Quality Metric for Approximating Subjective Evaluation of 3-D Objects

Yixin Pan, Irene Cheng, *Student Member*, and Anup Basu, *Senior Member*

***Abstract*—Many factors, such as the number of vertices and the resolution of texture, can affect the display quality of three-dimensional (3-D) objects. When the resources of a graphics system are not sufficient to render the ideal image, degradation is inevitable. It is, therefore, important to study how individual factors will affect the overall quality, and how the degradation can be controlled given limited resources. In this paper, the essential factors determining the display quality are reviewed. We then integrate two important ones, resolution of texture and resolution of wireframe, and use them in our model as a perceptual metric. We assess this metric using statistical data collected from a 3-D quality evaluation experiment. The statistical model and the methodology to assess the display quality metric are discussed. A preliminary study of the reliability of the estimates is also described. The contribution of this paper lies in: 1) determining the relative importance of wireframe versus texture resolution in perceptual quality evaluation and 2) proposing an experimental strategy for verifying and fitting a quantitative model that estimates 3-D perceptual quality. The proposed quantitative method is found to fit closely to subjective ratings by human observers based on preliminary experimental results.**

***Index Terms*—3-D graphics, image quality, perceptual metric, subjective evaluation.**

I. Introduction

THREE-dimensional (3-D) computer graphics were traditionally used in high-end graphics workstations for specific applications such as computer-aided design (CAD) and feature movie production. In recent years, 3-D graphics has taken an important role in interactive, networked applications such as computer games, e-commerce, and educational software. Although the rapid development in graphics hardware has made realistic 3-D display possible on personal computers (PCs), the increase in data complexity for high resolution models still surpasses the average PC and normal network capabilities for online applications. A decade ago, most 3-D models were carefully designed and composed of relatively small number of polygons, in order to speed up processing and rendering. Today, highly complex models are required in many applications. In computer vision, range data on an object is acquired via 3-D scanning systems. In CAD, polygonal models are produced by the subdivision of curved parametric surfaces. In medicine, organs and tissues are reconstructed from radiological and nuclear images. In remote sensing, terrain data is obtained from satellite photographs. These applications often demand 3-D models containing millions of polygons. Since the processing and display of high-resolution 3-D objects require substantial computer resources, trade-off has to be made between display quality and efficient interactivity.

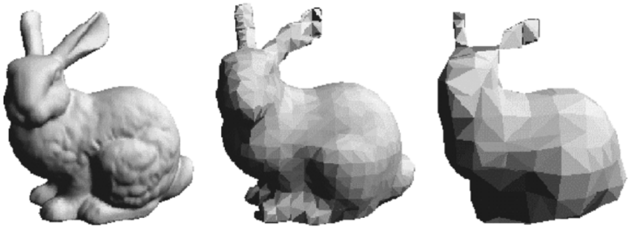


Fig. 1. Stanford Bunny at various resolution levels.

The constraints that can determine the display quality of a 3-D image in online applications fall into two main categories: computational constraint and network bandwidth constraint. Computational constraint includes the resources for displaying 3-D objects, which are determined by the number of polygons, shading, lighting, texture resolution etc. network constraint includes the available bandwidth of a network, such as the Internet, which can affect the transmission speed significantly depending on the current traffic. It is, thus, unwise in an interactive application to transmit a high-resolution 3-D object over a congested network. An adaptive approach can be applied through compression and simplification of 3-D data to make the transmitted size 10–20 times smaller than the original without noticeable distortions [5]. An example of geometric simplification is shown in Fig. 1, in which the Stanford Bunny is simplified to various resolution levels (number of triangles is 69 451 left, 1919 middle, and 462 right).

When a graphics system is under computational and/or network constraints discussed above, the appropriate size can be determined depending on the available resources in order to speed up interactivity. Fortunately, a high-resolution representation is not always required. A simplified version of the 3-D object can reduce temporary storage, memory utilization, as well as processing and rendering time. An essential consideration in designing effective interactive 3-D systems is to adaptively adjust the model representation, while preserving satisfactory quality as perceived by a viewer. Whether or not the perceived quality is satisfactory is a subjective decision and can only be determined by the viewers. While most research in the literature focus on geometric compression and use only synthetic

Manuscript received February 18, 2003; revised July 8, 2003. This work was supported by the Alberta Science and Research Authority (ASRA) and by the National Science and Engineering Research Council (NSERC). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ryoichi Komiya.

The authors are with the Department of Computing Science, University of Alberta, Edmonton, AB T6G 2E8, Canada (e-mail: yixin@cs.ualberta.ca; lin@cs.ualberta.ca; anup@cs.ualberta.ca).

Digital Object Identifier 10.1109/TMM.2005.843364

texture or color, we address both *geometry resolution* and *realistic texture resolution*, and analyze how these factors affect the overall perceptual quality and fidelity. Our analysis is based on experiments conducted on human observers. The perceptual quality metric derived from the experiments allows the appropriate level of detail (LOD) to be selected given the computation and bandwidth constraints. A detailed survey on simplification algorithms can be found in [9]. These algorithms try to control the complexity of a wireframe by developing various strategies for simplifying the LOD in different parts of a 3-D object. In order to easily control the details on a 3-D object we will follow a simple model approximation strategy based on multiresolution representation of texture and wireframe. More complex LOD models as perceived by human observers will be included in our future work. Our main contribution is in proposing and evaluating a quantitative metric that measures perceptual quality variations in a restricted online environment. For example, given limited bandwidth [24] our model can give multimedia developers some insight into how to reduce the texture and wireframe details before transmission, and what is the relative importance of these two factors.

The remainder of this paper is organized as follows. Section II reviews past work on perceptual quality evaluation. Section III examines the factors that control the fidelity of 3-D images. Section IV presents the user interface and environment used for conducting human evaluation experiments and proposes a quantitative metric for estimating subjective evaluations. In Section V, the quantitative metric is estimated through experimental results on 3-D objects; the reliability of the subjective evaluations is also discussed and measured in this section. Finally, the Conclusion and future work are summarized in Section VI.

II. Review of Perceptual Quality Evaluation

In the area of image compression, the mean square error (mse) is commonly used as a quality predictor. The mse is defined as

$$\text{mse} = \sum_{\text{i}=0}^{M-1}\sum_{j=0}^{N-1}\frac{(P_0(i,j)-P_c(i,j))^2}{MN} \qquad (1)$$

where P_0 is the original pixel value, P_c is the compressed pixel value, M and N are the width and height of the image in pixels, respectively. However, past research has shown that mse does not correlate well to perceived quality based on human evaluation [12]. Since this study, a number of new quality metrics based on the human visual system have been developed [11], [18], [4], [19], [10]. Limb [11] originally looked at fitting an objective measure that closely estimated impairment ratings on five test pictures. Effects of distortion based on masking and filtering were considered in these works. The new models incorporate distortion criteria, psychophysical rating, spatio-temporal frequency response, color perception, contrast sensitivity and contrast masking results, achieved from different psychovisual experiments. One of the models was extended to also evaluate the distortion of digital color video [23]. An error metric, to approximate perceptual video quality, was proposed by Webster *et al.* [22]. Our work focuses on the 3-D-display quality evaluation of geometric as well as texture data, and is different from prior work on image and video compression assessment.

In the study on image synthesis, different shading and global illumination algorithms have been introduced to simulate photo-realistic effect. Since mathematical models cannot solely determine the accuracy of the display quality, human perceptual evaluation has to be taken into account. A number of perception-driven rendering algorithms were developed to incorporate the human visual system (HVS) as a factor to compute global illumination so as to improve perceptual accuracy [6], [2].

In research on 3-D-model simplification, a predefined error bound is often used to measure the deviation between the original and simplified models. While such a measure can control the deviation, it does not estimate the perceptual quality of the simplified models. Most researchers leave it up to the readers to evaluate quality by showing a number of images at various simplified stages. Only recently, a few authors started to develop perceptual experiments to examine how well error metrics can reflect perceptual quality. However, their results are not encouraging.

The criteria used to evaluate perceptual quality in the various areas mentioned above are rather different. In image compression, the process is simpler in the sense that it only deals with two-dimensional (2-D) images, and the images can be preprocessed. In contrast to image compression, image synthesis computes visibility, shading and global illumination at run-time. While preserving perceptual quality, the underlying mechanism has to reduce computational complexity, i.e., to avoid displaying cull-faces and intersecting invisible objects. Unlike image compression, in which an overall fidelity metric is desired, the image synthesis process may use ray tracing, radiosity or other sampling techniques to determine visual distortion. The sample set may need to be adjusted adaptively until such distortion is negligible. Because of the processing time involved, image synthesis is often an offline process. Although the image is rendered as a 3-D object, the perspective is actually represented by translating the coordinates on a 2-D plane. Thus, the quality evaluation is based on a view-dependent 2-D image.

III. Factors Controlling 3-D Image Degradation and Perceptual Quality Estimation

There are many strategies available to provide smooth degradation in the fidelity of 3-D images. In an early pioneering paper [13], [14], Nagata discussed the evaluation of subjective depth of 3-D objects as a function of the variation of texture conditions between sharp and blurred images, and texture patterns of form, density, shade, or polish. The evaluations were based on depth sensitivities of various cues for depth perception as a function of distance to the viewer. Three subjects two male and one female with normal stereoscopic vision were used as subjects in the experiments. Depth thresholds and viewing conditions were varied depending on a number of factors. The author extended the work [17] with Siegel to include studies of stereoscopy with very small inter-occular disparities, called "micro-stereopsis." Implications of the research in developing "zoneless" auto-stereoscopic displays were also discussed. Our research differs from the above study in the following aspects.

1) We do not evaluate the depth perception on 3-D objects *per se*.
2) We perform an overall quality evaluation of a 3-D object based on wireframe resolution and texture resolution, wireframe resolution was not considered in the previous study.
3) We attempt to estimate the perceptual quality depending on wireframe and texture quality using a quantitative model, in our opinion this model fitting is a new contribution.

An overview of some of the related factors that influence perceptual quality include the following.

A. Geometric Representation

The shape of a 3-D object is commonly represented by a polygonal mesh, in which the vertices of the polygons provide depth information on the object surface. While smooth surfaces are represented by a smaller number of vertices, surfaces with more detail are represented by a higher number of vertices. Given an adequate number of polygons, we are able to describe complex geometry with high precision. Nevertheless, complex models demand large storage, and long computing and rendering time, which may not be suitable for interactive visualization. The performance of graphics systems can thus be measured in terms of the number of polygons (or vertices) updated per second. In other words, the number of vertices in a model is a good evaluation metric.

B. Texture Resolution

Construction of high-quality 3-D objects requires high-resolution texture images for mapping. The question is how much the texture resolution can be reduced in order to achieve satisfactory interactivity without affecting perceptual quality. In experiments on substitution of geometry with texture, Rushmeier *et al.* observed that covering colored geometric models with very low-resolution texture decreases the perceived quality [16]. This is because the spatial frequency of an overly simplified texture is significantly lower than the minimum level required by the human visual system. Based on this observation, we should not use very low-resolution texture. Moderate resolution texture maps [8] generally provide acceptable quality. One simple way to obtain a lower resolution texture is by averaging. In addition to decreasing texture resolution, increasing image compression ratio can also reduce the size of transmitted data. Texture compression is thus another metric related to quality evaluation.

C. Shading

Shading complexity is determined by the shading model, the number and positions of lights illuminating the scene, the textures of objects, the use of shadows and so on. Rogowitz *et al.* concluded that visual quality varies significantly depending on the directions of illumination [15], and thus, comparison of different shading models should be based on the same illumination environment. Consequently, constant illumination is used in our experiments.

D. Frame Rate

Many systems automatically decrease the frame rate to allow enough computation time to render a scene. Since the refresh rate is inversely proportional to computation time of a single frame, reducing frame rate is simple and allows smooth control of the quality of an individual frame, but will lead to the problem of flickering. In interactive applications, even a slight mismatch of refreshing rate is not acceptable. For instance, in a shooting game, users may not be able to aim at an object at a particular position or moment. It is, therefore, important to use the correct frame rate, and update the scene adaptively to optimize performance [20].

E. Distance

The impact of visual stimuli depends on the distance between the stimuli and the viewer. The contrast sensitivity function (CSF) [12] describes how sensitive the human visual system is to various frequencies of visual stimuli. Even though distance is an important factor in perceived quality, we eliminate this factor in the current version of our proposed evaluation metric by keeping a fixed distance. However, as mentioned by Siegel and Nagata [17], we scale objects to the largest possible size on a large monitor to allow observers to have better depth perception.

F. Visual Masking and Adaptation

Visual texture may hide faceting because of the tessellation of a curved surface. Ferwerda *et al.* developed a comprehensive model of visual masking that can predict whether the presence of one visual pattern affects the perceptibility of another visual pattern when one is imposed over another [11], [6]. Visual adaptation is the phenomenon that the sensitivity of human eye changes for varying luminance, for example, we sometimes need several minutes to see well when entering a dark theatre.

G. Other Factors

Some other important factors are discussed in psychology, such as the degree of concentration from the viewers. The fovea, the region of highest sensitivity on the retina occupies roughly a central angle of one degree of vision. Visual acuity, measured as the highest perceptible spatial frequency, is significantly lower in the visual periphery than in the fovea [1]. Thus, if within an image different objects have different resolutions, the perceived quality largely depends on the resolution of the object in focus. Since these factors vary from person to person, and from situation to situation, we will eliminate their influence in our experimental environment.

1) Perceptual Quality Estimation: Automatic measures based on mathematical theory were used to evaluate perceptual quality, but few studies have been performed on psycho-visual experiments to evaluate 3-D object quality. There are two basic reasons: First, interactive visualization has a short history of less than ten years since the time high-performance graphics accelerators became available. Second, perceptual experiments are time-consuming and expensive. Naming time, the response time to recognize a given object, has been used in cognitive psychology research as a measure of recognition for a long time. However, naming time becomes ineffective with prior knowledge on objects. Thus, Watson *et al.* included ratings and forced choices as perceptual quality measures in their later work [21]. They suggested that ratings and forced choices are better measures than naming time for a small number of participants. In addition, BMP [2], mse, and MetroMN [3] are excellent predictors of fidelity as measured by ratings. MetroMN is a notion and measure from the Metro tool [3]. Watson *et al.* used the mean of the values in BMP difference images as a measure in their experiments [21]. Nevertheless, Watson *et al.* only used still images in their experiments, thus their results only apply to static scenes with a fixed viewpoint. Can their results be extended to dynamic scenes with variable viewpoints? Rogowitz *et al.* provided a negative answer in their experiments [15].

Although more experiments are needed to draw robust conclusions, our initial observation is that neither the geometry-based metric Metro nor the image-based metric mse is a good predictor for view-independent perceptual quality of 3-D objects. Nevertheless, we have learnt from previous experiments that rating is a better measure than naming time for a smaller number of participants, and illumination and animation of objects are important factors in perceptual quality. We will consider these factors in our experiments.

2) Incorporating Various Factors in Designing a Perceptual Metric: Since variable frame rate is expensive to implement and may be uncomfortable for viewers to experience, we do not adjust frame rates in our experiments. We choose geometric and texture resolution as our focus because they are the fundamental components of a 3-D object, and their simplification can have significant impact on both computation-constrained and bandwidth-constrained graphics systems. Other factors such as shading, illumination, and distance are also important in determining perceptual quality, and will be incorporated in our experiments in future work.

IV. Experimental Environment and Derivation of a Perceptual Metric

Fig. 2 shows the orientation of a virtual world (right) with respect to the viewer (left), with the view platform in the middle.

In our interface, we start by rotating an object at a slow speed, and the users can use a scrollbar to adjust the rotation speed or stop rotation. Rotation speed is an important factor relating depth perception to motion parallax, as described in ([17, Fig. 3]). A green grid is drawn in the background to represent the floor. Two text fields and one pull down menu are available to control the object geometry resolution in the X and Y directions, and the resolution of texture image respectively. Fig. 3 shows the user interface and the virtual world rendering a 3-D object *Nutcracker*. Fig. 4 shows snapshots of the object from different viewpoints.

A. 3-D Data Acquisition Hardware and Processing

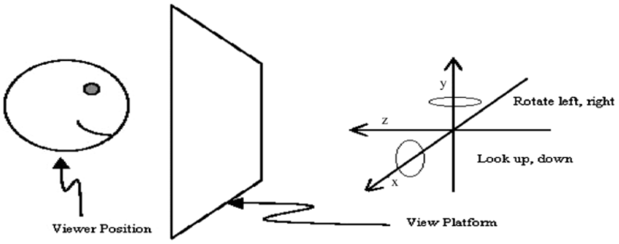
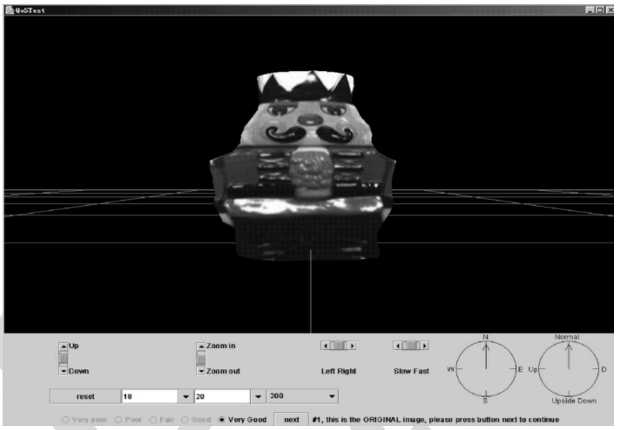


Fig. 2. View platform in virtual world.



18 20 200

○ Very poor ○ Poor ○ Fair ○ Good ● Very Good

#1, this is the ORIGINAL image, please press button next to continue

Fig. 3. User interface and the 3-D virtual world.

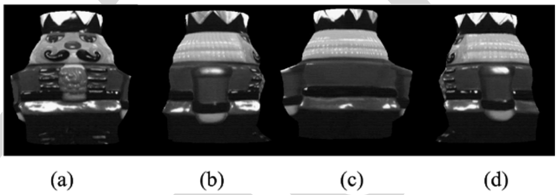


Fig. 4. Views of Nutcracker from various angles. (a) Front. (b) Left. (c) Rear. (d) Right.

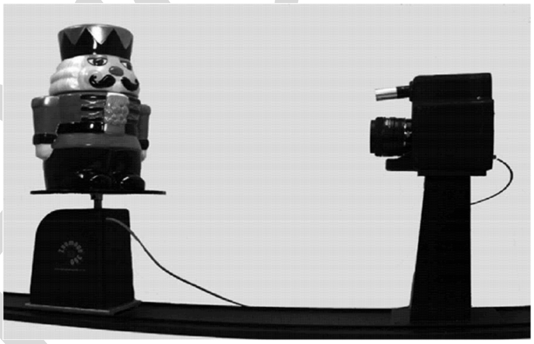


Fig. 5. Zoomage 3-D scanner.

Five 3-D objects (*Doll, Nutcracker, Pot, Head, and Dog*) were used as stimuli in the experiments. These objects were acquired with the *Zoomage* 3-D scanner. The scanned objects were rotated and captured by the range scanner. The wireframe resolution can be up to 10 000 × 30 polygons for a full 360 degrees, 31 laser-lines scanned model, and texture resolution is up to 5300 pixels (vertical) × 40 000 pixels (horizontal). Since all objects were lit from the front during scanning, and the viewpoint is fixed at the front of objects, the rendered scene simulates illumination from the front. Fig. 5 illustrates the scanning process, and Fig. 6 shows the texture, wireframe, and the canonical view of object Nutcracker. The other objects (dog, doll, head and pot) used in the experiments are shown in Fig. 7.

B. Simplification of Scanned Models

The original scanned objects were intentionally over-sampled with respect to both geometry and texture resolution. In order to study the quality degradation related to geometry and texture, we simplified the models at different stages until further simplification produced visible degradation.

When the distance between the viewer and an object is fixed, the best display quality is achieved when one texel (pixel on texture image) maps exactly to one pixel in the rendered image. If the texture resolution is higher, a number of texels are averaged and the resulting value is mapped onto a pixel. In other words, extra resolution is reduced. When the texture is of a lower resolution, a texel is duplicated and the same values are mapped onto the corresponding pixels. In the experiments, all objects were normalized to 2.0-m height and the distance from the object to view platform was about 4.0 m in the virtual world. The view platform was placed at a distance so that the whole object was visible and occupied most of the space in the vertical direction, which was approximately 750 pixels in height given a display monitor with resolution of 1280 × 1024. The actual distance between an observer and the image plane (CRT display) was 0.4572 m (1.5 ft). We reduced the resolution of all texture images to 750 pixels in height. The width of texture images was decreased in proportion to the height.

Many algorithms are available for mesh simplification. We focus on the change of visual quality resulting from reduced geometry and texture resolution, using a *Regular Grid Mesh*. Grid sampling can easily be obtained from range scanners. The objects were initially composed of 360 × 30 polygons. Then simplification was performed in the X and Y axes directions. Each simplification reduced 5% of the vertices in one direction. Each object and its next 5% simplified version were grouped into pairs. Three human viewers were requested to compare each pair of simplified objects and the process did not stop until the first pair of visually different objects was found and agreed upon by all three parties. The simplest polygonal mesh preserving fidelity identified in this process was used as the full-resolution model in our subsequent experiments.

To ensure consistent physical luminance, all evaluation experiments were performed on a DELL 1.8-GHZ, 512-MB, Geforce3 workstation, which was equipped with a 21-in Trinitron monitor. The resolution of the display was 1280 × 1024.

C. Experimental Process

In the experiments, there were five different visual objects, each object was represented by six levels of wireframe resolution and three levels of texture resolution, giving a total of $5 \times 6 \times 3$ rating stimuli. These stimuli were evaluated by ten participants, all of whom were Computing Science students with no prior knowledge of these objects. Two more referential stimuli were displayed side by side with the rating stimulus for comparison. One referential stimulus had the highest geometry and texture resolution, and the other had the lowest geometry and texture resolution. The highest quality referential stimulus was assigned a rating of 5 (very good), and the lowest quality one was assigned 1 (very poor). The referential stimuli were rotated at the same speed as the target stimulus to be rated, and the rotation speed of all three could be adjusted simultaneously. The participants (judges) were asked to compare the target stimulus with the two referential stimuli and assign it one of the following ratings: ***very poor*** (1), ***poor*** (2), ***fair*** (3), ***good*** (4), ***very good*** (5).

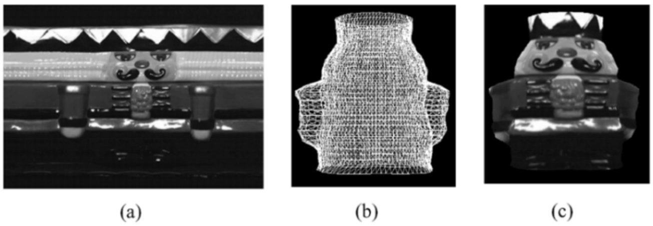


Fig. 6. Texture, wireframe, and the canonical view of Nutcracker. (a) Texture. (b) Wireframe. (c) Canonical view.

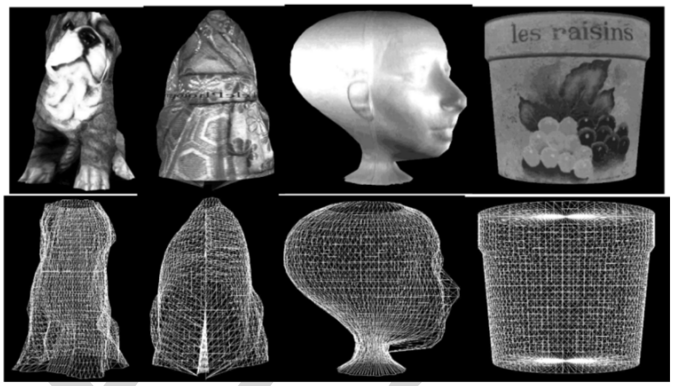


Fig. 7. Other objects (dog, doll, head, and pot) used in qualitative experiments.

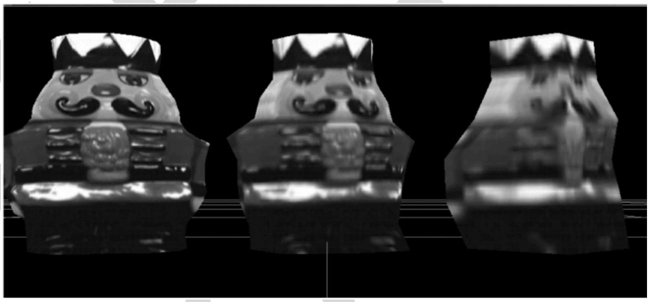


Fig. 8. Evaluation example.

Fig. 8 illustrates two referential stimuli (left and right) and one target stimulus (center) in the experiment. In order to avoid the effect of the temporal sequencing factor, the sequence of 90 target stimuli was randomly generated so that no two participants shared the same sequence, and each participant made his/her decision independently. Only one object was displayed at a time to the viewer.

1) A Metric for Estimating Perceptual Quality: Given the same texture resolution, the image quality improves with the augmentation of wireframe resolution, which creates a finer geometry. When the wireframe resolution is low, a marginal increase in resolution shows a relatively significant improvement in quality. On the other hand, when the resolution is high, an augmentation in resolution is not as significant. When the wireframe resolution reaches a particular density at the high end, further increase is no longer perceptible to the human eyes. We assume that, given the same texture resolution, the image quality curve of different wireframe resolutions vary exponentially. This assumption is represented by (2), which is a function of g and will be tested by our experimental results.

$$\text{Quality} = \frac{1}{a + be^{-cg}} \tag{2}$$

We define g as a variable representing the level of detail of geometric data, which is implemented as the square root of the number of vertices, and a, b, c are constant coefficients. We define the minimum quality to be m, and the maximum quality to be M. When $\text{g} \to \infty, Q(\infty)$ denotes the image quality with an optimal wireframe, i.e., $Q(\infty) = 1/a = \text{M}$. When $\text{g} \to 0, Q(0) = 1/(a+b) = m$. Thus, we can deduce the constant coefficients a and b as follows:

$$a = 1/Q(\infty), \quad b = 1/m - 1/Q(\infty). \tag{3}$$

Since quality also varies with texture, $Q(\infty)$ is also a function of texture (say $Q(\infty) = f(t)$), where t represents texture resolution. Substituting a and b from (3), and for $Q(\infty)$ into (2), we get

$$\text{Quality } (g, t) = \frac{1}{\frac{1}{f(t)} + \left(\frac{1}{m} - \frac{1}{f(t)}\right) e^{-cg}}. \tag{4}$$

In (4), $g \in [0, \infty]$, when g is equal to or larger than the optimal geometry G_0, the extra detail is not perceptible to the human visual system, thus the optimal quality can be written as

$$\text{Quality } (\infty, t) = \text{Quality } (G_0, t). \tag{5}$$

For convenience, it is desirable to normalize the value of geometry to the interval $[0, 1]$. We map g from interval $[0, G_0]$ to g' in the interval $[0, 1]$ using the function

$$g' = \frac{1 - e^{-g}}{1 - e^{-G_0}}. \tag{6}$$

Note that the transformation from g to g' with scaling based on G_0, eliminates the problem of varying units and scales for different objects and different experimental setup.

Combining (4) and (6) gives

$$\text{Quality } (g', t) = \frac{1}{\frac{1}{f(t)} + \left(\frac{1}{m} - \frac{1}{f(t)}\right) (1 - g'K)^c} \tag{7}$$

where $K = 1 - e^{-G_0}$. Note that even if G_0 is as small as 5, e^{-G_0} is about 0.007. Hence, K is very close to 1. Thus, (7) can be simplified for practical situations to (8)

$$\text{Quality } (g', t) = \frac{1}{\frac{1}{f(t)} + \left(\frac{1}{m} - \frac{1}{f(t)}\right) (1 - g')^c}. \tag{8}$$

For the experimental curve fitting in the next section, we use (8) and the assumption that the scaled geometry parameter (g') varies between 0 and 1. (For example, see (10) and Fig. 10, geometry resolution axis.)

TABLE I
QUALITY VERSUS TEXTURE RESOLUTION (100% GEOMETRY RESOLUTION)

Texture / Object	25%	37.5%	50%	62.5%	75%	87.5%
Nutcracker	2.2	2.5	3.2	3.6	4.1	4.5
Head	2.3	2.6	3.1	3.3	3.9	4.3
Dog	2.2	2.6	3.2	3.7	4.2	4.7
Doll	2.0	2.5	3.2	3.6	4.1	4.6

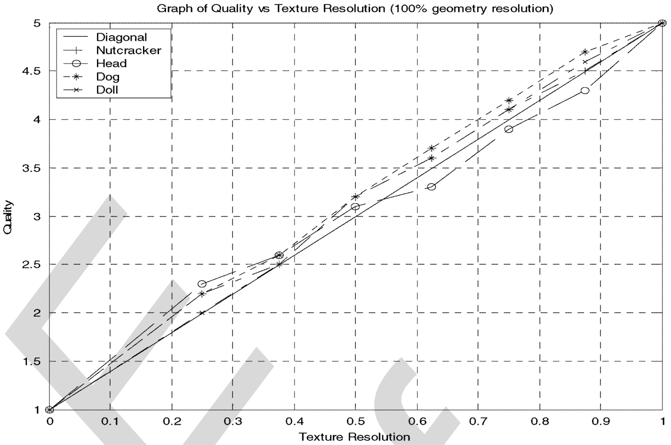


Fig. 9. Quality versus texture resolution (100% geometry resolution).

There are three problems remaining in this quality metric.

1) How closely the quality curve for images under the same texture resolution and different wireframe resolutions follow the exponential property?
2) How to resolve the function of texture resolution $f(t)$?
3) How to determine coefficient c?

We are going to answer the first two questions through the analysis of evaluation results in the next section, and discuss the other question for a specific set of objects.

V. EXPERIMENTAL RESULTS AND ANALYSIS

From the experiments, it was observed that results for the objects Head, Dog, and Doll are very close to the Nutcracker and are not shown separately to reduce the length of this document. The qualitative evaluations for these objects fitted (8) quite well for different values of c. Results for the Pot object did not fit an exponential curve well; this occurred because people found a square (distorted) pot acceptable compared to the original (round) pot. Once the exponential relation between geometry parameter and visual quality is verified, the problem that remains is to study the property of quality related to texture resolution. That is, how to resolve $f(t)$ in (8). $f(t)$ is the visual quality for maximum geometry resolution and various texture resolutions. More samples are collected to resolve this function. For each object other than Pot, four more stimuli are tested. All of these stimuli have maximum geometry resolution and different texture resolutions. Combined with samples from former experiments, there are six samples along the quality axis direction for each object with 100% geometry resolution. The data is listed in Table I and plotted in Fig. 9.

In Fig. 9, all samples points are distributed close to the diagonal, that is, $f(t)$ approximately follows a linear regression

$$f(t) = m + (M - m)t, \quad t \in [0, 1]. \tag{9}$$

Substituting the expression for $f(t)$ in (9) into (8) gives

$$\text{Quality } (g, t) = \frac{1}{\frac{1}{m+(M-m)t} + \left(\frac{1}{m} - \frac{1}{m+(M-m)t}\right)(1-g)^c}. \tag{10}$$

At first, curves were fitted based on observed values of $f(t)$. The graph for fitting based on experimental $f(t)$ for the Nutcracker object is shown in Fig. 10. However, determining $f(t)$ from experiments is not feasible before estimating perceptual quality in an experimental setup in general. Thus, $f(t)$ was estimated based on linear regression [(9)] after observing the data in Fig. 9. [Note that $(ai, bi, ci), i = 1, 2, 3$ denotes the fitting parameters for the various texture resolutions, and $\text{ssei}, i = 1, 2, 3$ denotes the sum of square errors for the respective cases.]

In Table II, the first three rows are the standard deviation of the vertical distances from the sample points to the corresponding curves fitted at different texture resolutions. The quality values of best (100% texture and 100% geometry) and worst (25% texture and lowest geometry) stimuli are fixed at 5 and 1, respectively, and produce no error, thus the sizes of the groups for 100%, 50%, and 25% texture are counted as 5, 6, and 5, respectively. Let r be the distance between a point and the fitting curve. SD is the standard deviation—shown in (11)—of r at each texture level, where N is the size of the group

$$\text{SD} = \sqrt{\frac{\sum_i r_i^2}{N-1}}. \tag{11}$$

The overall standard deviation SD_{all} is calculated as in (12).

$$\text{SD}_{\text{all}} = \sqrt{\frac{\sum_i r_i^2}{df}}. \tag{12}$$

Here, df is the degree of freedom and is equal to the number of sample points minus the number of parameters fitted. $df = 16 - 2 = 14$ in our experiments.

We perform a curve fitting based on (10) for validation. The results are listed in Table III and plotted in Figs. 11 and 12 for two of the objects. Since the values of $f(t)$ are estimated from (9) and substituted into (10), we use the notation of "estimated $f(t)$" as opposed to the results in which the values of $f(t)$ are obtained from experiments.

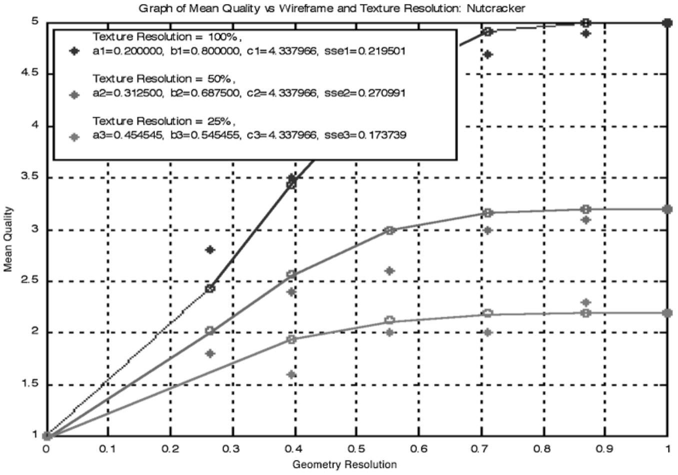


Fig. 10. Comparing best fitting with experimental $f(t)$ to average perceptual evaluations (Nutcracker).

TABLE II
STANDARD DEVIATION OF RESIDUALS WITH EXPERIMENTAL $f(t)$

Object / Texture	Nutcracker	Head	Dog	Doll	Pot
SD (Texture 100%)	0.21	0.29	0.30	0.08	0.47
SD (Texture 50%)	0.21	0.18	0.24	0.15	0.13
SD (Texture 25%)	0.18	0.45	0.31	0.24	0.37
SD_{all}	0.22	0.34	0.32	0.20	0.36
C	4.34	4.39	4.59	4.19	7.03

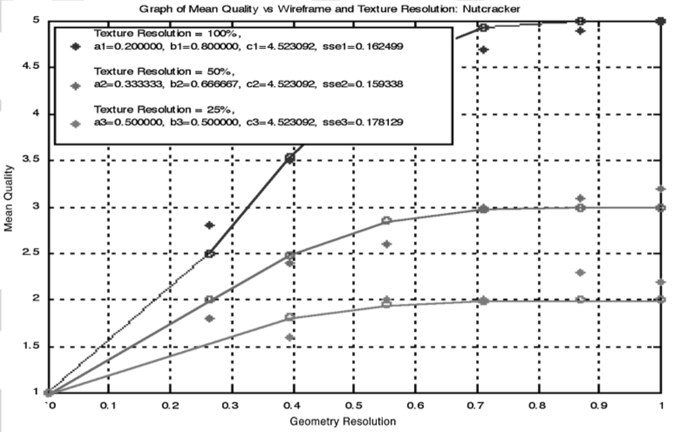


Fig. 11. Comparing best fitting with estimated $f(t)$ to average perceptual evaluations (Nutcracker).

The discrete points in Figs. 10, 11, and 12 are the averages taken from ten participants' perceptual quality evaluation values based on a maximum of 18 stimuli for each texture resolution on each of the two objects, Nutcracker and Pot. An additional four stimuli were used at the highest geometry level to better estimate the linear regression given by $f(t)$. (There were only three stimuli, three rotating objects, shown at a time for a given level of geometry resolution level.) Results for the other objects (Head, Dog, and Doll) are very close to the Nutcracker and are not shown separately. From the figures, we can observe that perceptual quality roughly follows an exponential distribution for geometry as we predicted. At each level of texture resolution, there is a diminishing return; the same percentage increase in geometry resolution will show more visible improvement in quality at the low-resolution end than at the high-resolution end. Some exceptional points exist in these figures, for example, in Fig. 11, the sample point $\text{Q(g, t)} = \text{Q}(87\%, 25\%)$ has the value 2.30, and Q(100%, 25%) has the value 2.20, but Q(100%, 25%) is supposed to have a better quality than Q(87%, 25%). Several factors combined to produce such exceptional points. First, stimuli were evaluated in

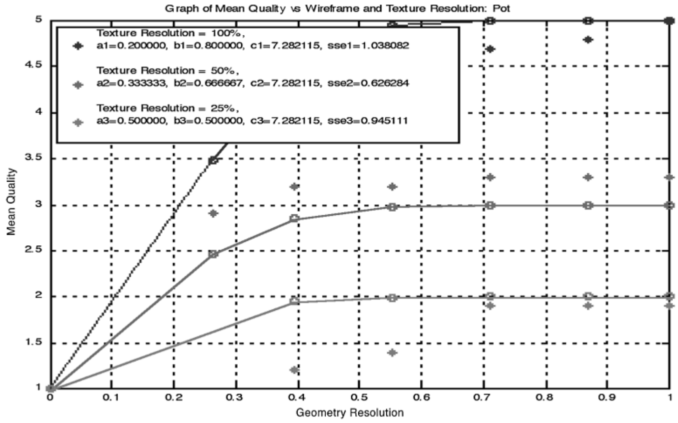


Fig. 12. Comparing best fitting with estimated $f(t)$ to verage perceptual evaluations (Pot).

TABLE III
STANDARD DEVIATION OF RESIDUALS WITH ESTIMATED $f(t)$ FOR ALL OBJECTS

Object Texture	Nutcracker	Head	Dog	Doll	Pot
SD (Texture 100%)	0.18	0.25	0.27	0.11	0.47
SD (Texture 50%)	0.16	0.16	0.18	0.20	0.13
SD (Texture 25%)	0.19	0.29	0.23	0.27	0.37
SD_{all}	0.19	0.25	0.26	0.22	0.36
c	4.52	4.59	4.41	4.38	7.03

a random sequence, and each stimulus was compared to the referential stimuli independently, with one target appearing at a time. Viewers were asked to evaluate the quality according to the proximity between target and referential stimuli. If the viewer felt that the target was closer to the left stimulus, he/she would give a rating of 5 or 4 depending on the degree of quality. If the target stimulus was closer to the other end, a rating of 1 or 2 would be assigned. If the viewer had no preference, he/she would probably give a rating of 3. Therefore, there was no direct comparison among the target stimuli themselves, and an evaluation error was likely to occur for two stimuli with close quality. We notice that these exceptional points are located at the high geometry resolution regions, because visual quality changes only marginally with the variation of geometry resolution, causing the human visual system difficulty in rating two high resolution stimuli accurately. Second, the number of participants, ten, is relatively small, and any exceptions such as viewer illusion or operational mistakes (values are recorded by choosing one of five buttons) can easily cause an error of ± 0.1. Although such error cannot be totally eliminated, with a larger group of participants the effect of the error can be reduced. Fortunately, evaluation errors in the results are small and do not change the overall property of the curves.

In Section IV, the quality metric was given in (8). We perform curve fitting minimizing the sum of square error over each of the five datasets. The results for the Nutcracker and the Pot are illustrated in Figs. 11 and 12. In the legends, $a = 1/f(t), b = 1/m - 1/f(t)$, and c is the constant in (8).

Despite of the poor performance of Object "Pot", all of the four other objects have a good fit. The standard deviation ranges from 0.20–0.34, which is less than 10% of the dynamic range. The exponential coefficient c has a reasonable value around 4.40. More importantly, sample points are fairly evenly distributed around the fitting curve. The $\mathrm{SD}_{\mathrm{all}}$ value of Pot is not large either, but an exponential coefficient as high as 7.03 describes a curve that is nearly constant in the intermediate to high geometry resolution, and drops abruptly for low geometry resolution. A closer examination reveals uneven deviation of sample points around the fitted curve. The points are close to a straight line in higher geometry resolution area and rather poorly fitted in lower geometry resolution area. Thus, the fitting does not serve much better than a piecewise linear function. Given more samples in low geometry resolution end, the $\mathrm{SD}_{\mathrm{all}}$ for this example will become too large to be acceptable.

Note the sample point $Q(g,t) = Q(26\%, 25\%)$, which is used as the referential stimulus of worst quality, is excluded in the curve fitting. In the derivation of (6), we assume $Q(0,t) = m$; however, since an object without geometry has no visual appearance $Q(26\%, 25\%) = m$ is set for comparison purpose. Consequently, this sample point does not follow (6). As a matter of fact, if $c = 4.4$ and $f(t) = 2.0, Q(26\%, 25\%) = 1.58$. With a finer-granularity rating system in which $x.5$ is available, 1.5 is a better setup value for the worst quality referential stimulus. In our experiments, the worst quality reference is under-estimated; for this reason, sample points close to it are likely to be under-estimated as well. This is also evident in the figures and Table II, where the values of sample points at 25% texture resolution decrease faster than predicted, resulting in a relatively higher value of standard deviation (Row 3 in Table II).

Comparing Table III with Table II, the $\mathrm{SD}_{\mathrm{all}}$ value of Doll remains nearly the same, but the $\mathrm{SD}_{\mathrm{all}}$ values of the first three objects become smaller, which means a comparatively better fitting. This phenomenon can be explained as follows. The former fitting ensures exact matches along 100% geometry resolution axis, and matching in other regions is not as good. However, the experimental values of $f(t)$ may have accidental errors. Therefore, the entire sample space rather than samples in 100% geometry resolution axis alone conform to (11). We notice that the exponential coefficient c, although not a constant for different objects, varies inside a narrow interval. For those applications that do not require high accuracy, 4.50 can be used as an estimate for c; otherwise, some samples should be collected to better determine this parameter.

Our metric does not estimate well the quality of object Pot. From the feedback of viewers, we found that while users were requested to rate based on comparing the target stimuli to two referential stimuli, not all of them followed the guidelines equally well. The first impression played an important role in their evaluations. Since square pots exist in the real world, when the geometry resolution of Pot is decreased, some of them felt "pot can be square, not necessarily round," and unconsciously gave a higher than expected rating to some distorted stimuli. Psychologically, this proves prior knowledge is an important issue in quality evaluation. For instance, the quality degradation of a face is easier to detect than that of a rock because people are more familiar with the structure of a face. In our experiments, besides the referential stimuli, prior knowledge also serves as a pattern in the evaluation. Although such effects cannot be bypassed in any psycho-visual experiment, cautious selection of stimuli may reduce this problem. In conclusion, the Pot object is not an appropriate candidate for our experiments.

Tables IV–VII give more details on the exact average user evaluations for the objects (Nutcracker and Pot) plotted in the figures, along with the standard deviations of the evaluations.

1) Reliability of the User Evaluations: It is important to discuss issues relating the reliability of our perceptual evaluations. The reliability discussions are based on studies described by Guilford [7]. As stated on page 279 in [7] a "reliability rating of 0.90 can be obtained with 10 to 50 judges, and a reliability rating of 0.95 can be obtained with 21 to 106 judges." It was noted that reliability "increases with the number of judges."

It was observed in the book that reliability of measurement depends on "self-correlation" where repeated sets of measurements are correlated to each other. Forming groups for self-correlation is however a difficult and time consuming task that is beyond the scope of our preliminary study. A group of judges for self-correlation needs to have comparable judging behavior. Thus, extensive psychological behavior tracking over time needs to be conducted to create consistent groups of judges. Given that we do not have groups of judges to measure self-correlation, we performed the following experiments instead. In addition to the experimental results already described we conducted tests with an additional 10 judges after a time interval of about 18 months. The results obtained using the second group for the Nutcracker object are summarized in Tables VIII and IX, and the overall results for the two groups (i.e., 20 judges) for the same object are shown in Tables X and XI. Observe in Table X that with 20 judges there are no inconsistencies in the mean evaluations; i.e., the means are nondecreasing for increasing geometry resolution at each level of texture resolution.

Note that the results for the second group are very close to the first, and that the variations of the combined group are close to the individual groups. Instead of computing correlation between similar judges, we will consider correlations of the average evaluations of the two groups of judges. Consider the average ratings in Tables IV and VIII for the Nutcracker object for the two groups of ten judges. If we pair ratings at corresponding texture and wireframe resolution levels, we get the (X, Y) pairs given in Table XII.

The correlation between these pairs of average evaluations is equal to 0.9938.

The (X, Y) pairings for the dog and head objects are given in Tables XIII and XIV).

For the dog and head objects the corresponding correlations computed to 0.9891 and 0.9947, respectively.

This study, though not strictly according to psychometric guidelines shows a strong association between repeated sets of measurements on a "similar" group of judges, leading us to believe that the reliability of our study should be fairly high, possibly higher than 0.95 (or 95%).

TABLE IV
MEAN QUALITY OF USERS' EVALUATIONS FOR NUTCRACKER

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	2.80	3.50	4.60	4.70	4.90	5.00
50%	1.80	2.40	2.60	3.00	3.10	3.20
25%	1.00	1.60	2.00	2.00	2.30	2.20

TABLE V
STANDARD DEVIATION OF USERS' EVALUATIONS FOR NUTCRACKER

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	1.15	0.85	0.70	0.42	0.48	0.00
50%	0.82	0.82	0.92	0.67	0.97	0.79
25%	0.00	0.84	0.82	0.82	0.67	0.79

TABLE VI
MEAN QUALITY OF USERS' EVALUATIONS FOR POT

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	3.80	3.90	4.30	4.70	4.80	5.00
50%	2.90	3.20	3.20	3.30	3.30	3.30
25%	1.00	1.20	1.40	1.90	1.90	1.90

TABLE VII
STANDARD DEVIATION OF USERS' EVALUATIONS FOR POT

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	1.03	1.06	0.67	0.67	0.95	0.00
50%	1.03	0.71	0.71	0.74	0.94	0.63
25%	0.00	0.52	0.52	1.10	0.99	0.74

TABLE VIII
MEAN QUALITY OF USERS' EVALUATIONS FOR NUTCRACKER (GROUP 2)

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	2.80	3.70	4.60	4.70	4.80	5.00
50%	1.70	2.30	3.00	3.20	3.10	3.10
25%	1.00	1.70	2.00	2.10	2.20	2.30

The number of stimuli was also noted as an important factor in [7]. In [7, p. 280], it was observed that: "ranking becomes difficult and irksome when there are more than 30 to 40 stimuli." In consideration of this factor it should be noted that we use only three stimuli at a time (Fig. 8) in our evaluation experiment, and total number of stimuli per judge is limited to 22 per object. We believe that the simplicity of our experiments resulting in fewer stimuli is much more likely to produce more reliable results than experiments trying to evaluate more factors which will obviously result in more stimuli per object.

2) Comparing Results to Recent Perceptual Experiments: Compared with recent related perceptual experiments [15], [21], in which none of the metrics model perceptual

TABLE XII
COMPARING AVERAGE RATINGS AT SAME RESOLUTION LEVELS FOR NUTCRACKER

X	1.0	1.6	2.0	2.0	2.3	2.2	1.8	2.4	2.6	3.0	3.1	3.2	2.8	3.5	4.6	4.7	4.9	5.0
Y	1.0	1.7	2.0	2.1	2.2	2.3	1.7	2.3	3.0	3.2	3.1	3.1	2.8	3.7	4.6	4.7	4.8	5.0

TABLE XIII
COMPARING AVERAGE RATINGS AT SAME RESOLUTION LEVELS FOR DOG

X	1.0	1.4	1.7	2.0	2.1	2.2	1.8	2.2	2.6	3.0	3.0	3.2	2.9	3.6	4.4	4.6	4.8	5.0
Y	1.0	1.4	1.7	2.1	2.2	2.2	1.9	2.6	3.1	3.0	3.1	3.1	2.8	3.7	4.1	4.5	4.9	5.0

TABLE XIV
COMPARING AVERAGE RATINGS AT SAME RESOLUTION LEVELS FOR HEAD

X	1.0	1.3	1.7	1.9	2.0	2.3	1.7	2.3	2.8	3.1	3.0	3.1	2.9	3.8	4.3	4.7	4.9	5.0
Y	1.0	1.5	1.6	1.9	2.1	2.2	1.8	2.4	2.8	2.8	3.0	3.2	2.7	3.9	4.1	4.6	4.9	5.0

TABLE IX
STANDARD DEVIATION OF USERS' EVALUATIONS FOR NUTCRACKER (GROUP 2)

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	1.08	0.67	0.70	0.67	0.42	0.00
50%	0.82	0.79	0.84	0.79	0.70	0.74
25%	0.00	0.82	0.74	0.88	1.08	0.67

TABLE X
MEAN QUALITY OF USERS' EVALUATIONS FOR NUTCRACKER (COMBINED)

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	2.80	3.60	4.60	4.70	4.85	5.00
50%	1.75	2.35	2.80	3.10	3.10	3.15
25%	1.00	1.65	2.00	2.05	2.25	2.25

TABLE XI
STANDARD DEVIATION OF USERS' EVALUATIONS FOR NUTCRACKER (COMBINED)

Geometry Texture	26%	39%	55%	71%	87%	100%
100%	1.09	0.76	0.68	0.54	0.44	0.00
50%	0.81	0.80	0.90	0.73	0.82	0.76
25%	0.00	0.82	0.76	0.84	0.87	0.72

quality well and only comparative accuracies of different metrics are provided, our research provides a perceptually based metric that accurately fits the qualitative evaluation results. In addition, while previous experiments only consider nontextured model, our metric describes how texture resolution as well as geometry resolution control the overall quality of 3-D images. The experiments in [21] use still 2-D images as stimuli, and Rogowitz *et al.* proved that still views are not sufficient by comparing the results of view-dependent 2-D and view-independent 3-D stimuli evaluations [15]. We also advance a step further from Rogowitz *et al.* to provide a rotation speed control for visual stimuli. Given this flexibility in the experimental interface, the spatial factor can be reduced. Viewers have a more dynamic impression of the object and are more confident in rating the quality.

The fitting result with one of the objects, Pot, was not satisfactory because a distorted Pot was considered acceptable by some viewers. It is, therefore, important to note that our quantitative model should only be used to estimate perceptual quality of objects for which geometry is an important component of perceived shape. For example, stones or deformable objects such as flags and balloons are not suitable objects for the proposed metric.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we first discussed factors controlling 3-D image degradation and quantitative error measures approximating qualitative evaluations. A review of literature on evaluating depth perception [13] was given, followed by discussion of some of the methods for evaluating visual quality of objects not considering depth estimation *per se*. In previous perceptual experiments modeling visual evaluation, the results suggest that proposed qualitative error measures are not always good indicators of perceptual quality. The correlation of qualitative measures, such as "naming time", to existing standard error measures (such as BMP, mse, MetroMN) was also compared in prior research. However, new quantitative measures that are designed to model 3-D quality have not been proposed. In order to extend prior research, we first examined the factors that determine the quality of 3-D images including geometry resolution, texture resolution, shading complexity, frame rate and other psycho-visual factors. Of these factors, two (texture and geometry resolution) that affect bandwidth requirements were considered in our initial research. We designed a perceptual experiment and derived from the results a quantitative metric that approximates perceptual quality and reflects how geometry and texture resolution control the overall quality of 3-D images. The strengths and limitations of this metric were also analyzed. A preliminary study suggested that the reliability of the evaluations is possibly higher than 0.95.

From the figures showing experiment results, we observe that the degradation of visual quality follows an exponential model for the geometry resolution parameter, and a linear model for the texture resolution parameter. This suggests that human viewers are far more sensitive to the distortion of texture than to that of geometry.

We plan to increase the strength of our findings by increasing the number of parameters in the model. We also want to check whether a finer-granularity rating system provides better results than the current experiments. Automatic simplification methods should be adopted to obtain the simplest ideal model, which can reduce preprocessing effort, especially if a large number of models are used as stimuli. Finally, it is important to incorporate more factors such as distance, shading, and visual masking into the proposed metric.

ACKNOWLEDGMENT

The support of the Alberta Science and Research Authority (ASRA) and the National Science and Engineering Research Council (NSERC) in making this work possible is gratefully acknowledged, as are helpful discussions with Prof. S. Nagata. A note on the contributions of various authors: Y. Pan performed the detailed modeling of the quantitative metric, conducted user evaluations, and prepared the initial literature survey; I. Cheng produced the 3-D data and models used in the experiments and revisions to the draft report; A. Basu proposed the original problem and suggestions for modeling, prepared the first journal draft and addressed the statistical issues relating to estimating reliability of the evaluations.

REFERENCES

[1] A. Basu and K. J. Wiebe, “Enhancing videoconferencing using spatially varying sensing,” *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 28, no. 2, pp. 137–148, Mar. 1998.

[2] M. R. Bolin and G. W. Meyer, “A perceptually based adaptive sampling algorithm,” in *Proc. ACM SIGGRAPH*, 1998, pp. 299–309.

[3] P. Cignoni, C. Montani, and R. Scopigno, “A comparison of mesh simplification algorithms,” in *Computers & Graphics*. New York: Pergamon, 1997, vol. 22.

[4] S. Daly, “The visible differences predictor: An algorithm for the assessment of image fidelity,” in *Digital Images and Human Vision*, A. B. Watson, Ed. Cambridge, MA: MIT Press, 1993, pp. 179–206.

[5] M. Deering, “Geometry compression,” in *Proc. ACM SIGGRAPH*, 1995, pp. 13–19.

[6] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg, “A model of visual masking for computer graphics,” in *Proc. ACM SIGGRAPH*, 1997, pp. 143–152.

[7] J. P. Guilford, *Psychometric Methods*. New York: McGraw-Hill, 1936.

[8] P. Haeberli and M. Segal. (1993) Texture Mapping as a Fundamental Drawing Primitive. [Online]. Available: http://www.sgi.com/grafica/texmap

[9] P. S. Heckbert and M. Garland, “Survey of Polygonal Surface Simplification Algorithms,” School of Comput. Sci., Carnegie-Mellon Univ., , Pittsburgh, PA, Tech. Rep., 1997.

[10] C. J. van Den, B. Lambrecht, and O. Verscheure, “Perceptual quality measure using a spatio-temporal model of the human visual system,” in *Proc. SPIE*, 1996, pp. 450–461.

[11] J. O. Limb, “Distortion criteria of the human viewer,” *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 12, pp. 778–793, Dec. 1979.

[12] J. L. Mannos and D. J. Sakrison, “The effects of a visual fidelity criterion on the encoding of images,” *IEEE Trans. Inform. Theory*, vol. IT-20, no. 4, pp. 525–535, Jul. 1974.

[13] S. Nagata, “How to reinforce perception of depth in single 2D pictures, -comparative study on various depth cues,” in *Proc. SID*, vol. 25, 1984, pp. 239–247.

[14] ——, “How to reinforce perception of depth in single 2D pictures-Comparative study on various depth cues,” in *Pictorial Communication in Virtual and Real Environments*, S. Ellis, Ed. London, U.K.: Taylor and Francis, 1991, pp. 527–545.

[15] B. Rogowitz and H. Rushmeier, “Are image quality metrics adequate to evaluate the quality of geometric objects?,” in *Proc. SPIE*, vol. 4299, 2001, pp. 340–349.

[16] H. Rushmeier, B. Rogowitz, and C. Piatko, “Perceptual issues in substituting texture for geometry,” in *Proc. SPIE Human Vision and Electronic V*, vol. 3959, 2000, pp. 372–383.

[17] M. Siegel and S. Nagata, “Just enough reality: Comfortable 3D viewing via micro-stereopsis,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 3, pp. 387–396, Apr. 2000.

[18] C. S. Stein, A. B. Watson, and L. E. Hitchner, “Psychophysical rating of image compression Techniques,” in *Proc. SPIE*, vol. 1977, 1989, pp. 198–208.

[19] P. Teo and D. Heeger, “Perceptual image distortion, human vision,” in *Proc. SPIE—Visual Processing, and Digital Display V*, vol. 2179, Feb. 1994, pp. 127–141.

[20] J. Torborg and J. T. Kajiya, “Talisman: Commodity realtime 3D graphics for the PC,” in *Proc. SIGGRAPH’94, Computer Graphics*, Aug. 1996, pp. 353–363.

[21] B. Watson, A. Friedman, and A. McGaffey, “Measuring and predicting visual fidelity,” in *ACM SIGGRAPH*, 2001, pp. 213–220.

[22] A. Webster *et al.*, “An objective video quality assessment system based on human perception,” in *Proc. SPIE Human Vision, Visual Processing, Digital Display TV*, San Jose, CA, 1993, pp. 15–26.

[23] S. Winkler, “A perceptual distortion metric for digital color video,” in *Proc. f SPIE Human Vision and Electronic Imaging*, San Jose, CA, Jan. 1999, pp. 23–29.

[24] Y. Yu, I. Cheng, and A. Basu, “Optimal adaptive bandwidth monitoring for QoS based retrieval,” *IEEE Trans. Multimedia*, vol. 5, no. 3, pp. 466–472, Sept. 2003.

Yixin Pan was born in Quanzhou, China, in 1977. He received the B.S. degree in computer science from Tsinghua University, Beijing, China, in 2000 and the M.Sc. degree in computing science from the University of Alberta, Edmonton, AB, Canada, in 2002.

He is currently a Software Developer at ****AUTHOR: PLS. SUPPLY COMPANY NAME****, Buffalo, NY. His research interests include three-dimensional graphics animation, distributed graphics systems, and perceptual image quality evaluation.

Irene Cheng (S’03) is currently an NSERC Ph.D. scholar in the Department of Computer Science, University of Alberta, Edmonton, AB, Canada.

She has worked with Lloyds Bank on financial databases and with TelePhotogenics Inc., Edmonton, as the Chief Software Architect. Her research interests include optimal monitoring of distributed networks and efficient retrieval of super-high-resolution three-dimensional visual information. She has over ten publications in international conferences and journals.

Anup Basu (S’89–M’90) received the Ph.D. degree in computer science from the University of Maryland, College Park, in 1990.

He has worked with Tata Consultancy Services, Mumbai, India, and Strong Memorial Hospital, Rochester, NY, on networked banking software and graphical visualization of medical data, respectively. Since 1990, he has been with the Department of Computer Science, University of Alberta, Edmonton, AB, Canada, where he is currently a Professor. He has also collaborated in starting up companies, including TelePhotogenics Inc., Edmonton. He has over 80 publications in journals and conferences, and holds a patent on super-high-resolution stereo imaging. His research interests include super-high-resolution three-dimensional imaging, multimedia communications, and foveated multimedia systems.