



Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces

Georges Nader, Kai Wang, Franck Hétroy-Wheeler, Florent Dupont

► To cite this version:

Georges Nader, Kai Wang, Franck Hétroy-Wheeler, Florent Dupont. Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 2016, 22 (11), pp.2423-2436. 10.1109/TVCG.2015.2507578 . hal-01242271

HAL Id: hal-01242271

<https://hal.science/hal-01242271>

Submitted on 19 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces

Georges Nader, Kai Wang, Franck Hétroy-Wheeler, and Florent Dupont

Abstract—It is common that a 3D mesh undergoes some lossy operations (e.g., compression, watermarking and transmission through noisy channels), which can introduce geometric distortions as a change in vertex position. In most cases the end users of 3D meshes are human beings; therefore, it is important to evaluate the visibility of introduced vertex displacement. In this paper we present a model for computing a Just Noticeable Distortion (JND) profile for flat-shaded 3D meshes. The proposed model is based on an experimental study of the properties of the human visual system while observing a flat-shaded 3D mesh surface, in particular the *contrast sensitivity function* and *contrast masking*. We first define appropriate local perceptual properties on 3D meshes. We then detail the results of a series of psychophysical experiments where we have measured the threshold needed for a human observer to detect the change in vertex position. These results allow us to compute the JND profile for flat-shaded 3D meshes. The proposed JND model has been evaluated via a subjective experiment, and applied to guide 3D mesh simplification as well as to determine the optimal vertex coordinates quantization level for a 3D model.

Index Terms—Just noticeable distortion, human visual system, psychophysical experiments, contrast sensitivity function, contrast masking, 3D mesh

1 INTRODUCTION

THREE-DIMENSIONAL (3D) meshes may be subject to various lossy operations such as compression and watermarking. These operations introduce geometric distortions in form of vertex displacement. Since computer graphics applications are intended towards human subjects, it is important to study the visibility of those distortions, taking into account the properties of the human visual system. Geometric measures like Hausdorff distance or root mean square error (RMS) [1], [2] reflect the physical variation of the mesh geometry. They do not correlate with the human vision [3] and therefore cannot be used to predict whether a distortion is visible or not. The visibility of the geometric distortions is also affected by the lighting conditions, the viewpoint, the surface's material and the rendering algorithm. Figure 1 gives two examples of how the scene parameters can affect the visibility of vertex noise. A slightly visible noise is injected onto the 3D model (Fig. 1.(a)). When the viewing distance is increased, the noise becomes invisible (Fig. 1.(b)). When the light direction is changed from front to top-left, the noise on the mesh becomes more visible (Fig. 1.(c)).

The Just Noticeable Distortion (JND) threshold refers to the threshold above which a distortion becomes visible to the majority of observers [4]. If a distortion is below the JND value, it can be considered as invisible. JND models reflect the local properties of the human visual system, in particular the *contrast sensitivity function* and *contrast masking*. On one hand, in image/video processing, JND models of 2D

images and videos have proven to be helpful for various applications such as evaluating the image visual fidelity [5] and optimizing image compression [6], [7], [8]. On the other hand, while perceptually driven graphics are popular within the computer graphics community [3], [9], [10], there has been little effort given to study the visibility of vertex displacement and further to compute a JND profile for 3D models.

This is exactly what this paper proposes. More specifically, our contributions are the following:

- 1) We define local perceptual properties that are appropriate for a "bottom-up" evaluation of vertex displacement visibility on 3D meshes.
- 2) We design and conduct psychophysical experiments to study properties of the human visual system when observing a flat-shaded 3D mesh.
- 3) We propose a JND profile for 3D meshes that takes into consideration the various circumstances of mesh usage (display size, scene illumination, viewing distance).

The rest of this paper is organized as follows: Section 2 briefly introduces the properties of the human visual system that are essential to compute the JND profile, and discusses existing work on perceptually driven graphics techniques. Section 3 explains how perceptual properties are evaluated on a 3D mesh and presents a series of psychophysical experiments that were carried out in order to measure the visibility threshold and their results. Section 4 describes our method to compute the JND profile for a 3D mesh. In Section 5 we evaluate the proposed JND model's performance via subjective experiments. In Section 6 we apply our JND profile to guide mesh simplification and to automatically determine the optimal vertex coordinates quantization level. Finally, we discuss the limitations and conclude the paper.

- G. Nader and F. Dupont are with Université de Lyon, LIRIS UMR 5205 CNRS, France. E-mail: {georges.nader, florent.dupont}@liris.cnrs.fr
- K. Wang is with CNRS and Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble, France. E-mail: kai.wang@gipsa-lab.grenoble-inp.fr
- F. Hétroy-Wheeler is with Univ. Grenoble Alpes, LJK, F-38000 Grenoble, France and with Inria. E-mail: franck.hetroy@grenoble-inp.fr

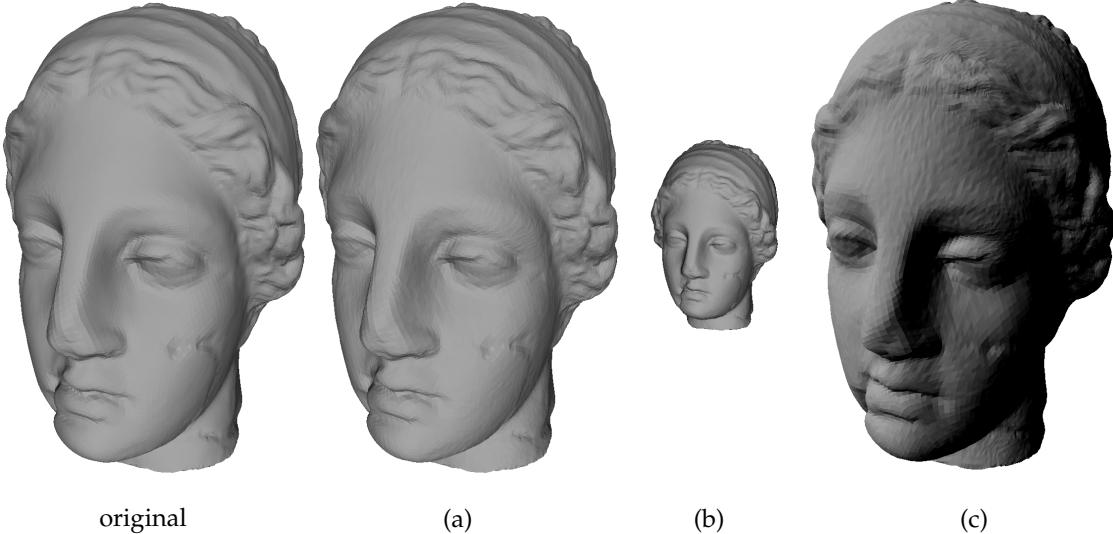


Fig. 1. (a) The noise injected onto the original 3D mesh is slightly visible. (b) Increasing the viewing distance makes the noise invisible. (c) Changing the light direction from front to top-left increases the perceived intensity of the noise.

2 RELATED WORK

Before presenting the existing works on perceptually driven graphics, we introduce the local properties of the human visual system on which a large number of the existing methods rely.

2.1 Local Properties of the Human Visual System

The visibility of a visual pattern depends on its local *contrast* and its *spatial frequency* [11]. Local contrast refers to the change of light intensity over the light intensity of its surrounding. The spatial frequency is defined as the size of light patterns on the retina. In the field of visual perception, the spatial frequency is expressed in terms of cycles per degree (cpd) which represents the number of light patterns on the retina in 1 degree of the visual angle. To show how local contrast and spatial frequency affect the visibility of a pattern, we briefly introduce the notions of *contrast sensitivity* and *contrast masking*. Both these concepts describe the basic human vision mechanisms and were largely exploited for the development of many image processing methods [12], [13] especially for JND profiles of 2D images [4]. We refer to [11], [14] for a more detailed treatment of the human visual perception.

Contrast sensitivity. A visual pattern can be detected by the visual system only if the amount of contrast in the pattern is above some threshold. This visibility threshold varies under different spatial frequencies of the visual pattern. This is mainly due to the optics in the eye and the size of the photoreceptors on the retina. The reciprocal of this detection threshold is the contrast sensitivity. The contrast sensitivity function (CSF) describes the visibility threshold with respect to the spatial frequency. The CSF represents the visual system's band-pass filter characteristics when it comes to contrast sensitivity. In general, it exhibits a peak at around 2 to 5 cpd then drops off to a point where no detail can be resolved. The shape of the CSF (peak location and drop off slope) depends on the nature of the visual stimulus (complexity, periodicity) [15]. Since the CSF is the

main component in many perceptual models [12], [13], there has been a great interest for measuring it in different circumstances [16]. Most notably, let us cite the Modelfest project [17], in which 10 laboratories collaborated on measuring the contrast threshold for 43 different types of visual stimuli.

Contrast masking. Contrast masking is the change in visibility of a stimulus (target) due to the presence of another stimulus (mask). The visibility of the target depends on different factors, in particular the spatial frequency and the contrast of the mask and the target. Legge and Foley [18] studied the contrast threshold necessary to detect the target when varying the contrast and frequency of the mask. One important observation was that this contrast threshold is increased when the contrast of the mask is increased. The effects of contrast masking can be described by a curve which possesses two asymptotic regions: the first one with a slope of zero and the second one with a positive slope of about 0.6 to 1 (depending on the stimulus) [18]. The zero slope occurs for mask contrast values below the mask's visibility threshold as given by the CSF, indicating that there is no masking effect. By contrast, as the mask contrast value increases beyond the CSF threshold in the second asymptotic region, the threshold for detecting the target increases too.

Despite being essential for the development of perceptual image processing algorithms, the *contrast sensitivity function* and the *contrast masking* have, to the best of our knowledge, never been carefully studied in the 3D setting.

2.2 Perceptually Driven Graphics

Over the last two decades, perceptually driven methods have drawn more and more attention in the computer graphics community. These methods have proven to be useful for evaluating the quality of 3D models and optimizing graphics applications. Perceptually driven methods can be divided into two groups: *top-down* methods and *bottom-up* ones.

Top-down methods. Top-down techniques do not rely on

the exact internal mechanism of the human visual system. Such techniques rather propose hypotheses, which are usually difficult to prove, about the overall behavior of the visual system in order to estimate how a specific visual artifact is perceived. Based on the observation that visual artifacts are less visible on rough regions than on smooth ones of a 3D mesh [19], several perceptual metrics have for instance been proposed [20], [21], [22]. Other features used by such top-down metrics include surface curvature [23], [24] and dihedral angle [25]. Perceptual methods have also been used to guide mesh simplification [26], [27], [28], [29] and compression [30]. In the previously mentioned methods, the perceptual analysis is carried out on the geometry of a 3D mesh. In general, these methods are neither easily applicable to models of different properties (size, details and density) nor capable of adapting to varying circumstances of mesh usage (display characteristics, scene illumination and viewing distance). In addition, they mainly study distortions that are above the visibility threshold. Near-threshold vertex displacement visibility is not the focus of these methods.

Bottom-up methods. A bottom-up approach explicitly takes into account the mathematical models describing the mechanisms of the human visual system. Bottom-up methods have been popular in computer graphics. They are usually based on the concepts of *contrast sensitivity* and *contrast masking* (Section 2.1). One of the most popular methods in image processing is Daly's Visual Difference Predictor (VDP) [31]. This algorithm outputs a map describing the visibility of the difference between two images. The VDP has recently been extended to 2D HDR images in [32]. Based on Daly's VDP algorithm, Ramasubramanian *et al.* [33] computed a 2D threshold map in which each pixel contains the value of the minimum detectable difference. This map is used to guide global illumination computations. This threshold model has later been improved in [34]. Bottom-up analysis has also been used for mesh simplification applications [35], [36], [37]. However, the visibility analysis in those methods is still carried out in the 2D space of rendered images.

In this paper, different from all the methods mentioned in the previous paragraph that conduct the visibility analysis in a 2D space, we present a method for studying the visibility of vertex displacement in the 3D space. This method allows us to define a model for computing a JND profile for 3D meshes. To the best of our knowledge, the only existing JND-like model for a 3D mesh is the one of Cheng *et al.* [38], [39]. However, their goal and approach are very different from those of our method. Their goal was to evaluate the visibility of removing a group of vertices from one level of detail (LOD) to another. To do so, they used a top-down approach where they assumed that the visibility of removing vertices is related to the change in distance to a predefined model skeleton. The limitation of this work is that the JND depends on the predefined skeleton and is only applicable for evaluating LOD techniques. By contrast, our method for computing the JND model is based on a bottom-up experimental study of the properties of the human visual system and explicitly takes into account its internal mechanisms. The proposed JND model can cope with different mesh properties (size, density) and different possible usage of a mesh (illumination, display characteristics).

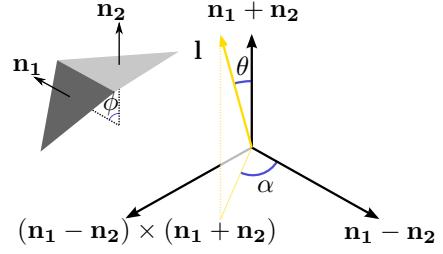


Fig. 2. The contrast between adjacent faces is computed using the angle between their normals and the spherical coordinates of the light direction in the local coordinate system defined by the face normals.

3 LOCAL PERCEPTUAL PROPERTIES AND PSYCHOPHYSICAL EXPERIMENTS

Existing top-down perceptual methods use surface roughness [20], [21], surface curvature [23], [24], dihedral angle [25] and face normal [40] as perceptually-relevant features of a 3D mesh. In this section we define new local perceptual properties for 3D meshes (*i.e.*, local contrast and spatial frequency) that are appropriate for a bottom-up evaluation of vertex displacement visibility. These perceptual properties allow us to study the effects of the contrast sensitivity and the contrast masking in the 3D setting. In the following, we start by explaining how the local contrast (Section 3.1) and the local spatial frequency (Section 3.2) are evaluated on a 3D mesh. We then present a series of psychophysical experiments that were conducted to measure the detection threshold (Section 3.3). In this study we consider 3D meshes that are rendered with a flat shading algorithm. We also limit our study to diffuse surfaces illuminated by a white directional light source.

3.1 Local Contrast Estimation

As explained above, contrast is the measure of difference of luminance. In the case of a flat-shaded rendering, each face of the 3D mesh is attributed a luminance value proportional to the cosine of the angle between its normal and the light direction. The luminance of a face is given by:

$$L = \max(1 \cdot \mathbf{n}, 0), \quad (1)$$

where \mathbf{n} is the unit face normal and 1 is the light direction. The Michelson contrast between two adjacent faces can then be defined by:

$$c = \frac{\| L_1 - L_2 \|}{L_1 + L_2} = \frac{\| \max(1 \cdot \mathbf{n}_1, 0) - \max(1 \cdot \mathbf{n}_2, 0) \|}{\max(1 \cdot \mathbf{n}_1, 0) + \max(1 \cdot \mathbf{n}_2, 0)}, \quad (2)$$

where \mathbf{n}_1 and \mathbf{n}_2 are the normals of the two adjacent faces. Under the circumstances where the inner products between the light direction and the two face normals are both positive, the above equation yields to the following equation:

$$c = \left\| \cos \alpha \times \tan \theta \times \tan \frac{\phi}{2} \right\|, \quad (3)$$

where α and θ are the spherical coordinates of the light direction in the local coordinate system defined by $\mathbf{n}_1 - \mathbf{n}_2$, $\mathbf{n}_1 + \mathbf{n}_2$ and their outer product (see Fig. 2). ϕ is the angle between the normals of the two faces.

Equation (3) shows how the contrast is affected by surface

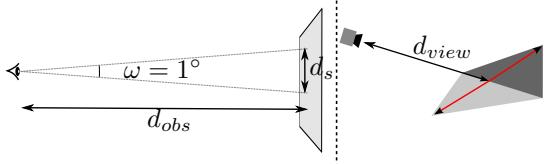


Fig. 3. The spatial frequency of a visual stimulus depends on the observer's distance (d_{obs}) to the screen and the view distance (d_{view}) of the virtual camera to the 3D model.

geometry and the scene illumination. The term $\tan \frac{\phi}{2}$ indicates the impact of surface geometry on the local contrast. On one hand, if the surface is locally smooth ($\phi \approx 0^\circ$), then the local contrast is minimal. On the other hand, if the surface is locally rough ($\phi \gg 0^\circ$), then the local contrast tends to be high. In addition, the term $\cos \alpha \times \tan \theta$ describes how the light direction affects the local contrast. A grazing light direction will maximize the value of the contrast where θ is close to 90° and α is close to 0° or 180° , while a light direction close to the normal direction ($\theta \approx 0^\circ$) makes the contrast minimal.

3.2 Local Frequency Estimation

Spatial frequency refers to the size of a visual stimulus on the retina. It is expressed in cycles per degree (cpd) [11]. The spatial frequency is affected by the physical size of the object and the observer's distance to the object. In our case, the visual stimulus is displayed on a screen and consists of the difference in luminance between a pair of adjacent faces. The perceived size of this stimulus depends then on the display's properties (resolution and size), the observer's distance to the display, the position of the model in the virtual 3D world and the size of the faces (see Fig. 3). As a consequence, in order to compute the spatial frequency we first need to evaluate the number of pixels that are occupied by the pair of adjacent faces. This is achieved by applying the perspective projection to the opposing vertices of the pair of adjacent faces. We then convert the computed number of pixels to cpd using the following approximation:

$$f = \frac{d_s}{n_{px}/ppcm} = \frac{2d_{obs} \times \tan \frac{1}{2} \frac{\pi}{180}}{n_{px}/ppcm} \approx \frac{d_{obs}}{n_{px}/ppcm} \times \frac{\pi}{180}, \quad (4)$$

where d_s is the spread of 1 cpd on the screen (see Fig. 3), d_{obs} is the observer's distance to the screen in cm, n_{px} is the number of pixels occupied by the visual stimulus obtained by applying perspective projection, and $ppcm$ is the number of pixels in 1 cm of the screen, computed as:

$$ppcm = \frac{\sqrt{r_X^2 + r_Y^2}}{s}, \quad (5)$$

with r_X the horizontal resolution in pixels, r_Y the vertical one and s the diagonal length of the screen in cm. $n_{px}/ppcm$ represents the size of the displayed stimulus in cm. The density of the 3D mesh is related to the size of its faces. This implies that a dense mesh will display high frequency visual stimuli while a coarse mesh will show low frequency ones under the same observation condition.

3.3 Threshold Measurement

Having defined the local contrast and spatial frequency on the 3D mesh, we now explain how to measure the local contrast threshold required to detect a change in the mesh.

3.3.1 Contrast Sensitivity Function

As detailed above, the contrast sensitivity function describes the human visual system's ability to detect contrast at different frequencies. The threshold given by the CSF refers to the amount of contrast required to distinguish a stimulus from its uniform surrounding (surrounding contrast is 0).



Fig. 4. Visual stimulus to measure the contrast sensitivity. Left: the reference plane. Right: a regular plane where a vertex is displaced.

Stimulus. In order to be able to measure the CSF in the 3D setting, the natural visual stimulus consists of a vertex displaced from the surface of a regular plane whose contrast is 0 (Fig. 4). The displacement of the vertex alters the normal of the adjacent faces and thus changes the contrast. In order to measure the threshold of different frequencies we change the density of the plane, which alters the size of its faces. The threshold is measured for 8 spatial frequencies (1.12, 2, 2.83, 4, 5.66, 8, 11.30 and 16 cpd, also considered in the Modelfest project [17]). The plane is tilted by 20° to give the observer a 3D feel.

Experimental setup. Experiments took place in a laboratory environment. The stimuli were displayed on an Asus 23-inch display in a low illuminated room. Screen resolution was 1920×1080 . The stimuli were observed from a distance of 1 m, which allowed us to measure the threshold for frequencies between 1 and 16 cpd.

Method. Two planes were displayed side by side on the screen, one of which exhibits a displaced vertex in its central area. The participants were then asked to answer by Yes or No whether they can see any difference between the displayed planes. According to our experience, this method is faster and less tiring for inexperienced subjects than 2AFC methods. We used the QUEST procedure [41] with a fixed number of trials (20 trials) to find the threshold. Each participant repeated the experiment 4 times each on a different day. An additional "dummy" frequency, whose data were not taken into account, was included at the beginning of each session to stabilize the subject's answers. In order to avoid any bias, frequency order was randomized for each observer in each session. No user interaction was allowed.

Participants. 5 subjects (3 males and 2 females) participated in our experiments. All had normal or corrected-to-normal vision and were 22 to 26 years old. One of the participants was experienced in perceptual subjective evaluations and the other 4 were inexperienced.

Results. The results of this experiment are shown in Fig. 5. The displacement of a vertex causes a variation in contrast for multiple face pairs. We save the maximum contrast between the affected face pairs. The left panel of Fig. 5 plots the

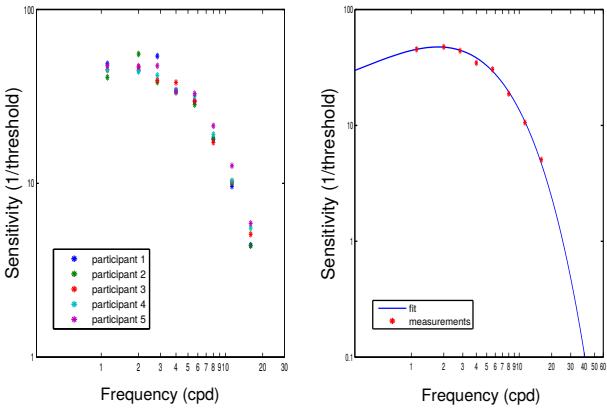


Fig. 5. Left: plot of the mean sensitivity for each observer over each frequency. Right: plot of the subjects' mean sensitivity over each frequency fitted using Manos and Sakrison's mathematical model.

mean sensitivity for each observer over each frequency. The plot shows a high consistency between the participants: All of them exhibit a peak in sensitivity at 2 cpd and the drop off in sensitivity on either side of the peak is similar for all participants. The right panel of Fig. 5 shows the subjects' mean sensitivity over each frequency, fitted using Mannos and Sakrison's mathematical model [42] that is defined by:

$$\text{csf}(f) = \left(1 - a + \frac{f}{f_0}\right) e^{-f^p}, \quad (6)$$

with $a = -15.13$, $f_0 = 0.0096$ and $p = 0.64$. The fit predicts a peak in sensitivity at around 2 cpd that drops rapidly at high frequencies. At low frequencies the drop in sensitivity is much slower than the one measured with a 2D contrast grating [15], [17]. This is due to the aperiodic nature of the visual stimulus [15]. Equation (6) shows the relation between the density of a mesh and the visibility of a local contrast alteration. As the density increases from a very low value, it would be slightly easier for the human visual system to notice the local contrast alteration on the mesh's surface until it reaches a certain limit (around 2 cpd) beyond which it would be harder to detect contrast alteration as the density increases.

3.3.2 Contrast Masking

Contrast masking refers to the ability of the human visual system to discriminate between two visible visual stimuli (a *target* and a *mask*). Since the visibility of a visual stimulus depends on its spatial frequency, the contrast masking threshold is different at each frequency. However, if we normalize the threshold values by the mask's CSF value, then the resulting threshold will be independent of the stimulus's spatial frequency [31]. Therefore, measuring the masking effect can be done by only changing the contrast value of a mask signal without the need to pay too much attention to its spatial frequency. Nevertheless, in our preliminary tests, we measured the normalized contrast masking effects for 3 different frequencies and found that the results were indeed the same (as stated in [31]), showing that measuring the contrast masking effect can be done independently from the spatial frequency of the visual stimulus.

Stimulus and method. In order to be able to measure

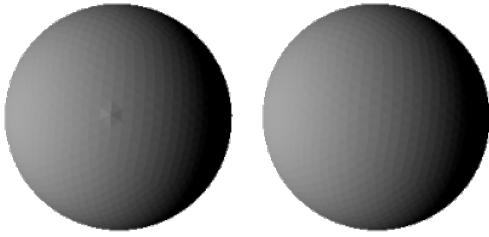


Fig. 6. Visual stimulus for measuring contrast masking. Left: a sphere approximated by an icosahedron subdivided 3 times from which a vertex is displaced. Right: the reference sphere.

the threshold relative to the contrast masking effect, the visual stimulus needs to exhibit a visible initial contrast (*i.e.*, above the CSF value). We then increase the initial contrast and measure the value needed to notice that change. In other words, if c is the initial contrast (mask signal) and c' is the increased value, we measure $\Delta c = c' - c$ (target signal) needed to discriminate between c and c' . Similarly to the method used for measuring the CSF, two models were displayed on the screen and the participants were supposed to decide whether they saw the difference between the two objects or not. The stimulus consists of a vertex displaced from a sphere approximated by a subdivided icosahedron (Fig. 6). The icosahedron is subdivided 3 times, which makes the contrast between two adjacent faces (stimulus of about 2 cpd) visible for an observer. This initial contrast represents the mask signal. Varying the light direction modifies the value of the initial contrast between two adjacent faces. We measured the threshold relative to 7 mask contrasts that were log-linearly spaced from 0.6 to 4 times the CSF threshold.

Experimental setup. The same experimental setup and the same method than for the CSF measurement experiment previously described have been used. The same 5 subjects also participated in the contrast masking experiments.

Results. The results of this experiment are shown in Fig. 7. The left panel plots for every participant the mean normalized threshold over the normalized contrast mask. For mask contrasts below the visibility threshold (normalized contrast mask lower than 1), the measured normalized threshold is close to 1. This indicates that the measured threshold refers to the one given by the CSF and that no masking has occurred. For mask contrasts above the visibility threshold, the measured normalized threshold is above the one given by CSF and lies close to the asymptotic region with a slope near 0.7. The right panel of Fig. 7 shows the subjects' mean threshold over each mask contrast fitted using Daly's mathematical masking model [31] that is defined by:

$$\text{masking}(\tilde{c}) = \left(1 + (k_1 \times (k_2 \times \tilde{c})^s)^b\right)^{1/b}, \quad (7)$$

with \tilde{c} the normalized threshold, and the fitted values $k_1 = 0.0078$, $k_2 = 88.29$, $s = 1.00$ and $b = 4.207$. The fit exhibits the two asymptotic regions that characterize the contrast masking effect with a transition between the two regions at the CSF visibility threshold. To some extent, Eq. (7) shows how an increase in surface roughness can hide local geometric distortions on the mesh's surface. A rough

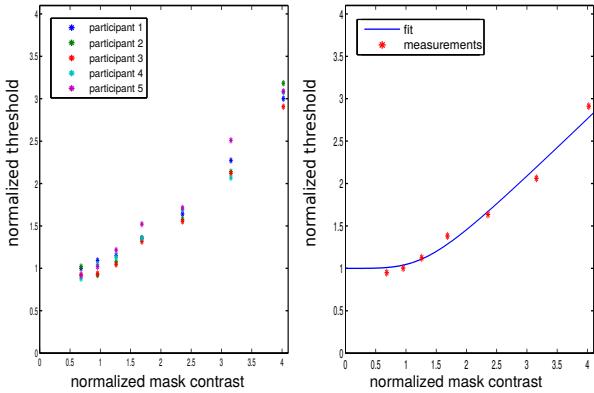


Fig. 7. Left: plot of the normalized mean threshold for each observer over normalized mask contrast. Right: plot of the subjects' mean normalized threshold over each normalized mask contrast, fitted using Daly's mathematical contrast masking model.

surface in general exhibits more contrast and thus is more likely to be able to mask the visibility of a local distortion.

3.3.3 Contrast Visibility Threshold

Having measured the effects of the contrast sensitivity function and the contrast masking, we can now compute the threshold T needed to detect the difference between a pair of adjacent faces. To do so, we first evaluate the spatial frequency f and the local contrast c of the initial face pair (Eqs. (4) and (3)). We then normalize the computed contrast by the corresponding CSF value (Eq. (6)) and finally obtain the threshold using the masking model (Eq. (7)). The threshold T is expressed using the following equation:

$$T = \frac{\text{masking}(c \times \text{csf}(f))}{\text{csf}(f)}, \quad (8)$$

where c is the initial local contrast and f is the corresponding spatial frequency. Accordingly, if a local geometric distortion causes a change in contrast that is higher than the computed threshold then it is classified as visible. In the next section we explain how the JND profile for a 3D mesh is obtained using the computed threshold T .

4 JUST NOTICEABLE DISTORTION PROFILE

The JND refers to the threshold beyond which a change in contrast becomes visible for the average observer. The JND profile that we propose in this section allows us to get this threshold for an arbitrary displacement direction.

4.1 Overview

In the 3D setting, the JND is evaluated by computing the maximum displacement each vertex can tolerate. On one hand, a vertex displacement in a given direction will probably cause a change in the normals of adjacent faces and a change in local density. On the other hand, we showed in Section 3 that the face normals and the local density affect the contrast and the spatial frequency, respectively. This means that the displacement of a vertex probably alters the local perceptual properties. The visibility of this alteration

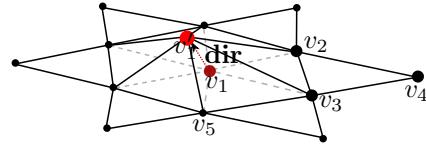


Fig. 8. The displacement of a vertex v_1 in a direction dir causes a change in contrast and spatial frequency for surrounding pairs of faces sharing a common edge in 1-ring and 2-ring of the displaced vertex.

can be evaluated using the perceptual models presented in Eqs. (6), (7) and (8). In this section, we present a numerical method for computing the maximum displacement beyond which the local distortion can be detected by an average human observer.

4.2 Visibility of Adjacent Face Pairs

The displacement of a vertex alters the local perceptual properties, *i.e.*, contrast and spatial frequency, of the surrounding pairs of adjacent faces (Fig. 8). In order to get the maximum displacement a vertex can tolerate, we need to evaluate the perceptual effect of any displacement on the set of its surrounding pairs of faces. In the following, we show how the visibility of a change in local perceptual property is evaluated. To this end, we first compute the change in contrast and spatial frequency and then evaluate the probability of detecting this change.

4.2.1 Change in Contrast

The displacement of a vertex v_1 in a direction dir causes the normals of its adjacent faces to change. This change in normals causes a variation in contrast for the surrounding pairs of adjacent faces. Therefore, evaluating the change in contrast requires evaluating the change in normal direction of these adjacent faces. For example, having two adjacent faces $\{v_1, v_3, v_2\}$ and $\{v_2, v_3, v_4\}$ (see Fig. 8) with normals \mathbf{n}_1 and \mathbf{n}_2 respectively, we express the new normal \mathbf{n}'_1 after displacing v_1 in a direction dir with a magnitude d by:

$$\begin{aligned} \tilde{\mathbf{n}}'_1 &= (v_1 - v_2) \times (v_3 - v_2) + d \cdot (\mathbf{dir} \times (v_3 - v_2)), \\ \mathbf{n}'_1 &= \frac{\tilde{\mathbf{n}}'_1}{\|\tilde{\mathbf{n}}'_1\|}. \end{aligned} \quad (9)$$

Since none of the vertices of the second face $\{v_2, v_3, v_4\}$ is displaced, its normal direction does not change. For the cases where the displacement of v_1 causes changes in the normal directions of both faces (*e.g.*, the pair of adjacent faces $\{v_1, v_3, v_2\}$ and $\{v_1, v_5, v_3\}$ in Fig. 8), their new normals are evaluated similarly, according to an adaptation of Eq. (9). The new contrast between adjacent faces is then evaluated using Eq. (3), with the new face normal(s).

4.2.2 Change in Spatial Frequency

Moving the vertex v_1 in the direction dir may cause a change in spatial frequency as well, because the size of the adjacent face pairs might be altered. Computing the new spatial frequency requires evaluating the distance between the opposing vertices v'_1 and v_4 , v'_1 being the position of

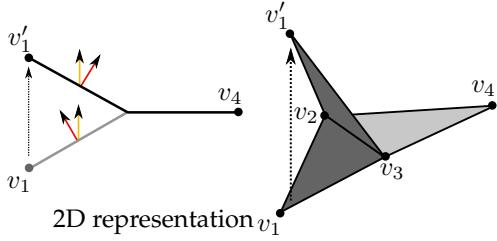


Fig. 9. When the displacement of a vertex alters the convexity of two adjacent faces, the contrast might remain the same as long as the angle between the light direction (yellow arrow) and the face normal (red arrow) does not change.

v_1 after its displacement. The new distance between the opposing vertices is expressed as:

$$\|v_4 - v'_1\| = \sqrt{\|v_4 - v_1\|^2 + d^2 - 2d \times (v_4 - v'_1) \cdot \text{dir}}. \quad (10)$$

To obtain the spatial frequency in cpd, we apply the perspective projection to $(v_4 - v'_1)$ in order to get the number of pixels that the face pair occupies on the screen and then apply Eq. (4).

4.2.3 Detection Probability

Having expressed how the displacement of a vertex in an arbitrary direction affects the local contrast and the spatial frequency of the surrounding pairs of adjacent faces, we now explain how to determine whether this change is visible or not. To do so, we compute the probability of detecting this change. The method for computing this probability is inspired by Daly's VDP [31].

The so-called *psychometric function* describes the probability for the human visual system to detect a visual stimulus. A common choice for the psychometric function is given by the following equation:

$$p(\Delta c) = 1 - e^{-(\Delta c/T)^\beta}, \quad (11)$$

where T is the threshold as computed in Eq. (8) and β is the slope of the psychometric function. β is set to 3.5 across many psychophysical experiments and perceptual studies [43]. Using the data collected from our psychophysical experiments, we evaluated the detection probability for contrasts near the measured threshold. The computed probabilities are fitted to the psychometric function and we obtain β of about 3.6. This fitted value will be used in our calculation. Δc is the change in contrast which corresponds to contrast difference before and after the displacement of a vertex and is evaluated as:

$$\Delta c = \begin{cases} \|c' - c\| & \text{if } \text{sgn}(\mathbf{n}_1 \cdot (v_4 - v_3)) \text{ does not change,} \\ c' + c & \text{if } \text{sgn}(\mathbf{n}_1 \cdot (v_4 - v_3)) \text{ changes,} \end{cases} \quad (12)$$

where c and c' are respectively the contrast of the adjacent faces before and after the vertex displacement. We test whether the vertex displacement causes a switch in sign of $\mathbf{n}_1 \cdot (v_4 - v_3)$, which implies a change in convexity between the adjacent faces. This allows us to detect the ambiguous case as shown in Fig. 9, where the displacement does not induce a change in the "conventional" contrast between the adjacent faces.

4.3 Vertex Displacement Threshold

In order to compute the threshold beyond which the displacement of a vertex v in a direction dir is visible, we proceed by the following steps. First, a list of the adjacent pairs of faces that are affected by the displacement of v is built. For each pair of faces, we start by computing their original perceptual properties and the corresponding contrast threshold using Eqs. (3), (4) and (8). In particular, the display and observation parameters are the inputs of the JND algorithm, therefore the proposed JND profile can be adaptively computed for different viewing distances and display sizes. Then we gradually increase the displacement magnitude of v and compute the change in frequency and contrast (Eqs. (10) and (9)) at each step. This allows us to evaluate the probability of detecting the vertex displacement (Eq. (11)) for each of the adjacent face pairs at different displacement steps. Note that when the displacement causes a change in spatial frequencies, we take into account the most sensitive frequency that results in a higher detection probability. Finally, the threshold is attributed to the displacement magnitude where the detection probability reaches a certain threshold for at least one of the face pairs. In practice we set the probability threshold at 0.95. To better understand this process, let us consider the two vertices v_1 and v_2 in Fig 10. Both vertices are displaced in their normal direction. The first vertex v_1 is situated on a rough region (initial contrast of all surrounding pairs of adjacent faces $>$ CSF threshold) and the second vertex v_2 on a smooth region (initial contrast $<$ CSF threshold). The displacement of v_1 and v_2 barely affects the spatial frequency of the surrounding face pairs as can be seen in the left plots. The middle plots show how displacing v_1 and v_2 in the normal direction affects the local contrast. The probability of detecting this change in contrast is shown in the right plots. These plots show that v_2 is more sensitive and can tolerate less displacement than v_1 . This is due to the different initial contrasts of the two vertices. The initial contrasts around v_1 is above the CSF threshold. This implies that the visibility threshold is increased due to the masking effect, which explains the slow increase in detection probability. For v_2 all initial contrasts are below the CSF threshold. No masking should occur which means that once the contrast is above the CSF threshold the displacement should be visible. This is exactly what we observe. When the contrast of "face pair 4" reaches the CSF level then the detection probability becomes close to 1.

In the description above, we explain how to compute the displacement threshold by brute-force incremental step searching only for clarity purposes. In practice, we instead use a half-interval search to find the threshold (as described in Algorithm 1), which is simple yet very fast and accurate. In our tests we have set the visibility threshold th to 0.95, the precision p to 0.005 and the parameter *very_high_value* to 1/10th of the mesh bounding box. In order to compute the value of *visibility*, we call the psychometric function (Eq. (11)) which again requires the evaluation of the change of contrast (Eq. (12)) and the contrast threshold (Eq. (8)).

Computing the displacement threshold requires an estimation of the spatial frequency and the local contrast. This makes the obtained threshold dependent on the display

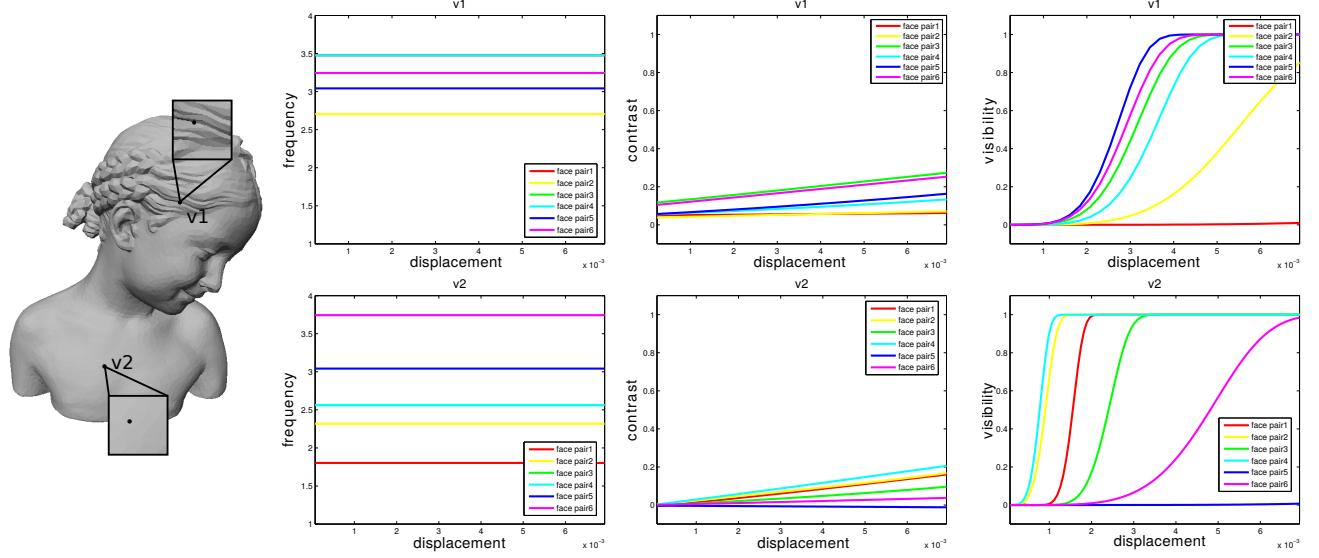


Fig. 10. The evolution of the local perceptual properties and visibility, for two displaced vertices v_1 and v_2 on the Bimba model. Left plots show the change in frequency, middle ones show the change in contrast and the right ones show the detection probability, of different pairs of affected adjacent faces of the two vertices. Note that some of the faces have the same spatial frequency, so the color curves overlap in the left plots.

Algorithm 1: Half-interval search algorithm.

Data: v : vertex, dir : noise direction, l : light direction, th : visibility threshold, p : precision
Result: $dist$: displacement threshold

```

min = 0;
max = very_high_value;
dist = max;
visibility = compute_visibility(v, dir, l, dist);
while || visibility - th || > p do
    dist = (max - min) / 2 + min;
    visibility = compute_visibility(v, dir, l, dist);
    if visibility > th then
        | max = dist;
    else
        | min = dist;
    end
end

```

parameters (size and resolution), the observer's distance to the display and scene's illumination. However, in an interactive setting where the light source is fixed relative to the viewpoint, the light direction varies with respect to the 3D mesh. It is therefore important to compute the displacement threshold independently of the light direction.

We hereby propose a *light independent mode* for computing the displacement threshold. To do so, we simply compute the threshold according to multiple light directions and then choose the smallest one. Since the contrast between two faces is defined only when the dot product between the light direction and the normals is positive, then we consider that the possible light directions are located and sampled on the local half sphere around the displaced vertex. In practice, it is not necessary to densely sample the entire half sphere; Eq. (3) implies that if the light direction is close to the normal, the contrast varies slowly. This means that the displacement threshold is bigger for light directions that are close to the normal direction than for light directions close to the tangent direction. This suggests that it would

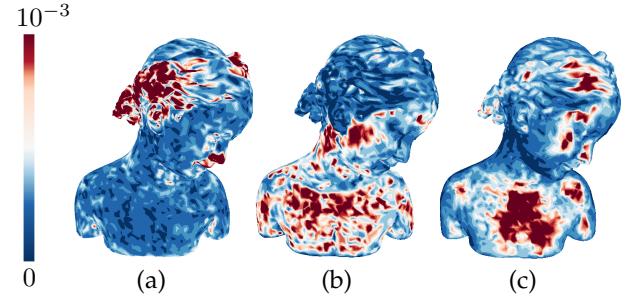


Fig. 11. The JND profile for the Bimba model under different circumstances. (a) The threshold is computed with respect to a displacement in the normal direction of each vertex in a light independent mode. (b) The threshold is computed with respect to a displacement in the tangent direction in a light independent mode. (c) The threshold is computed with respect to a displacement in the normal direction of each vertex with a light fixed in front of the model.

only be necessary to sample the half sphere near the tangent direction.

In order to obtain the JND profile of a 3D mesh, we compute for each of the vertices the displacement threshold relative to a certain direction. Figure 11 shows the JND profile for a mesh under different circumstances. Figure 11.(a) displays the JND profile relative to a displacement in the normal direction in a light independent mode. Due to the effects of contrast masking, the rough region of the model can tolerate more noise than the smooth part. This is not the case when the JND is computed relative to a displacement in the tangent direction (Fig. 11.(b)) where the smooth part can tolerate more displacement. This is because a displacement in the tangent direction for a smooth vertex will barely alter the normal of the surrounding faces and thus the local contrast will not be affected by the displacement. Figure 11.(c) shows the JND profile relative to a displacement in the normal direction when the light source is fixed. As expected, we can see that the obtained threshold is maximal when the surface normals are in the same direction of the light.

5 SUBJECTIVE VALIDATION

In order to test the performance of a Just Noticeable Distortion profile, it is common in the image or video JND context to perform a subjective experiment [44], [45], [46] where a JND modulated random noise is added to the images or videos. The participants should then rate the visibility of the displayed noise. A JND model should be able to inject noise into the image or video while keeping it invisible; the best JND model being the one that is able to add the largest amount of invisible noise. We have conducted a subjective experiment where we have tested the performance of the proposed JND model. We compared the visibility of noise on altered 3D meshes, which were obtained by adding three different types of noise to an initial mesh. The three types of noise are:

- uniform random noise without any modulation;
- random noise modulated by the surface roughness;
- random noise modulated by the proposed JND model.

Surface roughness is an important candidate to test our JND model against since it is accepted in the computer graphics community that noise is less visible in rough regions [19]. We also have reached this conclusion when the noise is in the normal direction (Fig. 11.(a)).

5.1 Mesh Alteration

We injected noise into 3D meshes according to the following equation:

$$v'_i = v_i + \text{rnd} \times M(v_i) \times \text{dir}_i, \quad (13)$$

where v_i is the i^{th} vertex of the initial mesh and v'_i is the corresponding noisy vertex. dir is the noise direction. rnd is a random value equal to either $+1$ or -1 and $M(v_i)$ represents the magnitude of the noise for v_i . It is defined as:

$$M(v_i) = \begin{cases} \beta_{unif} & \text{uniform noise,} \\ \beta_{rough} \times lr(v_i) & \text{roughness modulated noise,} \\ \beta_{jnd} \times jnd(v_i) & \text{JND modulated noise,} \end{cases} \quad (14)$$

where β_{unif} , β_{rough} and β_{jnd} regulate the global noise energy for each of the noise injection methods. $lr(v_i)$ is the local surface roughness as defined in [21] and $jnd(v_i)$ is the JND value computed as explained in Section 4.3. In order to allow user interaction during the experiment, the JND value was computed independently from any light direction.

For the subjective experiments we injected noises of two different energy levels: $\beta_{jnd} = 1$ and $\beta_{jnd} = 2$. These levels correspond to a near-threshold noise and to a supra-threshold noise, respectively. For $\beta_{jnd} = 1$ the injected noise is supposed to be difficult to notice while for $\beta_{jnd} = 2$ the noise is expected to be visible. We then fix β_{unif} and β_{rough} such that for the meshes altered using our JND model, the maximum root mean square error (MRMS) [1], [2], a widely used purely geometric distance, is the biggest for each noise level. Indeed, the objective here is to show that our JND model is able to inject the highest amount of noise onto the mesh among the three methods, while producing the least visible one. In addition, we tested the performance of the JND model for noise in a random direction for each vertex

and that in the normal direction for each vertex. To see the effects of light direction we ran the experiment twice: once with the light source in front of the model and another time with the light on top left of the model.

5.2 Method

Procedure. The subjective experiment followed the “adjectival categorical judgment method” [47]. This procedure consists of displaying two 3D meshes side by side, the reference on the left and the noisy one on the right. The participants were asked to rate the visibility of the noise on a discrete scale from 0 to 5, 0 being the score attributed when the noise cannot be seen and 5 when the noise is clearly visible. 5 “dummy” models were included at the beginning of each session to stabilize subjective scores. The models were presented in a randomized order. To avoid any memory-based bias, two meshes derived from the same reference model were never displayed consecutively.

Settings. The experiment was conducted in a low illuminated environment. We used a 23-inch Asus screen with a 1920×1080 resolution to display the 3D models. The participants viewed the models from a distance of 50 cm. During the experiment, the two displayed meshes had a synchronized viewpoint and subjects could freely rotate around the displayed meshes. To encourage close examination of the displayed mesh, no score could be registered before 10 seconds of interaction occur. The initial viewpoint was manually set for all models. The light source was fixed with reference to the camera position. A front and a top-left light directions were used.

Participants. 12 subjects (7 females and 5 males) participated in these experiments. All of them had normal or corrected-to-normal vision and were between the age of 20 and 29.

5.3 Results

After collecting the subjective scores, we have computed the mean score over each of the noise types. “JND 1” and “JND 2” refer to the models obtained by modulating the random noise with our JND model for near-threshold and supra-threshold levels, respectively. “Rough 1” and “Rough 2” refer to the ones obtained using the surface roughness measure and “Unif 1” and “Unif 2” to the ones with uniform random noise. Figure 12 displays the results of the subjective experiments. Plots (a) to (c) present the results for the noise in the normal direction and plots (d) to (e) the results for the noise in a random direction. Figures 12.(a) and 12.(d) show that the noise on the “JND 1” models was indeed difficult to detect as the mean subjective score is about 0.45. Interestingly, the participants rated “Unif 1” and “Rough 1” models similarly to “JND 2” which refers to the supra-threshold noise level models that contain twice the noise of “Unif 1” and “Rough 1”. Plots (b) and (e) also show that “JND 1” models were perceived almost identically both under front and top-left illumination conditions. This is not the case for “Unif 1” and “Rough 1” models where the grazing light direction of the top-left illumination made the noise more apparent. It is also important to note that the visibility of the noise for “JND 1” models was identical for all models. This is not the case for “Rough 1” and “Unif

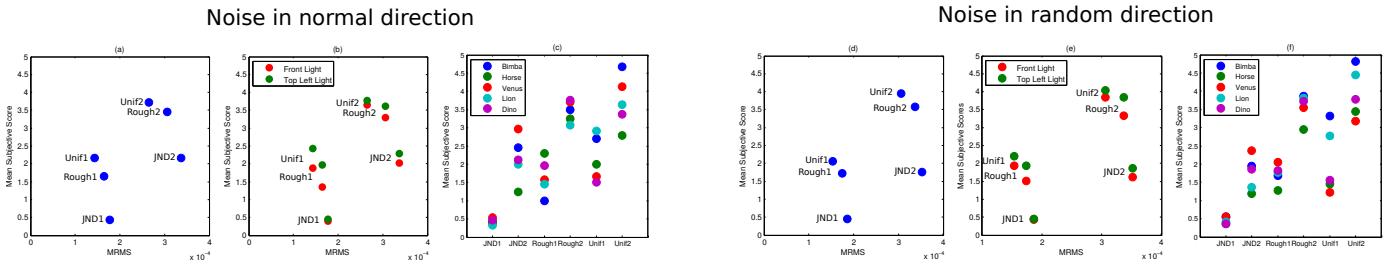


Fig. 12. Mean subjective score values versus MRMS distance values. Plots (a) and (d) present, for different noise injections, the mean subjective scores over all test models and the two illumination settings. Plots (b) and (e) show the difference in mean subjective scores between the experiments in the two illumination settings. Plots (c) and (f) compare the mean subjective scores for the different models used in the experiments.

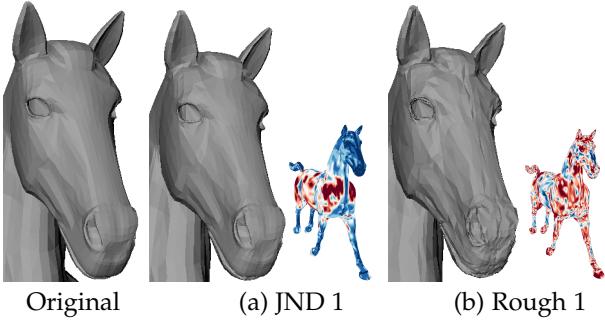


Fig. 13. The proposed JND model takes into account the density of mesh when computing the visibility threshold. When adding random noise modulated by the JND profile, the noise will be added to the coarse regions and avoid the dense area where the noise will be easily visible (a). This is not the case when adding random noise modulated by the surface roughness (b).

1" where the visibility of noise varied a lot for different models (see Figs. 12.(c) and 12.(f)). This is mainly due to the difference in mesh density between the models; high density models are in general more sensitive to noise than low density ones.

The main advantage of the proposed JND model over the surface roughness measures is that it adapts to the mesh characteristics (density, size), noise direction and scene illumination. Figure 13 illustrates the importance of mesh density. The Horse is a model with mostly smooth regions, the rough regions are packed in the head's features. In addition, the head is densely sampled while the body is coarsely sampled. The JND model avoids adding noise in the dense head and takes advantage of the coarse body, while surface roughness measures are not able to detect the difference in sampling. The noise is thus rather injected in the dense head features, which makes it visible.

These results show that the proposed JND model is indeed able to add the largest amount of invisible noise onto the mesh surface among the three methods. Furthermore, the proposed JND model can accurately predict the visibility threshold for 3D meshes, taking into account the noise direction, the mesh characteristics and the scene illumination. However, the proposed model cannot accurately describe how the supra-threshold noise visibility (or annoyance) is perceived since it has not been designed for this purpose; the noise was perceived differently for each model in "JND 2" (Figs. 12.(c) and (f)).

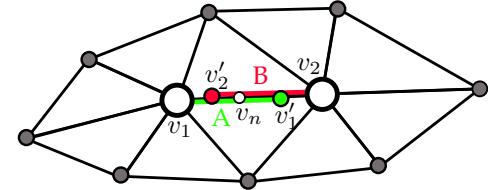


Fig. 14. If $v_1 v_2$ and $v'_1 v'_2$ are in opposite directions, then the edge (v_1, v_2) can be collapsed to v_n without causing any visible distortion.

6 APPLICATIONS

The JND models of 2D images and videos have been used extensively throughout the literature to guide and perceptually optimize several image and video processing algorithms [6], [7], [8]. In this section, we show how the proposed 3D JND profile can be integrated to mesh processing algorithms. We used the proposed JND profile to guide the simplification of 3D meshes and to automatically select the optimal vertex coordinates quantization level.

6.1 JND-Driven Mesh Simplification

The goal of mesh simplification algorithms is to reduce the number of vertices in a mesh to a certain degree by iteratively applying a simplification step (edge collapse or vertex removal). Mesh simplification is usually used to efficiently display highly detailed models or to create multiple levels of details (LOD) of a mesh, so it is required that the simplified mesh preserves the geometric features of the model as much as possible. To do so, a simplification cost is assigned to each of the mesh edges (or vertices), then the simplification step is applied to the edge (or vertex) with the lowest cost and finally the costs are updated prior to the next iteration. Several perceptual methods have been proposed to compute the simplification cost. However, existing perceptual methods either carry out the perceptual analysis on the rendered image [35], [36], [37] or rely on a top-down estimation of saliency [26], [28], [29]. Moreover, none of the existing algorithms propose a method to automatically control the quality of the resulting output; the simplification is usually carried out until a manually prescribed number of edges or vertices is reached.

We use our JND model to define both the simplification cost for each edge and a stopping criterion that automatically controls the quality of the simplified mesh.

Edge cost. In an edge collapse operation, an edge (v_1, v_2) is removed and is replaced by a vertex v_n (Fig. 14). This can

be seen as if the vertices v_1 and v_2 moved towards the new vertex v_n . Using our JND model we analyze the visibility of displacing v_1 and v_2 along the edge (v_1, v_2) . Let A (resp. B) be a part of (v_1, v_2) bounded by v_1 and v'_1 (resp. v_2 and v'_2) (see Fig. 14) where v'_1 (resp. v'_2) is the vertex obtained by displacing v_1 (resp. v_2) by exactly the JND value in the direction of $\mathbf{v}_1\mathbf{v}_2$ (resp. $\mathbf{v}_2\mathbf{v}_1$). This means that replacing v_1 (resp. v_2) by a vertex belonging to A (resp. B) will not cause any visible distortion. In order to apply an edge collapse that is invisible to a human observer, we need to find a new vertex v_n such that $v_n \in A \cap B$. This requires that the vectors $\mathbf{v}_1\mathbf{v}_2$ and $\mathbf{v}'_1\mathbf{v}'_2$ should be in opposite directions so that $A \cap B \neq \emptyset$. Otherwise, if $\mathbf{v}_1\mathbf{v}_2$ and $\mathbf{v}'_1\mathbf{v}'_2$ are in the same direction, then we have $A \cap B = \emptyset$, making the distortion caused by the edge collapse visible. This analysis leads us to define the simplification cost of an edge by:

$$c = \frac{\mathbf{v}_1\mathbf{v}_2 \cdot \mathbf{v}'_1\mathbf{v}'_2}{\|\mathbf{v}_1\mathbf{v}_2\|^2}. \quad (15)$$

The value of our simplification cost c varies between $[-1, 1]$. If $c < 0$ then the collapse operation does not affect the visual fidelity of the model. If $c > 0$ then the edge collapse will be visible. Figure 15 shows the simplification cost on a cube where we have injected a random noise on each of its sides. The simplification cost of the edges belonging to the top side is below 0 as the injected noise is under the JND threshold.

Vertex placement. Having defined the simplification cost of an edge, we now should decide how the position of the new vertex v_n is computed. In order to get the "optimal" position we have found that minimizing the following quadratic energy produce very good results:

$$\arg \min \left\{ \left(\frac{\|\mathbf{v}_1\mathbf{v}_n\|}{\text{jnd}_{v_1}} \right)^2 + \left(\frac{\|\mathbf{v}_2\mathbf{v}_n\|}{\text{jnd}_{v_2}} \right)^2 \right\}, \quad (16)$$

where jnd_{v_1} (resp. jnd_{v_2}) is the JND threshold of v_1 (resp. v_2) in the direction of $\mathbf{v}_1\mathbf{v}_2$ (resp. $\mathbf{v}_2\mathbf{v}_1$). This yields to:

$$\|\mathbf{v}_1\mathbf{v}_n\| = \|\mathbf{v}_1\mathbf{v}_2\| \times \frac{\text{jnd}_{v_1}^2}{\text{jnd}_{v_1}^2 + \text{jnd}_{v_2}^2}, \quad (17)$$

where $\|\mathbf{v}_1\mathbf{v}_n\|$ and $\|\mathbf{v}_2\mathbf{v}_n\|$ represent respectively the distances by which v_1 and v_2 are being displaced. The idea behind minimizing this quadratic energy is to make the displacement of v_1 and v_2 adaptive to their corresponding JND values.

Stopping criterion. The value of the defined simplification cost varies between $[-1, 1]$. For edges with a cost greater than 0 the collapse operation will be visible. So in order to have a simplified mesh that is visually similar to the original version, we collapse all the edges whose cost is less than or equal to 0. This allows us to define a stopping criterion which consists in stopping the simplification process once all edges have a simplification cost above 0. Figure 16 shows a highly dense 3D mesh. The model is then simplified with the JND-driven simplification method. The resulting simplified mesh (Fig. 16.(a)) has 80% less vertices and is visually very similar to the original version. Removing 5% more vertices beyond the JND level introduces slightly visible distortions to the model (Fig. 16.(b)). In addition, simplifying the model using Lindstrom and Turk's method [48] (edge collapse with a different cost) to the

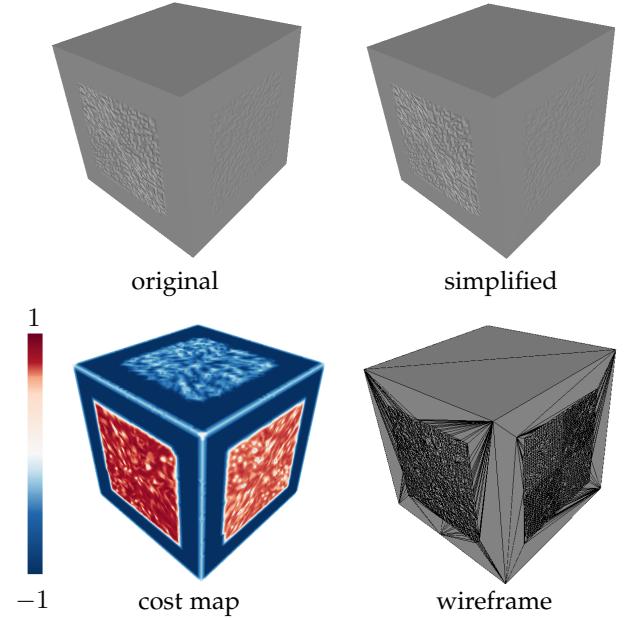


Fig. 15. A random noise of different intensities is injected to different sides of a dense cube mesh. The noise on the top side is below the JND threshold. On the right side, the noise is barely visible as it is just above the JND threshold and on the left side is injected a visible noise. The JND driven simplification process will keep all of the visible noise and simplify the top side with noise that is below the visibility threshold.

same number of vertices as the JND-driven simplification also results in slightly visible distortions (Fig. 16.(c)). The mesh LOD results of the simplification application can be found in the supplementary material submitted along with the manuscript.

6.2 Vertex Coordinates Quantization

Vertex coordinates quantization is an important step in many mesh processing algorithms, especially compression. This operation may introduce visible distortion to the original mesh. It is thus important to find the optimal quantization level (in bits per coordinate, bpc), which is different for each mesh due to differences in geometric complexities and details. We define the optimal quantization level as the one with the highest quantization noise energy that remains visually indistinguishable from the original mesh.

The proposed JND model provides a simple and automatic way to determine the optimal quantization level independently of the nature of the mesh. The idea is to compute a score allowing us to compare the model's JND profile to the magnitude of introduced noise. To do so, we start by computing the displacement vectors as:

$$\mathbf{disp}_i = v'_i - v_i, \quad (18)$$

where v'_i and v_i are the i^{th} vertices of respectively the distorted mesh and the original one. The direction of \mathbf{disp}_i represents the quantization noise direction. We then compute the JND profile of the original mesh with respect to the computed displacement direction. We finally compute

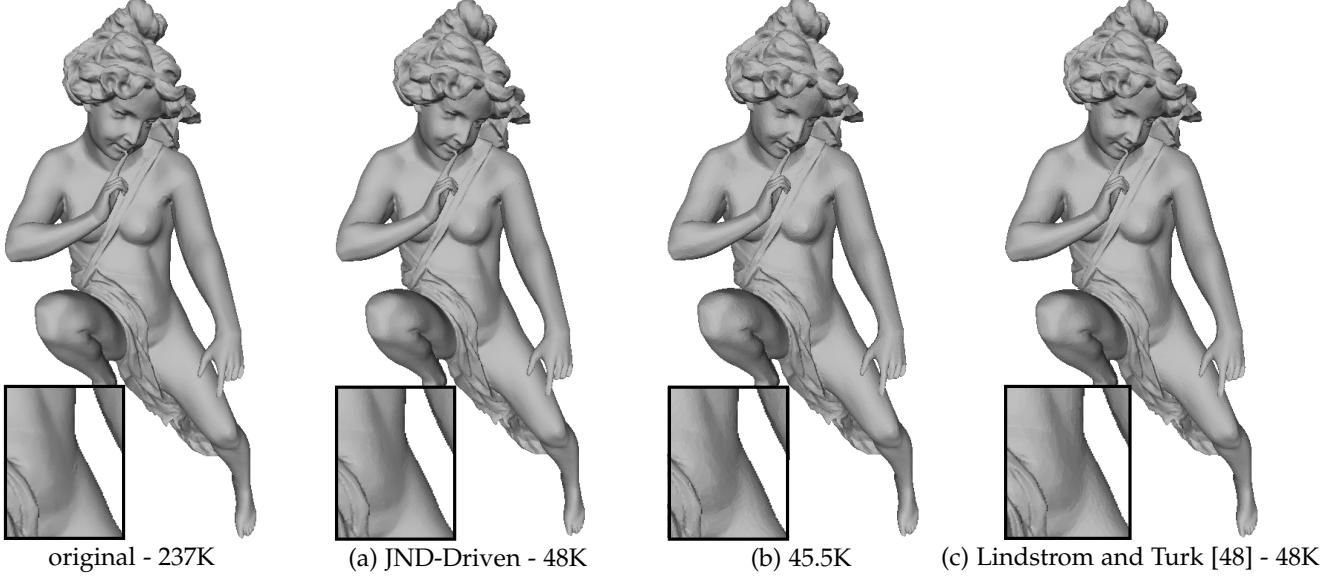


Fig. 16. (a) The JND-driven mesh simplification process outputs a model that is visually very similar to the original model. (b) Removing 5% more vertices will introduce slightly visible distortions to the simplified model. (c) The simplified model by using the method of Lindstrom and Turk [48] to the same number of vertices as the JND-driven simplification. (b) and (c) contain slightly visible distortions, especially on the belly and thighs.

the score as the mean of the ratio between the norm of the displacement vector and the JND value:

$$S = \frac{1}{n} \sum_{i=1}^n \frac{\|\text{disp}_i\|}{\text{jnd}(v_i)}, \quad (19)$$

where n is the number of vertices in the mesh. This score allows us to test whether the introduced distortion is visible. If $S < 1$, the noise magnitude is globally below the visibility threshold, which means that the distortion is not visible. On the other hand if $S > 1$, the distortion becomes visible as the noise magnitude is above the visibility threshold. Figure 17 shows the JND comparison scores versus the level of coordinates quantization for three different meshes. According to the defined score the optimal quantization level is respectively 12, 11 and 10 bpc for the Venus, Head and Bimba models. These results are consistent with human observations as shown in Fig. 18. Figure 17 shows also the FMPD [21] scores versus the level of coordinates quantization for the three meshes. We cannot define a proper threshold on the FMPD [21] scores that gives the same optimal quantization levels for the Venus, Head and Bimba models. One possible explanation is that the FMPD metric has difficulties in producing consistent evaluation results on meshes of different densities and geometric complexities.

7 LIMITATIONS

One of the limitations of the proposed JND model is that it currently only works for diffuse surface that is illuminated by a white directional light and rendered with a flat shading algorithm. This is due to the simplified contrast definition under that circumstance which is proposed in Section 3.1. However, the JND threshold is based on the estimation of visibility which is obtained using low-level proprieties of human visual system (CSF and contrast masking) and relies heavily on an estimation of local contrast. This means that extending the JND model to different types of surfaces and

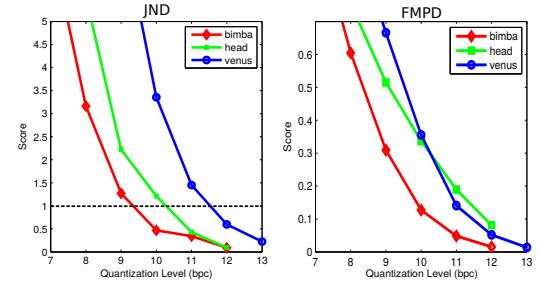


Fig. 17. Left: The JND comparison score versus the quantization levels (in bpc) of three meshes. Right: The FMPD [21] score versus the quantization levels (in bpc) of the three meshes.

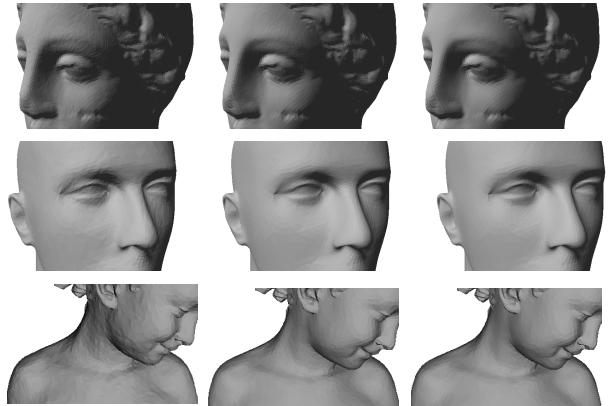


Fig. 18. Quantized meshes with different quantization levels. The middle mesh corresponds to the optimal quantization level (12, 11 and 10 bpc for Venus, Head and Bimba respectively). The right mesh corresponds to a one bit higher than the optimal level while the one on left corresponds to a one bit lower. (For better comparison between the models please refer to the electronic version of this manuscript.)

lights such as specular surfaces and point light illumination requires a generalization of the definition of contrast under the corresponding surface and lighting condition, probably via an appropriate analysis of the rendering algorithm. This will be the focus of our future work.

Another limitation of the proposed JND model is that it relies on low-level properties of the human visual system such as the contrast sensitivity function and contrast masking. These low-level properties allow us to predict whether a distortion is visible or not. If a distortion is visible, the current JND model cannot accurately predict to which extent this distortion affects the visual fidelity of the model. Taking this into consideration requires adding to the model some higher-level properties of the human visual system such as entropy masking [49] or the free energy principle [50]. We think that both of these properties could be properly defined by analyzing the local contrast of the surface in a certain neighborhood.

8 CONCLUSION AND FUTURE WORK

In this paper, we have presented a model for computing a Just Noticeable Distortion (JND) profile for flat-shaded 3D meshes. The proposed model takes into consideration the different mesh properties (size, density) and the varying circumstances of mesh usage (display parameters, light direction, viewing distance). Our JND profile is based on an experimental study of the local perceptual properties of the human visual system, *i.e.*, the local contrast and the spatial frequency. These perceptual properties have been defined for 3D meshes. They were then used to experimentally measure the effects of the *contrast sensitivity function* and *contrast masking* when the displacement of a vertex occurs. The results of these experiments have been utilized to evaluate the probability of detecting the displacement of a vertex in an arbitrary direction, which allows us to define a JND profile for flat-shaded 3D meshes. We have tested the performance of the proposed JND model via a subjective experiment where the participants had to rate the visibility of JND modulated random noise added to a series of models. The results show that our model can accurately predict the visibility threshold of vertex noise.

We have used the proposed JND model to guide the simplification of 3D meshes. The JND-driven simplification method relies on a perceptual simplification cost assigned to each edge, and it can automatically stop the simplification process in order to obtain a visually very similar simplified mesh. Finally, we have proposed a method to automatically obtain the optimal vertex coordinates quantization level.

Our future work will first focus on generalizing the contrast definition of 3D meshes. This will broaden the usage of the proposed JND model to include smooth shaded surface and different types of illumination. We will also work to add higher aspects of the human visual system to the JND, which will allow us to predict the visibility/annoyance of supra-threshold geometric distortions. Interestingly, there have also been several studies on the perception of dynamic 3D meshes recently [51], [52]. By incorporating the dynamic aspects of the human visual system we may be able to extend the JND model to dynamic meshes.

ACKNOWLEDGMENTS

We would like to thank all the subjects who participated in the subjective experiments. This work is supported by the ARC6 program of the "Région Rhône-Alpes" through the PADME project. We thank the anonymous reviewers for their very helpful comments and suggestions.

REFERENCES

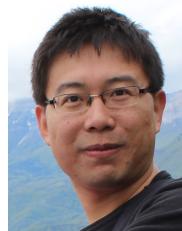
- [1] P. Cignoni, C. Rocchini, and R. Scopigno, "Metro: Measuring error on simplified surfaces," *Computer Graphics Forum*, vol. 17, no. 2, pp. 167–174, 1998.
- [2] N. Aspert, D. Santa-Cruz, and T. Ebrahimi, "MESH: Measuring error between surfaces using the Hausdorff distance," in *Proc. of IEEE International Conference on Multimedia & Expo*, 2002, pp. 705–708.
- [3] M. Corsini, M.-C. Larabi, G. Lavoué, O. Petřík, L. Váša, and K. Wang, "Perceptual metrics for static and dynamic triangle meshes," *Computer Graphics Forum*, vol. 32, no. 1, pp. 101–125, 2013.
- [4] W. Lin, "Computational models for just-noticeable difference," in *Digital Video Image Quality and Perceptual Coding*, H. R. Wu and K. R. Rao, Eds. London, UK: CRC Press, 2006, pp. 281–303.
- [5] W. Lin, L. Dong, and P. Xue, "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 7, pp. 900–909, 2005.
- [6] C.-H. Chou and Y.-C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 467–476, 1995.
- [7] Z. Liu, L. J. Karam, and A. B. Watson, "JPEG2000 encoding with perceptual distortion control," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1763–1778, 2006.
- [8] Z. Wei and K. Ngan, "Spatio-temporal just noticeable distortion profile for grey scale image/video in DCT domain," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 3, pp. 337–346, 2009.
- [9] R. W. Fleming and M. Singh, "Visual perception of 3D shape," in *Proc. of ACM SIGGRAPH Courses*, 2009, pp. 24:1–24:94.
- [10] G. Lavoué and M. Corsini, "A comparison of perceptually-based metrics for objective evaluation of geometry processing," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 636–649, 2010.
- [11] B. A. Wandell, *Foundations of Vision*. Sunderland, MA, USA: Sinauer Associates, 1995.
- [12] A. Beghdadi, M.-C. Larabi, A. Bouzerdoum, and K. M. Iftekharuddin, "A survey of perceptual image processing methods," *Signal Processing: Image Communication*, vol. 28, no. 8, pp. 811–831, 2013.
- [13] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.
- [14] S. E. Palmer, *Vision Science: Photons to Phenomenology*. Cambridge, MA, USA: MIT Press, 1999.
- [15] C. T. Blakemore and F. Campbell, "On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images," *The Journal of Physiology*, vol. 203, no. 1, pp. 237–260, 1969.
- [16] D. G. Pelli and P. Bex, "Measuring contrast sensitivity," *Vision Research*, vol. 90, pp. 10–14, 2013.
- [17] A. B. Watson and A. J. Ahumada, "A standard model for foveal detection of spatial contrast ModelFest experiment," *Journal of Vision*, vol. 5, no. 9, pp. 717–740, 2005.
- [18] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *Journal of Optical Society of America*, vol. 70, no. 12, pp. 1458–1471, 1980.
- [19] G. Lavoué, "A local roughness measure for 3D meshes and its application to visual masking," *ACM Transactions on Applied Perception*, vol. 5, no. 4, pp. 1–23, 2009.
- [20] M. Corsini, E. Drelie Gelasca, T. Ebrahimi, and M. Barni, "Watermarked 3D mesh quality assessment," *IEEE Transactions on Multimedia*, vol. 9, no. 2, pp. 247–256, 2007.
- [21] K. Wang, F. Torkhani, and A. Montanvert, "A fast roughness-based approach to the assessment of 3D mesh visual quality," *Computers & Graphics*, vol. 36, no. 7, pp. 808–818, 2012.

- [22] L. Dong, Y. Fang, W. Lin, and H. S. Seah, "Objective visual quality assessment for 3D meshes," in *Proc. of International Workshop on Quality of Multimedia Experience*, 2014, pp. 1–6.
- [23] G. Lavoué, "A multiscale metric for 3D mesh visual quality assessment," *Computer Graphics Forum*, vol. 30, no. 5, pp. 1427–1437, 2011.
- [24] F. Torkhani, K. Wang, and J.-M. Chassery, "A curvature-tensor-based perceptual quality metric for 3D triangular meshes," *Machine Graphics & Vision*, vol. 23, no. 1–2, pp. 59–82, 2014.
- [25] L. Váša and J. Rus, "Dihedral angle mesh error: a fast perception correlated distortion measure for fixed connectivity triangle meshes," *Computer Graphics Forum*, vol. 31, no. 5, pp. 1715–1724, 2012.
- [26] C. H. Lee, A. Varshney, and D. W. Jacobs, "Mesh saliency," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 659–666, 2005.
- [27] X. Chen, A. Saparov, B. Pang, and T. Funkhouser, "Schelling points on 3D surface meshes," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 29:1–29:12, 2012.
- [28] J. Wu, X. Shen, W. Zhu, and L. Liu, "Mesh saliency with global rarity," *Graphical Models*, vol. 75, no. 5, pp. 255–264, 2013.
- [29] R. Song, Y. Liu, R. R. Martin, and P. L. Rosin, "Mesh saliency via spectral processing," *ACM Transactions on Graphics*, vol. 33, no. 1, pp. 1–17, 2014.
- [30] S. Marras, L. Váša, G. Brunnert, and K. Hormann, "Perception-driven adaptive compression of static triangle meshes," *Computer-Aided Design*, vol. 58, pp. 24–33, 2015.
- [31] S. Daly, "The visible differences predictor: An algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, A. B. Watson, Ed. Cambridge, MA, USA: MIT Press, 1993, pp. 179–206.
- [32] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Transactions on Graphics*, vol. 30, no. 4, pp. 40:1–40:13, 2011.
- [33] M. Ramasubramanian, S. Pattanaik, and D. Greenberg, "A perceptually based physical error metric for realistic image synthesis," in *Proc. of ACM SIGGRAPH*, 1999, pp. 73–82.
- [34] G. Ramanarayanan, J. Ferwerda, B. Walter, and K. Bala, "Visual equivalence: towards a new standard for image fidelity," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 76:1–76:11, 2007.
- [35] N. Williams, D. Luebke, J. D. Cohen, M. Kelley, and B. Schubert, "Perceptually guided simplification of lit, textured meshes," in *Proc. of ACM Symposium on Interactive 3D Graphics*, 2003, pp. 113–121.
- [36] L. Qu and G. W. Meyer, "Perceptually guided polygon reduction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 5, pp. 1015–1029, 2008.
- [37] N. Menzel and M. Guthe, "Towards perceptual simplification of models with arbitrary materials," *Computer Graphics Forum*, vol. 29, no. 7, pp. 2261–2270, 2010.
- [38] I. Cheng and P. Boulanger, "A 3D perceptual metric using just-noticeable-difference," in *Proc. of Eurographics Short Papers*, 2005, pp. 97–100.
- [39] I. Cheng, R. Shen, X. D. Yang, and P. Boulanger, "Perceptual analysis of level-of-detail: The JND approach," in *Proc. of IEEE International Symposium on Multimedia*, 2006, pp. 533–540.
- [40] Y. Yang, N. Peyerimhoff, and I. Ivrissimtzis, "Linear correlations between spatial and normal noise in triangle meshes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 1, pp. 45–55, 2013.
- [41] A. B. Watson and D. G. Pelli, "QUEST: a Bayesian adaptive psychometric method," *Perception & Psychophysics*, vol. 33, no. 2, pp. 113–120, 1983.
- [42] J. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 525–536, 1974.
- [43] M. J. Mayer and C. W. Tyler, "Invariance of the slope of the psychometric function with spatial summation," *Journal of Optical Society of America A*, vol. 3, no. 8, pp. 1166–1172, 1986.
- [44] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, "Just noticeable difference for images with decomposition model for separating edge and textured regions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1648–1652, 2010.
- [45] Y. Zhao, Z. Chen, C. Zhu, Y.-P. Tan, and L. Yu, "Binocular just-noticeable-difference model for stereoscopic images," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 19–22, 2011.
- [46] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," *IEEE Transactions on Multimedia*, vol. 15, no. 7, pp. 1705–1710, 2013.
- [47] International Telecommunication Union, Rec. BT.500: Methodology for the Subjective Assessment of the Quality of Television Pictures, 2012.
- [48] P. Lindstrom and G. Turk, "Fast and memory efficient polygonal simplification," in *Proc. of IEEE Visualization Conference*, 1998, pp. 279–286.
- [49] L. Dong, Y. Fang, W. Lin, C. Deng, C. Zhu, and H. S. Seah, "Exploiting entropy masking in perceptual graphic rendering," *Signal Processing: Image Communication*, vol. 33, pp. 1–13, 2015.
- [50] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.
- [51] L. Váša and V. Skala, "A perception correlated comparison method for dynamic meshes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 2, pp. 220–230, 2011.
- [52] F. Torkhani, K. Wang, and J.-M. Chassery, "Perceptual quality assessment of 3D dynamic meshes: Subjective and objective studies," *Signal Processing: Image Communication*, vol. 31, pp. 185–204, 2015.

Georges Nader received his B.S. degree in 2011 from the Lebanese University (Lebanon) and his M.S degree in 2013 from the "Université de Bordeaux" (France). He is now a Ph.D. student at the "Université Claude Bernard Lyon 1" (France) under the supervision of Prof. Florent Dupont, Dr. Kai Wang and Dr. Franck Hétry-Wheeler. His research interests include perceptually driven computer graphics and geometry processing.



Kai Wang received the Ph.D. degree in Computer Science in 2009 from the University of Lyon, Lyon, France. Following a ten-month post-doctoral position at Inria Nancy, he joined in 2011 GIPSA-Lab, Grenoble, France, as a full-time CNRS researcher. His current research interests include multimedia security and surface analysis.



Franck Hétry-Wheeler received an Eng. degree and a M.S. degree in applied mathematics both in 2000 from the University of Grenoble Alpes, as well as a PhD in computer science from the same university in 2003. Since 2004 he has been an assistant professor at the University of Grenoble Alpes (Grenoble INP - Ensimeag). He has been a member of the Morphéo team at Inria Grenoble since 2011, where he conducts research on shape analysis and understanding, digital geometry and topology processing.



Florent Dupont received his B.S. and M.S. degree in 1990, and his Ph.D. in 1994 from "Institut National des Sciences Appliquées" of Lyon, France. He became Associate Professor in 1998. He is now Professor in the Multiresolution, Discrete and Combinatorial Models (M2DisCo) team of the LIRIS Laboratory in the "Université Claude Bernard Lyon 1", France. His technical research concerns 3D digital image processing and computational geometry.

— Supplementary Material —

Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces

Georges Nader, Kai Wang, Franck Hétroy-Wheeler, and Florent Dupont

This Supplementary Material is organized as follows. Section 1 presents the results of additional psychophysical experiments that were carried out in order to verify the accuracy and robustness of the measured thresholds obtained by our original experiments. Section 2 provides some details about the accuracy and performance of computing the JND threshold, especially for the light independent mode. Section 3 presents an additional subjective experiment that validates the proposed JND profile. In Section 4 we give some details and show additional results for the mesh simplification application, including the generation of mesh levels of details (LODs) and a subjective validation. Section 5 provides some comparison results with mesh perceptual quality metrics for the application of optimal vertex coordinates quantization, as well as an additional subjective validation. Finally, in Section 6 we discuss the difference between the proposed JND model and mesh saliency measures.

1 Psychophysical Experiments

We have conducted additional psychophysical experiments in order to make sure that our previous contrast sensitivity function (CSF) and contrast masking measurements were accurate and robust. The results of this new set of experiments show that the previous measurements are indeed accurate and stable. We present the results of the new experiments in this Supplementary Material and not in the manuscript because we have validated the JND profile in Section 5 of the manuscript using the models fitted by the data from the first set of experiments. In fact, it would be quite time consuming to redo the subjective validation using the new models, and the corresponding results would be very close to the ones presented in the manuscript. Therefore, we would like to present in the manuscript the original results, which however are proven to be accurate and stable.

1.1 Experimental Procedure

We used the same experimental procedure as described in Section 3 of the manuscript. We display two models on the screen one of which has a displaced vertex. The subjects have to answer by “yes” or “no” whether they see a difference between the two models on the screen. The magnitude of the vertex displacement is then regulated using the QUEST procedure [WP83]. 5 new subjects participated in the experiment. None of them was a participant in any of our previous experiments.

1.2 Results

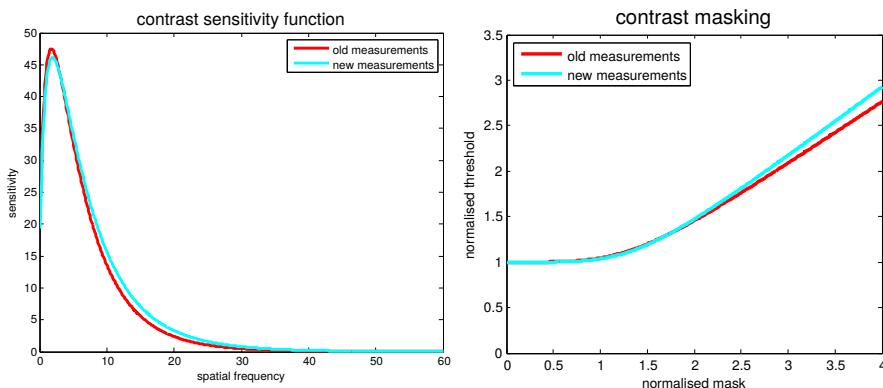


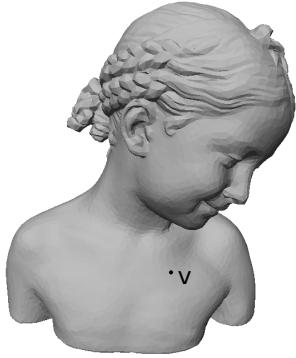
Figure 1: Plot of the fitted contrast sensitivity function and contrast masking models for the data obtained by the old and new experiments.

Figure 1 shows the fitted models from the old and new sets of experiments. The new fit of the CSF model is obtained using Eq. (6) of the manuscript with $a = -13.59$, $f = 0.001$ and $p = 0.62$ while the old fit was computed with $a = -15.13$, $f = 0.0096$ and $p = 0.64$. For the contrast masking model, the new fit is obtained from Eq. (7) of the manuscript with $k_1 = 0.006$, $k_2 = 90.66$, $s = 1.05$ and $b = 4.53$ while the old fit was computed with $k_1 = 0.0078$, $k_2 = 88.29$, $s = 1.00$ and $b = 4.207$.

2 Light Independent JND and Algorithm Speed

Here we present some details that would help readers efficiently implement the proposed JND model in particular for the light independent mode (indeed we plan to freely deliver a reference implementation in the near future). We will also report some theoretical and practical results concerning execution time of the JND computation.

2.1 Threshold Accuracy



local light direction (α, θ)	JND threshold
(10, 85)	0.00195312
(10, 55)	0.003125
(10, 25)	0.0078125
(100, 85)	0.0015625
(100, 55)	0.00390625
(100, 25)	0.0069725
(190, 85)	0.00107422
(190, 55)	0.00234375
(190, 25)	0.0046875
(280, 85)	0.00107422
(280, 55)	0.0021875
(280, 25)	0.004375
(0, 0)	0.015625

Figure 2: The JND threshold of a vertex v computed for different light directions.

The algorithm presented in Section 4.3 of the paper computes the JND threshold for a given light direction. However, in an interactive setting where the light source is fixed relative to the viewpoint, the light direction varies with respect to the 3D mesh. It is therefore important to compute the displacement threshold independently of the light direction.

To do so, we compute the threshold according to multiple light directions and then choose the smallest one. The light independent threshold can then be seen as the one corresponding to the worst possible illumination (*i.e.*, the light direction that makes distortions the most visible). The set of all possible light directions belongs to the sphere around a vertex. However, the contrast between two faces is only defined when the dot product between the light direction and the normals is positive. This means that the set of all possible light can be reduced to the local half sphere in the direction of the unit normal. In practice, we do not need to densely sample all the half sphere. Figure 2 shows the JND threshold obtained from different light directions belonging to the half sphere of a vertex v . We notice that as the light direction approaches the base of the half sphere, the threshold gets smaller. This implies that the worst possible illumination is at the most of time found near the base of the half sphere. This observation can also be deduced from Eq. (3) of the manuscript. Through our testing we notice that it is actually not necessary to densely sample the half sphere in order to obtain an accurate solution. It is observed that the algorithm begins to converge to an accurate JND value with 8 samples as it can be seen in Fig. 3, where the normalized root mean square error (RMSE) is computed with regard to the JND profile obtained with 64 light direction samples (shown in the rightmost of Fig. 3). In practice, in order to obtain the results presented in Sections 4 and 5 of the manuscript, we have used the 12-points sampling, as shown in Fig. 2 (excluding the point $(0, 0)$), which ensures a very good trade-off between threshold accuracy and algorithm speed.

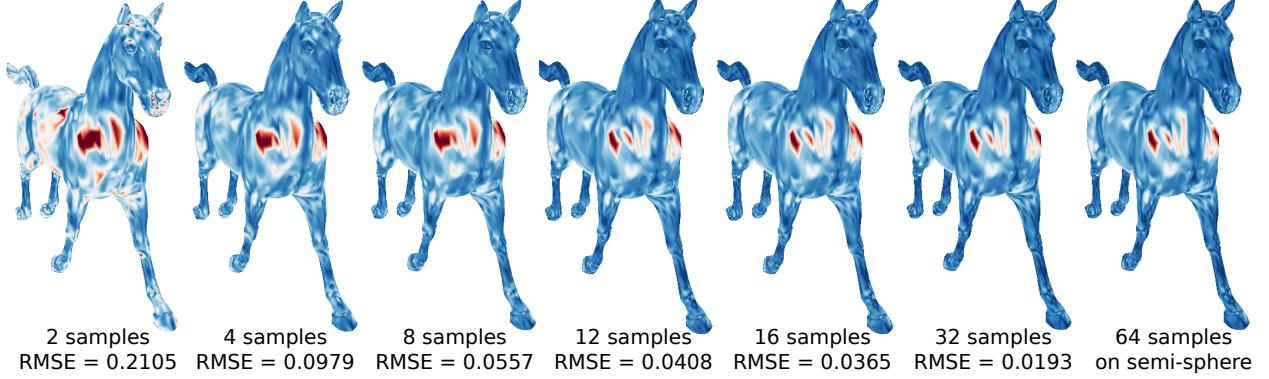


Figure 3: The effect of the number of light samples on the JND computation accuracy. The illustrated JND profiles are computed for noise in the normal direction for each vertex.

2.2 Algorithm Complexity

A theoretical analysis of the proposed JND algorithm shows that the complexity of computing the light independent JND for one vertex is equivalent to:

$$O\left(L \times \log\left(\frac{x_{max}}{x_{precision}}\right)\right), \quad (1)$$

where L is the number of light samples and x_{max} and $x_{precision}$ are respectively the upper displacement bound and the precision used in the half-interval search algorithm (Algorithm 1 of the paper).

This means that the complexity for computing the JND profile of a mesh is :

$$O\left(V \times L \times \log\left(\frac{x_{max}}{x_{precision}}\right)\right), \quad (2)$$

where V is the number of vertices. This shows that as the number of vertices increases the execution time should increase in a linear way at a rate relative to the number of light samples and the precision of the search procedure.

2.3 Execution Time

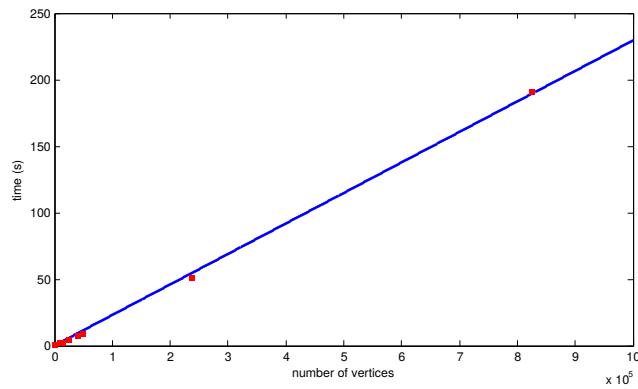


Figure 4: JND profile execution time.

Having adopted a half-interval search algorithm makes finding the JND threshold a very efficient operation. On average computing the JND threshold for a vertex in the light independent mode takes about 7×10^{-4} s. We have used an HP EliteBook 8570w with an i7-3740QM cpu (4 cores) and 16GB of RAM in our computation.

When the number of vertices increases, we have observed that the execution time increases approximately in a linear way (see Fig. 4, also analyzed in Section 2.2 of this Supplementary Material). In addition, since the JND threshold of a vertex is independent from the threshold of other vertices then the JND profile on a mesh can be computed in a parallel way. Using OpenMP, the algorithm performs about three to five times

faster. For a model with 237K vertices, the JND profile took about 52 seconds to compute. Indeed, the vertex displacement threshold searching consumes the most part of the computation time. However, with the simple but efficient half-interval search, this step can be accomplished in a very reasonable time and thus does not harm the efficiency of the whole algorithm.

3 JND Validation

We have conducted an additional subjective experiment in order to validate the proposed JND model. In the manuscript we have presented a subjective experiment where the participants rate the visibility of an injected noise on scale of 0 to 5. It showed that the models with a JND modulated noise were rated the lowest on the visibility scale but meanwhile could tolerate the biggest amount of distortions. However, in the following new experiment we gradually increase the intensity of the injected noise until the subject notices it.

3.1 Experimental Procedure

Similarly to the experiment presented in Section 5 of the manuscript, the noise is injected into the 3D meshes according to the following equation:

$$v_i' = v_i + \text{rnd} \times M(v_i) \times \vec{dir}_i, \quad (3)$$

where v_i is the i^{th} vertex of the initial mesh and v_i' is the corresponding noisy vertex. \vec{dir} is the noise direction. rnd is a random value equal to either +1 or -1 and $M(v_i)$ represents the magnitude of the noise for v_i . It is defined as:

$$M(v_i) = \begin{cases} \beta_{unif} & \text{uniform noise,} \\ \beta_{rough} \times lr(v_i) & \text{roughness modulated noise,} \\ \beta_{jnd} \times jnd(v_i) & \text{JND modulated noise,} \end{cases} \quad (4)$$

where β_{unif} , β_{rough} and β_{jnd} regulate the global noise intensity for each of the noise injection methods. $lr(v_i)$ is the local surface roughness as defined in [WTM12] and $jnd(v_i)$ is the JND value computed as explained in Section 4 of the manuscript.

The idea behind this experiment is to find the minimum noise intensity (β_{unif} , β_{rough} and β_{jnd}) starting from which the participants notice the noise in the model. To do so, we have adapted the same experimental procedure that we have used to measure the local contrast threshold in the studies of CSF and contrast masking. Two models were displayed on the screen, one of which has noise injected. The subjects had to answer by either “yes” or “no” whether they saw the noise on one of the model. The intensity of the noise (β_{unif} , β_{rough} and β_{jnd}) is then adjusted using the QUEST procedure [WP83]. The subjects were allowed to interact with the displayed models by rotating the camera around them. 5 new subjects participated in the experiment.

3.2 Results

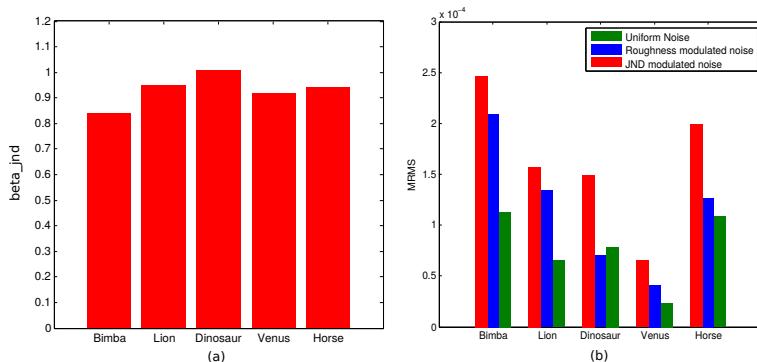


Figure 5: (a) Plot of the measured noise intensity relative to the JND modulated noise. (b) Plot of the MRMS induced by noise injection for three different types of noise at the same visibility level.

Figure 5 displays the results of this subjective experiment. Plot (a) shows the mean measured intensity required to make JND modulated noise visible on a 3D mesh. We see that the measured β_{jnd} is close to 1 for

all of the models, meaning that the proposed JND profile is able to accurately detect the threshold beyond which a noise is visible. Plot (b) shows that the MRMS value of the mesh model with JND modulated noise of just noticeable level is higher than those of the corresponding models with uniform noise or roughness modulated noise at the same visibility level. This means that the JND model is able to tolerate the highest amount of noise among the three candidates, which is what we expected.

4 JND-Driven Mesh Simplification

Integrating the JND model into the mesh simplification process allows us to define a perceptual simplification cost for the edge collapse operation. In order to obtain a simplified model that is visually very similar to the input, we run the simplification process until all the edges have a cost whose value is greater than 0. Moreover, the defined JND profile takes into consideration the size of the display, the viewing distance, and the position of the model in the virtual 3D world. This means that the JND-driven simplification can be useful for generating model level of details (LOD) as it will automatically stop the simplification at different stages for different settings. Figure 6 shows some mesh LODs generated by the JND-driven simplification method at different viewing distances.

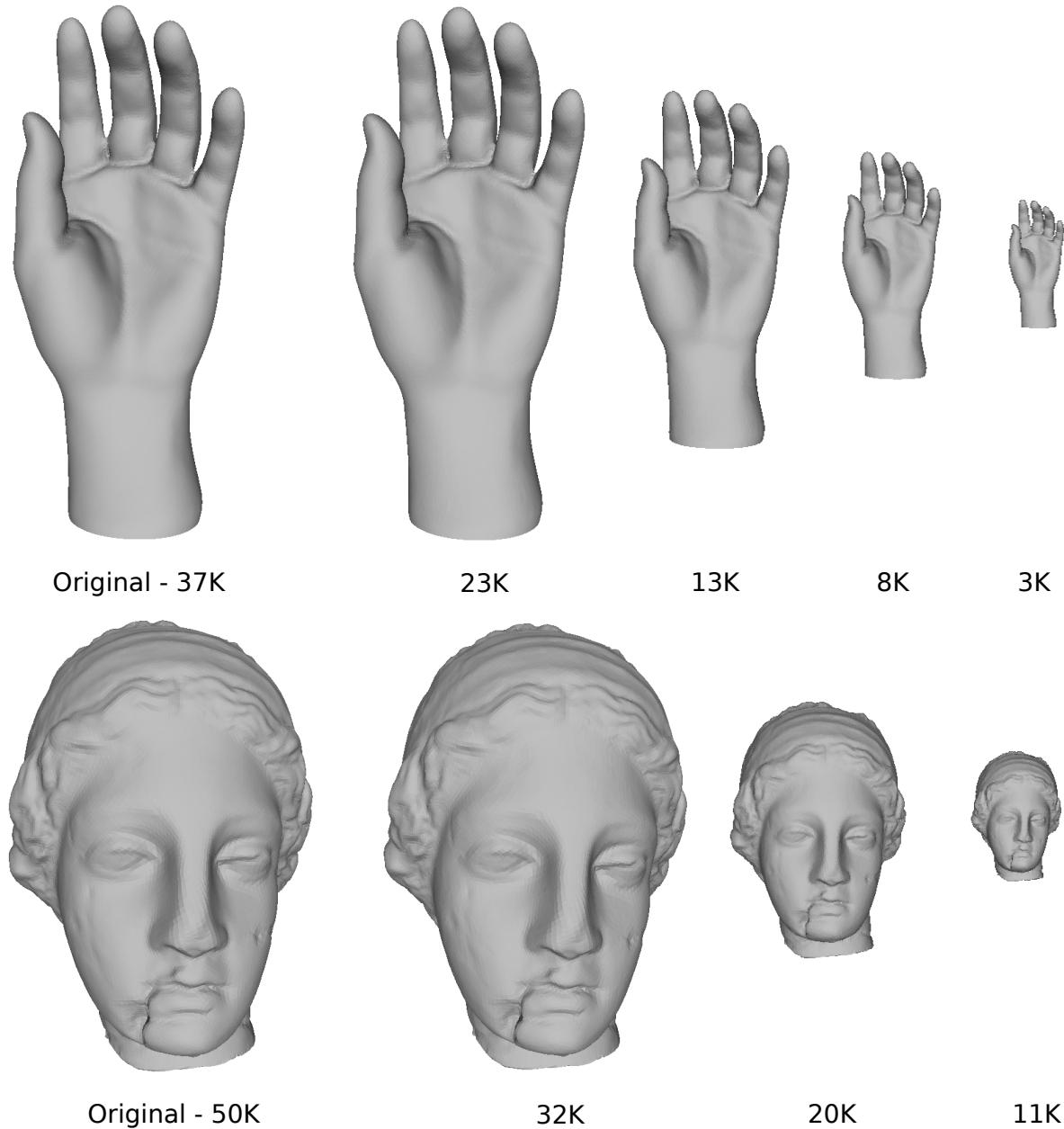


Figure 6: LODs generated using the JND-driven simplification method at different viewing distances. Different adaptive numbers of vertices are obtained under different viewing distances.

4.1 Subjective Validation

We have performed a subjective experiment in order to verify that the proposed JND-driven simplification outputs a simplified mesh that is visually very similar to the original one. 10 subjects participated in the experiment. We have adopted the same experimental procedure as the JND validation experiment presented in Section 5 of the manuscript. We displayed two models on the screen and the participants had to rate the visibility of the distortion on a scale from 0 to 5. 0 corresponds to the case where the subject cannot see any distortion and 5 to the case where the noise is clearly visible. For this experiment we included three models (Angel, Venus and Hand) with three simplified versions each: a JND-driven simplified mesh, a mesh that is further simplified to 5% beyond the JND level and a mesh simplified with Lindstrom and Turk's method to the same number of vertices as the JND-driven simplified mesh. The results of this experiment are shown in Fig. 7. It is clear that the geometric distortion that is due to the simplification process is not visible for the models simplified with the JND-driven method. The average subjective score is below 0.2 for the three models meaning that about 80% of the participants were unable to notice the distortion. For the models where we removed 5% more vertices than the JND level the geometric distortion is visible as the average subjective score given by the participants is greater than 1. As for meshes simplified with Lindstrom and Turk's method to the same number of vertices as the JND level, the mean subjective score is above 0.85 for the three models meaning that most subjects were able to see the distortions introduced by the simplification process.

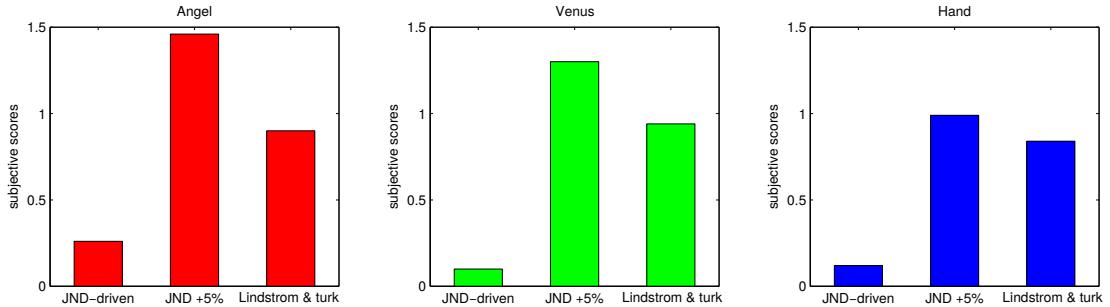


Figure 7: Plot of the subjective scores of the simplified meshes obtained by three different methods.

4.2 Implementation Details

We have implemented a prototype of the proposed perceptual mesh simplification using CGAL's mesh simplification module [Cac15]. The JND-driven simplification algorithm can be summarized as follows. The collapse cost is first computed for all the edges of a mesh and then added to a priority queue. As long as the stopping criterion has not been reached, we pick the edge with the lowest cost and perform the edge collapse operation. After each collapse operation the algorithm will update the simplification costs of the affected edges of the last edge collapse. Accordingly, we update the priority queue. A rather straightforward theoretical complexity analysis shows that the complexity of this family of mesh simplification algorithms is dominated by the complexity of managing the internal data-structure that handles the order by which the edges are being collapsed. This means that both the proposed perceptual mesh simplification and Lindstrom and Turk's simplification methods perform similarly since they are implemented using the same edge-collapse-based mesh simplification framework. From a computation perspective the only difference between the two methods consists in the computation of the edge cost. For a suitable number of light samples (in mesh simplification application we use 8 samples, number from which the JND computation starts to converge, see Section 2.1 of this Supplementary Material), the execution time for computing both the JND-based and Lindstrom and Turk's edge costs is quite similar (see Table 1 of this Supplementary Material).

Table 1: Execution time (in seconds) for computing the JND-based (with 8 light direction samples) and Lindstrom and Turk's edge costs

# of edges	9124	16056	50000	162641
JND-based	0.4	0.76	1.93	3.65
Linstrom and Turk's	0.31	0.65	1.89	3.39

5 Vertex Coordinates Quantization

For the application to optimal vertex coordinates quantization, we compare the JND scores to those of FMPD [WTM12], MSDM2 [Lav11] and MRMS. The comparison results are shown in Fig. 8 of this Supplementary Material (The comparison results of JND scores and FMPD are also illustrated in Fig. 17 of the paper). As shown by this figure, MRMS is not correlated with human perception. For FMPD and MSDM2, it is not possible to define a proper threshold that gives the correct and perceptually relevant optimal quantization levels for all the three models. One possible explanation is that FMPD and MSDM2 have difficulties in producing consistent evaluation results on meshes of different densities and geometric complexities. In addition, the main advantage of the JND model is that it does not require manually defining a threshold. Instead, it can automatically determine the optimal quantization level.

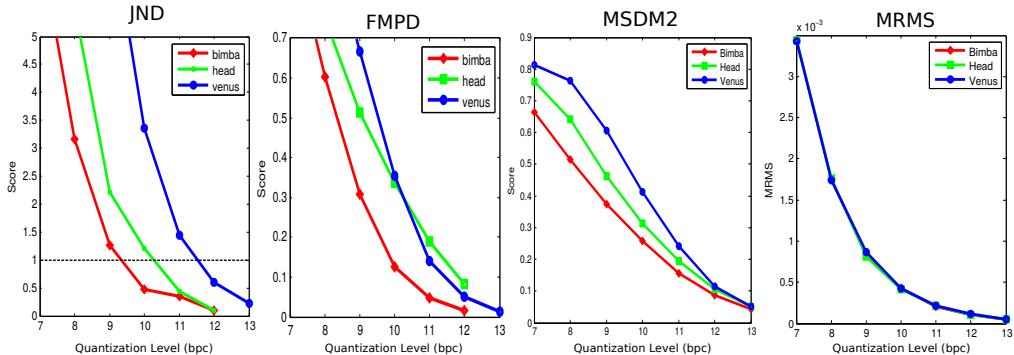


Figure 8: Comparison between the JND scores and those of FMPD [WTM12], MSDM2 [Lav11] and MRMS, versus the vertex coordinates quantization level (in bpc, bits per coordinate).

5.1 Subjective Validation

We have performed a subjective experiment in order to verify that the optimal quantization level obtained using the JND profile is indeed the one beyond which quantization noise becomes visible. 10 subjects participated in the experiment. We have adopted the same experimental procedure as the JND validation experiment presented in Section 5 of the paper. We displayed two models on the screen and the participants had to rate the visibility of the distortion on a scale from 0 to 5. 0 corresponds to the case where the subject cannot see any distortion and 5 to the case where the noise is clearly visible. For this experiment we included three models with five quantization levels each: the optimal quantization level, two immediate higher levels and two immediate lower levels. The results of this experiment are shown in Fig. 9. It is clear that the geometric distortion that is due to vertex quantization becomes visible when the quantization level in bpc become even 1 bit lower than the optimal level. For the quantization levels that are higher than the optimal one, the quantization noise is invisible as the participants rated its visibility by 0. As for the optimal quantization level the average subjective score is between 0 and 1 meaning that some participants were able to barely see the distortions while others were unable to notice it.

As a final remark concerning vertex coordinates quantization application, it is worth mentioning that the coordinates of the original models are never quantized and are represented by high-precision floating numbers.

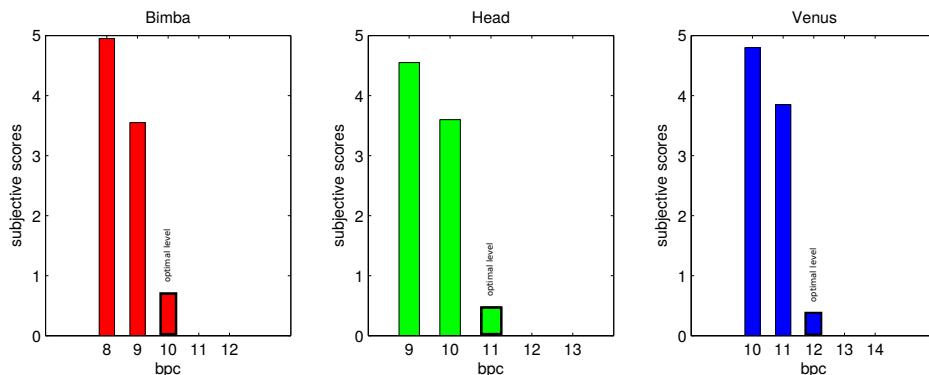


Figure 9: Plot of the subjective scores of quantized meshes with different levels of bpc.

6 JND vs Saliency

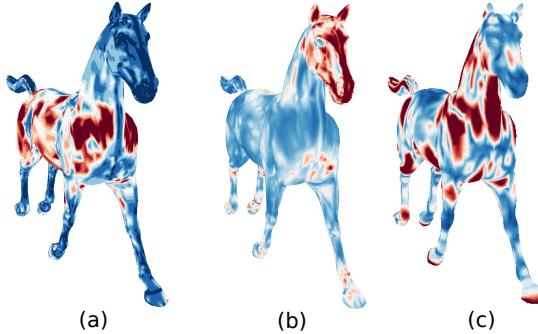


Figure 10: (a) JND profile relative to the normal direction in a light independent mode. (b) JND profile relative to a tangent direction in a light independent mode. (c) Mesh saliency value computed with the method of Lee *et al.* [LVJ05]

Mesh saliency has been the basis of many perceptual geometry processing methods [LVJ05, WSZL13, SLMR14]. By definition saliency is a measure of whether an area is visually attractive for the human visual system or not. For example, Fig 10.(c) shows that the Horse’s neck, extremities of its legs and part of its head are the most visually important features (saliency computed by the method of Lee *et al.* [LVJ05]). A human observer will most likely focus his attention on these areas while observing the Horse model.

On the contrary, the proposed JND model computes the threshold beyond which a distortion becomes visible using low-level properties of the human visual system such as the contrast sensitivity function and the contrast masking effect. The JND profile in Fig. 10 indicates that the Horse’s body can tolerate the most noise in the normal direction. This is because in that region the mesh is coarse, so the visibility threshold is higher due to the CSF property of the human visual system. By contrast, if the noise is in a tangent direction, then the head can tolerate the most noise. This is because in that area the geometry is relatively flat, so the displacement of a vertex in a tangent direction does not cause any change in contrast.

It is clear from the example in Fig. 10 that the saliency and JND profile measure different properties of a 3D mesh. The former points out the visually important regions that are more likely to be observed by a human being, while the latter computes the threshold beyond which a displacement of a vertex becomes visible. Integrating the JND model into a geometry processing application will allow us to (automatically) control the visibility of the introduced distortion. In JND model the main components are low-level properties of the human visual system such as CSF and contrast masking, while in mesh saliency higher-level properties such as visual attention should be taken into account.

However, it would be interesting and possible to use the low-level properties studied in the proposed JND model for the purposes of mesh saliency derivation, since a better understanding of the low-level properties would be helpful for the development of accurate higher-level properties. In particular, a salient region is by definition an area that stands out from its surrounding. It can be attributed to regions where a big change of local contrast occurs. Such regions usually attract human’s visual attention. Having defined a measure of contrast in Section 3 of the manuscript, we think that it would be possible to use it in order to define a saliency measure. In addition, in perceptually oriented mesh processing, it would be beneficial to combine both low- and high-level properties of the human visual system (*e.g.*, both the concept of JND and that of mesh saliency), so as to achieve better performance or to reach a good trade-off.

References

- [Cac15] F. Cacciola. Triangulated surface mesh simplification. In *CGAL User and Reference Manual*. CGAL Editorial Board, 4.7 edition, 2015.
- [Lav11] G. Lavoué. A multiscale metric for 3D mesh visual quality assessment. *Computer Graphics Forum*, 30(5):1427–1437, 2011.
- [LVJ05] C. H. Lee, A. Varshney, and D. W. Jacobs. Mesh saliency. *ACM Transactions on Graphics*, 24(3):659–666, 2005.
- [SLMR14] R. Song, Y. Liu, R. R. Martin, and P. L. Rosin. Mesh saliency via spectral processing. *ACM Transactions on Graphics*, 33(1):1–17, 2014.

- [WP83] A. B. Watson and D. G. Pelli. QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2):113–120, 1983.
- [WSZL13] J. Wu, X. Shen, W. Zhu, and L. Liu. Mesh saliency with global rarity. *Graphical Models*, 75(5):255–264, 2013.
- [WTM12] K. Wang, F. Torkhani, and A. Montanvert. A fast roughness-based approach to the assessment of 3D mesh visual quality. *Computers & Graphics*, 36(7):808–818, 2012.