

Watermarked 3D Mesh Quality Assessment

Massimiliano Corsini, Elisa Drelie Gelasca, Touradj Ebrahimi and Mauro Barni

Abstract—This paper addresses the problem of assessing distortions produced by watermarking 3D meshes. In particular, a new methodology for subjective evaluation of the quality of 3D objects is proposed and implemented. Two objective metrics derived from measures of surface roughness are then proposed and their efficiency to predict perceptual impact of 3D watermarking are assessed and compared with the state of the art. Results obtained show good correlations between the proposed objective metrics and subjective assessments by human observers.

Index Terms—3D watermarking, mesh watermarking, objective metrics, perceptual metrics, subjective evaluation, 3D objects quality assessment.

I. INTRODUCTION

In the last decade, digital watermarking has become a very active research topic with important applications in the fields of copyright protection for multimedia content, digital rights management, document authentication and conditional access to enhanced services. Until few years ago the research efforts in digital watermarking have been mainly focused on the watermarking of audio, image and video data; only relatively recently watermarking of 3D objects has gained attention, due to the ever increasing diffusion of such objects in many areas including architecture, design, mechanical engineer, Cultural Heritage and entertainment. Hence, watermarking of 3D objects has not reached the same level of maturity as for still images and video.

One of the main requirements of any watermarking algorithm is the *imperceptibility* of the watermark. This is also a fundamental constraint for 3D watermarking. It is then very important that suitable methodologies are developed to measure the quality of the watermarked objects as judged by human observers. Moreover, study of human perception of geometric artifacts on 3D surfaces due to 3D watermarking can aid the design of efficient watermarking schemes.

Generally speaking, there are two major classes of quality criteria evaluations: *objective* and *subjective*. In fact, one obvious way of determining the quality of any visual information is to measure it by means of subjective experiments with human observers. After all, these signals are meant for human consumption. However, such subjective evaluations are not

only time-consuming and expensive (e.g, the use of a standard evaluation room and a large panel of human observers are mandatory requirements), they can not even be incorporated into automatic systems. Due to the above inherent drawbacks of subjective measures, there has been a great interest in developing objective metrics for 3D models quality assessment.

The goal of this paper is to develop objective measures that can automatically predict the quality of watermarked 3D mesh as perceived by human subjects. More specifically, we propose two perceptual metrics for the quality assessment of watermarked 3D objects. Note that, admittedly, our new metric will not represent a general tool to evaluate the quality of a 3D model, but it will only serve the purpose of judging the quality of watermarked meshes. The final aim of our study is to provide a mean whereby watermarking researchers can evaluate the degradation introduced by the watermark and take appropriate countermeasures to minimize it. At the same time, the proposed metric will be particularly useful for comparing the performance of different 3D watermarking algorithms on the basis of the artifacts perceived on the 3D mesh.

As a matter of fact, some progress has been achieved in the study of perceptual metrics for images and video. A possible approach, then, could be to simply apply such perceptual metrics to the final rendered views 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 [1]. Hence, the approach we chose is to evaluate the human perception of geometric distortions in watermarked models and then to build *ad-hoc* perceptual metrics that work directly on the model's surface. In such a case, subjective experiments dealing directly with the 3D models are needed.

The paper is organized as follows. The state-of-the-art related to the research addressed in this work is reviewed in Section II. Our experimental methodology to carry out subjective experiments on 3D model quality evaluation is described in Section III. In Section IV we describe the subjective experiments and the artifacts introduced by common 3D watermarking algorithms. Section V describes the proposed metrics. Results are presented and discussed in Section VI before drawing the conclusions in Section VII.

II. RELATED WORK

It is widely known among researchers working in image and video watermarking that the characteristics of the Human Visual System (HVS) have to be carefully considered in order to minimize the visual degradation introduced by the watermarking process while maximizing robustness [2] [3]. Many methods have been proposed so far to exploit the models of the HVS to improve the effectiveness of existing watermarking

M.Corsini is with ISTI-CNR of Pisa, Italy. mailto: massimiliano.corsini@isti.cnr.it. Phone:+390503152880. Fax:+390503152811

M. Barni is with Dept. of Information Engineering, University of Siena, Italy. mailto: barni@dii.unisi.it. Phone:+390577234621. Fax:+390577234632

E. Drelie Gelasca is with Department of Electrical and Computer Engineering, University of California, Santa Barbara, California mailto: elisa.drelie@a3.epfl.ch, Phone: +18058932526

T. Ebrahimi is with the Signal Processing Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. mailto: touradj.ebrahimi@epfl.ch, Phone:+41793319993. Fax:+41216937600

schemes. We can divide the approaches proposed so far into theoretical [4] [3] and heuristic [5] [6].

Concerning quality evaluation, in the past years, a set of techniques have been defined and proposed both subjectively [7] [8], and objectively [9] to be used for watermarked video quality evaluation. On the contrary, no agreed solution exists for watermarked 3D meshes. 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 paid to establishing a reliable procedure for subjective evaluation of 3D data and results show the complexity of such a task [1].

Mesh simplification reduces the number of vertices and triangles of a polygonal mesh model while 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 are the Maximum Geometric Error and Mean Geometric Error based on the Hausdorff distance [10] [11]. Concerning perceptual image-based simplification, Lindstrom and Turk [12] 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, Williams *et al.* [13] developed a view-dependent simplification algorithm based on a simple model of Contrast Sensitivity Function that takes into account texture and lighting effects.

The aim of *perceptually-guided rendering* is to accelerate photo-realistic rendering algorithms to avoid computations that do not impact the perceived final result. Some remarkable works in this field include Bolin and Meyer [14], and the work of Ferwarda *et al.* [15], 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.

It is worth to mention that a possible approach to evaluate the visual quality of watermarked 3D objects could be to simply apply image-based perceptual metrics to the final rendered images of the 3D model. This method was investigated in the work by Lindstrom and Turk [12] just mentioned. 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 as concluded by the subjective experiments of Rogowitz and Rushmeier [1]. Another possible approach is to evaluate how the human visual system perceives geometric distortions on the model surface and to build an *ad-hoc* perceptual metric for geometric artifacts. The latter approach is more interesting from a research viewpoint, since no similar studies have been conducted so far. The potential field of applications is not limited to 3D watermarking, but other Computer Graphics applications can also benefit from it. For these reasons, this work has adopted the second approach, i.e. to work directly

on the geometry of the 3D model.

III. 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 watermarked 3D objects based on the criteria usually followed for video and multimedia content quality evaluation [7], [8].

In designing subjective experiments for quality evaluation of 3D objects, 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. For instance, 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. In our investigations the rendering conditions have been set as follows.

- *Non-uniform background.* The 3D model is visualized on a non-uniform background in order to not overestimate the importance of the contour of the model with respect to its overall shape. The background color fades from blue (at the top) to white (at the bottom).
- *Light source.* All models are illuminated with a single white point light source since multiple lights can confuse the observer [16]. To be more specific, each model is illuminated with one white point light source located in a top corner of the Object Bounding Box (OBB) of the 3D object.
- *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 kinds of texturing usually produce masking effects on the perceived geometry [15] and consequently on the perception of watermarking artifacts. Since we leave visual masking studies specific for 3D models as a future research area we do not account for such visual effects in our metrics. For some theoretical considerations about visual masking and texturing we refer to Sec. VII-A.
- *Material properties.* The color of a surface is determined by the parameters of the light source that illuminates the surface, by the lighting model used and by the properties of the surface's material. We set the material properties to obtain a *stone-like* effect. This choice is made for different reasons: first, if all models are seen as "statues" the subjects perceive them in a natural way; second, in this way, the *memory color* phenomenon is avoided [17]. The memory color phenomenon describes the fact that an object's characteristic color influences the human perception of that object's color, e.g. shapes such as heart and strawberries are characteristically red.
- *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 about 80% of such window and the resolution of the models ranges between 50'000 and 100'000 triangles allowing a good visualization of the details.

- *Interaction.* The studies of Rogowitz and Rushmeier [1] suggest that an experimental methodology to evaluate the perceived alterations of 3D objects should rely on the interaction with the model. For this reason, in our experimental method, we decided 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 an uncalibrated 17-inch LCD monitor, with participants sitting approximately 0.4 meters from the display. The experiments 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 subjects had to give a score to indicate how much the distortions were evident. The subjects were instructed to enter a numerical value greater than 0 proportional to the distortion noticed. 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, test subjects were asked to provide a qualitative description of the perceived artifacts.

IV. SUBJECTIVE EXPERIMENTS

Two sets of subjective experiments were carried out with different purposes. The first set of experiments (Experiment I), was performed to tune the two *objective metrics* (proposed in the next section) with psychovisual data in order to transform them into two *perceptual metrics*.

In this first set of experiments, test subjects evaluated differently watermarked models ranging from severe down to weak visual impairments. Those different distortion strengths were generated using a specific watermarking algorithm, i.e. the algorithm by Ucheddu *et al.* (UCB) [18]. 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) [19], the Normal Bin Encoding (NBE) [20], and the method by Kanai *et al.* (KDK) [21]. Concerning the kind of impairments introduced in the model, the UCB algorithm produces a uniform noise that can be described as an increase of the roughness of the watermarked surface. VFA produces a noise similar to a marble streak, depending on the viewpoint. The artifacts of the KDK algorithm are similar to those obtained by 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 are not uniformly distributed over the model surface. Concerning NBE, the visual aspect (crack-like) of its artifacts is different from that of UCB, VFA and KDK and more difficult to perceive. Figure 1 shows a detail of the model “Horse” after application of the four different watermarking algorithms.

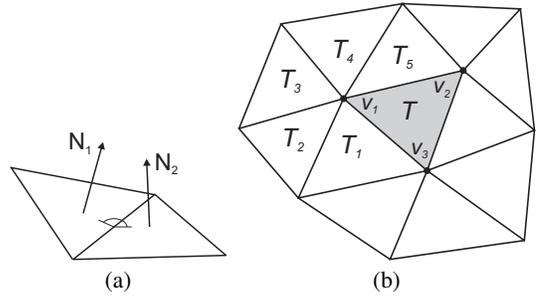


Fig. 2. (a) Dihedral angle. (b) Triangles involved in the evaluation of $G(v_1)$ and $V(v_1)$, used in equation (2).

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.

V. PROPOSED PERCEPTUAL METRICS

Thanks to the intuition and to previous studies about visual aspects of 3D watermarking [22], we argued that a good measure of the visual artifacts produced by watermarking should be based on the amount of roughness introduced on the surface. Moreover, the interviews in Experiments I confirmed that the different types of noise on the surfaces produced by the watermarking can be described essentially with the term *roughness*. Hence, we chose to measure the strength of the artifacts on the basis of an estimation of the surface roughness. In particular, two objective metrics based on roughness estimation have been developed.

A. Multi-scale Roughness Estimation

The first roughness measure we propose is a variant of the method by Wu *et al.* [23]. This metric measures the per-face roughness by making statistical considerations about the dihedral angles associated to each face. Wu *et al.* developed this measure in order to preserve significant shape features in mesh simplification algorithms.

The *dihedral angle* is the angle between two planes. For a polygonal mesh, the dihedral angle is the angle between the normals of two adjacent faces (Fig. 2 (a)). The basic idea of this method [24] is that the dihedral angle is related to the surface roughness. In fact, the face normals of a smooth surface vary slowly over the surface, consequently the dihedral angles between adjacent faces are close to zero. To be more specific, Wu *et al.* associated to each dihedral angle an amount of roughness given by the quantity

$$\rho_d = 1 - (\vec{N}_1 \cdot \vec{N}_2). \quad (1)$$

where \cdot denotes the scalar product between two vectors, and the subscript d was added to explicitly indicate that ρ_d measures the roughness of a dihedral angle.

Given a triangle T with vertices v_1 , v_2 and v_3 , its roughness is computed as:

$$\rho_1(T) = \frac{G(v_1)V(v_1) + G(v_2)V(v_2) + G(v_3)V(v_3)}{V(v_1) + V(v_2) + V(v_3)} \quad (2)$$

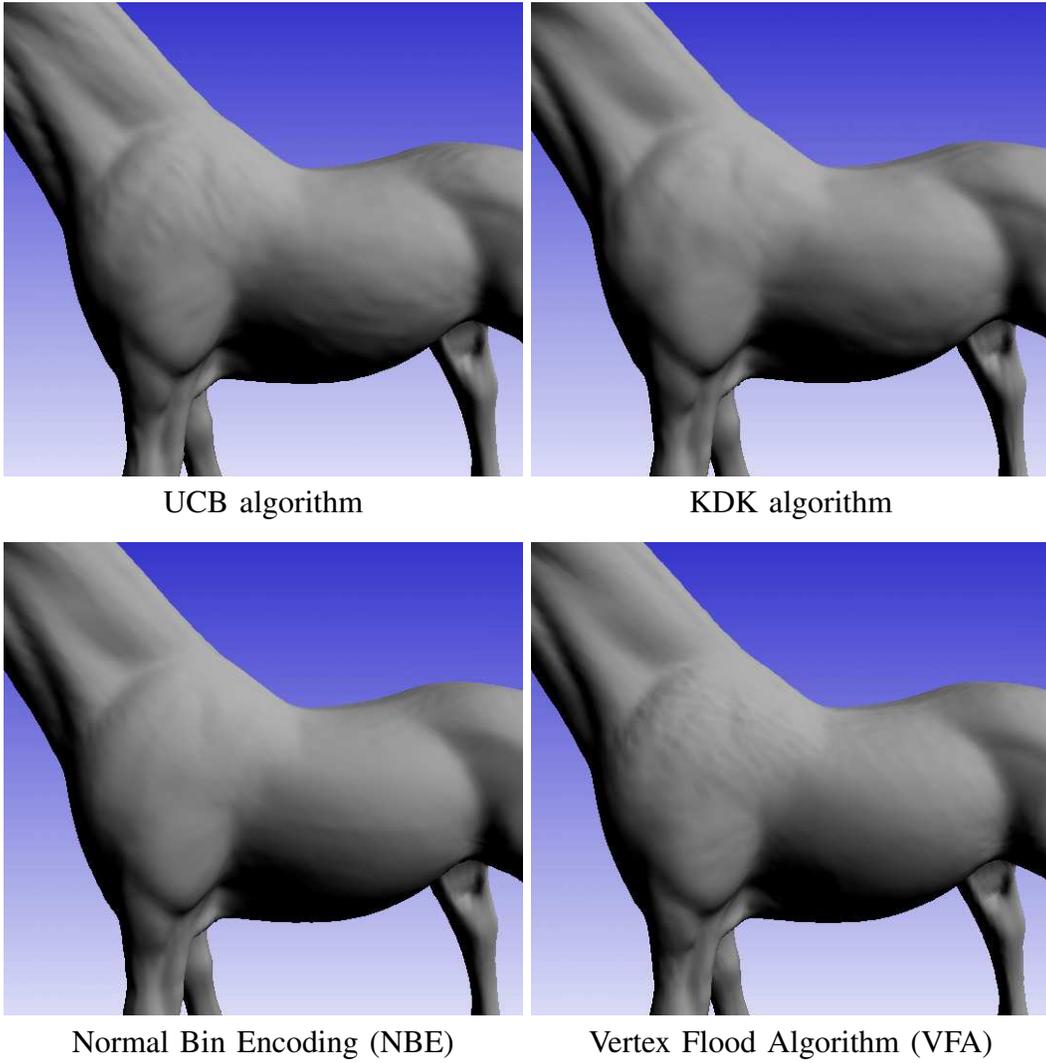


Fig. 1. Geometric defects introduced by 3D watermarking algorithms.

Referring to Fig. 2 (b), $G(v_1)$ is the mean of the roughness associated to the dihedral angles $T - T_1$, $T_1 - T_2$, $T_2 - T_3$, $T_3 - T_4$, $T_4 - T_5$ and $T_5 - T$. In the same way $G(v_2)$ and $G(v_3)$ are the mean roughness associated to the dihedral angles of the faces adjacent to the vertices v_2 and v_3 . Instead, $V(v_1)$, $V(v_2)$ and $V(v_3)$ are the variances of the roughness ρ_d of the dihedral angles of the faces adjacent to the vertices v_1 , v_2 and v_3 .

A rough surface can be considered as a surface with a high concentration of bumps of different sizes. The roughness measure expressed in Eq. (2) is able to measure the 'bumpiness' of surfaces at the face level, but, if the granularity of the surface roughness, i.e. the size of the bump, is bigger than the average size of one face, this metric fails to measure it correctly. In other words, this measure does not take into account the *scale* of the roughness. Our idea is to modify equation Eq. (2) so as to account for different bump scales. The first step to achieve this goal is to transform the per-face roughness estimation in a per-vertex roughness in the following way:

$$\rho_1^N(v) = \frac{1}{|S_T^N|} \sum_{i \in S_T^N} \rho_1(T_i) \mathcal{A}_{T_i} \quad (3)$$

where S_T^N is the set of the faces of the N -ring of the vertex v , $|\cdot|$ is the cardinality operator and \mathcal{A}_{T_i} is the area of the face T_i . Considering the N -ring in the roughness evaluation accounts

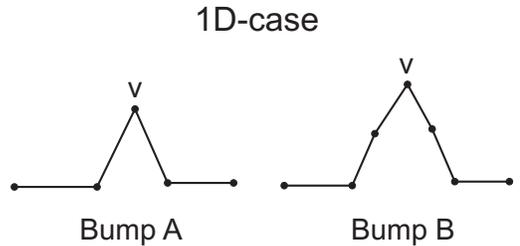


Fig. 3. Bumps with different scale.

for different scales of bumpiness. By referring to Fig. 3, the bump of size equivalent to the 1-ring (A) is well measured by $\rho_1^1(v)$, a correct value of roughness for the vertex v in the case (B) is provided by $\rho_1^2(v)$. Approximately, we can state that the roughness of a vertex v centered on a bump whose area is close to the area of the faces that form the N -ring is well measured by $\rho_1^N(v)$. This approximation could still create

invalid estimates in certain cases. For example for high values of N , or when a surface presents high curvatures. Despite possible impairments, the proposed metric provides acceptable results in most cases. In order to obtain a single value of roughness for each vertex that accounts for the roughness evaluated at several scales, the maximum value produced by N -rings of different sizes is retained. In our objective metric, after several tests, 3 scales of roughness: the 1-ring, the 2-ring and the 4-ring were retained. Hence, the final per-vertex roughness metric becomes:

$$\rho_1(v) = \max\{\rho_1^1(v), \rho_1^2(v), \rho_1^4(v)\} \quad (4)$$

A metric using a linear scale of roughness, i.e. the 1-ring, the 2-ring and the 3-ring, provides similar results.

The total roughness of the 3D object is the sum of the roughnesses of all vertices:

$$\rho_1(M) = \sum_{i=1}^{N_v} \rho_1(v_i) \quad (5)$$

where N_v is the total number of mesh vertices. In the following, we describe how to transform this multi-scale roughness estimation into an objective metric that correlates well to the human perception of geometric artifacts.

B. Smoothing-based Roughness Estimation

The second method [25] is based on the consideration that artifacts are better perceived on smooth surfaces. This was confirmed by interviews with subjects. Hence, a *smoothing-based roughness estimation* was developed. 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. A sketch of the smoothing-based roughness estimation is depicted in Fig. 4.

The first step is to build a smoothed version of the model (M^S) by applying a *smoothing* algorithm to the input model (M). Several possibilities for smoothing exist [26] [27] [28]. Here, we decided to use the *Taubin filter* [26] for its simplicity of implementation. For the Taubin filter we used the parameters $\lambda = 0.6307$, $\mu = -0.6352$ that give a strong smoothing effect. The filter is iterated 5 times. Once the smoothed model is obtained, the distance between each vertex of M and M^S is computed in the following way:

$$d_{OS}(v, v^S) = \text{proj}_{\vec{n}_v^S}(v - v^S) \quad (6)$$

where $\text{proj}(\cdot)$ indicates the projection of the vector ($v - v^S$) on the vertex normals of the smoothed surface (\vec{n}_v^S). At this point the per-vertex roughness is computed by evaluating the local variance of the distances $d_{OS}(\cdot)$ 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 smoothing-based roughness is computed by:

$$\rho_2(v) = \frac{V(S_d^2(v))}{A_{S^2}} \quad (7)$$

where A_{S^2} is the area of the faces that form the 2-ring of v . This area is used as the denominator since surfaces with

the same local variance of the distances but smaller area are assumed to be rougher. The roughness of the input model is the sum of the roughnesses of all the vertices in the model:

$$\rho_2(M) = \sum_{i=1}^{N_v} \rho_2(v_i) \quad (8)$$

where N_v is the number of vertices of the model.

C. Objective Metrics

Now, we describe how to use the roughness estimation to predict the visual distortion produced by a certain 3D watermarking algorithm. 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, leading to:

$$\mathcal{R}(M, M^W) = \log\left(\frac{\rho(M^W) - \rho(M)}{\rho(M)} + k\right) - \log(k) \quad (9)$$

where $\rho(M)$ is the total roughness of the original model and $\rho(M^W)$ is the total roughness of the watermarked model. Both $\rho_1(\cdot)$ and $\rho_2(\cdot)$ can be used to obtain two different objective metrics for the evaluation of the quality of watermarked 3D objects. Equation (9) is a sort of “*Weber Law*” in which the amount of roughness substitutes the grey levels. The logarithm is employed to better discriminate small values of relative roughness increments. The constant k is used to avoid the numerical instability of Eq. (9) since the logarithm tends to $-\infty$ for $\rho(M^W)$ very close to $\rho(M)$. Since in experiments, subjects used values between 0 and 10, the value k has been set to normalize the metric’s outputs to the same range. In the following, we indicate with $\mathcal{R}_1(M, M^W)$ the objective metric based on the multi-scale roughness estimation and with $\mathcal{R}_2(M, M^W)$ the objective metric based on the smoothing-based roughness estimation.

D. From objective to perceptual metric

Basically, there are two approaches [29] to model psychophysical quantities: *performance modeling* and *mechanistic modeling*. Although the distinction is more a continuum than a strict dichotomy, the performance models tend to treat the entire visual system as a “*black box*” for which input/output functions need to be specified. For the mechanistic model, physiological and psychophysical data are used to open the black box. For our metrics, we opted for the black box approach since a model of visual perception of geometric artifacts could become too complex to be handled in practice. In addition, the application-oriented nature of our metrics makes the black box approach to be a more appropriate solution.

Before using the results of Experiment I to obtain our perceptual metric, we summarized the *subjective scores*.

The subjective scores have to be condensed by statistical techniques used in standard methods [8] to yield results which summarize the performance of the system under test. The averaged score values, *Mean Opinion Score (MOS)*, are considered as the amount of distortions that anyone can perceive

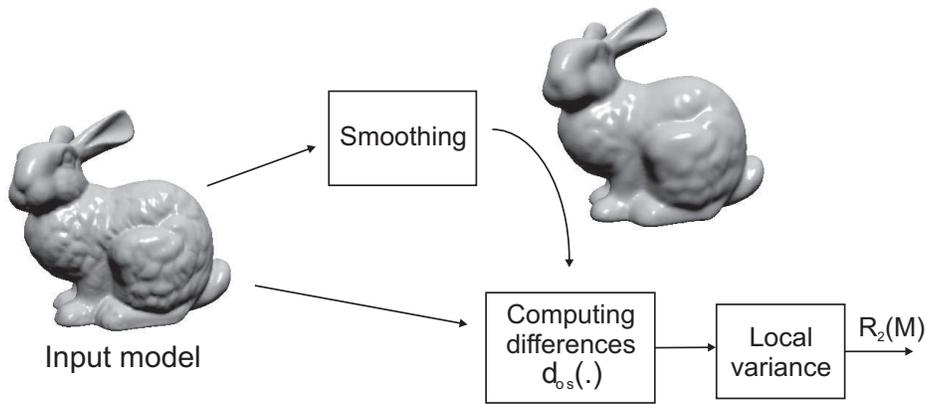


Fig. 4. Smoothing-based Roughness Estimation.

on a particular watermarked 3D object. However, impairment is measured according to a certain scale, and such a scale may vary from person to person. For this reason, we used standard methods to normalize and to screen the judgments provided by the subjects [7]. The *MOS* values obtained following the above mentioned approach were used to derive a perceptual metric by fitting them with a psychometric curve.

The purpose of a psychometric curve is to associate the values given by the objective metric to the subjective score values provided by the subjects. This step is always necessary in order to take into account the saturation effects typical of human senses. For these reasons, psychometric curves exhibit a typical sigmoid shape that penalizes the strongest stimuli. Through psychometric mapping, a match between the human perception of geometric artifacts and the values provided by the objective metric is established. In particular, we use the Gaussian psychometric function:

$$g(a, b, \mathcal{R}) = \frac{1}{2\pi} \int_{a+b\mathcal{R}}^{\infty} e^{-\frac{t^2}{2}} dt \quad (10)$$

where a and b are parameters to be estimated by fitting the objective metric values versus the subjective data, and \mathcal{R} is the objective metric used to measure the visual distortions. We opt for this psychometric curve since it provided the best fit for our data among the commonly used psychometric curves, i.e. Gaussian, logistic and Weibull curves [30]. To estimate the parameters a and b , we used a nonlinear least-squares data fitting by the Gauss-Newton method.

VI. EXPERIMENTAL RESULTS

In this section, we analyze the performance of the two proposed objective metrics and compare them with geometric metrics usually adopted in literature for quality evaluation of 3D models.

Evaluations are performed by assessing the correlations with the *Mean Opinion Scores* described previously. The correlations between the subjective *MOS* collected in Experiment I and the distances given by two geometric metrics based on the Hausdorff distance for model similarity are evaluated. In this way we obtain a term of comparison for the evaluation of our metrics. The data of the Experiment II are used to *validate*

the developed perceptual metrics on different watermarking algorithms.

First, the performance of the state of the art metrics currently used to evaluate the differences between 3D models are analyzed. These state of the art metrics are Mean and Maximum Geometric Error based on Hausdorff distance [10] [11]. Second, the performance of the proposed metrics are assessed and compared to those of the state of the art metrics.

A. Hausdorff distances

As previously stated (Section II) two of the most common geometric metrics used to measure the similarity between two 3D objects, $d_{\infty}(\cdot)$ and $d_1(\cdot)$, are based on the Hausdorff distance between surfaces. Here, we want to evaluate if the Hausdorff distance between the original and the watermarked models could be a reliable metric for perceptual watermarking impairments prediction. To do this, the Hausdorff distances between watermarked models and the originals are plotted versus the *MOS* provided by Experiment I. At this point, the linear correlation coefficient of Pearson (r_P) [31] or the non-linear (rank) correlation coefficient of Spearman (r_S) [32] are calculated in order to evaluate the global performance of the geometric and proposed metrics obtained by fitting these geometric data with the cumulative Gaussian in Eq. (10). The geometric metric results do not correlate well with subjective *MOS*. This underlines the fact that $d_{\infty}(\cdot)$ and $d_1(\cdot)$ do not reflect how humans perceive geometric artifacts due to watermarking. The results are summarized in the first two columns of Table I, and will be used as a reference to compare the performance of our perceptual metrics.

B. Roughness-based Metrics in Experiment I

As stated before, the goal of the first experiment was to make an initial study on the perception of the geometric artifacts caused by watermarking algorithms and to lay the basis for the development of perceptual metrics to measure the visual impact of such artifacts. The experimental data confirm that the subjective perception of the impairments is well-described by the measure of roughness. The subjective data of this experiment are used to obtain two perceptual metrics, named $\mathcal{R}_1^*(M, M^W)$ and $\mathcal{R}_2^*(M, M^W)$, from the corresponding

proposed objective metrics $\mathcal{R}_1(M, M^w)$ and $\mathcal{R}_2(M, M^w)$. These perceptual metrics are obtained by fitting the subjective data with the Gaussian psychometric curve in Eq. (10). This leads to two perceptual metrics: one based on the multi-scale roughness measure and another based on smoothing-based roughness estimation. The parameters of the Gaussian psychometric curve after the fitting are ($a = 1.9428, b = -0.2571$) for $\mathcal{R}_1(M, M^w)$ and ($a = 2.0636, b = -0.2981$) for $\mathcal{R}_2(M, M^w)$. The smoothing-based metric provides a better fit ($r_P = 0.82, r_S = 0.89$) than the multi-scale ($r_P = 0.67, r_S = 0.86$) as depicted in Fig. 5. On the right top of the graphs it is possible to notice some points outside the fitting curve. Most of these outliers correspond to the Venus model. This is due to the fact that the Venus model represents a human face. Human face images are well-known in subjective experiments as a high level factor attracting human attention, i.e. people perceive human faces differently, so the distortions on the Venus head are perceived as more visible and annoying with respect to other models.

C. Perceptual Metrics Performance

As discussed in the previous section, the two proposed objective metrics have been transformed into two corresponding perceptual metrics using the data from Experiment I. In order to evaluate such metrics, Experiment II was carried out with three other watermarking algorithms: KDK, NBE and VFA. The validation is very simple: the perceptual metrics obtained in Experiment I are used to predict the *MOS* obtained in the second experiment and their correlation coefficients are computed. The correlation coefficients r_P and r_S are reported in Table I. The rows indicate which watermarking algorithms were applied to the 3D models. The first two columns of this table report the Spearman correlation coefficients of the Maximum and Mean Geometric Errors for comparison. The third and the fourth show the values of r_P and r_S for $\mathcal{R}_1^*(M, M^w)$, while the last two columns are the r_P and r_S values for $\mathcal{R}_2^*(M, M^w)$. Referring to this table the following considerations are in order:

- Overall, both geometric metrics based on the Hausdorff distance (reported in the first two columns) do not correlate well with the subjective data. On the other hand the developed metrics exhibit good correlations with the subjective data, in particular concerning the Spearman's coefficient.
- The first row reports the UCB's correlation coefficients values for reference.
- The Spearman's coefficients for the NBE and VFA algorithms (second and third rows respectively) demonstrate that both metrics are able to predict impairment introduced by these two algorithms.
- The worst performance of the proposed metrics are obtained for the KDK algorithm. This can be explained by considering that the distortions produced by this algorithm are *non-uniform*.
- The fifth row summarizes the overall performance of the proposed metrics for the other three watermarking algorithms tested in Experiment II. The values of correlation

coefficients ($r_S = 0.71$ for the first metric, $r_S = 0.6929$ for the second metric) outperform the results provided by the state-of-the-art metrics ($r_S = 0.3759$ for the Maximum Hausdorff metric, $r_S = 0.4853$ for the Mean Hausdorff metric).

- The overall performance of the perceptual metrics for the watermarking algorithms that introduce *uniform* distortions are reported in the sixth row of the table. The values of the correlation coefficients ($r_S = 0.8416$ for the first metric and $r_S = 0.8954$ for the second metric) are very good. Hence, one can claim that the developed metrics provide good prediction of the impairment caused by 3D watermarking where artifacts are uniform.

TABLE I
PERCEPTUAL METRICS PERFORMANCES.

Algo.	Hausdorff Distance		$\mathcal{R}_1^*(M, M^w)$		$\mathcal{R}_2^*(M, M^w)$	
	Max (r_S)	Mean (r_S)	r_P	r_S	r_P	r_S
UCB	0.67	0.66	0.67	0.87	0.83	0.90
NBE	0.71	0.70	0.56	0.80	0.62	0.81
VFA	0.50	0.88	0.75	0.94	0.78	0.91
KDK	0.70	0.32	0.62	0.72	0.55	0.71
KDK NBE VFA	0.38	0.49	0.49	0.71	0.50	0.69
UCB NBE VFA	0.52	0.62	0.65	0.84	0.74	0.90

In order to visualize the results of Experiment II, the graphs in Fig. 5 show the values of the objective metrics plotted versus the subjective *MOS* for several watermarking algorithms. It is important to underline that the curves drawn on this figure do not represent the results of a fit as the same psychometric Gaussian curve obtained with the subjective data of the Experiment I are drawn for all the pictures. In other words, these graphs visualize the behavior of the KDK, the NBE and the VFA algorithms with respect to the developed perceptual metrics, that is represented by the solid curve (the dashed line is the confidence interval for that curve).

VII. FINAL REMARKS

In this work, our investigations about the extension of the perceptual image watermarking to 3D watermarking have been presented. In particular, a new experimental methodology for subjective quality assessment of watermarked 3D objects has been proposed. The analysis of the data collected by two subjective experiments demonstrates that the proposed methodology is appropriate and provides reliable subjective data about quality evaluation of watermarked 3D objects. Moreover, two perceptual metrics for 3D watermarking impairments prediction have been developed by combining roughness estimation with subjective data. The performance of these metrics have been analyzed. The results of the analysis demonstrate the effectiveness of the proposed perceptual metrics with respect to the state-of-the-art geometric metrics commonly used for 3D models comparison. More importantly, the experimental results show that the proposed metrics provide a good prediction of the human perception of the distortions introduced by 3D

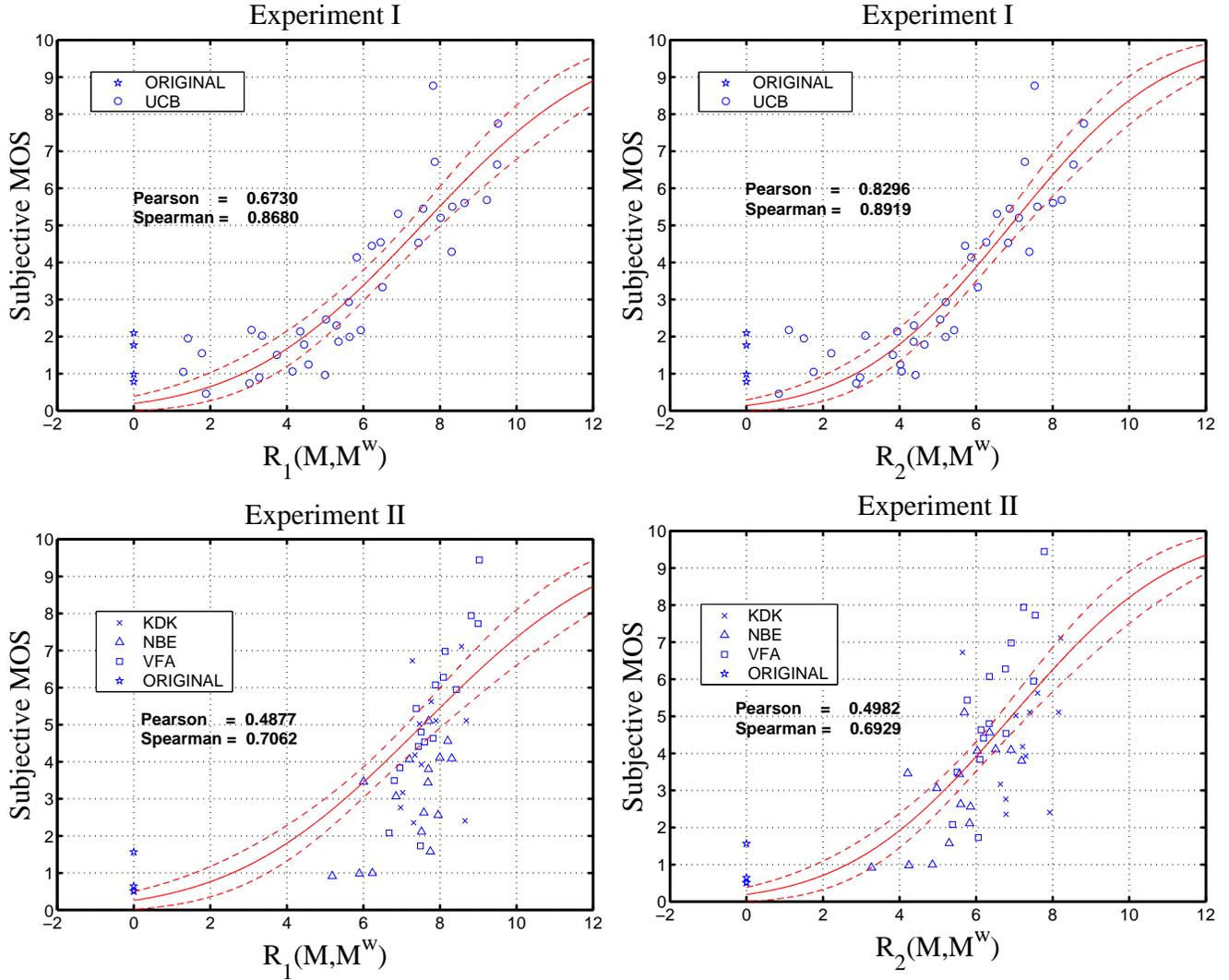


Fig. 5. Experiment I and Experiment II: Subjective *MOS* versus objective metrics curves fits. In Experiment II the parameters of the fitting curve are the same of the Experiment I.

watermarking. Hence, these metrics could be used in a feedback mechanism to tune the watermarking parameters of 3D watermarking algorithms optimizing the watermark insertion. For example, referring to the UCB algorithm, for each level of resolution, the maximum amount of watermark strength before reaching watermark perceptibility could be computed by using these metrics, thus improving the robustness of the algorithm while ensuring imperceptibility.

A. Visual Masking and Texturing

Masking is a well-known perceptual effect of both human auditory and visual systems caused by the fact that the frequency content in certain channels suppress the perceptibility of other frequencies in that channel. In 3D model watermarking, this effect can be caused by the application of image texture on the models' surfaces or by the local peculiarity of the surface. For example, a surface that reproduces a particular fine pattern could mask the watermark during visualization. The proposed metrics compensate well

the visual masking effect caused by geometry since they rely on the *relative* roughness of the model. However, the developed metrics do not take into account the visual masking caused by texture. Due to the importance of textured 3D models, one may want to further analyze the role of the proposed metrics when objects with texture mapping are considered. Ferwerda *et al.* [15] were the first to deal with this issue. They demonstrate that texture images could make geometric artifacts of a surface less perceptible. In their work, the degree of masking of a texture is evaluated by taking into account the contrast, orientation and spatial frequency content of the final rendered image. Yixin Pan *et al.* [33] propose a different interesting approach related to the quality assessment (in terms of geometric and texture resolution) of 3D models. Their work underline that the perceptual contribution of image texture is, in general, more important than the model's geometry. According to their results, we can state that, in the general case, the perceptibility of the watermark on a textured model should be lower w.r.t. the same model rendered without textures. Consequently, our metrics can

be considered as an “upper bound” of the detectability of the watermark on a textured model. In other words, if the impairment predicted by our metrics is low on a 3D mesh, such impairment will be even low(er) for a 3D model with textures.

Concluding, we can state that, despite the fact that the perceptual evaluation of geometric artifacts is a very difficult task due to the enormous number of influencing factors, the results obtained are encouraging. Further research can regard the evaluation of the performance of the proposed metrics under different rendering conditions and their extension to take into account the influence of local properties of the surface (e.g. curvature, protrusions) on the perception of geometric artifacts as well as visual masking effects due to texture mapping.

ACKNOWLEDGMENTS

The authors would like to thank John M. Foley for his valuable inputs, particularly in subjective experiments design.

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*, B. E. Rogowitz and T. N. Pappas, Eds., vol. 4299. SPIE Proc., 2001, pp. 340–348.
- [2] I. Cox and M. L. Miller, “A review of watermarking and the importance of perceptual modeling,” in *Proceedings of SPIE: Vol. 3016. Human Vision and Electronic Imaging II*, Feb. 1997, pp. 92–99.
- [3] R. Wolfgang, C. I. Podilchuk, and E. J. Delp, “Perceptual watermarks for digital images and video,” *IEEE PROC.*, vol. 87, no. 7, pp. 1108–1126, 1999.
- [4] C. I. Podilchuk and W. Zeng, “Image-adaptive watermarking using visual models,” *IEEE JSAC*, vol. 16, pp. 525–539, Apr. 1998.
- [5] R. G. V. Schyndel, A. Z. Tirkel, and C. F. Osborne, “A digital watermark,” in *Proceedings of IEEE International Conference on Image Processing '94*, vol. 2, Austin, Texas, Nov. 1994, pp. 86–90.
- [6] F. Bartolini, M. Barni, V. Cappellini, and A. Piva, “Mask building for perceptually hiding frequency embedded watermarks,” in *Proceedings of the 5th IEEE International Conference on Image Processing, ICIP98*, vol. 1, Chicago, IL, USA, Oct. 1998, pp. 450–454.
- [7] *Subjective Video Quality Assessment Methods for Multimedia Applications Recommendation P.910*. International Telecommunication Union, Geneva, Switzerland, 1996.
- [8] *Methodology for Subjective Assessment of the Quality of Television Pictures Recommendation BT.500-11*. International Telecommunication Union, Geneva, Switzerland, 2002.
- [9] S. Winkler, E. Drelie Gelasca, and T. Ebrahimi, “Toward perceptual metrics for video watermark evaluation,” in *Applications of Digital Image Processing*, ser. Proc. of SPIE, vol. 5203, SPIE. SPIE, August 2003, pp. 371–378.
- [10] N. Aspert, D. Santa-Cruz, and T. Ebrahimi, “Mesh: Measuring error between surfaces using the hausdorff distance,” in *Proceedings of the IEEE International Conference on Multimedia and Expo 2002 (ICME)*, vol. 1, 2002, pp. 705–708.
- [11] R. S. P. Cignoni, C. Rocchini, “Metro: measuring error on simplified surfaces,” *Computer Graphics Forum*, vol. 17, no. 2, pp. 167–174, 1998.
- [12] P. Lindstrom and G. Turk, “Image-driven simplification,” *ACM Transaction on Graphics*, vol. 19, no. 3, pp. 204–241, 2000.
- [13] N. Williams, D. Luebke, J. D. Cohen, M. Kelley, and B. Schubert, “Perceptually guided simplification of lit, textured meshes,” in *Proceedings of the 2003 symposium on Interactive 3D graphics*. Monterey, California: ACM Press, 2003, pp. 113–121.
- [14] M. R. Bolin and G. W. Meyer, “A perceptually based adaptive sampling algorithm,” in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. ACM Press, 1998, pp. 299–309.
- [15] J. A. Ferwerda, P. Shirley, S. N. Pattanaik, and D. P. Greenberg, “A model of visual masking for computer graphics,” in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1997, pp. 143–152.
- [16] R. Shacked and D. Lischinki, “Automatic lighting design using a perceptual quality metric,” *Computer Graphics Forum*, vol. 20, no. 3, 2001.
- [17] E. B. Goldstein, *Sensation and Perception*. 4th Edition: International Thomson Publishing Company, 1996.
- [18] 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*. Magdeburg, Germany: ACM Press, 2004, pp. 143–154.
- [19] O. 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.
- [20] Benedens, “Watermarking of 3D polygon based models with robustness against mesh simplification,” in *Proceedings of SPIE: Security and Watermarking of Multimedia Contents*, vol. 3657, 1999, pp. 329–340.
- [21] S. Kanai, H. Date, and T. Kishinami, “Digital watermarking for 3D polygons using multiresolution wavelet decomposition,” in *Proceedings of the Sixth IFIP WG 5.2 International Workshop on Geometric Modeling: Fundamentals and Applications (GEO-6)*, Tokyo, Japan, Dec. 1998, pp. 296–307.
- [22] F. Uccheddu, M. Corsini, M. Barni, and V. Cappellini, “A roughness-based algorithm for perceptual watermarking of 3D meshes,” in *Proceedings of the Tenth International Conference on Virtual System and Multimedia*, Ogaki City, Japan, Nov. 2004.
- [23] J.-H. Wu, S.-M. Hu, J.-G. Sun, and C.-L. Tai, “An effective feature-preserving mesh simplification scheme based on face constriction,” in *Proceedings of the 9th Pacific Conference on Computer Graphics and Applications*. IEEE Computer Society, 2001, p. 12.
- [24] 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, Montreux, Switzerland, April 2005*. WIAMIS Proc., 2005.
- [25] E. Drelie Gelasca, T. Ebrahimi, M. Corsini, and M. Barni, “Objective evaluation of the perceptual quality of 3d watermarking,” in *IEEE International Conference on Image Processing, Genoa, Italy, September 2005*, vol. 1. ICIP Proc., 2005, pp. 241–244.
- [26] G. Taubin, “A signal processing approach to fair surface design,” in *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM Press, 1995, pp. 351–358.
- [27] L. Kobbelt, “Discrete fairing,” in *Proceedings of the Seventh IMA Conference on the Mathematics of Surfaces '97*, 1997, pp. 101–131.
- [28] M. Desbrun, M. Meyer, P. Schröder, and A. H. Barr, “Implicit fairing of irregular meshes using diffusion and curvature flow,” in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1999, pp. 317–324.
- [29] J. Lubin, “The use of psychophysical data and models in the analysis of display system performance,” in *Digital Images and Human Vision*, A. B. Watson, Ed. Cambridge, Massachusetts: MIT Press, 1993.
- [30] P. G. Engeldrum, *Psychometric Scaling: A Tool for Imaging Systems Development*. Imcotek Press, Winchester, MA, 2000.
- [31] E. W. Weisstein, “Correlation coefficient,” From *Math-World*— A Wolfram Web Resource. [Online]. Available: <http://mathworld.wolfram.com/CorrelationCoefficient.html>
- [32] E. L. Lehmann and H. J. D’Abrera, *Nonparametrics: Statistical Methods Based on Ranks*. Englewood Cliffs, NJ: Prentice Hall, 1998.
- [33] Y. Pan, I. Cheng, and A. Basu, “Quality metric for approximating subjective evaluation of 3-d objects,” *Multimedia, IEEE Transactions on*, vol. 7, no. 2, pp. 269–279, April 2005.