OBJECTIVE EVALUATION OF THE PERCEPTUAL QUALITY OF 3D WATERMARKING

Elisa Drelie Gelasca, Touradj Ebrahimi

Signal Processing Institute
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne
Switzerland

Massimiliano Corsini[1], Mauro Barni[2]

[1] Dept. of Electronics and Telecom.
University of Florence, Italy
[2] Dept. of Information Engineering
University of Siena, Italy

ABSTRACT

In this paper an objective metric to measure the perceptual quality of watermarked 3D meshes is presented. The metric, which is based on a black-box approach, relies on the measurement of the roughness of 3D meshes before and after the insertion of the watermark. To calibrate the metric and to validate it, a set of psychovisual experiments has been carried out. Due to the lack of prior work in this field, a new methodology for the subjective evaluation of the quality of watermarked 3D objects is introduced. The validity of the proposed metric has been tested against a number of different 3D watermarking algorithms, showing an excellent match with the subjective evaluation of the quality stemming from the pshycovisual experiments.

1. INTRODUCTION

In the last decade, digital watermarking has become a very active research topic with important applications in the fields of copyright protection of multimedia content, digital rights management, document authentication and conditional access to enhanced services. Although in these years research efforts have been mainly focused on the watermarking of image and video, and to a lesser extent, of audio data, it has been recently realized that watermarking of 3D objects is also very important, due to the ever increasing diffusion of such objects in many areas including virtual reality and computer aided design (CAD).

As usual, one of the main requirements of 3D watermarking is that the changes brought to the watermarked model must be imperceptible to a human observer. It is then very important that suitable methodologies are developed to measure the quality of the watermarked objects as judged by a human observer.

Generally speaking, there are two major classes of quality criteria evaluations: objective and subjective. Due to the relative novelty of quality evaluation of 3D data, no standardized procedures exist and current studies [1] show this lack of generalization. Nevertheless, since subjective quality measures exhibit some inherent drawbacks (the use of a normalized evaluation room, a large panel of human observers, etc.), there has been a great interest in developing objective metrics for 3D model quality assessment.

Our purpose is to find out an objective metric that can be consistently used as a substitute of subjective evaluation. Specifically, by relying on some preliminary psychovisual experiments, we decided to base our metric on a comparison of the roughness in 3D objects. By interviewing the subjects that participated to the experiments, it was found that most based their judgments on the roughness of the 3D surface.

In this paper we describe the newly proposed metric according to the following scheme. In Sec. 2, a review of the few existing methods to evaluate the quality of 3D meshes is given. Then, Sec. 3 presents the new objective metric. As it will be evident from the discussion in this section, the proposed metric needs to be tuned to psychovisual data. To do so we had to define a methodology for the subjective evaluation of the quality of 3D objects. Since no standard protocol has ever been defined on this respect, we introduced a new methodology which overcomes most of the problems reported in the literature (Sec. 4). This methodology has been used both to tune the parameters of the proposed metric and to validate it. In particular, the procedure we followed is reported in Sec. 5. In Sec. 6, we present and discuss the results obtained by comparing the objective quality measure provided by our new metric and the subjective quality evaluation resulting from the psychovisual experiments. Section 7 draws the conclusions.

2. RELATED WORK

In the past years, a set of techniques have been defined and proposed both subjectively [2, 3] and objectively [4] to be used for watermarked video quality evaluation. To the best of our knowledge no similar standards or objective techniques have been proposed for 3D mesh watermarking. In recent years, in order to properly evaluate *mesh simplification* and *perceptually guided rendering* of 3D objects, few objective metrics have been proposed but even less attention has been payed to establish a reliable procedure for 3D data subjective evaluation and results [1] show the complexity of such a task.

Mesh simplification regards the reduction of the number of vertices and triangles of a polygonal mesh model preserving its visual appearance. In general, the simplification process is driven by a similarity metric that measures the impact of the changes of the model after each simplification step. Two kinds of metrics are usually adopted for simplification: geometric metrics and (perceptual) image-based metrics. The most used global geometry-based metrics for off-line quality evaluation of 3D models [5] are the Maximum Geometric Error and Mean Geometric Error based on the Hausdorff distance. Concerning perceptual image-based simplification, Lindstrom and Turk [5] propose an image -driven approach for guiding the simplification process: the model to be simplified is rendered by considering several viewpoints and an image quality metric is used to evaluate the perceptual impact of the simplification operation. More recently, Luebke et al. [6] developed a view-dependent simplification algorithm based on a simple model of Contrast Sensitivity Function that takes into account texture and
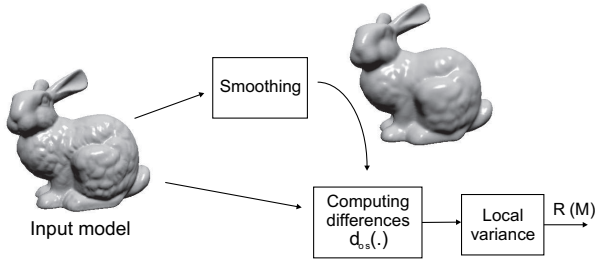
Fig. 1. Smoothing-based Roughness Estimation.

lighting effects.

The aim of perceptually-guided rendering is to accelerate photo-realistic rendering algorithms to avoid computation for which the final result will be imperceptible. Some remarkable works in this field include Bolin and Meyer [7] and the work of Ferwarda et al. [8], where a sophisticated perceptual metric for the evaluation of how much a visual pattern, i.e. a texture, hides geometry artifacts is proposed. The visual masking effect caused by texturing is taken into account by analyzing the final rendered images. In this work we did not account for visual masking, leaving that as an important and interesting area for future.

It is worth to mention that a possible approach to evaluate the visual quality of 3D watermarked objects could be to simply apply image-based perceptual metrics to the final rendered images of the 3D model. The main problem of this approach is that the perceived degradation of still images may not be adequate to evaluate the perceived degradation of the equivalent 3D model. In Rogowitz et al. [1], subjective experiments are conducted on the use of still images of rendered 3D objects for quality assessment and results show that this methodology is view dependent and does not perform well.

3. AN OBJECTIVE METRIC

Our approach relies on the observation, corroborated by the results of the interviews of the subjects involved in a set of preliminary experiments [1], that in most cases the judgment depends on the perceived roughness of the watermarked metric compared to the roughness of the original 3D object [2]. Hence we developed an *ad-hoc perceptual metric* that relies on the computation of the roughness of the 3D objects under investigation.

Once again, the method we developed to measure the perceived roughness relies on the considerations that arose during the preliminary subjective tests. Since during the interview following the experiments, most of the subjects said that the defects were easier to perceive on smooth surfaces, we decided to develop a *perceptual smoothing-based roughness estimation*. The basic idea of this approach is to apply to the model a smoothing algorithm and then to measure the roughness of the surface as the variance of the differences between the smoothed version of the model and the original one. A sketch of our smoothing-based roughness technique is depicted in Figure 1.

---

[1] A more detailed description of the experiments we carried out is given in Sec. 4.

[2] It has to be noted that in many cases the effect of watermark insertion is just an increase of the roughness of the model

The first step of our approach consists in building a smoothed version of the model (M^S) by applying a *smoothing* algorithm to the input model (M). Several possibilities for smoothing exist. Here, we decided to use the *Taubin filter* [9] for its simplicity of implementation (with parameters: $\lambda = 0.6307$, $\mu = -0.6352$, number of iterations = 5) to obtain a medium smoothing effect. When the smoothed model is obtained, the distance between each vertex of M and M^S is computed as follows:

$$d_{OS}(v, v^S) = \mathtt{proj}_{n_v^S}(v - v^S) \qquad (1)$$

where $\mathtt{proj}(.)$ indicates the projection of the vector $(v - v^S)$ on the vertex normals of the smoothed surface (n_v^S). At this point the per-vertex roughness is computed by evaluating the local variance of the distances $d_{OS}(.)$ around each vertex. To be more specific, for each vertex v, the set of distances associated to its *2-ring* ($S_d^2(v)$) is built and the variance of this set evaluated. Then, the per-vertex smooth-based roughness, $\mathcal{R}(v)$ is computed by:

$$\mathcal{R}(v) = \frac{V(S_d^2(v))}{\mathcal{A}_{S^2}} \qquad (2)$$

where $\mathcal{A}_{S^2}$ is the area of the faces that form the 2-ring of v. This area is used at the denominator since surfaces with the same local variance of the distances but smaller areas are assumed to have a higher roughness. The roughness of the input model is the sum of the roughness of all vertices of the model:

$$\mathcal{R}(M) = \sum_{i=1}^{N_v} \mathcal{R}(v_i) \qquad (3)$$

where N_v is the number of vertices of the model. On the basis of several evaluations we decided to define our objective metric as the increment of roughness between the original and the watermarked model. This increment is normalized with respect to the roughness of the original model. In formula:

$$\mathcal{R}(M, M^w) = \log\left(\frac{\mathcal{R}(M) - \mathcal{R}(M^w)}{\mathcal{R}(M)} + k\right) - \log(k) \qquad (4)$$

where $\mathcal{R}(M)$ is the total roughness of the original model and $\mathcal{R}(M^w)$ is the total roughness of the watermarked model. The logarithm is employed to better discriminate between low values of relative roughness increments. The constant k is used to avoid numerical instability of equation (4) since the logarithm tends to $-\infty$ for M^w very close to M. In our case, the value of k has been used to normalize the values between 0 and 10, that is the same range of values given by the subjective tests.

In order to transform this multi-scale roughness estimation in an objective metric that correlates well to human perception of geometric defects we need to collect a set of subjective data. These data have been obtained through the procedure described in the following two sections.

4. A METHODOLOGY FOR SUBJECTIVE QUALITY EVALUATION

Since no standards exist for the evaluation of the quality of 3D objects with impairments, we propose a method for subjective evaluation of 3D watermarked objects based on the criteria usually followed for video and multimedia content quality evaluation [2, 3].

In designing subjective experiments for 3D objects quality evaluation, a first crucial problem is to decide the way the object under examination is rendered. This is not a trivial task hence an accurate study has been carried out. The rendering conditions should not bias the human perception of the 3D model by privileging, for example, one view of the 3D object rather than another one. In our investigations the rendering conditions have been set as listed below.

- *Light source*. All models are illuminated with one white point light source since multiple lights can confuse the observer.

- *Lighting*. We use a simple local illumination lighting model where only the diffusive component of the reflected light is considered to avoid the dependence on camera's position.

- *Texturing*. Image texture mapping, bump mapping, and other kind of texturing may mask the watermarking artifacts [8] so they are not included in the model.

- *Material properties*. The color of a surface is determined by the parameters of the light source that illuminate the surface, by the lighting model used and by the properties of the surface's material. To make the objects as natural as possible, we choose a uniform gray color (*stony-like* objects).

- *Screen and Models Resolution*. The monitor resolution used in the experiments is 1280×600 and each model is displayed in a window of 600×600 pixels. The model occupies around 80% of such window and the resolution of all models ranges between 50.000 and 100.000 triangles allowing for a a good visualization of the model details.

- *Interaction*. In our experimental method, we decide to allow the subject to interact with the model by rotation and zoom operations.

Eleven test subjects (one female, ten males) were drawn from a pool of students aged between 24 and 30. The 3D models were displayed on a 17-inch LCD monitor, with participants sitting approximately 0.4 meters from the display. The experiment followed a five-stage procedure. The stages were: (1) oral instructions, (2) training, (3) practice trials, (4) experimental trials, (5) interview. In the first stage, the subjects were verbally given instructions and made familiar with the task and the graphic interface. In the training the original models and the watermarked models were shown to establish the range for the impairment scale. The practice trials stage was used to familiarize subjects with the experimentation. In the experimental stage, the subject had to give a score to indicate how much the distortions were evident. The subject was instructed to enter a numerical value greater than 0 proportional to the noticed distortion. The value of 10 had to be assigned to the most evident distortion representing the worst cases shown in the training phase. Finally, in the interview stage, we asked the test subjects for a qualitative description of the perceived defects.

5. SUBJECTIVE EXPERIMENTS

We carried out two sets of subjective experiments with different purposes. The first set of experiment (Experiment I), was carried out to fit the roughness-based metric described in the previous section to psychovisual data so to transform it in a perceptual metric. In this first set of experiments, test subjects evaluated differently watermarked models ranging from severely down to weakly visual impairments. Those different distortion strengths

were generated using a specific watermarking algorithm, i.e. the algorithm of Uccheddu et al. (UCB) [10]. The second set of experiments, (Experiment II) was conducted to validate the proposed metrics. Specifically, in Experiment II, we implemented and adopted three other different watermarking algorithms: the Vertex Flood Algorithm (VFA) [11], the Normal Bin Encoding (NBE) [12], the method of Kanai et al. (KDK) [13]. Concerning the kind of impairments introduced in the model, the UCB algorithm produces an uniform kind of noise that can be described as an increase of the roughness of the watermarked surface. VFA produces a kind of noise that looks like marble streak, depending on the viewpoint. The artifacts of the KDK algorithm are the same of the UCB algorithm but due to the geometric tolerance introduced by Kanai to limit the visual impact of the watermarking the final visual effects of such distortions is not uniformly distributed over the model surface. Concerning NBE, the visual aspect (crack-like) of its artifacts is different from that of the UCB, VFA and KDK and more difficult to perceive.

The test models used for both the experiments were: *"Bunny"*, *"Horse"*, *"Venus"* and *"Feline"*. A total of 40 (4 originals $\times$ 3 watermarking strength $\times$ 3 resolution level + 4 originals) test models were used in Experiment I. A total of 48 (4 models $\times$ 11 watermarking settings + 4 originals) test models were used in Experiment II.

5.1. From roughness measure to perceptual metric

Before using the results of Experiment I, we used standard methods to normalize and screen the judgments provided by the test subjects [2]. From the data gathered we calculated the Mean Opinion Score (MOS) of each model. Then these data were used to obtain a perceptual metric $\mathcal{R}^*(M, M^w)$ from Eq. 4 by fitting it with a *Gaussian psychometric function, g*:

$$\mathcal{R}^*(M, M^w) = g(a, b, \mathcal{R}(M, M^w)) = \frac{1}{2\pi} \int_{a+b \cdot \mathcal{R}}^{\infty} e^{-\frac{t^2}{2}} dt$$

where a and b are the parameters to be estimated by fitting the objective metric values versus the subjective data. To estimate such parameters we used a nonlinear least-squares data fitting by the Gauss-Newton method.

The parameters of the Gaussian psychometric curve for the fitting in Experiment I are: $a = 2.0636, b = -0.2981$ for our metric $\mathcal{R}(M, M^w)$. Note that the same fitting parameters were used to validate the proposed metric by means of the results stemming from Experiment II.

6. EXPERIMENTAL RESULTS

The final validation of the proposed metric has been carried out by comparing the results it provides with the data stemming from subjective quality evaluation. More specifically, Experiment II was carried out with three watermarking algorithms (KDK, NBE and VFA) different from that used in Experiment I. The validation procedure is very simple: the perceptual metric is used to predict the MOS obtained in Experiment II and then the correlation among the predicted and the true MOS is computed. The Spearman correlation coefficient [14], r_S is used and reported in Table 1. The rows indicate the watermarking algorithm group. The first two columns of this table also report the performance of two common geometric metrics for comparison purposes: the Maximum and Mean Hausdorff distance. The third column shows the values of

r_S for the presented *perceptual smoothing-based roughness metric*, $\mathcal{R}^*(M, M^w)$. The first row represents the results for Experiment I. From this table, it can be noticed that overall, both geometric metrics based on the Hausdorff distance do not correlate with the subjective data as well as the proposed objective metric. In particular, r_S for NBE and VFA algorithms (third and fourth rows respectively) validate the ability of the proposed metric in predicting impairments introduced by different watermarking algorithms ($r_S = 0.81460$ and 0.9147 respectively). On the other hand, the proposed metric does not perform so well for the KDK algorithm, it gives a correlation $r_S = 0.7111$. This can be explained by considering that the distortion produced by the KDK algorithm on the surface is *non-uniform*. For a better understanding of the metric performance in Experiment II, the results of the perceptual metrics for the watermarking algorithms that introduce only *uniform* distortions on the surface (NBE and VFA) are reported in the sixth row of the table. The value of the correlation coefficient, $r_S = 0.8309$, is high so resulting in good prediction of the developed metric for the quality evaluation. The overall performance of the perceptual metrics considering all the *uniform* and *non-uniform* watermarking algorithms tested in Experiment II, are reported in the last row of the table ($r_S = 0.6929$). Despite the presence of the KDK algorithm, for which the performance are not so good, the global prediction of the metric remains exceedingly better that those of the two geometry-based metrics usually adopted in 3D object quality evaluation.

	Hausdorff Distance		$\mathcal{R}^*(M, M^w)$
Algorithms	Max (r_S)	Mean (r_S)	r_S
UCB	0.6672	0.6595	**0.8919**
KDK	0.6904	0.3230	**0.7111**
NBE	0.7087	0.7026	**0.8146**
VFA	0.4951	0.8815	**0.9147**
NBE+VFA	0.3991	0.6945	**0.8309**
KDK+NBE+VFA	0.3759	0.4853	**0.6929**

Table 1. Perceptual metrics performances.

7. CONCLUSIONS

In this work, an original study about the subjective and objective evaluation of the quality of watermarked 3D objects has been conducted. In particular, a perceptual metric for 3D watermarking impairment prediction has been developed by combining a smoothing-based roughness estimation with subjective data and show very good performance. The experimental results demonstrated the effectiveness of the proposed perceptual metric with respect to the state of the art geometric metrics. We also introduced a new experimental methodology for subjective quality assessment of watermarked 3D objects to fill a gap in the current literature.

8. REFERENCES

[1] B. Rogowitz and H. Rushmeier, "Are image quality metrics adequate to evaluate the quality of geometric objects?," in *Human Vision and Electronic Imaging VI*, Bernice E. Rogowitz and Thrasyvoulos N. Pappas, Eds. 2001, vol. 4299, pp. 340–348, SPIE Proc.

[2] *Subjective Video Quality Assesment Methods for Multimedia Applications Recommendation P.910*, International Telecommunication Union, Geneva, Switzerland, 1996.

[3] *Methodology for Subjective Assesment of the Quality of Television Pictures Recommendation BT.500-11*, International Telecommunication Union, Geneva, Switzerland, 2002.

[4] Winkler S., Drelie Gelasca E., and Ebrahimi T., "Toward perceptual metrics for video watermark evaluation.," in *Applications of Digital Image Processing*. SPIE, August 2003, vol. 5203 of *Proc. of SPIE*, pp. 371–378, SPIE.

[5] Peter Lindstrom and Greg Turk, "Image-driven simplification," *ACM Transaction on Graphics*, vol. 19, no. 3, pp. 204–241, 2000.

[6] Nathaniel Williams, David Luebke, Jonathan D. Cohen, Michael Kelley, and Brenden Schubert, "Perceptually guided simplification of lit, textured meshes," in *Proceedings of the 2003 symposium on Interactive 3D graphics*. 2003, pp. 113–121, ACM Press.

[7] Mark R. Bolin and Gary W. Meyer, "A perceptually based adaptive sampling algorithm," in *Proc. of SIGGRAPH'98*. 1998, pp. 299–309, ACM Press.

[8] James A. Ferwerda, Peter Shirley, Sumanta N. Pattanaik, and Donald P. Greenberg, "A model of visual masking for computer graphics," in *Proc. of SIGGRAPH'97*. 1997, pp. 143–152, ACM Press/Addison-Wesley Publishing Co.

[9] Gabriel Taubin, "A signal processing approach to fair surface design," in *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. 1995, pp. 351–358, ACM Press.

[10] F. Uccheddu, M. Corsini, and M. Barni, "Wavelet-based blind watermarking of 3d models," in *Proceedings of the 2004 multimedia and security workshop on Multimedia and security*. 2004, pp. 143–154, ACM Press.

[11] Oliver Benedens, "Two high capacity methods for embedding public watermarks into 3D polygonal models," in *Proceedings of the Multimedia and Security-Workshop at ACM Multimedia 99*, Orlando, Florida, 1999, pp. 95–99.

[12] Oliver Benedens, "Watermarking of 3D polygon based models with robustness against mesh simplification," in *Proceedings of SPIE: Security and Watermarking of Multimedia Contents*, 1999, vol. 3657, pp. 329–340.

[13] Satoshi Kanai, Hiroaki Date, and Takeshi Kishinami, "Digital watermarking for 3D polygons using multiresolution wavelet decomposition," in *Proc. Sixth IFIP WG 5.2 (GEO-6)*, Tokyo, Japan, Dec. 1998, pp. 296–307.

[14] R. P. Runyon and A. Haber, *Fundamental of Behavioral Statistics*, Addison-Wesley Publishing Co., 1984.