c^2 is the contrast power), e_n is the stimulus noise contrast power, and ε is the eccentricity in degrees (see Supplement to Najemnik & Geisler, 2005). (We note that d' is monotonically related to the proportion of correct responses in the detection task.)

The value of the constant α is a function of the narrow band of background noise that affects the responses of the target-matched template, and it can be estimated by measuring the mean and the variance of the template responses to a large number of samples of the target embedded in the actual background noise used in the experiments. The value of α is approximately 0.022.

In the gaze-contingent display, the transfer function at each eccentricity will attenuate both the target and the relevant narrow band of the noise background by an equal factor, and thus,

$$d'(c, e_n, \varepsilon, f)^2 = \frac{|T(f, \varepsilon)|^2 c^2}{\alpha |T(f, \varepsilon)|^2 e_n + \beta(c, e_n, \varepsilon)}$$

or

$$d'(c, e_n, \varepsilon, f)^2 = \frac{c^2}{\alpha e_n + \beta(c, e_n, \varepsilon)/|T(f, \varepsilon)|^2}, \quad (\text{B2})$$

where $T(f, \varepsilon)$ is given by substituting Equation 3 into Equation 2.

Consistent with Equation B1 and the literature on noise masking (e.g., Burgess et al., 1981; Pelli & Farell, 1999), Najemnik and Geisler (2005) found that the slope and intercept of the noise masking functions measured at different eccentricities for a 6-cpd target are described by a linear equation of the form:

$$c_T^2(e_n, \varepsilon) = a(\varepsilon)e_n + b(\varepsilon), \quad (\text{B3})$$

where c_T is the contrast threshold for the detection. Furthermore, we found that the slopes and intercepts of these linear functions vary systematically with eccentricity:

$$\ln[a(\varepsilon)] = m_a \varepsilon + b_a \quad (\text{B4})$$

and

$$\ln[b(\varepsilon)] = m_b \varepsilon + b_b. \quad (\text{B5})$$

Finally, we found that the steepness parameter of the psychometric functions (Weibull functions) varied systematically with eccentricity:

$$s(\varepsilon) = \frac{2.8\varepsilon}{\varepsilon + 0.8} + 2. \quad (\text{B6})$$

To generalize these results to the current experiment, we first make use of the fact that the logarithm of the

masking function intercept is linear with eccentricity (Equation B5). This fact is consistent with a well-known formula for the human contrast sensitivity function. In the absence of a noise background, the human CSF for brief stimulus presentations is well approximated by

$$c_s = c_s(0) \exp\left(-kf \frac{e_2 + \varepsilon}{e_2}\right), \quad (\text{B7})$$

where $c_s(0)$ is the contrast sensitivity in the fovea, k is a constant whose value is typically in the neighborhood of 0.1, f is spatial frequency in cpd, and e_2 is the half-resolution eccentricity, which is typically around 2.0–2.5 deg (see, e.g., Geisler & Perry, 1998). Writing Equation B7 in terms of the threshold contrast power, we have

$$c_T^2 = b_0^2 \exp\left(2k_b f \frac{e_2 + \varepsilon}{e_2}\right), \quad (\text{B8})$$

where b_0 is the threshold contrast in the fovea at zero spatial frequency. Therefore, we predict from Equation B8 that $\ln[b(\varepsilon)]$ should be a linear function of eccentricity:

$$\ln[b(\varepsilon)] = \frac{2k_b f}{e_2} \varepsilon + 2k_b f + 2\ln[b_0]. \quad (\text{B9})$$

Setting e_2 to a representative value from the literature ($e_2 = 2.3$ deg), we can estimate k_b and b_0 by setting Equation B9 equal to Equation B5: $k_b = 0.168$, $b_0 = 0.0295$. To generate estimates of the slopes of the masking functions for other spatial frequency targets, we assume that an equation similar to Equation B9 holds for slopes:

$$\ln[a(\varepsilon)] = \frac{2k_a f}{e_2} \varepsilon + 2k_a f + 2\ln[a_0]. \quad (\text{B10})$$

Setting Equation B10 equal to Equation B4, we find that $k_a = 0.088$ and $a_0 = 0.44$.

Equations B9 and B10 are guaranteed to be consistent with the masking functions we have measured for 6-cpd targets, and they can be used to provide estimates of the intercepts and slopes for other target spatial frequencies. Finally, we assume that the steepness parameters of the psychometric functions for all spatial frequencies are the same as they are for 6-cpd targets (Equation B6). Obviously, there is potential for error here, and thus, the ideal searcher predictions shown here are most trustworthy for the 6-cpd target.

To generate estimates of the visibility maps, we use the estimated masking functions and psychometric function slopes to directly compute values of d' for each point in the display relative to the current point of fixation. We

substitute these d' values into Equation B1 to estimate the equivalent internal noise power $\beta(c, e_n, \varepsilon)$ in the unfoveated display. The equivalent internal noise power in the foveated display is obtained by dividing $\beta(c, e_n, \varepsilon)$ by the square of the foveated transfer function (see Equations 2, 3, and B2). The equivalent external noise power is obtained by multiplying the contrast noise power by the estimated constant α . In simulating ideal search performance, the external noise was taken to be static noise and the internal noise was taken to be dynamic noise that was statistically independent in space and time. For more details, see Najemnik and Geisler (2005).

Acknowledgments

This research was supported by National Institutes of Health Grant EY-02688.

Commercial relationships: none.

Corresponding author: Wilson S. Geisler.

Email: geisler@psy.utexas.edu.

Address: Center for Perceptual Systems, 1 University Station, University of Texas, Austin, TX 78712, USA.

References

- Beard, B. L., & Ahumada, A. J. (1998). A technique to extract relevant image features for visual tasks. *SPIE Proceedings: Human Vision and Electronic Imaging III*, 3299, 79–85.
- Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science*, 214, 93–94. [PubMed]
- Burton, G. J., & Moorehead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, 26, 157–170.
- Duchowski, A. T., Cournia, N., & Murphy, H. (2004). Gaze-contingent displays: A review. *Cyberpsychology & Behavior*, 7, 621–634. [PubMed]
- Eckstein, M. P., Beutter, B. R., & Stone, L. S. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception*, 30, 1389–1401. [PubMed]
- Engel, F. L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research*, 17, 95–108. [PubMed]
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and Image Science*, 4, 2379–2394. [PubMed]

- Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research*, 37, 617–631. [PubMed]
- Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision*. Oxford: Oxford University Press.
- Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review*, 102, 356–378. [PubMed]
- Geisler, W. S., & Perry, J. S. (1998). A real-time foveated multi-resolution system for low-bandwidth video communication. *SPIE Proceedings*, 3299, 294–305.
- Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research*, 38, 125–137. [PubMed]
- He, P., & Kowler, E. (1989). The role of location probability in the programming of saccades: Implications for “center-of-gravity” tendencies. *Vision Research*, 29, 1165–1181. [PubMed]
- Hooge, I. T. C., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research*, 39, 1567–1575. [PubMed]
- Irwin, D. E. (1991). Information integration across saccadic eye movement. *Cognitive Psychology*, 23, 420–456. [PubMed]
- Jacobs, A. M., & O’Regan, J. K. (1987). Spatial and/or temporal adjustments of scanning behavior to visibility changes. *Acta Psychologica*, 65, 133–146. [PubMed]
- Juday, R. D., & Fisher, T. E. (1989). Geometric transformations for video compression and human teleoperator display. *SPIE Proceedings: Optical Pattern Recognition*, 1053, 116–123.
- Kortum, P. T., & Geisler, W. S. (1996a). Implementation of a foveated image-coding system for bandwidth reduction of video images. *SPIE Proceedings*, 2657, 350–360.
- Kortum, P. T., & Geisler, W. S. (1996b). Search performance in natural scenes: The role of peripheral vision. *Investigative Ophthalmology & Visual Science*, 37/3(Suppl.), S297.
- 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, 99–117. [PubMed]
- Loschky, L. C., & McConkie, G. W. (2005). How late can you update? Detecting blur and transients in gaze-contingent multi-resolution displays. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 1527–1530.
- Loschky, L. C., McConkie, G. W., Yang, J., & Miller, M. E. (2005). The limits of visual resolution in natural scene viewing. *Visual Cognition*, 12, 1057–1092.
- Lu, Z. L., & Doshier, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 16, 764–778. [PubMed]
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578–586.
- Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research*, 38, 1805–1815. [PubMed]
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434, 387–391. [PubMed]
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40, 1227–1268. [PubMed]
- Parkhurst, D. J., & Niebur, E. (2002). Variable-resolution displays: A theoretical, practical, and behavioral evaluation. *Human Factors*, 44, 611–629. [PubMed]
- Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America A, Optics, Image Science and Vision*, 16, 647–653. [PubMed]
- Perry, J. S., & Geisler, W. S. (2002). Gaze-contingent real-time simulation of arbitrary visual fields. *SPIE Proceedings*, 4662, 57–69.
- Rajashekar, U., Cormack, L. K., & Bovik, A. C. (2002). Visual search: Structure from noise. In A. T. Duchowski (Ed.), *Eye tracking research & applications* (pp. 119–123). New Orleans: ACM SIGGRAPH.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. [PubMed]
- Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multi-resolutional displays: An integrative review. *Human Factors*, 45, 307–328. [PubMed]
- Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, 53, 245–277. [PubMed]
- Riggs, L. A., Ratliff, F., Cornsweet, J. C., & Cornsweet, T. N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America*, 43, 495–501. [PubMed]
- Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, 31, 523–535. [PubMed] [Article]

- Warner, H. D., Serfoss, G. L., & Hubbard, D. C. (1993). *Effects of area-of-interest display characteristics on visual search performance and head movements in simulated low-level flight* (No. AL-TR-1993-0023). Armstrong Laboratory, Human Resources Directorate, Aircrew Training Division, Williams AFB, AZ.
- Weiman, C. F. R. (1990). Video compression via log polar mapping. *SPIE Proceedings: Real Time Image Processing II*, 1295, 266–277.
- Wilson, H. R., Levi, D., Maffei, L., Rovamo, J., & DeValois, R. L. (1990). The perception of form: Retina to striate cortex. In L. Spillmann & J. S. Werner (Eds.), *Visual perception: The Neurophysiological Foundations* (pp. 231–272). San Diego: Academic Press.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–74). East Sussex, UK: Psychological Press.
- Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.
- Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, 36, 2177–2187. [[PubMed](#)]
- Zelinsky, G. J., Rao, R. P., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, 8, 448–453.