A Feasibility Test for Perceptually Adaptive Level of Detail Rendering on Desktop Systems

Derrick Parkhurst and Ernst Niebur
The Zanvyl Krieger Mind/Brain Institute

The Johns Hopkins University, Baltimore, Maryland
$\{derrick.parkhurst \mid niebur\}$@jhu.edu

Abstract

Level of detail (LOD) rendering techniques reduce the geometric complexity of 3D models, sacrificing visual rendering quality in order to increase frame rendering rates. Perceptually adaptive LOD rendering techniques take into account the characteristics of the human visual system to minimize visible artifacts attributable to the reduced LOD. While these techniques have been previously examined in the context of high-performance rendering systems, it is not clear whether the benefits will necessarily overcome the behavioral costs associated with a reduced LOD on ordinary desktop systems. To answer this question, two perceptually adaptive rendering techniques, one velocity-dependent and one gaze-contingent, were implemented in the Unreal^{TM} rendering engine on a standard desktop computer and monitor. These techniques were evaluated in separate experiments where participants were required to perform a virtual search for a target object among distractor objects in a perceptually rendered virtual home interior using a mouse to rotate the viewport. In the first experiment, objects moving across the observer's field of view were rendered in less detail than stationary objects, taking advantage of the fact that visual sensitivity to the details of moving objects is substantially reduced. Reaction times to detect the target remained constant with decreasing detail, whereas reaction times to localize a target decreased. In a second experiment, an eye tracker was used to render objects at the point of gaze in more detail than objects in the periphery, taking advantage of the fact that visual sensitivity is greatest at that location. Reaction times to detect the target increased with decreasing detail, whereas reaction times to localize a target decreased. The results from these experiments suggest that a reduced LOD can impede target identification, however, the resultant increase in frame rates facilitates virtual interaction. Overall, the behavioral costs associated with perceptually adaptive LOD techniques can be offset by the behavioral performance gains on desktop systems. However, we show that the nature of the task is important in determining the exact cost-benefit trade-off.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques;

Keywords: Velocity-dependent, Gaze-Contingent, Visual Search

1 Introduction

The visual quality of rendered virtual environments is strongly dependent on the computational resources available in the rendering process. When resources are plentiful, computationally expensive lighting and shading techniques can be combined with complex geometric models to generate a very realistic rendering. When resources are scarce, compromises must be made in visual quality in order to reduce the computational requirements of the rendering process. Although the computational power of today's most advanced rendering systems is impressive, even this hardware is severely underpowered with respect to the most realistic rendering algorithms currently available. This is especially true for real-time systems where environments must be rendered at frame rates suitable for natural human interaction. This fact has motivated research into techniques that reduce the computational requirements of the rendering process.

While many variables influence the computational requirements of the rendering process, for example, lighting and texture-mapping techniques, the geometric complexity of three-dimensional models has received a great deal of attention. This is because, in general, the greater the complexity of a model, the more computational resources are required to render that model.

A technique known as level of detail (LOD) rendering takes advantage of the fact that much of a model's geometric detail is unnecessary under certain circumstances [Funkhouser and Sequin 1993]. For example, distance-based LOD manipulations reduce model complexity when the model is distant from the viewer and thus subtends only a small portion of the visual field. Distance-based LOD reduction was originally conceived for use in flight simulators [Clark 1976] but has now seen widespread use in most rendering engines.

Perceptually adaptive LOD rendering techniques are similar but involve, for example, selectively reducing model complexity when such reductions are not likely to be perceived by the viewer [Reddy 2001]. These techniques exploit the fact that the sensitivity of the visual system to detail can vary in different situations. One such technique, view-dependent LOD reduction exploits view-dependent illumination, visibility and frame-to-frame coherence to optimize rendering [Xia et al. 1997; Hoppe 1997; Luebke and Erikson 1997]. More recently, view-dependent techniques have also incorporated models of contrast and spatial frequency sensitivity of the human visual system to help guide LOD reduction [Reddy 2001; Luebke and Hallen 2001; Williams et al. 2003]. Other, similar techniques utilize models of attention [Yee et al. 2001; Brown et al. 2003] or perception [O'Sullivan and Dingliana 2001].

Another such technique, gaze-contingent LOD reduction, takes advantage of the fact that sensitivity to detail falls off rapidly in the visual periphery [Virsu and Rovamo 1979]. This technique utilizes an eye-tracking system to render models near the center of gaze in more detail than models in the periphery [Levoy and Whitaker 1990; Ohshima et al. 1996; Luebke et al. 2000; Danforth et al. 2000; Murphy and Duchowski 2001]. Finally, a less studied technique, velocity-based LOD reduction, renders moving objects in less detail than stationary objects [Reddy 1994; Reddy 1997]. This technique exploits the fact that the visual system has a reduced sensitivity to the details of moving stimuli [Murphy 1978; Burr and Ross 1982].

Figure 1: Four different viewpoints of one virtual home interior used in the experiment. These four views of 90° each span the whole room (360°). Each view is drawn using the highest LOD setting (1.0).

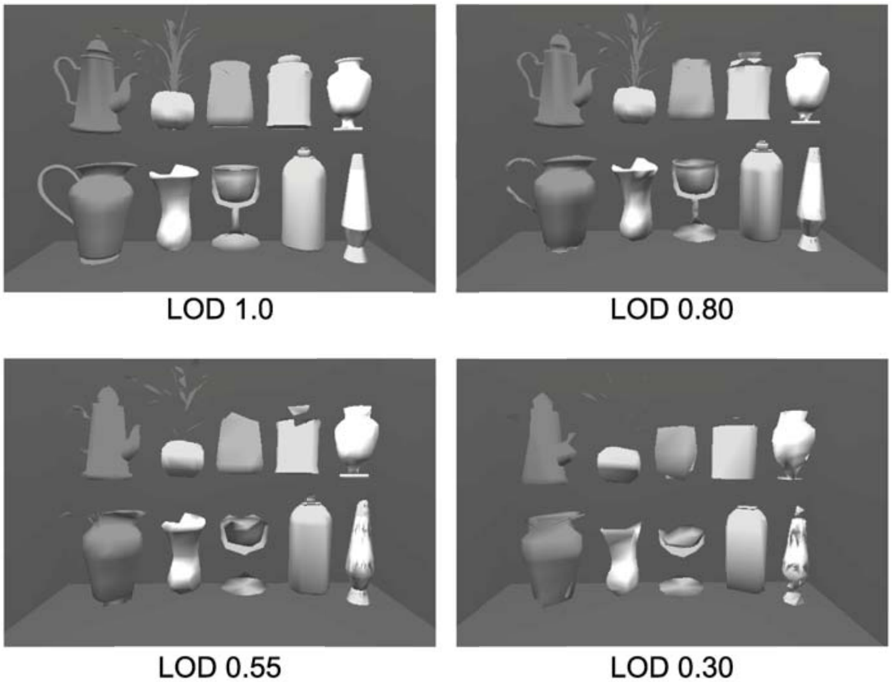

Figure 2: The 10 target objects are shown at 4 LODs. Note that these LODs are for example and that LOD took on a range of values in the experiments (see methods).

2 Motivation and Contribution

The majority of perceptually adaptive rendering techniques have not yet seen widespread application in spite of their potential to significantly reduce the computational requirements of rendering. One reason is that the greatest benefits are obtained when used in combination with large, high-resolution displays [Parkhurst and Niebur 2002]. While these techniques are potentially very useful for virtual-reality systems using head-mounted displays or CAVEs [Cruz-Neira et al. 1993], it is not clear how useful these techniques will be for smaller field-of-view desktop rendering systems of the sort typically used in applications such as architectural walk-through, computer-assisted design and video gaming. This question has gone unanswered because the majority of perceptually adaptive LOD research has used head-mounted displays [Ohshima et al. 1996; Watson et al. 1997; Danforth et al. 2000; Murphy and Duchowski 2001] and/or high-end rendering systems [Xia et al. 1997; Luebke et al. 2000; Reddy 2001; Yee et al. 2001].

The first contribution of the current research is the implementation of two perceptually adaptive rendering techniques on a typical desktop system, a velocity-based LOD reduction and a gaze-contingent LOD reduction. We utilized a standard 1Ghz desktop computer with an attached 17-inch monitor driven by a consumer-grade video card under the control of a modified version of the Unreal^{TM} rendering engine. An ISCAN eye tracker was used in order to implement gaze-contingent LOD reduction.

Although some perceptually adaptive LOD techniques aim at reducing detail without noticeable artifacts, this is not always possible. In many cases, an aggressive LOD reduction is intentionally applied in order to obtain a high frame rate. Typically this is the case in real-time rendering applications where high frame rates are necessary. Low frame rates can be detrimental when interacting with virtual environments [Watson et al. 1998]. In our desktop implementation, we use aggressive LOD reductions to maximize frame rates. This induces LOD artifacts that are perceptually visible but not necessarily perceptually disturbing. Our own anecdotal evidence with perceptually adaptive displays suggests that such artifacts become less disturbing with experience.

Another factor limiting the widespread implementation of perceptually adaptive rendering techniques is that the behavioral costs associated with use of these displays have not been extensively studied with realistic tasks and complex environments. The consequences of, for example, explicitly reducing the geometric complexity of models in different behavioral tasks is largely unknown. The majority of the behavioral research with perceptually adaptive displays has been anecdotal and reported only in conjunction with research primarily focused on the mechanics and algorithms necessary for implementation of the display.

A significant amount of research has focused on the behavioral consequences of variable-resolution displays. These displays present an image on the screen using a number of different pixel resolutions and are primarily used in image-transmission applications. For a review of this literature, see Parkhurst and Niebur [2002]. While some of the lessons learned from this literature may generalize to perceptually adaptive LOD rendering, LOD reduction cannot be strictly characterized as a decrease in spatial resolution, and thus an investigation of the particular artifacts associated with LOD rendering is warranted.

The second contribution of the current research is to determine if, on desktop systems, the behavioral artifacts associated with aggressive LOD reduction in a real-time rendering system can be overcome by the benefits associated with the obtainable increase in frame rendering rates. To this end, we examined the behavior of users performing a "virtual search" task in complex virtual environments rendered using two different perceptually adaptive LOD techniques. The first of two experiments examines a simple velocity-based LOD reduction technique. The second experiment examines the more technically challenging gaze-contingent LOD reduction technique.

The virtual search task is a paradigm similar to traditional visual search paradigm in that participants are required to search for a specified target item in a display. However, the task requires that participants search a complex three-dimensional virtual environment for a target object. Participants view the entire room from a single central location by rotating the virtual viewport using a mouse. Single-room home interiors were used as virtual environments (see Figure 1) with common objects such as vases or cups used as search targets (See Figure 2).

We chose to use this natural task with rich and complex stimuli to evaluate the perceptually adaptive LOD reduction techniques rather than a more simplistic virtual search task. This is the first experiment that we are aware of that uses a natural task with complex stimuli to evaluate a perceptually adaptive display. Other experiments have utilized much simpler tasks and stimuli, for example, searching for a target letter among distractor letters painted on the walls of an empty room [Pausch et al. 1997]. While simple tasks and stimuli can be better controlled for the purposes of examining detailed behavioral effects, the results obtained in our experiments will be more likely to generalize to applied settings, for example, architectural walk-through and video games.

3 Level of Detail

The LOD of models was manipulated using a simple edge-collapsing algorithm. The LOD reduction algorithm iteratively collapses the edges that introduces the least amount of error into the model. A wide variety of parameters can be used to determine the error associated with collapsing an edge including the edge length, local curvature, local color or texture differences. In general, these parameters can greatly affect the visual quality of a simplified mesh. There exists no unambiguous measure that captures all aspects of the visual quality of a simplified mesh, however, most techniques focus on some aspect of geometric fidelity.

In this study, an edge collapse technique was used where the error introduced by a collapsed edge is determined strictly by edge length. This technique is one of the simplest and most widely used techniques [Schroeder et al. 1992; Turk 1992; Hoppe et al. 1993]. Although recent work shows that this technique is inferior to other more complex techniques, for example, image-driven simplification [Lindstrom and Turk 2000], we chose to use it for three reasons. First, the focus of our research is on real-time LOD management techniques rather than advanced LOD reduction techniques. Second, the implementational details of the Unreal^{TM} rendering system required the use of a linked list of vertices to indicate the successive order of edge collapses to obtain a LOD ranging anywhere from the highest to the lowest LOD. This list is precomputed prior to run time to afford rapid selection of a LOD and its use is advantageous for real-time performance on a desktop system. Third, the use of this simple algorithm allows us to be conservative and examine the worse case scenario wherein if a better LOD reduction technique is used, system performance will only improve.

In order to implement an edge-collapse LOD reduction technique useful for our virtual environments, where model size, model complexity (total number of vertices), and geometric construction vary significantly, we developed a normalized LOD measure. The goal of this measure was to allow us to produce a perceptually constant simplification across a wide variety of models. The normalized LOD measure is calculated by dividing the total error for any given reduced model (i.e., the sum length of collapsed edges) by the sum length of all edges in the full model. This measure ranges between 0 (lowest LOD) and 1 (highest LOD). This normalization procedure serves to equate LOD reductions on models of different complexity. Models at a low LOD tend to have few vertices and models at a high LOD tend to have more vertices.

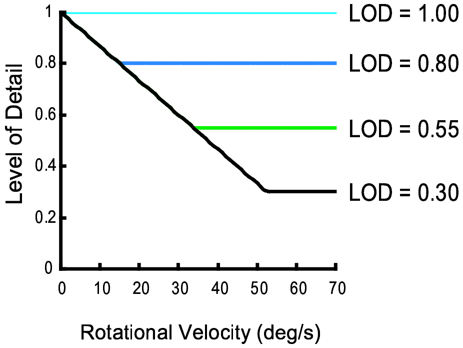

Figure 3: The relationship between the LOD of rendered models and the rotational velocity of the viewport in each of the four conditions of Experiment 1.

While edge length is a simple approximation of error, it's use in conjunction with the normalization procedure produces LODs that strongly correlate with perceived image-based differences in the rendering of the models used in our experiments. This can be clearly seen in Figure 2 where the target objects are shown at four different LODs. Note that its use may not generalize well to all models, such as those that are highly tessellated.

4 Experiment 1: Velocity-Based LOD

Given the relatively slow time constants of biological vision, objects in motion create a blur which obscures the fine details of those objects. This experiment examined the behavioral consequences of using a velocity-based LOD reduction technique taking advantage of this physiological fact. The models in the environment were rendered using a LOD proportional to the relative angular velocity of the object across the visual field; higher velocities induced a lower LOD. To limit the complexity of this experiment, all objects were kept stationary in the environment. Thus, the relative velocity of an object moving across the visual field was determined purely by the angular velocity of the viewport. As the viewport was rotated, the LOD of the object dropped, but as the viewport slowed and stopped, the LOD of objects increased to its maximum level. As shown in Figure 3, LOD was continuously and linearly related, above a threshold, to the rotational velocity of the viewport. Four experimental conditions were examined corresponding to four different threshold LODs (1.00, 0.80, 0.55, and 0.30).

4.1 Methods

4.1.1 Participants

Six Johns Hopkins students were paid for participation in the experiment. All participants had normal or corrected-to-normal vision and all were naive with respect to the purpose of the study.

4.1.2 Apparatus

Virtual environments were rendered using the Unreal^{TM} rendering engine on an 1Ghz Intel Pentium III based personal computer using an Elsa Gladiac Graphics Adapter with 32MByte of video memory. All environments were presented full-screen on a standard 17-inch computer screen at a resolution of 800 by 600 pixels at a video refresh rate of 60hz. Participants were seated at normal viewing distance $(60cm)$ from the computer screen, the viewable portion of which subtended 30.0° of visual angle horizontally and 22.4° vertically. Participants interacted with the environment using a standard right-handed, three-button mouse. The viewport of the virtual environment could be rotated by moving the mouse left or right, controlling yaw, as well as by moving the mouse forward or backward, controlling pitch. The rendered field of view spanned 90° of visual angle in the virtual environment.

4.1.3 Stimuli

A total of six different home interiors were created. Each home interior contained the same set of objects, but arranged differently. All home interiors were of the same physical dimensions and the participant's virtual position was always in the exact center of the room. Participants could control the viewport yaw and pitch but not their virtual position. Objects were placed along each of the walls in a way that gave as near as possible a normal home interior appearance. No objects were placed in the center of the room. Because the complexity, and therefore the rendering time, of each object varied widely, objects were distributed across the room with the goal of maintaining an approximately constant frame rate from any view. The distribution of objects can be seen in figure 1 where four different views of the same home interior are shown.

The virtual home interiors were constructed by using polygonal mesh objects obtained from a wide range of sources on the Internet. These models were presumably generated using different methods and/or software. The complexity of the models differed greatly, with the number of vertices in the models ranging from 100 to 6005 with a mean of 1730 vertices (SD=1617). In addition, the topology of the meshes varied. Some meshes were a singular piece with no holes (e.g. the vase in figure 2) while others had a number of disconnected parts (e.g. the leaves of the plant). One consequence of this fact is that heavily reduced meshes sometimes appeared to have a number of disconnected parts (e.g. the teapot and its handle in Figure 2). Note that this effect was not strictly a consequence of the chosen simplification technique.

4.1.4 Task

The virtual search task requires participants to search for a target object in a virtual environment. Our virtual search task differs from traditional visual search paradigms because it involves not only a sensory component, i.e. detecting a visual stimulus, but also a motor component, i.e. controlling the viewport. Ten objects of similar visual appearance were selected from those that were contained in each home interior and used as target objects. These targets are shown in Figure 2. The target on each trial was randomly chosen.

Psychophysical evidence indicates that color can be efficiently used to guide attention in visual search paradigms [Egeth et al. 1984]. If the task examined does not critically depend on the visual form of the target but rather on the color of that target, little or no effect of varying the LOD should be found. By removing the ability of color to guide the search, the task becomes more dependent on form, and more likely to be sensitive to LOD manipulations. To limit the ability of participants to use color information rather than form information, all objects in the environments were rendered in shades of gray. Although the range of object shades varied from target to target, no target was uniquely defined by its shade of gray.

4.1.5 Procedure

At the beginning of each trial a target object was randomly selected from 10 possible target objects (see Figure 2) and presented in full detail (LOD=1.0) to the participant on a blank screen. The target was continually rotated about all three axes such that participants could view it from all directions. The target was presented until the participant pressed the left mouse button at which time the target disappeared and was replaced by a centrally located fixation cross. Participants were instructed to fixate the cross and press the left mouse button to begin the trial, at which time the virtual environment was displayed.

Participants were instructed to find the target as quickly as possible and respond that they had detected the object by clicking the left mouse button. The time to make this response is referred to as the detection time. After this first response was given, a green cross hair appeared at the center of the display. Participants were required to localize the target by panning the display so that the cross hair was aligned with the target object and then press the left mouse button. Participants were told that accuracy was important for the

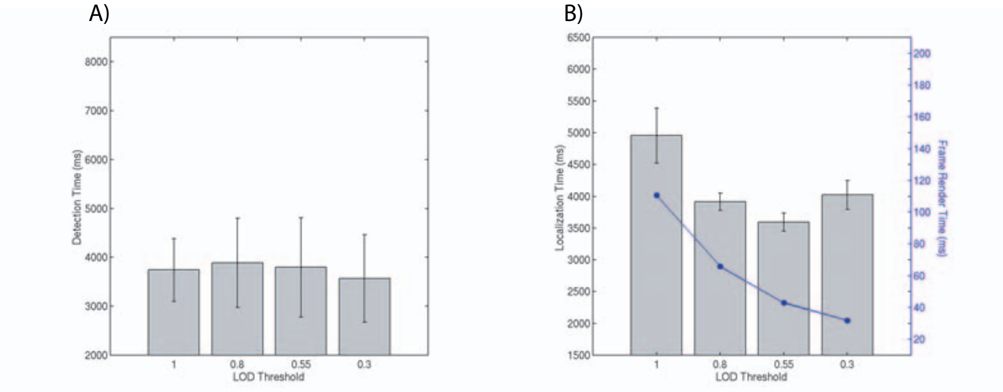

Figure 4: A) Mean time to detect the target (Error bars=±1SE). B) Mean time to localize the target are indicated as bars (Error bars=±1SE). Median frame render times are indicated with a line (Error bars smaller than symbols).

localization response. The time to make this response is referred to as the localization time.

4.1.6 Experimental Design

A total of four experimental conditions with different LOD thresholds (1.00, 0.80, 0.55, 0.30) were tested. Note that the distance-based LOD reduction normally implemented in the Unreal^{TM} rendering engine was disabled and irrelevant to this study given that all objects were approximately equidistant from the viewer. Each participant completed 120 trials. Each of the six different home interiors was presented in each of the four different experimental conditions a total of five times. Experimental conditions were selected for each trial in a random order.

4.2 Results

Mean detection times were calculated based only on those trials where the subsequent localization response was made within plus or minus 2 degrees of visual angle from the target. Any trials where the participant make a detection response in less than one second were excluded. Any trials where participants required more than 5 seconds to localize the target after signaling detection were also excluded. These criteria eliminated trials where participants may have erroneously given a premature detection response and subsequently tried to localize the correct target. Accuracy, detection time and localization time were each subjected to a one-way repeated-measures ANOVA with LOD threshold as the within-subjects independent variable.

Accuracy was, on average, 95 percent and did not significantly vary with LOD threshold $(F(3, 15) = 1.24, p > 0.10)$. No effect of LOD threshold on detection times was observed either $(F(3, 15) = 0.51, p > 0.10)$. The detection times are shown in Figure 4A as the mean taken across participants. However, a significant effect of LOD threshold on localization times was observed $(F(3, 15) = 4.74, p < 0.05)$, where localization times tended to decrease with the decreasing LOD threshold. The localization times are shown as bars in Figure 4B as the mean taken across participants. Median frame rendering times were tabulated and are shown as a line in Figure 4B. Frame render times decreased substantially with decreasing LOD. Median frame rendering times were preferred over the mean times to counteract the heavy skew induced by the frame render time for high LODs that occur when rotational velocity is low. Note that the pattern of results is similar using the mean times, however, the mean rendering times are biased towards longer times due to the skew.

4.3 Discussion

The goal of this experiment was to examine the effect on virtual search performance of using a velocity-based LOD reduction technique to render an interactive virtual environment on a desktop system. Participants performed an interactive search for target objects in a rendered home interior. Target detection and localization times were measured. We found no evidence that detection times were affected by the LOD reduction used in this experiment. This can be seen in Figure 4A. On the other hand, target localization times were affected by the LOD reduction. As can be seen in Figure 4B, localization times tended to decrease with LOD reduction. This improvement in performance was due not to the reduced detail, but rather to the decreased frame rendering time that was the direct consequence of using a lower LOD. As can be seen by comparing the frame rendering times with the localization times in Figure 4B, there is a good correlation between the two. This suggests that the decrease in rendering time lead to an increased system responsiveness to user input, allowing participants to more efficiently interact with the environment.

These results suggest that a velocity-based LOD rendering technique which trades model detail for decreased frame rendering times can be used effectively in a desktop rendering system without hurting detection performance in complex tasks such as visual search. Furthermore, the resulting decrease in frame rendering times can facilitate localization performance when manual interaction with a virtual environment is required.

It is important to note that the technique we used in this experiment related LOD to the velocity of the object across the *display*. We relied on the fact that if the eyes are fixated on a given location, the movement of the object across the display corresponds to the movement of the object across the retina, which induces a blur. However, if the eyes are smoothly tracking an object that is in motion, the reduced LOD of the object will be visible because the movement of the object across the display corresponds to a stationary object on the retina. This fact may be especially problematic given that motion can strongly attract attention and induce visual tracking of the eye. While this simplification apparently did not adversely affect performance in this experiment, it may be more critical in environments with moving stimuli, especially when the viewport is stationary and a single object is in motion. Our future research will investigate whether this simplification will be sufficient under less restricted conditions or whether active tracking of eye movements will be necessary. The next experiment begins to address this question by examining the behavioral effects of relating LOD to the position of the eye.

5 Experiment 2: Gaze-Contingent LOD

It is well established that the sensitivity of the human visual system to detail falls off rapidly in the periphery. This experiment examines the behavioral consequences of using a gaze-contingent LOD reduction technique that takes advantage of this physiological fact to minimize the computational resource required to render a virtual environment. The technique utilizes an eye-tracking system in order to render models near the center of gaze in more detail than

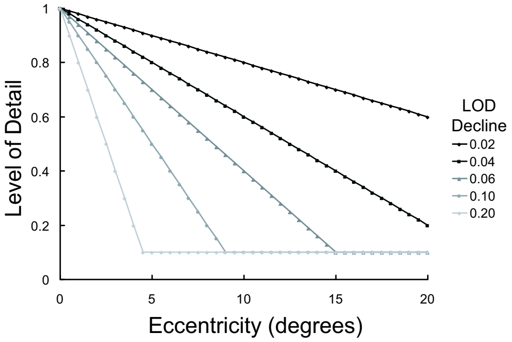

Figure 5: The relationship between the LOD and the distance from the point of gaze in each experimental condition.

models in the visual periphery.

Gaze-continent LOD reduction was implemented by relating the LOD of each object to the distance (in degrees of visual angle) of that object to the instantaneous point of gaze. As shown in Figure 5, LOD decreased linearly as the distance from the rendered object to the point of gaze increased. A minimum LOD of 0.10 was maintained in all conditions. Six experimental conditions were examined corresponding to different rates of LOD decline. Detail dropped as the distance from the point of gaze increased at a rate of 0.00, 0.02, 0.04, 0.06, 0.10 or 0.20 LOD units per degree.

Pilot studies indicated that noise in the measurement of the point of gaze, ($<$1 deg of visual angle) in combination with the gaze-contingent LOD rendering caused a visually disturbing effect. The fluctuations in the measured point of gaze (due to noise) caused the LOD of each object to fluctuate rapidly. This fluctuation resulted in vertices ‘popping’ in and out of meshes. The effect, best characterized as a ‘shimmering’, was most significant in the periphery where sensitivity to motion signals (or temporal transients) is greatest. To reduce this visually disturbing effect a threshold was imposed on the point of gaze. The point of gaze, as relevant to the gaze-contingent rendering, was considered to have moved only if it had surpassed a criterion distance of 1.5 degrees of visual angle. This threshold for the most part eliminated the disturbing effect in our pilot study.

5.1 Methods

5.1.1 Participants

Six Johns Hopkins students were paid for participation in the experiment. All participants had normal or corrected-to-normal vision and all were naive with respect to the purpose of the study. None of the participants in this experiment participated in Experiment 1.

5.1.2 Eye Tracking

An ISCAN model RK-416 eye tracker was used to monitor eye position. This model is a real-time digital image processor that tracks the center of the participant’s pupil and measures its size from an infrared video image of the participant’s eye. The unit automatically computes the position of the pupil over the two-dimensional matrix of the eye imaging camera. Pupil coordinates are computed at a rate of 60Hz and with a spatial accuracy of 0.5 deg. Note that the frame-rendering rate and the eye-tracking rate are independent of each other. A bi-cubic nonlinear interpolation (cubic in both horizontal and vertical dimensions) between a grid of nine calibration points was used to calibrate the eye tracker [Stampe 1993]. The calibration was adjusted using a procedure where an eye sample from the fixation point at the beginning of each trial, just after viewing the target object and just prior to entering each virtual room, was used to re-align the original nine point interpolation. A chin rest was used to minimize eye tracking artifacts due to head movements.

5.1.3 Experimental Design

A total of six experimental conditions were examined, each with a different rate of LOD decline. Frame rates were allowed to vary and increased with increasing rates of LOD decline (see Results). Each participant completed a total of 180 trials. Each of six different home interiors was presented in each of the six different experimental conditions a total of five times. Experimental conditions were selected for each trial in a random order. Note that the stimuli, task and procedure were these same as in Experiment 1 with one exception; the experiment was divided into two experimental sessions, each consisting of 90 trials.

5.2 Results

Mean reaction times were calculated using the same criteria as in Experiment 1. Accuracy, detection time and localization time were each subjected to a two-way repeated-measures ANOVA with rate of LOD decline and experimental session as within-subjects independent variables.

Accuracy was, on average, 90 percent and did not significantly vary with the rate of LOD decline ($F(5,25) = 0.41, p > 0.10$), but was higher in the first session (93%) than in the second session (86%) ($F(5,25) = 9.93, p < 0.05$).

Detections times increased with increasing rates of LOD decline ($F(5,25) = 2.61, p < 0.05$). The detection times are shown in Figure 6A for session 1 and Figure 6B for session 2 as the mean taken across participants. The best-fitting linear regression to the detection times from all participants are shown for each session. Detections times differed significantly across session ($F(1,5) = 26.62, p < 0.005$) being faster in session 2 that session 1. Furthermore, the average slopes obtained from the best-fitting linear regression to the detection times of each participant differed significantly across sessions ($t(10) = 1.90, p < 0.05$). The increase in detection times with increasing rate of LOD decline was greater in session 1 (112 ms/0.01 units of LOD decline) than in session 2 (42 ms/0.01 units of LOD decline).

A significant effect of rate of LOD decline on localization times was observed ($F(5,25) = 2.61, p < 0.05$), where localization times tended to decrease with the decreasing rates of LOD decline. Localization times did not significantly differ across session ($F(1,5) = 1.01, p > 0.10$). The localization times are shown as bars in Figure 6C as the mean taken across participants. Median frame rendering times were tabulated and are shown as a line in Figure 6C. Frame render times decreased substantially with decreasing LOD. Median frame rendering times were used to allow comparison with the times from Experiment 1.

5.3 Discussion

The goal of this experiment was to examine the effect on virtual search performance of using a gaze-contingent LOD reduction technique to render an interactive virtual environment on a desktop system. Participants performed an interactive search for target objects in a rendered home interior. Target detection and localization times were measured.

We found that detection time was a function of the rate of LOD decline with higher rates of LOD decline leading to slower detection times. It is known that as item similarity increases, so does visual search difficulty [Duncan and Humphreys 1989]. As can be seen in Figure 2, target similarity increases with decreasing detail. The decreased detail of objects in the periphery made the visual search more difficult and resulted in increased detection times. This relationship is seen in session 1 (Figure 6A) and 2 (Figure 6B). The detection times were faster and less affected by the reduced LOD in the second session. This result suggests that the adverse effects of the reduced LOD can be minimized with experience.

Target localization times were also affected by the LOD reduction. As can be seen in Figure 6C, localization times tended to decrease with LOD reduction. This is a similar result to that obtained in Experiment 1. The decrease in rendering time again lead to an increased system responsiveness to user input, allowing participants to more efficiently interact with the environment. As can be seen

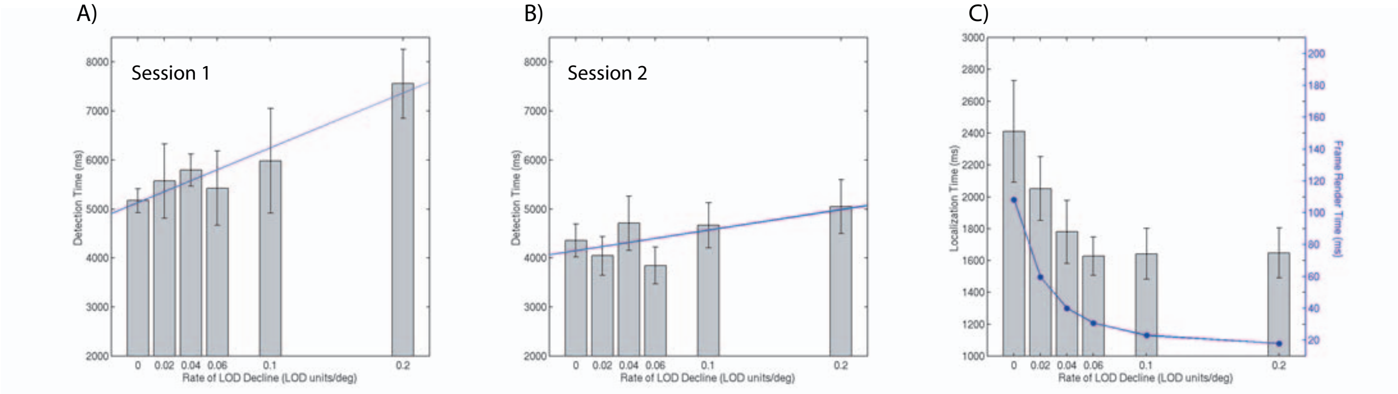

Figure 6: A) Mean times to detect the target in the first session are indicated with bars (Error bars=±1SE). Best linear fit is shown as a line. B) Mean times to detect the target in the second session are indicated with bars (Error bars=±1SE). Best linear fit is shown as a line. C) Mean times to localize the target (across sessions) are indicated as bars (Error bars=±1SE). Median frame render times are indicated with a line (Error bars are smaller than symbols).

by comparing the frame rendering times with the localization times in Figure 6C, there is a strong correlation between the two.

It should be noted that while the decreased detail of objects in the periphery made the the visual search more difficult, another factor may have also contributed to the adverse effects. In this experiment, the LOD of rendered objects was linked to the point of gaze. However, eye movements can be very rapid, requiring only a few tens of milliseconds (depending on the distance moved). With a sampling period of 16 ms, the eye tracker was sufficiently fast to follow these eye movements, but the frame rendering times were often too slow to keep up with the measured eye movements. This means that for a short period (tens of milliseconds) after some eye movements, the objects at the point of fixation could be in an inappropriately low LOD, making them difficult to identify. This rendering delay may have contributed to an overall slowing of task performance. The rendering delay cannot account for the increase in detection times with increasing rates of LOD decline given that the frame rendering time, and thus the rendering update delay after an eye movement, decreases significantly with increasing rates of LOD decline (See Figure 6C). While a faster eye tracker would help minimize the delay to some degree, the frame rendering rates are the limiting factor.

6 General Discussion

The goal of this study was to determine if LOD rendering techniques could be feasibly implemented and used on desktop rendering systems. Two experiments were conducted using two different perceptually adaptive LOD reduction techniques on a desktop computer. Participants were required to search for target objects in a virtual home interior. The primary finding was that target detection was in some cases impaired from reduced LOD but that the resultant reduction in frame rendering times aided manual interaction with the environment. In both experiments, using a medium degree of LOD reduction resulted in a decrease of overall reaction time, i.e., detection plus localization time. These results indicate that perceptually adaptive LOD reduction techniques can be effectively used even on desktop systems.

While there is a optimal trade-off between task performance benefits and costs, we did not make an explicit measure of the perceptual quality of the display. Did participants notice the LOD reduction? The optimal LOD reduction levels in both experiments were aggressive, in that the reduced LOD was within the acuity limits of the visual system, but this does not mean that the LOD reductions were necessarily perceived. We did ask participants if they notice anything unusual about the display in a written, post-experiment questionnaire. Interestingly, none of the participants reported anything unusual even after being verbally prompted by the experimenter to try and answer the question. However, a number of the participants reported that they did notice something, most frequently in the gaze-contingent displays, when explicitly told about the experimental manipulation during the debriefing. This anecdotal evidence suggests that although visible, the LOD reduction may not significantly harm the perceptual quality of the display.

Why didn't the reduced LOD adversely affect localization times as they did the detection times? We believe that the effects of LOD reduction only influence tasks where the fine grain details of objects are critical. Detection in the visual search task requires participants to differentiate similar targets, and is thus affected by a reduced LOD. On the other hand, target localization depends on the large scale features of the object, and thus is not adversely affected by the reduced LOD. However, target localization does depend on interactivity and visual feedback. The results from both experiments show that the increased frame rendering rates facilitated localization performance.

A final result of interest is that the localization times in the Experiment 1 are approximately 2.5 seconds slower than in Experiment 2. We speculate that this slowing is a general strategy adopted by participants in Experiment 1 to deal with the highly variable frame rates that occur as participants position the virtual viewport. Given the importance of visual feedback in the localization task, this variability may have made interaction more difficult and even unnatural. This effect needs more investigation.

We used a simple edge-collapse technique to reduce LOD. The ability of this reduction technique to maintain a high degree of perceptual quality for heavily reduced models is limited. However, edge collapse is a widely used technique and serves as the basis for many more sophisticated reduction algorithms. Thus, the overall results of this study will likely generalize to other reduction techniques. LOD was also reduced isotropically which means that no preference was given to a particular portion of a model in the geometric simplification process. An improvement could be made by reducing model complexity anisotropicly. For example, with gaze-contingent rendering, the portions of the model closest to the point of gaze could have a higher complexity than farther portions (e.g., see Murphy and Duchowski 2001). This type of improvement would be especially important when objects occupy a large portion of the display. It is likely that the use of more sophisticated reduction techniques will limit the costs associated with LOD rendering techniques.

The results of this study complement the results of a few other notable studies. Watson et al. [2000; 2001] investigated the influence of LOD on the time required to name a model. They found that a strong LOD reduction lead to an *increase* in the time required for participants to name a model, indicating that geometric distortion of a model decreases identifiability. While these studies did not examine a real-time LOD display, they clearly indicate the potential for reduced LOD to negatively impact user performance with such displays. Interestingly, moderate amounts of LOD reduction actually *decreased* naming times for some models. They hypothesized that this effect was due to fine-grained detail hampering a coarse-grain perceptual categorization. The complexity and task dependence of these effects indicate that more behavioral research will be needed before we can begin to take full advantage of perceptually adaptive rendering using LOD reduction techniques.

7 Acknowledgements

We thank Tim Sweeney at Epic Games for providing access to the Unreal^{TM} rendering engine. We also thank Irwin Law for aid in setting up the virtual environments and an anonymous reviewer for helpful comments. This research was supported by an NSF CAREER grant IBN9876271 and a NIMH fellowship to DP.

References

BROWN, R., COOPER, L., AND PHAM, B. 2003. Visual attention-based polygon level of detail management. In *Proceedings of the 1st international conference on Computer graphics and interactive techniques in Austalasia and South East Asia*, ACM Press, 55–62.

BURR, D. C., AND ROSS, J. 1982. Contrast sensitivity at high velocities. *Vision Research 22*, 479–484.

CLARK, J. H. 1976. Heirarchical geometric models for visible surface algorithms. *Communications of the ACM 19(10)*, 547–554.

CRUZ-NEIRA, C., SANDIN, D. J., AND DEFANTI, T. A. 1993. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, ACM Press, 135–142.

DANFORTH, R., DUCHOWSKI, A. T., GEIST, R., AND MCALILEY, E. 2000. A platform for gaze-contingent virtual environments. In *Proceedings of the American Association for Artificial Intelligence, Smart Graphics Symposium*, AAAI press, vol. 4, 66–70.

DUNCAN, J., AND HUMPHREYS, G. 1989. Visual search and stimulus similarity. *Psychological Review 96*, 433–458.

EGETH, H., VIRZI, R., AND GARBART, H. 1984. Searching for conjunctively defined targets. *J. Experimental Psychology 10*, 1, 32–39.

FUNKHOUSER, T., AND SEQUIN, C. H. 1993. Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In *Computer Graphics (SIGGRAPH'93 proceedings)*, ACM, vol. 27, 247–254.

HOPPE, H., DEROSE, T., DUCHAMP, T. MCDONALD, J., AND STUETZLE, W. 1993. Mesh optimization. In *Computer Graphics (SIGGRAPH'93)*, ACM press, New York,NY, 19–26.

HOPPE, H. 1997. View-dependent refinement of progressive meshes. In *Computer Graphics and Interactive Techniques (SIGGRAPH'97)*, ACM press, New York,NY, 189–198.

LEVOY, M., AND WHITAKER, R. 1990. Gaze-directed volume rendering. In *Proceedings of the 1990 symposium on Interactive 3D Graphics*, ACM, 217–223.

LINDSTROM, P., AND TURK, G. 2000. Image-driven simplification. *ACM Trans. Graph. 19*, 3, 204–241.

LUEBKE, D., AND ERIKSON, C. 1997. View-dependent simplification of arbitrary polygonal environments. In *Computer Graphics and Interactive Techniques (SIGGRAPH'97 proceedings)*, ACM, 199–208.

LUEBKE, D., AND HALLEN, B. 2001. Perceptually driven simplification for interactive rendering. *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, 223–234.

LUEBKE, D., HALLEN, B., NEWFIELD, D., AND WATSON, B. 2000. Perceptually driven simplification using gaze-directed rendering. *University of Virginia Technical Report CS-2000-04*.

MURPHY, H., AND DUCHOWSKI, A. T. 2001. Gaze-contingent level of detail rendering. In *EuroGraphics 2001*, EuroGraphics Association.

MURPHY, B. J. 1978. Pattern thresholds for moving and stationary gratings. *Vision Research 18*, 521–530.

OHSHIMA, T., YAMAMOTO, H., AND TAMURA, H. 1996. Gaze-directed adaptive rendering for interacting with virtual space. In *VRAIS'92*, vol. 26(2), 103–110.

O'SULLIVAN, C., AND DINGLIANA, J. 2001. Collisions and perception. *ACM Trans. Graph. 20*, 3, 151–168.

PARKHURST, D., AND NIEBUR, E. 2002. Variable resolution displays: a theoretical, practical and behavioral evaluation. *Human Factors 44*, 4, 611–29.

PAUSCH, R., PROFFITT, D., AND WILLIAMS, G. 1997. Quantifying immersion in virtual reality. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 13–18.

REDDY, M. 1994. Reducing lags in virtual reality systems using motion-sensitive level of detail. In *Proceedings of the 2nd UK VR-SIG Conference*, 25–31.

REDDY, M. 1997. Specification and evaluation of level of detail selection criteria. *Virtual Reality: Research, Development and Applications 3(2)*, 132–143.

REDDY, M. 2001. Perceptually optimized 3d graphics. *IEEE Computer Graphics and Applications 21(5)*, 68–75.

SCHROEDER, W. J., ZARGE, J. A., AND LORENSEN, W. E. 1992. Decimation of triangle meshes. In *Computer Graphics (SIGGRAPH'92)*, vol. 26(2), 65–70.

STAMPE, D. M. 1993. Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, and Computers 25*, 2, 137–142.

TURK, G. 1992. Re-tiling polygonal surfaces. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, ACM Press, 55–64.

VIRSU, V., AND ROVAMO, J. 1979. Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research 37*, 3, 475–494.

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 Trans. Comput.-Hum. Interact. 4*, 4, 323–346.

WATSON, B., WALKER, N., RIBARSKY, W., AND SPAULDING, V. 1998. Effects of variation in system responsiveness on user performance in virtual environments. *Human Factors 40(3)*, 403–414.

WATSON, B., FRIEDMAN, A., AND MCGAFFEY, A. 2000. Using naming time to evaluate quality predictors for model simplification. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, 113–120.

WATSON, B., FRIEDMAN, A., AND MCGAFFEY, A. 2001. Measuring and predicting visual fidelity. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, ACM Press, 213–220.

WILLIAMS, N., LUEBKE, D., COHEN, J. D., KELLEY, M., AND SCHUBERT, B. 2003. Perceptually guided simplification of lit, textured meshes. In *Proceedings of the 2003 symposium on Interactive 3D graphics*, ACM Press, 113–121.

XIA, J. C., EL-SANA, J., AND VARSHNEY, A. 1997. Adaptive real-time level-of-detail-based rendering for polygonal models. *IEEE Transactions on Visual. Comput. Graph. 3(2)*, 171–183.

YEE, H., PATTANAIK, S., AND GREENBERG, D. P. 2001. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics 20*, 1, 39–65.