

Perceptually driven 3D distance metrics with application to watermarking

Guillaume Lavoué^a, Elisa Drelie Gelasca^b, Florent Dupont^a, Atilla Baskurt^a and Touradj Ebrahimi^c

^aLIRIS UMR 5205 CNRS, University Claude Bernard Lyon 1 and INSA Lyon, France;

^bVision Research Lab, Electrical and Computer Engineering Department, University of California, Santa Barbara;

^cSignal Processing institute, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

ABSTRACT

This paper presents an objective structural distortion measure which reflects the visual similarity between 3D meshes and thus can be used for quality assessment. The proposed tool is not linked to any specific application and thus can be used to evaluate any kinds of 3D mesh processing algorithms (simplification, compression, watermarking etc.). This measure follows the concept of structural similarity recently introduced for 2D image quality assessment by Wang *et al.*¹ and is based on curvature analysis (mean, standard deviation, covariance) on local windows of the meshes. Evaluation and comparison with geometric metrics are done through a subjective experiment based on human evaluation of a set of distorted objects. A quantitative perceptual metric is also derived from the proposed structural distortion measure, for the specific case of watermarking quality assessment, and is compared with recent state of the art algorithms. Both visual and quantitative results demonstrate the robustness of our approach and its strong correlation with subjective ratings.

Keywords: 3D mesh, perceptual quality, structural similarity, curvature analysis, subjective experiment

1. INTRODUCTION

The technological advances in the fields of telecommunication, computer graphics and multimedia during the last decade, have contributed to an evolution of the digital data being manipulated, visualized and transmitted over the Internet. Thus, after sound and 2D images in the eighties, and video in the nineties, three-dimensional data (mostly represented by polygonal meshes) constitute the actual emerging multimedia content. In this context, 3D models are subject to a wide variety of processing operations such as compression, simplification, indexing or watermarking, which can introduce degradations of the visual quality. Since these processes generally deal with visual information that are meant for human consumption, the need of an efficient tool to measure the loss of quality or the visual difference between 3D objects become even more acute.

However, classical metrics based on geometric differences like the Hausdorff distance, available in many softwares^{2,3} do not match well with the human visual perception. This perceptual gap was much investigated in the field of image processing.

1.1. Review of 2D quality metrics

A review of 2D existing perceptual metrics has been done by Eckert and Bradley.⁴ Basically there are two different approaches: *computational* and *ad-hoc*. *Computational* metrics, like the Visible Difference Predictor (VDP) of Daly,⁵ consist in complex numerical models taking into account psychophysical and physiological evidences. These models often relies on the same perceptual attributes:⁴

- The Contrast Sensitivity Function (CSF) which defines the contrast at which frequency components become just visible.

Further author information: (Send correspondence to Guillaume Lavoué.)

Guillaume Lavoué: E-mail: glavoue@liris.univ-lyon1.fr, Telephone: +33 (0)4 26 23 44 64

- The Channel Decomposition, claiming that the human vision consists of several channels selective to spatial frequency and to orientation.
- The Masking Effect which defines the fact that a signal can be masked by the presence of another signal with similar frequency or orientation.

These attributes lead to filtering operations (according to the CSF), filter bank decompositions, errors normalizations (masking effect) and error summation across frequency bands and space (usually using the minkowski metric).

Ad-hoc metrics consider simpler mathematical measures, intuitively relating to the visual perception, and/or introduce penalties for specific artifacts. A typical example is the work of Marziliano *et al.*⁶ which aims at detecting and quantifying blocking and ringing artifacts of JPEG compression.

Recently, Wang *et al.*¹ have introduced an alternative framework for image quality assessment, not relying on a summation of kinds of perceptual errors, but on the degradation of the structural information.

1.2. Review of 3D quality metrics

Transposing complex *computational* metrics from 2D images to 3D objects is quite complex, and to our knowledge, was only investigated by Ferwerda *et al.*⁷ They propose a masking model, extending the Daly *VDP*, which demonstrates how surface texture can mask the polygonal tessellation.

Most of the work on 3D perceptual distances has focused on three specific applications: Realistic rendering, mesh simplification and evaluation of specific processes (compression or watermarking).

The objective of perceptually driven rendering is to determine, according to the location of the observer, which Level Of Details (LOD) to use to satisfy frame rate and image quality requirements. In fact these methods are based on 2D image perceptual models described in the previous paragraph. Reddy⁸ analyses the frequency content in several pre-rendered images to determine the best LOD. In a different way, Bolin and Meyer⁹ use perceptual models to optimize the sampling for ray tracing algorithms. Most of these works concern offline rendering; Dumont *et al.*¹⁰ present a view dependant simplification algorithm for real time rendering, based on the worst cases imperceptible contrast and spatial frequency changes. Although the existing work on realistic rendering is mostly based on 2D image metrics, several authors have considered kinds of 3D metrics like Tian *et al.*¹¹ and Pan *et al.*¹² Their metrics rely respectively on geometry and texture deviations¹¹ and on texture and mesh resolutions.¹²

3D metrics (*ad-hoc*) are used to control mesh simplification algorithms, which consist in reducing the number of vertices while preserving the visual appearance. Kim *et al.*¹³ state that the human vision is sensitive to curvature changes and propose a *Discrete Differential Error Metric* (DDEM), between two vertices v and v' , which is basically the following:

$$DDEM(v, v') = Q(v, v') + T(v, v') + C(v, v') \quad (1)$$

with Q a quadratic distance, T a normal vector difference and C a discrete curvature difference. In a different way, Howlett *et al.*¹⁴ lead their simplification so as to emphasize visually salient features, determined through an eye tracking system. Lee *et al.*¹⁵ follow a similar approach but extract automatically the saliency from the input mesh by computing kinds of multiresolution curvature maps.

Recently several authors have investigated the use of perceptual metrics (*ad-hoc*) for the evaluation of specific applications. Karni and Gotsman,¹⁶ in order to evaluate properly their compression algorithm, introduce the *Geometric Laplacian* (GL), which measures the smoothness of a vertex v :

$$GL(v) = v - \frac{\sum_{i \in n(v)} l_i^{-1} v_i}{\sum_{i \in n(v)} l_i^{-1}} \quad (2)$$

where $n(v)$ is the set of indices of the neighbors of v , and l_i the Euclidean distance from v to v_i . $GL(v)$ represents the difference vector between v and its new position after a Laplacian smoothing step, thus it represents a measure of smoothness: the lower it is, the smoother is the surface around v . Using this *Geometric Laplacian*, Karni and Gotsman introduce a metric between two meshes X and Y , with the same connectivity, containing n vertices:

$$GLD(X, Y) = \frac{1}{2n} \left(\sum_{i=1}^n \|v_i^x - v_i^y\| + \sum_{i=1}^n \|GL(v_i^x) - GL(v_i^y)\| \right) \quad (3)$$

where v_i^x and v_i^y , the respective i^{th} vertices from X and Y . With the same idea, Drelich Gelasca *et al.*¹⁷ propose a perceptual metric based on global *roughness* variation, to measure the quality of a watermarked mesh. They define the *roughness* as the variance of the difference between a 3D model and its smoothed version, similarly to the *Geometric Laplacian* from Karni and Gotsman.¹⁶ Corsini *et al.*¹⁸ present a similar *roughness* based measure. Finally Rondao-Alface *et al.*¹⁹ present two other metrics to benchmark watermarking schemes, one based on a measure of distortion between several 2D views, and the second based on the distortion of energy calculated using 2D parameterization of the meshes.

1.3. Our objective

3D metrics used for simplification¹³⁻¹⁵ are rather vertex-to-vertex local error measurements and have not been considered to produce a single distance value between two meshes. On the other hand, existing global 3D metrics are designed to measure specific artifacts produced by watermarking¹⁷⁻¹⁹ or compression¹⁶ algorithms; these artifacts are mostly like uniform noise thus these existing metrics are not really suitable to evaluate smoothing, simplification or other non-uniform processings against a mesh.

In this context, we introduce a Structural distortion Measure between 3D meshes which reflects their visual perceived distance and is suitable for any kinds of processing applications. This measure is not based on complex psychophysical and physiological models, since these notions are too difficult to manage for 3D objects, but relies on the concept of structural similarity recently introduced for 2D image quality assessment by Wang *et al.*¹ Section 2 provides details about the proposed 3D distortion measure. Most of existing 3D quality measures (presented in the previous section) like those of Rondao-Alface *et al.*,¹⁹ or the *Geometric Laplacian* from Karni and Gotsman¹⁶ do not incorporate a subjective experiment, but are just visually evaluated. However, it seems quite necessary to match a quality measure with psychovisual data collected from subjective experiments, since it allows to (1) evaluating and verifying the correlation of the proposed measure with subjective data and (2) fitting a quantitative model in order to obtain a *real* perceptual metric that estimates quantitatively the perceived quality. In this paper, two subjective experiments are considered (see Section 3), in order to (1) evaluate the correlation of our structural distortion measure with subjective data in a general-purpose context and (2) derive a quantitative perceptual metric in the specific context of watermarking quality assessment.

2. THE PROPOSED 3D DISTORTION MEASURE

This measure was mostly inspired by the work of Wang *et al.*,¹ thus we will firstly describe their 2D image quality assessment algorithm, before presenting in details our measure.

2.1. The 2D image Structural SIMilarity (SSIM) Index

Starting with the assumption that the human visual system is highly adapted for extracting the structural information of a scene, Wang *et al.*¹ define a quality measure between two images, based on visual degradation of the structural information (at the opposite of traditional approaches which rather consider a sum of visible errors). Their structural similarity index $SSIM(x, y)$ between two image signals x and y is defined as follows:

$$SSIM(x, y) = l(x, y) \times c(x, y) \times s(x, y) \quad (4)$$

l is a luminance comparison function, defined as follows:

$$Existingmetricl(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (5)$$

μ_x and μ_y are respectively the mean intensities of signals x and y and C_1 is a constant to avoid instability; c is a contrast comparison function, defined as follows:

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (6)$$

σ_x and σ_y are respectively the standard deviations of signals x and y ; s is a structure comparison function, defined as follows:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad with \quad \sigma_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \quad (7)$$

σ_{xy} is the covariance between x and y , and n the number of pixels of images x and y .

For image quality assessment, this structural quality measure is of course not applied globally between two entire images; local statistics and $SSIM$ measures are computed on local windows (11×11) which move pixel by pixel on the entire image. In other words, for each pixel, the $SSIM$ is calculated on its 11×11 neighborhood. The global structural similarity measure $MSSIM$ between an image X and its distorted version Y is the following:

$$MSSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j) \quad (8)$$

M is the number of local windows in the images and x_j and y_j are the local contents of the j^{th} local windows. $MSSIM(X, Y)$ quantifies how much X and Y are different, with $MSSIM(X, Y) = 1$ if X and Y are identical and $MSSIM(X, Y)$ tends toward 0 when X and Y are very different.

2.2. Our 3D Mesh Structural distortion Measure (MSDM)

2.2.1. Curvature calculation

Our objective is to transpose the quite efficient structural similarity index from Wang *et al.* to 3D objects. The first step is to find a 3D measure analogous to the 2D concept of luminance. The geometric information (i.e. the coordinates of the vertices) does not seem to be relevant for this objective since the human eye is not really sensible to this information. Several authors have considered discrete curvature for 3D quality assessment,^{13,15} indeed this information well describes the visual characteristics of a 3D model, particularly the smoothness, sharpness or roughness of a region.

A triangle mesh is a piecewise linear surface, thus the calculation of its curvature is not trivial. We have implemented the work of Cohen-Steiner *et al.*,²⁰ based on the Normal Cycle, to estimate the curvature tensor at each vertex of the mesh. This estimation procedure relies on solid theoretical foundations and convergence properties. In order to remain independent of the connectivity of the input mesh, the tensor is averaged over a geodesic region, like in.²¹ For each vertex, the curvature tensor is thus calculated and then the principal curvature values $kmin$, $kmax$ are extracted, they correspond to the Eigenvalues of the curvature tensor. Figure 1 presents the mean curvature $kmean = \frac{kmax+kmin}{2}$ of the *Dyno* mesh.

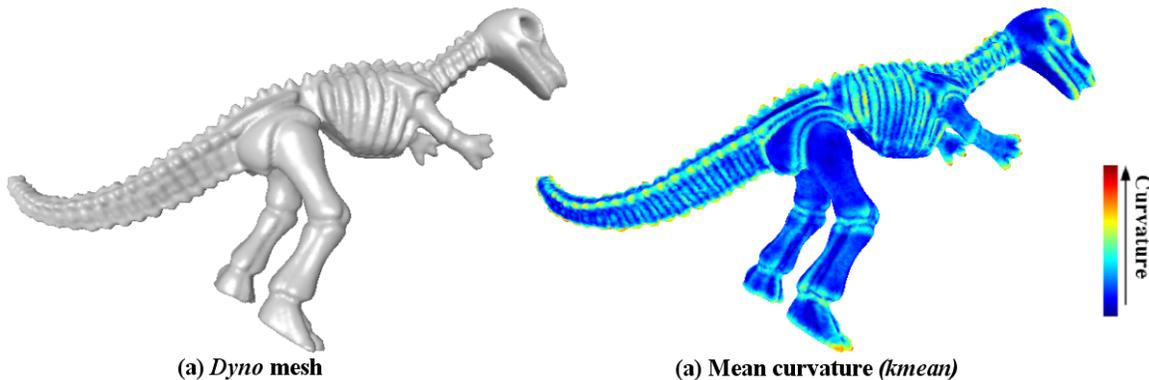


Figure 1. (a) *Dyno* 3D mesh and (b) corresponding mean curvature $kmean$.

2.2.2. Definition of a 3D local window

Following the framework of Wang *et al.*, the proposed measure is based on the computation of statistics (mean, standard deviation and covariance of the curvature) on local windows of the meshes. However, even if the concept of *local window* is trivial in the field of 2D image (a 11×11 neighborhood for each pixel for Wang *et al.*), it becomes quite more complex for 3D objects with non-regular connectivity. For each vertex v , we define its associated *local window*, as the connected set of vertices belonging to the sphere with center v and radius r (see blue points in figure 2.a). We also integrate intersections between this sphere and edges of the mesh, we call

them *edge points* (see green points in figure 2.a), their curvature value is interpolated from their neighbors ones. In figure 2.b, the *edge point* v_e is added to the local window of vertex v , and its curvature $C(v_e)$ is calculated as follows:

$$C(v_e) = \frac{d_2}{d_1 + d_2}C(v_1) + \frac{d_1}{d_1 + d_2}C(v_2) \quad (9)$$

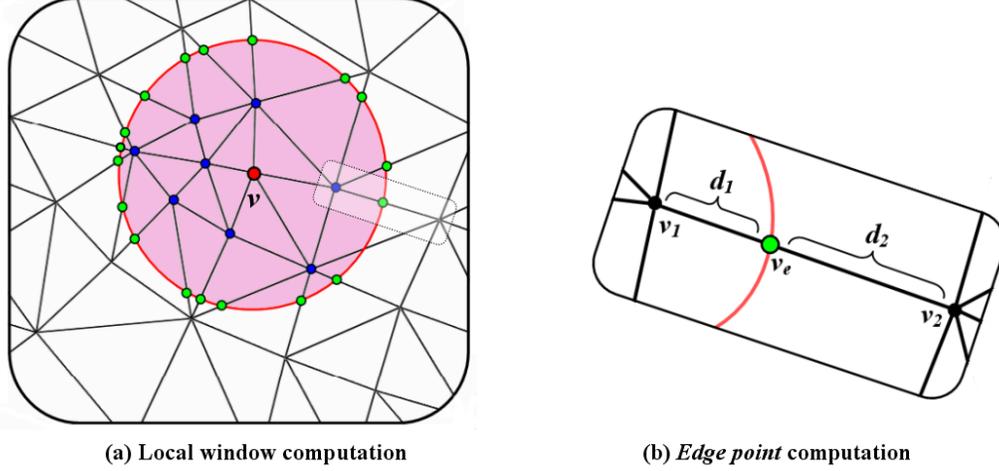


Figure 2. (a) Example of local window computation, for a vertex v . (b) Example of *edge point*.

2.2.3. Comparison functions calculation

For a *local window* x of a 3D mesh, we define the local mean curvature μ_x and standard deviation σ_x as follows:

$$\mu_x = \frac{1}{n} \sum_{v_i \in x} C(v_i) \quad \text{and} \quad \sigma_x = \sqrt{\frac{1}{n} \sum_{v_i \in x} (C(v_i) - \mu_x)^2} \quad (10)$$

with v_i the i^{th} vertex of the local window x and n the number of vertices of this local window.

For two local windows x and y from two 3D meshes, we have to define the concept of covariance (see σ_{xy} in equation 7). Whereas its calculation is trivial for 2D image local windows, the problem is much more complex for 3D local windows, since their connectivity and vertex number and position can be different. σ_{xy} is defined by the following equation:

$$\sigma_{xy} = \frac{\sigma_{xy}^x + \sigma_{xy}^y}{2} \quad (11)$$

with σ_{xy}^x the x -based covariance defined as follows:

$$\sigma_{xy}^x = \frac{1}{n} \sum_{v_i \in x} (C(v_i) - \mu_x)(C(u_i) - \mu_y) \quad (12)$$

u_i is the vertex from local window y , the closest (Euclidean distance) to vertex v_i from x . The y -based covariance σ_{xy}^y is defined similarly.

We then introduce three comparison functions between two corresponding local windows x and y from two meshes X and Y : L (curvature comparison), C (contrast comparison) and S (structure comparison). They are quite similar to those from Wang *et al.*, except that they really refer to *distances* since they tend toward 1 (theoretical limit) when x and y are different and are equal to 0, for identical x and y .

$$L(x, y) = \frac{\|\mu_x - \mu_y\|}{\max(\mu_x, \mu_y)}, \quad C(x, y) = \frac{\|\sigma_x - \sigma_y\|}{\max(\sigma_x, \sigma_y)}, \quad S(x, y) = \frac{\|\sigma_x \sigma_y - \sigma_{xy}\|}{\sigma_x \sigma_y} \quad (13)$$

These distance functions are intuitively strongly linked with psychovisual concepts and several existing works on 3D perception. L represents a normalized curvature distance. The curvature distance, also considered by Kim

et al.,¹³ is strongly linked with normal directions which lead the rendering and therefore the visual appearance of the 3D object. C is based on the standard deviations σ_x and σ_y which reflect the *roughness* of the surfaces. The relevance of *roughness* distance for perceptual metrics was emphasized by several authors.^{16–18} Finally S , by considering the covariance between the local windows, aims at detecting changes in salient features, similarly to the work of Howlett *et al.*¹⁴ and Lee *et al.*¹⁵

It is admitted in the psychovisual research community⁴ that there exist three principal relevant categories of regions in an image or a 3D object, they are presented on figure 3: Edge (a), textured (b), or smooth (c) region. These categories are linked with the concept of *masking*, indeed a textured (or *rough*) region exhibit a high degree of masking, whereas a geometric change on edge or smooth regions is much more visible. The proposed metric well reflects these phenomena: A geometric change on a smooth region will modify the roughness degree and thus will be highlighted by the C coefficient, whereas a modification on an edge region will be shown up by the structural S coefficient.

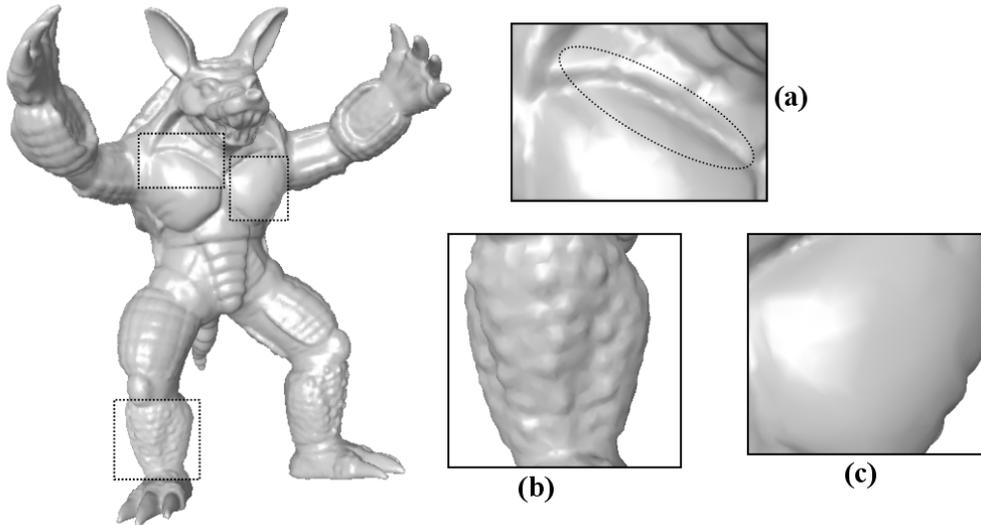


Figure 3. Different relevant regions of the mesh *Armadillo*. (a) Edge, (b) textured and (c) smooth.

2.2.4. Distance computation

Once we have the comparison functions L , C , and S , we have to combine them to obtain a local distance measure $LMSDM$ between two 3D local windows x and y . Wang *et al.*¹ consider a product function (see equation 4), however after some experiments we have found that a Minkowski metric is more appropriate, and increases the stability of the measure:

$$LMSDM(x, y) = (\alpha \times L(x, y)^a + \beta \times C(x, y)^a + \gamma \times S(x, y)^a)^{\frac{1}{a}} \quad (14)$$

Concerning the value of a , many authors admit that values between 2.5 and 4 are strongly linked with psychophysical results.⁴ After several experiments, we have chosen $a = 3$, which seems to provide the best results. α , β and γ were set respectively to 0.4, 0.4 and 0.2. We argue that γ must be smaller since $S(x, y)$ has a larger disparity than both other functions. Of course these parameters (a , α , β and γ) could be adjusted, *a posteriori*, regarding specific experimental results.

Finally, our Mesh Structural distortion Measure $MSDM$, between two meshes X and Y , is defined by a Minkowski sum of their local windows distances:

$$MSDM(X, Y) = \left(\frac{1}{M} \sum_{j=1}^M LMSDM(x_j, y_j)^a \right)^{\frac{1}{a}} \in [0, 1[\quad (15)$$

a has the same value than for equation 14; M is the number of local windows in the meshes and x_j and y_j are the local contents of the j^{th} 3D local windows. In this paper we have considered one local window per vertex,

associated with a radius $r = 0.015$ (meshes are normalized in a unit bounding box). Thus, meshes to compare must have the same connectivity.

3. SUBJECTIVE EXPERIMENTS AND RESULTS

3.1. General-purpose experiment

3.1.1. Experimental measure and corpus

In order to be evaluated, the proposed structural distortion measure *MSDM* has to be matched with a 3D quality evaluation experiment, which consists in assessing the visual quality of several 3D objects (from an experimental corpus) by human observers.

Watson *et al.*²² provides details about the different existing experimental measures: *Naming times*, *rating* and *forced choice preference*. With rating, the observers assign to each stimulus object a number reflecting its visual quality or visual distance with a reference object. Since it is suggested in CCIR Recommendation 500-3,²³ and like much of existing experimental assessment protocols^{12, 14, 17, 18} we have chosen this experimental measure.

The experimental corpus (i.e. the set of distorted 3D objects to evaluate) is quite critical in such subjective experiments. Existing corpus are quite dependant of a specific application: Drelie Gelasca *et al.*¹⁷ and Corsini *et al.*¹⁸ consider a set of objects distorted by different watermarking algorithms associated with different parameters. Pan *et al.*¹² consider several 3D objects with different texture and mesh resolutions, that strongly refer to a rendering application. Finally Howlett *et al.*¹⁴ evaluate a set of simplified objects. Our goal is to evaluate our measure in a general-purpose context thus we have created a corpus with the objective of basically reflecting every possible distortions occurring in common 3D object processing operations. We thus have considered three models: *Armadillo*, *Dyno* and *Venus* (see figure 4), and we have applied two types of distortions:

- **Noise addition:** We modify the three coordinates of vertices of the mesh, according to a randomly chosen offset between 0 and a maximum deviation.
- **Smoothing:** We apply the smoothing filter from Taubin²⁴ to the vertices of the mesh.

These distortions were applied according to three strengths (visually chosen): high, medium and low (these strengths correspond to a number of iterations for smoothing and a value of maximum deviation for noise addition). Finally these distortions were applied according to four situations: Uniformly (on the complete object), on smooth areas, on rough areas and on intermediate areas. The roughness was defined by the variance of the curvature. Since the smoothing was not applied on smooth areas, 21 degraded versions were produced per model (3 noise strengths \times 4 situations + 3 smoothing strengths \times 3 situations), and thus the experimental corpus contains 66 models (3 originals + 3×21 degraded versions). Figure 5 presents some samples of the corpus. These non-uniform noise addition and smoothing basically reflect a lot of possible distortions occurring on common 3D object processing operations:

- Simplification, which finally has a low-pass effect on a mesh, like smoothing does.
- Compression, of which many algorithms are based on a geometric quantization which introduce noise-like deformations.
- Watermarking, which also introduces mainly noise-like deformations.

3.1.2. Subjective evaluation protocol

Few protocols exist for subjective assessment of 3D objects similarity, we have basically followed the evaluation protocol developed by Massimiliano Corsini and Elisa Drelie Gelasca.^{17, 18} Twelve subjects were selected from a pool of students from the Swiss Federal Institute of Technology (Lausanne, Switzerland) and from the University Claude Bernard of Lyon (France). They viewed the 3D objects from comfortable sitting distances, and interaction was allowed (rotation, scaling, translation). Firstly, the original models were displayed together with some distorted ones and with the worst cases (uniform, maximum strength) for noise and smoothing (see figures 5.c and 5.d) in order to establish a referential range for each object. It is quite important to notice that for each object, worst cases for noise and smoothing were displayed and the subject was asked to remember the one he found the worst among them. Finally, the 66 objects of the corpus were displayed iteratively during 20 seconds, and the subjects were asked to provide a score reflecting the degree of perceived distortion, between 0 (identical to the original) and 10 (worst case). In order to avoid the effect of the temporal sequencing factor, the sequence of 66 objects was randomly generated for each participant.

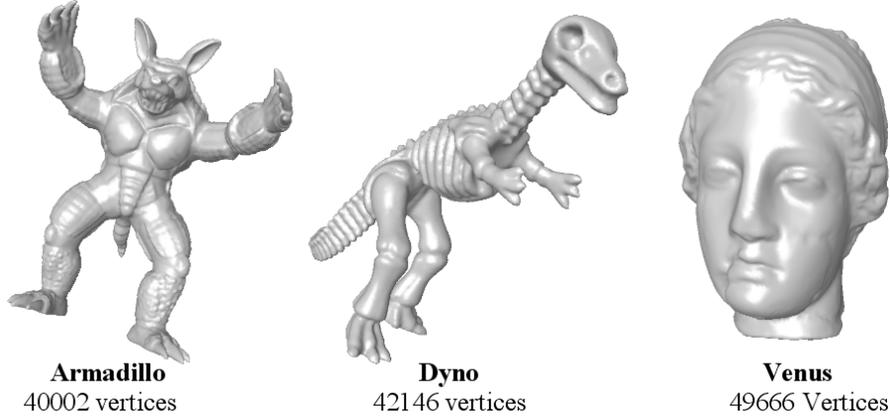


Figure 4. The 3D original models of the experimental corpus.

3.1.3. Subjective data analysis protocol

The judgment scores from each subject depend on some scale and some perception acuity which may vary from person to person, thus these scores have to be matched together by statistical techniques. They are normalized by using mean and variance of each subject. The mean opinion score (MOS) is then computed for each object of the corpus:

$$MOS_i = \frac{1}{n} \sum_{j=1}^n m_{ij} \quad (16)$$

MOS_i is the mean opinion score of the i^{th} object, n is the number of test subjects, and m_{ij} is the normalized score given by the j^{th} subject to the i^{th} object.

In order to validate properly our measure, we basically follow the performance evaluation procedure employed in the video quality expert group (VQEG).²⁵ A cumulative Gaussian psychometric curve is used to provide a non-linear mapping between objective and subjective scores. This fitting just allows here to evaluate the performance of our measure, but could also be considered to construct a real perceptual metric in the case of a specific application (see section 3.2.2). The non linear (rank) correlation of Spearman (r_s) is also calculated, it is considered as a measure of prediction monotonicity.

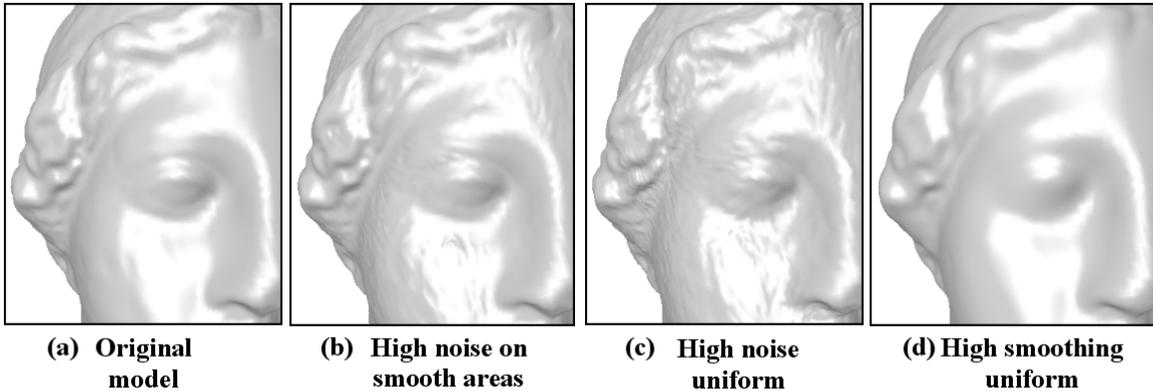


Figure 5. Samples of the corpus for the venus model. (a) Zoom on the original object, (b) High noise (maximum deviation = 0.012) on smooth regions (MOS = 8.8, MSDM = 0.64, Hausdorff mean = 0.16), (c) High noise on the whole object (MOS = 9.4, MSDM = 0.70, Hausdorff mean = 0.26), (d) High smoothing (30 iterations) on the whole object (MOS = 8.1, MSDM = 0.58, Hausdorff mean = 0.25)

3.1.4. Results

Our objective here, is just to evaluate the performance of the proposed structural distortion measure in terms of correlation with subjective scores and in terms of possible prediction capacity after Gaussian psychometric curve fitting. We cannot provide or calibrate a real quantitative perceptual metric in this general purpose context since it strongly depends on the application. Indeed, subjective measures ($\in [0, 10]$) depend on reference bounds (the worst cases displayed at the beginning of the experiment) while the proposed structural distortion measure is *absolute* ($\in [0, 1]$). For this evaluation the performance of our measure is considered object-per-object since providing visually coherent bounds (i.e. worst cases) for different objects is quite impossible in this general purpose context. We have compared the MSDM with Hausdorff distances (maximum, mean and root mean square) calculated using the MESH³ software.

Figure 5 provides a first visual example for the *Venus* model: the *noised* object in (b) is associated with a higher MOS than the *smoothed* object in (d) (MOS=8.8 vs 8.1). That seems intuitively normal since the smoothed model appears visually less distorted than the noised one. The proposed measure well reflects this subjective opinion since the MSDM for (b) is also higher than for (d) (MSDM=0.64 vs 0.58). On the contrary, the geometric mean distance does not reflect at all this subjective opinion (Hausdorff mean=0.16 vs 0.25).

Table 1 presents the performances of our measure and Hausdorff metrics in term of Spearman correlation C_S with subjective results. Overall the proposed MSDM outperforms Hausdorff distances, particularly for Venus and Dyno objects. We have to notice the high stability of MSDM results (C_S is around 75 – 85%), whereas Hausdorff distances performances are very dependant of the considered object: For instance Hausdorff RMS distance gives an excellent correlation value for Armadillo ($C_S = 89\%$) but is quite weak for Venus ($C_S = 42.2\%$) and Dyno ($C_S = 12.1\%$).

Table 1. Spearman correlations (C_S) of subjective results with several metrics: Our structural distortion measure (MSDM), Hausdorff maximum, mean and root mean square (RMS) distances.

	Hausdorff Max	Hausdorff Mean	Hausdorff RMS	MSDM
Armadillo	0.723	0.863	0.892	0.867
Venus	0.006	0.689	0.422	0.861
Dyno	0.289	0.245	0.121	0.735

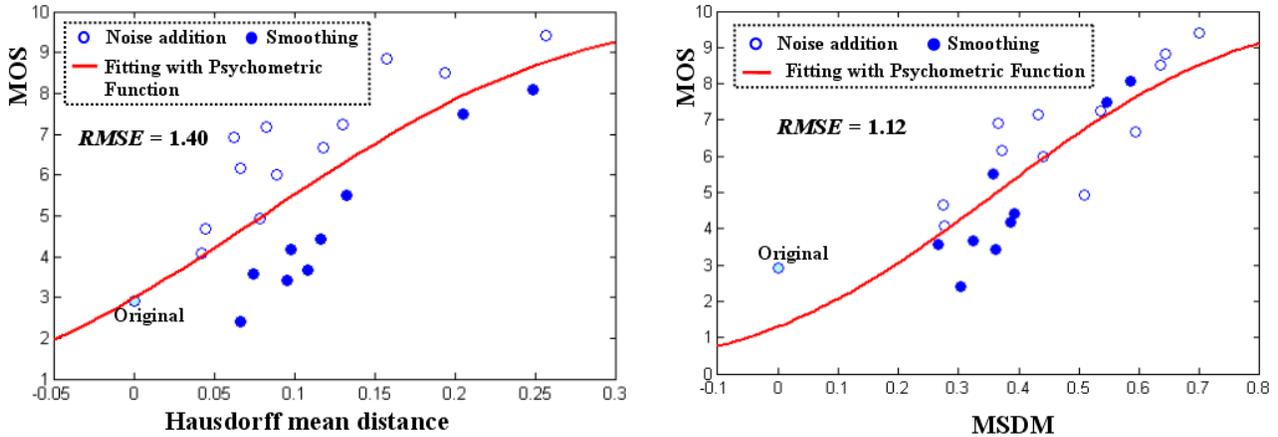


Figure 6. Subjective MOS vs our MSDM and the Hausdorff mean distance for the *Venus* object. Root mean square prediction errors (RMSE) are displayed. Each point represents a test object issued from noise addition (empty circle) or smoothing (full circle).

Figure 6 presents results of the psychometric curve fitting (cumulative Gaussian) between subjective (MOS) and objective scores (Hausdorff mean and MSDM) for the *Venus* model. The Spearman correlation was better for MSDM (see table 1: $C_S = 86.1\%$ vs 68.9%), these results are confirmed since the fitting (i.e. the possible

prediction) is more efficient for MSDM. The prediction error (RMSE) is higher for the Hausdorff mean distance. The main drawback of Hausdorff distance is that results associated with smoothing (full circles) and noise addition (empty circles) are not well merged together and form basically two distinct blocks. This geometric distance basically overestimates the visual error introduced by smoothing while underestimating the visual error associated with noise addition.

3.2. Watermarking experiment

3.2.1. MSDM evaluation and comparison

The proposed MSDM has been evaluated on our general-purpose database, now we will compare this measure on the existing watermarking subjective data issued from the subjective experiment jointly used by Drelié Gelasca *et al.*¹⁷ and Corsini *et al.*¹⁸ Their experiment follows a similar protocol to that which is described in section 3.1. They consider 48 test objects, issued from the application of three watermarking algorithms on 4 different models, with different parameters. The watermarking algorithms are the wavelet based method from Kanai *et al.*,²⁶ the Normal Bin Encoding²⁷ and the Vertex Flood algorithm.²⁸

Our first objective is to compare the proposed structural distortion measure with their two respective roughness based measures, R_1 and R_2 , specifically designed for watermarking quality assessment. Table 2 details results in term of Spearman correlation and figure 7 presents the psychometric curve fitting between the objective/subjective scores.

The correlation associated with MSDM is about 13% better than for their roughness measures ($C_S = 80.1\%$ vs respectively 69.3% and 70.6%). That represents a quite good result since R_1 and R_2 are specifically designed for watermarking whereas the MSDM is for general-purpose.

The prediction associated with the cumulative Gaussian curve fitting is also better for our structural measure, with a prediction error equal to 1.11 against 1.40 for R_2 . Particularly, the dynamic of the MSDM is better spread over its range.

Table 2. Spearman correlations (C_S) of subjective results with several metrics: Our structural similarity distance (MSDM), Hausdorff maximum and mean distances, and roughness metrics from Drelié Gelasca *et al.*¹⁷ (R_1) and Corsini *et al.*¹⁸ (R_2)

Hausdorff Max	Hausdorff Mean	Roughness R_1	Roughness R_2	MSDM
0.376	0.485	0.693	0.706	0.801

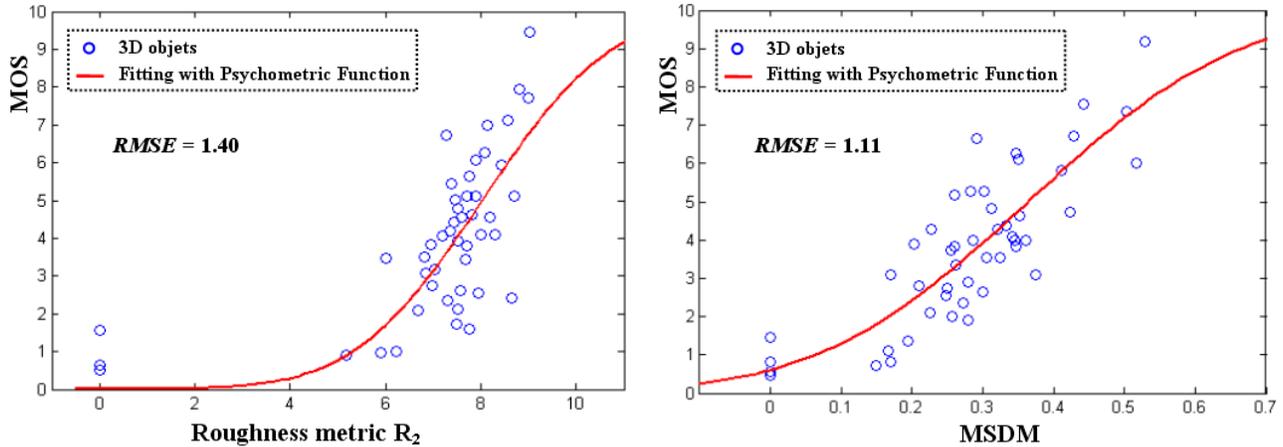


Figure 7. Subjective MOS vs our MSDM metric and the Roughness metric from Corsini *et al.*¹⁸ Root mean square prediction errors (RMSE) are displayed.

3.2.2. Perceptual metric construction for watermarking quality assessment

The proposed MSDM has been evaluated on the watermarking subjective database, now we will use this measure to derive a quantitative perceptual metric. Indeed, on the contrary to the previous general-purpose experiment (see section 3.1), the fact of considering the specific watermarking application allows to determine absolute reference bounds (the worst cases) by using the principal existing 3D watermarking algorithms.

We construct the quantitative perceptual metric M_{MSDM} (between two meshes x and y) by fitting the proposed MSDM outputs with the subjective scores by using a psychometric function g , a cumulative Gaussian.

$$M_{MSDM}(x, y) = g(a, b, MSDM(x, y)) = \frac{1}{2\Pi} \int_{a+b.MSDM}^{\infty} e^{-\frac{t^2}{2}} dt \quad (17)$$

Parameters a and b are estimated by fitting the MSDM values versus the objective data from the experiment. We obtain $a = 1.5521$ and $b = -4.2613$, the corresponding psychometric curve is displayed on figure 7, on the right. With these parameters, the M_{MSDM} metric is then able to predict efficiently the subjective opinions, and can thus be used to benchmark any watermarking schemes.

4. CONCLUSION

This paper has introduced a distortion measure designed for 3D meshes and which follows the concept of structural similarity recently introduced for 2D image quality assessment by Wang *et al.*¹ The 3D Mesh Structural distortion Measure (MSDM) is not linked to a specific application and can be used to evaluate any kinds of 3D mesh processing algorithms (simplification, compression, watermarking etc.). This measure mainly relies on curvature analysis (mean, standard deviation, covariance) on local windows of the mesh. A subjective experiment has been conducted to quantitatively evaluate this distance and to compare with Hausdorff metrics. The proposed distance was also compared with recent state of the art roughness based metrics specifically designed for watermarking assessment, using existing subjective data. Results have proven the effectiveness and the robustness of the MSDM, which provides a high correlation with subjective data. We have also derived from the MSDM, a quantitative perceptual metric that approximates directly the perceived quality in the specific watermarking domain.

The proposed structural distortion measure can be used to *evaluate* 3D mesh processing algorithms, but since it is based on local windows it could also be quite interesting to consider locally this MSDM, to *control* simplification, compression or watermarking algorithms for instance.

ACKNOWLEDGMENTS

We would like to acknowledge Massimiliano Corsini, for the Matlab implementation of the Psychometric curve fitting, for the subjective experiments software and for the tool to compute the roughness metrics.

REFERENCES

1. Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing* **13**(4), pp. 1–14, 2004.
2. P. Cignoni, C. Rocchini, and R. Scopigno, "Metro : Measuring error on simplified surfaces," in *Computer Graphics Forum*, **17**(2), pp. 167–174, 1998.
3. N. Aspert, D. Santa-Cruz, and T. Ebrahimi, "Mesh: Measuring error between surfaces using the hausdorff distance," in *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 705–708, 2002.
4. M. Eckert and A. Bradley, "Perceptual quality metrics applied to still image compression," *Signal Processing* **70**(3), pp. 177–200, 1998.
5. S. Daly, *The visible differences predictor: An algorithm for the assessment of image fidelity*, pp. 179–206. MIT Press, Cambridge, MA, 1993.
6. P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to jpeg2000," *Signal Processing: Image Communication* **19**(2), pp. 163–172, 2004.

7. J. Ferwerda, S. Pattanaik, P. Shirley, and D. Greenberg, "A model of visual masking for computer graphics," in *ACM Siggraph*, pp. 143–152, 1997.
8. M. Reddy, "Perceptually optimized 3d graphics," *IEEE Computer Graphics and Applications* **21**(5), pp. 68–75, 2001.
9. M. Bolin and G. Meyer, "A perceptually based adaptive sampling algorithm," in *ACM Siggraph*, pp. 299–309, 1998.
10. R. Dumont, F. Pellacini, and J. Ferwerda, "Perceptually-driven decision theory for interactive realistic rendering," *ACM Transactions on Graphics* **22**(2), pp. 152–181, 2003.
11. D. Tian and G. AlRegib, "Fqm: A fast quality measure for efficient transmission of textured 3d models," in *ACM Multimedia*, pp. 684–691, 2004.
12. Y. Pan, I. Cheng, and A. Basu, "Quality metric for approximating subjective evaluation of 3-d objects," *IEEE Transactions on Multimedia* **7**(2), pp. 269–279, 2005.
13. S. Kim, S. Kim, and C. Kim, "Discrete differential error metric for surface simplification," in *Pacific Graphics*, pp. 276–283, 2002.
14. S. Howlett, J. Hamill, and C. O'Sullivan, "Predicting and evaluating saliency for simplified polygonal models," *ACM Transactions on Applied Perception* **2**(3), pp. 286–308, 2005.
15. C. Lee, A. Varshney, and D. Jacobs, "Mesh saliency," in *ACM Siggraph*, pp. 659–666, 2005.
16. Z. Karni and C. Gotsman, "Spectral compression of mesh geometry," in *ACM Siggraph*, pp. 279–286, 2000.
17. E. Drelie Gelasca, M. Corsini, and T. Ebrahimi, "Objective evaluation of the perceptual quality of 3d watermarking," in *IEEE International Conference on Image Processing*, pp. 241–244, 2005.
18. M. Corsini, E. Drelie Gelasca, and T. Ebrahimi, "A multi-scale roughness metric for 3d watermarking quality assessment," in *Workshop on image analysis for multimedia interactive services*, 2005.
19. P. Rondao-Alface, M. De Craene, and B. Macq, "Three-dimensional image quality measurement for the benchmarking of 3d watermarking schemes," in *SPIE Electronic Imaging: Security, Steganography, and Watermarking of Multimedia Contents*, pp. 230–240, 2005.
20. D. Cohen-Steiner and J. Morvan, "Restricted delaunay triangulations and normal cycle," in *19th Annu. ACM Sympos. Comput. Geom.*, 2003.
21. P. Alliez, D. Cohen-Steiner, O. Devillers, B. Levy, and M. Desbrun, "Anisotropic polygonal remeshing," *ACM Transactions on Graphics* **22**(3), pp. 485–493, 2003.
22. B. Watson, A. Friedman, and A. McGaffey, "Measuring and predicting visual fidelity," in *ACM Siggraph*, pp. 213–220, 2001.
23. CCIR, "Method for the subjective assessment of the quality of television pictures, recommendation 500-3," in *Recommendations and reports of the CCIR, International Telecommunication Union, Geneva, Switzerland*, 1986.
24. G. Taubin, "A signal processing approach to fair surface design," in *ACM Siggraph*, pp. 351–358, 1995.
25. VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment (<http://www.vqeg.org/>)," tech. rep., 2000.
26. S. Kanai, H. Date, and T. Kishinami, "Digital watermarking for 3d polygons using multi-resolution wavelet decomposition," in *IFIP WG 5.2 International workshop on geometric modeling: fundamental and application (GEO-6)*, pp. 296–307, 1998.
27. O. Benedens, "Watermarking of 3d polygon based models with robustness against mesh simplification," in *SPIE Electronic Imaging: Security, Steganography, and Watermarking of Multimedia Contents*, pp. 329–340, 1999.
28. O. Benedens, "Two high capacity methods for embedding public watermarks into 3d polygonal models," in *Multimedia and Security Workshop at ACM Multimedia*, pp. 95–99, 1999.