# Chapter 5

# *Results*

---

*'To tell us that every species of things is endowed with an occult specific quality by which it acts and produces manifest effects, is to tell us nothing'*

(Sir Isaac Newton, *Opticks*)

---

A prototype implementation was developed for the perceptually-based LOD model that we have just described. This chapter will consequently present an empirical evaluation of the prototype system. To fulfil this objective, a test suite of experiments was devised to evaluate:

1. the accuracy of the model,
2. the effects of the model on the user, and
3. the effects of the model on system performance.

These three goals were achieved respectively via a number of psychophysical studies, user studies, and computational studies. This chapter is therefore structured into three sections, covering each of these topics in turn.

## 5.1 Psychophysical Studies

This section describes a number of psychophysical experiments which were performed as part of the evaluation suite. Psychophysics is the branch of psychology which is concerned with establishing quantitative relations between physical stimulation and perceptual events. Therefore, the following experiments attempt to analyse the effect of our LOD modulation on the user's perception of the scene. That is, to assess the accuracy of our model in its goal of modulating detail without the user perceiving any visual change.

As with most psychophysical experiments, the stimuli used must be very simple. This is necessary in order to isolate the experiment to the particular condition under analysis, and to reduce the influence of any extraneous or obfuscatory variables.

We will evaluate the three facets of our perceptual model independently so that we may gain some insight into the efficiency of each individual component. We will therefore look at how well our model predicts the visibility of stimuli based upon their size, their eccentricity, and their velocity.

### 5.1.1 Object Size

#### 5.1.1.1 Objective

This first experiment aims to establish how well our perceptual model can predict the optimal LOD switching thresholds based upon the size of an object in screen coordinates. That is, to efficiently circumvent the popping effect which is often associated with traditional size or distance LOD techniques.

#### 5.1.1.2 Method

For each object's LOD, the optimal switching threshold was determined in terms of object size. This was then compared with the estimated results from our perceptual model. The following definitions were made:

- The size of an object was defined as the length of the radius of the object's smallest bounding sphere once it had been projected into screen coordinates.
- The 'optimal' switching threshold was evaluated using an automated brute force method as follows: the LOD model being tested and the next highest

LOD were displayed under identical conditions. An interval halving algorithm was used to converge on the shortest distance at which there was no discernible difference between the display of the two LOD models, i.e. if the two images were perceptually dissimilar, then the object was moved further away (object becomes smaller). If the two images were perceptually identical, then the object was moved closer (object becomes larger). The distance translated was halved at each iteration and the process terminated as this step size tended towards zero.

- The question of how to decide whether two images are 'perceptually identical' is an interesting one. As Rushmeier *et al*. (1995) note, the simplest procedure would be to conclude that two images are identical if and only if every corresponding pair of pixels in each image is identical. This will of course provide a somewhat over cautious measure because very small differences between a pixel's colour may not be discernible.

  The metric which was adopted here was therefore to define two images as being perceptually identical if and only if every corresponding pair of pixels were perceptually identical. We used the CIELUV colour difference equation as a measure of whether two pixels were perceptually identical, i.e. images $I_1(x, y)$ and $I_2(x, y)$, both of dimensions $M \times N$, were deemed perceptually identical if:

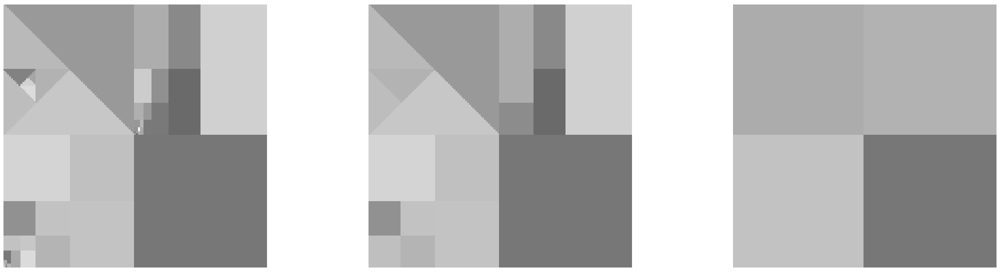
  $$\Delta E^*_{uv}(x, y) < \epsilon, \quad \text{where} \quad 0 \leq x < M, 0 \leq y < N, \tag{5.1}$$

  where $\Delta E^*_{uv}(i, j)$ represents the CIELUV colour difference between pixels $I_1(i, j)$ and $I_2(i, j)$. We chose a threshold value of $\epsilon = 5$ CIELUV units to provide a more stringent measure of colour difference than we have used previously. The value of $\Delta E^*_{uv}(x, y)$ was size corrected for a stimulus size of 0.1 degrees using Silverstein and Merrifield's (1985) formulae (this represented a display region of $\sim$2 pixels). Finally, the computer predicted thresholds were confirmed by visual inspection.

- The estimated results from our perceptual model were calculated in the manner described in the previous chapter. Each LOD was displayed so that it occupied the entire display (a display resolution of 700 $\times$ 700 pixels was used). The size of the object and all of its spatial frequencies were recorded in each instance, and then transition spatial frequencies were calculated to characterise switching conditions for each LOD.

#### 5.1.1.3 Results

Two objects were used for this experiment. The first object was a very simple 2D test card which contained a number of randomly coloured triangles, squares, and rectangles at various scales. This object was chosen to investigate our model's basic ability to extract perceptual features and to subsequently predict their visibility. Figure 5.1 illustrates three different levels of detail for the test card object.



**(a)** **(b)** **(c)**

Figure 5.1: *The test card pattern at three different levels of detail. Eight levels of detail were generated for this model. The figure shows models 8, 4, and 2, respectively.*

In total, eight levels of detail were produced for the test card object. All of the lower levels of detail were generated automatically from the highest LOD using the perceptually-driven LOD generation paradigm advocated in Section 3.3. The results from the experiment are presented in Figure 5.2.

The second object was a simple 3D representation of a die. This was used in order to test our model with a more realistic, three-dimensional stimulus; and also to show the model performing on different viewpoints of the same stimulus. The die object consisted of three levels of detail. Sample results are provided in Figure 5.3.

#### 5.1.1.4 Discussion

From the results of this experiment we can observe that the predicted switching thresholds compare very favourably with the optimal switching thresholds.

We note that there is an occasional discrepancy between the predicted and optimal thresholds. This could be due to small inaccuracies in the brute force

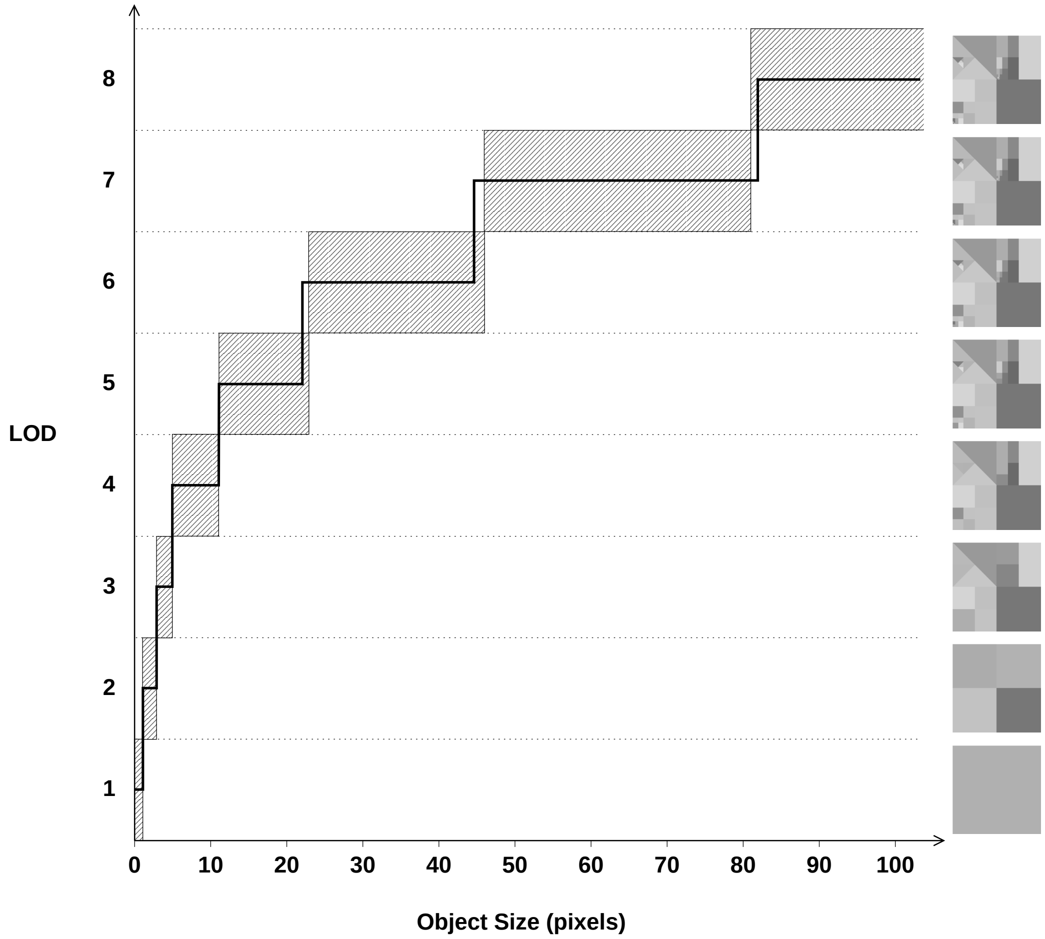
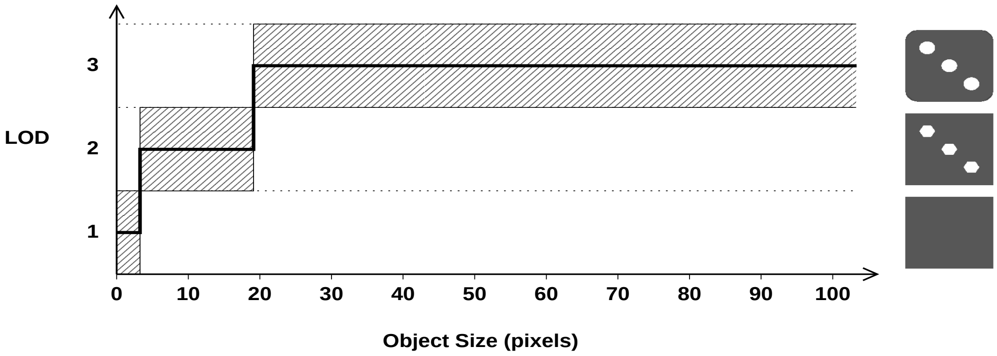


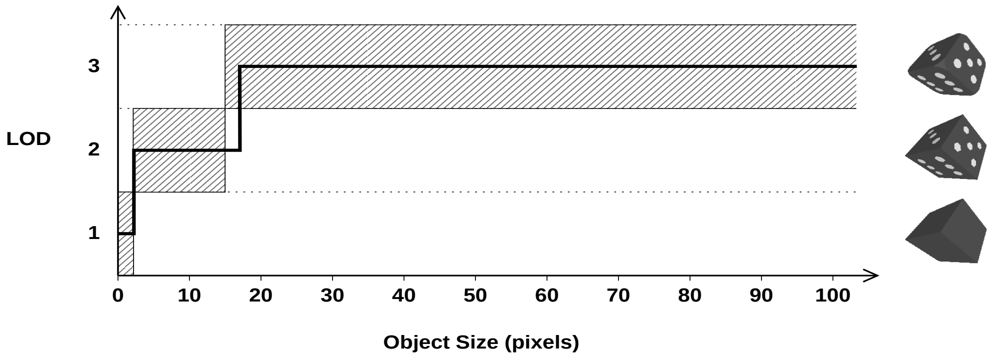

Figure 5.2: *Actual and projected switching thresholds (based upon object size) for a test card pattern. The shaded regions represent optimal switching ranges as found by the brute force algorithm. The solid line represents the projected switching thresholds estimated from the spatial frequency profile (object analysed at size = 493 pixels).*

algorithm and/or the spatial frequency extraction process. However, having noted this, the discrepancies are only in the order of 1 to 2 pixels: around 0.2% of the display resolution. This is certainly an acceptable error margin and the author can attest that no significant or distracting popping effects were experienced.

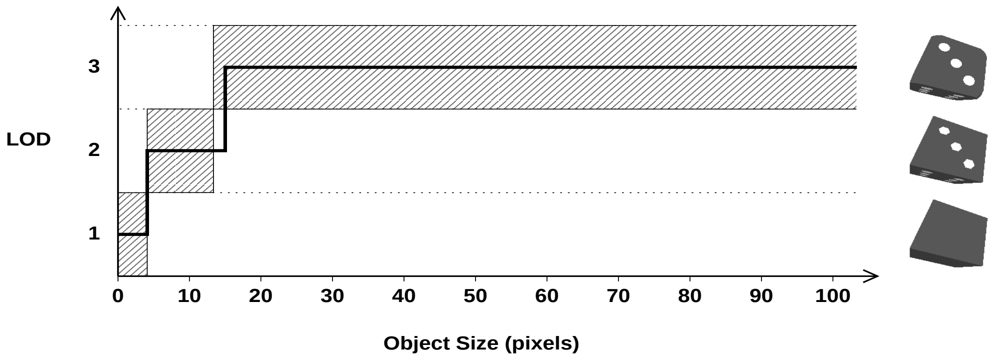
We can therefore conclude that our model is able to efficiently and accurately locate optimal switching sizes for different LODs of an object with minimal visual artifacts.




**(a)**




**(b)**




**(c)**

Figure 5.3: *Actual and projected switching thresholds (based upon object size) for a model of a die at three orientations. The shaded regions represent optimal switching ranges as found by the brute force algorithm. The solid line represents the projected switching thresholds estimated from the spatial frequency profile (object analysed at size = 360 pixels).*

### 5.1.2 Object Eccentricity

#### 5.1.2.1 Objective

This next experiment was devised to assess the effective threshold eccentricity for a number of 2D, aperiodic stimuli at a fixed contrast and spatial frequency. This was performed in order to establish how well our perceptual model can predict the user's ability to resolve detail in their peripheral field.

#### 5.1.2.2 Overview

The user was presented with a number of simple stimuli at various points in their peripheral field. The experiment was devised to locate the user's threshold eccentricity for a stimulus of fixed size and contrast. This was then compared with the predicted threshold from our perceptual model.

**The 2AFC Method**

The design of this type of experiment is crucial. It has been shown that presenting a single stimulus to a subject and asking whether they can see it—replying 'yes' or 'no'—is a flawed approach. This is because most subjects appear to have a predisposed readiness to report positive responses (Lamming, 1991a). Therefore, the generally recommended method for determining signal detection thresholds is the **2 Alternative Forced Choice** (2AFC) method. This involves presenting the subject with the test stimulus in one of two successive observation intervals and asking them to decide which interval it occurred in, guessing if necessary (Lamming, 1991b). For example, presenting the stimulus to the left or to the right of the user's gaze.

When the stimulus is above threshold then the user will always identify the stimulus interval correctly. However, when the stimulus is below threshold the user will have to guess and their performance will drop to chance level ($\sim$50%). Again, when guessing, the observer may exhibit a predisposed tendency to choose the first (i.e. left) interval. It has been shown that this can be countered by randomly varying the stimulus interval (Simpson, 1989).

**The Staircase Method**

There are two principal procedures for executing a 2AFC experiment. These are the **constant stimuli** method and the **staircase** method. In our present context, the former would mean selecting a number of eccentricities (normally five) and only testing the subject with these values. Alternatively, the staircase method adaptively selects new eccentricities to analyse based upon the subject's previous responses; eventually converging around the threshold value. The staircase methodology is often considered the more efficient and robust approach (Levitt, 1971), and so we will adopt this technique.

The operation of the staircase method (also known as the **Transformed Up–Down** method) can be described as follows (Wetherill and Levitt, 1965):

1. Choose an initial eccentricity (our best *a priori* estimate, e.g. the predicted result from our model).
2. If the subject gets two observations correct, then make the next condition more difficult (i.e. larger eccentricity).
3. If the subject gets one observation wrong, then make the next condition more easy (i.e. smaller eccentricity).
4. The experiment terminates after 6–8 reversals, where a reversal is defined as the turning point after a sequence of stimulus changes in the same direction.

The final result is then found by averaging the values at each of the reversals. The asymmetric nature of the algorithm (i.e. two correct responses induce an increase in difficulty, but only one incorrect response is required to decrease the difficulty) is designed to locate the eccentricity at which, statistically, 70.7% positive responses are obtained. This is normally taken as a measure of detection threshold.

#### 5.1.2.3 Method

*Stimuli*. The stimuli were displayed at full contrast (a white stimulus on a 2 deg black patch) on a mid-grey background. The display consisted of two black patches, one to the left and one to the right of the crosshair (both patches were equidistant from the crosshair). The stimulus was displayed randomly in either the left or the right interval and retained the same contrast and spatial frequency throughout the experiment; only varying in the eccentricity at which it was displayed to the user. Figure 5.4 illustrates the stimuli used for this experiment.

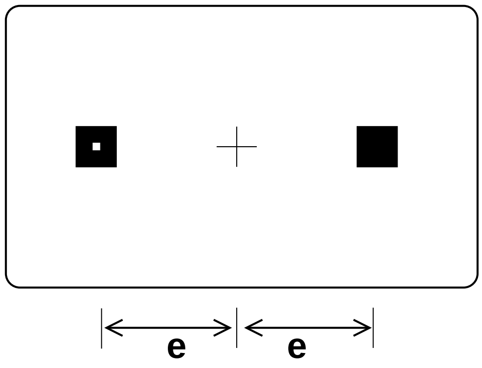



Figure 5.4: *The display used for the eccentricity 2AFC test. The stimulus was presented at eccentricity* e*, in either the left or the right patch (shown here on the left).*

The experiment was run for a number of different sized stimuli (1 and 2 pixels square) at a number of different viewing distances (10, 20, 30, 40, 50, and 60 cm). Overall, seven different spatial frequencies were analysed: 2.52, 5.04, 7.11, 9.50, 12.05, 14.68, and 17.36 c/deg.

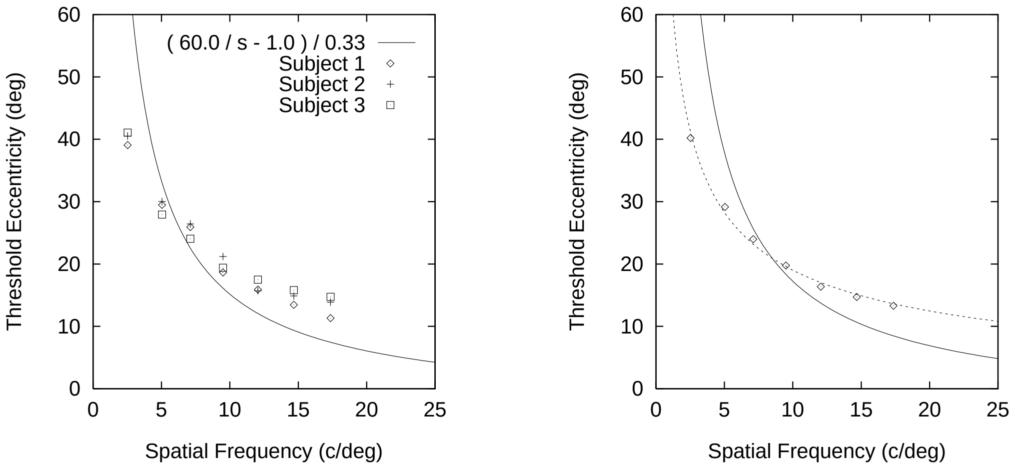
A staircase algorithm was employed to converge upon the user's threshold eccentricity for each stimulus. The initial step size was 1 deg. This was reduced to 0.5 deg after the first reversal, and further reduced to 0.25 deg after 5 reversals. The experiment terminated after 10 reversals.

*Procedure*. Three subjects, with normal or corrected-to-normal vision, performed each experiment twice. All subjects were unpaid male postgraduate students in the Department of Computer Science at the University of Edinburgh.

Each subject fixated on a crosshair displayed at eye level on the computer monitor. Eye movements were not monitored, but a chin rest was used in order to restrict head movement and to preserve the viewing distance. The subject was asked to chose the correct interval by pressing one of two buttons on the keyboard. They could abort any observation with another button if desired and this data would not be included in the final result.

Subjects were allowed a number of test runs beforehand in order to acquaint themselves with the experimental technique. The final threshold figures were found by averaging each pair of results from all subjects.

#### 5.1.2.4 Results




**(a)** **(b)**

Figure 5.5: *Results of the eccentricity experiment. (a) presents the individual responses of each of the three subjects, while (b) presents the averaged response for all subjects (the data points) and the best fit curve to this average response (the dotted line: Equation 5.3). The continuous line in (a) and (b) represents the predicted response from our model, given by Equation 5.2.*

The results from this experiment are presented in Figure 5.5. In order to calculate the predicted threshold response, we used our cortical magnification factor, defined by Equation 3.20. This can be written in terms of spatial frequency ($\alpha$) as:

$$E = (M_0/\alpha - 1)/0.29, \tag{5.2}$$

where $M_0 = 60$ c/deg: the highest spatial frequency which can be resolved at the fovea (Campbell and Gubisch, 1966).

The experiment produced results which were in the correct order of magnitude, but somewhat displaced from our predicted curve. A best fit curve was therefore calculated which represented the data more closely. This curve is defined in Equation 5.3 below, and plotted in Figure 5.5(b).

$$E = (21.2/\alpha^{0.5} - 1)/0.3. \tag{5.3}$$

We can use this curve to formulate a new model for our cortical magnification factor by first rewriting Equation 5.3 in terms of spatial frequency, e.g.

$$\begin{aligned} \alpha &= (21.2/(0.3E+1))^2 \\ &= 449.44/(0.3E+1)^2, \end{aligned} \tag{5.4}$$

and then normalising this so that a value of 1.0 is returned at the fovea and values tending towards 0.0 are returned for highly eccentric locations. We have already reported that the highest spatial frequency visible at the fovea is 60 c/deg. Using Equation 5.3 we find that our best fit curve does not go below this spatial frequency until an eccentricity of 5.79 deg is exceeded. We can therefore present our new equation for cortical magnification as follows. This equation is plotted in Figure 5.6.

$$M = \begin{cases} 1.0, & \text{when } E \leq 5.79 \\ 449.44/((0.3E+1)^2 \times 60) & \\ = 7.49/(0.3E+1)^2, & \text{when } E > 5.79. \end{cases} \tag{5.5}$$

#### 5.1.2.5 Discussion

The results from this experiment are very encouraging. They display a clear and smooth decline in spatial frequency sensitivity with increasing eccentricity. The results for all subjects lie consistently within 2 deg of the average result, and mostly within 1 deg. Given the inherent variability of individuals' vision systems, and taking into consideration any experimental error, this is a very good result. However, it is apparent that the empirical data, although in the correct order of magnitude, does not exactly match the theoretical threshold. This could be due to inaccuracies in our theoretical model (Equation 3.20), however, the most likely reason for the discrepancy is that, being based solely on retinal ganglion cell distributions, Rovamo and Virsu's (1979) model may not encapsulate the total processing performed by the entire visual system.

We therefore produced a best fit curve (Equation 5.3) which better models the observers' decline of peripheral sensitivity, and subsequently formulated a new model of cortical magnification for our application (Equation 5.5).

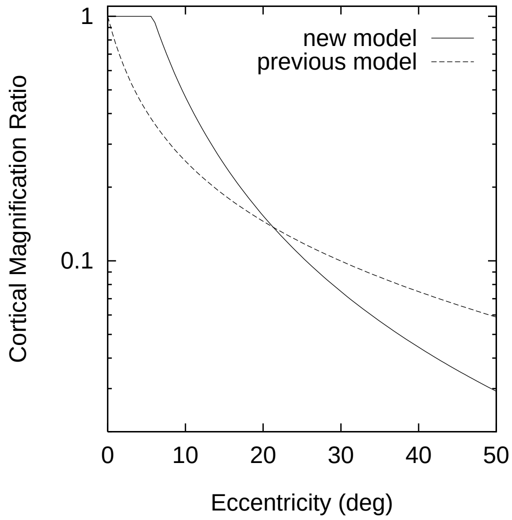



Figure 5.6: *Comparison between our new and previous definitions of the cortical magnification factor, $M$. The solid curve represents our new empirically derived model (Equation 5.5), whilst the broken curve represents our previous definition (Equation 3.20).*

### 5.1.3 Object Velocity

#### 5.1.3.1 Objective

This final psychophysical experiment was devised to assess various subjects' threshold velocity for a number of 2D, aperiodic stimuli at a fixed contrast and spatial frequency. This was performed in order to establish how well our perceptual model can predict the user's ability to resolve detail based upon the velocity of a stimulus.

#### 5.1.3.2 Method

*Stimuli*. Two black 2 deg patches were animated vertically past the observer's viewpoint (at 72 Hz) with a constant angular velocity; one to the left of the crosshair and one to the right. A white (full contrast) stimulus was randomly displayed in either the left or right interval during each trial and the observer had to chose which interval they thought the stimulus had appeared in. The stimuli were always presented at a horizontal angular distance of 2.5 deg from the crosshair in order to minimise the effect of eccentricity on detection (recall from Section 1.4.4.2 that our vision is maximally sensitive within the central $\sim$5 deg region. Also, from our results in Figure 5.6 we can observe that within 0 to

5.79 deg all potential stimuli would be visible under static conditions). Figure 5.7 illustrates the display used for the experiment.

A 2AFC staircase method was adopted as before. The initial step size for the staircase algorithm was 5 deg/s. This was reduced to 2.5 deg/s after the first reversal, 1 deg/s after 3 reversals, and 0.5 deg/s after 6 reversals. The experiment terminated after 10 reversals. Six different spatial frequencies were analysed: 5.04, 7.11, 9.50, 12.05, 14.68, and 17.36 c/deg.

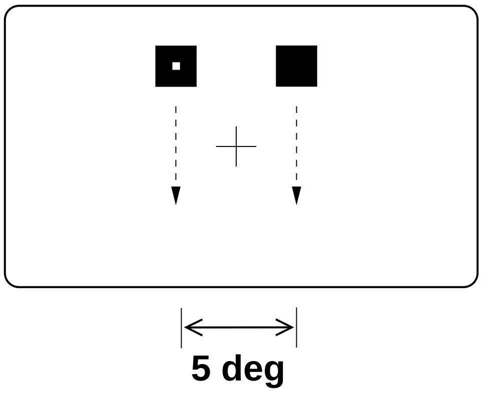


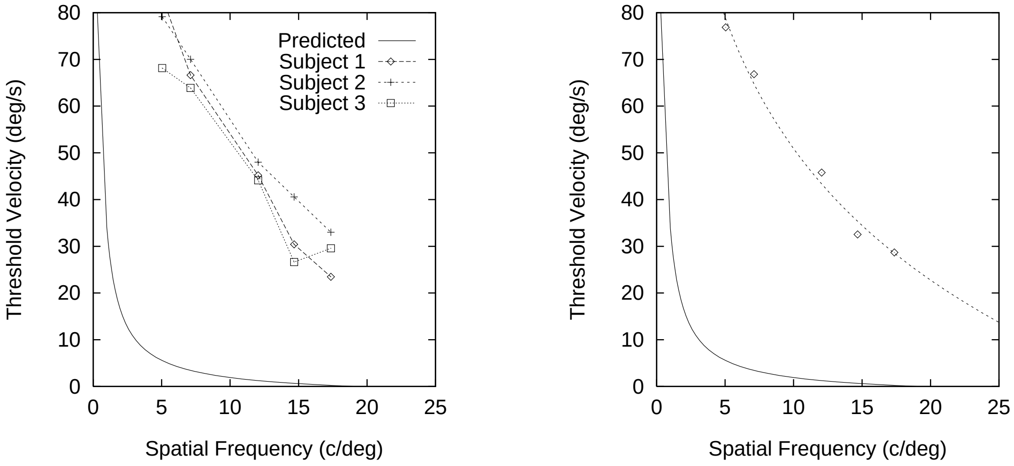

Figure 5.7: *The display used for the velocity 2AFC test. Two patches moved past the observer's fixation point at a constant angular velocity, with the stimulus being present in either the left or the right patch (shown here on the left).*

*Procedure*. The same three subjects who performed the previous experiment were solicited for this experiment. Each subject fixated upon a crosshair, positioned at eye level, with their head movement constrained by a chin rest. Once again, subjects were allowed a number of test runs beforehand in order to acquaint themselves with the experimental technique.

The experiment was performed twice for each spatial frequency. The final threshold figures were found by averaging each pair of results from all subjects.

#### 5.1.3.3 Results

The results of the velocity experiment are presented in Figure 5.8. In order to calculate the predicted threshold response, we used our spatiotemporal contrast sensitivity model (Equation 3.15) and found the upper root of the relation, $G(\alpha, v) - 1 = 0$, using an interval halving algorithm.




**(a)** **(b)**

Figure 5.8: *Results of the velocity experiment. (a) presents the individual responses of each of the three subjects, while (b) presents the averaged response for all subjects (the data points) and the best fit curve to this average response (the dotted line: Equation 5.6). The continuous line in (a) and (b) represents the predicted response from our model.*

As can be observed from the data in Figure 5.8, the results which were obtained are substantially deviant from our theoretical model. A best fit curve was therefore calculated for the experimental data (using a computer graphing package). An exponential curve fit was used in order to retain as much as possible the general characteristics of Kelly's (1975) abstract model (see Section 3.2). The resulting curve is presented in Equation 5.6 below, and plotted in Figure 5.8(b).

$$v = 119.353 \times 10^{-0.036\alpha}. \tag{5.6}$$

We can use this curve to help us formulate a new spatiotemporal threshold model. This requires that we find the corresponding form of Equation 5.6 in terms of spatial frequency, $\alpha$. That is,

$$\begin{aligned} \alpha &= \log_{10}(v/119.353)/-0.036 \\ &= (\log_{10}(v) - \log_{10}(119.353))/-0.036 \\ &= -27.78\log_{10}(v) + 57.69. \end{aligned} \tag{5.7}$$

Once again we can note that this equation only goes below 60 c/deg when $v > 0.825$ deg/s. We can therefore present the final definition for our spatiotemporal threshold model (for $v \geq 0$ deg/s) as follows. This equation is plotted in Figure 5.9 along with the previous model for comparison.

$$\alpha = \begin{cases} 60.0, & \text{when } v \leq 0.825 \\ -27.78\log_{10}(v) + 57.69, & \text{when } v > 0.825. \end{cases} \tag{5.8}$$

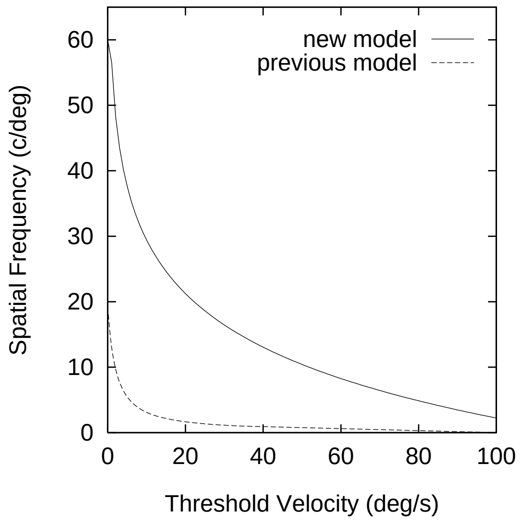



Figure 5.9: *Comparison between our new and previous definitions for the spatiotemporal threshold surface. The solid curve represents our new empirically derived model (Equation 5.8), whilst the broken curve represents our previous definition (using Equation 3.15).*

#### 5.1.3.4 Discussion

The results from this experiment are not as smooth as those obtained from the eccentricity experiment, or as consistent with the theoretical response. However, this was anticipated given the discussion in Section 3.2.4 on the problems of defining a standard spatiotemporal threshold surface. The most probable cause for the observed discrepancy is due to the fact that we are using very localised stimuli (in terms of the field of view occupied), whereas vision scientists normally deal with extended stimuli which fill a large proportion of the FOV. Also, the experiment was more complicated, both for the author to devise, and for the subjects to perform. In the first instance, it would be impossible to completely isolate the effect of eccentricity using a spatial 2AFC test because the stimulus must be presented at some displacement left or right from the fixation point;

also, because the stimulus is moving, its eccentricity is constantly changing. Secondly, the experiment is more difficult for the subjects to perform accurately because it is an instinctive reflex to fixate upon and track moving objects, thus altering their effective angular velocity.

We therefore produced a best fit curve (Equation 5.6) which models the observed decline of temporal sensitivity more closely; and subsequently, we formulated a new model of spatiotemporal threshold for our application (Equation 5.8). It may be noted that this equation, as well as providing a more accurate and practical model, is also significantly less complex to compute than our original model.

## 5.2 User Study

So far we have considered the accuracy of our model in the task of predicting the visibility of various stimuli. This next study will undertake an analysis of the effect of our model on the user when it is used to reduce detail in a VE. This will be done by employing an objective analysis of the user's ability to perform a certain task, under normal and optimised conditions.

### 5.2.1 Overview

The measure which was elected to assess the user's performance was a way-finding task, based upon the **simulated pursuit fixation** work of Cutting *et al.* (1992) and others (e.g. Warren and Hannon, 1990; Cutting, 1986; Rieger and Toet, 1985). This task involves the subject being passively transported through an environment of objects. The display is updated so that the observer is always looking towards a certain fixation point which is deviant from their heading direction. To illustrate this task in terms of a real world scenario, imagine that you are on the back of a jeep manning a TV camera which is free to pan left and right. The jeep is being driven through an environment cluttered with various objects, e.g. trees. You look through the camera's viewfinder and track one specific tree as the jeep moves through the environment, keeping the tree always in the centre of the viewfinder. Your gaze direction is now different from your heading direction. This concept is illustrated in Figure 5.10.

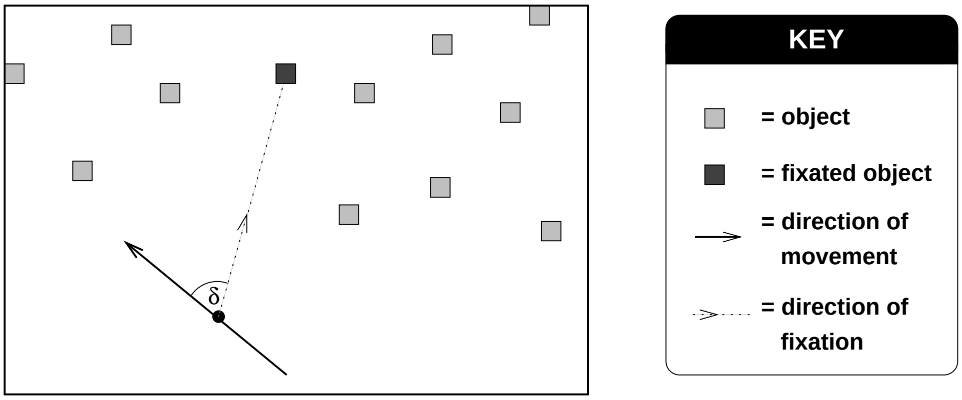



Figure 5.10: *Overview of the wayfinding task in which the subject is passively navigated through an environment with their direction of fixation oriented differently from their direction of movement. Angle $\delta$ represents the gaze/movement angle at one point on the navigated course.*

The task of the subject is to deduce whether they are being navigated to the left or to the right of the fixation point (based upon the radial motion cues which they acquire from the surrounding objects). For example, in Figure 5.10 the heading direction is to the left of the target. Intuitively, this task will become more difficult as the differential between the fixation and the heading vector (referred to as the **gaze/movement angle**) decreases. Typically, the gaze/movement angle will increase as one progresses along the navigated path. The maximum value of this angle for any trial is the independent variable of interest, i.e. the final gaze/movement angle.

From a number of subjects' responses to different navigation scenarios, Cutting *et al*. (1992) produced psychometric curves which plot the final gaze/movement angle against the percentage of correct responses. The point at which this curve drops below a certain threshold can be used to compare the subject's task performance under different situations. The reader should note that once again we are using a 2AFC paradigm (i.e. the subject is asked, 'did you move left or right?', as opposed to, 'did you move left?' or 'did you move right?').

This simulated pursuit fixation task was chosen as a measure of user performance because it offers a number of advantages which are particularly pertinent to our circumstances, e.g.

1. The subject must constantly fixate upon the centre of the display and so their physical gaze direction is always known. This therefore conveniently abates the requirement for a suitable eye tracking technology (which was not available to the author).

2. The environment is inherently motion-rich with many objects of different sizes and a good proportion of peripherally located features. It therefore offers good opportunity to exercise all aspects of our model.

3. This type of experiment, as well as having been used by numerous vision scientists in the past, has also been used by various researchers in the field of VR to assess the performance of subjects within a VE (e.g. Wann *et al.*, 1995). The technique is therefore an accepted and stable metric for assessing user performance.

### 5.2.2 Method

*Stimuli*. The test VE was generated using a Silicon Graphics, Inc. (SGI) Onyx RealityEngine$^2$ computer with one 200 MHz R4400 micro-processor, 128 MB of RAM and one Raster Manager (RM4) card with 4 MB of texture memory. The experiment was developed using IRIS Performer V2.0 for peak performance and used 'free running' phase mode (i.e. the application output a new frame as soon as it had completed the previous frame and was not constrained to a fixed frame rate).

The content and dimensions of the environment were modeled in order to replicate Cutting *et al.*'s (1992) experimental setup. Specifically, each environment contained 175 objects, randomly positioned at ground level (across the x–z plane), and randomly rotated (about the y axis). The fixation object was coloured light purple and a crosshair was positioned over it to guide the user's fixation.

Subjects were navigated through the environment for up to 5 seconds, but could submit their decisions at any point. The fixated object was initially positioned at a distance of 50 m from the viewpoint. Each object was 4.32 m high and the simulated forward velocity was 4.36 m/sec. The object which was used to populate the environment was the 'temple' model supplied by SGI. This was chosen because it was available in a convenient format and also because it contains a

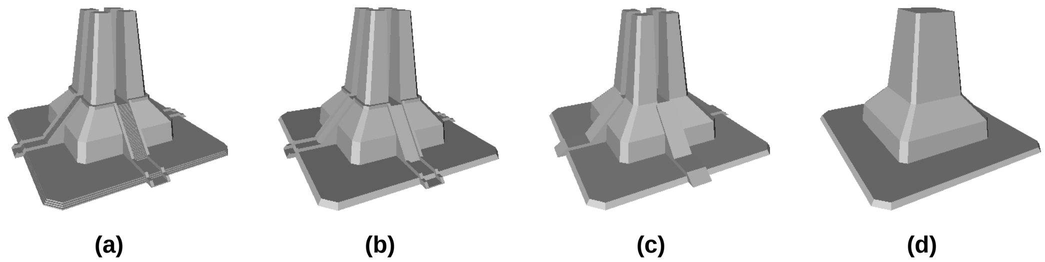


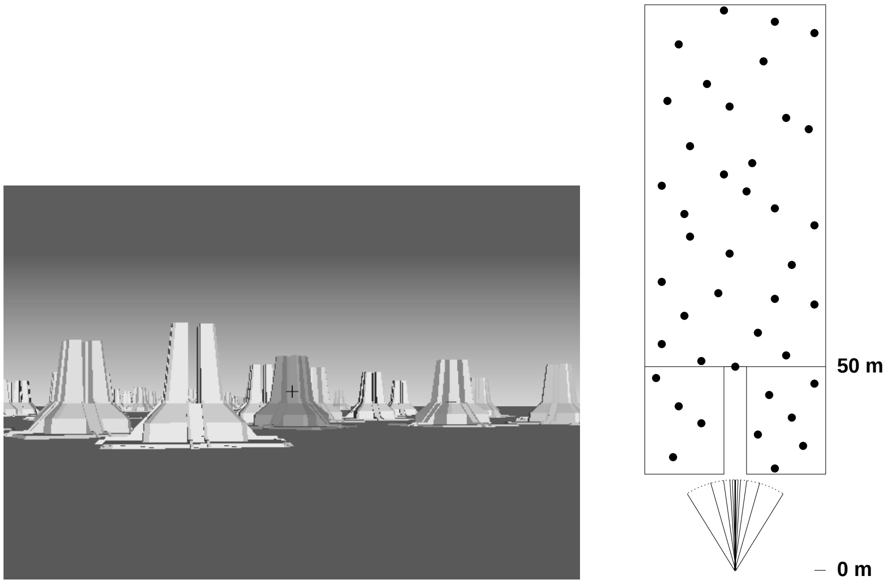

Figure 5.11: *The four levels of detail used for the object in the simulated pursuit fixation task. From left to right, these contain 3928, 834, 254, and 76 triangles respectively.*

broad range of spatial frequencies. Three further levels of detail for the test object were generated. These are illustrated in Figure 5.11 and a description of their spatial frequency content is provided in Appendix B. An overview and example screenshot of the test environment are provided in Figure 5.12.

Images were displayed at a resolution of 1280 $\times$ 1024 pixels. The subject viewed the screen such that it occupied 43.6 $\times$ 33.4 degrees of their field of view (and these FOV values were used by the graphics renderer for all perspective calculations). Steps were taken to ensure that no other users could remotely log into the workstation during the course of an experiment. This was done in order to limit the effect of any background processes which might interfere with the update rate of the experiment.

Sixty-four trials were randomly presented to each subject. These were composed of 8 final gaze/movement angles (0.25, 0.5, 1, 2, 4, 8, 16, and 32 degrees) $\times$ 2 gaze directions (left and right of the fixated object) $\times$ 2 LOD filtering conditions (with and without our perceptual optimisations) $\times$ 2 attempts. The model of threshold visibility which was used for the perceptual optimisations was based upon our new empirically-derived results (from earlier in this chapter), rather than the theoretical model we developed in Chapter 3.

*Procedure*. Twenty subjects participated in the study, drawn from a breadth of backgrounds including undergraduate students, postgraduate students, staff, and graduates of the University of Edinburgh. All subjects had normal or corrected-to-normal vision and were naïve to the experimental hypothesis. They were encouraged to keep their gaze fixated on the crosshair at the centre of the screen, but no attempt was made to monitor eye movements. A chin rest was used to restrict subjects' head movement and to maintain the viewing distance




**(a)** **(b)**

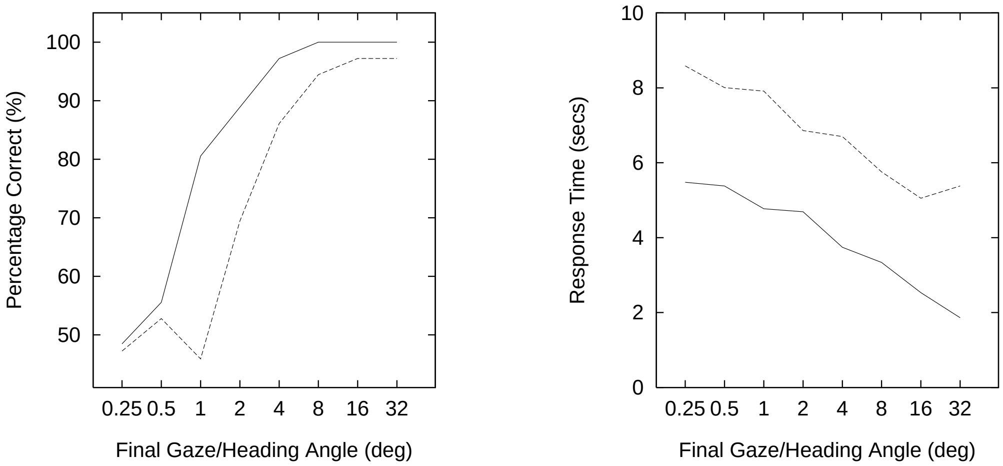
Figure 5.12: *An overview of the simulated pursuit fixation task. (a) presents an example screen shot of the test environment, and (b) provides a plan overview of the environment layout. The fan shape in (b) illustrates the various paths through the environment which were traversed. The fixation object can be noted at the centre of the 50m line.*

(which was set to 50 cm).

Subjects pressed either the left or right mouse button to indicate whether they felt that they had been transported to the left or to the right of the fixation point. Reaction times were recorded for each trial. No feedback was given to the subject on their success rate during the experiment to reduce the chances of them learning some nonsense visual task unrelated to wayfinding. Subjects were given a number of practice trials beforehand until they were satisfied that they understood the task. The experiment lasted about 20 minutes and participants were paid £5.00.

### 5.2.3 Results

Figure 5.13 presents the results of the simulated pursuit fixation experiment; averaged over the 20 subjects. These are compared for the case when our perceptual LOD optimisations are employed, and when no optimisations are used. On average, the act of implementing our optimisations produced a five-fold increase in frame rate (from 2.3 Hz to 11.5 Hz). It was anticipated that this performance increment would improve the subjects' ability to perform the heading task.




**(a)** **(b)**

Figure 5.13: *Averaged results from the simulated pursuit fixation experiment. (a) shows the overall proficiency of subjects to perform the task. (b) presents the overall response times of subjects. In both graphs, the solid line represents the case where perceptual LOD optimisations were employed; whilst the broken line represents the case where no optimisation was performed.*

Figure 5.13(a) shows the average proficiency of subjects to correctly ascertain their heading. We can see from this graph that at large gaze/heading angles, the user is very proficient at resolving their heading direction; but this ability drops to chance level (~50%) for smaller angles. It is immediately evident that subjects' efficiency was maintained higher for longer when our perceptual optimisations were employed.

Warren and Hannon (1990) used the interpolated final gaze/heading value corresponding to a 75% performance level in order to assess performance; whereas Cutting *et al*. (1992) proposed a more strict 95% criterion. We can therefore describe the subject's increase in performance in terms of these two definitions by referring to the data in Figure 5.13(a). Using Warren and Hannon's 75% threshold, subjects could discriminate final gaze/heading angles which were 3.0 times smaller (2.7 deg / 0.9 deg) using the perceptual LOD optimisations. Using Cutting *et al*.'s 95% threshold, subjects could discriminate angles which were 2.8 times smaller (9.5 deg / 3.4 deg) under optimised conditions.

Figure 5.13(b) shows the average response time of subjects during the task, i.e. how long it took them to resolve their heading direction. Again it is obvious from a cursory inspection that subjects had a distinctly faster response for cases when our perceptual LOD filtering was employed. From these data we find that, on average, users performed 1.67 times faster under optimised conditions.

### 5.2.4 Discussion

The findings of this task performance experiment show that for the prototype implementation of our perceptually-based LOD system (in comparison to the normal unoptimised case):

1. The frame rate of the simulation was increased (by around 5 times).
2. The subjects' ability to perform the heading task improved (they were able to resolve final gaze/heading angles up to 3 times smaller).
3. The time subjects took to perform each task decreased (they performed the task 1.67 times faster on average).

It is also worth noting that subjects did not report any visual discrepancies in the scenes which had perceptual LOD filtering applied, even when explicitly asked. It would therefore appear that we achieved our goal of reducing detail without affecting the user's percept of the VE.

## 5.3 Computational Studies

We summarised the goal of this doctoral thesis on Page 2 by stating that we would attempt to reduce lag in VEs by investigating perceptually-based LOD. We must therefore evaluate whether our system achieves this goal, i.e. to what degree does our prototype implementation reduce lag? An important issue which we should also address is that of the cost of implementing our perceptually-based system. Finally, it would also be interesting to investigate the extent to which each of the components of our model contribute towards any increase in performance.

All of the following computational studies used the simulated pursuit fixation experiment as their test application. These were run on the same platform as before (i.e. an Onyx RealityEngine$^2$ with one 200 MHz processor, 128 MB of RAM, and one RM4 card). All remote access to the machine was disabled for the duration of each study and all disk writes were made to the local disk, only upon completion of each trial.

All trials were performed under the same visual conditions as the previous task performance study, i.e. a display FOV of 43.6 $\times$ 33.4 degrees was assumed and the simulated forward velocity was 4.36 m/sec. All timings were recorded using a sub-microsecond resolution clock.

## 5.3.1 Analysis of Performance Speedup

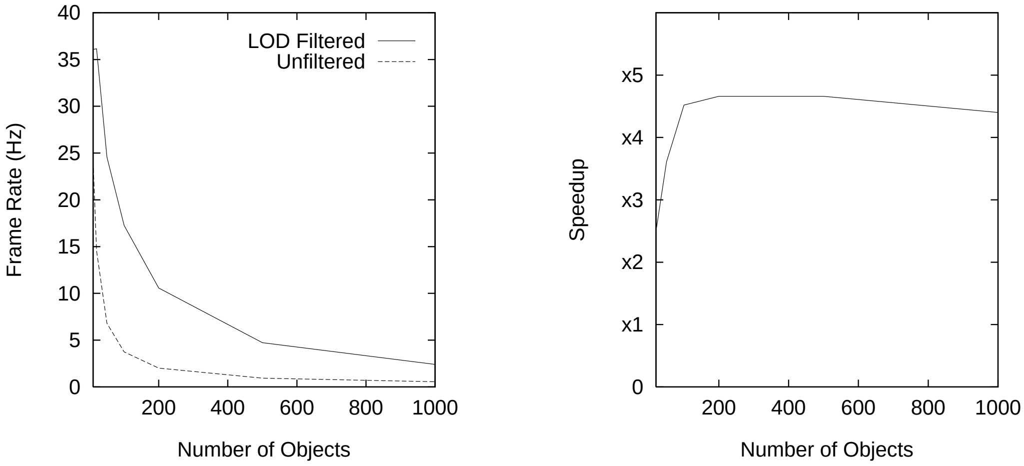
### 5.3.1.1 Objective

The aim of our first computational study was to assess the extent to which our prototype implementation reduces the latencies in a VE (i.e. the extent to which it improves performance). This was contrasted for a number of VEs with different numbers of objects in order to give an indication of how speedup varies with VE complexity for the test application.

### 5.3.1.2 Method

*Procedure*. The average frame rate (Hz) for a trial was found by dividing the number of frames rendered by the total time for the trial. Twenty trials were performed where the number of objects in the VE were kept constant (but their positions randomly varied). The final frame rate figure for such a set of trials was found by averaging the results of all 20 trials. Each set was repeated for the case were no LOD filtering was applied and when full perceptual LOD filtering was applied. Finally, a number of these sets of trials were performed for environments with different numbers of objects (10, 20, 50, 100, 200, 500, and 1000 objects). That is, 280 trials were performed in total (7 complexities $\times$ 2 LOD cases $\times$ 20 trials).

#### 5.3.1.3 Results & Discussion




**(a)** **(b)**

Figure 5.14: *Results from the performance speedup analysis. (a) contrasts the average frame rate for the test application under normal unfiltered conditions (the broken line) and when perceptual LOD optimisations were employed (the solid line). (b) presents the data from (a) in terms of the relative speedup which was achieved when using the perceptual LOD optimisations over the normal unfiltered case.*

The results from this study are presented in Figure 5.14. From Figure 5.14(a) we can observe a consistent and marked increase in frame rate when our perceptual optimisations were employed. To describe the degree of this increment, Figure 5.14(b) illustrates the relative increase in performance which occurred when our perceptual optimisations were used. From this we can see that an average speedup of over $\times 4.5$ was quickly achieved (for environments with around 50 objects) and that this depreciated very gradually for more complex environments ($> 500$ objects).

Care should be taken when interpreting the initial sharp rise in speedup for environments with $< 50$ objects. This does not necessarily imply that less complex environments offer fundamentally smaller speedup values. We must also consider the fact that under the current implementation of IRIS Performer, the frame rate is constrained to be an integer multiple of the video refresh rate. Our video refresh rate was 72 Hz and so if the simulation was unable to run at 72 Hz, then it would drop down to the next integer multiple of 72, i.e. 36 Hz. Therefore, what we are most likely observing in the sharp speedup rise for $< 50$ objects is the point where the optimised environment is capable of being rendered faster than 36 Hz, but not as fast as 72 Hz, and so it is restricted to only 36 Hz. The relative speedup is therefore confounded by this additional factor for simple environments.

## 5.3.2 Analysis of Scheduler Complexity

### 5.3.2.1 Objective

Having investigated the performance gain that our prototype offers, we now wish to gain some insight into the cost of implementing our perceptually-based LOD model. That is, how much CPU time is spent evaluating the perceptual content and consequent optimal LOD for each object? Also, how does this change as the complexity of the VE changes?

### 5.3.2.2 Method

*Procedure*. Code was added to the prototype implementation to record the time before and after the LOD modulation routine. This was taken as a pragmatic measure of the time spent implementing our perceptual model for each frame. The final figure for any one trial was found by averaging the times for every frame of the simulation. Twenty trials were performed where the number of objects in the VE were kept constant (but their positions randomly varied). The final figure for such a set of trials was found by averaging the results of all trials. A number of sets of trials were performed in order to find the average time complexity for environments with different numbers of objects (10, 20, 50, 100, 200, 500, and 1000 objects). That is, 140 trials were performed in total (7 complexities $\times$ 20 trials).

### 5.3.2.3 Results & Discussion

Figure 5.15 presents the results from the analysis of our prototype scheduler. We can consequently observe that the scheduler exhibits an essentially linear increase in processing requirements as the VE becomes more complex. For the host architecture which was used, these data imply that our scheduler would not begin to interfere with a 30 Hz frame rate until around 450 objects were introduced (assuming that no other intensive processing is performed per frame),

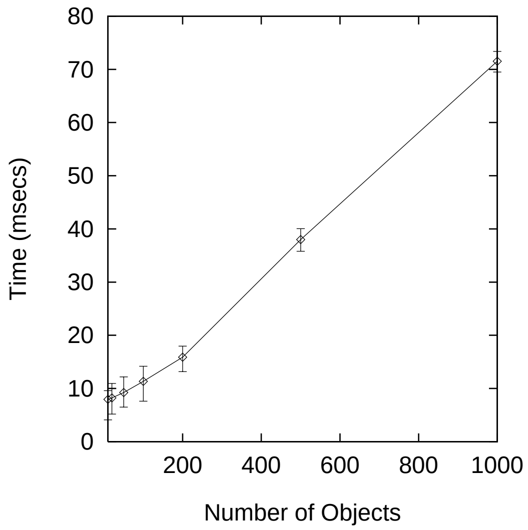



Figure 5.15: *The overhead of implementing our perceptually-based LOD system for a range of VE complexities. This includes the spatial frequency estimation as well as the calculation of object visibility and the subsequent LOD selection.*

and it would take over 900 objects to begin to interfere with a 15 Hz frame rate.

It is worth noting that when we say 900 objects, this does not necessarily mean a VE with only 900 objects. We are of course referring to *degradable* objects here (i.e. objects which have multiple levels of detail) and not all objects in a VE will be degradable. The number of non-degradable objects in a VE will have no effect on the computational resources consumed by the scheduler because these objects should never be considered by the scheduler.

It is also worth noting that these figures are for our prototype implementation. Being a prototype, the author was more concerned with the functionality of the implementation rather than its efficiency. It is therefore highly likely that a more efficient implementation could be developed. (For example, the results from our model of threshold visibility could be placed into a look up table rather than being computed directly each frame. Also, the author used the IrisGL character string routines to calculate the projected screen coordinates for each object. These return screen coordinates as a side-effect but also perform a number of other functions which are unrelated to our task. A more efficient implementation would therefore calculate these projected coordinates directly.)

### 5.3.3 Breakdown of Model Components

#### 5.3.3.1 Overview

In this final computational study we will attempt to gauge the extent to which each of the three principal components of our model are utilised. That is, how much does size LOD, eccentricity LOD, and velocity LOD contribute towards the total increase in performance for our test application.

It would be an unfair assessment to analyse the velocity and eccentricity components of the model in isolation because most of the benefit of these two components is realised through their synergic combination (we will show this in the next chapter). Therefore we shall compare the case where all optimisations are used, against the case where only the velocity and eccentricity based optimisations are used. In effect, this provides us with an indication towards the additional benefit which can be accrued over traditional size LOD when we also incorporate optimisations based upon the velocity and eccentricity of objects.

#### 5.3.3.2 Method

*Procedure*. The average level of detail (1–4) for a frame was found by dividing the sum of all objects' LOD by the number of objects. An average figure for an entire trial was found by dividing all of the frame averages by the total number of frames displayed. As before, 20 such trials were performed, and the average LOD figure from all of these was found. This measure was used as an assessment of the degree to which detail was optimised; with higher values indicating a greater degree of optimisation.

Two level of detail cases were compared: one where full optimisations were employed (size, velocity, and eccentricity), and the other where detail was reduced based only upon the velocity and eccentricity of objects. This relationship was analysed for environments of various complexities (10, 20, 50, 100, 200, 500, and 1000 objects). 280 trials were therefore performed in total (7 complexities $\times$ 2 LOD cases $\times$ 20 trials).

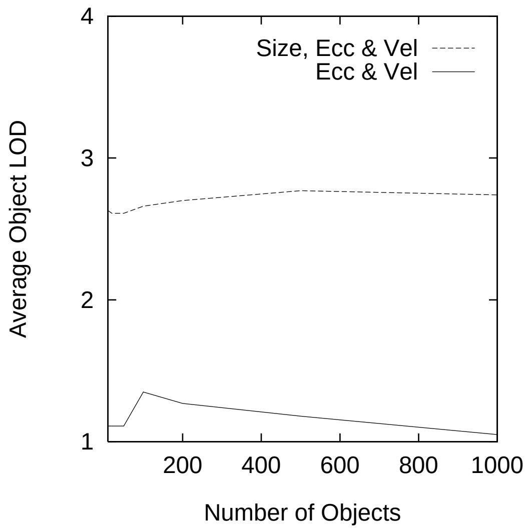



Figure 5.16: *Comparison of average LOD during a trial with and without size optimisations. (Recall that the four levels of detail for each object are illustrated in Figure 5.11.)*

#### 5.3.3.3 Results & Discussion

Figure 5.16 presents the data from this study. The most obvious result from this graph is that size optimisations appear to account for the bulk of any reduction of detail (and hence the increase in performance). The combined contribution of velocity and eccentricity optimisations rose to a maximum of 21% of the total LOD reduction for around 100 objects. This dropped down to 10% for 500 objects, and to as little as 3% of the total reduction for 1000 objects[1].

Of course, this analysis was performed for environments with a fixed FOV and (relatively slow) forward velocity. If we were to increase the forward velocity through the environment then the velocity component of our model would be given more chance to contribute towards the overall reduction in detail and performance should improve as a result.

It would be tempting to make the same analogy for FOV, but this would be a specious assumption. Although increasing the FOV of the display would allow objects to exist at greater eccentricities (and hence theoretically allow them to have their detail reduced further), we must also consider the angular resolution of the display. If we take the same display and simply stretch it over a wider

[1]This drop in the contribution of the velocity and eccentricity components might explain the gradual decline in speedup which was found for environments with > 500 objects, as shown in Figure 5.14(b).

field of view, then the individual pixels will become larger and so the angular resolution—and hence the highest visible spatial frequency—will drop. So, although the field of view is larger, objects must move further into the periphery before their detail drops below threshold and a lower LOD can be selected[2].

## 5.4 Summary

We have now completed the evaluation of our visibility model and prototype implementation, and can therefore assess the implications of our findings. (N.B. a deeper discussion of the various implications of these results will be performed in the next chapter.)

We began by investigating the accuracy of our model; employing various psychophysical studies to compare the perceived threshold of different stimuli with the predicted thresholds from our model. The resulting data showed that our model predicted the size threshold for objects with sufficient accuracy. The predicted eccentricity thresholds were very close to those found through experiment; however we produced a new equation (Equation 5.5) to better model the empirical data. In the case for velocity thresholds, we found quite a large discrepancy between the predicted and observed response. This was accounted for and a new, more simple model was devised (Equation 5.8) based upon the observed results.

Our formulation of these new models of visual acuity is of course valid because we are concerned with sufficiently different stimuli from those which normally preoccupy vision scientists. For example, we are concerned with local, aperiodic, non-harmonic stimuli whilst vision scientists normally deal with contrast gratings which are harmonic, periodic, and extend over a large field of view. We have therefore tailored our model to the genus of stimuli commonly found in computer-generated imagery.

The second section of this chapter presented a comparative analysis of a user's ability to perform a prescribed task. For the particular simulated pursuit fixation

[2]This problem poses a common dilemma in the field of head-mounted display design, where manufacturers must decide whether they wish to produce an HMD with high angular resolution or wide FOV, because the pixel resolution of current LCD technology is limited.

task chosen, we found that the act of implementing our perceptual LOD model improved the frame rate of the simulation markedly. As a result, the user's ability to accurately perform the task improved, as did the speed with which they could complete the task. All of this was achieved with no reported drop in image fidelity. These results vividly illustrate the merits of this research effort.

The final investigation focussed on the basic computational benefits which resulted from the implementation of our system. These studies were performed using the simulated pursuit fixation experiment and so the results pertain only to that particular application. However, this application is not completely artifical; for example, the behaviour and operation of the experiment are very similar to that of a driving simulator. The author therefore feels that these results provide a good representative indication of the general utility and performance which can be offered by perceptually-based LOD. To reinforce this, it is worth noting that various other researchers have achieved comparable performance increments using size and eccentricity LOD, e.g. Ohshima *et al*. (1996) experienced a five-fold improvement in frame rate for their LOD system, and Levoy and Whitaker (1990) experienced an improvement in rendering performance of $\times$ 4.6 for their perceptually optimised ray tracer.

We found that the frame rate of the simulation was improved dramatically by our system (around four to five-fold speedup), and that this improvement was maintained for VEs over a large range of complexities. We also saw that the implementation of our system can be made very efficient, such that it would not affect the update rate of the simulation until an excessively large number of objects is introduced (in the order of 1000 degradable objects). Finally, we found that out of the three major components of our model (size, velocity, and eccentricity), the accommodation of size-based optimisations provided the largest contribution towards the subsequent improvement in performance.