

Hybrid Image/Model-Based Gaze-Contingent Rendering

HUNTER A. MURPHY, ANDREW T. DUCHOWSKI, and RICHARD A. TYRRELL
Clemson University

A nonisotropic hybrid image/model-based gaze-contingent rendering technique utilizing ray casting on a GPU is discussed. Empirical evidence derived from human subject experiments indicates an inverse relationship between a peripherally degraded scene's high-resolution inset size and mean search time, a trend consistent with existing image-based and model-based techniques. In addition, the data suggest that maintaining a target's silhouette edges decreases search times when compared to targets with degraded edges. However, analysis suggests a point of diminishing returns with an inset larger than 15° when target discrimination is a component of visual search. Benefits of the hybrid technique include simplicity of design and parallelizability, both conducive to GPU implementation.

Categories and Subject Descriptors: I.3.3 [**Computer Graphics**]: Picture/Image Generation—*Display algorithms*; I.3.6 [**Computer Graphics**]: Methodology and Techniques—*Ergonomics*

General Terms: Gaze-Contingent Display

Additional Key Words and Phrases: Eye tracking, Level of Detail

ACM Reference Format:

Murphy, H. A., Duchowski, A. T., and Tyrrell, R. A. 2009. Hybrid image/model-based gaze-contingent rendering. ACM Trans. Appl. Percept. 5, 4, Article 22 (January 2009), 21 pages. DOI = 10.1145/1462048.1462053 <http://doi.acm.org/10.1145/1462048.1462053>

1. INTRODUCTION

The main contribution of this paper is the development of a hybrid image/model-based nonisotropic gaze-contingent rendering technique. This method, based on ray casting, is particularly well suited to implementation on modern programmable Graphics Processing Units, or GPUs. The technique is meant for the investigation of the effects of spatial peripheral degradation of arbitrary geometric objects, e.g., meshes, without the need for mesh reparameterization. A demonstration of this approach to gaze-contingent rendering is given by degrading meshes as a function of foveal eccentricity in relation to the dynamically translating point of regard, as measured by a real-time (50 Hz) eye tracker. Results from a human subject experiment involving a visual search task are compared to previous gaze-contingent rendering experiments.

The idea of gaze-contingent rendering is not new, with military applications assuming some of the earliest instances [Duchowski 2007]. Significant progress has been made in exploring both the extent of peripheral degradation permitted by the Human Visual System (HVS) and the algorithms used to perform that degradation. The field has become dominated by two approaches in particular:

Authors' addresses: H. Murphy, and A. Duchowski, Clemson University, School of Computing, 100 McAdams Hall, Clemson, SC 29634, USA; R. Tyrrell, Clemson University, Dept. of Psychology, 418 Brackett Hall, Clemson, SC 29634, USA.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2009 ACM 1544-3558/2009/01-ART22 \$5.00 DOI 10.1145/1462048.1462053 <http://doi.acm.org/10.1145/1462048.1462053>

image-based rendering and model-based rendering. This paper introduces a third technique which combines the strengths of both image-based and model-based rendering in a manner conducive to implementation on modern graphics hardware.

Our empirical results suggest a potential benefit of the gaze-contingent display to decreasing visual search time, but only when the gaze-contingent high-resolution window is restricted to a certain extent ($10\text{--}15^\circ$) and when the search task involves parafoveal discrimination. The experiment conducted was designed with the intention of replicating previous findings. Although we were able to replicate the well-known inversely proportionate relationship between search time and gaze-contingent window size, the apparent benefit of the mid-sized gaze-contingent window was unexpected.

The paper is organized as follows. Section 2 explores the two divergent approaches taken in gaze-contingent rendering. The unification of these disparate techniques is discussed in Section 3. Details of the visual search experiment are presented in Section 4, with results and discussion appearing in Sections 5 and 6, respectively. Conclusions appear in Section 7.

2. BACKGROUND

Gaze-contingent rendering exploits the human visual system's reduction of peripheral sensitivity to several visual attributes with the goal of improving rendering performance. The measurement of task performance can be ambiguous, but usually falls into two categories: visually imperceptible degradation and functionally imperceptible degradation.

Visually imperceptible degradation attempts to reduce peripheral information in a manner that cannot be distinguished from a nondegraded version. This stringent approach requires strict adherence to all aspects of the HVS simultaneously. Use of visual imperceptibility as a performance metric is typically applied to evaluation of the HVS itself; e.g., see Loschky et al. [2005].

Functionally imperceptible degradation takes a more practical approach by interpreting imperceptible degradation to mean that the reduction in peripheral information does not result in reduced task performance. This less rigid stance allows much greater flexibility in manipulating the stimulus as it is focused on exploiting, rather than deriving, fundamental HVS knowledge. This situation is analogous to *functionally realistic* computer graphics images, which provide the same visual information as the scene they are modeling, rather than producing the same visual response expected by *photo-realistic* representations [Ferwerda 2003]. The goal of achieving functionally imperceptible rendering requires that visually imperceptible rendering, such as achieved by strict adherence to the Human Visual System's Contrast Sensitivity Function (CSF), e.g., as demonstrated by Loschky et al. [2005] and replicated by Loschky and Wolverton [2007], be mimicked at a level of degradation discernible by the viewer. The CSF is thus only used as a starting point for a degradation function that is meant to minimize the impact of the necessarily visible degradation (the use of the CSF is detailed below).

2.1 Image-Based GCD

An image-based Gaze-Contingent Display (GCD) is particularly well suited to exploring the perceptual limits of the HVS. Implementation typically involves application of a convolution filter to a pre-rendered full resolution image. This simplifies not only the programming aspect, as 2-D filtering is relatively simple, but also the experimental aspect, since starting with an ideal image allows the experimenter to focus on manipulating a single perceptual variable of the HVS.

Sophisticated approaches have been developed for image and video coding [Parkhurst and Niebur 2002; Reingold et al. 2003; Geisler et al. 2006]. For screen-based rendering, the work of Watson et al. [1997], who studied the effects of Level Of Detail (LOD) peripheral degradation on visual search performance, is particularly relevant. Both spatial and chrominance detail degradation effects were evaluated in head mounted displays. To sustain acceptable frame rates, two polygons were texture

mapped in real-time to generate a high resolution inset within a low resolution display field. Watson et al. suggested that visual spatial and chrominance complexity can be reduced by almost half without degrading performance. More recently, Watson et al. [2004] pointed out that speed improvement may be limited if only a sub-threshold approach is used to reduce detail.

In a similar approach, Reddy [1998] used a view-dependent screen-based LOD technique to evaluate both perceptual effects and system performance gains. The author reported a perceptually modulated LOD system which provided a 4.5 factor of improvement in frame rate. It is not entirely clear how the LOD model was constructed, i.e., the method of degradation.

By filtering out high frequency information in the periphery, image-based GCDs can also offer a benefit for image compression. The increase in image compressibility gained through peripheral degradation can be exploited to improve bandwidth utilization for image transmission [Bergström 2003].

Despite recent interest (see Duchowski and Çöltekin [2007] for a review of image-based approaches), real-time computational benefits of image-based GCD are lacking. Unless used for remote display, there can be no improvement in render time when the first step is to render the scene in full resolution, and the second is to apply a full screen convolution filter. Indeed, outside its uses in psychophysical research and image transmission, image-based GCD offers no practical benefit per se. Model-based GCDs attempt to address the desire for render time efficiency.

2.2 Model-Based GCD

A model-based GCD can significantly improve render time by removing mesh geometry which is not perceived by the user. The reduction of geometric complexity is achieved either through a series of edge collapses, or vertex decimations applied to the original mesh, or through the reconstitution of a base mesh to a higher complexity form which satisfies the requirements of the HVS. Whether using a top-down or bottom-up approach, model-based GCD rendering typically requires significant processing, either as a preprocessing step or at run-time, but can provide a net speedup in rendering time [Luebke and Erikson 1997].

Simplification of geometric objects to reflect perceptual limits is a widely used technique, dating back to Clark's [1976] description of LOD rendering, wherein an object's screen coverage was used as a metric to select one of several precomputed reduced resolution meshes for display. This technique is currently used in applications ranging from animation rendering to virtual reality; however, the various approaches used to generate the reduced resolution meshes typically result in isotropic, or uniform, object simplification.

LOD rendering has also been used in model-based GCD. Ohshima et al. [1996] proposed a scheme considering three visual characteristics: central/peripheral vision, kinetic vision, and fusional vision. The LOD algorithm generated isotropically degraded objects at different visual angles. Although the use of a binocular eye tracker was proposed, the system as discussed used only head tracking as a substitute for gaze tracking.

Traditional LOD is clearly beneficial when rendering for a standard display paradigm; the use of isotropic object degradation for GCD is suboptimal. Uniform mesh degradation assumes uniform perceptual discrimination across the entire field of view. In this case, traditional LOD schemes will display an LOD mesh at its full resolution even though the mesh may cover the entire field of view. Since acute resolvability of human vision is limited to the foveal 5°, object resolution need not be uniform within a gaze-contingent context. This is the central tenet of gaze-contingent systems.

An alternative to isotropic LOD, multiresolution mesh modeling techniques suitable for gaze-contingent viewing have been developed [Zorin and Schröder 2000]. Multiresolution modeling allows the possibility of rendering meshes that are nonisotropically degraded by selecting the geometry to display from a hierarchy of meshes based on the angle subtended from the Point Of Regard (POR).

Techniques range from multiresolution representation of arbitrary meshes to the management of LOD through peripheral degradation within an HMD, where gaze position is assumed to coincide with head direction [Lindstrom et al. 1996; MacCracken and Joy 1996; Hoppe 1997; Zorin et al. 1997; Schmalstieg and Schaufler 1997].

Danforth et al. [2000] developed a nonisotropic gaze-contingent multiresolution terrain navigation environment. A surface, represented as a quadrilateral mesh, was divided into fixed-size (number of vertexes) subblocks, allowing rendering for variable LOD on a per-subblock basis. Resolution level was chosen per subblock, based on viewer distance. The resolution level was not discrete; it was interpolated between the precomputed discrete levels to avoid “popping” effects. The approach used was reasonably effective, but it is not clear whether the technique is applicable to arbitrary meshes.

Rather than decimate existing geometry, Murphy and Duchowski [2001] constructed a mesh in real-time from a base mesh. Distance from the POR indicated how much geometry was to be added to the mesh, with additions being made from a hierarchy of intermediate forms.

Unfortunately, model-based GCDs lack the simplicity of implementation inherent to image-based GCDs. Reconstruction of mesh geometry, either through decimation of the original mesh, or reconstitution of a base mesh, is an issue of local mesh connectivity; however, maintaining the global perceptual requirements makes the utilization of model-based GCD a nontrivial task. Moreover, model-based LOD manipulation may not be applicable to arbitrary meshes and may require constraints on mesh connectivity (e.g., 2-manifold) or mesh reparameterization [Luebke et al. 2000].

In an effort to provide image-based GCD’s ease of implementation, while offering hope of render time efficiencies corresponding to a model-based GCD, a hybrid approach is discussed in the following section. The technique is similar to an early implementation by Levoy and Whitaker [1990], in which the authors ray traced volumetric data in a perceptually adaptive manner. Although rays were not cast in accordance with any particular HVS function, the number of rays cast increased at the POR and decreased with increasing visual angle according to a Gaussian function. The resulting samples were used to index a 2-D mipmap of images and combined to construct the final image.

3. HYBRID IMAGE/MODEL-BASED GCD

The goal of the hybrid technique is to nondestructively sample scene geometry in a manner consistent with the limits of the HVS. This is achieved using ray casting, with ray distribution conforming to the angular frequency dictated by a Contrast Sensitivity Function (CSF), although any decaying degradation function may be used for different effects. The ray casting and CSF combination allows nonisotropic (within mesh) degradation without directly manipulating mesh geometry (edge collapse or vertex decimation). An intermediate mesh existing between the eye position and the scene geometry provides both a direction vector for the rays and storage for the resulting ray/primitive intersection data. Further refinement of the intermediate mesh is performed to maintain silhouette edges. The intermediate mesh is ultimately rendered in place of the scene geometry.

3.1 Contrast Sensitivity Function

A contrast sensitivity function attempts to describe the amount of detail visible at increasing angular separation from the point of regard (e.g., see Geisler and Perry [1998]; Luebke and Hallen [2001]; Loschky et al. [2005]). Contrast is a difference in luminance, typically the difference in reflected light levels between adjacent triangles. A CSF is usually expressed in units of cycles per degree, which in the case of the hybrid technique refers to the number of pixels that can be discerned at a particular distance from the POR.

Creation of a perceptually based intermediate mesh which defines points through which rays are cast is achieved using a discretized approximation of an empirically derived contrast sensitivity function.

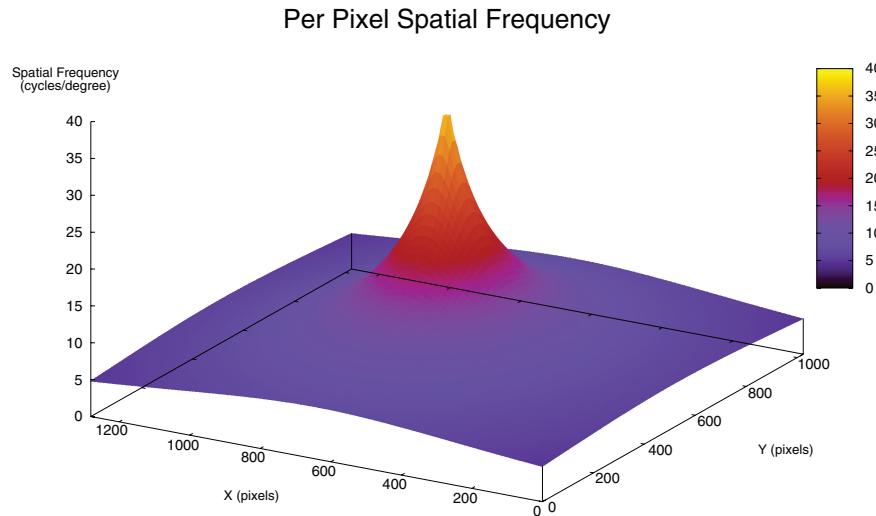


Fig. 1. Spatial frequency with POR at center of screen.

This is created using a rearrangement of the contrast threshold function explored by Geisler and Perry [1998]. The derivation starts with the contrast threshold equation:

$$CT(f, e) = CT_0 \exp\left(\alpha f \frac{e + e_2}{e_2}\right), \quad (1)$$

where f is the spatial frequency in cycles per degree, e is the visual angle with respect to the POR (eccentricity), CT_0 is the minimum contrast threshold, α is the spatial frequency decay constant, and e_2 is the half-resolution eccentricity. Geisler and Perry determined values of $CT_0 = 1/64$, $\alpha = 0.106$, and $e_2 = 2.3$, empirically.

By setting the left hand side of (1) to the maximum contrast possible, i.e., 1 (unity), and rearranging terms we obtain: $e = (e_2/\alpha f) \ln(1/CT_0) - e_2$. The resulting function provides the eccentricity at which a given spatial frequency f no longer contributes to perception. In order to construct a spatial frequency map for a monitor's display it is necessary to compute e at coordinates (x, y) . Rearranging terms again gives

$$f(x, y) = \frac{e_2 \ln(1/CT_0)}{\alpha(e(x, y) + e_2)}. \quad (2)$$

The resulting spatial frequency map is shown in Figure 1. This map guides the sampling rate of the intermediate mesh by dictating the minimum cycles per degree at a particular point on the screen. The map, however, is only used as a starting point for degradation—other variants (e.g., based on minimum angle of resolution or visual acuity), may also be suitable.

It is necessary to discretize (2) along pixel alignment in order to generate the intermediate mesh. While a fractional pixel mapping would provide a closer representation mathematically, the visual impact was deemed unacceptable. A set of equations was established to represent the spatial frequency resulting from casting a ray through a monotonically increasing set of pixel locations (every pixel, every two pixels, every three pixels, etc.) The equations are given by $PPD(n) = \tan(1.0)/n$, $n \in [1..17]$. Calculating the intersections between this set and (2) provides a series of thresholds dictating pixel spacing along an axis. A 1-D array is populated by inserting every integer from zero to the first threshold

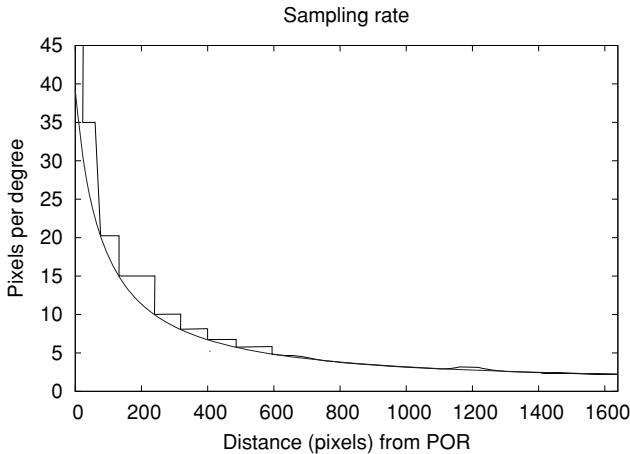


Fig. 2. Plot of minimum sampling rate derived from the CSF, with superimposed discretized version used to guide ray casting (at 53 cm viewing distance).

point, then every second integer from the first threshold to the second threshold, and so forth until a value in the 1-D array exceeds the maximum screen extent. Figure 2 illustrates this process, showing the smooth spatial frequency equation given by (2) and the sampling rate used in the ray mask.

Given the 1-D array of pixel locations, the final 2-D intermediate mesh is generated by rotating the pixel locations by some angular displacement. The displacement is determined by calculating the angular separation between the two outermost entries in the 1-D array. This angular displacement is used to rotate the points, whose positions are recorded in the 2-D ray mask, until a full 360° is attained (see Figure 3). All pixels within the single pixel separation distance indicated by the first threshold encountered in the 1-D array are filled to mask any discretization errors caused by rotation.

The intermediate mesh is completed by connecting concentric circles of rays outside the contiguous region to form quad strips.

3.2 Ray Casting

Ray casting, a subclass of ray tracing, is a rendering technique that tests for intersections between a ray and primitives in the scene. In order to achieve efficient render times for any nontrivial scene it is necessary to reduce the number of ray/primitive intersection tests. The most obvious algorithmic approach is to use a ray casting acceleration structure, which places boundaries on the search space to be traversed in determining if an intersection has occurred. Modern ray casting acceleration structures fall into two bodies of solutions: Bounding Volume Hierarchies (BVHs) and spatial subdivision (see Purcell et al. [2002] as well as Popov et al. [2007]). In the present implementation, a uniform grid spatial subdivision structure is used, selected for its suitability for the type of scene displayed in the demonstration experiment and because of the type of GPU hardware used in development.

Uniform grid representations partition space into a 3-D array of identical rectangular prisms. Intersection tests traverse the array in much the same way that a line drawing algorithm traverses a 2-D array of pixels. The traversal algorithm used involved simple repetitive arithmetic, which allowed optimization of both GPU memory usage and GPU computational units.

A uniform grid was constructed for each mesh by iteratively increasing the number of subdivisions in the mesh bound box until no cell exceeded an empirically derived binning size (40 triangles per cell worked well for the meshes used in the experiment). Primitives that spanned more than one cell were

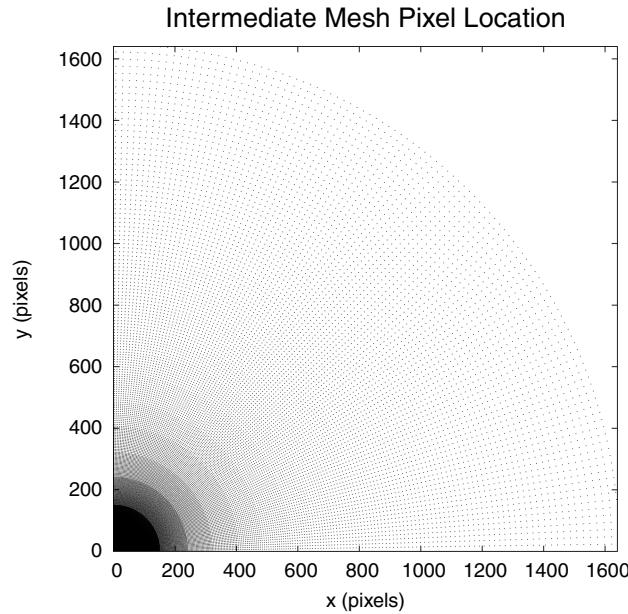


Fig. 3. Intermediate mesh (2-D ray mask) used for ray casting.

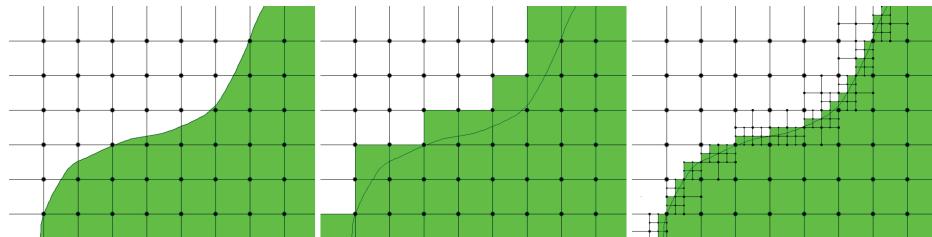


Fig. 4. Steps in adaptive construction of an intermediate mesh.

placed in all intersected cells. Pertinent information, such as bounding box and cell dimensions, was stored as part of the uniform grid structure.

3.3 Hybrid Rendering

With both the CSF-based ray mask and uniform grid acceleration structure in place, application of the hybrid technique proceeds by casting rays from the eye point through each pixel in the ray mask. The rays pass through cells in the uniform grid, where ray/primitive intersection tests are performed for all primitives present in the cell. If an intersection occurs the information is stored in the intermediate mesh and the ray's progress is halted. The process is repeated until all rays have been cast.

Following the initial scene sampling, the intermediate mesh is searched for quads with heterogeneous intersection data, described by a mix of intersected and nonintersected corners (see Figure 4). Heterogeneous quads indicate the presence of an external silhouette edge and are rendered by casting rays through all contained pixels. Homogeneous quads are rendered using standard rasterization.

Ray casting on the GPU was accomplished by reformatting the mesh geometry, intermediate mesh, and uniform grid acceleration structures to fit in OpenGL textures. Custom OpenGL 2.0 vertex and

fragment shaders were used to perform ray casting, subsequent to which the framebuffer contained normal information for each point in the intermediate mesh. Due to GPU architectural constraints, the framebuffer was copied into system memory (incurring a substantial penalty in framerate), where the mapping between framebuffer and intermediate mesh was completed. The complete intermediate mesh was then rendered and displayed to the subject.

4. DEMONSTRATION OF THE HYBRID GCD IN A VISUAL SEARCH TASK

The demonstration of the GCD was motivated by prior gaze-contingent studies. The goal of the experiment was therefore to replicate previous results reported from experiments that generally involved some form of visual search. To demonstrate the current rendering method's ability to vary the spatial resolution of meshes rendered at varying eccentricity, geometric meshes of faces were used. The task given to participants, however, called for target *identification*, beyond its mere *detection*.

We have previously reported that search tasks may benefit from our hybrid gaze-contingent display when discrimination is involved [Murphy and Duchowski 2007]. Our conclusions were based on the notion of narrowing the display to the extent of the viewer's "attentional spotlight" [Posner et al. 1980], thus aiding the viewer in limiting the number of features needed for target identification. One troubling aspect of the attentional spotlight hypothesis is that the mean search time should increase as the high resolution window increases to 15°, the first point at which all distractors would be subtended peripherally. This effect was not previously observed.

4.1 Method

In keeping with previous gaze-contingent visual search findings, it was hypothesized that varying the size of a high-resolution inset at the POR, while reducing peripheral detail in accordance with the CSF, would affect localization time in a visual search task. Moreover, it was believed that maintaining silhouette edges would provide a performance benefit for target localization due to the reduction of peripheral noise. Therefore, no significant difference was expected between search times among the factors under consideration (preservation of silhouette edges and reduction of foveal window size).

4.2 Equipment

A Tobii ET-1750 video-based binocular eye tracker was used for real-time gaze sampling. The Tobii samples at 50 Hz with an accuracy typically better than 0.3° over a ±20° horizontal and vertical range using the pupil/corneal reflection difference [Tobii Technology AB 2003] (in practice, measurement error ranges roughly ± 10 pixels). The eye tracker's 17" LCD monitor was set to 1280×1024 resolution. The eye tracking server ran on a dual 2.0 GHz AMD Opteron 246 PC (2 G RAM) running Windows XP. Although the Tobii allows limited head movement (30×15×20 cm volume), a chin rest was used to maintain constant distance (53 cm) from the monitor.

The client display application ran on a 2.2 GHz AMD Opteron 148 Sun Ultra 20 running the CentOS operating system. The client/server PCs were connected via the departmental 1 Gb Ethernet (both connected to a switch on the same subnet). A keyboard attached to the client system provided user input. The physical setup is shown in Figure 5 (viewer is running a demonstration program).

The client PC was equipped with an Nvidia 8800GTX GPU with 768Mb of texture memory and 128 stream processors operating at 1.35 GHz. All ray casting occurred on-GPU; however, construction of the intermediate mesh necessitated CPU involvement.

4.3 Experimental Design & Procedure

The independent variables used in this 2×5 repeated measures experiment were high-resolution inset size and the presence or absence of object edges. Inset size diameter spanned 2°, 5°, 10°, 15°, and 20°



Fig. 5. Example of equipment setup.

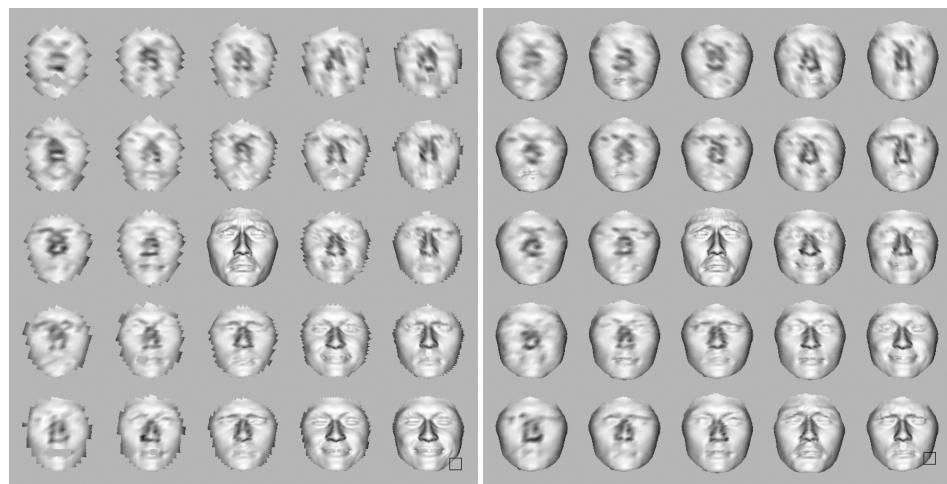


Fig. 6. Sample search task (left), sample search task with silhouette edges preserved (right), each with box in lower right indicating POR with a 5° high-resolution inset size.

visual angle. For each inset size, object edges were either preserved or discarded, resulting in 10 combinations. Addition of a control (full screen ray casting) resulted in a total of 11 scene types to be tested.

The experiment utilized a within-subjects design, with each participant exposed to the aforementioned 11 scene types. A Latin square was used to control order effects. Each subject performed the 11 search tasks four times, resulting in 44 total trials.

The dependent variable in this experiment was the visual search time required to locate a target object. The user was asked to find a replica of the centrally located target object in a field of distractors of similar appearance. Users were first familiarized with an isolated image of the target object then, when ready (no time limit imposed), the screen changed to display a 5×5 object search field with the target located at center and simultaneously at a random position in the grid of distractors. Search time started when the field of distractors was displayed and ended when the subject pressed the space bar. Participants were instructed to identify the target in the field of distractors that matched the target in the center of the screen as quickly as possible, favoring speed over accuracy. Participants were instructed to refer back to the target in the center of the screen if they thought it necessary.

4.4 Human Subjects

There were 11 participants in the experiment, four female and seven male. Subject ages ranged from 25 to 34, all with self-reported normal or corrected to normal vision. Although contact lenses were permitted, potential subjects wearing glasses were excluded from the experiment, as were others who failed the screening process (a total of two) due to the tracker's inability to track their gaze at a tolerance accepted by the experimenter.

4.5 Stimulus

A nine point calibration was used at the beginning of a session, with a scene confirming calibration accuracy displayed immediately thereafter. The calibration accuracy confirmation scene was also displayed at the midpoint of the experiment to assure that accuracy was within tolerances.

Four topologically consistent meshes displaying different facial expressions were used as targets. The original silhouette edges were consistent across all four meshes. Each target subtended approximately 3.5 degrees of visual angle at the screen distance used in the experiment. The entire scene subtended approximately 20 degrees of visual angle.

The scenes used for the visual search task consisted of a 5×5 grid of the meshes described above. The scenes were constructed by first selecting a target from the set of four meshes. Each mesh was selected at random from a pre-seeded pool, assuring that although the presentation order of each mesh was random, each mesh assumed the role of target the same number of times. The target mesh was then placed both in the center of the scene and at a random location in the grid. The remaining positions in the 5×5 grid were then filled at random with duplicates of the three nontarget meshes.

4.6 Presentation

A scripted verbal explanation describing the display program commands and execution flow was read to familiarize subjects with the demonstration program. The chair, chin rest, and keyboard were then adjusted for user comfort. At this point a demonstration program was run to reinforce the scripted explanation. The subjects were informed that the demonstration program functioned in exactly the same manner as the experimental program, with only the control stimulus being displayed. The demonstration program was also used to introduce the meshes used in the experiment. Subjects were informed that time was not a factor in the demonstration, and were encouraged to examine the different facial expressions in detail. Upon completion of the demonstration program the participants were permitted to ask questions and stretch.

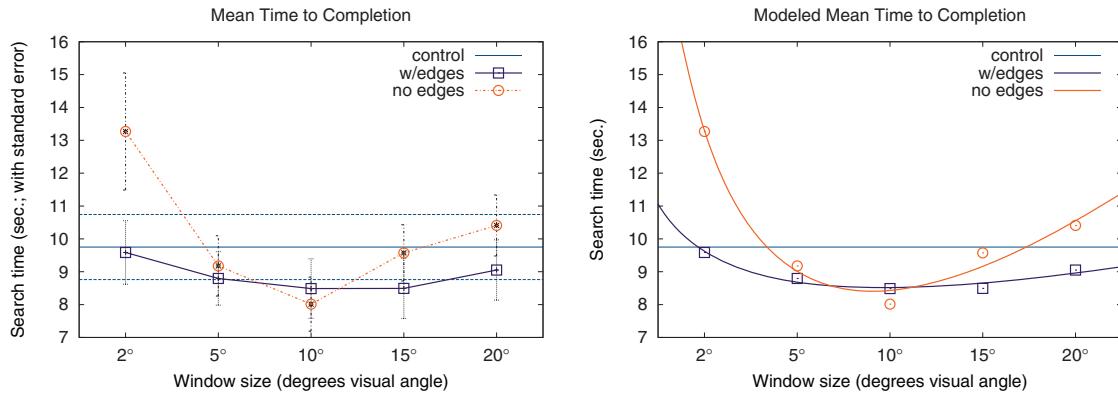


Fig. 7. Comparison of mean search times per window size condition (left), modeled by $y = ax^b + cx + d$ (right, with reduced $\chi^2 = 0.05$ and reduced $\chi^2 = 0.39$ for data with edges and no edges, resp.). Mean search time for the control condition is shown as a constant line in both graphs (with standard error range).

Subjects started the experimental program when ready. A nine point calibration was followed by a scene designed to allow the experimenter to judge the quality of the calibration. Recalibration was performed if the calibration was deemed to lack accuracy or precision. Following calibration, the target was displayed in the center of the screen. The subject was given as much time as desired to examine the target. The subject then triggered the search task, which displayed the 5×5 grid of distractors. Upon localizing the target within the grid, the participant pressed the space bar to finish the search task and load the next target. A break to check the calibration or stretch was allowed after six search tasks were completed. Upon completion of all 11 search tasks, users were encouraged to stand and stretch. The experimental program was run four consecutive times for each subject.

5. RESULTS

Recorded search time data was systematically culled with the removal of outliers falling beyond three standard deviations.

5.1 Effects on Search Time

Omitting the control condition and using edge presence and window size as fixed factors (using subject as the random factor, see Baron and Li [2007]), repeated-measures two-way ANOVA indicates a significant main effect of window size on time to completion ($F(4,40) = 4.58$, $p < 0.01$).¹ Averaging over the edge presence conditions, pair-wise t-tests with pooled SD indicate significantly slower search times with the 2° window than with the window of 10° ($p < 0.05$, with Bonferroni correction). No other significant differences were detected.

Repeated-measures ANOVA also indicates a significant main effect of edge presence on time to completion ($F(1,10) = 6.84$, $p < 0.05$), with edge \times window interaction not significant ($F(4,40) = 1.16$, $p = 0.34$, n.s.).

Plotting the mean time to completion with standard error against gaze-contingent window size, as shown in Figure 7, suggests a U-shaped performance function where an improvement is seen with a window of 10° that tends to diminish when the window size increases to 15° and beyond.

Performance data can be fit to quadratic learning curves typically created for task performance over time [Card et al. 1983]. Here we use the generalized quadratic function $y = ax^b + cx + d$. As indicated

¹Assuming sphericity as computed by R, the statistical analysis package used throughout.

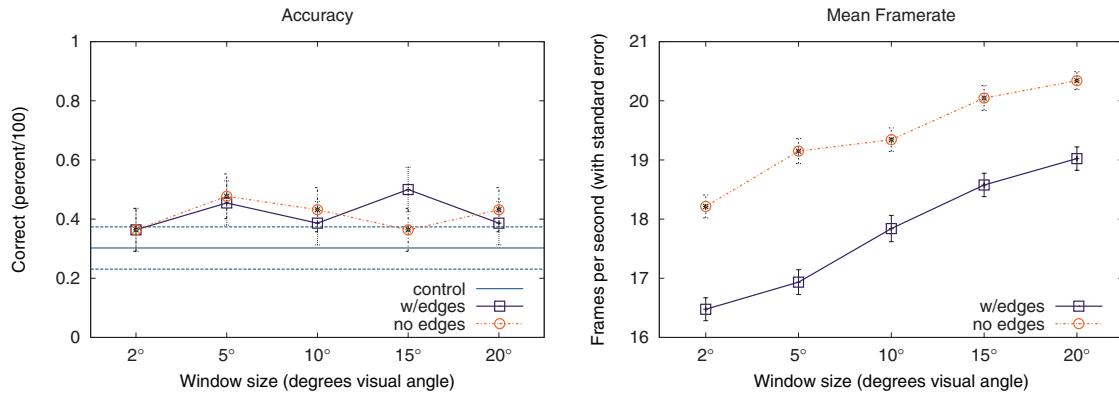


Fig. 8. Percent accuracy (left) and mean framerate (right) per window size condition. Accuracy for the control condition is shown as a constant line (with standard error range). Mean framerate for the control condition was just under 40 fps and for clarity is omitted from the graph.

in Figure 7, the curves suggest a slightly lower minimum for the display with a 10° window size. The tighter fit of the curve to the response times with silhouette edges suggests that preservation of edges is important for performance with larger foveal windows.

5.2 Effects on Accuracy and Potential Effects of Framerate

We also considered the potential effects of the GCD on accuracy and the effects of display framerate on search performance. Plausible explanations for a difference in average performance due to lack of edge information may stem from a speed-accuracy tradeoff or a faster framerate due to the smaller number of rays cast to preserve edges.

No discernible effect is seen in terms of accuracy (Figure 8 (left)), in general, although edge preservation appears to offer a slight benefit with a 15° foveal window.

Omitting the control condition and using edge presence and window size as fixed factors, repeated-measures two-way ANOVA shows no significant effect of window size ($F(4,40) = 0.80, p = 0.53$, n.s.) or of edge presence ($F(1,10) = 0.00, p = 0.99$, n.s.).

Although the effect is not significant, both gaze-contingent conditions appear to provide some accuracy benefit over the control condition. We should note that our definition of accuracy reflects a rather stringent requirement of the target stimulus being fixated during the user's key press upon target identification. The instructions given to the participants indicated that they were to press the button while fixating the target, but we cannot reasonably guarantee that participants adhered to the instructions.

To estimate the effect of gaze-contingent processing on framerate, a single experimental run was conducted with the program reinstrumented to output the instantaneous framerate (on every frame) during display. Data capture was performed by simulating gaze with the mouse, traversing over each target in the visual field alternating from left-to-right then right-to-left, from top-to-bottom in each condition. Mouse movement, and hence timing, was unpaced, resulting in an arbitrary number of total fps measurements collected (1259 in this instance) over all eleven condition combinations. The average frame rate during the control (no peripheral geometry processing) averaged just under 40 fps.

As expected, frame rate with gaze-contingent processing is on average lower than the control, see Figure 8 (right). Omitting the control and using edge presence and window size as fixed factors, two-way ANOVA indicates a significant effect of window size on framerate ($F(4,1254) = 44.70, p < 0.01$). Averaging over the edge presence conditions, pair-wise t-tests with pooled SD indicate a significant

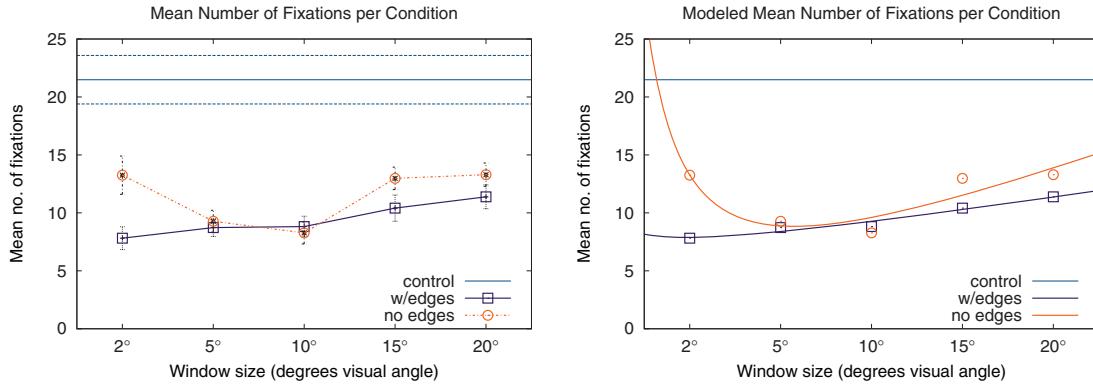


Fig. 9. Measured (left) and modeled (right) mean number of fixations per window size condition, modeled by $y = ax^b + cx + d$ (right, with reduced $\chi^2 = 642.02$ and reduced $\chi^2 = 8539.17$ for data with edges and no edges, resp.). Mean number of fixations for the control condition is shown as a constant line in both graphs (with standard error range).

difference between every pair of window sizes ($p < 0.05$, in all cases, with Bonferroni correction) except the pair of windows of 5° and 10° and the pair of windows of 15° and 20° .

Two-way ANOVA also indicates a significant effect of edge presence on framerate ($F(1,1257) = 160.55$, $p < 0.01$). Although both window size and edge processing significantly affects framerate, the mean rate never fell below 16 fps, suggesting its effect on visual search performance is negligible (display updates as late as 60 ms after eye movement completion do not significantly increase the detectability of image blur and/or motion transients due to the update [Loschky and Wolverton 2007]).

5.3 Validity of Results in Terms of Eye Movement Measures

To gage the validity of our data we decided to follow Geisler et al.'s [2006] fairly comprehensive examination of various eye movement measures during gaze-contingent visual search. More specifically, they considered search time variation in relation to the number of fixations and fixation durations. The median number of fixations was found to mirror average search time quite closely. Indeed, if fixation durations are normally distributed with a fairly constant mean across conditions, then the correlation between search time and numbers of fixations should not be surprising: search time increases because more fixations are required to perform the search.

We define fixations as segments of the eye movement signal delineated by saccades. Thus, our fixation detection algorithm relies on velocity-based saccade detection. Specifically, a simple two-tap velocity filter was used to identify saccades [Duchowski et al. 2002]. Because saccadic velocity can range anywhere from 20° – $130^\circ/\text{s}$ [Schumacher et al. 2004; Smeets and Hooge 2003; Crawford et al. 1998; Shafiq-Antonacci et al. 2003; Castet et al. 2002] we set the velocity threshold to an average value of $60^\circ/\text{s}$.

5.3.1 Effects on Number of Fixations. As seen in Figure 9, the number of fixations tends to increase with larger window sizes.

Omitting the control condition and using edge presence and window size as fixed factors, repeated-measures ANOVA indicates a significant effect of window size on the number of fixations ($F(4,40) = 5.80$, $p < 0.01$). Averaging over the edge presence conditions, pair-wise t-tests with pooled SD indicate a significantly larger number of fixations with the 20° window than with windows of 10° and 5° ($p < 0.01$, $p < 0.05$, resp., with Bonferroni correction), but not the 2° window.

Repeated-measures ANOVA also indicates a significant effect of edge presence on the number of fixations ($F(1,10) = 9.74$, $p < 0.05$) with edge \times window interaction not significant ($F(4,40) = 2.00$, $p = 0.11$, n.s.).

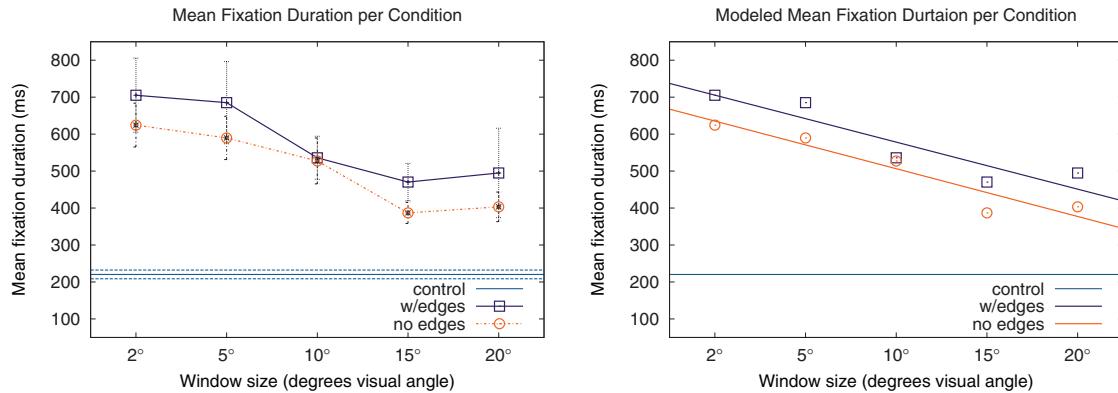


Fig. 10. Measured (left) and modeled (right) mean fixation duration per window size condition, modeled by $y = a - be^{-x/c}$ (right, with reduced $\chi^2 = 3778.82$ and reduced $\chi^2 = 2306.90$ for data with edges and no edges, resp.). Mean fixation duration for the control condition is shown as a constant line in both graphs (with standard error range).

Because the number of fixations can generally be expected to correlate with search time, once again we can fit the data to the generalized quadratic learning curve $y = ax^b + cx + d$. The fits, shown in Figure 9 (right), suggest fewer fixations with smaller window sizes, with an increasing trend toward the number of fixations observed in the control condition as the window size increases.

5.3.2 Effects on Fixation Durations. According to Geisler et al.'s [2006] findings, search time varies across stimulus conditions both because of the number of fixations and due to the duration of the fixations, with the number of fixations tending to dominate (cf. Loschky and McConkie [2002]). In our case, we observe an inverse relationship between fixation durations and window size, as seen in Figure 10, suggesting that as window sizes increase, viewers make shorter fixations than with smaller window sizes.

Omitting the control condition and using edge presence and window size as fixed factors, repeated-measures ANOVA indicates a significant effect of window size on fixation durations ($F(4,40) = 3.80$, $p < 0.05$). Averaging over the edge presence conditions, pair-wise t-tests with pooled SD indicate significantly longer fixation durations with the 2° window than with windows of 15° and 20° ($p < 0.05$, with Bonferroni correction).

Repeated-measures ANOVA also indicates a significant effect of edge presence on fixation durations ($F(1,10) = 6.58$, $p < 0.05$) with edge \times window interaction not significant ($F(4,40) = 0.11$, $p = 0.98$, n.s.).

Fixation durations seem to be better modeled by exponential decay $y = a + be^c$ rather than by the quadratic. The fits, shown in Figure 10 (right), suggest a decrease in mean fixation durations with increased window size.

Our data in terms of the number of fixations and fixation durations also disagrees with the findings of Geisler et al. [2006], who reported the number of fixations decreasing with larger window sizes. Geisler et al. reported a fairly narrow distribution of fixation durations, ranging from about 200 ms to a little more than 300 ms. Based on our estimation of fixations, we find a wider duration range, including durations shorter than 100 ms and longer than 600 ms. Figure 11 shows the range of mean fixation durations captured across all sessions and participants.

5.3.3 Effects on Saccade Length. According to Loschky and McConkie's [2002] observations, larger gaze-contingent windows produce larger saccades, since with small windows the viewer tends to make more saccades to targets inside the window of high resolution. We find a contrary relationship between

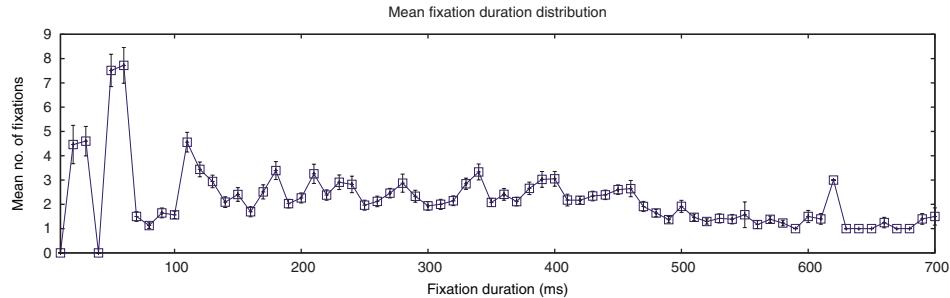


Fig. 11. Overall mean fixation duration distribution.

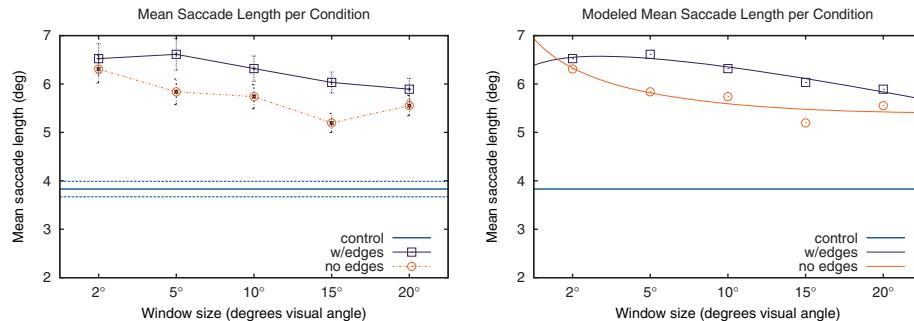


Fig. 12. Measured (left) and modeled (right) mean saccade lengths per window size condition, modeled by $y = ax^b + cx + d$ (right, with reduced $\chi^2 = 0.02$ and reduced $\chi^2 = 0.12$ for data with edges and no edges, resp.). Mean saccade length for the control condition is shown as a constant line in both graphs (with standard error range).

saccade lengths and window size, as seen in Figure 12, suggesting that as window sizes increase, viewers make smaller saccades than with smaller window sizes. Smaller saccades may be an indication of better use of peripheral preview, made available with larger window sizes. The effect is only marginally significant.

Omitting the control condition and using edge presence and window size as fixed factors, repeated-measures ANOVA indicates a marginally significant effect of window size on saccade lengths ($F(4,40) = 2.25$, $p = 0.08$). Averaging over the edge presence conditions, pair-wise t-tests with pooled SD indicate marginally significantly larger saccades with the 2° window than with window of 15° and 20° ($p < 0.05$, with Bonferroni correction).

Repeated-measures ANOVA also indicates a marginally significant effect of edge presence on saccade lengths ($F(1,10) = 7.63$, $p = 0.07$) with edge \times window interaction not significant ($F(4,40) = 0.50$, $p = 0.73$, n.s.).

5.3.4 Descriptive Statistics. Finally, we consider the mean number of fixations and saccades as a function of eccentricity. Figure 13 indicates that most fixations fall within the central 15° . This is not surprising given that fixations generally begin at the display center, as Geisler et al. observed in their study. Similarly, saccade length decreases with mean number of saccades, but only after peaking at about 5° magnitude. Both mean number of fixations and saccades over all conditions and subjects suggests a normal distribution, modeled well by the Gaussian distribution, $y = ae^{-(x+b)^2/c}$.

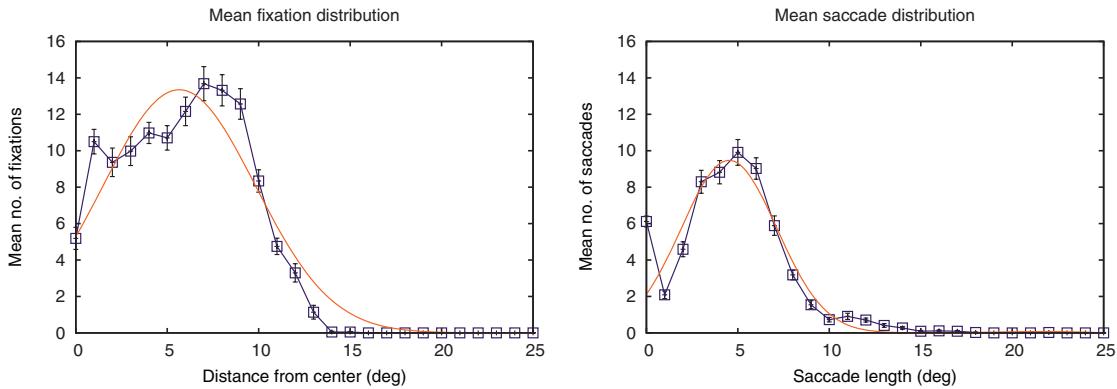


Fig. 13. Measured and modeled mean fixation and saccade distributions. The mean number of fixations is plotted against the distance from the screen center (in degrees visual angle). The mean number of saccades is plotted against saccade length (in degrees visual angle). Both measures are modeled by the Gaussian $y = ae^{-(x+b)^2/c}$ (reduced $\chi^2 = 0.12$ and reduced $\chi^2 < 0.01$ fixations and saccades, resp.).

6. DISCUSSION

As expected, search time decreased as window size increased (up to a window size of 10°) and was lower still when silhouette edges were maintained. Search time was significantly prolonged with a 2° window with no silhouette edge information. Results suggest that performance is degraded with a very small foveal window (2°) when edge information is not preserved.

These results are similar to those of other gaze-contingent rendering experiments, e.g., those of Geisler et al. [2006] and Loschky and McConkie [2002]. The downward trend of the curve's left slope of Figure 7, an inverse relationship between high-resolution inset and visual search time, is consistent with previous image-based GCD results, e.g., those of Watson et al. [2004], who evaluated a visual search task over face images with a $30^\circ \times 30^\circ$ high-resolution inset. Peripheral degradation was uniform within a scene (as opposed to the continuously varied peripheral degradation used in the hybrid technique), but varied between trials.

Our results are also similar to those of Parkhurst and Niebur [2004], obtained for their model-based visual search task. Their results are based on peripheral degradation generated through edge collapse, guided by normalized error in the collapse process. It should be noted that the visual search task utilized objects that varied widely in topology and profile.

In our case, a typical search would start at the center of the display and sweep out concentrically in search of the target. For example, as shown in Figure 14, targets are fixated in counter clockwise order until the target is located at upper left, when it is visually verified against the reference mesh at center (note the multiple refixations).² Not surprisingly, fixations tend to cover the mouth region, a particularly distinguishing feature of these meshes.

The relatively poor performance of the control group was unexpected. In the absence of any peripheral degradation, subjects were expected to localize targets rapidly. Anecdotal evidence, acquired through informal exit interviews, indicates that the control scenes were “overwhelming” and provided “too many possibilities”. It appears that the lack of overt visual cues, coupled with the rich informational environment, precluded rapid localization of peripheral targets. There is some support in the literature for this

²Fixation are detected via velocity-based analysis where velocity $< 15^\circ/\text{s}$ denotes a fixation; the centroid of all congruent fixations determines the mean fixation coordinates.

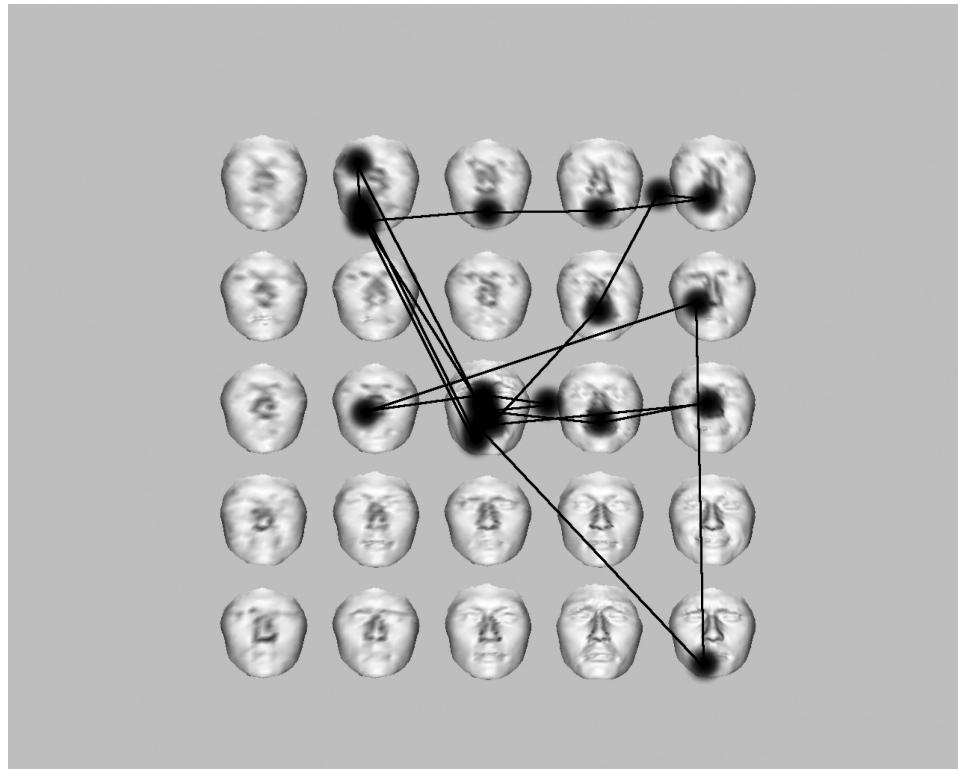


Fig. 14. Example scanpath exhibited during visual search, in this case starting at center and sweeping out concentrically in a counter clockwise direction until target is located at upper left (when it is visually verified against the reference mesh at screen center).

interpretation. Specifically, Cave and Bichot [1999] suggest that for discrimination tasks, peripheral objects may increase reaction time if they distract the viewer during discrimination. Furthermore, the authors point out that the importance of attention in suppressing competing information from distractors is supported by neurophysiological studies showing that attentional modulation of neural responses is greatest when target and distractor both fall in the receptive field of a neuron, thus competing for representation by that neuron. By masking peripheral distractors, a GCD may be beneficial to visual search where discrimination is a significant task component. If discrimination is parafoveal, then the GCD may be masking distracting elements in the near periphery, an effect known as *crowding* [Pelli et al. 2004; Louie et al. 2007].

Our analysis of eye movement measures (e.g., number of fixations, fixation durations, and mean fixation and saccade distributions) shows performance within normal limits (cf. the saccade “main sequence” [Bahill et al. 1975; Knox 2001]). Analysis of search times following a reduction of outliers tends to highlight the poorer performance with larger gaze-contingent window sizes, not previously reported [Murphy and Duchowski 2007]. Although a performance benefit appears at the 10° window, it seems to reach a point of diminishing returns at about the 15° window size.

In general, the gaze-contingent display appears to affect performance by altering natural eye movements in terms of the number of fixations and fixation durations. Smaller GCD window sizes tend to evoke fewer but longer fixations. Conversely, larger window sizes produce more shorter fixations. If longer fixation durations can be associated with increased task difficulty (e.g., larger cognitive load),

small window sizes can be considered particularly detrimental to performance. Meanwhile, larger window sizes tend to produce search performance approaching that of a full-resolution display. However, this form of linear reasoning may unintentionally “average out” the potential benefit of mid-sized gaze-contingent windows. Our analysis suggests a possible inflection point in terms of search performance with a window of 10°.

Although our study was initially designed as a visual search task, in retrospect, participants’ behavior suggests a more deliberate component beyond that of visual search and target detection. The representative scanpath shown in Figure 14 suggests a more engaging decision-making aspect that complicates simple analysis based on time to target detection. Ideally, analysis of search time in the present circumstance should strive to tease apart the time taken to search the field from the time taken to discriminate the similarity of the potential target against the reference.

Nevertheless, the GCD’s apparent speed benefit with a mid-sized window seems to lend credence to the interpretation of benefit toward discrimination (as drawn from Cave and Bichot’s [1999] findings), particularly, we conjecture, during target discrimination in the middle periphery (considering the effect of crowding [Pelli et al. 2004; Louie et al. 2007]). The GCD may be beneficial by masking peripheral distractors just at the critical time of identification. The decrease in fixation durations with larger window sizes (seen in Figure 10) supports this observation if fixation durations could be reasonably interpreted as indicators of viewers’ deliberation of the target’s semblance to the reference in the presence of peripheral distractors. If so, then, on the one hand, large GCD windows offer greater potential for influence from peripheral distractors, diminishing viewers’ concentration during discrimination. Small GCD windows, on the other hand, preclude peripheral preview benefit thereby degrading search performance. Mid-sized GCD windows may provide the happy medium by providing peripheral preview benefit that is sufficiently attenuated so as to not impede (parafoveal) discrimination.

Our analysis leads us to conclude that the gaze-contingent display with a mid-sized window may provide a benefit in both speed and accuracy, but only if the visual search task is comprised of both search and discrimination. However, the reader should be cautioned that this interpretation is somewhat tenuous. An experiment specifically designed to explore the attentional spotlight premise should be carried out to substantiate this claim.

7. CONCLUSIONS

A non-isotropic gaze-contingent rendering technique was presented utilizing aspects of both imaged-based and model-based rendering. There are several areas of potential improvement that may increase the efficiency of the hybrid technique. The decision to use a uniform grid was partially based on the architectural restrictions specific to the GPU used during development (an older graphics card). Use of the much more capable 8800GTX graphics card alleviates some of those concerns and should allow the use of more computationally efficient acceleration structures, such as Bounding Volume Hierarchies.

Another GPU-based improvement is to exploit the on-chip geometry instantiation made available on the 8800GTX. Currently, the entire frame buffer is copied back into system memory for mapping to the intermediate mesh. It may be possible to create the intermediate mesh directly on the video card through the use of geometry shaders, simultaneously reducing traffic on the system bus and further exploiting the parallel nature of the GPU. This would also allow the possibility of adaptive subdivision of the intermediate mesh; reducing the number of rays cast to maintain silhouette edges.

An alternative to more efficient exploitation of hardware would be greater utilization of the HVS. The CSF function used was intentionally forced into its most conservative form. It is possible to increase peripheral degradation either by setting (1) to a non-maximal value of contrast, or by taking into account the kinetic portion of the CSF (cf. Daly et al. [2001]).

Dynamic generation of CSF-based ray masks would also benefit the hybrid technique. The use of a chin rest during the experiment was necessary to keep the subject's eyes at the proper distance from the screen for the CSF function used. If the eye/screen separation could be determined on-the-fly, distance-correct ray masks could offer improved performance by sampling a scene less frequently.

The experimental results of the hybrid technique generally conform to prior results obtained using both image-based and model-based techniques. It would be instructive to duplicate the experiment using image-based and model-based degradation in order to directly compare performance, but until the most recent generation of GPU, the geometry manipulation required by model-based techniques forced the algorithms to run on the CPU.

Acknowledgments

We wish to thank Robert Sumner and Jovan Popović of the Computer Graphics Group at MIT for providing the mesh data used in this experiment and to Jacob Richards at Blue Sky Studios for Maya services “rendered”. Would also like to thank the reviewers of the manuscript for providing their thorough reviews and for alerting us to the *crowding* phenomenon.

REFERENCES

- BAHILL, A. T., CLARK, M., AND STARK, L. 1975. The main sequence, a tool for studying human eye movements. *Mathematical Biosciences* 24, 3/4, 191–204.
- BARON, J. AND LI, Y. 2007. Notes on the use of R for psychology experiments and questionnaires. Online Notes. URL: <http://www.psych.upenn.edu/baron/rpsych/rpsych.html> (last accessed December 2007).
- BERGSTRÖM, P. 2003. Eye-movement Controlled Image Coding. Ph.D. thesis, Linköping University, Linköping, Sweden.
- CARD, S. K., MORAN, T. P., AND NEWELL, A. 1983. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum, Hillsdale, NJ.
- CASTET, E., JEANJEAN, S., AND MASSON, G. S. 2002. Motion perception of saccade-induced retinal translation. *Proc Natl Acad Sci USA* 99, 23, 15159–63.
- CAVE, K. R. AND BICHOT, NARCISSE, P. 1999. Visuospatial attention: beyond a spotlight model. *Psychonomic Bulletin & Review* 6, 2, 204–223.
- CLARKE, J. H. 1976. Hierarchical geometric models for visible surface algorithms. *Communications of the ACM* 19, 10 (October), 547–554.
- CRAWFORD, T. J., SHARMA, T., PURI, B. K., MURRAY, R. M., BERRIDGE, D. M., AND LEWIS, S. W. 1998. Saccadic eye movements in families multiply affected with schizophrenia: the maudsley family study. *Am J Psychiatry* 155, 12, 1703–10.
- DAILY, S., MATTHEWS, K., AND RIBAS-CORBERA, J. 2001. As plain as the noise on your face: adaptive video compression using face detection and visual eccentricity models. *Journal of Electronic Imaging* 10, 1, 30–46.
- DANFORTH, R., DUCHOWSKI, A., GEIST, R., AND MCALILEY, E. 2000. A platform for gaze-contingent virtual environments. In *Smart Graphics (Papers from the 2000 AAAI Spring Symposium, Technical Report SS-00-04)*. AAAI, Menlo Park, CA, 66–70.
- DUCHOWSKI, A., MEDLIN, E., COURNIA, N., GRAMOPADHYE, A., NAIR, S., VORAH, J., AND MELLOY, B. 2002. 3D eye movement analysis. *Behavior Research Methods, Instruments, Computers (BRMIC)* 34, 4 (November), 573–591.
- DUCHOWSKI, A. T. 2007. *Eye Tracking Methodology: Theory & Practice*, 2nd ed. Springer-Verlag, Inc., London, UK.
- DUCHOWSKI, A. T. AND ÇÖLTEKİN, A. 2007. Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging. *Transactions on Multimedia Computing, Communications and Applications* 3, 4 (December).
- FERWERDA, J. A. 2003. Three varieties of realism in computer graphics. In *Human Vision and Electronic Imaging*. SPIE, Bellingham, WA, 290–297.
- GEISLER, W. S. AND PERRY, J. S. 1998. Real-time foveated multiresolution system for low-bandwidth video communication. In *Human Vision and Electronic Imaging*. SPIE, Bellingham, WA.
- GEISLER, W. S., PERRY, J. S., AND NAJEMNIK, J. 2006. Visual Search: The Role of Peripheral Information Measured Using Gaze-Contingent Displays. *Journal of Vision* 6, 9, 858–873.
- HOPPE, H. 1997. View-dependent refinement of progressive meshes. In *Computer Graphics (SIGGRAPH '97)*. ACM, New York.
- KNOX, P. C. 2001. The parameters of eye movement. Lecture Notes, URL: <http://www.liv.ac.uk/~pcknox/teaching/Eymovs/params.htm> (last accessed October 2001).

- LEVOY, M. AND WHITAKER, R. 1990. Gaze-directed volume rendering. In *Computer Graphics (SIGGRAPH '90)*. ACM, New York, 217–223.
- LINDSTROM, P., KOLLER, D., RIBARSKY, W., HODGES, L. F., FAUST, N., AND TURNER, G. A. 1996. Real-time, continuous level of detail rendering of height fields. In *Computer Graphics (SIGGRAPH '96)*. ACM, New York, 109–118.
- LOSCHKY, L. C. AND 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.
- LOSCHKY, L. C., MCCONKIE, G. W., YANG, J., AND MILLER, M. E. 2005. The limits of visual resolution in natural scene viewing. *Visual Cognition* 12, 6, 1057–1092.
- LOSCHKY, L. C. AND WOLVERTON, G. S. 2007. How late can you update gaze-contingent multiresolutional displays without detection? *Transactions on Multimedia Computing, Communications and Applications* 3, 4 (December).
- LOUIE, E. G., BRESLER, D. W., AND WHITNEY, D. 2007. Holistic crowding: selective interference between configural representations of faces in crowded scenes. *Journal of Vision* 7, 2, 1–11.
- LUEBKE, D. AND ERIKSON, C. 1997. View-dependent simplification of arbitrary polygonal environments. In *Computer Graphics (SIGGRAPH '97)*. ACM, New York.
- LUEBKE, D. AND HALLEN, B. 2001. Perceptually driven simplification for interactive rendering. In *Proceedings of the 2001 Eurographics Workshop on Rendering (Rendering Techniques)*, S. Gortler and K. Myszkowski, Eds. Eurographics.
- LUEBKE, D., HALLEN, B., NEWFIELD, D., AND WATSON, B. 2000. Perceptually Driven Simplification Using Gaze-Directed Rendering. Tech. rep. CS-2000-04, University of Virginia.
- MACCRACKEN, R. AND JOY, K. 1996. Free-from deformations with lattices of arbitrary topology. In *Computer Graphics (SIGGRAPH '96)*. ACM, New York, 181–188.
- MURPHY, H. AND DUCHOWSKI, A. T. 2001. Gaze-contingent level of detail. In *EuroGraphics*. EuroGraphics, Manchester, UK.
- MURPHY, H. AND DUCHOWSKI, A. T. 2007. Hybrid image-/model-based gaze-contingent rendering. In *APGV '07: Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization*. ACM, New York, 107–114.
- OHSHIMA, T., YAMAMOTO, H., AND TAMURA, H. 1996. Gaze-directed adaptive rendering for interacting with virtual space. In *Proceedings of VRAIS'96*. IEEE, 103–110.
- PARKHURST, D. J. AND NIEBUR, E. 2002. Variable resolution displays: a theoretical, practical, and behavioral evaluation. *Human Factors* 44, 4, 611–629.
- PARKHURST, D. J. AND NIEBUR, E. 2004. A feasibility test for perceptually adaptive level of detail rendering on desktop systems. In *Applied Perception, Graphics & Visualization (APGV) Symposium*. ACM, New York, 49–56.
- PELLI, D. G., PALOMARES, M., AND MAJAJ, N. J. 2004. Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision* 4, 1136–1169.
- POPOV, S., GÜNTHER, J., SEIDEL, H.-P., AND SLUSALLEK, P. 2007. Stackless KD-tree traversal for high performance GPU Ray Tracing. *Computer Graphics Forum* 26, 3, 415–424.
- POSNER, M. I., SNYDER, C. R. R., AND DAVIDSON, B. J. 1980. Attention and the detection of signals. *Experimental Psychology: General* 109, 2, 160–174.
- PURCELL, T. J., BUCK, I., MARK, W. R., AND HANRAHAN, P. 2002. Ray tracing on programmable graphics hardware. *ACM Transactions on Graphics* 21, 3 (July), 703–712.
- REDDY, M. 1998. Specification and evaluation of level of detail selection criteria. *Virtual Reality: Research, Development and Application* 3, 2, 132–143.
- REINGOLD, E. M., LOSCHKY, L. C., MCCONKIE, G. W., AND STAMPE, D. M. 2003. Gaze-contingent multi-resolutional displays: an integrative review. *Human Factors* 45, 2, 307–328.
- SCHMALSTIEG, D. AND SCHAUFLER, G. 1997. Smooth levels of detail. In *Proceedings of VRAIS'97*. IEEE, 12–19.
- SCHUMACHER, J., ALLISON, R. S., AND HERPERS, R. 2004. Using saccadic suppression to hide graphic updates. *Eurographics Symposium on Virtual Environments*.
- SHAFIQ-ANTONACCI, R., MARUFF, P., MASTERS, C., AND CURRIE, J. 2003. Spectrum of saccade system function in alzheimer disease. *Arch Neurol* 60, 9, 1272–8.
- SMEETS, J. AND HOOGE, I. 2003. Nature of variability in saccades. *Journal of Neurophysiology* 90, 12–20.
- TOBII TECHNOLOGY AB. 2003. Tobii ET-17 Eye-tracker product description (v1.1). URL: <http://www.tobii.se/> (last accessed January 2007).
- WATSON, B., WALKER, N., AND HODGES, L. F. 2004. Supra-threshold control of peripheral LOD. *Transaction on Graphics (SIGGRAPH'04 Proceedings)* 23, 3, 750–759.

- WATSON, B., WALKER, N., HODGES, L. F., AND WORDEN, A. 1997. Managing level of detail through peripheral degradation: effects on search performance with a head-mounted display. *ACM Transactions on Computer-Human Interaction* 4, 4 (December), 323–346.
- ZORIN, D. AND SCHRÖDER, P. 2000. *Course 23: Subdivision for Modeling and Animation*. ACM, New York, URL: <http://www.mrl.nyu.edu/dzorin/sig00course/>
- ZORIN, D., SCHRÖDER, P., AND SWELDENS, W. 1997. Interactive multiresolution mesh editing. In *Computer Graphics*. ACM, New York.

Received November 2007; revised April 2008; accepted July 2008