Investigating the Rate-Distortion Performance of a Wavelet-Based Mesh Compression Algorithm by Perceptual and Geometric Distortion Metrics

Maja Krivokuća[1] Burkhard Wuensche[2] Waleed Abdulla[3] Guillaume Lavoué[4]
[1,2,3] The University of Auckland, New Zealand
[4] Université de Lyon, France
[1] mkri012@aucklanduni.ac.nz [2] b.wuensche@auckland.ac.nz [3] w.abdulla@auckland.ac.nz
[4] glavoue@liris.cnrs.fr

ABSTRACT

The rate-distortion performance of wavelet-based mesh compression algorithms is usually only evaluated in a purely geometric sense, by measures such as the Root Mean Square Error and the Hausdorff distance, which do not capture the human visual perception of distortion. This lack of quantitative information about the perceptual effects of wavelet compression has prompted us to present a more complete evaluation of a classic wavelet-based mesh compression algorithm, by measuring its rate-distortion performance both with geometric metrics (the Hausdorff distance and Root Mean Square Error) and perceptual metrics (the recently introduced Mesh Structural Distortion Measure 2 (MSDM2) and the Mean Opinion Scores (MOS) that we obtained by conducting a subjective experiment with human observers), where the rate is measured as the percentage of wavelet coefficients used in reconstruction. The MSDM2 has already been proven to outperform other existing perceptual metrics for several different types of distortions, but this is the first time that it has been tested for this type of geometric distortion in a real-use case scenario. We found that, in this context, the MSDM2 generally correlates well with the MOS but seems to under-estimate the perceptual error in cases of low-frequency (large scale) shape distortion. Due to the disparities in the distortion values produced by the tested distortion metrics, we also conclude that a complete evaluation of any mesh compression algorithm should include several different distortion metrics, to allow the developers and users of these compression algorithms to make more informed decisions about the applicability of those algorithms in different application areas.

Keywords

Distortion metrics, quality metrics, wavelets, mesh compression, rate-distortion evaluation.

1. INTRODUCTION

Since the influential work of Lounsbery [Lou94], who introduced the notion of multiresolution analysis on surfaces, wavelet-based mesh compression has been a topic of much interest in the research community. This is reflected by the variety of wavelet-based mesh compression algorithms proposed in the literature, notably [HP05, KG02, KSS00, SS95, VP04]. The main idea in wavelet-based mesh compression is to decompose a high-resolution input mesh into a coarse representation called a *base mesh* and a set of detail coefficients termed *wavelet coefficients,* which can be used to refine the base mesh at multiple levels of detail [Lou94]. Geometry compression may then be obtained either by discarding small (unimportant) wavelet coefficients at each resolution level, and/or by quantizing and entropy coding the remaining coefficients that are to be transmitted. The aim in wavelet-based mesh compression is to optimise the trade-off between data size and approximation accuracy, which is measured by the *rate-distortion* (*R-D*) curve. The *rate* refers to the amount of information transmitted in a compressed mesh, and the *distortion* refers to a quantified measure of difference between the reconstructed and original meshes. While the rate is a relatively straightforward measure to quantify (usually represented as the number of bits transmitted per vertex, or the number of wavelet coefficients transmitted), the term *distortion* still lacks a formal definition. Without such a definition, it is difficult to claim that any lossy mesh compression algorithm developed to date (including wavelet-based algorithms) has been fully evaluated. While wavelet-based mesh compression has shown some very promising results and is still an active area of research, the performance of existing algorithms has usually only been reported based on a single error metric. In addition, the distortion metrics

that are currently used to evaluate these algorithms are normally purely geometric measures, such as the Root Mean Square Error and the Hausdorff Distance, which are not designed to capture the *visual* disparity between two 3D models. Due to the previous lack of availability of an objective perceptual distortion metric, the visual distortion caused by discarding wavelet coefficients in a wavelet-based mesh compression system has not yet been documented in a quantitative manner. However, this information is important for many graphics applications where the ultimate judge of the transmitted model is a human. The surge in recent years to design perceptually-based distortion metrics (a survey and extensive comparison can be found in [LC10]) has produced some promising perceptual metrics, particularly the *Mesh Structural Distortion Measure 2* (*MSDM2*) recently introduced by Lavoué [Lav11]. While the *MSDM2* has been proven to outperform other existing perceptual metrics for several different types of classical distortions [LC10, Lav11], it has not been tested in a real-use case and has not been investigated for the type of geometric distortions that result from discarding wavelet coefficients in a wavelet-based mesh compression system.

The objective of this paper, therefore, is twofold: (1) to evaluate how the *MSDM2* metric compares to human perception of distortion, in the real-use case of wavelet-based mesh compression, where the distortion is caused by discarding different percentages of wavelet coefficients in mesh reconstruction; and (2) to offer a more complete evaluation of the effects of discarding wavelet coefficients in a wavelet-based mesh compression system. The latter is achieved by measuring the rate-distortion performance of a classic wavelet mesh compression algorithm with several different error metrics - two commonly used geometric measures (the Hausdorff distance (d_H) and the Root Mean Square Error (*RMSE*)) and two perceptual metrics (the subjective Mean Opinion Score (*MOS*) from a group of human observers, and the objective visual distortion metric, *MSDM2*). In this way, we also aim to demonstrate that, due to the disparities that exist between the performance results generated by these distortion metrics, a complete evaluation of a lossy mesh compression algorithm is not possible with only one distortion metric. We use our investigation of the wavelet compression technique as a case study to suggest how an appropriate error metric may be chosen for evaluating a lossy mesh compression algorithm based on different application needs.

Section 2 of this paper introduces the wavelet-based compression method that we have implemented and describes our evaluation procedure, Section 3 discusses the basic concepts behind the distortion metrics that we have investigated, Section 4 describes the perceptual distortion test that we carried out, Section 5 presents our results and discusses the insights gained in relation to our objectives, and the conclusion ties up the key messages of this paper.

2. WAVELET COMPRESSION CASE STUDY

The progressive compression algorithm that we have investigated is an own implementation of the fundamental Lounsbery *subdivision wavelet* method [Lou94]. We only cover here the basic concepts of this method that are necessary for an understanding of our investigation. For further details, we refer the interested reader to the original thesis [Lou94].

2.1 Background and Implementation

The subdivision wavelet method applies the notion of multiresolution analysis to subdivision surfaces. The basic idea is that we can take a high-resolution input mesh with subdivision connectivity and decompose it into a lower-resolution mesh, together with a set of detail coefficients that can be added to the lower-resolution mesh to reconstruct the higher-resolution mesh. This decomposition, or *analysis*, is done by two separate filtering operations on the original mesh: a low-pass filtering where we compute weighted averages of the vertices in the higher-resolution mesh to obtain the (sparser) set of vertices in the resulting lower-resolution mesh (we call this lower-resolution mesh an *approximation* of the higher-resolution mesh that it was obtained from), and a high-pass filtering where we compute weighted differences of the higher-resolution mesh vertices to obtain the detail coefficients, called *wavelet coefficients*. If we continue these filtering operations on each successive, lower-resolution mesh, we eventually obtain the coarsest possible mesh (called the *base mesh*), together with wavelet coefficients at multiple levels of detail. If we add back the wavelet coefficients to the corresponding mesh at each level, we can progressively refine the base mesh and thereby obtain a multiresolution representation of the original mesh. *Figure 1* illustrates a simple example of this process, where the mesh on the left is the input mesh, the mesh on the right is the base mesh, and the middle mesh is the model at an intermediate resolution level.

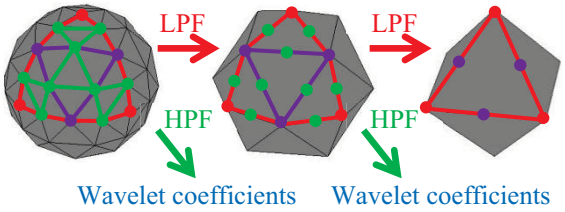

Figure 1: Illustration of wavelet decomposition. LPF=low-pass filter, HPF=high-pass filter. The reconstruction process proceeds in reverse. The purple and green circles represent vertices added at the midpoints of edges in the refinement process.

The reconstruction, or *synthesis*, process involves two more filtering operations: a refining operation where each triangular face of a lower-resolution mesh is subdivided into four sub-triangles by introducing new vertices at edge midpoints, and a perturbing operation where the new vertices are perturbed (displaced) to their new positions according to the wavelet coefficients. *Figure 1* highlights the refinement process of one face of the base mesh (outlined in red), by showing how the new vertices (purple) are added to the midpoints of that face and are connected together in the next higher resolution level (the middle mesh), in order to refine the one base mesh face into four smaller sub-faces. Similarly, in the middle mesh in *Figure 1*, the next set of vertices (green) is added at the midpoints of the edges of the new faces, and these vertices are connected together to produce 16 sub-faces in the next higher resolution level (leftmost mesh). Each new set of vertices is, in this case, projected onto the surface of a sphere (perturbing operation), which is how we get from an octahedron base mesh with flat faces (rightmost mesh in *Figure 1*) to an n-sided figure (leftmost mesh in *Figure 1*) that begins to approximate the shape of a sphere. Since the connectivity of the mesh at each resolution level is obtained by a common refinement process, we only need to transmit the base mesh and the set of wavelet coefficients in order to reconstruct the original mesh geometry. Due to the orthogonality property of the wavelets (see [Lou94] for an in-depth explanation), the mesh approximation at each resolution level is guaranteed to be the approximation at that level that is the closest to the next, finer-level mesh in a least-squares sense. The orthogonality property also means that the wavelet coefficients are a good indicator of where the approximation is not close to the finer-level mesh that it is approximating. For example, if a wavelet coefficient is zero, this means that the approximation is locally perfect as it indicates that there is no information (detail) missing in that region. If the wavelet coefficient is large, however, then it indicates that there is a large amount of information missing at that position and scale. Since the approximation at each resolution level is guaranteed to be the best possible approximation for that level, however, most of the wavelet coefficients are usually distributed around zero (not much information missing), so we can discard many of the (small) coefficients (thereby achieving compression) and still be able to reconstruct a close approximation of the original mesh.

In our implementation, we order the coefficients by magnitude and then select the largest wavelet coefficients at each level. The percentage of chosen coefficients at each level constitutes our *rate* of transmission. The wavelet coefficients are deliberately not quantized or entropy coded, in order to isolate the distortion effects (and compression) that are caused solely by eliminating wavelet coefficients. Our implementation also incorporates a *naïve* reconstruction, where every location on the mesh is refined in a uniform manner (every face is split into 4 sub-faces), regardless of whether or not additional detail is ever added to that location. This means that the reconstructed mesh always has the same number of triangles as the input mesh, but the quality of geometry reconstruction (the accuracy of the final vertex positions) depends on the number of wavelet coefficients that we use in the reconstruction.

2.2 Experimental Procedure

We tested the *rate-distortion* performance of the subdivision wavelet compression method on six different meshes with subdivision connectivity - shown in *Figure 2*, along with their associated base meshes. The base meshes are presented in faceted form, to demonstrate their different levels of complexity.

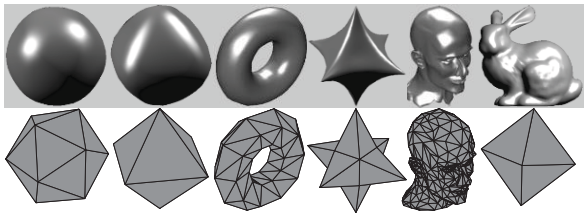

Figure 2: The subdivision models used for testing, and their associated base meshes (underneath each model). From left: Sphere (res. level 4), Rounded Octahedron (res. level 4), Torus (res. level 3), Star (res. level 4), Mannequin (res. level 3), and Bunny (res. level 6).

2.2.1 Choice and Preparation of Test Models

The models were chosen based on the amount of complexity, in terms of visual detail, that they possessed. We wished to investigate a range of models, from those that have very smooth surfaces (e.g., the Sphere) to those with quite a large amount of visual detail on the surface (e.g., Mannequin and Bunny). We also chose the models so that they would possess different *types* of details (e.g., smooth corners in the Rounded Octahedron, sharp points in the Star, sharp curves on the Bunny's ears and thighs) because we wished to see how the wavelet method would handle these different features as measured by the different distortion metrics. The associated base mesh for each model was chosen as the simplest (lowest-resolution) version of that mesh available in the model library. Each base mesh is therefore considered resolution level 1 in our experiments, and the resolution level for the corresponding input model is always considered relative to its base mesh. The resolution levels for each of the input models, relative to their base meshes, are stated in the caption of *Figure 2*.

2.2.2 Rate-Distortion Testing

Each of these test meshes was first decomposed (similar to *Figure 1*) to obtain its corresponding base mesh and wavelet coefficients. Then the reconstruction process was carried out at 11 different *rates*, where the *rate* is defined as the percentage of largest wavelet coefficients used at each resolution level, from 0% to 100% coefficients inclusive, in 10% steps. The mesh *distortion* at each of these rates was then measured separately with four different error metrics, which are discussed below.

3. STUDIED DISTORTION METRICS

We selected two traditional geometric error metrics - the Hausdorff distance (d_H) and the Root Mean Square Error (*RMSE*) - and two perceptual error metrics - the *MSDM2* metric and the subjective Mean Opinion Scores (*MOS*) obtained through a subjective experiment. The basic concepts behind each metric are discussed below.

3.1 Hausdorff Distance (d_H)

The Hausdorff distance represents the maximum distance between a point on one mesh and the surface of another mesh. It is formulated as follows:

The distance, $e(p,S)$, between a point p in 3D space and a mesh surface S, is defined as:

$$e(p,S) = \min_{p' \in S} d(p,p')$$

where $d(p,p')$ is the Euclidean distance between points p and p', and p' is a point on surface S. Then the one-sided (asymmetric) distance between two surfaces (or meshes), S_1 and S_2, is defined as:

$$E(S_1,S_2) = \max_{p \in S_1} e(p,S_2).$$

A two-sided (symmetric) distance, which is termed the Hausdorff distance, d_H, is defined as the maximum of $E(S_1,S_2)$ and $E(S_2,S_1)$:

$$d_H(S_1,S_2) = \max\{E(S_1,S_2), E(S_2,S_1)\}.$$

We used the *Metro* tool [CRS98] with the default settings, to compute the symmetric Hausdorff distance in all our experiments. The obtained d_H values were then normalized to fit into the range [0,1], to enable a comparison between the rate-distortion trends obtained from the different error metrics. The normalized Hausdorff distance, $d_H N^i$ for distorted model *i,* was computed as:

$$d_H N^i = \frac{(Current\ {d_H}^i - Minimum\ {d_H}^M)}{(Maximum\ {d_H}^M - Minimum\ {d_H}^M)}$$

Where *Current* ${d_H}^i$ is the current (un-normalized) d_H value for mesh *i*, and *Minimum* ${d_H}^M$ and *Maximum* ${d_H}^M$ are, respectively, the smallest and largest Hausdorff values produced for any of the distorted versions of mesh *M*.

3.2 Root Mean Square Error (RMSE)

The Root Mean Square Error *(RMSE)* measures how far, on average, the difference between the original and reconstructed vertex positions is from 0. The *RMSE* between two meshes, S_1 and S_2, is computed as:

$$RMSE(S_1,S_2) = \sqrt{\frac{\sum_{i=1}^{n}(v_i^{S_1} - v_i^{S_2})^2}{n}},$$

where n is the number of vertices in the meshes (both meshes must have the same number of vertices), and $v_i^{S_2}$ is the vertex in mesh S_2 corresponding to vertex $v_i^{S_1}$ in mesh S_1.

The *RMSE* was also normalized in our experiments to fit into the range [0,1], in a similar way to the Hausdorff normalization.

3.3 Mesh Structural Distortion Measure 2 (MSDM2)

The Mesh Structural Distortion Measure 2 (*MSDM2*), descendant of the earlier Mesh Structural Distortion Measure (*MSDM*) [LGDBE06], was recently introduced by Lavoué [Lav11] as a multiscale metric for objective visual quality assessment of a 3D mesh. The *MSDM2* is based on the 2D image metric, *SSIM* (Structural SIMilarity index), from Zhou et al. [ZBSS04], and works by measuring the differences in curvature statistics between two meshes, which are computed on local corresponding spherical neighbourhoods. These neighbourhoods vary in size according to the *scale*, h, which is related to the maximum length of the bounding box of the model. For example, for 3 scales, $h_i \in \{2\varepsilon, 3\varepsilon, 4\varepsilon\}$, where $\varepsilon = 0.5\%$ of the maximum length of the bounding box.

The symmetric *MSDM2* measure between a reference mesh, M_r, and a distorted mesh, M_d, is computed as the average of the two asymmetric *global multiscale distortion* (*GMD*) measures, $GMD_{M_d \to M_r}$ and $GMD_{M_r \to M_d}$, where $GMD_{M_d \to M_r}$ is defined as:

$$GMD_{M_d \to M_r} = \left(\frac{1}{|M_d|} \sum_{v \in M_d} MLD(v)^3 \right)^{1/3}.$$

$MLD(v)$ is the *multiscale local distortion* measure defined as:

$$MLD(v) = \frac{\sum_{i=1}^{n} LD^{h_i}(v)}{n}$$

which is the average of *local distortion* (*LD*) values at single scales. The *LD* at a given scale *h* is defined for each vertex *v* from M_d as:

$$LD^h(v) = \frac{\alpha L^h(v) + \beta C^h(v) + \gamma S^h(v)}{\alpha + \beta + \gamma}$$

where α, β and γ are set to 1, 1 and 0.5, respectively, and $L^h(v)$, $C^h(v)$ and $S^h(v)$ are, respectively, the 3D mesh equivalents of the luminance comparison function, the contrast comparison function, and the structure comparison function defined for images in [ZBSS04]. Lavoué defines these functions for meshes in [Lav11].

In our experiments, we obtained the symmetric *MSDM2* values and the associated distortion maps from the MEPP platform [LTD12], using 3 scales. These values are in the range [0,1], where 0 indicates that the reconstructed mesh is identical to the original, and values closer to 1 correspond to increasingly larger visual differences between the two meshes.

3.4 Mean Opinion Score (MOS)

The Mean Opinion Score (*MOS*) is a subjective measure of distortion, where a group of human observers is asked to give a score to some distorted objects, which reflects the observer's degree of perceived distortion on these objects, in relation to an original, undistorted object. The *MOS* for a distorted model, *i*, is computed as:

$$MOS_i = \frac{1}{n}\sum_{j=1}^{n} m_{ij}$$

where MOS_i is the mean opinion score of the i^{th} distorted model, n is the number of test observers, and m_{ij} is the distortion score given by the j^{th} observer to the i^{th} model.

The next section describes the subjective experiment that we conducted to obtain these *MOS* values.

4. SUBJECTIVE EXPERIMENT

Due to a lack of information in the literature about the perceptual effects of discarding different percentages of wavelet coefficients, and from a desire to evaluate the *MSDM2* in this context, we carried out our own perceptual distortion test on three of our test models. The details of this test are explained below.

4.1 Assessment Procedure

A group of 13 human observers were shown 3 different 3D models (the Torus, Star and Bunny – see *Figure 2*), where each original model was printed on paper together with its 11 distorted versions (reconstructions with 0%-100% wavelet coefficients), and the printouts were given to each individual observer. The original model was labelled, but the 11 distorted versions were arranged in random order around the original. The observers were asked to give a distortion score between 0 and 10 to each distorted model (using whole numbers only), which would reflect the degree of perceived distortion on this model in relation to the original. Participants were told that a score of 0 meant that, in their opinion, the distorted model was identical to the original (i.e., they could not perceive any distortion on this model) and a score of 10 was the worst case scenario (i.e., the distorted model was "very different" to the original). Participants were *not* told to assign a 10 to the worst model and then assign lower scores to all the other models; in fact, the observers were told that if they felt that no model "deserved" a 10, for example, they did not need to use the full range of distortion scores. The observers were also told to consider the distorted models together and give them relative scores. They were further asked, for all the models to which they gave a score greater than zero, to circle the area(s) on those models that they thought were the "worst distorted" areas. There was no specific time limit for the test, but the task took approximately 20 minutes for the whole group.

4.1.1 Choice and Preparation of Test Models

The Torus, Star and Bunny were chosen for the subjective experiment because these models have a range of interesting surface details: the smooth curves on the Torus, the sharp points on the Star, and the curves of varying degrees of sharpness on the Bunny. We wished to investigate how the *MSDM2* compares with the *MOS* for capturing such different types of surface detail. The viewing angle for each model (same viewing angle as in *Figure 2*) was chosen so as to portray that model in what we subjectively judged to be both its most discriminative angle and the angle that offered the most familiar viewpoint of the object. The lighting and surface reflectance conditions were also chosen specifically for each model, based on the selected viewpoint, so as to produce what we believed was the best visibility of the shape and surface detail of each model, for the test observers. For all three models we used interpolated shading and *Phong* face lighting, but the strengths of diffusion, ambient lighting and specular highlights were chosen individually for each model so that, for the chosen viewpoint, no surface details (or, as little as possible) would be hidden by the specular highlights or shadows. The printed models were made to be of similar size, so that 6 models could fit on one side of an A4 sheet of paper in landscape orientation, organized side by side in two rows and three columns. While this use of a fixed viewpoint and fixed viewing distance for the test models does limit the generality of the obtained results, this step was an attempt at reducing the amount of variance between the test subjects. If the observers were free to zoom in and out of the models and rotate them, it would have been difficult to ensure that all observers saw the same parts of each model; perhaps some people would have used more viewpoints than other people, or a wider range of viewpoints, to make their decision. Having a fixed

viewing angle and distance ensured that this type of variation between the subjects was eliminated and so, since everyone was looking at exactly the same images, this made it easier to isolate the features of each model which were responsible for the different distortion scores given. This was ultimately the goal of the subjective experiment: to determine how the *MOS* compares to the *MSDM2* for different models with different types of features, distorted in the same way.

4.2 Normalizing the Distortion Scores

Before computing the *MOS* for each distorted model, the individual distortion scores were normalized to be in the range [0,1]. This was simply done by dividing each score by 10. Since the observers were instructed to only use the entire available scoring range (0-10) if they actually saw the need for it (i.e., if they could see enough difference between the different levels of distortion to require them to use the entire scoring range), they did not necessarily have to assign a 10 to the worst model and a 0 to the best model. Rather than scaling all the scores to fit into the 0-10 range, we were interested in the differences in the actual range of values that would be assigned for each model (a discussion of this is provided with *Table 1*, opposite). For this reason, it was not appropriate in our experiment to correct for the differences in gain and offset among the observers (as was done in [LC10], for example). The suitability of our experimental protocol was assessed by computing three Intraclass Correlation Coefficients (ICC), each of which measures the variation between the different observers in their subjective ratings of all the distorted versions of one of three test models (Torus, Star, and Bunny), respectively. The ICCs were computed from a two-way ANOVA test with no repetitions, and using the ICC equation related to *Model 2* in [SF79]. The ICC value for the Torus models was computed as 0.81, for the Star models as 0.79, and for the Bunny models as 0.94. Since ICC values close to 0 indicate poor agreement while values close to 1 indicate almost perfect agreement, the computed values show that there was strong agreement between the observers on the distortion scores they assigned to the Torus and Star models, and almost perfect agreement on the distortion values assigned to the Bunny models. These high levels of agreement indicate that our subjective experiment protocol was correct as it produced meaningful, consistent scores from the test observers.

5. RESULTS AND ANALYSIS

The rate-distortion curves for all six models were plotted and are displayed in *Figure* 3.

5.1 Variability between the Observers

While the ICC values reported above suggest strong agreement between the overall subjective distortion scores given to all three test models, looking at the distortion scores for the models at each rate of wavelet coefficients individually leads to some interesting observations. In particular, for the Torus model, only 4 out of 13 observers used the full distortion scoring range (0-10), for the Star 5 out of 13 used the full range, and for the Bunny 9 out of 13 used the full range. *Table 1* shows the distortion scores provided by the 13 observers for the Torus, Star and Bunny, at a rate of 0% wavelet coefficients, which is where the highest distortion scores were given.

Observer	Distortion score for **Torus**	Distortion score for **Star**	Distortion score for **Bunny**
1	6	7	10
2	10	10	10
3	10	10	10
4	8	7	10
5	7	8	10
6	8	8	10
7	8	9	9
8	5	8	10
9	7	6	10
10	10	10	10
11	10	10	10
12	10	10	10
13	9	10	10

Table 1: Subjective distortion scores for the Torus, Star and Bunny models reconstructed with 0% wavelet coefficients.

At a rate of 0% wavelet coefficients, the mesh reconstruction just produces the base mesh *shape* (but with the same number of triangles as the input mesh, due to the naïve reconstruction). Since the Bunny's base mesh is perceptually much more different to the input Bunny (see *Figure 2*) than the Torus and Star base meshes are to their respective originals, this explains why for a rate of 0%, the Bunny received the highest possible distortion score (10) from 12 out of 13 observers, whereas the Torus at this rate received a 10 from only 5 observers and the Star from 6 observers. This observation indicates that the choice of base mesh is a critical factor in the perceptual quality of the reconstructed mesh, especially at low rates of wavelet coefficients. Furthermore, the table above shows that the range of distortion scores for the Torus model at a rate of 0% coefficients is the largest (10-5=5), the Bunny the smallest (10-9=1), and the Star in between the two (10-6=4). A similar relationship holds for other low rates – from 10% to around 40% wavelet coefficients – but this is not shown in *Table 1*. The wider ranges for the Torus and Star indicate that the observers

found it more difficult to decide on appropriate distortion scores for these models, especially at the lower rates of wavelet coefficients. The greater agreement between the scores for the Bunny in general (indicated by both the smaller range and the higher ICC value) implies that the different levels of distortion were perceptually more obvious on the Bunny than on the Torus or Star. In fact, when the 13 observers were asked which model they found it the easiest and hardest to judge different levels of distortions on, the Bunny was almost unanimously selected (by 12/13 observers) as the easiest, and the Torus and Star were both said to be equally difficult. The observers felt that this was because the Bunny is quite a familiar, natural object, and so it was easy to tell when the model looked 'wrong'.

5.2 Comparison of MOS and MSDM2

We were able to make several important comparisons between the *MOS* and *MSDM2* distortion results for each tested model.

5.2.1 Bunny

The most significant difference between the *MOS* and the *MSDM2* R-D curves for all three tested models is that the *MSDM2* has a more stable, almost linear rate of decrease with increasing percentages of wavelet coefficients, whereas the *MOS* curves are more irregular. The most obvious example of this is in the Bunny R-D plots, where the *MOS* curve indicates a sharp drop in perceived distortion between 40% and 50% wavelet coefficients, but the *MSDM2* curve does not reflect this as a large change. Indeed, the *MSDM2* curve seems to decrease at a nearly constant (slower) rate, from 0% right up to around 70% wavelet coefficients. If we compare the Bunny reconstructions with 40% and 50% wavelet coefficients (see *Figure 4*) and compare these to the original Bunny model (see *Figure 2*), we notice that the Bunny at 50% looks much more similar to the original model than the Bunny at 40% does. Perceptually, the worst distortion in the 40% model (as judged by the majority of the test observers) is the area circled in *Figure 4*, and this distortion is not present in the 50% model. The reason why the *MSDM2* does not perceive this as a large difference between the two models may be because the *MSDM2* is designed to capture differences in curvature, but the circled area in the 40% Bunny and the corresponding area in the 50% Bunny have almost the same curvatures (i.e., the Bunny 'bulges out' in nearly the same places). The perceptual difference, as reflected by the *MOS*, is purely geometric – the circled area bulges out much further in the 40% model than in the 50% model, which makes the Bunny look out of proportion and unnatural, whereas the 50% Bunny looks much closer to what we might imagine a bunny to look like and it is much closer, visually, to the original Bunny model.

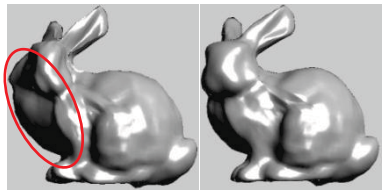

Figure 4: Bunny reconstructed with 40% wavelet coefficients (left) and 50% coefficients (right).

Interestingly, the opposite result occurs for the Bunny at 50%-60% wavelet coefficients: the *MOS* curve in this range changes very little (it is nearly flat), while the *MSDM2* has a sharper rate of decrease. Looking at *Figure 5,* which shows the reconstructed Bunny models with 50% and 60% wavelet coefficients, we can see that the most noticeable difference between them is the area circled in red (almost all observers circled this area as looking the most distorted). The *MSDM2* indicates this as a large distortion (see the area circled in red in the *MSDM2* distortion map in *Figure 5*, where "warmer colours" indicate higher distortion [Lav11]). This is because this 'bump' consists of rather a sharp curve compared to the smooth corresponding area in the 60% model (notice the corresponding area in the *MSDM2* distortion map for the 60% model, which indicates a considerably smaller error). However, the human observers did not perceive the removal of this bump in the 60% model as a very significant improvement, presumably because this was only a small 'glitch' on an otherwise good-looking bunny.

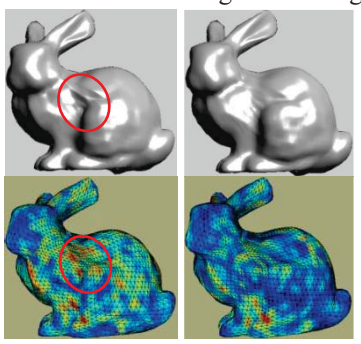

Figure 5: Bunny reconstructed with 50% wavelet coefficients (top left) and 60% wavelet coefficients (top right). Corresponding *MSDM2* distortion maps are beneath each model.

5.2.2 Star

In the case of the Star model, the *MOS* curve seems to follow the *MSDM2* curve more closely than in the Bunny's case. A good reason for this might be that the perceptual distortions in the Star are mainly due to the changes in curvature, which is what the *MSDM2* is designed to capture. For example, comparing the Star models reconstructed with 10%, 20% and 50% wavelet coefficients (see *Figure 6*), it appears that the perceived distortions on all these models are due to the differences in the sharpness of the Star's points and in the concavity of the Star's

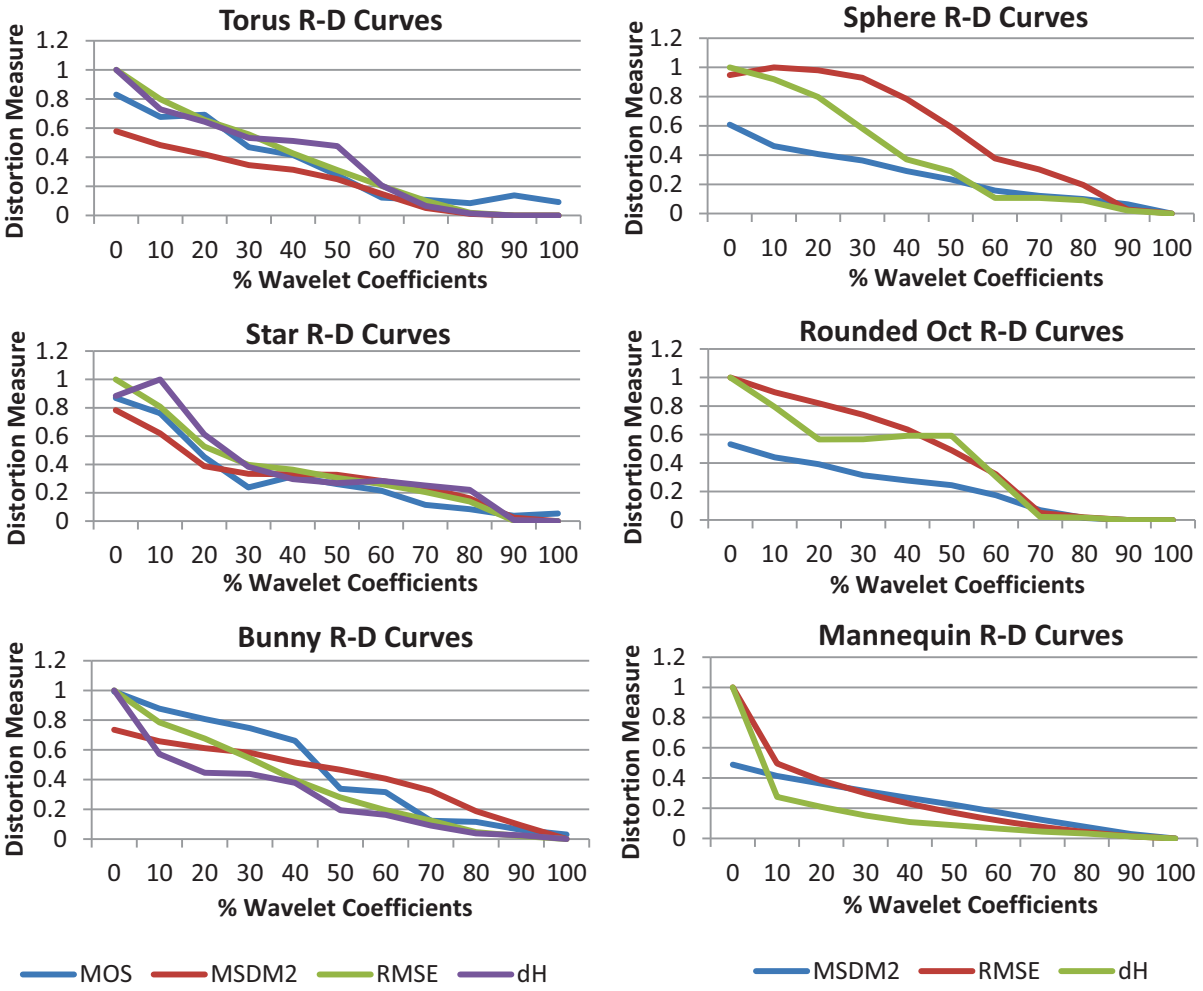

Figure 3: (Left) Rate-distortion curves for the 3 models used in the subjective test, and (Right) Rate-distortion curves for the remaining 3 models.

surface (circled on the 20% model, below, as an indication). These are the areas that the 13 observers circled as having the worst (most noticeable) distortion on most of the Star models. The corresponding *MSDM2* distortion maps indicate that the worst *MSDM2* distortions are in the same regions.

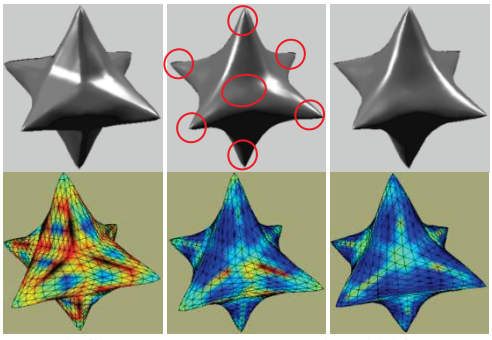

Figure 6: Star reconstructed with 10% wavelet coefficients (top left), 20% coefficients (top middle), and 50% coefficients (top right). Corresponding *MSDM2* distortion maps are beneath each model. Red circles indicate areas of worst distortion, as judged by the majority of the test observers.

5.2.3 Torus

The Torus model seems to have the largest disparity out of the three models tested, between the *MOS* and *MSDM2* rate-distortion curves, especially at the lower rates. A reason for this might be that, because the perceptual distortions in the Torus usually manifested themselves as small 'bumps' on the surface (for example, see the Torus reconstructed with 30% wavelet coefficients, in *Figure 7*), the *MSDM2* does not perceive these as very large differences in curvature to the original model, as these distortions are, in fact, not very sharp curves. This can be seen in the *MSDM2* distortion map in *Figure 7*, where the areas of the Torus with the worst perceived distortion (circled in red) do correspond to the areas of highest distortion in the *MSDM2* colour map, but these colours are mostly still in the lower (blue-green) range of distortion values, which, compared with the bright reds in the distortion map for the 10% Star model in *Figure 6,* are not very high. However, because the human eye expects a smooth, uniform shape, any small 'bumps' on the surface are readily visible and annoying as they

interrupt the smooth flow of the surface. This observation was confirmed in the perceptual distortion test, where people seemed so aware of the little bumps that they even perceived the Torus models reconstructed with 100% wavelet coefficients as being different to the original (the two models are, in fact, geometrically identical). Only 5 observers out of 13 gave the Torus model at 100% a distortion score of 0, compared to the 10/13 zeros which were given to the 100% Star model and the 9/13 zeros which were given to the 100% Bunny model.

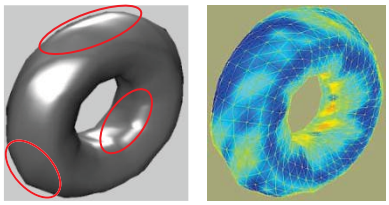

Figure 7: Torus reconstructed with 30% wavelet coefficients (left) and the corresponding *MSDM2* distortion map (right). Red circles indicate areas of worst distortion on this model, as judged by the majority of the test observers.

5.2.4 Overall Performance of MSDM2 vs MOS

Even though there are disparities between the *MOS* and *MSDM2* rate-distortion curves, the locations of various distortions on our test models and the general levels of distortion have been captured well by the *MSDM2*, compared to the human perception of these distortions. Evidence of this can be seen in *Figures 5, 6* and *7*, where the human-circled areas of "worst" distortion correspond to the highest error values on the *MSDM2* distortion maps, and in the R-D curves in *Figure 3*, where the *MSDM2* values, like the *MOS* values, gradually decrease as the percentage of wavelet coefficients increases.

In general, the largest differences between the *MOS* and *MSDM2* rate-distortion curves for all the three test models in our experiments seemed to occur at the lower rates (smaller percentages of wavelet coefficients). Because the largest wavelet coefficients are responsible for reconstructing the overall, global shape of a model, and the smaller wavelet coefficients are used to reconstruct the details, these results seem to suggest that the human eye is more critical of large-scale distortions (which affect the overall shape of a model) than the *MSDM2* predicts. On a low-detail model like the Torus, the human eye notices differences in local distortions more easily. This is to be expected, as this detail is not masked on a smooth surface whereas it might be masked on a highly-detailed surface. This last point agrees with the observations of the *visual masking effect* demonstrated by Lavoué [Lav11].

The disparities between the *MOS* and *MSDM2* for different models are an indication that human observers use more visual cues than just curvature differences to perceive visual distortion. For the type of distortion that we investigated, the removal of low- or medium-frequency wavelet coefficients produces large-scale localized distortions that have a high impact on the visual quality (see the left Bunny in *Figure 4*). However, the *MSDM2* metric fails to detect them; it underestimates these degradations, especially on very smooth, low-detail models like the Torus.

5.3 Using a Combination of Distortion Metrics for R-D Performance Evaluation

The performance of a wavelet-based mesh compression algorithm has not previously been reported with a combination of geometric and perceptual error metrics, and our investigation resulted in four main observations:

1. Both the geometric and perceptual distortion values decrease with increasing percentages of wavelet coefficients. This confirms that both the geometric and visual quality of a mesh improve with a greater number of wavelet coefficients.
2. The d_H and *RMSE* curves are quite closely correlated with the *MOS* curves for the Torus and Star models, but for the Bunny model the d_H drops more rapidly than the *MOS* at low rates. This suggests that large-scale distortions on detailed models like the Bunny affect perceptual quality more than geometric quality.
3. Across all six test models, the d_H curves generally seem less stable than the *RMSE* curves. This shows that the quality of maximum error induced by the surface reconstruction (measured by the d_H) is not always proportional to the average quality of vertex reconstruction (measured by the *RMSE*) as the number of wavelet coefficients increases.
4. The R-D performance of a wavelet-based mesh compression algorithm, and the suitability of a distortion metric to measure this performance, is dependent on the 3D model used as input.

More generally, we are able to conclude that, due to the disparities in distortion measurements produced by the different error metrics, it is not sufficient to evaluate the rate-distortion performance of a lossy mesh compression algorithm, such as the wavelet method, with a single error metric. The use of multiple distortion metrics for evaluation would benefit both the developers and users of mesh compression algorithms. The developers would benefit as it would be easier to accurately compare the performance of different algorithms, and the users would benefit because it would be easier to select the right compression method based on their application needs. For example, users that are only concerned with the *look* of the reconstructed model may choose a compression method with the best

MSDM2 or *MOS* performance, users that desire a close surface reconstruction (geometrically) within a certain tolerance regarding the original surface may be interested in the Hausdorff performance, while users that require the mesh vertices to be reconstructed exactly may consider the *RMSE* values. Our evaluation of the classic wavelet mesh compression method with several different error metrics aims to provide the first small step in this direction.

6. CONCLUSION

We evaluated the rate-distortion performance of the Lounsbery wavelet mesh compression scheme by using four different distortion metrics – the Hausdorff distance (d_H), the Root Mean Square Error (*RMSE*), the Mesh Structural Distortion Measure 2 (*MSDM2*), and the Mean Opinion Scores (*MOS*) obtained through a subjective experiment. We used the *MOS* to evaluate how well the *MSDM2* metric compares to the human perception of distortion caused by discarding different numbers of wavelet coefficients in the mesh reconstruction. The *MSDM2* has been found to correlate well with the *MOS* in terms of capturing the locations and general levels of these distortions, but has been shown to under-evaluate the perceptual effects of low-frequency (large-scale) shape distortions, especially on very smooth, low-detail models. We have further shown, through our evaluation of the wavelet compression method with different error metrics, that there exist disparities between the performance results produced by the existing distortion metrics, and for this reason it is important to measure and report the performance of a (lossy) mesh compression algorithm with several different distortion metrics. This would make it easier for developers and users of these compression algorithms to make more informed decisions about the applicability of those compression algorithms in different application areas and for different types of 3D models.

Promising future work in perceptual distortion metrics for wavelet-based mesh compression might include an investigation of the visual optimization tools used in JPEG 2000 [ZDL02], to determine whether some of the perceptual models used there to steer 2D image compression might be useful for 3D mesh compression.

7. ACKNOWLEDGEMENTS

The authors would like to thank Michael Lounsbery for the correspondence relating to his subdivision wavelet method, Gabriel Peyré and AIM@SHAPE for providing the subdivision models, and the anonymous reviewers whose comments helped us to improve the quality of this paper.

8. REFERENCES

[CRS98] Cignoni, P., Rocchini, C., Scopigno, R.: Metro: Measuring Error on Simplified Surfaces. In *Computer Graphics Forum* (1998). 17(2): pp. 167-174.

[HP05] Hoppe, H., Praun, E.: Shape Compression using Spherical Geometry Images. In *Advances in Multiresolution for Geometric Modelling*, N.A. Dodgson, M.S. Floater, and M.A. Sabin, Editors. 2005, Springer Berlin Heidelberg. pp. 27-46.

[KG02] Khodakovsky, A., Guskov, I.: Compression of Normal Meshes. In *Geometric Modeling for Scientific Visualization*, G. Brunnett et al., Editors. 2002, Springer Verlag. pp. 189–206.

[KSS00] Khodakovsky, A., Schröder, P., Sweldens, W.: Progressive geometry compression. In *Proc. 27th International Conference on Computer Graphics and Interactive Techniques*, New Orleans, LA, USA, July 2000, pp. 271-278, ACM/Addison-Wesley.

[Lav11] Lavoué, G.: A Multiscale Metric for 3D Mesh Visual Quality Assessment. In *Computer Graphics Forum* (2011). 30(5): pp. 1427-1437.

[LC10] Lavoué, G., Corsini, M.: A Comparison of Perceptually-Based Metrics for Objective Evaluation of Geometry Processing. In *Multimedia, IEEE Transactions on* (2010). 12(7): pp. 636-649.

[LGDBE06] Lavoué, G., Gelasca, E.D., Dupont, F., Baskurt, A., Ebrahimi, T.: Perceptually driven 3D distance metrics with application to watermarking. In *Proc. SPIE Applications of Digital Image Processing XXIX*, San Diego, CA, USA, August 2006, vol. 6312, pp. 63120L.1–63120L.12.

[Lou94] Lounsbery, J.M.: Multiresolution analysis for surfaces of arbitrary topological type. PhD Dissertation. Dept. Comput. Sci. and Engineering, University of Washington, Washington, USA, 1994.

[LTD12] Lavoué, G., Tola, M., Dupont, F.: MEPP - 3D MEsh Processing Platform. International Conference on Computer Graphics Theory and Applications, Rome, Italy, February 2012.

[SF79] Shrout, P.E., Fleiss, J.L.: Intraclass correlations: uses in assessing rater reliability. In *Psychological Bulletin* (1979). 86(2): pp. 420-428.

[SS95] Schröder, P., Sweldens, W.: Spherical wavelets: efficiently representing functions on the sphere. In *Proc. 22nd Annual Conference on Computer Graphics and Interactive Techniques*, Los Angeles, CA, USA, August 1995, pp. 161-172, ACM.

[VP04] Valette, S., Prost, R.: Wavelet-based Progressive Compression Scheme for Triangle Meshes: Wavemesh. In *Visualization and Computer Graphics, IEEE Transactions on* (2004). 10(2): pp. 123-129.

[ZBSS04] Zhou, W., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. In *Image Processing, IEEE Transactions on* (2004). 13(4): pp. 600-612.

[ZDL02] Zeng, W., Daly, S., Lei, S.: An overview of the visual optimization tools in JPEG 2000. In *Signal Processing: Image Communication* (2002). 17(1): pp. 85-104.