

Using Naming Time to Evaluate Quality Predictors for Model Simplification

Benjamin Watson

Dept. Computing Science
615 General Services Bldg.
University of Alberta
Edmonton, Alberta
Canada T6G 2H1
+1 780 492 9918
watsonb@cs.ualberta.ca

Alinda Friedman

Dept. Psychology
University of Alberta
Edmonton, Alberta
Canada T6G 2E9
+1 780 492 2909
alinda@ualberta.ca

Aaron McGaffey

Dept. Psychology
University of Alberta
Edmonton, Alberta
Canada T6G 2E9
+1 780 492 2909
mcgaffey@gpu.srv.ualberta.ca

ABSTRACT

Model simplification researchers require quality heuristics to guide simplification, and quality predictors to allow comparison of different simplification algorithms. However, there has been little evaluation of these heuristics or predictors. We present an evaluation of quality predictors. Our standard of comparison is naming time, a well established measure of recognition from cognitive psychology. Thirty participants named models of familiar objects at three levels of simplification. Results confirm that naming time is sensitive to model simplification. Correlations indicate that view-dependent image quality predictors are most effective for drastic simplifications, while view-independent three-dimensional predictors are better for more moderate simplifications.

Keywords

Model simplification, simplification metrics, image quality, naming time, human vision.

INTRODUCTION

As the number of methods available for constructing or capturing three dimensional (3D) polygonal models proliferates [8,32], so do the models themselves. Often these models are quite large (containing many polygons), allowing high fidelity representation of real world objects. Unfortunately, large 3D models can be quite difficult to display interactively. This has given rise to a significant body of research [14] that attempts to simplify the models; that is, to reduce the number of polygons in the model, while preserving as much as possible the model's fidelity, or quality.

For some applications, purely geometric definitions of quality suffice. But for many other applications, the simplified model must "look like" the original - and by touching upon perceptual issues, the notion of quality

becomes much more complex. Ultimately, perceptual quality can only be determined in controlled studies with human observers. But the demands of model simplification do not often allow such involved experimentation. As a simplification algorithm runs, a quality heuristic must predict which of many alternative basic simplifications will have the least impact on visual appearance. During application development, similar quality predictors can be used to predict the perceptual quality of complete simplifications, which represent the repeated application of a heuristic.

Which predictors and heuristics are the best indicators of perceptual quality? Model simplification researchers have understandably avoided this difficult question, typically presenting an array of images depicting the simplifications produced using their algorithms. In the related fields of image compression and rendering, researchers have begun working on a two dimensional (2D) version of this problem. Image compression can be viewed as a 2D analogue of model simplification, with effort directed toward preserving quality while removing pixels, rather than polygons. The result of this research has been several quality predictors [7,23], some of which are also used as quality heuristics [3,26]. Since models have no perceptual qualities until they are turned into images, these results are of particular interest to model simplification researchers.

In this paper, we present an evaluation of various quality predictors for model simplification. As a standard for comparison of various predictions, we use mean naming times for a set of 30 models of common objects. We believe we are the first to use this standard, which is already well established in the field of cognitive psychology.

In the following section, we survey the typical sorts of predictors and heuristics used in model simplification and image compression. We then review the naming time measure, and provide a brief history of its use in cognitive psychology. We continue with a description of the experiments that obtained naming times, and a comparison of these times to several quality predictors. We conclude

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI '2000 The Hague, Amsterdam

Copyright ACM 2000 1-58113-216-6/00/04...\$5.00

with a summary and some possible avenues for future research.

QUALITY PREDICTORS AND HEURISTICS

There are three related fields of research that are concerned with visual quality: model simplification, image compression, and image rendering. Below we briefly review the efforts and concerns of each of these fields.

Researchers creating model simplification algorithms often endeavor to preserve the appearance of the models they simplify. To guide the simplifications they make, they employ quality heuristics that predict which of many basic, local simplifications will affect appearance least. To enable comparison and evaluation of simplification results, they may use quality predictors that measure the global similarity in appearance between a simplified model and the unsimplified original. These quality predictors may in turn find use as complex heuristics.

Quality heuristics for model simplification fall primarily into two classes. View-independent heuristics judge quality without reference to the eventual viewing conditions. Most of these heuristics use geometric measurements of distance [5,28] or curvature [15,31]. View-dependent heuristics use a two-stage process to gain more control over appearance. The first view-independent stage produces a nested hierarchy of simplifications. Given knowledge of the current view, the second stage selects an appropriate simplification from the hierarchy, using 2D distance heuristics [21,22] and knowledge of human perception [27,34].

Model simplification researchers have expended little effort on quality prediction, contenting themselves instead with informal examination of a series of views of the simplifications they produce. One notable exception is the view-independent Metro tool from Cignoni et al. [4], which overlays the original and simplified models in 3D space, and uses global geometric measures of the difference between the surfaces in distance and volume. These measures do not reference a particular view or viewpoint. We are not aware of any formal efforts at perceptual evaluation of either quality heuristics or predictors for model simplification.

With some recent exceptions, quality heuristics for image compression [2] have been fairly simple, and often ignore issues of spatial locality and contrast on the image in favor of various measures of distance in color space.

For quality prediction, the field of image compression has long used mean squared error (MSE):

$$\sum \sum ((P_o - P_c)^2 / P_o^2)$$

where P_o is the original pixel value and P_c the compressed pixel value, summed over the image. However, in recent years the shortcomings of MSE have become clear [13], and several predictors based on models of the early stages of the human visual system have been constructed [7,23].

Some initial perceptual evaluation of the Daly predictor was performed by Martens and Myszkowski in [24], who found a high ($R^2 = .83$) correspondence between the predictor and subjective ratings of stimuli masked with texture.

The field of image rendering has long felt the need for a quality heuristic (loosely speaking, rather than knowing which pixel to remove, rendering must know which one to add), and has adopted and extended these quality predictors and used them as quality heuristics [3,26]. Since 3D models have no perceptual qualities until they are viewed as images, image quality predictors might also be used as predictors and heuristics for model simplification.

The goal of our research is to provide a rigorous, perceptually-based evaluation of existing quality predictors and heuristics, especially as applied to the field of model simplification. Our work to date has focussed primarily on evaluation of quality predictors, but it nevertheless has some implications for the design and use of heuristics.

NAMING TIME AS A MEASURE OF QUALITY

As a standard of comparison for our evaluation, we chose naming time, a measure of object identification with a long history of use in cognitive psychology research. Since most interactions with objects begin with their identification, we believed it would be a good perceptual measure of the quality of a simplified model.

Research in cognitive psychology has already shown that the time it takes to name an object can index a number of factors that affect object identification. For example, some linguistic factors that affect naming time are the frequency of an object's name in print, the proportion of individuals who call it by a particular name (percent name agreement), the number of different names it is given, and the age at which the name was acquired (see [33] for review). Among the non-linguistic factors that affect naming times are whether or not it is displayed in its canonical [25] or upright [18,19] orientation, how much prior practice an individual has had naming the same stimulus [19,20], and the degree to which something is visually or structurally similar to other things [1,9,10,16,17].

In a particularly interesting example of the effects of visual similarity, researchers have consistently found that individuals can name manmade artifacts (e.g. furniture, vehicles, musical instruments) faster than they can name natural objects (e.g. animals, fruits, vegetables) [11,17]. Humphreys et al. [16,17] hypothesized that this was due to the similarity of structures existing in nature when compared to the diversity of structures among manmade artifacts. In their explanatory model, information about an object is presumed to be accessed and retrieved in three sequential but overlapping cognitive stages. Initially, visual input undergoes perceptual analyses whose output is used to access one or more representations in a structural description system. Activation of structural descriptions then cascades "forward" to the semantic system, where



Figure 1: One of the experimental stimuli, a standard (0%) bunny model seen in the canonical view.

semantic, categorical, and associative characteristics become available. This activation eventually cascades to the phonological name system, where a name or names are activated. An object with many perceptual neighbors (e.g. 4-legged mammals) may activate not only its own structural description, but also structural descriptions of its perceptual neighbors (at least partially). These competing activations then propagate through the semantic to the naming level, leading to naming time delays for members of structurally similar categories.

This structural similarity effect and others like it [9,10,17] led us to expect that naming times would be sensitive to model quality. Since simplification generally results in a stimulus that is less distinctive and more similar to other stimuli, naming times should increase as quality decreases.

For a number of reasons, we believe that naming times provide a more psychologically valid measure of the effects of model or image quality than the subjective ratings sometimes used in image compression research [6,24]. First, it is not known what dimensions people use to construct their ratings, whereas one cannot name a displayed object correctly without having identified it. Second, whereas cognitive variables that are relevant to object recognition affect naming times, variables that are irrelevant to object recognition may affect subjective ratings. For example, ratings may be influenced by identity-irrelevant factors like instructions, task demands, and idiosyncratic strategies; it is unlikely that naming times are affected by these factors.

OBTAINING NAMING TIMES FOR 3D MODELS

Our experiment had two goals. First, we wanted to confirm that naming time was an effective and a valid measure of



Figure 2: Another experimental stimuli, the same bunny model, this time 80% simplified.

model quality. Second, having obtained this confirmation, we wanted to use the differences in naming times for models at various levels of simplification to evaluate several quality predictors.

Method

The experiment used a 2 x 3 x 3 x 3 design. There were two types of models: manmade artifacts and animals. The models were presented at three levels of simplification. All participants saw each model at all three levels of simplification, so there were three repetitions. The between-participants independent variable was a counterbalancing factor that controlled for this repeated exposure to the same model.

Stimuli

Stimuli were created from 30 3D polygonal models in the public domain. None of the models contained color, texture, material or vertex normal information. 15 of these models represented manmade artifacts, 15 were representations of animals. Each of these models was then assigned a viewing position, simplified, and displayed, with the resulting digital images being saved to a file for later experimental display.

Model coordinate systems and poses can vary widely, and therefore some viewing parameters had to be defined interactively by the authors. However, certain viewing parameters were constant. Views were always directed towards the mean of a model's vertices, and the virtual field of view was always 40 degrees. The virtual eye point was always at a distance of twice the length of the longest dimension of the bounding box (the smallest axis aligned box that contained a model). The authors then interactively rotated each model about its mean vertex so that it was

upright and facing left. Models were next rotated 21 degrees positively in the XZ plane (where X and Y span the image plane, and Z is scene depth), and 18 degrees positively in the YZ plane. The resulting viewpoint was to the right of and slightly above the model. This “canonical” view revealed a reasonable level of detail across the models (see also [25]). Figure 1 shows an example view.

Models were simplified in two stages. First, we ensured that all models had the same number of polygons by using the Qslim algorithm [12] to simplify each model to 3700 polygons (± 50). (3700 was the number of polygons in the smallest model in the original set). We will refer to the models resulting from this stage of simplification as the “standards” (0%). Second, the models were further simplified using a vertex clustering [28] algorithm, with each standard model reduced by two levels: 50% and 80% of the original polygons were removed, respectively. Figure 2 shows a standard model with 80% of its polygons removed.

Images were created from the standard and simplified models, creating three exemplars of each object, at 0%, 50%, and 80% simplification, for a total of 90 stimuli. The models were displayed using OpenGL on a Silicon Graphics Crimson RE workstation running the IRIX operating system. Models were illuminated with a single white (RGB = (1,1,1)) light located at the eye point. All models were assigned the same white color as the light source, and flat shaded. The resulting images were saved to a file and regularized for size, with each image scaled down to 591 pixels in the longest dimension, while maintaining aspect ratio.

During experimentation, the images were displayed in grayscale on a black background. All images were centered on the screen.

Design

As described above, we used 30 models (of two types: 15 animals and 15 artifacts) and 90 stimuli (3 levels of simplification for each model). All stimuli were displayed once per session, for a total of 90 trials. Sessions were organized into three blocks of 30 trials, with each model presented once during each block. Blocks were organized into three groups, with each group containing stimuli at the same level of simplification (that is, each block contained 10 stimuli at 0% simplification, 10 at 50%, and 10 at 80%). Models were assigned to groups randomly, and thus model type was not balanced within these groups.

The 30 participants were divided into three participant-groups of 10 individuals. Within each block, the order of presentation of the stimulus-groups was balanced across the three participant-groups. Thus in block 1, depending on which participant-group he was in, a given participant saw a model (e.g. a dolphin) at either 0, 50, or 80 percent simplification, then at the alternate levels in the subsequent two blocks. So, each participant saw all 30 models three times, once at each of the three levels of simplification.

Table 1: Results of experimental statistical analysis.

Variable	Across	ANOVA
simplification	participants	F(2,58) = 39.81, p<.001
simplification	items	F(2,54) = 12.84, p<.001
model type	participants	F(1,29) = 21.95, p<.001
model type	items	F(1,27) = 2.86, p<.10
repetition	participants	F(2,58) = 67.72, p<.001
repetition	items	F(2,54) = 41.10, p<.001
simp x rep	participants	F(4,116) = 10.42, p<.001
simp x rep	items	F(4,108) = 5.90, p<.001

There were six practice trials. Stimuli for the practice trials were created in a similar manner to the experimental stimuli, however, their original and degraded face counts were not standardized: all the practice models fell below the 3700 original face threshold. The practice stimuli were degraded to the same relative percentages as the experimental stimuli. Specifically, two had 0%, two had 50% and two had 80% of their original faces removed. The practice stimuli were presented in a random order for each participant.

Apparatus and Procedure

The participants were seated approximately 0.7 m from the display, with the stimuli subtending a visual angle of 15 degrees. They performed the task by viewing the pictured stimuli on a computer screen and speaking the names of the modeled objects into a hand-held microphone. The pictures were displayed on a 17-inch Microscan CRT driven by a 166MHz Pentium PC. The experiment was executed under the control of the Microcomputer Experimental Laboratory (MEL) software [29]. Accuracy of the responses was recorded offline by the experimenter.

Participants were told that on each trial of the experiment, they would see a picture of an object, and their task was to name that object as quickly and accurately as they could. They were also told that some pictures would be simplified representations. The experimenter controlled the pace of the trials by pressing the space bar at the start of each trial. After the practice trials, participants were asked if they had any questions regarding the task. The participants then performed the 90 experimental trials.

A trial consisted of the following events: the experimenter pressed a the space-bar, a fixation cross appeared for 750ms, the picture appeared on the screen, the participant named the picture, the picture disappeared as soon as a name was said, the experimenter scored the response and pressed the space-bar again to begin the next trial. Naming times were recorded from stimulus onset to the participant's response.

Participants

Thirty paid undergraduate volunteers from University of Alberta pool participated in the experiment.

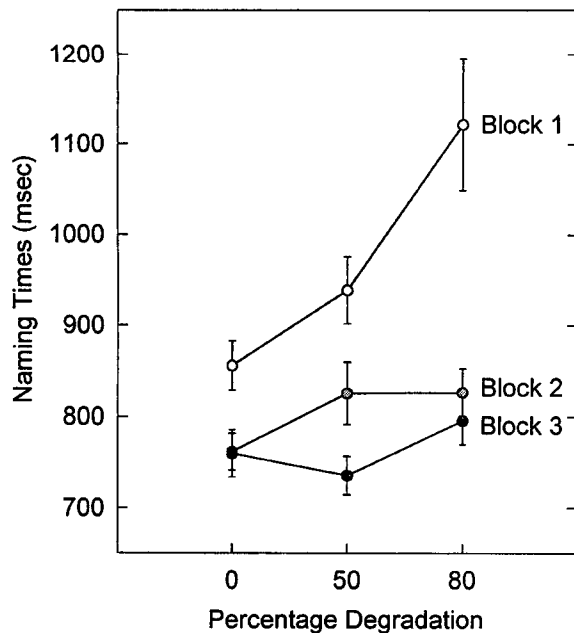


Figure 3: Naming times as a function of model simplification (degradation) and viewing repetition (block).

Results

Two kinds of trials were excluded from analysis. First, we excluded naming times measured during spoiled trials (e.g., trials in which participants failed to trigger the microphone with their first vocalization). Spoiled trials made up 5.1% of all trials. We also excluded naming times from trials in which a participant's response was an error (e.g., calling a picture of a table a "statue"). Only 0.2% of all trials were errors. Near misses inside a semantic category were permissible as correct names (e.g., calling a "hippo" a "rhino"). This is common practice in psychological research [19,20]. However, preliminary analyses showed that the image of the computer monitor was called a "monitor" or "computer monitor" by 47% of the participants, and a "television" or "TV" by 57% of the participants, giving it a percent name agreement score of 57%. Because percent name agreement has been shown to be a strong predictor of naming times [33], and because the computer monitor was the only model with a low percent agreement, we excluded it from further analyses. Only 2.8% of the remaining trials were categorical near misses.

We performed two types of analyses. For the analysis with items as the random factor, we averaged naming times across participants for each combination of simplification and repetition (block); each item thus had nine scores contributing to a mixed design analysis of variance (ANOVA). The within-item factors were repetition and simplification, and the between-item factor was object type. For the analyses with participants as the random factor, we averaged naming times across items for each combination of repetition and simplification for the first ANOVA, and

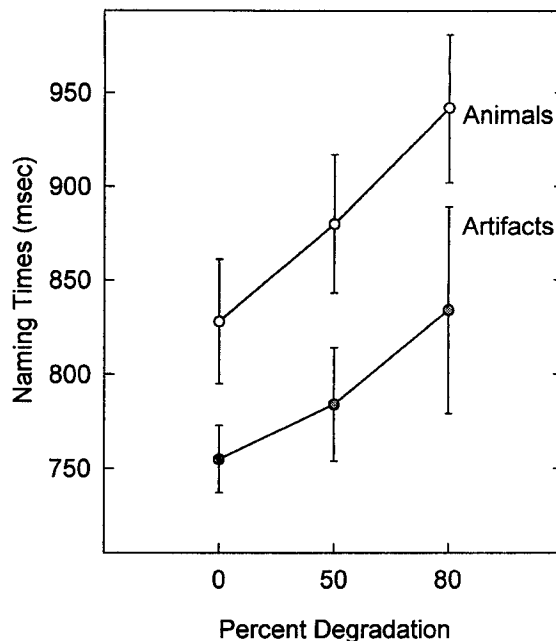


Figure 4: Naming times as a function of model simplification (degradation) and model type.

for each combination of object type and simplification for the second.

The results of these analyses are shown in Table 1. The effects of simplification, model type, and repetition were reliable, as was the interaction between percentage simplification and block. No other interactions approached reliability.

Discussion

Figure 3 shows the effects of simplification and repetition on naming time, averaged across participants. Clearly, naming times are sensitive to simplification and model quality. Naming times decrease with repeated viewings of the model. However, simplification has little effect on naming times if the model has previously been seen (explaining the interaction). This finding replicates earlier research in which a manipulation that increases the time it takes to access the structural description of an object (such as rotating the object in the picture plane; [9,10,18,19]) influences naming time on the first viewing, but has a lesser effect on later viewings.

Figure 4 shows the effects of model type and simplification on naming times, averaged across participants. Animals took longer to name than artifacts, replicating results of the psychological research we reviewed earlier. Interestingly, there was no interaction between model type and simplification. This is evidence that the animal-artifact effect is unlikely to be occurring at the perceptual level, but rather, occurs later in the cognitive system. That is, if simplification and model type both affected early perceptual processes, they would be expected to interact within the same "module" (cf. [30]). It is unlikely that

other measures of quality, such as subjective ratings, would resolve perceptual/semantic effects in such detail.

These results are compelling evidence that naming time is an effective and psychologically valid measure of model quality. Naming times are strongly affected by model simplification. We also replicated some well-known results from previous research using polygonal models as stimuli. The slight naming time differences between 50% simplified models and standards may be an indication that the task of simplification increases in difficulty as the number of polygons decreases. This may also indicate that naming times are not an ideal standard of comparison for limited application of a quality heuristic, especially when the number of polygons is large.

We should sound a note of caution about these inferences. First, we have obtained naming times only for one almost optimal view. We expect that other, less optimal, viewpoints will increase the effect of simplification on naming times, and thus, the generality of our results. Second, our naming times were obtained with a set of 30 models, which is a moderately small number compared to other naming time studies (although it is within an appropriate range). Finally, models are often used in complex interactive applications, which contain motion, additional color and texture information, and make simultaneous use of several models, not just one in isolation. All of these additional factors are not reflected in our results.

Subject to these concerns, our naming time experiment produced some interesting additional implications for designers of interactive applications. First, designers may wish to take advantage of the simplification-repetition interaction by allowing model quality to degrade when an object has been visible for some time. Also, the animal-artifact effect may indicate that user performance can be improved by simplifying models of natural objects less than models of manmade objects (while keeping in mind that simplification on both types of models has similar effects). Both of these suggestions are early indications that will have to be tested in practice.

EVALUATING MODEL QUALITY PREDICTORS

Having determined that naming times are an effective measure of model quality, we turn to an evaluation of predictors of model quality. We compare two view dependent image quality predictors from image compression and rendering research to a single 3D, view independent predictor.

The image quality predictors we evaluate are MSE [13] and an implementation of the predictor from Bolin and Meyer [3] (BMP), which models the early, perceptual stage of the human visual system. Both compare a standard image to an image that approximates it. We say these two predictors are “view dependent” because the input images must be created from a certain viewpoint. While MSE returns a single difference (quality) measure, BMP returns a

Table 2: Correlations of naming time differences to predictors.

Predictor	50% Simp		80% Simp	
	Corr	Sig?	Corr	Sig?
BMP	0.040	no	0.379	yes
MSE	0.145	no	0.315	marg
Metro max	0.207	no	0.317	marg
Metro mse	0.245	no	0.196	no
Metro mean	0.129	no	0.182	no

difference image. The value at each location of the image predicts the ability of human observers to perceive the difference between the two input images. For our evaluations we need a single quality rating. We use the mean of the values in BMP’s difference image.

The 3D predictor we evaluate is the Metro tool from Cignoni et al [4]. Metro compares a 3D model standard to an approximating model. We say this predictor is “view independent” because by using models rather than images, Metro does not reference a certain viewpoint. Metro samples the surface of each model at multiple points, and determines the shortest distance from the surface of one model to the next. Distances are signed depending on which of the two surfaces is outermost. It then returns the mean, mean squared and the maximum of these distances, normalized by the length of the diagonal of the standard model’s bounding box.

We used MSE, BMP and Metro to predict the quality difference between the standard (0% simplified) stimuli and each of the corresponding 50% and 80% simplified stimuli, for a total of 60 quality predictions. We compared each of these predictions to differences in mean naming times for corresponding models at the first viewing. For example, we compared the predictions for the quality difference between the standard dolphin and the 80% simplified dolphin to the difference between the mean naming times at first viewing for these two models. We used only the first viewing differences because these were the most affected by simplification.

Table 2 shows the correlations between the predictors and the naming time differences. None of the predictors accounted for much of the difference in naming times between the standard and 50% simplified models. However, several predictors did a reasonable job with the difference between the standard and 80% times, including BMP, MSE and the Metro maximum distance. A step-wise regression with all the predictors as independent variables confirms what is seen in the table; BMP was the best predictor of the difference in naming times, $F(1, 27) = 4.54$, $p < .05$, and none of the other predictors added significantly to the amount of variance accounted for. Each of the MSE and Metro maximum distance predictors accounted for a marginally significant amount of the variance in names times ($p < .10$) if used as the first term in the regression.

Which predictor was most effective? Predictor effectiveness varied with the size of the quality difference being predicted. Below, we present our two recommendations, based on the best available evidence. We follow these recommendations with some discussion of other possible interpretations of our results.

BMP is the best predictor of larger quality differences. BMP predicted the quality of the 80% simplifications reliably, but failed miserably at predicting the quality of the 50% simplifications. This is quite a contrast and it merits further research. It may indicate instead that the naming time measure cannot resolve smaller quality differences. But we think this is unlikely, because the naming time measure is quite sensitive to small featural differences between instances of the same “basic-level” object (e.g., two different cars; Friedman & Rabiau, 1999). However, other predictors did show better correlations to naming times. Perceptually speaking, we feel comfortable calling the standard-50% difference “small”. However, from a systems viewpoint, 50% of all polygons is not a small number, a reliable predictor for such quality differences would be very useful.

For smaller quality differences, Metro’s view independent predictors were most reliable. Although no predictors reliably accounted for standard-50% variance, it is intriguing to note that Metro’s view independent geometric measures had the highest correlations for these smaller quality differences. This may be an indication that geometric measures are good quality heuristics, but this requires confirmation in further experimentation.

Interestingly, for the larger standard-80% quality differences, Metro’s maximum distance predictor was fairly effective, while its mean predictors were not. It may be that maximums are more sensitive than means to the changes in large model features that occur during drastic simplifications. However, this possibility also needs experimental confirmation.

CONCLUSIONS AND FUTURE WORK

Model simplification researchers need quality heuristics to guide simplification, and quality predictors to aid in selection of appropriate heuristics and simplification algorithms. We have presented an evaluation of several quality predictors for model simplification using a new perceptual experimental standard, naming times. Naming times have long been used in cognitive psychology research as a measure of recognition. We found that naming times are sensitive to simplification, and we were able to duplicate several well known results from cognitive psychology using polygonal models as stimuli.

We then compared naming times to several model quality predictors. Predictors differed in effectiveness depending on the size of the quality difference they were being asked to predict. Image based predictors such as that from Bolin and Meyer [3] were quite effective for large differences and much less effective for small differences. View

independent geometric predictors like Metro were most effective for smaller quality differences, but were even then unreliable.

We plan to increase the reliability of these results by increasing the number of models to be named, and by using additional, non-optimal viewpoints. We will also evaluate the cumulative effect of different quality heuristics embedded in the same simplification algorithm. Ultimately, we also hope to consider colored and textured models in our evaluations.

ACKNOWLEDGEMENTS

Oscar Meruvia wrote indispensable 3D viewing software. Roman Kotovych provided help with simplifications and quality predictors. We thank Greg Turk for his models and geometry filters. Peter Lindstrom participated with a thought provoking correspondence, and by assisting in finding relevant code. Oleg Veryuvka and Lisa Streit provided useful implementations of simple image quality metrics. Peter Johnstone also provided assistance. Stanford University was the source of the bunny model. This research was supported by NSERC grants to the first two authors, and by NSERC Undergraduate Fellowships to the third and Mr. Kotovych.

REFERENCES

1. Bartram, D.J. (1976). Levels of coding in picture-picture comparison tasks. *Memory and Cognition*, 4, 593-602.
2. Baxes, G. (1994). *Digital Image Processing*. John Wiley & Sons, New York.
3. Bolin, M. & Meyer, G. (1998). A perceptually based adaptive sampling algorithm. Proc. of SIGGRAPH 98. In *Computer Graphics Proceedings, Annual Conference Series*, 1998, ACM SIGGRAPH, 299-309.
4. Cignoni, P., Rocchini, C. & Scopigno, R. (1998). Metro: measuring error on simplified surfaces. *Computer Graphics Forum*, 17, 2, 167-174. Available at: <http://vcg.iei.pi.cnr.it/metro.html>.
5. Cohen, J., Varshney, A., Manocha, D., Turk, G., Weber, H., Agarwal, P., Brooks, F. & Wright, W. (1996). Simplification envelopes. Proc. of SIGGRAPH 96. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 119-128.
6. Cosman, P., Gray, R. & Olshen, R. (1993). Evaluating quality of compressed medical images: SNR, subjective rating and diagnostic accuracy. *Proc. IEEE*, 82, 6, 919-932.
7. Daly, S. (1993). The visible differences predictor: an algorithm for the assessment of image fidelity. In Watson, A. B. (ed.). *Digital Images and Human Vision*. MIT Press, Cambridge, MA, 179-206.
8. DeRose, T., Kass, M. & Truong, T. (1998). Subdivision surfaces in character animation. Proc. of SIGGRAPH 98. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 85-94.

9. Friedman, A., & Rabiau, M. (1999). Lions and tigers and bears: The role of structural similarity and visual detail in naming disoriented objects. Manuscript submitted for publication.
10. Friedman, A., & Vuong, Q. (1999). Cats, cows, cameras, and cars: The role of structural similarity in naming and categorizing upright and disoriented pictures. Manuscript submitted for publication.
11. Gaffen, D., & Heywood, C.A. (1993). A spurious category-specific visual agnosia for living things in normal human and nonhuman primates. *Journal of Cognitive Neuroscience*, 5, 118-128.
12. Garland, M. & Heckbert, P. (1997). Surface simplification using quadric error metrics. Proc. SIGGRAPH 97. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 209-216. Available at: <http://www.cs.cmu.edu/~garland/quadrics/qslim.html>
13. Girod, B. (1993). What's wrong with mean-squared error? In Watson, A. B. (ed.). *Digital Images and Human Vision*. MIT Press, Cambridge, MA. 207-220.
14. Heckbert, P. & Garland, M. (1997) Survey of polygonal surface simplification algorithms. Technical report, CS Dept., Carnegie Mellon, U. Available at: <http://www.cs.cmu.edu/~garland/papers/simp.pdf>.
15. Hinker, P. & Hansen, C. (1993). Geometric optimization. *Proc. IEEE Visualization '93*, 189-195.
16. Humphreys, G. W., Lamote, C., & Lloyd-Jones, T. J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory*, 3, 535-586.
17. Humphreys, G. W., Riddoch, M. J., & Quinlin, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, 5, 67-103.
18. Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition*, 13, 289-303.
19. Jolicoeur, P. (1988). Mental rotation and the identification of disoriented objects. *Canadian Journal of Psychology*, 42, 461-478.
20. Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: Effects of context of prior presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 200-210.
21. Leubke, D. & Erikson, C. (1997). View dependent simplification of arbitrary polygonal environments. Proc. of SIGGRAPH 97. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 199-208.
22. Lindstrom, P., Koller, D., Ribarsky, W., Hodges, L., Faust, N., & Turner, G. (1996). Proc. of SIGGRAPH 96. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 109-118.
23. Lubin, J. (1993). A visual discrimination model for imaging system design and evaluation. In Peli, E. (ed.). *Vision Models for Target Detection and Recognition*, World Scientific, New Jersey, 245-283.
24. Martens, W. & Myszkowski, K. (1993). Psycho-physical validation of the visible differences predictor for global illumination applications. *IEEE Visualization '93, Late Breaking Topics*, 49-52. Also available at: <http://wwwsv1.u-aizu.ac.jp/labs/csel/vdp/>.
25. Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddelay (Eds.), *Attention & Performance IX*, Hillsdale, NJ : Erlbaum, 135-151
26. Ramasubramanian, M., Pattanaik, S. & Greenberg, D. (1999). A perceptually based physical error metric for realistic image synthesis. Proc. of SIGGRAPH 99. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 73-82.
27. Reddy, M. (1998). Specification and evaluation of level of detail selection criteria. *Virtual Reality: Research, Development and Application*, 3, 2, 132-143.
28. Rossignac, J. & Borrel, P. (1993). Multi resolution 3D approximations for rendering complex scenes. In Falcidieno, B. & Kunii, T. (eds.), *Geometric Modeling in Computer Graphics*. Springer Verlag, 455-465.
29. Schneider, W. (1988). Micro Experimental Laboratory: an integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation, and Computers*, 20, 206-217.
30. Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, 30, 276-315.
31. Turk, G. (1992). Re-tiling polygonal surfaces. Proc. of SIGGRAPH 92. In *Computer Graphics*, 26, 2 (July), ACM SIGGRAPH, 55-64.
32. Turk, G. & Levoy, M. (1994). Zippered polygon meshes from range images. Proc. of SIGGRAPH 94. In *Computer Graphics Proceedings, Annual Conference Series*, ACM SIGGRAPH, 311-318.
33. Vitkovitch, M., & Tyrell, L. (1995). Sources of name disagreement in object naming. *Quarterly Journal of Experimental Psychology*, 48A, 822-848.
34. Watson, B., Walker, N., Hodges, L. & Worden, A. (1997). Managing level of detail through peripheral degradation: effects on search performance with a head-mounted display. *ACM Trans. Computer-Human Interaction*, 4, 4, 323-346.