Visual Quality Assessment of 3D Models: On the Influence of Light-Material Interaction

KENNETH VANHOEY, Computer Vision Laboratory, ETH Zürich, Switzerland
BASILE SAUVAGE and PIERRE KRAEMER, ICube, Université de Strasbourg, CNRS, France
GUILLAUME LAVOUÉ, Univ Lyon, CNRS, LIRIS, France

Geometric modifications of three-dimensional (3D) digital models are commonplace for the purpose of efficient rendering or compact storage. Modifications imply visual distortions that are hard to measure numerically. They depend not only on the model itself but also on how the model is visualized. We hypothesize that the model's light environment and the way it reflects incoming light strongly influences perceived quality. Hence, we conduct a perceptual study demonstrating that the same modifications can be masked, or conversely highlighted, by different light-matter interactions. Additionally, we propose a new metric that predicts the perceived distortion of 3D modifications for a known interaction. It operates in the space of 3D meshes with the object's appearance, that is, the light emitted by its surface in any direction given a known incoming light. Despite its simplicity, this metric outperforms 3D mesh metrics and competes with sophisticated perceptual image-based metrics in terms of correlation to subjective measurements. Unlike image-based methods, it has the advantage of being computable prior to the costly rendering steps of image projection and rasterization of the scene for given camera parameters.

CCS Concepts: • **Computing methodologies** → **Appearance and texture representations**; **Perception**; *Mesh models*;

Additional Key Words and Phrases: Quality assessment, 3D object, perceptual study, surface mesh reflectance, appearance

ACM Reference format:
Kenneth Vanhoey, Basile Sauvage, Pierre Kraemer, and Guillaume Lavoué. 2017. Visual Quality Assessment of 3D Models: On the Influence of Light-Material Interaction. *ACM Trans. Appl. Percept.* 15, 1, Article 5 (October 2017), 18 pages.
https://doi.org/10.1145/3129505

1 INTRODUCTION

Three-dimensional (3D) content authoring often has the final goal of visualizing objects in a virtual or augmented world. For that purpose, surfaces represented by meshes are popular. The authoring pipeline generally consists of three successive stages as illustrated in Figure 1: creation, processing, and rendering. Creation is achieved

This work is supported by the European Research Council, under the ERC grant "VarCity" (#273940), and by the Auvergne-Rhône-Alpes region, under the COOPERA grant "ComplexLoD."
Authors' addresses: K. Vanhoey, Computer Vision Laboratory, Sternwartstrasse 7, ETH Zentrum, CH-8092, Switzerland; email: kenneth@research.kvanhoey.eu; B. Sauvage and P. Kraemer, ICube, Pôle API, 300, Boulevard Sébastien Brant, CS 10413, F-67412 Illkirch Cedex; emails: {sauvage, kraemer}@unistra.fr; G. Lavoué, INSA de Lyon, Bât. J. Verne, 20, Avenue Albert Einstein, F-69621 Villeurbanne Cedex; email: glavoue@liris.cnrs.fr.

https://doi.org/10.1145/3129505

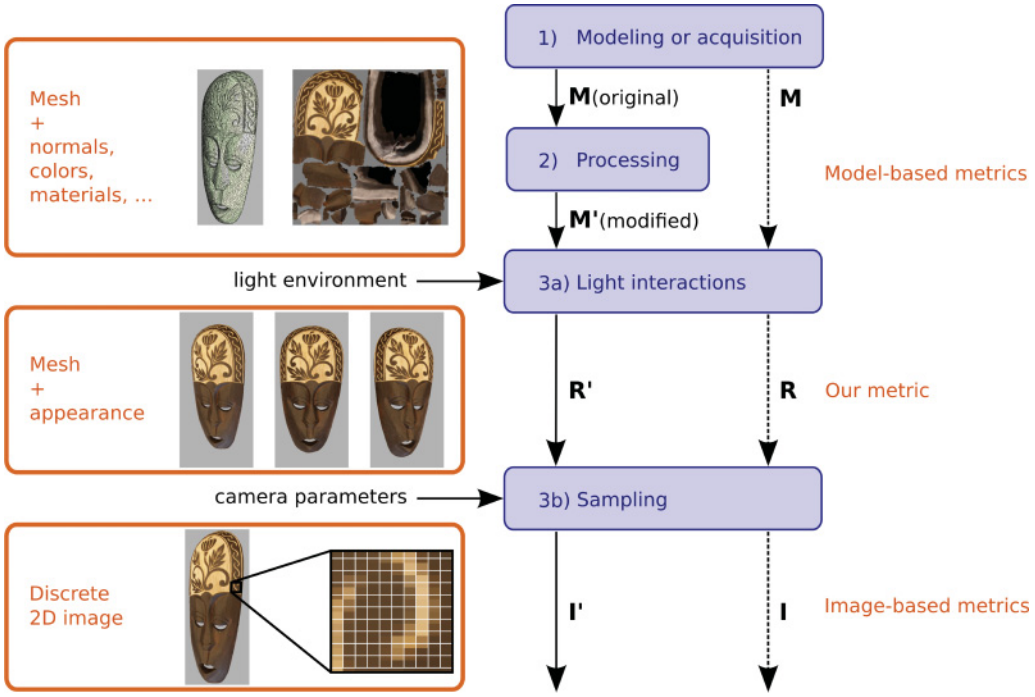

Fig. 1. 3D content authoring pipeline (top-down): (1) Creation by modeling or acquisition generates a mesh with intrinsic properties. (2) Processing distorts this data. (3) Rendering. (a) The light environment interacts with the object, defining the object's omni-directional appearance, and (b) sampling for a particular viewpoint and image-space sampling frequency results in a discrete raster image (Best viewed on screen and zoomed in.).

by modeling the object by hand or through reconstruction from acquired data. The object is represented by its topology, its geometry, and possibly by surface material attributes such as color or reflectance. Processing prepares the object for distribution and rendering and is application dependent. It can include simplification, compression, watermarking, and so on, and generally distorts the initial appearance to some extent. Finally, rendering the final image can be decomposed into two stages: (i) simulate the interaction between light and surface, which defines the object's radiance, that is, its color emitted in all directions and (ii) take the surface's color for a virtual camera setting (viewpoint, projection, sampling frequency, etc.), which finally defines the rendered image.

To guide the mesh processing stage, one should minimize quality loss, that is, maximize visual fidelity to original data. The question to answer is as follows: how do we tune the processing parameters so the process is as lossless as possible w.r.t. the final rendered appearance? Therefore, several *mesh* quality metrics have been designed for measuring visual distortion [5, 17, 35, 40]. They assess distortion as a function of intrinsic mesh properties, that is, geometry, normals, and sometimes color or texture. None of them are aware of light-material interactions, which may strongly influence the final appearance. Another class of quality metrics operate directly on the rendered *images* [1, 3, 14], that is, after sampling (viewpoint selection and rasterization). They are thus view dependent and implicitly integrate both light-material interactions and camera parameters. However, these methods cannot predict quality from the data alone, but rather require costly rendering of final images that in turn requires the camera parameters and resolution to be fixed. Moreover, these image metrics focus on detecting artifacts from global illumination approximation and rasterization and may not be suited for the assessment of mesh processing artifacts.

Our objective is to integrate light-material interactions, while being independent of particular viewpoints and camera parameters. We are inspired by Vanhoey et al. [39], who effectively drive local mesh simplification by exploiting the mesh's outgoing radiance in the particular context of acquired surface light fields. We hypothesize that lighting has a strong influence on perception of surface meshes and thus should be integrated to mesh quality metrics. The intuition is that the lighting and the object's material change the perception of a given modification. In this work, we conduct a user study that corroborates our hypothesis and propose a new metric.

To do so, we reason about *visual* distortion, defined on a surface's appearance. Figure 1 illustrates the general context setting. Processing (e.g., mesh simplification) operates in the space of surface meshes M and their intrinsic properties (left, top box) and so do state of the art mesh metrics. We argue that ignoring the light and material interaction cannot lead to a good perceptual metric. Hence, for our metric, we consider the space of surface appearances R (left, middle box), in which the light environment's contribution has been integrated, but virtual camera parameters (e.g., viewpoint, sampling frequency) have not (left, bottom box). Our experiments confirm that this space is more suited for evaluating the distortions caused by processing operations than the space of surface meshes M.

In summary, our contributions are as follows:

- We design and conduct a user study on both hand-crafted and acquired data (Section 3) and show that perception can dramatically change with the mesh properties (material) and light environment.
- We propose a new metric that predicts the perceived distortion of rendered meshes given a light environment, while being independent of camera parameters (Section 4).
- We validate our metric by showing strong correlation with human perception (Section 4.3). It outperforms state-of-the-art 3D mesh metrics and even competes with state-of-the-art image quality metrics.

2 RELATED WORK

We are interested in visual quality assessment in computer graphics. We can distinguish two main classes of perceptually based quality metrics according to the space they operate in: two-dimensional image metrics and 3D mesh metrics. We also present Just Noticeable Difference (JND) models introduced for 3D meshes in more depth, as well as metrics dedicated to materials.

2.1 2D Image Metrics

In the field of 2D image processing, research into objective image quality assessment metrics is substantially developed [43]. Bottom-up techniques try to mimic the low-level mechanisms of the human visual system (HVS) [6, 23, 25]. In contrast, top-down approaches do not take into account any HVS models but instead operate based on some intuitive hypotheses of what the HVS attempts to achieve when shown a distorted image. The most well-known example is the Structural SIMilarity index (SSIM) [44] and its derivatives, for example, Multi-Scale SSIM [45]. The latest techniques are data driven, that is, the metrics are learned based on user studies [10, 22, 49]. Most of these techniques have been designed and evaluated on *natural* images. However, recently, image metrics have been introduced for the specific case of *synthetic* images [1, 3, 14]: They focus on detecting artifacts introduced in the rendering step. Some mesh processing algorithms have been driven by such 2D image metrics applied on a set of 2D projections, following the example of mesh simplification [21]. Such approaches have the advantage of taking into account light interactions.

However, they are computationally very expensive, because they require us to sample the viewpoints and, for each viewpoint, to rasterize an image. A recent study [18] pointed out that these image-based approaches are only suited for the evaluation of simple processing operations (e.g., simplification). Moreover, they have shown particularly bad results on textured meshes [12].

2.2 3D Mesh Metrics

As opposed to *image* metrics presented above, several *mesh* metrics try to evaluate the visual quality based directly on the 3D geometry. Lavoué [17] and Torkhani et al. [36] proposed metrics based on local differences of curvature statistics, while Váša and Rus [40] considered the dihedral angle differences. These metrics consider local variations of attribute values at vertex or edge level, which are then pooled into a global score. In contrast, Corsini et al. [4] and Wang et al. [42] compute global roughness values per model and then derive a simple global roughness difference. A recent survey [5] details these works and compares their performance in terms of correlation with mean opinion scores derived from subjective rating experiments. Several authors have proposed metrics for the quality assessment of textured meshes [12, 29] by combining geometry and texture image quality measurements. As stated in the Introduction, these mesh metrics ignore the light-material interactions whose role is predominant in the final rendered appearance.

In contrast to the image and mesh metrics presented above, our method considers the light emitted by the surface. Hence, it takes into account the light-material interactions while avoiding both viewpoint sampling and image rasterization, thus still allowing prediction from raw data prior to rendering.

2.3 Just-Noticeable-Difference Models for 3D Meshes

Besides the algorithms presented above on global visual fidelity assessment (suited for supra-threshold distortions), several local perceptual models of JND have been introduced mostly to order the priority of edges to collapse for simplification. Williams et al. [47] map the change resulting from a local simplification operation to a worst-case contrast and a worst-case frequency and then determine whether this operation will be perceptible. This approach is view dependent. Menzel and Guthe [27] propose a similar approach but attempt to be view independent by moving the computation from the 2D space to the vertex-level. More recently, Nader et al. [28] introduced a bottom-up visibility threshold predictor for 3D meshes associated to flat-shaded rendering. Qu and Meyer [32] do not introduce a JND model in the strict sense of the term but quantify local visual masking induced by an object's appearance in texture space using the 2D Sarnoff Visual Discrimination Model (VDM) of Lubin [24]. Mesh density is then locally reduced to the limit where the appearance can mask mesh coarseness. However, no quality metric—even local—is defined.

These works provide models able to predict the near-threshold visibility of artifacts (i.e., visible or not visible), locally, at the vertex level. On the contrary, we aim at proposing a global quality metric, able to evaluate and rank the global perceived quality of distorted 3D models, subject to supra-threshold distortions. Hence, near-threshold JND models are not suited to our purpose as transforming them into a spatially aggregated metric suited for supra-threshold distortions is far from trivial. As noted in Reference [28], if a distortion is visible, a JND model cannot accurately predict to which extent this distortion affects the global visual quality of the model.

2.4 Material Metrics

Several authors have tackled the problem of quality assessment of *material* (mostly represented by Bidirectional Texture Functions) processed by approximation or compression, for example, References [7–9, 13, 15]. These methods try to predict the visual impact of distortions applied on the material only, and do not take into account possible distortions on the 3D shape or even the influence of the 3D shape on material artifact visibility.

Our method operates on the space of surface mesh appearances, which incorporates geometry and the result of the interaction between material and illumination. It can thus predict distortions applied on these three domains altogether.

3 SUBJECTIVE EXPERIMENT

We performed two subjective experiments in which the subject is asked to assess the quality of meshes, using either synthetic or acquired data, respectively. The purpose of these experiments was twofold. First, we wanted to

Fig. 2. The stimuli we used. From left to right: Four synthetic models are two geometric models (*Dragon* and *LionVase*) times two different materials (flat and Phong shading); two physical objects reconstructed with their surface light field (*AfricanMask* and *DoubleDragon*).

show that light and material interactions significantly influence the perception of geometric artifacts. Synthetic data allowed us to create similar meshes with varying materials for this purpose. Second, subjective data were used to evaluate the performance of our metric. For this purpose, we leverage our synthetic data as well as acquired data. We chose a paired comparison methodology, where observers are shown two stimuli side by side and are asked to choose the one that is most similar to the reference. This protocol was shown to be more accurate than others (e.g., single stimulus rating) due to the simplicity of the subjects' task [26]. This section provides details on the subjective studies and their results.

3.1 Stimuli Generation

For our purpose, we gathered two datasets: a synthetic and an acquired one. Synthetic data provides control on the rendering process and allows us to build models with identical geometry but extremely different material. We use it to confirm our assumption that light interaction influences perceived distortion. Acquiring and rendering real-world data validates the relevance of our findings. We use both sets to evaluate our metric by showing correlation with subjective judgments in Section 4.

3.1.1 Synthetic Data. We selected two triangle meshes containing a wide range of geometric characteristics (i.e., smooth and rough regions, sharp features): *Dragon* (114K vertices) and *LionVase* (39K vertices), see Figure 2. The meshes are enlightened by an ambient light and a white point light source bound to the camera (at infinity on the focal axis). This scenario is typical for surface inspection, such as in CAGD applications.

Each mesh is assigned two different appearances by changing the way light is reflected: flat or Phong [30] shading (detailed in Appendix A). In our flat scenario, a unique normal per triangle is used to compute the surface's outgoing colors, and the material is diffuse, that is, there are no specular reflections. Conversely, in our Phong shading, normals are bilinearly interpolated within the triangle prior to color computation, and the material is glossy, that is, specular highlights are included. Due to normal interpolation, Phong shading has a low-pass effect that is beneficial when geometry is imprecise (e.g., contains a high-frequency noise, as introduced by quantization). Conversely, this effect is harmful when geometry is sparse (e.g., when polygon count is low, as after simplification): It creates a blurry aspect. We will see that these effects modify the perceived quality.

These two variants correspond to different light interaction rules in Figure 1, thus we consider it as distinct test objects in our experiments. The four resulting reference models are shown in Figure 2 (four leftmost). The variation in appearance (flat or Phong) for a given geometry is exploited in the remainder of this article to prove our point: light interaction influences the perception so it is relevant to evaluate the perceived distortion in terms of appearance.

Table 1. Details on the Distortions Applied to Each Reference Model

ID	Distortion	*Dragon* (114K vert.)	*LionVase* (38K vert.)	*DoubleDragon* (28K vert.)	*AfricanMask* (48K vert.)
S1	Simplification	48K vert.	21K vert.	14K vert.	24K vert.
S2	Simplification	20K vert.	12K vert.	7K vert.	12K vert.
S3	Simplification	8K vert.	3K vert.	3K vert.	6K vert.
Q1	Quantization	1536	1536	256	256
Q2	Quantization	1024	1024	192	192
Q3	Quantization	768	768	128	128

Note: For quantization, the size of the grid is given.

3.1.2 Acquired Data. We selected digital representations of two physical objects that were acquired under natural lighting, both geometrically and photographically, in a fixed and arbitrary light environment. Both come from Vanhoey et al. [38] and are shown in Figure 2 (two rightmost): *AfricanMask* (48K vertices) and *DoubleDragon* (28K vertices). They are high-fidelity models used for visual inspection and chosen for the range of visual effects they cover. *AfricanMask* exhibits low-frequency reflections and higher-frequency textural motifs, while *DoubleDragon* shows higher-frequency specular peaks and a variety of colors. In this scenario, the result of the light interaction is captured: A surface light field [48] is reconstructed from the photos, and each surface point is assigned a function (called "radiance") defining color per viewing direction [38]. Such data are typically used in the cultural heritage community, because radiance conveys material properties very well [33].

3.1.3 Modifications. The reference models' geometries have been corrupted by two types of distortions, typical of compression/transmission scenarios: simplification and geometry quantization. Simplification was done using the *quadric error metric* (QEM) algorithm by Garland and Heckbert [11]. Geometry quantization was processed on a uniform grid of varying resolution in $\mathbb{R}^3$, snapping each vertex to its closest grid point. The size of the grid is given in Table 1 and corresponds to the direction of the largest side of the object's bounding box. The definition of the two other directions are deduced to generate a regular hexahedral grid.

Each modification was applied with three different strengths that we adjusted to span a whole range of visual distortions, as shown in Figure 3: Level 1 is barely perceptible, level 2 has a low visual impact, and level 3 has a large visual impact. To choose the appropriate distortions, a large set of modification strengths was generated and viewed by the authors, and a subset of them spanning the desired visual quality was chosen to be included in the database. Table 1 lists the complete test set. Note that modifications are applied on geometry only: the light-material interaction plays no role at all in this step. This means, for example, that the models *LionVase* Phong and *LionVase* flat are geometrically identical for a given modification type and level.

3.1.4 User Interaction. Different ways have been used to display the 3D models to the observers in subjective studies, from the most simple such as static images [7, 15, 46] to the most complex by allowing free rotation, zoom, and translation [4, 37]. Though accessing different viewpoints is important for the observer, free interaction causes a cognitive overload that may alter the results of the study. A good compromise is to use animations [12, 29]. For each object in our database, we generate a low-speed rotation animation around the vertical axis and show the resulting video.

3.2 Experimental Procedure

As stated above, we chose a paired comparison methodology that has been demonstrated to be more reliable than rating methods [26]. Participants were shown two videos of distorted models at a time, side by side, and were asked to choose the one that is most similar to the reference. The reference was displayed together with the stimuli, either on top or at the center of them. The best layout was chosen according to the video aspect ratio to span the whole screen. Figure 4 illustrates our rating interface.

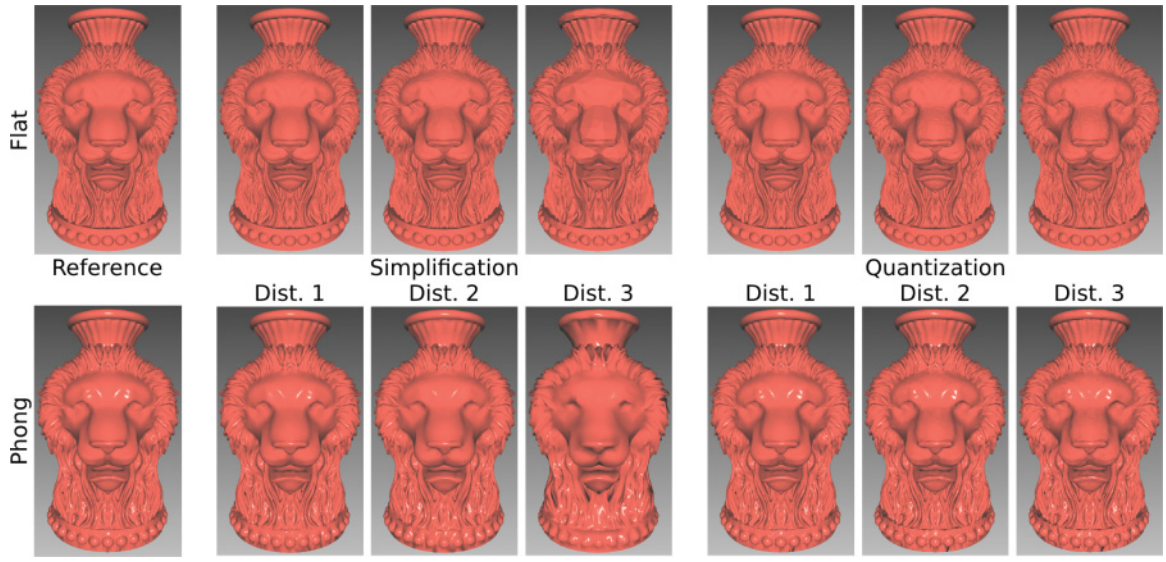

Fig. 3. All the modifications (simplification and quantization, three levels each) are applied on the models *LionVase* with flat shading (top row) and *LionVase* with Phong shading (bottom row).

We opted for a *full* paired-comparison design, meaning that each subject evaluates *all possible* comparison pairs taken from the test set. Practically for each experiment, each subject is presented with $\binom{n}{2} = n(n-1)/2$ pairs to evaluate for each object, with n the number of distorted versions. In our experiments, $n = 6$ versions generate 15 pairs per object: A subject is presented sequentially either the acquired data (2 objects, 30 pairs) or the synthetic data (4 objects, 60 pairs). This full design, as opposed to incomplete ones, is time consuming but allows a complete evaluation of observer agreement and consistency [19].

3.3 Participants

A total of 36 subjects took part in the experiment, aged between 20 and 40, all with normal or corrected-to-normal vision. Participants were graduate students and academic staff not related to our research. On average, it took 25 minutes for one observer to finish the experiment. The experiments were conducted in a dark room on 22in screens (1680×1050 pixels). In total, each video was judged by 16 and 22 observers in the first (synthetic data) and second (acquired data) experiment, respectively.

3.4 Computing Scores

For each experiment, the votes from all subjects performing the task are combined into a single preference matrix per object. In this matrix P, each element $P_{i,j}$ represents the number of times the stimulus i was judged to be of higher quality than stimulus j. Table 2 illustrates the preference matrix for the *LionVase* object with Phong shading from the first experiment. As in Ledda et al. [19] and Mantiuk et al. [26], we consider the number of votes received by each stimuli as its quality score. More sophisticated statistical methods exist for inferring scale values from a preference matrix. We computed scores using Thurstone's Law of Comparative Judgments, Case V [34], which assumes that observers' choices can be thought of as sampled from a normal distribution of underlying quality scores. However, obtained scores were very close to the simple vote counts described above (more than 0.98 Pearson correlation between them for all objects in both experiments), so we decided to use the simpler vote counts.

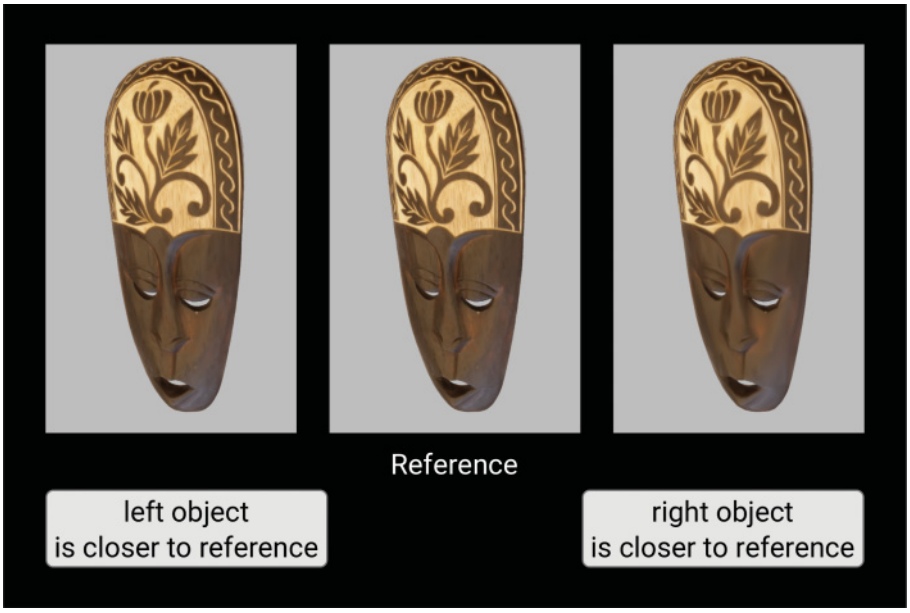

Fig. 4. Our rating interface: the subject is forced to choose the modification that induces the least visual distortion.

Table 2. Preference Matrix for the *LionVase* Object, with Phong Shading

	Q1	S1	Q2	S2	Q3	S3	**Vote count**	**Thurstone**
Q1 = quantization 1536	0	12	21	20	19	22	**94**	**0.00**
S1 = Simplification 21k	10	0	12	22	17	22	**83**	**−0.33**
Q2 = quantization 1024	1	10	0	18	15	21	**65**	**−0.85**
Q3 = quantization 768	3	5	7	14	0	22	**51**	**−1.20**
S2 = Simplification 12k	2	0	4	0	8	22	**36**	**−1.65**
S3 = Simplification 3k	0	0	1	0	0	0	**1**	**−3.45**

Note: Quality scores can be deduced as the Number of times each stimuli has been preferred over others (vote count) or using Thurstone's Law of Comparative Judgments [34].

3.5 Analysis and Discussion

3.5.1 Observer Consistency and Agreement. As in Banterle et al. [2] and Ledda et al. [19], to assess the accuracy of our experimental results, we first analyze the intra –and inter-observer consistencies.

The intra-observer consistency is evaluated by calculating the number of cyclic triads d occurring in an observed configuration of preferences. A cyclic triad occurs when a pair comparison is intransitive, (e.g., A is preferred to B, B is preferred to C and C is preferred to A). The coefficient of consistency ζ is then given, for each observer, as [16]:

$$\zeta = \begin{cases} 1 - \dfrac{24c}{n^3 - n} & \text{if } n \text{ is odd,} \\ 1 - \dfrac{24c}{n^3 - 4n} & \text{if } n \text{ is even,} \end{cases}$$

Table 3. Results of the Experiments

Model	Type	Coef. of Agr. u	Coef. of Cons. ζ (ave.)	1st	2nd	3rd	4th	5th	6th
Dragon Flat	synth.	0.63	0.89	S2	S1	S3	Q1	Q2	Q3
Dragon Phong	synth.	0.63	0.97	Q1	S1	Q2	S2	Q3	S3
LionVase Flat	synth.	0.83	0.96	S1	Q1	S2	Q2	Q3	S3
LionVase Phong	synth.	0.57	0.89	Q1	S1	Q2	Q3	S2	S3
AfricanMask	acq.	0.42	0.80	S1	Q1	S2	Q2	S3	Q3
DoubleDragon	acq.	0.58	0.88	Q1	Q2	S1	Q3	S2	S3

Note: Kendall Coefficient of Agreement u, Average Coefficient of Consistency ζ, and rankings of distortions. The significant differences in rankings between flat and Phong settings illustrate the impact of light-material interactions on the perceived visual quality.

where n is the number of stimuli (i.e., 6 in our experiment) and c the number of cyclic triads. $\zeta = 1$ when there is no circular triads (i.e., perfect consistency) and will decrease to zero as the number of circular triads, and thus the inconsistency, increases.

The inter-observer consistency (also referred to as *agreement*) reflects the similarity of votes between observers. The coefficient of agreement u was introduced by Kendall and Smith [16] as

$$u = \frac{2\Sigma}{\binom{s}{2}\binom{n}{2}} - 1, \tag{1}$$

where s is the number of subjects and Σ is the sum of the number of agreements between all $\binom{s}{2}$ possible pairs of observers and $\binom{n}{2}$ possible pairs of stimuli. It ranges from 1 (perfect agreement) to $-1/(s-1)$ if s is even and $-1/s$ if s is odd.

Table 3 presents the agreement and average consistency results. Overall, values are high as compared to the values typically reported in similar experiments, for example, Ledda et al. [19], Banterle et al. [2], and Leloup et al. [20] reported average consistency values between 0.7 and 0.9 and average agreement values around 0.3. This strong and consistent agreement implies that observers had no difficulty to perceive differences between stimuli and to rank them. The *AfricanMask* object presents slightly lower u and ζ values, implying that the task was more difficult for this specific object.

3.5.2 Observations. Figure 5 (top row) illustrates the quality scores obtained for the two shading methods (Phong and flat) for the synthetic *Dragon* and *LionVase* objects. We observe from this figure and from rankings detailed in Table 3 that (i) the shading method (i.e., the light-material interactions) strongly influences the subjective scores and (ii) the strength of this impact depends on the 3D shape (the impact is stronger for *Dragon*).

To understand the impact of the shading method, we compare the scores and rankings for both shading methods on each object in Figure 5 (top row). We observe that simplified versions are better ranked with flat shading than with Phong shading (see, for instance, lines 1 and 2 in Table 3), and this is reversed for quantized versions. It means that simplification artifacts are less perceived with flat shading than with Phong shading. Conversely, flat shading tends to highlight quantization artifacts as compared with Phong shading. We believe this is due to several perceptual effects. From inspection of our numerical results and our dataset videos illustrated in Figure 6, we can now hypothesize on it.

First, we believe that one of the predominant effects negatively influencing perceived quality is the lack of sharpness, that is, over-smoothness, which is strongly induced by geometric simplification guided by the QEM metric. Phong shading by definition generates a smooth aspect due to its normal interpolation property, while flat shading is not interpolating. As a result, when the polygon count is low, for example, after simplification, Phong shaded meshes look over-smoothed, while flat shaded meshes still show shading discontinuities (see

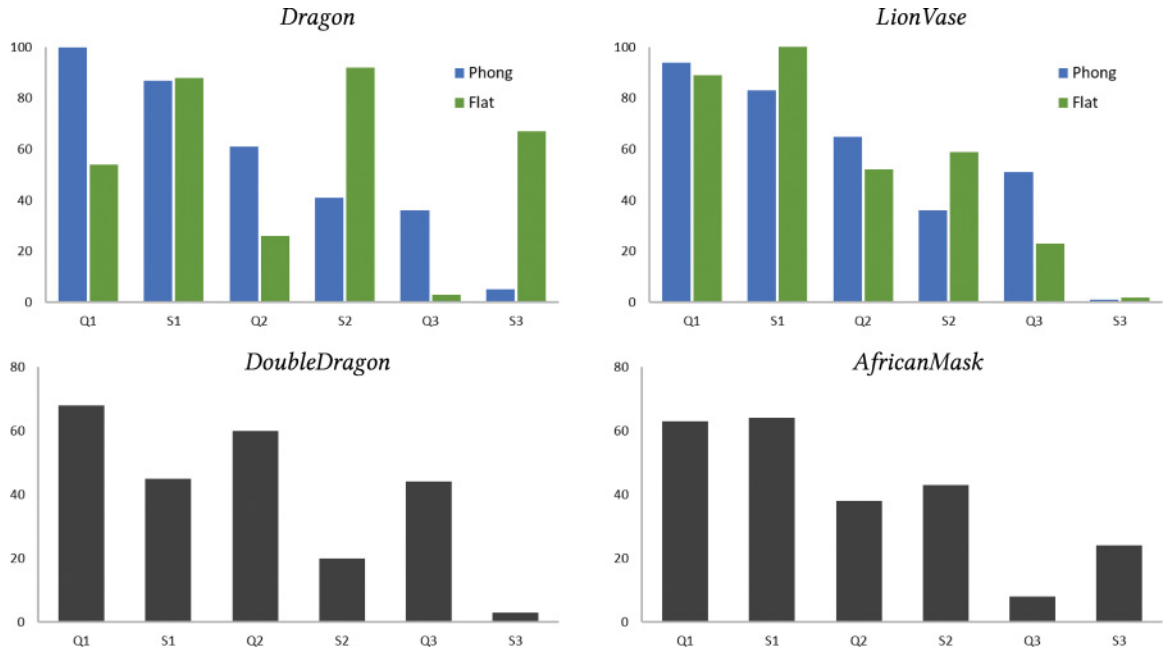

Fig. 5. *Top row:* Comparisons of quality scores obtained for both shading methods for the synthetic *Dragon* and *LionVase* objects. Simplification is well tolerated (high scores) with flat shading, as opposed to Phong shading. The opposite is true for quantization. *Bottom row:* Quality scores obtained for the acquired objects *DoubleDragon* and *AfricanMask* (acquired under natural lighting).

Figure 6). We even hypothesize that flat shading masks simplification artifacts. In Table 3 (first row), one can see that simplification levels 1 and 2 are reversed: Their subjective scores are high and very similar (Figure 5, left), which indicates that these distortions are barely perceptible.

Quantization introduces geometric mid and high frequencies. Normal interpolation (as calculated within Phong shading) is thus beneficial comparatively to flat shading because it has a smoothing effect, while flat shading highlights any discontinuity in the geometry. The material we used for Phong shading being very glossy, we can observe in Figure 6 that this glossiness increases the contrast of the rendered images and thus makes all distortions more perceptible.

It is more difficult to draw strong conclusions for real-world data, that is, the models *AfricanMask* and *DoubleDragon*. Subjective ratings (Figure 5 bottom row and Table 3) show that *AfricanMask* better resists to simplification. Possibly this is because the *AfricanMask* model has smooth geometry, texture, and reflections. Hence simplification, which induces smoothing, is hardly perceived (see Figure 7). Conversely, *DoubleDragon* better resists to quantization. We hypothesize that its higher-frequency specular and texture content allows to mask artifacts—quantization in particular—by exploiting visual masking [32, 41].

To conclude, we have gained insight in the influence of different light-material interactions on the perception of distortions. Our analyzes hint that a mesh quality metric taking the latter into account will perform better. This is the subject of Section 4.

4 METRICS ON RADIANCE DATA

In this section, we show that a metric based on appearance is more correlated to human perception. We formally define the mesh's appearance attribute we use (Section 4.1), introduce a new metric (Section 4.2), and, finally, analyze correlation with the subjective assessments (Section 4.3).

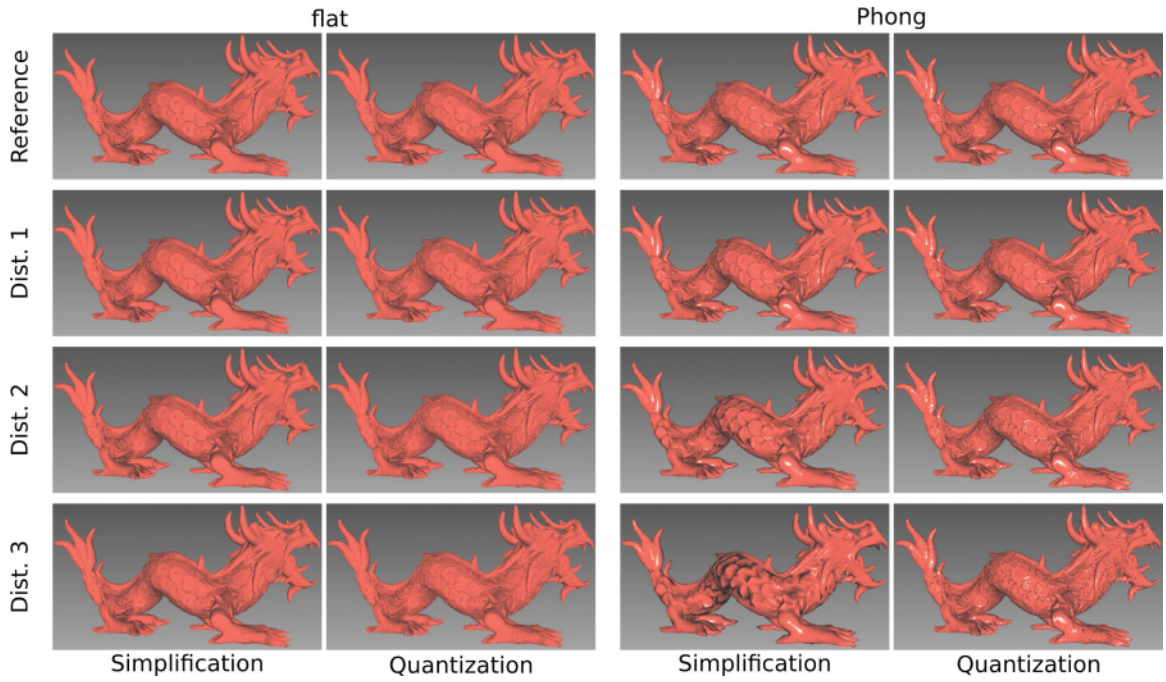

Fig. 6. All distortions on the *Dragon* flat and *Dragon* Phong models.

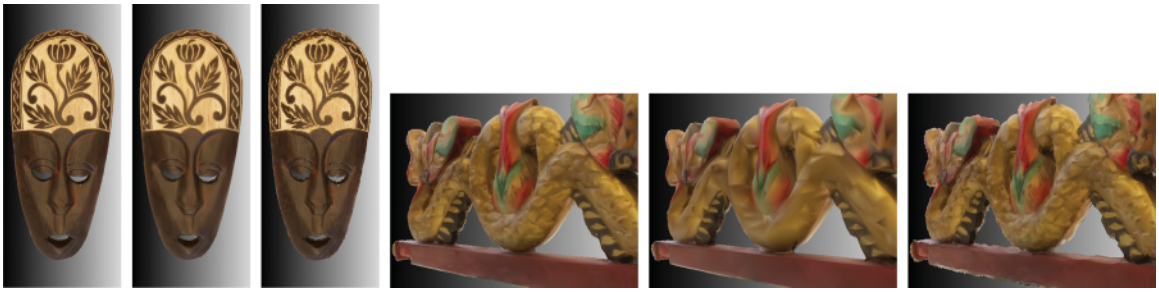

Fig. 7. Real-world data and their level-3 simplification ($S3$) followed by their level-3 quantization ($Q3$). *AfricanMask* is smooth in terms of geometry, texture, and reflections, and, hence, additional smoothing artifacts induced by simplification are not strongly perceived. *DoubleDragon* has a more complex and high-frequency texture compared to *AfricanMask*. Hence, we believe texture masking explains why it resists better to quantization artifacts, as shown quantitatively in Table 3.

4.1 Appearance as a Directional Color

We aim at measuring what an observer perceives when looking at a 3D object in a non-participative environment. From a rendering point of view, he or she perceives, for any point $\mathbf{p}$ on the object's surface, a color that may depend on both the input lighting direction (e.g., due to shading and occlusions) and the output viewing direction (e.g., due to specular highlights). This is a 6D function: 3 parameters, 2D each. In this article, we model the appearance by a color function $R(\mathbf{p}, \omega)$ with 2 parameters only, namely the point $\mathbf{p} \in M$ and the viewing direction ω in the visible hemisphere Ω. The problem is thus reduced to 4 dimensions, which allows for developing a

tractable metric. In the following, we show that the two realistic scenarios presented in Section 3 (synthetic and acquired data) are compliant with this simplification.

In the first scenario, a 3D mesh is enlightened by an ambient light and a point light source bound to the camera (at infinity on the focal axis). The camera moves on a sphere around the object and is oriented toward the center. So the lighting environment is not fixed but lighting and viewing directions are related so the color $R(\mathbf{p}, \omega)$ is unique for a given viewing direction ω.

In the second scenario, the lighting environment is fixed. $R(\mathbf{p}, \omega)$ is called a Surface Light Field (SLF) [48] while the hemispherical function at a fixed $\mathbf{p}$ is called the outgoing radiance. $R(\mathbf{p}, \omega)$ is stored in a spherical harmonics basis for each vertex $\mathbf{p}$ and is further interpolated over the triangles [39].

4.2 Distance Measurements

We aim at measuring the distance between an original mesh M and a modified version M'. As shown in Figure 1, we are interested in the perceived difference for a given light interaction, that is, an error between the appearances R (original) and R' (modified). We first define a projection $\pi : M \to M'$ that maps any point $\mathbf{p} \in M$ onto $\pi\mathbf{p} \in M'$, defined as the closest point on M'. We then derive the following asymmetric error:

$$E^2(R \to R') = \frac{1}{Area(M)} \int_{\mathbf{p} \in M} \frac{1}{SolidAngle(\Omega_\mathbf{p})} \int_{\omega \in \Omega_\mathbf{p}} \|R(\mathbf{p}, \omega) - R'(\pi\mathbf{p}, \omega)\|^2 d\omega d\mathbf{p}. \quad (2)$$

That is, the squared color error is first integrated over the directions ω in the visible hemisphere $\Omega_\mathbf{p}$ at point $\mathbf{p}$ and then over the positions $\mathbf{p}$ on the mesh M. This is actually a Mean Square Error (MSE) $L_{2,2}(\Omega, M)$ over both the directional and spatial domains.

We experimented several variants of this error formula. First, we tried different metrics using the L_1 instead of L_2 norm over Ω (inner integral) and/or M (outer integral). Second, we tried different directional domains of integration as alternatives to the hemisphere $\Omega_\mathbf{p}$: (i) We integrated over the intersection of the hemispheres defined at $\mathbf{p}$ and $\pi\mathbf{p}$; (ii) we removed the occluded parts of $\Omega_\mathbf{p}$, that is, the viewing directions that intersect the object and thus represent viewing directions from which $\mathbf{p}$ is not visible; (iii) we restricted $\Omega_\mathbf{p}$ to the viewing directions that correspond to the exact camera path used in our experiments. All these variants systematically produced similar results, that is, the rankings and relative differences between error values computed for different modified versions were identical.

Numerical integration on Ω is based on Lebedev quadrature formulas. Numerical integration on M is performed by uniformly sampling each triangle as follows. Two steps of 1-to-4 subdivision provide 16 sub-triangles. For Phong shading and acquired data, $R(\mathbf{p}, \omega)$ is continuous (with respect to $\mathbf{p}$) at the vertices so the samples are the vertices of the sub-triangles. For flat shading, the discontinuities at edges and vertices cause numerical instability, so the samples are the centroids of the sub-triangles.

4.3 Results

In this section we evaluate the performance of our metric against two geometry-based approaches: the MSE, which is comparable to our metric but considers geometry only, and MSDM2 [17], which is one of the best performing perceptually motivated metrics. We also consider three image quality assessment metrics: SSIM [44], HDRVDP2 [25], and CID [31]. SSIM and HDRVDP2 consider the luminance only, whereas CID also takes the chrominance and hue into account. These image-based metrics have been applied on 30 images sampled regularly on each generated video and obtained values have been averaged.

The performance of these metrics is evaluated using the Kendall and Pearson correlations between the objective metric values and the subjective scores. The Pearson correlation is computed after a logistic regression which provides a non-linear mapping between the objective and subjective scores. Tables 4 and 5 detail the performance results. To assess the superiority of metrics among others, we conducted a pairwise Student's t-test on Spearman correlation values. Practically, for each dataset, we created 200 versions of the scores by

Table 4. Pearson Correlation Coefficients on Our Dataset

Model	type	Model-based			Image-based		
		Our metric	MSE	MSDM2	SSIM	CID	HDRVDP2
Dragon flat	synth.	0.87	0.90	0.84	0.90	0.90	0.83
Dragon Phong	synth.	0.90	0.49	0.57	0.99	0.96	0.97
LionVase flat	synth.	0.99	0.95	0.99	0.99	0.98	0.99
LionVase Phong	synth.	0.88	0.79	0.89	0.99	0.98	0.93
Average synthetic		**0.91±0.06**	0.78±0.20	0.82±0.18	**0.97±0.04**	0.95±0.04	0.93±0.07
AfricanMask	acq.	0.93	0.11	0.35	0.97	0.97	0.90
DoubleDragon	acq.	0.93	0.54	−0.60	0.77	0.87	0.88
Average acquired		**0.93±0.01**	0.33±0.31	−0.12±0.67	0.87±0.14	**0.92±0.07**	0.89±0.01

Table 5. Spearman Correlation Coefficients on Our Dataset

Model	type	Model based			Image based		
		Our metric	MSE	MSDM2	SSIM	CID	HDRVDP2
Dragon flat	synth.	0.71	0.89	0.60	0.89	0.89	0.89
Dragon Phong	synth.	0.89	0.49	0.83	0.94	0.94	0.94
LionVase flat	synth.	0.94	1.00	0.94	1.00	1.00	1.00
LionVase Phong	synth.	0.83	0.77	0.83	1.00	1.00	0.94
Average synthetic		**0.84±0.10**	0.79±0.22	0.80±0.14	**0.96±0.06**	**0.96±0.06**	0.94±0.05
AfricanMask	acq.	0.89	0.77	0.71	0.94	1.00	0.89
DoubleDragon	acq.	0.94	0.54	−0.60	0.77	0.89	0.94
Average acquired		**0.91±0.04**	0.66±0.16	0.06±0.93	0.86±0.12	**0.94±0.08**	0.91±0.04

Fig. 8. The 95% statistical similarity groups for both datasets. Computed with pairwise t-test (Bonferroni-corrected) between all metrics (using bootstrapped Spearman correlation). Details of all p-values are provided in an online supplementary document.

bootstrapping the subjects (selection with replacement). The resulting similarity groups at the 95% significance level are reported in Figure 8.

Our metric outperforms other mesh-based metrics (MSE and MSDM2) at a statistically significant level (see Table 5 and Figure 8) on both datasets. This means that taking into account the light-material interaction, even with a simple metric (our metric can be seen as a simple MSE integrated over the spatial and directional domains), improves greatly on geometric metrics (even sophisticated perceptual ones). As could be expected, the improvement is even larger on *acquired* data that exhibits colors.

Image-based metrics applied on rendered images provide better performance on *synthetic* data. However, they have been applied on the exact rendered images shown to the observers, that is, in very favorable conditions. In most real applications, the visual quality of the 3D model has to be evaluated without knowing the future camera parameters, image resolution, or rendering algorithm. For *acquired* data, the performance of our metric is very close to the best-performing image-based metric. This demonstrates the robustness of our approach that still works very well in such complex cases.

Note that the lower Spearman correlation given by our metric for the Dragon Flat model (0.71) is due to the broken order $S2 > S1$, which can be seen in the first row of Table 3. As stated in Section 3.5.2, it is due to the fact that both simplified versions appear of similar excellent quality. However, this inversion is very harmful for the Spearman correlation, which is based on ranks.

5 CONCLUSION

In this article, we investigated the impact of light-material interactions on the perception of geometric distortion of 3D models. Therefore, we considered the full pipeline that goes from acquisition or modeling to the rendering of final images. In this pipeline, early data processing that modifies (e.g., compresses) the geometry is a common task that can induce visual distortions possibly perceived as quality loss by the viewer. Being able to quantify this distortion is beneficial for tuning data modifications to achieve a lower visual impact. We conducted a user study to obtain quality evaluations from human perception for different types of light-material interactions and distortions and then elaborated on a metric that shows improved correlation with these measures.

Thanks to our user study, we were able to conclude that the interaction between light and material strongly influences how a given distortion is perceived. As it is impossible to build a perceptual metric for each material and light environment, we argue that such measures should be calculated in the space of directional colors when possible: The colors that 3D objects emit in all directions after having reflected incoming light. Hence, we built a metric in this space: measuring appearance differences in terms of directional color functions at each surface point. We show improved correlation compared to metrics that are unaware of light and material, that is, that use only intrinsic properties of the 3D model, like geometry or texture. Our metric even competes with image quality metrics that can be computed for particular camera parameters (i.e., viewpoint, projection, etc.) only.

This new metric can be used when the light environment and surface reflectance are known. A rendered mesh is necessarily computed with these parameters being given. However, many 3D shape visualizers render geometric meshes without questioning this part: a "generic" material and a point light source is often applied. Since we have discussed hypotheses for explaining the effects shown by our results, for example, that over-smoothness is an important artifact that humans perceive as very disturbing, we are one step closer to elaborating recommendations on how to properly visualize a mesh for geometric inspection. One could, for example, say that not interpolating normals (i.e., flat shading) and contrast enhancement (i.e., through glossiness) highlights geometric mesh details, thus conveying the shape better in the scenario of visual inspection. A more in-depth study would be beneficial for understanding and showing this in detail.

Finally, our metric can be used to measure distortions for tuning geometry processing. It would be worth investigating its use directly within the optimization. It is, however, still relatively simple and could be improved by incorporating more sophisticated statistics as used in image-based metrics, or higher-level phenomena like visual masking effects induced by texturing or lighting for example.

APPENDIX

A EXPERIMENTAL DATA

Acquired data have been provided by Vanhoey et al. [38]. We provide synthetic data (i.e., the mesh files) through online additional material: Together with the shader codes given below, our experiments can be reproduced. We also provide as online supplementary material (i) a set of image matrices containing reference and all degradations (as shown in, e.g., each row in Figure 3) of all six models, for both the synthetic and acquired datasets, and (ii) the p-values behind Figure 8.

FLAT vertex shader:

```
in vec3 vertex_pos;
uniform mat4 projection_matrix;
uniform mat4 model_view_matrix;
out vec3 pos;
void main()
{
   vec4 pos4 = model_view_matrix * vec4(vertex_pos, 1.0);
   pos = pos4.xyz;
   gl_Position = projection_matrix * pos4;
}
```

FLAT fragment shader:

```
in vec3 pos;
uniform vec4 front_color;
uniform vec4 back_color;
uniform vec4 ambiant_color;
uniform vec3 lightPosition;
out vec4 fragColor;
void main()
{
   vec3 N = normalize(cross(dFdx(pos), dFdy(pos)));
   vec3 L = normalize(lightPosition - pos);
   float lambert = dot(N, L);
   fragColor = vec4(ambiant_color.rgb + (lambert * front_color.rgb), front_color.a);
}
```

PHONG vertex shader:

```
in vec3 vertex_pos;
in vec3 vertex_normal;
uniform mat4 projection_matrix;
uniform mat4 model_view_matrix;
uniform mat3 normal_matrix;
uniform vec3 lightPosition;
out vec3 Normal;
out vec3 LightDir;
out vec3 EyeVector;
void main()
{
   Normal = normal_matrix * vertex_normal;
   vec4 pos4 = model_view_matrix * vec4(vertex_pos, 1.0);
   vec3 pos = pos4.xyz;
   LightDir = lightPosition - pos3;
   EyeVector = -pos3;
   gl_Position = projection_matrix * pos4;
};
```

PHONG fragment shader:

```
in vec3 Normal;
in vec3 LightDir;
in vec3 EyeVector;
uniform vec4 front_color;
uniform vec4 spec_color;
uniform vec4 ambiant_color;
uniform float spec_coef;
out vec4 fragColor;
void main()
{
   vec3 N = normalize(Normal);
   vec3 L = normalize(LightDir);
   vec4 finalColor = ambiant_color;
   float lambertTerm = clamp(dot(N, L), 0.0, 1.0);
   finalColor += front_color * lambertTerm;
   vec3 E = normalize(EyeVector);
   vec3 R = reflect(-L, N);
   float specular = pow(max(dot(R, E), 0.0), spec_coef);
   finalColor += spec_color * specular;
   fragColor = finalColor;
}
```

ACKNOWLEDGMENTS

We thank Oliver Mattausch for his useful comments and suggestions.

REFERENCES

[1] Tunç Ozan Aydin, Martin Čadík, Karol Myszkowski, and Hans-Peter Seidel. 2010. Video quality assessment for computer graphics applications. *ACM Trans. Graph.* 29, 6 (Dec. 2010), 1.

[2] Francesco Banterle, Patrick Ledda, Kurt Debattista, Marina Bloj, Alessandro Artusi, and Alan Chalmers. 2009. A psychophysical evaluation of inverse tone mapping techniques. *Comput. Graph. Forum* 28, 1 (2009), 13–25.

[3] Martin Čadík, Robert Herzog, Rafal Mantiuk, Radoslaw Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. 2013. Learning to predict localized distortions in rendered images. *Comput. Graph. Forum* 32, 7 (2013), 401–410. DOI: https://doi.org/10.1111/cgf.12248

[4] Massimiliano Corsini, Elisa Drelie Gelasca, Touradj Ebrahimi, and Mauro Barni. 2007. Watermarked 3-D mesh quality assessment. *IEEE Trans. Multimedia* 9, 2 (Feb. 2007), 247–256.

[5] Massimiliano Corsini, Mohamed-Chaker Larabi, Guillaume Lavoué, Oldrich Petrík, Libor Vása, and Kai Wang. 2013. Perceptual metrics for static and dynamic triangle meshes. *Comput. Graph. Forum* 32, 1 (Feb. 2013), 101–125. DOI: https://doi.org/10.1111/cgf.12001 Improved version of Eurographics State-of-the-Art Report.

[6] Scott Daly. 1993. The visible differences predictor: An algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*, Andrew B. Watson (Ed.). MIT Press, Cambridge, 179–206.

[7] Jiří Filip, Michael J. Chantler, Patrick R. Green, and Michal Haindl. 2008. A psychophysically validated metric for bidirectional texture data reduction. *ACM Trans. Graph.* 27, 5, Article 138 (Dec. 2008), 11 pages. DOI: https://doi.org/10.1145/1409060.1409091

[8] Jirí Filip, Radomír Vávra, Michal Havlícek, and Mikulás Krupicka. 2017. Predicting visual perception of material structure in virtual environments. *Comput. Graph. Forum* 36, 1 (2017), 89–100. DOI: https://doi.org/10.1111/cgf.12789

[9] Adria Fores, James Ferwerda, and Jinwei Gu. 2012. Toward a perceptually based metric for BRDF modeling. In *Proceedings of the 20th Color and Imaging Conference*. 142–148.

[10] F. Gao, D. Tao, X. Gao, and X. Li. 2015. Learning to rank for blind image quality assessment. *IEEE Trans. Neur. Netw. Learn. Syst.* 26, 10 (Oct. 2015), 2275–2290. DOI: https://doi.org/10.1109/TNNLS.2014.2377181

[11] Michael Garland and Paul S. Heckbert. 1998. Simplifying surfaces with color and texture using quadric error metrics. In *Proceedings of the Conference on Visualization'98 (VIS'98)*. IEEE Computer Society Press, Los Alamitos, CA, 263–269.

[12] Jinjiang Guo, Vincent Vidal, Irene Cheng, Anup Basu, Atilla Baskurt, and Guillaume Lavoue. 2016. Subjective and objective visual quality assessment of textured 3D meshes. *ACM Trans. Appl. Percept.* 14, 2, Article 11 (Oct. 2016), 20 pages. DOI: https://doi.org/10.1145/2996296

[13] Michael Guthe, Gero Müller, Martin Schneider, and Reinhard Klein. 2009. BTF-CIELab: A perceptual difference measure for quality assessment and compression of BTFs. *Comput. Graph. Forum* 28, 1 (Mar. 2009), 101–113.

[14] Robert Herzog, Martin Čadík, Tunç O. Aydın, Kwawng In Kim, Karol Myszkowski, and Hans-Peter Seidel. 2012. NoRM: No-reference image quality metric for realistic image synthesis. *Comput. Graph. Forum* 31, 2 (2012), 545–554. DOI: https://doi.org/10.1111/j.1467-8659.2012.03055.x

[15] Adrian Jarabo, Hongzhi Wu, Julie Dorsey, Holly Rushmeier, and Diego Gutierrez. 2014. Effects of approximate filtering on the appearance of bidirectional texture functions. *IEEE Trans. Visualiz. Comput. Graph.* XX, Xx (2014), 1–14.

[16] Maurice G. Kendall and B. Babington Smith. 1940. On the method of paired comparisons. *Biometrika* 31, 3/4 (1940), 324–345.

[17] Guillaume Lavoué. 2011. A multiscale metric for 3D mesh visual quality assessment. *Comput. Graph. Forum* 30, 5 (2011), 1427–1437.

[18] Guillaume Lavoué, Mohamed-Chaker Larabi, and Libor Vasa. 2016. On the efficiency of image metrics for evaluating the visual quality of 3D models. *IEEE Trans. Visualiz. Comput. Graph.* 22, 8 (Aug. 2016), 12.

[19] Patrick Ledda, Alan Chalmers, Tom Troscianko, and Helge Seetzen. 2005. Evaluation of tone mapping operators using a high dynamic range display. *ACM Trans. Graph.* 24, 3 (Jul. 2005), 640.

[20] Frédéric B. Leloup, Michael R. Pointer, Philip Dutré, and Peter Hanselaer. 2010. Geometry of illumination, luminance contrast, and gloss perception. *Int. J. Occupat. Safety Ergonom.* 27, 9 (2010), 2046–2054.

[21] Peter Lindstrom and Greg Turk. 2000. Image-driven simplification. *ACM Trans. Graph.* 19, 3 (Jul. 2000), 204–241. DOI: https://doi.org/10.1145/353981.353995

[22] Yiming Liu, Jue Wang, Sunghyun Cho, Adam Finkelstein, and Szymon Rusinkiewicz. 2013. A no-reference metric for evaluating the quality of motion deblurring. *ACM Trans. Graph.* 32, 6, Article 175 (Nov. 2013), 12 pages. DOI: https://doi.org/10.1145/2508363.2508391

[23] Jeffrey Lubin. 1993. The use of psychophysical data and models in the analysis of display system performance. In *Digital Images and Human Vision*, A. B. Watson (Ed.). 163–178.

[24] Jeffrey Lubin. 1995. A visual discrimination model for imaging system design and evaluation. *Vis. Models Target Detect. Recogn.* 2 (1995), 245–357.

[25] Rafat Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. 2011. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. In *ACM SIGGRAPH 2011 Papers (SIGGRAPH'11)*. ACM, New York, NY, Article 40, 14 pages. DOI: https://doi.org/10.1145/1964921.1964935

[26] Rafał K. Mantiuk, Anna Tomaszewska, and Radosław Mantiuk. 2012. Comparison of four subjective methods for image quality assessment. *Comput. Graph. Forum* 31, 8 (Dec. 2012), 2478–2491.

[27] Nicolas Menzel and Michael Guthe. 2010. Towards perceptual simplification of models with arbitrary materials. *Comput. Graph. Forum* 29, 7 (2010), 2261–2270.

[28] Georges Nader, Kai Wang, Franck Hétroy-Wheeler, and Florent Dupont. 2016. Just noticeable distortion profile for flat-shaded 3D mesh surfaces. *IEEE Trans. Visualiz. Comput. Graph.* 22, 11 (Nov. 2016), 2423–2436. DOI: https://doi.org/10.1109/TVCG.2015.2507578

[29] Yixin Pan, I. Cheng, and A. Basu. 2005. Quality metric for approximating subjective evaluation of 3-D objects. *IEEE Trans. Multimedia* 7, 2 (apr 2005), 269–279.

[30] Bui Tuong Phong. 1975. Illumination for computer generated pictures. *Commun. ACM* 18, 6 (June 1975), 311–317. DOI: https://doi.org/10.1145/360825.360839

[31] Jens Preiss, Felipe Fernandes, and Philipp Urban. 2014. Color-image quality assessment: From prediction to optimization. *IEEE Trans. Image Process.* 23, 3 (2014), 1366–1378.

[32] L. Qu and G. W. Meyer. 2008. Perceptually guided polygon reduction. *IEEE Trans. Visualiz. Comput. Graph.* 14, 5 (Sept 2008), 1015–1029. DOI: https://doi.org/10.1109/TVCG.2008.51

[33] Filippo Stanco, Sebastiano Battiato, and Giovanni Gallo. 2011. *Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks*. CRC Press.

[34] Louis L. Thurstone. 1927. A law of comparative judgment. *Psychol. Rev.* 34, 4 (1927), 273.

[35] Dihong Tian and G. AlRegib. 2008. Batex3: Bit allocation for progressive transmission of textured 3-d models. *IEEE Trans. Circ. Syst. Vid. Technol.* 18, 1 (2008), 23–35.

[36] Fakhri Torkhani, Kai Wang, and Jean-Marc Chassery. 2012. A curvature tensor distance for mesh visual quality assessment. In *Proceedings of the International Conference on Computer Vision and Graphics (ICCVG 2012)*, Vol. 7594. Springer-Verlag, New York, NY, 253–263. DOI: https://doi.org/10.1007/978-3-642-33564-8_31

[37] Fakhri Torkhani, Kai Wang, and Jean Marc Chassery. 2015. Perceptual quality assessment of 3D dynamic meshes: Subjective and objective studies. *Sign. Process. Image Commun.* 31 (2015), 185–204.

[38] K. Vanhoey, B. Sauvage, O. Génevaux, F. Larue, and J.-M. Dischler. 2013. Robust fitting on poorly sampled data for surface light field rendering and image relighting. *Comput. Graph. Forum* 32, 6 (2013), 101–112. DOI: https://doi.org/10.1111/cgf.12073

[39] Kenneth Vanhoey, Basile Sauvage, Pierre Kraemer, Frédéric Larue, and Jean-Michel Dischler. 2015. Simplification of meshes with digitized radiance. *Vis. Comput.* 31, 6–8 (2015), 1011–1021. DOI: https://doi.org/10.1007/s00371-015-1124-9

[40] Libor Váša and Jan Rus. 2012. Dihedral angle mesh error: A fast perception correlated distortion measure for fixed connectivity triangle meshes. *Comput. Graph. Forum* 31, 5 (2012), 1715–1724. DOI: https://doi.org/10.1111/j.1467-8659.2012.03176.x

[41] Bruce Walter, Sumanta N. Pattanaik, and Donald P. Greenberg. 2002. Using perceptual texture masking for efficient image synthesis. In *Computer Graphics Forum*, Vol. 21. Wiley Online Library, 393–399.

[42] Kai Wang, Fakhri Torkhani, and Annick Montanvert. 2012. A fast roughness-based approach to the assessment of 3D mesh visual quality. *Comput. Graph.* 36, 7 (Nov. 2012), 808–818. DOI: https://doi.org/10.1016/j.cag.2012.06.004

[43] Zhou Wang and Alan C. Bovik. 2006. *Modern Image Quality Assessment.* Morgan & Claypool. DOI: https://doi.org/10.2200/S00010ED1V01Y200508IVM003

[44] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *Trans. Img. Proc.* 13, 4 (April 2004), 600–612. DOI: https://doi.org/10.1109/TIP.2003.819861

[45] Z. Wang, E. P. Simoncelli, and A. C. Bovik. 2003. Multi-scale structural similarity for image quality assessment. *Conf. Rec. Asilomar Conf. Sign. Syst. Comput.* 2, 1398–1402.

[46] Benjamin Watson, Alinda Friedman, and Aaron McGaffey. 2001. Measuring and predicting visual fidelity. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'01).* ACM, New York, NY, 213–220. DOI: https://doi.org/10.1145/383259.383283

[47] Nathaniel Williams, David Luebke, Jonathan D. Cohen, Michael Kelley, and Brenden Schubert. 2003. Perceptually guided simplification of lit, textured meshes. In *Proceedings of the 2003 Symposium on Interactive 3D Graphics (I3D'03).* ACM, New York, NY, 113–121. DOI: https://doi.org/10.1145/641480.641503

[48] Daniel N. Wood, Daniel I. Azuma, Ken Aldinger, Brian Curless, Tom Duchamp, David H. Salesin, and Werner Stuetzle. 2000. Surface light fields for 3D Photography. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'00).* ACM Press/Addison-Wesley Publishing Co., New York, NY, 287–296. DOI: https://doi.org/10.1145/344779.344925

[49] Hojatollah Yeganeh and Zhou Wang. 2013. Objective quality assessment of tone-mapped images. *IEEE Trans. Image Process.* 22, 2 (2013), 657–667. DOI: https://doi.org/10.1109/TIP.2012.2221725

Received January 2017; revised April 2017; accepted June 2017