A 3D Perceptual Metric using Just-Noticeable-Difference

Irene Cheng and Pierre Boulanger
Department of Computing Science, Univ. of Alberta, Edmonton, CANADA,{lin,pierreb}@cs.ualberta.ca

Abstract
In multimedia applications, it is essential to distribute resources efficiently among different types of data in order to optimize overall quality. We propose a perceptual metric using Just-Noticeable-Difference (JND) to identify redundant mesh data so that available bandwidth can be allocated to improve texture resolution. Evaluation of perceptual impact during runtime is based on statistics in a lookup table generated during preprocessing. If the impact is less than the JND, no mesh refinement is performed. We apply Weber's fraction to compute the JND threshold, which is verified by perceptual evaluations. Experimental result shows that our JND model can accurately predict perceptual impact based on the human visual system.

1. Introduction

When transmitting 3D textured mesh (TexMesh) over a shared network, limited resources such as bandwidth has to be allocated between both mesh and texture data. Progressive refinement strategies [Hop96][KSS03] often assume that visual quality improves as the mesh resolution increases, ignoring the experimental finding that texture resolution has more significant impact on quality after the mesh resolution has reached a certain threshold [PCB05][RRP00]. Geometric metrics were commonly used in previous simplification techniques [HH93] [GH98]. However, perceptual metrics [OHM*04] have been gaining increased attention among researchers for two main reasons: First, visual fidelity is ultimately determined by the Human Visual System (HVS), and thus using perceptual metrics is expected to be more accurate. Second, assessment relying on geometric criteria, such as mean square error (MSE) or quadric error [GH98] is not sufficient because geometrically different objects can be visually indistinguishable to the HVS. Transmitting redundant mesh data without improving visual quality is a waste of resources [CB04].

In this paper, we present a mathematical model to measure the perceptual values associated with 3D vertices, which are used to predict the benefit to visual fidelity when refining a coarse mesh to a denser version. In order to maximize the overall quality, the server decides, based on the statistics gathered during preprocessing, whether mesh refinement should terminate, allocating the remaining bandwidth to increase texture resolution. Our goal is to locate a perceptual threshold (Just-Noticeable-Difference or JND), where the HVS can just distinguish the difference between two levels-of-detail (LOD). We locate and verify the JND by conducting perceptual evaluation experiments with texture mapped on to the mesh. We consider texture mapping for two reasons: (a) it is easy to visually identify differences in mesh only, and (b) our goal is to optimize the perceptual quality of photo-realistic 3D objects given bandwidth limitations. Online transmission of 3D TexMesh can then be more efficient, by suppressing imperceptible geometric data, which have dimension below the JND.

We use Scale-Space Filtering (SSF) to extract 3D features [CB05]. Traversal between the different scales is achieved by varying the standard deviation parameter σ; the higher the value of σ the more is the smoothing [Wit83][KF01]. SSF is based on locating the zero-crossings of a signal at multiple scales. Zero-crossings are used to detect the degree of persistence of a structure (feature) on a 3D surface. Minor structures tend to diminish as σ increases, and only major structures survive at higher scales (Fig. 1).

Decimation and refinement are performed using edge collapse and vertex split operations. A detailed discussion of various mesh simplification approaches can be found in [Lue02]. There are two main differences between our edge collapse/vertex split and that used in progressive meshes [Hop96]: (1) There is no vertex relocation between different LOD in our TexMesh; all vertices at a coarse level is a subset of those at a finer level. (2) In progressive meshes, the minimum energy cost, recalculated each time a new vertex is introduced in an edge collapse operation, affects the choice of the next collapsing edge. In our TexMesh model, the order of collapsing edges follows the priority predetermined by applying SSF on the original 3D surface. A vertex is removed by integrating with its closest neighbor, collapsing the edges associated with it.

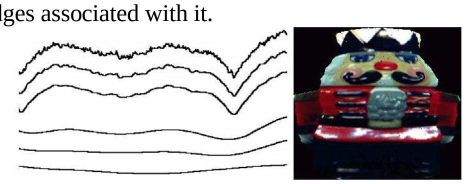

Fig. 1: ***Increasing scale S_i from top to bottom. S_0 is the original signal with 360 vertices near the bottom of the Nutcracker object.***

In related work, Reddy approximates the contrast sensitivity function (CSF) in dynamic scenes to optimize the amount of detail removed from the scene without the user noticing [Red01]. By contrast, our method is designed for comparatively static 3D objects. Perceptual metrics derived from CSF are used to measure the perceptibility of visual stimuli [LH01]. Only simplification operations inducing imperceptible contrast and spatial frequency are performed. However, the technique is not designed to suppress perceptually redundant data. Williams *et al.* [WLC*03] improves prior approaches by accounting for textures and dynamic lighting. The above techniques are view-dependent, while our approach is view-independent. In addition to the reduced navigation costs associated with view-independent algorithms, our perceptual model provides a systematic way, instead of heuristics, to predict visual fidelity. The JND defined in our mathematical model follows the same spirit as Weber's Law on contrast, computed as the change relative to the original value. Experimental results confirm that JND is a constant and is independent of viewing distance.

The rest of this paper is organized as follows: Section 2 extends our SSF model to incorporate perceptual values and JND. Section 3 applies perceptual evaluation experiments to locate and verify the JND. Section 4 outlines the use JND on 3D TexMesh transmission. Finally, Section 5 gives the conclusion and future work.

2. Perceptual value and JND

When a 3D object moves closer to the viewer in a virtual scene, the mesh needs to be refined only if the resulting mesh improves visual quality. To determine whether mesh refinement should be performed requires measuring perceptual impact on the HVS. Adding or deleting a vertex or surface structure from a mesh generates a stimulus to human vision. To compare the perceptual impacts of these stimuli, the dimension of a structure is used as a visual cue in our model. We follow the argument that humans naturally describe an object as consisting of parts and infer 3D shapes of these parts [AS04][ZN99], and segment the object into corresponding parts (skeletonization). In each edge collapse operation during preprocessing, when a vertex V_R is removed and integrated with its closest neighbor V_C, we record the surface change as the difference $\Delta\rho$ between R_R and R_C. R_i is the shortest distance between vertex i and the skeleton. For a spherical object, the skeleton is represented by the center of the object (Fig. 2). $\rho_R = (R_R - R_C)/R_C$ is defined as the *perceptual value* of V_R. If edge $V_Q V_C$ collapses after $V_P V_Q$, the perceptual value of the combined operation is $(R_P - R_C)/R_C$. Our model is designed for view-independent simplification. In a given view, when a 3D object is projected onto a 2D display, the stimulus can be interpreted by Weber's fraction on shape. Also, note that visual impact of a stimulus is dictated locally by the closest adjacent vertex and the closest distance to the skeleton. For example, collapsing $V_R V_C$ has higher impact than collapsing $V_Q V_C$, and we can disregard the overall shape and dimension of the object. Instead of representing the stimulus linearly, an alternative is to use the area of the quadric error generated by removing V_R, but experimental results show that our perceptual metric predicts visual quality well, closely following human perception.

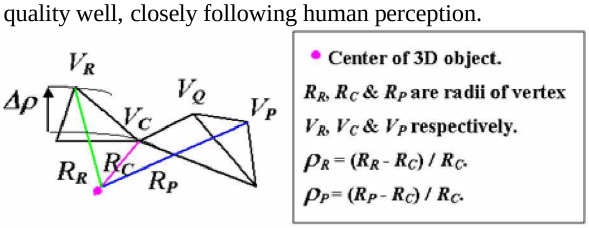

Fig. 2: *V_R and V_P have perceptual values ρ_R and ρ_P respectively.*

Let $\Delta\wp$ be the change when removing V_R and $\wp$ be the distance of V_C from the skeleton. When viewed on the display device, the difference $\Delta\wp$ generates a stimulus to the retina (Fig. 3). The Just-Noticeable-Difference (JND) is the minimum change in perceptual value in order to produce a noticeable variation in visual experience. Weber's Law [GW02] states that at the JND threshold,

$$\frac{\Delta\wp}{\wp} = K \qquad (1)$$ where K is a constant.

A value which is greater than K generates a significant perceptual impact on the HVS. Weber's Law has been applied to a variety of stimuli, including 2D images, brightness, loudness, mass, line length, size, etc., verified by psychovisual experiments. In this paper we extend Weber's fraction to evaluate perceived similarity in 3D TexMesh.

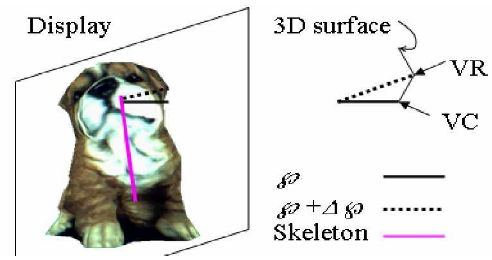

Fig. 3: *An example of perceptual impact generated by the removal of vertex V_R.*

We performed SSF on the nutcracker object (Fig. 1), with 1260 faces at S_0. For each scale change, the perceptual values of vertices removed were recorded. At each scale the average value was used to represent the perceptual impact when refining from S_i to S_{i-1} (Table 1). The cumulated perceptual values were also computed and stored in a lookup table (LUT), so that the perceptual impact between S_i and S_j can be retrieved (Table 2).

Scale	# of faces	Perceptual value	
S_{i-1} - S_i		Avg.	Std.
0-1	1162	0.0410	0.0308
1-2	1118	0.0412	0.0294
2-3	1074	0.0478	0.0390
3-4	1040	0.0468	0.0288
4-5	1002	0.0678	0.0491

Table 1: *$\Delta\wp/\wp$ of the nutcracker mesh between adjacent scales.*

Previous refinement techniques assume that visual quality is proportional to the number of vertices. Our preliminary finding shows that not every set of vertices has significant impact on visual quality [CB04]. Note that the average perceptual value column in the tables indicates that change of scale generates stimuli of different magnitudes. The HVS is insensitive to stimulus below a certain dimension. In the next section, we use perceptual experiments to locate and verify the JND for mesh refinement.

From scale	To scale	Perceptual value
0	1	0.0410
0	2	0.0616
0	3	0.0677
0	4	0.0759
0	5	0.1080

Table 2: *An example of cumulated perceptual values.*

3. Perceptual Evaluation Experiments to estimate JND

Psychovisual experiments were conducted to establish some thresholds for human sensitivity [ODG*03], but they explore the factors that affect the perception of dynamic events, while we focus on relatively static objects. Our experiments were conducted through a user interface. In the initial set of experiments, we used an 8" x 11" monitor of resolution 768 x 1024 pixels. We used indoor incandescent lighting and 360° automatically rotating objects as visual stimuli. By using rotating objects, the judges were able to examine all silhouettes, which is more accurate than selecting a limited number of views [WFM01]. Since the goal is to evaluate the visual impact resulting from geometry change, the same texture was mapped onto both stimuli under comparison. Four sets of experiments were conducted. Randomly generated ellipsoids of different dimensions were used in the first two. Irregular quadrics were used in the third, and a 3D object was used in the fourth. We started with ellipsoids because 3D surfaces can be approximated by ellipsoids [KT96] defined by the polynomial equation with parameters *a*, *b* and *c*.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \qquad (2)$$

Each 3D object has its unique surface property and thus perceptual values. Not every perceptual range can be found in a 3D object. For initial estimation, it is easier to apply scaling factors on ellipsoids to narrow down the range where the JND is located.

3.1 Experiment 1 – An initial estimation of JND

In each test, a pair of ellipsoids (original and scaled versions) was displayed to a judge. The stimuli could be zoomed in and out, and rotated in any direction in a synchronized manner for examination. The original version was generated by randomly selecting the parameter set $\{a, b, c\}$. The scaled version is defined by the parameters $\{fa, fb, fc\}$ where f is a scaling factor in the range [0.7, 1.3]. One, two or none of the three parameters *a*, *b* and *c* were randomly exempted from scaling. The left and right positions were randomly assigned to the ellipsoids. We applied the two-alternative forced-choice (2AFC) strategy [Web05], and asked judges to choose the larger ellipsoid. After 34 tests with one judge, we eliminated the scaling factors which could be recognized correctly 100% of the time. We also eliminated the scaling factors for which judges relied on guessing (correctly judged approximately 50% of the time). The range was then refined to [0.9, 1.1].

3.2 Experiment 2 – Locating JND for regular ellipsoids

Experiment 1 was repeated within the refined range [0.9, 1.1]. Each correct or wrong answer was recorded under the ten sub-ranges $\Re_i$ ($i \in [1,10]$), corresponding to the set of values $0 < \chi_1 \leq 0.01, \ldots, 0.09 < \chi_{10} \leq 0.10$, with $f = 1 \pm \chi_i$. After 439 tests with two judges, for each sub-range k the percentage Ω_k for which the judge had chosen the correct ellipsoid was computed. It was noticed that in a low sub-range, Ω_k is also low, implying that it was more difficult to distinguish between the ellipsoids. Experimental results show that in the sub-range $\Re_4$, the judge could choose the correct answer 75% of the time. Thus, 0.04 was determined as the JND for discriminating ellipsoids.

3.3 Experiment 3 – JND for irregular quadric surfaces

In the virtual world, 3D objects are often more complex than a smooth ellipsoid. To verify the JND for more general 3D shapes, Experiment 2 was repeated but the ellipsoids were randomly distorted to generate irregular quadrics of random dimensions.

For each evaluation, a texture was selected randomly from six different patterns to avoid possible texture masking effect, but the same texture was mapped onto each pair of stimuli, and indoor incandescent lighting was used in the experimental environment. Prior knowledge and familiarity are compelling factors affecting how the HVS perceive. It is believed that the irregular quadrics would be more difficult than the regular ellipsoids for the HVS to discriminate. To accommodate this factor, a broader range of [0.8, 1.2] and 20 sub-ranges were used. One thousand tests were assigned to seven judges, and each judging session did not exceed 3 minutes to avoid fatigue. To ensure unbiased result, at least 30 tests were completed in each sub-range. The line of best fit was solved by regression. JND 0.10 (sub-range 10) was the threshold where the judgement was correct 75% of the time. The experimental points were fitted by a regression line instead of a psychometric curve because in the selected interval the function appears to be approximately linear [BKT86].

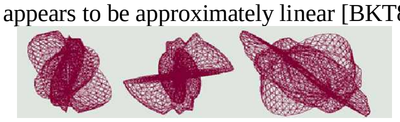

Fig. 4: *Examples of randomly generated irregular quadric.*

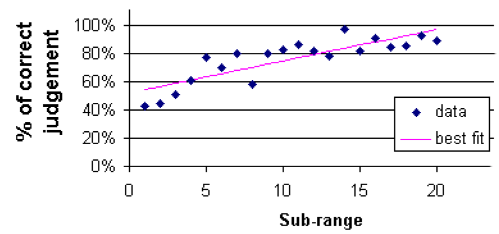

Fig. 5: *JND for quadrics based on data obtained from 1000 tests.*

Note that the JND is higher for irregular quadrics than regular ellipsoids. Since the appearance of 3D objects are close to quadrics then ellipsoids, 0.10 was used as the benchmark in Experiment 4 to evaluate the perceptual impact when refining the nutcracker object from a coarse to a denser version.

3.4 Experiment 4 – Verifying JND with 3D TexMesh

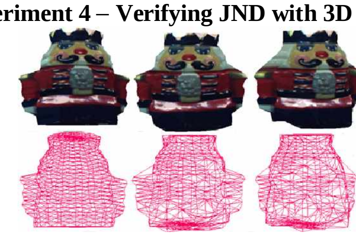

Fig. 6: *An example of different scales of the nutcracker object, S_0, S_6 and S_8 from left to right*.

In this experiment, we verified the JND by testing pairs of simplified meshes randomly selected from S_0 to S_{20} of the nutcracker object (Fig. 6). The original mesh S_0 was displayed as a reference in the upper part of the interface. Two stimuli were displayed side by side in the bottom part. We followed the 2AFC with reference strategy, and a judge was asked to decide which one (left or right) was a finer version closer to the original. The perceptual values in the LUT (explained in Section 3) were grouped into 10 sub-ranges. 361 tests were conducted by twenty judges on three monitors of different dimension and resolution, and the percentage of correct judgement in each sub-range was recorded. A JND of 0.096 (Fig. 7) was obtained by locating the 75% correct judgement, with a correlation coefficient of 95%. Note that 0.096 (which can be refined by increasing the number of tests) is slightly lower than the JND (0.10) obtained in Experiment 3, but is sufficient to show that our perceptual metric is consistent with the HVS.

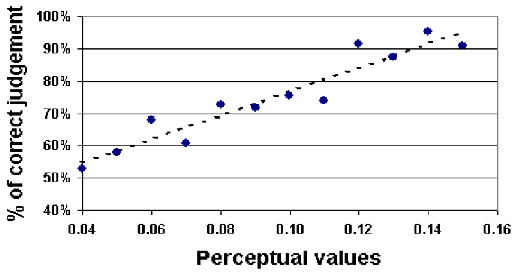

Fig. 7: *Verification of JND using the nutcracker object.*

4. Efficient Mesh Refinement

Based on the established JND (0.10), the scales of the nutcracker can be divided into tiers as shown in Fig. 8. For simplicity, we assume an application using 20 mesh scales in a distance range of 20 units. We define virtual distance as the distance between the object and the viewing platform in the virtual world. Instead of a linear relationship (pink Δ), the JND indicates that scales relate to virtual distance following a step function (blue ◊). For example, at distance unit 1, S_4 is used instead of S_1 because it requires a smaller number of vertices and has perceived similarity (Table 2). We define the scale where the pink and blue symbols meet as a major scale, and the others as minor scales. Only changes from one major scale to another adjacent major scale have significant impact on visual perception.

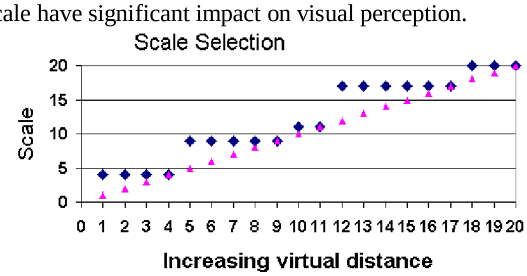

Fig. 8: *Perceptual function of the nutcracker object, relating scale to virtual distance.*

Since geometrically different objects can be perceptually similar [CB04], it is important during online transmission to suppress redundant mesh data, which do not improve visual quality. Major scales can be identified from the LUT during runtime to perform this task. For example, in Fig. 8 the major scales are 4, 9, 11, 17 and 20.

5. Conclusion and Future Work

In this paper, we proposed using perceptual value as a metric for efficient online visualization of 3D TexMesh. We used JND and Weber's fraction to evaluate the perceptual impact on the HVS resulting from changing mesh detail. Our approach is view-independent. Differing from previous techniques, which measure the spatial frequency generated by the stimulus affecting the visual field, our approach is independent of viewing distance. The novelty of our approach lies in integrating perceptual and geometric metrics to select scale in mesh refinement. In future work, we will use more 3D objects and monitors of different dimension and resolution in our evaluation experiments. Our second set of experiments using randomly selected 3D objects is already underway. We will also apply the JND computation model to other simplification techniques, and compare the efficiency of different mesh refinement approaches.

References

[AS04] M. Arguin and D. Saumier, "Independent processing of parts and of their spatial organization in complex visual objects," Psychological Science 2004, 15(9), p 629-633.

[BKT86] K. Boff, L. Kaufman and J. Thomas, "Handbook of perception and human performance," A Wiley-Interscience Publication pp 1-24.

[CB04] I. Cheng and P. Boulanger, "Perception of scale with distance in 3D visualization," Siggraph 2004 Poster.

[CB05] I. Cheng and P. Boulanger, "Feature Extraction on 3D TexMesh Using Scale-space Analysis and Perceptual Evaluation," to appear IEEE Transaction on Circuits and Systems for Video Technology Special Issue 2005.

[GH98] M. Garland and P. Heckbert, "Simplifying surfaces with color and texture using quadric error metrics", IEEE Visualization 98.

[GW02] R. Gonzalez and R. Woods, "Digital Image Processing," second edition, 2002 Prentice Hall.

[HH93] P. Hinker and C. Hansen, "Geometric optimization," Proc. Visualization 1993,.

[Hop96] H.Hoppe, "Progressive meshes," Proceedings of Siggraph 1996, Los Angeles, CA USA.

[KF01] A. Kuijper and L. Florack, "Logical filtering in scale space", Institute of Information and Computing Sciences, Utrecht University, Technical Report, 2001.

[KSS03] A. Khodakovsky, P. Schroder and W. Sweldens, "Progressive geometry compression," Siggraph 2000.

[KT96] A. Kalvin and R. Taylor, "Superfaces: polygonal mesh simplication with bounded error," IEEE Computer Graphics and Appl., 16(3), May 1996.

[LH01] D. Luebke and B. Hallen, "Perceptually driven simplification for interactive rendering", 12th Eurographics Workshop on Rendering Techniques 2001, London, UK.

[Lue02] D. Luebke et al., "Level of Detail for 3D Graphics", Morgan Kaufmann Publishers, 1st ed., 2002.

[ODG*03] C. O'Sullivan, J. Dingliana, T. Giang and M. Kaiser, "Evaluating the visual fidelity of physically based animations," ACM Trans. on Graphics Vol22 Issue3, Jul 2003.

[OHM*04] C. O'Sullivan, S. Howlett, R. McDonnell, Y. Morvan and K. O'Conor, "Perceptually adaptive graphics", Eurographics 2004.

[PCB05] Y. Pan, I. Cheng, and A. Basu, "Quantitative metric for estimating perceptual quality of 3D objects," IEEE Transactions on Multimedia, April 2005 vol. 7 No. 2.

[Red01] M. Reddy, "Perceptually optimized 3D graphics," IEEE Applied Perception, September/October 2001.

[RRP00] H. Rushmeier, B. Rogowitz and C. Piatko, "Perceptual issues in substituting texture for geometry," Proceeding of SPIE Vol.3935, p372-383.

[Web05] http://www.usd.edu/psyc301/WebersLaw.htm

[WFM01] B. Watson, A. Friedman and A. Mcgaffey, "Measuring and predicting visual fidelity", Siggraph 2001.

[Wit83] A. Witkin, "Scale-space filtering," International Joint Conference on AI, 1983.

[WLC*03] N. Williams, D. Luebke, J. Cohen, M. Kelly and B. Schubert, "Perceptually guided simplification of lit, textured meshes," ACM Symp. on Int. 3D Graphics, 2003.

[ZN99] M. Zerroug and R. Nevatia, "Part-based 3D descriptions of complex objects from a single image," IEEE Trans. on PAMI, Vol 21, no. 9, Sep. 1999.