

The limits of visual resolution in natural scene viewing

Lester C. Loschky

Kansas State University, Manhattan, KS, USA

George W. McConkie

University of Illinois at Urbana-Champaign, IL, USA

Jian Yang and Michael E. Miller

Eastman Kodak Company, Rochester, NY, USA

We examined the limits of visual resolution in natural scene viewing, using a gaze-contingent multiresolutional display having a gaze-centred area-of-interest and decreasing resolution with eccentricity. Twelve participants viewed high-resolution scenes in which gaze-contingent multiresolutional versions occasionally appeared for single fixations. Both detection of image degradation (five filtering levels plus a no-area-of-interest control) in the gaze-contingent multiresolutional display, and eye fixation durations, were well predicted by a model of eccentricity-dependent contrast sensitivity. The results also illuminate the time course of detecting image filtering. Detection did not occur for fixations below 100 ms, and reached asymptote for fixations above 200 ms. Detectable filtering lengthened fixation durations by 160 ms, and interference from an imminent manual response occurred by 400–450 ms, often lengthening the next fixation. We provide an estimate of the limits of visual resolution in natural scene viewing useful for theories of scene perception, and help bridge the literature on spatial vision and eye movement control.

SCENE PERCEPTION AND RESOLUTION ACROSS THE VISUAL FIELD

Although the eyes tend to be directed toward objects of current interest, parafoveal and peripheral visual information is generally believed to play several important roles in the free viewing of natural scenes: Aiding in perceiving the gist of the image, perceiving the layout of the image, recognizing objects, and helping select the object to which the eyes are sent next. It is, therefore, important to understand the nature of information available

Please address all correspondence to: Lester Loschky, Assistant Professor, Department of Psychology, 471 Bluemont Hall, 1101 Mid-Campus Drive, Manhattan, KS 66506-5302, USA.
Email: loschky@ksu.edu

from extrafoveal regions and its absolute limits (cf. similar ideas in reading Rayner, 1998).

Parafoveal and peripheral information appears to exert a large influence on the gist viewers acquire from a scene. Scene gist appears to be acquired from information across the entire image, not only in the foveal region (Boyce & Pollatsek, 1992; Boyce, Pollatsek, & Rayner, 1989; Li, VanRullen, Koch, & Perona, 2002; Schyns & Oliva, 1994). Thus, the nature of the information used to acquire gist becomes important. Some evidence suggests that the spatial frequency information for gist may be mostly low frequency, such as oriented blobs, but occasionally, it is also higher frequency, such as outlines and edges (Oliva & Schyns, 2000; Schyns & Oliva, 1994). Low frequency information may be especially useful for conveying the layout of a scene, which may be important not only for acquiring its gist, but also for understanding the spatial relations among its constituent objects (Sanocki, 2003; Sanocki & Epstein, 1997). Yet, so far, no accurate characterization of the spatial frequency information available from each part of the visual field during extended free viewing of natural scenes has been provided.

Parafoveal and peripheral vision is also useful for object recognition (Henderson, Pollatsek, & Rayner, 1987, 1989; Henderson, McClure, Pierce, & Schrock, 1997). This raises the question of how far into the visual periphery objects can be recognized. In fact, no simple distance measure is likely to provide an adequate answer since object recognition in the parafovea and periphery varies with task difficulty, image size, contrast, and eccentricity. In demanding tasks, such as using long-term memory to detect a change between two scenes involving a single critical object, good task performance is only possible when the eyes fixate within roughly 2° of the critical object (Nelson & Loftus, 1980; see also Henderson, Williams, Castelano, & Falk, 2003). However, in simple tasks, such as determining whether a large natural scene contains an animal or not, performance is quite good at even 50° eccentricity (Thorpe, Gegenfurtner, Fabre-Thorpe, & Buelthoff, 2001). Importantly, much research has shown that peripheral visual performance can be equated with the fovea by scaling the size and contrast of stimuli such as gratings, and also natural objects such as faces (Maekelae, Naesaenen, Rovamo, & Melmoth, 2001). Since recent studies have proposed an ideal spatial frequency range for foveal object recognition (Braje, Tjan, & Legge, 1995; Gold, Bennett, & Sekuler, 1999; Olds & Engel, 1998), the limits of visual resolution across the visual field must critically influence peripheral object recognition.

Parafoveal and peripheral information is essential in guiding the eyes during free viewing. As the eyes explore a scene, they move several times per second, with attention guiding the eyes to different scene regions (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Irwin & Gordon, 1998; Kowler, Anderson, Doshier, & Blaser, 1995; for review see Henderson & Hollingworth, 1998). Attentional selection involves preattentively processing parafoveal and

peripheral information (Levin, Takarae, Miner, & Keil, 2001; Treisman & Gelade, 1980; Wolfe, 1994). But what information is available from parafoveal and peripheral regions to guide attention? A few studies have varied the information content in scenes available on each eye fixation, in ways that can be related to variation in visual sensitivity across the visual field, and shown effects on eye guidance (Loschky & McConkie, 2002; Shioiri & Ikeda, 1989; van Diepen & Wampers, 1998). However, none of these studies varied image content to test models of visual sensitivity across the visual field. Thus, the relationship between the basic limits of parafoveal and peripheral sensitivity and eye guidance in scene viewing is still largely unknown.

All of the above-mentioned aspects of scene perception require parafoveal and peripheral visual information, sensitivity to which varies with retinal eccentricity. Thus, understanding these processes in scene perception first requires understanding the limits of visual resolution across the visual field. While there is a long history of psychophysical research on the sensitivity of the visual system as a function of distance from the centre of vision (for reviews, see Levi, 1999; Wilson, Levi, Maffei, Rovamo, & DeValois, 1990), only recently has such research been extended to natural scene stimuli and free viewing.

VARIABLE RESOLUTION OF THE VISUAL SYSTEM AND IMAGE FILTERING

At present, the most theoretically motivated measure of visual resolution is in terms of contrast sensitivity (Campbell, 1983). The resolution of any visual stimulus, from the simple alternating bars of a grating to a complex natural scene, is analysable in terms of its spatial frequency content, with fine detail (e.g., edges) comprised of high spatial frequencies, and coarse information (e.g., blobs) comprised of lower spatial frequencies. The range of visible spatial frequencies, measured by sensitivity to contrast at each frequency, is described by a contrast sensitivity function (CSF; Campbell & Robson, 1968; DeValois & DeValois, 1988), which varies as a function of distance from the fovea, or *retinal eccentricity* (Anderson, Mullen, & Hess, 1991; Banks, Sekuler, & Anderson, 1991; Cannon, 1985; Peli, Yang, & Goldstein, 1991; Pointer & Hess, 1989; Robson & Graham, 1981). Contrast thresholds increase as a function of retinal eccentricity much faster for higher spatial frequencies than for lower spatial frequencies.

A model of eccentricity-dependent contrast sensitivity

Based on much previous research using grating stimuli in simple detection or discrimination tasks, it is possible to mathematically model the drop-off of visual sensitivity across the visual field. Peli et al. (1991) described the contrast

threshold for detecting a grating patch of spatial frequency f at an eccentricity E as:

$$C_t(r, f) = C_t(0, f) \exp(kfE) \quad (1)$$

where $C_t(0, f)$ is the contrast threshold at the fovea, and k determines how fast sensitivity drops off with eccentricity. Peli et al. (1991) fit this equation to data from six experiments investigating the contrast thresholds of monochromatic gratings. These fits showed that the k value ranged from 0.030 to 0.057. In Equation 1, contrast thresholds increase rapidly with eccentricity for high spatial frequencies, suggesting that this information is only useful in central vision. Since the foveal contrast threshold was not defined in Equation 1, Yang et al. (2001) used an equation to estimate it based on previous research (Yang, Qi, & Makous, 1995; Yang & Stevenson, 1998):

$$C_t(0, f) = [N + \eta^2/(f^2 + \sigma^2)] \exp(\alpha f) \quad (2)$$

where, N , η , σ , and α are parameters (for more detail, see Yang et al., 1995). Using Equations 1 and 2, we can estimate the limits of resolution across the visual field in terms of the *spatial frequency cutoff* for perception at any given eccentricity. For a set of nominal k , N , and α values being 0.045, 0.0024, and 0.14° , respectively, the cutoff spatial frequency f_c can be approximated as:

$$f_c = 43.1 * E_2/(E_2 + E) \quad (3)$$

with $E_2 = 3.11^\circ$. E_2 is the retinal eccentricity at which the spatial frequency cutoff drops to half its foveal maximum (from 43.1 cpd to 21.55 cpd). (In other psychophysical studies, E_2 has often been defined as the retinal eccentricity at which foveal thresholds double, e.g., Beard, Levi, & Klein, 1997). Note that the cutoff frequency is based on the assumption of 100% contrast at each frequency. Thus, at lower contrast (e.g., in natural images) the cutoffs would drop to lower spatial frequencies.

Testing the limits of visual resolution using static multiresolutional displays

A key question is whether the above model scales up to natural images and viewing conditions. To test the model with natural images, we filter a scene's image resolution to match the resolution of the human visual system at each retinal eccentricity. If image filtering removes, at each eccentricity, only imperceptible information (e.g., frequencies higher than the threshold), the image should produce a normal perceptual experience for a viewer. On the other hand, if filtering removes perceptible information (e.g., frequencies lower than the threshold) at each eccentricity, viewers may perceive image degradation.

Several recent studies using briefly flashed (150–300 ms) multiresolutional photographic images have tested the limits of visual resolution in natural scenes (Peli & Geri, 2001; Sere, Marendaz, & Herault, 2000; Yang et al., 2001). Peli and Geri (2001) and Yang et al. (2001) tested models of eccentricity-dependent contrast sensitivity by varying parameter values producing different filtering frequency cutoffs, and then measuring viewers' perceived image degradation. In both studies, the effects of the resolution drop-off parameters on perceived image quality were both robust and close to predictions. Yet, in both studies, viewers were somewhat less sensitive to image filtering than predicted based on previous studies using isolated gratings or Gabor patches (see also Sere et al., 2000). Peli and Geri explain the latter unexpected finding in terms of masking by natural scene detail.

Importantly, because these studies used briefly flashed images, their results may not apply to dynamic, multifixation, free-viewing conditions. Attention focus may begin broad and become narrow during extended scene viewing, as suggested by viewers' initial long saccades (eye movements) with short fixations (still periods) and, later, short saccades with long fixations (Antes, 1974). Since attention focus affects visual resolution (Yeshurun & Carrasco, 1999), a rigorous test of the limits of visual resolution across the visual field in scene viewing should allow free eye movements.

Testing the limits of visual resolution using gaze-contingent multiresolutional displays

We propose to find the limits of visual resolution in dynamic natural scene viewing with a gaze-contingent display-change method (e.g., McConkie & Rayner, 1975; Rayner, 1975) in which image resolution varies with distance from the point of gaze, and the centre of highest resolution is aligned with gaze. We will use this gaze-contingent multiresolutional display (Parkhurst & Niebur, 2002; Reingold, Loschky, McConkie, & Stampe, 2003) to test predictions of the model of eccentricity-dependent contrast sensitivity (Yang et al., 2001) described in Equations 1–3, by measuring the perception of image degradation. The model predicts that image filtering matching the high frequency cutoff function in Equation 3 will be the degradation detection threshold; less filtering should be undetectably degraded, and more filtering should be more detectably degraded. We will test these predictions by varying the E_2 parameter in Equation 3 to produce a range of filtering levels, and finding the degradation threshold E_2 value. Besides testing the model's predicted threshold, the frequency cutoff function of the empirical threshold filtering level will provide an estimate of the limits of visual resolution in dynamic natural scene viewing.

To date, we know of one published study that has attempted to test the limits of visual resolution across the visual field, using natural images, under dynamic free-viewing conditions (Sere et al., 2000, Exp. 2). However, that experiment's

only dependent measure was participants' reported "subjective distinctness of the image" at the end of each trial (p. 1410), and its only reported results were (1) that "13 out of 14 observers reported a subjective visual change of the image," when filtered and unfiltered images were presented on alternating fixations, although (2) "none of the observers reported that the images were filtered" (p. 1410). Thus, there is clearly room for more investigation of this question.

The above study raises a critical methodological question. What is a good way to measure viewers' detection of image degradation when using a gaze-contingent multiresolutional display? If the point of highest resolution is updated on *every* fixation, and one asks viewers to report detection of image filtering only at the end of a trial, one does not know *when* the filtering was detected during the trial. In fact, a viewer could report detection when the filtering was actually only perceived on the last fixation of a trial. A more informative measure is to ask viewers to detect instances of image filtering that only occur *occasionally*, for single critical fixations, several times per trial, with a normal constant high-resolution image present at all other times (see Figure 1). This occasional gaze-contingent multiresolutional display method measures how frequently removal of specified spatial frequencies is detected, using the single fixation, rather than the entire trial, as the test unit.

The effects of above-threshold filtering on eye movements. We will also examine the effects of above-threshold filtering on eye behaviour. Several studies have shown that image degradation increases fixation durations (reduced luminance: Loftus, 1985; Loftus, Kaufman, Nishimoto, & Ruthruff, 1992; low-pass filtering: Loschky & McConkie, 2002; Mannan, Ruddock, & Wooding, 1995; Parkhurst, Culurciello, & Niebur, 2000; high-pass filtering: Mannan et al., 1995). This effect has been interpreted as reflecting disrupted processing (Loschky & McConkie, 2002), in particular, increased time needed to reach an information threshold (Loftus, 1985; Loftus et al., 1992).

Importantly, no studies have investigated the effects of removing *above-threshold* frequencies on fixation durations. None of the previous studies directly investigating resolution limits across the visual field (Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001) measured eye movements. Conversely, of those studies investigating the effects of image degradation on fixation durations (Loftus, 1985; Loftus et al., 1992; Loschky & McConkie, 2002; Mannan et al., 1995; Parkhurst et al., 2000), only Loschky and McConkie first assessed the perceptibility of the degradation. However, they did not use an image filtering method based on a model of eccentricity-dependent contrast sensitivity, so we cannot use their observed effects to predict the limits of resolution across the visual field. Thus, we will relate such limits, as measured by image filtering detection, to the effects of such filtering on eye movements, creating a bridge between the literatures on spatial vision and eye movement control.

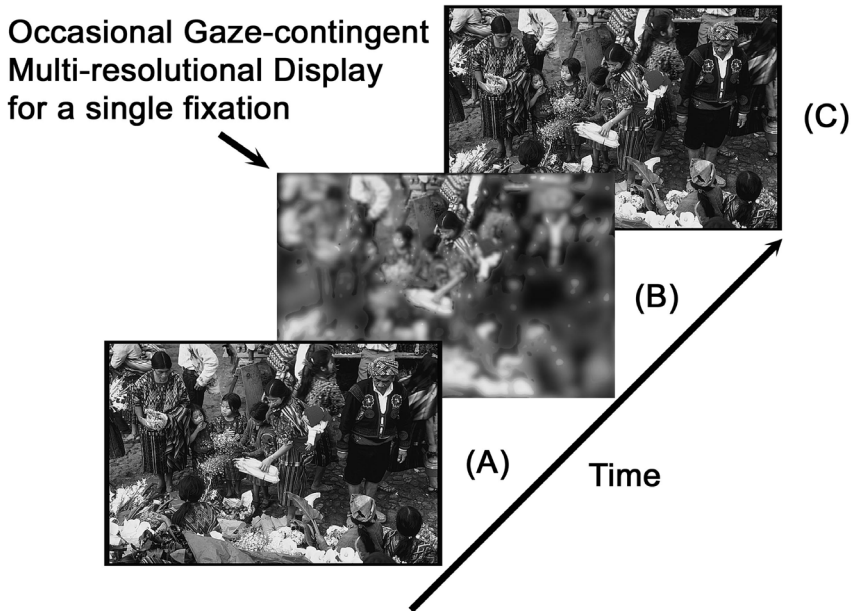


Figure 1. A schematic of an occasional GCMRD over time. (A) The viewer sees a constant high-resolution image for a prespecified number of fixations (unknown to the viewer). (B) For a single fixation, a gaze-contingent multi-resolutional image is presented with the centre of high-resolution at the viewer's point of gaze. The viewer is requested to press a button whenever he/she detects that this has happened. (C) At the start of the next fixation, the constant high-resolution image is again put on the screen for another prespecified number of fixations.

The time course of processes in detecting image filtering. We will also study the time course of perception of image filtering by investigating (1) the effects of fixation durations on the perceptibility of image filtering, (2) the effects of the perceptibility of image filtering on fixation durations, and (3) the effects, if any, of making a manual response (e.g., a button press) on fixation durations. First, we will determine if longer critical fixations, and thus, longer filtered stimulus durations, increase the likelihood of detecting image filtering, as suggested by studies of stimulus duration effects on perception and memory (e.g., Bundesen & Harms, 1999; Loftus & Ruthruff, 1994). Second, we will use the temporal structure of fixation duration distributions to elucidate the time course of processes involved in extracting scene information at or below the limits of visual resolution, using similar methods to those used to study the time course of perceptual processes in reading (Yang, 2002; Yang & McConkie, 2001, 2004). Third, a complication of our occasional gaze-contingent multi-resolutional display detection paradigm is that, since the observer must make a motor response in the detection task (e.g., a button press), the motor response

may itself affect eye behaviour (e.g., Fischer, 1999). Thus, we will also investigate the nature of the effect, if any, of making a manual response on fixation durations.

METHOD

Participants

There were 12 paid participants (6 female, 6 male), ages 19–34 ($M = 21.6$), all but one of whom had 20/20 uncorrected near vision in both eyes. The other participant had 20/30 uncorrected near vision in the right eye and 20/20 uncorrected near vision in the left.

Stimuli

The stimuli consisted of a set of 18 monochrome, 8-bit resolution photographs shown on a computer monitor (described below). The subject matter of the images was varied (from street scenes to building interiors), and contained a large amount of visual detail (see Figure 1). Images measured 768 pixels \times 512 lines and subtended $18^\circ \times 12^\circ$ visual angle.

The images were filtered using an algorithm developed at Eastman Kodak Company, which is a modified version of Geisler and Perry's (1998) foveated multiresolution pyramid. In the Geisler and Perry approach, different levels of the pyramid are circularly truncated, based on the estimated spatial frequency cutoffs of the visual system at different eccentricities. The reconstructed image from the zone-limited pyramid contains fine structure at the point of gaze, and becomes more blurred with eccentricity. In the modified method, we compute a contrast threshold map based on the peak frequency of the pyramid level, and use the contrast threshold map to threshold (i.e., reduce to zero) the image content. Whenever the image contrast of a specific frequency band is below the corresponding visual contrast threshold/cutoff, the image contrast is thresholded. All other image content is kept unchanged.

We used five levels of image filtering (see Figure 2) produced by varying a parameter in the filtering algorithm corresponding to E_2 for grating resolution (see Equation 3). The filtering E_2 values are 6.22, 3.11, 1.55, 0.78, and 0.39° . Importantly, the filtering cutoff for an E_2 value of 3.11° was predicted to be at contrast threshold at each retinal eccentricity. The largest E_2 value, 6.22° , twice the predicted threshold value, removed only higher spatial frequencies than the predicted contrast threshold at each eccentricity (i.e., below-threshold frequencies), and was, therefore, predicted to produce imperceptible image filtering. Conversely, the smaller filtering E_2 values, 1.55° , 0.78° , and 0.39° , removed incrementally lower spatial frequencies than the predicted contrast threshold (i.e., above-threshold frequencies), and were, therefore, predicted to

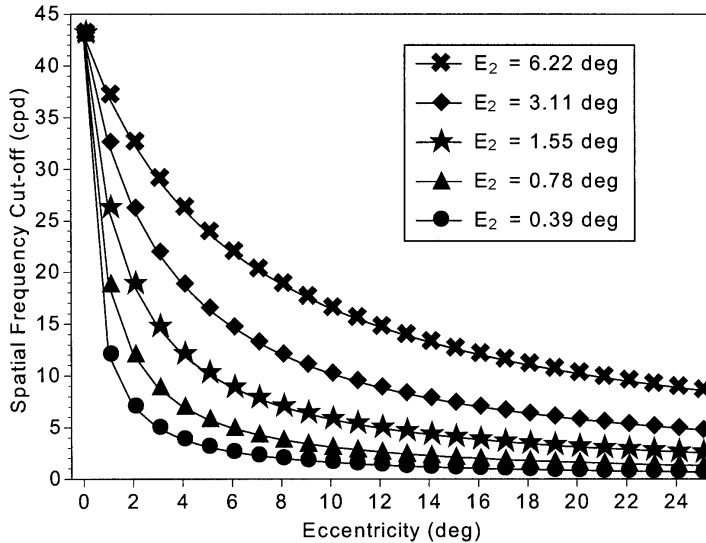


Figure 2. Spatial frequency cutoff as a function of retinal Eccentricity and filtering level for the five filtering conditions in the study. Contrast is set to 100%. cpd = cycles per degree. deg = visual angle in degrees.

produce perceptible image degradation. Example images in two filtering conditions and the control condition are shown in Figure 3.

There was a total of eighteen experimental images (three sets of six images each). Each set had an image in each of the five filtering conditions and one in a constant high-resolution, no-area-of-interest condition (i.e., a normal, high-resolution image). The six conditions were counterbalanced across images and participants.

Apparatus

Spatial alignment of the area-of-interest. In a static presentation paradigm, if the viewer carefully fixates the fixation cross, the point of highest resolution, or *area of interest*, will be accurately aligned with the centre of vision. In a dynamic presentation paradigm, imprecisely aligning the area-of-interest may result in degraded imagery being put at the fovea, even if the image resolution drop-off rate matches that of the retina. To avoid this problem, we used an eyetracker with a high level of spatial accuracy (a Dual Purkinje Image Generation 5 tracker with less than 10' error; Crane & Steele, 1978), and an array of area-of-interest locations with relatively fine spatial resolution. The array had 330 area-of-interest locations based on a 22×15 imaginary grid over

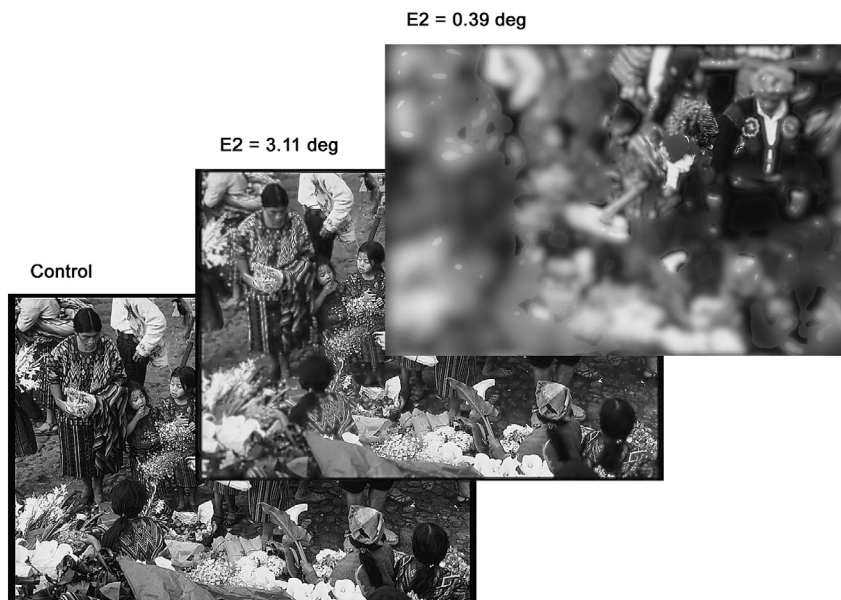


Figure 3. A set of three example images. (A) A constant high-resolution control image; (B) the same image filtered at a level that removes spatial frequencies predicted to be at or below contrast threshold ($E_2 = 3.11^\circ$); (C) the same image filtered at a level that removes spatial frequencies predicted to be 2 octaves above contrast threshold ($E_2 = 0.39^\circ$).

the $18^\circ \times 12^\circ$ image, with $\sim 0.82^\circ$ between grid points. We created 330 versions of each image, one area-of-interest version per grid point. Thus, wherever viewers looked, there was an image version whose centre of high-resolution was within 0.41° (horizontally or vertically) of their centre of gaze.

Speed of moving the area-of-interest. Unlike with a static presentation, in a gaze-contingent multiresolutional display, the area-of-interest must be quickly updated with each gaze movement. If this updating occurs too slowly, the viewer may notice degradation at the new gaze location prior to the update, or motion caused by the change in resolution across the image. Thus, we designed a system to minimize image-update delays. First, eye position was sampled at 1000 Hz (1 sample/ms), providing high temporal resolution in identifying the ends of saccades. Second, to avoid on-line image generation time, all 330 versions of each image were pre-stored in a ViewGraphics ViewStore 1 GB randomly accessible image memory and display controller. Third, display changes were initiated within 5 ms of the end of each saccade, based on the results of a separate study (McConkie & Loschky, 2002) showing that this deadline

eliminated perception of image-update motion. Fourth, our Conrac Mars 9320 monitor had a 60 Hz refresh rate, and updates could be made at any point during the 17 ms refresh cycle. Thus, an area-of-interest update was completed within 22 ms of the start of a fixation.

These methods allow greater confidence that our observed effects are the result of the spatial characteristics of the images rather than the result of inaccurate area-of-interest placement or perceived image motion.

Procedures

The procedures were designed to measure participants' detection of peripheral filtering during dynamic image viewing. On most fixations, the normal, constant high-resolution image was displayed. The gaze-contingent multiresolutional display image was presented only on occasional, unpredictable, single fixations. Participants' two tasks were to (1) prepare for a long-term memory change detection test, and (2) do a concurrent peripheral blur detection task. The memory task was to motivate participants to make many eye movements. Participants were told they would be tested after seeing all 18 images; however, no final test was actually given.

Figure 1 is a schematic of the occasional gaze-contingent multiresolutional display paradigm. Each trial was a single picture presented for 30 s, beginning with a fixation point at the centre of the screen. When ready, the participant pushed a button to begin the trial. After 750 ms, a constant high-resolution picture appeared. After a randomly varying interval of 9–11 saccades, at the start of the next fixation, a gaze-contingent multiresolutional version of the image was presented with its point of highest resolution at the centre of gaze. The original constant high-resolution image then returned during the next saccade. Thus, the blurred image was present for a single *critical fixation*. The next critical fixation occurred after another 9–11 saccades, and this pattern continued until the end of the trial. Participants had to press a button as quickly as possible whenever they noticed image blur.

Each set of six images constituted a block, and the eyetracker was also recalibrated before each block. Participants were given detailed instructions and practice with the tasks prior to the experiment.

In the analyses of the data, a probability level of .05 was used for all inferential statistics.

RESULTS

Precursors

In this experiment, we tested the limits of visual resolution during free viewing of natural scenes, the effects of above-threshold filtering on eye movements, and the time course of those effects. We first cleaned the data by making several

exclusions. We excluded all data from one of the 18 images (5.25% of the total) because of an artifact (a slight vertical shift between image versions). We excluded blinks, and cases with fixation durations in the 99th percentile and saccade lengths in the 1st percentile (3164 ms, and 0.16° , respectively). We kept short fixation durations to study their effect on detection, and excluded the very shortest saccades to avoid possible display change errors. We also analysed area-of-interest positioning error and excluded all cases with errors $\geq 2^\circ$ (3.6% of the total). After all exclusions (12% of the original total), 925 critical fixations remained, with 94% having an area-of-interest positioning error $< 1^\circ$ ($M = 0.50^\circ$, median = 0.45° , mode = 0.32°).

Perceptibility of image degradation as a function of filtering level

A key question is whether the Yang et al. (2001) model of visual sensitivity, based on studies using simple grating and Gabor patch stimuli, predicts the perceptibility of image filtering during dynamic free viewing of natural scene images. To answer this question, we measured viewers' explicit detection task responses, and their eye fixation durations. The detection task is a direct measure of perceptibility, but viewers' fixation durations might reveal an implicit sensitivity to filtered spatial frequencies absent from their explicit detection responses (cf. Hayhoe, Bensinger, & Ballard, 1998; Hollingworth, Williams, & Henderson, 2001). This combination of explicit and implicit measures should give a more detailed picture of viewers' sensitivity across the field of view in natural scene viewing. As we will show below, the results are largely consistent with the Yang et al. (2001) model.

Detection as a function of filtering level. The Yang et al. (2001) model predicted that filtering level 1 would be undetectably degraded, that level 2 would be at detection threshold, and that levels 3–5 would be increasingly detectable. We analysed detection rates as a function of filtering level using a multivariate ANOVA (five levels of filtering) on participants' arcsine transformed proportion detection data (Table 1). Raw proportions are shown in Figure 4, indicating a robust and significant main effect for filtering level on detection. Linear and cubic trends for level of filtering on detection were significant, linear trend, $F(1, 11) = 619.58$, $p < .001$, effect size = .98; cubic trend, $F(1, 11) = 83.11$, $p < .001$, effect size = .88, indicating that as filtering E_2 decreased, detection increased, but with both low and high detection asymptotes. The error bars in this figure represent 95% confidence intervals for the difference of each point against all others. Though the Yang et al. model predicted that filtering level 2 would be at detection threshold, there was no significant difference between filtering levels 1 and 2 ($E_2 = 6.22$ vs. 3.11°), neither of which was statistically different from zero detection. However,

TABLE 1
Multivariate ANOVA on mean proportion detection as a function of filtering level
($E_2 = 0.39\text{--}6.22$)

<i>Effect</i>	<i>Pillai's trace</i>	<i>F</i>	<i>Hypothesis df</i>	<i>Error df</i>	<i>Sig.</i>	<i>Partial eta squared</i>
Filtering level	.986	139.515	4.000	8.000	.000	.986

consistent with the model, filtering level 1 ($E_2 = 6.22$) was essentially never detected—the five detections across all participants were essentially identical to the six false alarms in the no-area-of-interest condition. There was no significant difference in detection between filtering levels 4 and 5 ($E_2 = 0.78$ vs. 0.39°) because both were at ceiling, which is consistent with the model's predictions of high detection levels for both. Consistent with the model's predictions, filtering levels 2–4 showed increasing detectability with significant differences between the three middle levels ($E_2 = 3.11^\circ$ vs. 1.55° and $E_2 = 1.55^\circ$ vs. 0.78°).

Fixation durations as a function of filtering level. Although the Yang et al. (2001) model makes no predictions about eye movements, assuming fixation durations implicitly measure perception, it would predict the same or stronger effects of filtering on fixation durations as for explicit detection. We analysed mean fixation durations on the critical fixation as a function of filtering level

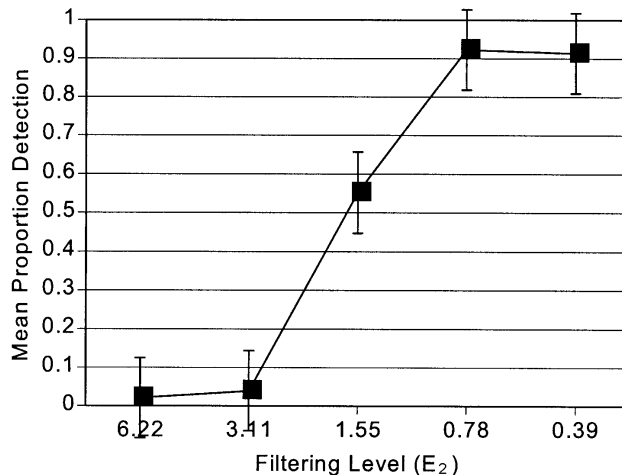


Figure 4. Proportion detection of occasionally presented gaze-contingent multiresolutional images as a function of peripheral filtering level ($E_2 = 0.39\text{--}6.22^\circ$). Error bars represent 95% confidence intervals for significance of all comparisons.

TABLE 2
Multivariate ANOVA on mean fixation duration as a function of filtering level ($E_2 = 0.39\text{--}6.22$) with the no-area-of-interest condition

<i>Effect</i>	<i>Pillai's trace</i>	<i>F</i>	<i>Hypothesis df</i>	<i>Error df</i>	<i>Sig.</i>	<i>Partial eta squared</i>
Filtering level	.871	13.524	4.000	8.000	.001	.871

(five levels plus no-area-of-interest condition) with a multivariate ANOVA (Table 2). The results showed a significant main effect of image filtering on fixation durations. Follow-up, planned comparisons of each filtering level versus the no-area-of-interest condition are shown in Figure 5, with error bars indicating 95% confidence intervals of the difference for each condition versus the no-area-of-interest condition. The pattern of results looks very similar to that for detection, with two differences. First, strangely, fixation durations in the no-area-of-interest (all high-resolution) condition were longer than those in the least-filtered condition (level 1, $E_2 = 6.22^\circ$), possibly reflecting different processing on these trials, compared with those experimental trials in which the area-of-interest actually appeared occasionally. To avoid this possible processing difference in the no-area-of-interest trials, we created a control condition

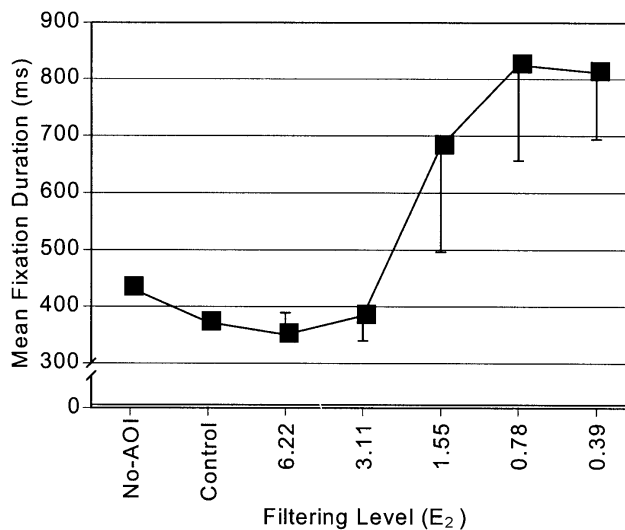


Figure 5. Mean fixation duration on occasionally presented gaze-contingent multiresolutional images as a function of peripheral filtering level ($E_2 = 0.39\text{--}6.22^\circ$), the no-area-of-interest condition, and the control condition (based on the sixth and seventh fixations following each occasional gaze-contingent multiresolutional image across all filtering conditions). no-AOI = no-area-of-interest.

TABLE 3
Multivariate ANOVA on mean fixation duration as a function of filtering level ($E_2 = 0.39$ – 6.22) with the control condition

<i>Effect</i>	<i>Pillai's trace</i>	<i>F</i>	<i>Hypothesis df</i>	<i>Error df</i>	<i>Sig.</i>	<i>Partial eta squared</i>
Filtering level	.952	27.966	5.000	7.000	.000	.952

from the experimental trials by taking the sixth and seventh fixations following each critical fixation (i.e., the fixation on which the gaze-contingent image filtering was present) from each filtering condition ($E_2 = 6.22$ – 0.39°). These selected fixations had no filtering or manual responses associated with them, came several fixations after both the critical fixation and any manual responses that may have been made, and were several fixations prior to the next critical fixation. Using this control condition (see Figure 5), we reanalysed participants' mean fixation durations with a multivariate ANOVA (Table 3). The results were the same as before, except that, consistent with the detection results and our predictions, there was no significant difference between the least filtered condition ($E_2 = 6.22$) and the control. Also consistent with the detection results, there was a significant difference between filtering levels 2 and 3 ($E_2 = 3.11^\circ$ vs. 1.55°). However, contrary to the detection results and our predictions, there was no significant difference between filtering levels 3 and 4 ($E_2 = 1.55^\circ$ vs. 0.78°).

The time course of processes in detecting image filtering

The effect of fixation durations on filtering detection rates. Our first question was whether the duration of a critical fixation affects the frequency of detecting image filtering. In particular, is there some minimal fixation duration (or stimulus presentation duration) below which no detections occur, and some fixation duration above which asymptotic detection performance is reached? The answer to both questions is "Yes." As shown below, in the most detectably filtered conditions, detections did not occur for fixations much below 100 ms, and reached asymptote by roughly 200 ms. These values were greater in the threshold filtering condition.

Three filtering level conditions were produced by combining data for conditions whose detection levels were very similar, filtering levels 1 and 2 ($E_2 = 6.22^\circ$ and 3.11°), and filtering levels 4 and 5 ($E_2 = 0.78^\circ$ and 0.39°): Filtering level 3 ($E_2 = 1.55^\circ$) constituted a third condition. The first condition, containing filtering levels 1 and 2, was virtually never detected at any fixation duration and is therefore uninformative for this analysis. Figure 6 shows, for the second and

third groups, the frequencies of detections and nondetections as a function of fixation duration in 10 ms bins for the first 500 ms of viewing.

The top panel of Figure 6 shows that for the detection threshold filtering level ($E_2 = 1.55^\circ$), the first detection occurred at roughly 200 ms, and detections increased thereafter. The bottom panel of Figure 6 shows that for the two highest filtering levels ($E_2 = 0.78$ and 0.39°) detections first occurred for fixations slightly less than 100 ms, and rapidly increased for fixations between 140 and 200 ms long. Figure 7 shows a more global representation of these trends, by combining the fixations into bins of 40 ms and showing the proportion detection for each bin.

The effect of filtering detectability on fixation durations. Our next question was, how does perceptible filtering affect the time course of fixation durations? As detailed below, perceptible image degradation had no effect on fixation durations before 160 ms, after which there was a gradual dampening of saccadic activity until roughly 400 ms, followed by a long lull in saccadic activity lasting up to 1000–2000 ms. Notably, differences between the threshold and highly detectable filtering conditions were much less than between those conditions and the undetectable filtering conditions.

Our approach was similar to research on the time course of perceptual and cognitive processes in reading, which showed how text manipulations affected fixation durations, using hazard curve analyses (Yang, 2002; Yang & McConkie, 2001, 2004). Hazard values give the probability of finality as a function of time, and can be compared across conditions. Example domains include demography (e.g., probability of death as a function of age, compared across gender, nationality, etc.) and engineering (e.g., probability of light bulb failure as a function of hours used, compared across filament type, brand, etc.). Here we analysed the probability of ending an eye fixation (i.e., making a saccade) as a function of its duration in milliseconds, compared across image filtering levels. Fixations were put into 40 ms intervals, with the interval labelled by the maximum value. Hazard values were calculated as the interval frequency divided by the number of remaining cases (total frequency minus cumulative frequency). Because successive values are based on fewer and fewer cases, the analysis was terminated at the 95th percentile. Data were smoothed using the average over a three-interval window, with the current interval doubly weighted. For comparability with our preceding analysis of the effects of fixation durations on detection, we used the same three filtering level groups: The undetectable filtering levels ($E_2 = 6.22$ – 3.11°), the detection threshold filtering level ($E_2 = 1.55^\circ$), and the well-above-detection-threshold filtering levels ($E_2 = 0.78$ – 0.39°), plus the control condition.

The overall results of this analysis are shown in Figure 8, with the top panel showing the entire time course and the bottom panel showing greater detail during the first 600 ms. The hazard curves for all conditions start very low and

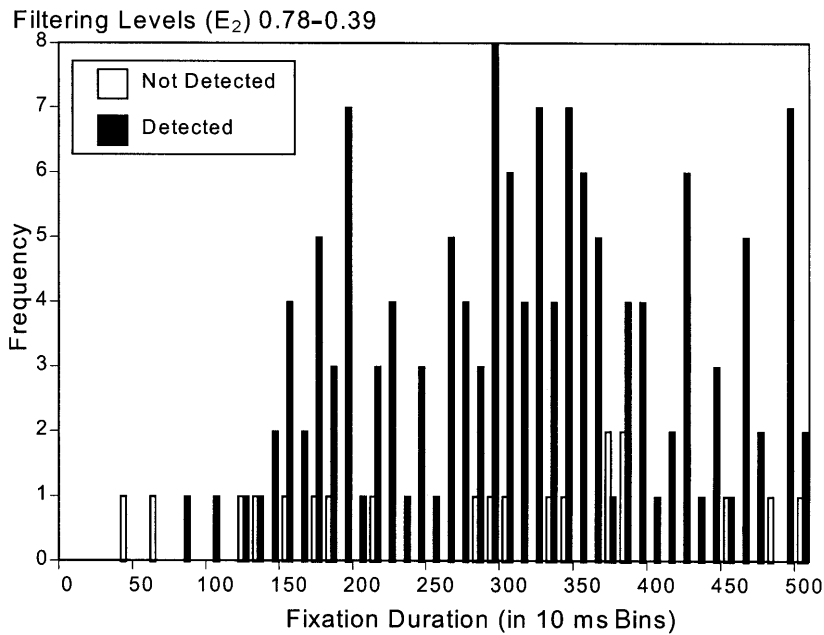
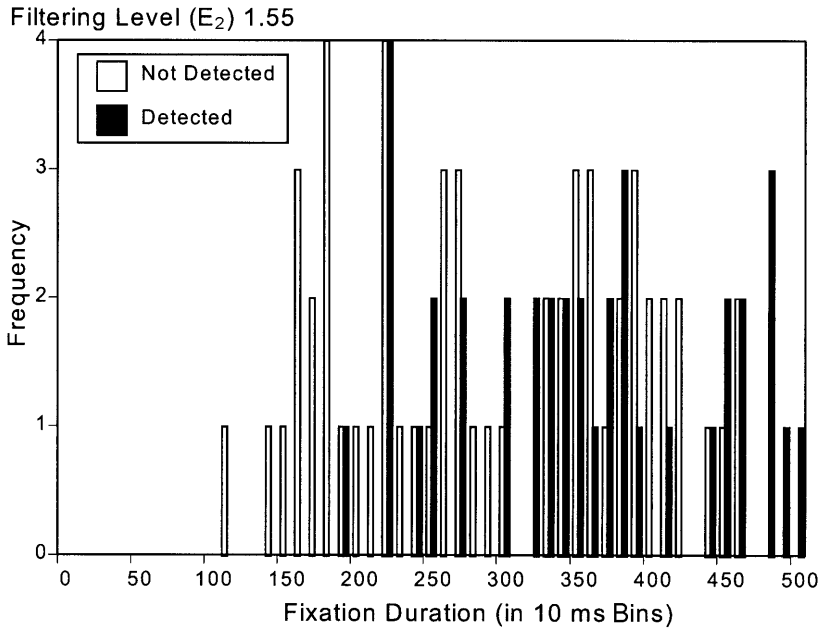


Figure 6. Frequencies of detected and undetected occasionally presented gaze-contingent multi-resolutional images as a function of grouped filtering level ($E_2 = 1.55^\circ$, and $0.78-0.39^\circ$) and fixation duration (in 10 ms bins).

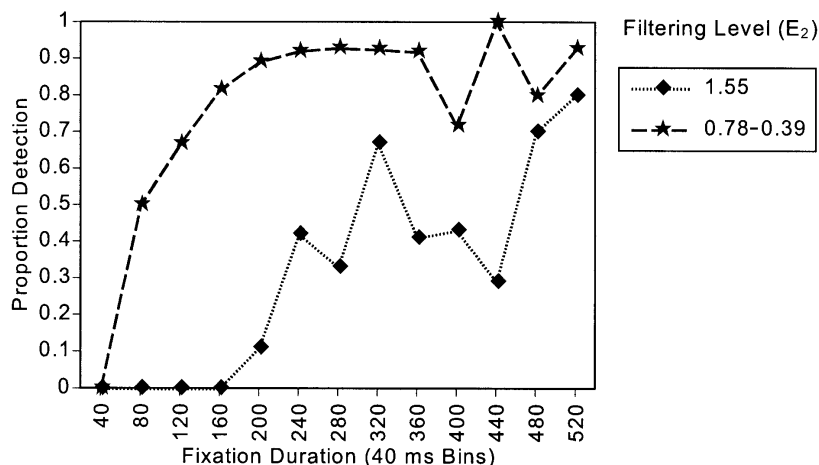


Figure 7. Proportion detection of occasionally presented gaze-contingent multiresolutional images as a function of grouped filtering level ($E_2 = 1.55^\circ$, and $0.78\text{--}0.39^\circ$) and fixation duration (in 40 ms bins).

then begin rising from around 120 ms (i.e., few saccades are made before then). In the undetectable filtering levels ($E_2 = 6.22\text{--}3.11^\circ$), the curve rises rapidly until about 320 ms, showing increasing saccadic activity. After that, the hazard curve is nearly asymptotic, suggesting a saturation of saccadic activity. Note that the rise in the curve around 520 ms occurs when 85% of the saccades have already been made and the data are less stable. In the highly detectable filtering levels ($E_2 = 0.78\text{--}0.39^\circ$), from 160 ms the hazard curve begins rising more slowly than in the control and undetectable conditions, indicating relatively lower saccadic activity than in the undetectably filtered conditions (i.e., fixation durations start lengthening). Then from 360 ms, the curve initially drops followed by a bumpy asymptote until 95% of fixations have ended by 2280 ms, indicating a drop in saccadic activity followed by an extended lull in activity (i.e., long fixation durations). The detection-threshold filtering level ($E_2 = 1.55^\circ$) shows a generally similar pattern. After separating from the undetectably filtered conditions at 160 ms, the hazard curve rises more slowly than those conditions until 400 ms, then drops, showing suppressed saccadic activity until roughly 920 ms. After that, the curve rises again and then reaches an unstable asymptote, suggesting a long period of fluctuating but relatively low saccadic activity.

A key question is whether or not the time courses of information accrual and eye movement control are the same. That is, how does the minimum effective fixation duration for detection compare with the time at which filtering first affects fixation durations? The bottom panel of Figure 8 shows that the first separation between the hazard curves for the detectably filtered conditions ($E_2 =$

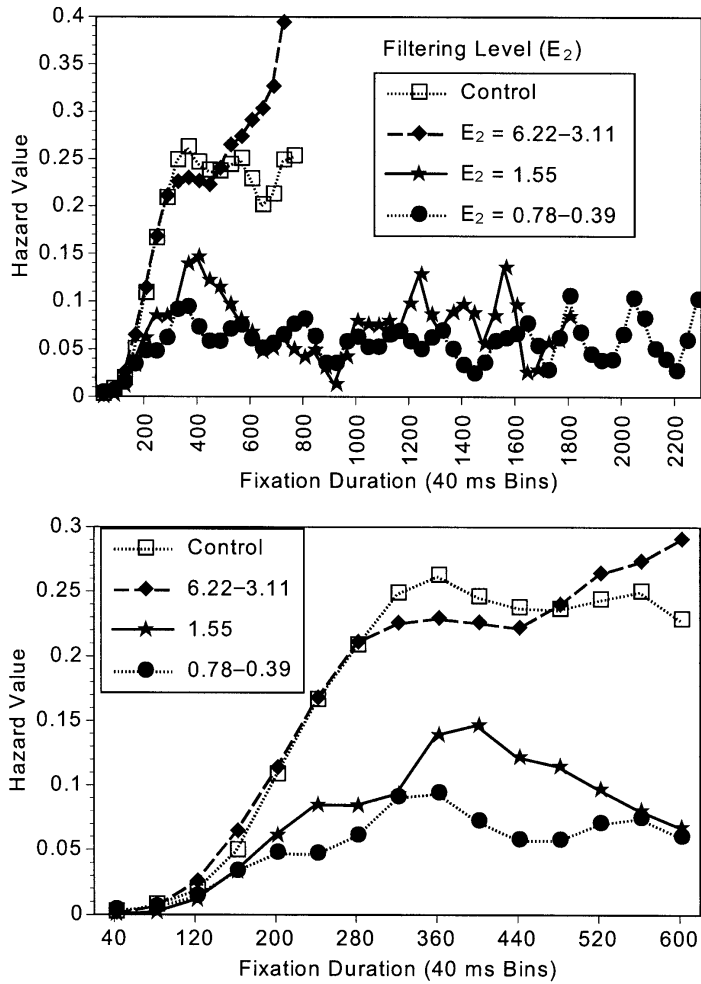
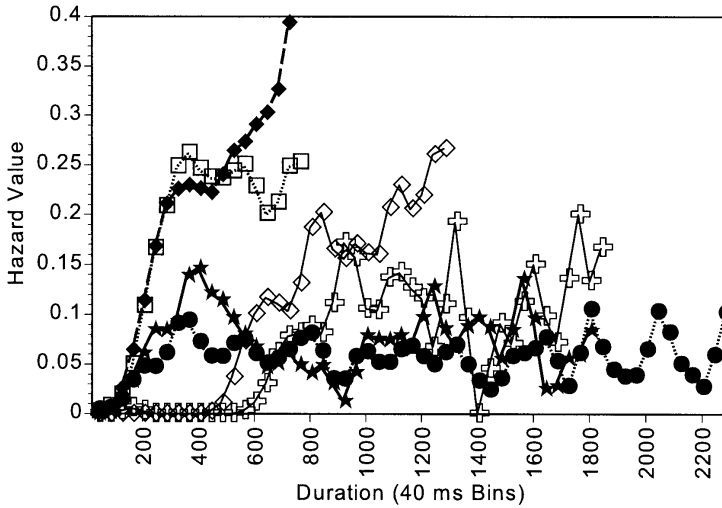


Figure 8. Top panel: Fixation duration hazard curves as a function of grouped filtering level ($E_2 = 6.22\text{--}3.11^\circ$, 1.55° , and $0.78\text{--}0.39^\circ$) with curves carried out to the 95th percentile for each filtering level group. Fixation duration intervals labelled by the maximum value of bins 40 ms in width, with the data smoothed (average over three-bin window with the centre bin weighted double). Control = control condition based on the sixth and seventh fixations following each occasional gaze-contingent multiresolutional image across all filtering conditions. Bottom panel: Same as top panel, but focused on the first 600 ms.

1.55–0.39°) and the control and undetectably filtered conditions ($E_2 = 6.22\text{--}3.11^\circ$) is at roughly 160–200 ms. Thus, in the highly detectable filtering conditions ($E_2 = 0.78\text{--}0.39^\circ$), although an 80–120 ms fixation duration might provide enough temporal integration of sensory information to detect image filtering (Figure 7), it would be too short to show an effect of filtering on eye movements (bottom panel of Figure 8). In the case of the threshold filtering condition ($E_2 = 1.55^\circ$), the timings are much closer together. The minimum effective fixation duration for detecting filtering is about 200 ms (Figure 7), which is about when fixation durations in this condition begin differing from the control and undetectable filtering conditions (bottom panel of Figure 8).

The effect of a manual response on fixation durations. Our next question was whether making a manual response had any effect on fixation durations (Fischer, 1999), and if so, at what point in time? Regarding the time course of detection, our most fundamental prediction would be that any response-related effects on fixation durations should occur after perceptual processing effects, since perception of filtering must logically precede any explicit detection response to it. However, we would also predict that response-related effects on fixation durations might start prior to actually executing the manual response, since both the detection decision process and motor programming of the manual response would occur prior to the actual button press, and either might affect fixation durations. In fact, as we will detail below, making a manual response did appear to lengthen fixation durations. This lengthening occurred in close temporal proximity to the manual response, either during a long fixation that preceded and continued beyond the manual response, or during the fixation immediately following the manual response.

To test the predictions listed above, we can look at Figure 9. This adds to the top panel of Figure 8 the manual response time (RT) hazard curves for the highly detectable ($E_2 = 0.78\text{--}0.39^\circ$) and threshold detection filtering levels ($E_2 = 1.55^\circ$) on the same time scale. Manual RTs show a flat hazard curve before 520 ($E_2 = 0.78\text{--}0.39^\circ$) to 600 ms ($E_2 = 1.55^\circ$), indicating no responses before then, with rapid increases in manual response activity thereafter. In contrast to these results, recall that the separation between fixation duration hazard curves in the detectable and undetectable conditions occurred at roughly 160–200 ms. It seems unlikely that this initial drop in saccadic activity at 160–200 ms is caused by an imminent manual response 360–400 ms later. Rather, the drop in the fixation duration hazard curves for the detectable filtering conditions after 400–440 ms seems more plausibly related to the increasing manual response hazard curves at 520–600 ms. We see the drop in the fixation duration hazard curve for the highly detectable filtering conditions ($E_2 = 0.78\text{--}0.39^\circ$) at 400 ms occurs 120 ms before the corresponding manual RT hazard curve first rises at 520 ms. Likewise, the drop in the fixation duration hazard curve for the detection-threshold filtering condition ($E_2 = 1.55^\circ$) at 440 ms occurs 160 ms before



Fixation Duration vs. Manual RT, and Filtering Level (E_2)

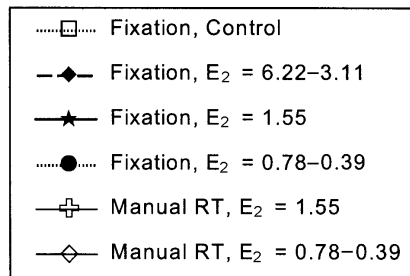


Figure 9. Fixation duration and manual response time hazard curves as a function of grouped filtering level ($E_2 = 6.22-3.11^\circ$, 1.55° , and $0.78-0.39^\circ$) with curves carried out to the 95th percentile for each filtering level group. Manual response time hazard curves given for those conditions having many detection responses ($E_2 = 1.55^\circ$, and $0.78-0.39^\circ$). Fixation duration intervals labelled by the maximum value of bins 40 ms in width, with the data smoothed (average over three-bin window with the centre bin weighted double). Fixation = fixation duration data. Manual RT = manual reaction time data. Control = control condition based on the sixth and seventh fixations following each occasional gaze-contingent multiresolutional image across all filtering conditions.

the rise in the corresponding manual RT hazard curve at 600 ms. These results are consistent with the above predictions regarding manual response-related effects. Specifically, the drop in saccadic activity at 400–440 ms is consistent with a manual response effect since it occurs after perceptual processing effects (at asymptote by 200 ms), but before the actual button press (starting at 520–600 ms).

If making a manual response interferes with making a saccade, then does this interference disappear as soon as the manual response is made? Interestingly, it does not. Figure 9 shows that in the detection threshold filtering condition ($E_2 = 1.55^\circ$) the longest latencies for both manual responses and eye movements are very similar (both around 1800 ms), whereas in the highly detectable filtering conditions ($E_2 = 0.78\text{--}0.39^\circ$), the longest eye movement latencies are about 1000 ms longer than the longest manual response latencies (roughly 2200 ms vs. 1200 ms, respectively). These findings suggest that, at least in the highly detectable filtering conditions, something besides the manual response is lengthening fixation durations, probably the image filtering itself.

To distinguish between the effect of the image filtering and the effect of the manual response on fixation durations, one can examine fixations immediately following the critical fixation. We will call these critical fixation +1, or n+1 fixations. Importantly, there was no image filtering present on those fixations. Our first interest is in that subset of n+1 fixations in which the previously viewed filtering was detected, but the manual response was *after* the critical fixation, n. Manual responses to detected filtering occurred *after* the critical fixation, n in 59% of cases in the highly detectable filtering conditions ($E_2 = 0.78\text{--}0.39^\circ$). Figure 10 compares the hazard curves in the highly detectable filtering conditions ($E_2 = 0.78\text{--}0.39^\circ$) for fixations n and n+1. Note that that in fixation n+1 a manual response was imminent (and occurred in 56% of these n+1 fixations), but there was no filtering present. As shown in Figure 10, even when there was no filtering present, the fact that a manual response occurred, or was imminent, was enough to reduce the fixation duration hazard curve relative to the control condition, suggesting interference with saccadic activity. On the other hand, this reduction in the fixation duration hazard value was only about half that shown for the critical fixation, suggesting that the manual response alone cannot account for all the reduction in saccadic activity. The remainder of this effect was assumedly due to the image filtering itself.

We can further explore the impact of the manual responses on fixation n+1 by comparing the fixation duration hazard curve for the highly detectable filtering conditions ($E_2 = 0.78\text{--}0.39^\circ$) with the corresponding manual RT hazard curve on the same fixation (Figure 11). Here, manual RTs are given relative to the start of fixation n+1. The earliest manual RTs are in the -40 ms bin, meaning they occurred during the saccade preceding fixation n+1. This figure shows that the drop in saccadic activity early in the fixation corresponds with rising manual response activity.

Finally, we compare this same fixation n+1 condition with the fixation n+1 condition in which the manual response *was* made during the critical fixation n (Figure 12). In this latter condition, there is no image filtering present, nor is a manual response made or even imminent on fixation n+1. Thus, we might expect normal saccadic activity in this condition. Surprisingly, *both* types of n+1 fixations show reduced saccadic activity for the first 360 ms. Thus, even in the

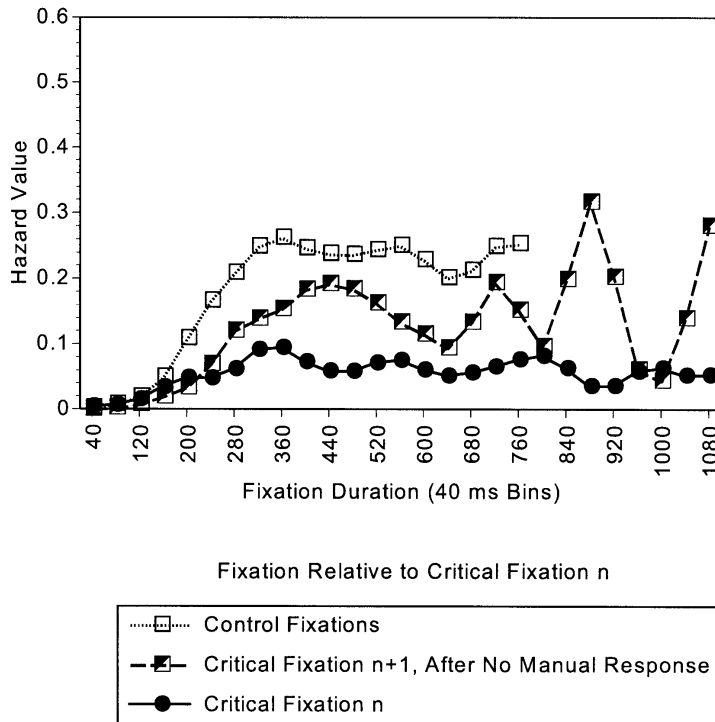


Figure 10. Fixation duration hazard curves as a function of fixation relative to the critical fixation n (for filtering levels $E_2 = 0.78-0.39^\circ$). Fixation duration intervals labelled by the maximum value of bins 40 ms in width, with the data smoothed (average over three-bin window with the centre bin weighted double). Control fixations = control condition based on the sixth and seventh fixations following each occasional gaze-contingent multiresolutional image across all filtering conditions. Critical fixation $n+1$, after no manual response = fixation immediately following a critical fixation on which no manual response was made, but a manual response was imminent. Critical fixation n = fixation on which filtering occurred.

absence of image filtering, and with the manual response having been made during the previous fixation, there is still reduced saccadic activity for 360 ms—an apparent spillover effect from the critical fixation n onto fixation $n+1$. However, after 360 ms, saccadic activity becomes normal in this condition, whereas it is still reduced for another 360 ms when a manual response is made or is imminent.

In sum, even in the absence of filtering, if a manual response occurs or is imminent, saccadic activity is reduced on fixation $n+1$. Furthermore, the suppression of saccadic activity found on fixation $n+1$ corresponds with the timeline for late manual responses made on that fixation. Finally, there appears to be a

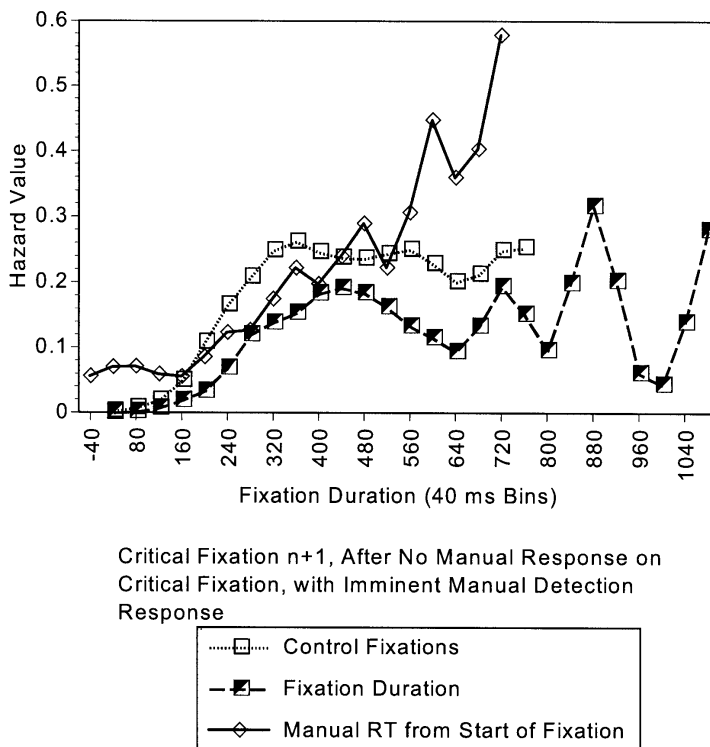


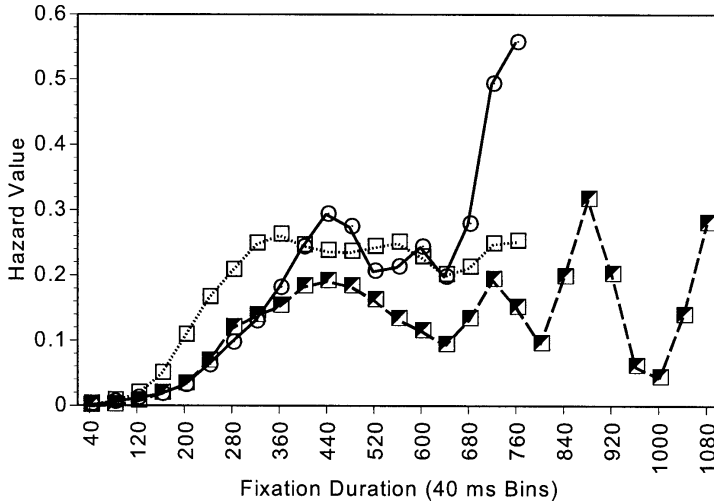
Figure 11. Fixation duration and manual response time hazard curves as a function of fixation relative to the critical fixation (for filtering levels $E_2 = 0.78-0.39^\circ$). Fixation duration intervals labelled by the maximum value of bins 40 ms in width, with the data smoothed (average over three-bin window with the centre bin weighted double). Control fixations = control condition based on the sixth and seventh fixations following each occasional gaze-contingent multiresolutional image across all filtering conditions. Fixation duration = fixation immediately following a critical fixation on which no manual response was made, but a manual response was imminent. Manual RT from start of fixation = manual response time relative to the beginning of the current fixation (in ms) for responses on the current fixation.

spillover effect reducing saccadic activity on fixation n+1 when both image degradation and a manual response occurred immediately prior to it.

DISCUSSION

The limits of resolution in dynamic natural scene viewing

A central goal of the current study was to determine the limits of visual resolution in free viewing of natural scenes, and in so doing to test the Yang et al. (2001) model of eccentricity-dependent contrast sensitivity. The detection and



Critical Fixation $n+1$, After Manual Response on Critical Fixation n or Not

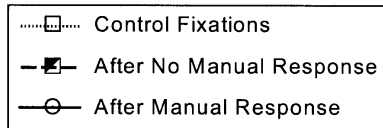


Figure 12. Fixation duration hazard curves as a function of whether a response was made on the preceding fixation (for filtering levels $E_2 = 0.78-0.39^\circ$). Fixation duration intervals labelled by the maximum value of bins 40 ms in width, with the data smoothed (average over three-bin window with the centre bin weighted double). Control fixations = control condition based on the sixth and seventh fixations following each occasional gaze-contingent multiresolutional image across all filtering conditions. After no manual response = fixation immediately following a critical fixation on which no manual response was made, but a manual response was imminent. After manual response = fixation immediately following a critical fixation on which a manual response was made.

fixation duration results were generally consistent with the predictions of the model, which was based on earlier studies that measured viewers' eccentricity-dependent CSFs using grating or Gabor patch stimuli (Anderson et al., 1991; Banks et al., 1991; Cannon, 1985; Peli et al., 1991; Pointer & Hess, 1989; Robson & Graham, 1981). Filtering level 1 ($E_2 = 6.22^\circ$), which removed only higher spatial frequencies than the predicted contrast threshold at each retinal eccentricity, was almost never detected, and fixation durations showed a similar pattern. Conversely, filtering levels 4-5 ($E_2 = 0.78-0.39^\circ$), which removed spatial frequencies predicted to be 2-3 octaves lower than the contrast threshold, were at asymptote for both detection (above 90%) and fixation durations.

The current results are also consistent with earlier studies using briefly flashed natural images with peripheral filtering based on eccentricity-dependent

CSFs (Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001). Those studies found that viewers were somewhat less sensitive to peripheral image filtering than predicted based on studies using isolated grating or Gabor patch stimuli. Consistent with those studies, our *predicted* threshold filtering level—level 2 ($E_2 = 3.11^\circ$)—was *almost never* detected, and the fixation duration results showed the same pattern (i.e., no difference from the control). Rather, the first filtering level to remove lower spatial frequencies than the predicted threshold (level 3— $E_2 = 1.55^\circ$)—was detected 55% of the time, and appears to be the actual threshold filtering level. Lateral masking from neighbouring natural image regions may account for this lower-than-predicted sensitivity (Peli & Geri, 2001; Yang et al., 2001). A related possibility involving attention is the well-known narrowing of the functional field of view by image clutter (Ball, Beard, Roenker, Miller, & Griggs, 1988; Bertera & Rayner, 2000; Mackworth, 1965; Sekuler, Bennett, & Mamelak, 2000).

The similarity of these results under dynamic viewing conditions to those found with statically presented multiresolutional images (Peli & Geri, 2001; Sere et al., 2000; Yang et al., 2001) is somewhat surprising, since, for the technical reasons discussed in our Apparatus section, one might predict higher detection rates under dynamic viewing conditions than under static ones. Our replication of lower-than-predicted sensitivity to image filtering, found in previous studies using static presentation of multiresolutional images, may be because we managed to eliminate low-level artifacts from our gaze-contingent multiresolutional display. As noted earlier, we updated our images as quickly as possible using an update initiation deadline of 5 ms after the end of a saccade (McConkie & Loschky, 2002), thus likely eliminating detections based on motion transients. We also used an eyetracking system (Crane & Steele, 1978) with high spatial resolution and a relatively fine grid of spatial locations for the centre of high-resolution, thus achieving a relatively small degree of error in placing the centre of high-resolution at the centre of gaze.

In sum, the Yang et al. (2001) model is largely supported by the data, though based on the current results we may need to adjust the parameters (e.g., k , representing the rate of visual resolution drop-off). We estimate that the resolution limit at each eccentricity in natural scene viewing is approximated by Equation 3 (restated below) with an E_2 value of 1.55, since that was the filtering threshold E_2 value for both detection and fixation durations:

$$f_c = 43.1 * 1.55 / (1.55 + E)$$

Again, f_c is the frequency cutoff, or resolution limit, and E is retinal eccentricity. Further research may more precisely pinpoint the perceptual threshold using more filtering levels or perhaps different underlying hypothesized CSFs, but the current estimate provides a good first approximation.

The limits of visual resolution across the visual field and scene perception. The above estimate of the limits of visual resolution in dynamic scene perception provides an important limit on the perceptible information assumed to be available in extrafoveal scene regions. This limit provides important constraints for theories of scene perception dealing with scene gist and layout acquisition, extrafoveal object recognition, and eye movement guidance in extended scene viewing. These limits are relevant to theories of scene processing whether or not they consider extrafoveal object recognition to be important.

Consider theories in which extrafoveal recognition of diagnostic objects plays an important role in scene gist acquisition (e.g., Biederman, Rabinowitz, Glass, & Stacy, 1974; Friedman, 1979). Studies have suggested an ideal spatial frequency band for object recognition (Braje et al., 1995; Gold et al., 1999; Olds & Engel, 1998). (While such studies tend to express these bandwidths in cycles/object, rather than cycles/degree, as in the current study, it is possible to translate between these scales, with the simplest case being an object 1° wide.) If object recognition is also isolated from scene semantic constraints (Hollingworth & Henderson, 1999), it would suggest that when a scene-diagnostic object is at an eccentricity where the central frequency of the critical bandwidth is below threshold, object recognition should suffer, and therefore scene gist acquisition should suffer as well.

Next, consider theories in which parafoveal or peripheral object information is used to guide the eyes in extended scene viewing, such as in a search task (Rao, Zelinsky, Hayhoe, & Ballard, 2002; Wolfe, 1994). If the eyes are targeted on the basis of discriminating target and nontarget objects, knowing the limits of resolution in scenes should be important for such a theory (e.g., Rao et al., 2002). Discrimination of a peripheral distractor from the target should be influenced by (1) the spatial frequency bandwidth necessary for discrimination, (2) the retinal eccentricity and contrast of the target and distractor, and (3) the limits of resolution in scenes. Thus, if diagnostic object recognition, or object feature discrimination, is assumed to play an important role in acquiring scene gist, or in guiding the eyes during extended scene viewing, knowing the limits of visual resolution across the visual field in scene viewing is critical.

However, even in theories of scene gist acquisition or eye guidance positing no role for object recognition or object feature discrimination, knowing the limits of visual resolution should be important. For example, Schyns and Oliva (1994) have argued that most frequently, low spatial frequency information is used early in determining scene gist. However, in roughly a third of their trials, high spatial frequency information was used for that purpose. A possible explanation for this asymmetry in information use is in terms of the visual resolution drop-off across the visual field. That is, high spatial frequency information is available in only a small part of the visual field, near the centre of vision, whereas low spatial frequency information is available across the entire

visual field. Thus, if high spatial frequency information useful for acquiring scene gist is only resolvable, because near the fovea, about one third of the time, the visual system would need to rely on the resolvable lower spatial frequency information the remaining two thirds of the time. In this way, the limits of visual resolution in scene viewing, as estimated in the current study, may explain a reliance on low spatial frequency information in acquiring gist (Schyns & Oliva, 1994).

Similarly, an interesting question for theories of gist acquisition that posit an important role for global texture discrimination (e.g., Walker & Malik, 2002) is whether the texture spatial frequency profile for a scene type varies with the point of fixation, or whether any initial fixation point would provide a match. In either case, an accurate description of the texture spatial frequency profiles used to identify scenes of various categories will need to incorporate the limits of visual resolution. A similar argument would apply to the spatial envelope model of gist acquisition (Oliva & Torralba, 2001).

In the case of eye movement guidance in scene viewing, an interesting possibility is that the limits of visual resolution, or more generally CSFs at different eccentricities, play a role in determining the salience, or attractiveness, of peripheral targets for eye movements with or without recognizable peripheral target objects (see Deubel, 1984; Loschky, 2003; Loschky & McConkie, 2002). An interesting question for further research is whether the filtering threshold for detecting image blur in the current study is the same as the filtering threshold for affecting saccade targeting in other studies (e.g., Loschky & McConkie, 2002). Such a determination would allow one to estimate the useful information for selecting saccade targets in natural scenes.

The limits of visual resolution and eye movements

The current study helps create a bridge between the literatures on spatial vision and eye movement control in the context of scene perception. A number of earlier studies have shown that image degradation of one kind or another lengthens fixation durations (Loftus, 1985; Loftus et al., 1992; Loschky & McConkie, 2002; Mannan et al., 1995; Parkhurst et al., 2000). However, the current study adds to an understanding of what determines such durations by relating these oculomotor effects to a model of visual sensitivity across the field of view. The results suggest a close correspondence between the conscious perceptibility of image degradation, when filtering goes beyond the limits of visual resolution, and the duration of fixations.

Comparison of detection rates and fixation durations. To provide a robust test of the limits of visual resolution in dynamic scene viewing, we used two very different measures, one explicit, blur detection, and the other implicit, fixation durations, and we found a similar pattern of results for both. One

interpretation of this similarity is that the networks of brain areas involved in both (1) making a detection response in our task, and (2) deciding when to move the eyes, may rely on the same set of low-level processes sensitive to the spatial frequency content of images, perhaps in the primary visual cortex (Itti, Koch, & Braun, 2000).

An alternative explanation for the similar pattern of responses for detection and fixation duration measures is that the one has influenced the other. Specifically, it is possible that making a manual response may have interfered with making saccades, thus increasing fixation durations whenever the filtering was detectable. This effect might help explain the chief difference between the results for the two measures, namely a significant difference in detection between the threshold filtering level ($E_2 = 1.55^\circ$) and the next higher level ($E_2 = 0.78^\circ$), but no such difference for fixation durations. If the manual detection responses lengthened fixation durations, it would reduce the variability of the fixation durations in all conditions having detectable filtering. Since our results are consistent with the idea that manual responses lengthened fixation durations (discussed below), we cannot exclude this explanation for the similar detection and fixation duration results. Note, however, that the two above explanations are not necessarily mutually exclusive. As suggested by our analyses of the time course of processes in detecting image filtering, there appear to be separate effects of filtering and manual response on fixation durations.

The time course of processes in detecting image filtering

The current study helps understand the time course of perceptual and oculomotor processes invoked by removing information beyond the limits of resolution, and making a manual response to it.

The effect of fixation durations on filtering detection rates. The current results suggest that fixation durations do affect perception of filtering, and in a way that is compatible with research on stimulus duration effects. Beyond a minimum effective fixation duration, we see a “liftoff effect” on detection rates similar to that found in experiments on the effects of stimulus duration (Bundesen & Harms, 1999; Loftus & Ruthruff, 1994). Furthermore, as the degree of image filtering is reduced to the detection threshold, the minimum effective fixation duration increases and the slope of the detection \times fixation duration function becomes shallower, again consistent with earlier research showing tradeoffs between stimulus contrast, duration, and visual performance (Loftus & Ruthruff, 1994). Nevertheless, our minimum effective fixation durations may be longer than the minimum effective stimulus durations in a static presentation, due to postsaccadic suppression, which can last for 20–50 ms from the start of a fixation (Diamond, Ross, & Morrone, 2000; McConkie &

Loschky, 2002; Shioiri, 1993). Interestingly, for the most detectable filtering levels ($E_2 = 0.78\text{--}0.39^\circ$), the minimum effective fixation duration is close to what many researchers consider the minimum fixation duration (~ 100 ms), and asymptotic near-perfect detection rates are achieved at normal-to-short fixation durations (~ 200 ms). Thus, the typical distribution of fixation durations in scene viewing seems well suited to perceiving the above-threshold spatial frequency profile of images. If other basic perceptual processes have similar minimum effective fixation durations of roughly 100–200 ms, then it might explain why fixation durations have sometimes been argued to have little impact on perception and memory (Loftus, 1972), since most variation in fixation durations will be for greater durations. However, due to the small number of observations in our data for these very short fixation durations, these proposals require further verification.

The effect of filtering detectability on fixation durations. The analysis of fixation duration hazard curves showed that filtering began lengthening fixations from roughly 160–200 ms onward for all detectably degraded conditions. For the most detectable filtering levels ($E_2 = 0.78\text{--}0.39^\circ$), this oculomotor response was somewhat slower than the minimum effective fixation duration of 80–120 ms, suggesting a delay between the information integration and eye movement response processes. However, in the filtering detection threshold condition ($E_2 = 1.55^\circ$), this delay was not found, most likely because the minimum effective fixation duration for detection was longer, about 200 ms, allowing the oculomotor system more time to respond to the filtering. This result suggests that there are rather fixed delays in registering peripheral degradation in the oculomotor system, which are somewhat independent of information thresholds for consciously registering image degradation.

The effect of a manual response on fixation durations. The results of the current study suggest that, if a fixation is long enough to be close in time to a manual response, it will interfere with executing a saccade, thus lengthening the fixation. Thus, motor programming for a manual response may interfere with that for a saccade. Furthermore, this interference can be long lasting and spill over to the next fixation. Thus, whatever interference is caused by both the manual response and the image degradation is not immediately erased by making an eye movement, but instead can persist for several hundred milliseconds. Such motor interference between hands and eyes could occur in the presupplementary motor area and/or the supplementary eye field, which have been shown in monkeys to be involved in both eye and hand movements (Fujii, Mushiake, & Tanji, 2002).

Further research is needed to investigate the role of response interference from hands to eyes. Indeed we have conducted a second experiment with the same participants using the same levels of image filtering but without a

detection task or manual response (Loschky, 2003, Exp. 3). That experiment showed a similar pattern of increased fixation durations as a function of filtering detectability, but the effects, while statistically significant, were roughly a level of magnitude smaller than those in the current experiment, consistent with the claim that the manual response in the current experiment caused longer fixations. However, the second experiment presented a multiresolutional image on every fixation, instead of only on occasional fixations, as in the current study. Thus, a future study might compare the effects of an occasional gaze-contingent multiresolutional display with and without a detection task and manual response.

Theoretical accounts of effects on fixation durations

There are two general approaches to accounting for the onset times of saccades. The first, often referred to as a cognitive account, assumes that the programming of a saccade is initiated when some critical cognitive event occurs (Henderson, 1993, 1996; Henderson & Ferreira, 1993; Rayner, 1998). In reading, this event may be the completion of word identification (Morrison, 1984) or the assessment of the familiarity of a word (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, 2003); in pictures, it may be the identification of an object (Loftus, 1981).

The second, sometimes referred to as an oculomotor inhibition account, assumes that saccades are generated at random intervals, with the average times controlled by parameters, and that they can be inhibited if processing difficulty is encountered (McConkie & Dyre, 2000; Yang, 2002; Yang & McConkie, 2001, 2004). The current results suggest that the cognitive account needs to be modified in some way to allow extrafoveal information on the current fixation to influence the saccade onset time (e.g., Kennedy, 1998, 2000); either by allowing characteristics of more peripheral information to influence the rate of foveal processing in order to delay the saccade-triggering event, or by allowing extrafoveal information to affect saccade onset time in some more direct fashion. The same is true for the effect of motor response interference such as pressing a button; this action must either compete for processing resources and thus delay the critical saccade-triggering processing event, or it must delay saccades in some other more direct fashion (Fischer, 1999).

Based on the current results, the oculomotor inhibition account would need a minor modification such that either the lack of normally available spatial frequencies is directly detected, or their lack creates difficulty in perception, either of which causes inhibition of saccadic activity. A more interesting modification of the account would be to allow for oculomotor inhibition caused by motor activity in a different effector modality (e.g., the hands), and for the inhibition produced in association with making a motor response to take some time after the response (several hundred ms or more) to dissipate. In both cases, the degree and duration of the inhibition would be revealed by fixation duration hazard

curves. The hazard curves in the current study show similarities to hazard curves from abnormal stimulus patterns encountered during reading (Yang, 2002; Yang & McConkie, 2001; Yang & McConkie, 2004), with the effect of the stimulus abnormality (e.g., a drop of hazard level below that of the control condition) only appearing after 160–200 ms after the onset of the fixation in both data sets. An interesting question for the current study is whether these effects reflect disruption of low-level perceptual processes, or whether the effect is due to a reduced ability to identify objects. Thus, the current study provides information useful for further developing current theories of eye movement control.

REFERENCES

- Anderson, S. J., Mullen, K. T., & Hess, R. F. (1991). Human peripheral spatial resolution for achromatic and chromatic stimuli: Limits imposed by optical and retinal factors. *Journal of Physiology (London)*, *442*, 47–64.
- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62–70.
- Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America*, *5*(12), 2210–2219.
- Banks, M. S., Sekuler, A. B., & Anderson, S. J. (1991). Peripheral spatial vision: Limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America*, *8*(11), 1775–1787.
- Beard, B. L., Levi, D. M., & Klein, S. A. (1997). Vernier acuity with non-simultaneous targets: The cortical magnification factor estimated by psychophysics. *Vision Research*, *37*(3), 325–346.
- Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception and Psychophysics*, *62*(3), 576–585.
- Biederman, I., Rabinowitz, J., Glass, A., & Stacy, E. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*, 597–600.
- Boyce, S., & Pollatsek, A. (1992). Identification of objects in scenes: The role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 531–543.
- Boyce, S., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 556–566.
- Braje, W. L., Tjan, B. S., & Legge, G. E. (1995). Human efficiency for recognizing and detecting low-pass filtered objects. *Vision Research*, *35*(21), 2955–2966.
- Bundesen, C., & Harms, L. (1999). Single-letter recognition as a function of exposure duration. *Psychological Research*, *62*(4), 275–279.
- Campbell, F. W. (1983). Why do we measure contrast sensitivity? *Behavioural Brain Research*, *10*(1), 87–97.
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, *197*, 551–566.
- Cannon, M. W. (1985). Perceived contrast in the fovea and periphery. *Journal of the Optical Society of America*, *2*(10), 1760–1768.
- Crane, H. D., & Steele, C. S. (1978). Accurate three-dimensional eyetracker. *Applied Optics*, *17*, 691–705.

- Deubel, H. (1984). *Wechselwirkung von sensorik und motorik bei sakkadischen augenbewegungen: Eine kybernetische beschreibung des visuellen systems*. Unpublished doctoral dissertation, Technical University of Munich, Germany.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837.
- DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.
- Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *Journal of Neuroscience*, *20*(9), 3449–3455.
- Fischer, M. H. (1999). An investigation of attention allocation during sequential eye movement tasks. *Quarterly Journal of Experimental Psychology*, *52*(A), 649–677.
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*(3), 316–355.
- Fujii, N., Mushiaki, H., & Tanji, J. (2002). Distribution of eye- and arm-movement-related neuronal activity in the SEF and in the SMA and pre-SMA of monkeys. *Journal of Neurophysiology*, *87*(4), 2158–2166.
- Geisler, W. S., & Perry, J. S. (1998). A real-time foveated multi-resolution system for low-bandwidth video communication. *Proceedings of the SPIE: The International Society for Optical Engineering*, *3299*, 294–305.
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, *39*(21), 3537–3560.
- Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research*, *38*(1), 125–137.
- Henderson, J., Pollatsek, A., & Rayner, K. (1987). The effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 449–463.
- Henderson, J., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics*, *45*(3), 196–208.
- Henderson, J. M. (1993). Visual attention and saccadic eye movements. In J. V. R. Gery d'Ydewalle (Ed.), *Studies in visual information processing: Vol. 4. Perception and cognition: Advances in eye movement research* (pp. 37–50). Amsterdam: North-Holland/Elsevier Science Publishers.
- Henderson, J. M. (1996). Visual attention and the attention-action interface. In K. A. Akins (Ed.), *Perception* (Vol. 5, pp. 290–316). New York: Oxford University Press.
- Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology*, *47*(2), 201–221.
- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (Vol. xi, pp. 269–293). Oxford, UK: Elsevier Science Ltd.
- Henderson, J. M., McClure, K. K., Pierce, S., & Schrock, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception and Psychophysics*, *59*(3), 323–346.
- Henderson, J. M., Williams, C. C., Castelhana, M. S., & Falk, R. J. (2003). Eye movements and picture processing during recognition. *Perception and Psychophysics*, *65*(5), 725–734.
- Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics*, *57*(6), 787–795.
- Hollingworth, A., & Henderson, J. M. (1999). Object identification is isolated from scene semantic constraint: Evidence from object type and token discrimination. *Acta Psychologica*, *102*(2–3), 319–343.
- Hollingworth, A., Williams, C. C., & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin and Review*, *8*(4), 761–768.

- Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, 5(1–2), 127–155.
- Itti, L., Koch, C., & Braun, J. (2000). Revisiting spatial vision: Toward a unifying model. *Journal of the Optical Society of America A—Optics and Image Science*, 17(11), 1899–1917.
- Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–179). Amsterdam: Elsevier Science.
- Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology A*, 53A(2), 429–455.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35(13), 1897–1916.
- Levi, D. M. (1999). Progress and paradigm shifts in spatial vision over the 20 years of ECVF. *Perception*, 28(12), 1443–1459.
- Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception and Psychophysics*, 63(4), 676–697.
- Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14), 9596–9601.
- Loftus, G., Kaufman, L., Nishimoto, T., & Ruthruff, E. (1992). Effects of visual degradation on eye-fixation durations, perceptual processing, and long-term visual memory. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 203–226). New York: Springer-Verlag.
- Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 3(4), 525–551.
- Loftus, G. R. (1981). Tachistoscopic simulations of eye fixations on pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 7(5), 369–376.
- Loftus, G. R. (1985). Picture perception: Effects of luminance on available information and information-extraction rate. *Journal of Experimental Psychology: General*, 114(3), 342–356.
- Loftus, G. R., & Ruthruff, E. (1994). A theory of visual information acquisition and visual memory with special application to intensity/duration trade-offs. *Journal of Experimental Psychology: Human Perception and Performance*, 20(1), 33–49.
- Loschky, L. C. (2003). *Investigating perception and eye movement control in natural scenes using gaze-contingent multi-resolutional displays*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.
- Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multi-resolutional display. *Journal of Experimental Psychology: Applied*, 8(2), 99–117.
- Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science*, 3, 67–68.
- Maekelae, P., Naesaenen, R., Rovamo, J., & Melmoth, D. (2001). Identification of facial images in peripheral vision. *Vision Research*, 41(5), 599–610.
- Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision*, 9(3), 363–386.
- McConkie, G. W., & Dyre, B. P. (2000). Eye fixation durations in reading: Models of frequency distributions. In A. Kennedy & R. Radach (Eds.), *Reading as a perceptual process* (pp. 683–700). Amsterdam: Elsevier Science Ltd.
- McConkie, G. W., & Loschky, L. C. (2002). Perception onset time during fixations in free viewing. *Behavioural Research Methods, Instruments, and Computers*, 34(4), 481–490.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, 17(6), 578–586.
- Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), 667–682.

- Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning and Memory*, 6(4), 391–399.
- Olds, E. S., & Engel, S. A. (1998). Linearity across spatial frequency in object recognition. *Vision Research*, 38(14), 2109–2118.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, 41(2), 176–210.
- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3), 145–175.
- Parkhurst, D., Culurciello, E., & Niebur, E. (2000). Evaluating variable resolution displays with visual search: Task performance and eye movements. In A. T. Duchowski (Ed.), *Proceedings of the Eye Tracking Research and Applications symposium* (pp. 105–109). New York: Association for Computing Machinery.
- Parkhurst, D. J., & Niebur, E. (2002). Variable-resolution displays: A theoretical, practical, and behavioural evaluation. *Human Factors*, 44(4), 611–629.
- Peli, E., & Geri, G. A. (2001). Discrimination of wide-field images as a test of a peripheral-vision model. *Journal of the Optical Society of America A—Optics and Image Science*, 18(2), 294–301.
- Peli, E., Yang, J., & Goldstein, R. B. (1991). Image invariance with changes in size: The role of peripheral contrast thresholds. *Journal of the Optical Society of America*, 8(11), 1762–1774.
- Pointer, J. S., & Hess, R. F. (1989). The contrast sensitivity gradient across the human visual field: With emphasis on the low spatial frequency range. *Vision Research*, 29(9), 1133–1151.
- Rao, R. P. N., Zelinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research*, 42(11), 1447–1463.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7(1), 65–81.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105(1), 125–157.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, 39(26), 4403–4411.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioural and Brain Sciences*, 26(4), 445–526.
- Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multi-resolutional displays: An integrative review. *Human Factors*, 45(2), 307–328.
- Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21(3), 409–418.
- Sanocki, T. (2003). Representation and perception of spatial layout. *Cognitive Psychology*, 47(1), 43–86.
- Sanocki, T., & Epstein, W. (1997). Priming spatial layout of scenes. *Psychological Science*, 8(5), 374–378.
- Schyns, P., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5, 195–200.
- Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of view. *Experimental Aging Research*, 26(2), 103–120.
- Sere, B., Marendaz, C., & Herault, J. (2000). Nonhomogeneous resolution of images of natural scenes. *Perception*, 29(12), 1403–1412.
- Shioiri, S. (1993). Postsaccadic processing of the retinal image during picture scanning. *Perception and Psychophysics*, 53(3), 305–314.
- Shioiri, S., & Ikeda, M. (1989). Useful resolution for picture perception as a function of eccentricity. *Perception*, 18, 347–361.

- Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Buelthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience*, *14*(5), 869–876.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.
- Van Diepen, P. M. J., & Wampers, M. (1998). Scene exploration with Fourier-filtered peripheral information. *Perception*, *27*(10), 1141–1151.
- Walker, L. L., & Malik, J. (2002). When is scene recognition just texture recognition? *Journal of Vision*, *2*(7). Abstract no. 255a. Retrieved from <http://journalofvision.org/2/7/255/>
- Wilson, H. R., Levi, D., Maffei, L., Rovamo, J., & DeValois, R. (1990). The perception of form: Retina to striate cortex. In L. Spillmann, J. S. Werner et al. (Eds.), *Visual perception: The neurophysiological foundations* (pp. 231–272). San Diego, CA: Academic Press, Inc.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202–238.
- Yang, J., Coia, T., & Miller, M. (2001). Subjective evaluation of retinal-dependent image degradations. In *Proceedings of PICS 2001: Image Processing, Image Quality, Image Capture Systems conference* (pp. 142–147). Springfield, VA: Society for Imaging Science and Technology.
- Yang, J., Qi, X., & Makous, W. (1995). Zero frequency masking and a model of contrast sensitivity. *Vision Research*, *35*(14), 1965–1978.
- Yang, J., & Stevenson, S. B. (1998). Effect of background components on spatial-frequency masking. *Journal of the Optical Society of America*, *15*(5), 1027–1035.
- Yang, S.-N. (2002). *Inhibitory control of saccadic eye movements in reading: A neurophysiologically based interaction-competition theory of saccade programming*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.
- Yang, S. N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research*, *41*(25–26), 3567–3585.
- Yang, S. N., & McConkie, G. W. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology*, *16*(1–2), 226–261.
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, *39*(2), 293–306.