



HAL
open science

A Genetically Based Combination of Visual Saliency and Roughness for FR 3D Mesh Quality Assessment: A Statistical Study

Anass Nouri, Christophe Charrier, Olivier Lézoray

► To cite this version:

Anass Nouri, Christophe Charrier, Olivier Lézoray. A Genetically Based Combination of Visual Saliency and Roughness for FR 3D Mesh Quality Assessment: A Statistical Study. *The Computer Journal*, 2022, 65 (3), pp.606-620. 10.1093/comjnl/bxaa089 . hal-02938455

HAL Id: hal-02938455

<https://hal.science/hal-02938455v1>

Submitted on 14 Sep 2020

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.

A Genetically Based Combination of Visual Saliency and Roughness for FR 3D Mesh Quality Assessment: A Statistical Study

Anass Nouri, Christophe Charrier, Olivier Lézoray, Hui Wang

► **To cite this version:**

Anass Nouri, Christophe Charrier, Olivier Lézoray, Hui Wang. A Genetically Based Combination of Visual Saliency and Roughness for FR 3D Mesh Quality Assessment: A Statistical Study. The Computer Journal, Oxford University Press (UK), 2020, 00. hal-02938455

HAL Id: hal-02938455

<https://hal.archives-ouvertes.fr/hal-02938455>

Submitted on 14 Sep 2020

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.

A Genetically Based Combination of Visual Saliency and Roughness for FR 3D Mesh Quality Assessment: A Statistical Study

ANASS NOURI^{1,2,*}, CHRISTOPHE CHARRIER³ AND OLIVIER LÉZORAY³

¹*ENSC, Ecole Nationale Supérieure de Chimie, Ibn Tofail University, BP 133, 14000 Kénitra, Maroc*

²*Laboratoire des Systèmes Électroniques, Traitement de l'Information, Mécanique et Énergétique, Département de Physique, Faculté des Sciences, Ibn Tofail University, BP 133, 14000 Kénitra, Maroc*

³*Normandie Univ, UNICAEN, ENSICAEN CNRS, GREYC, 14000 Caen, France*

**Corresponding author: anass.nouri@uit.ac.ma*

In this paper, we present a full-reference quality assessment metric based on the information of visual saliency. The saliency information is provided under the form of degrees associated to each vertex of the surface mesh. From these degrees, statistical attributes reflecting the structures of the reference and distorted meshes are computed. These are used by four comparisons functions genetically optimized that quantify the structure differences between a reference and a distorted mesh. We also present a statistical comparison study of six full-reference quality assessment metrics for 3D meshes. We compare the objective metrics results with humans subjective scores of quality considering the 3D meshes in one hand and the distortion types in the other hand. Also, we show which metrics are statistically superior to their counterparts. For these comparisons we use the Spearman Rank Ordered Correlation Coefficient and the hypothetic test of Student (ttest). To attest the pertinence of the proposed approach, a comparison with a ground truth saliency and an application associated to the assessment of the visual rendering of smoothing algorithms are presented. Experimental results show that the proposed metric is very competitive with the state-of-the-art.

Keywords: objective quality assessment; 3D mesh; visual saliency; patch; visual attention; graphs

Received 22 November 2019; Revised 12 May 2020; Accepted 16 June 2020

Handling editor: Hui Wang

1. INTRODUCTION

The intrinsic goal of each quality assessment algorithm is to come up with an objective scoring scheme for which the prediction error is minimal. This prediction error, which is a performance measure, can be seen as a function that represent the features used for the assessment of quality. Indeed, the effectiveness of a model is greatly linked to the pertinence of the used features. This induces the selection of suitable features in order to obtain a reliable tool for the quality assessment of 3D meshes. Perceptual metrics proposed in the literature, which came after the simple geometric ones, have embedded attributes correlated to the human perception such as curvature or the roughness of the 3D mesh surface. In this paper, we exploit the visual saliency information that represents a

primary feature of the human visual system (HVS). Visual saliency can be defined as the perceptual information that makes some regions of an object (or a scene), represented by 3D meshes, stand out from their surrounding and thus captures the visual attention of the human observer [9]. In this paper, we hypothesize that the distortions localized in perceptually salient regions will induce a strong drop in the global visual quality of the target 3D mesh, while distortions localized at the level of regions with low saliency will less decrease the global quality of the mesh (Fig. 1). Objective quality assessment metrics can be categorized into three families: (i) full-reference metrics, (ii) reduced-reference metrics and (iii) no-reference metrics. We are interested in this paper in full-reference metrics requiring the reference 3D mesh in order to assess the quality of its distorted version. Such a type of metrics

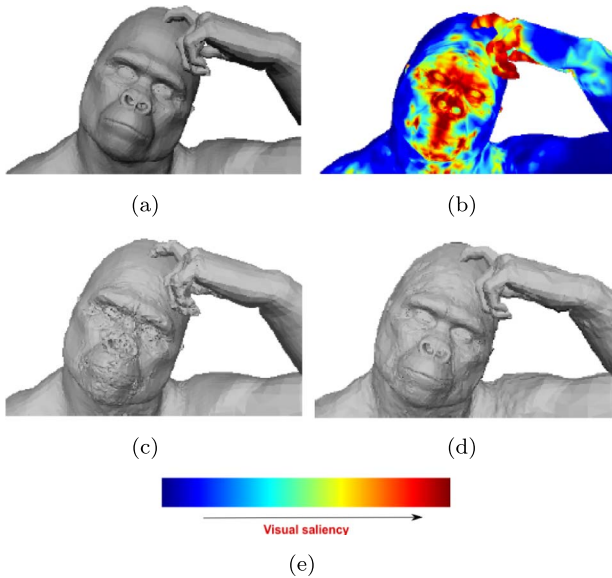


FIGURE 1. Comparison of 3D meshes with different perceptual qualities. (a) Original Gorilla 3D mesh. (b) Saliency map of (a) with [9]. (c) Gorilla 3D mesh noised in more visual attention areas. (d) Gorilla 3D mesh noised in less visual attention areas. (e) colormap.

can be involved into different 3D mesh processings where the quality of the rendering is important regarding to the reference signal. Among these processings, we can cite compression, smoothing, watermarking, remeshing, etc. In addition to the proposed metric named SMQI (Saliency-based Mesh Quality Assessment Index), this paper presents a statistical study where the performance of six well-known 3D full-reference quality assessment algorithms across various 3D meshes and distortions at different localizations are compared using two subject-rated databases. The index selected to measure performance is the SROOC. Moreover, an hypothesis test is conducted to evaluate the statistical significance of each tested algorithm performance. The paper is organized as follows: Section 2 presents the related work. Section 3 presents the motivation of this work as well as the pipeline of the proposed approach. In section 4, we detail the perceptual distance. Section 5 discusses the datasets used and the experimental results. In Section 6, we present an application where the proposed approach can be involved. Section 7 concludes.

2. RELATED WORK

In this section, we quote only viewpoint-independent (or model-based) approaches related to quality assessment and visual saliency of 3D meshes.

2.1. 3D mesh visual quality assessment metrics in the state-of-the-art

Early visual quality assessment metrics were purely geometric. We can cite for example the Hausdorff Distance [1] and the root

mean squared (RMS) error [2]. In the same vein, Karni and Gotsman [3] used the geometric laplacian and vertices coordinates to compute the distance between a reference 3D mesh and its distorted version. Thereafter, Sorkine *et al.* [4] assigned greater weights to the geometric Laplacian. Lately, Gelasca *et al.* [5] considered the roughness property of the surface mesh. Their approach is based on the variation of the roughness measure. The latter is computed as the difference between the initial 3D mesh and its smooth version. Corsini *et al.* [6] also considered the roughness property to assess objectively the quality of a 3D mesh. The roughness measure was derived from the variance of the dihedral angles. Differently, Bian *et al.* [7] proposed an approach based on the strain energy. Strain energy is defined as the energy required to deform the mesh geometry. In [8], Lavoué proposed an extension of the SSIM index developed for 2D images to the quality assessment of 3D meshes called MSDM (Mesh Structural Distorsion Measure). The latter considers the curvature of the mesh surface in order to compute the difference of statistics reflecting the structures of two considered 3D meshes. Thereafter, the same author extended the MSDM metric by considering the multi-scale aspect [9]. Vasa and Rus [10] measured the local variations of dihedral angles in order to provide a score of quality of a distorted mesh regarding its reference version. Wang *et al.* [11] proposed a metric based on the variation of the local roughness which is derived from the Laplacian of the discrete Gaussian curvature. Torkhani *et al.* [12] proposed a metric based on the comparison of the curvature tensors and the roughness character of the geometry. Dong *et al.* [13] proposed an approach based on the visual masking effect, the saturation effect and the curvature in order to assess the visual quality of a 3D mesh. Regarding model-based full-reference 3D mesh visual metrics that integrate 3D visual saliency in their pipeline, we can notice that there are relatively fewer works in the literature in comparison with 2D image metrics [14–18]. Indeed, only one approach was proposed [19]. This approach predicts the final quality score based on a spatial pooling strategy, visual saliency weighting and statistical descriptors obtained from a distortion map [12] along with a support vector regression model. Recently, Lin *et al.* [20] proposed a blind mesh quality assessment metric based on graph spectral entropy and spatial features where the score of quality is predicted by the use of a random forest regression.

2.2. 3D mesh visual saliency approaches in the state-of-the-art

In 2005, Lee *et al.* [21] proposed the first approach that detects visual saliency on 3D mesh surfaces. To do this, they used a center-surround operator on Gaussian-weighted curvatures computed in a local neighborhood at multiple scale. Wu *et al.* [22] proposed a method for visual saliency detection based on a local descriptor that measures the local height field in the neighborhood of each vertex [23], from which Zernik moments are computed. The local saliency is first computed after the

segmentation of the surface mesh. The global saliency of each vertex is then computed as the interpolation of the degrees of saliency of the closest patches. The final saliency degree of a vertex is obtained by combining and normalizing values of global and local saliency. Leifman *et al.* [24] proposed a descriptor that combines the vertex distinctness and shape extremities. 2D spin image histograms [25] are computed on each vertex to describe the mesh surface and a similarity measure between them is defined in order to obtain the vertex distinctness. Zhao *et al.* [26] presented a saliency model for attentional guided simplification of 3D meshes. The 3D target mesh is filtered using a Gaussian filter, then geometric attributes such as mean curvature and curvature directions are computed on multiples scales. Afterwards, the obtained maps are filtered using a median filter and combined in order to produce the final saliency map. Song *et al.* [27] suggested to use the Conditional Random Field (CRF) framework for saliency detection of 3D meshes. In another study, Zhao *et al.* [28] detects saliency by enhancing the local details with the Retinex theory [29]. The surface mesh is segmented into patches, and the saliency is quantified by computing the distance between the resulting segments. Song *et al.* [30], proposed a spectral method for mesh saliency detection by analyzing the irregularity spectrum. High frequencies represent in this context the saliency. Wang *et al.* [31] use low rank with a sparse analysis of features encoding both local and global features of the mesh surface. In a different manner, Chen *et al.* [32] proposed a regression model to detect the visual saliency of 3D meshes. Their model is based on subjective data collected in an online experimentation in which observers selected points that might be selected by other observers. Song *et al.* [33] proposed a full-reference deep learning approach by using a network weakly trained. The used ground truth refers to the class membership of the used 3D meshes in the training phase. Lau *et al.* [34] introduced the concept of *Tactile mesh saliency*. In this study, salient vertices are those that a human is more likely to grasp. To detect these tactile points, deep learning associated to learning-to-rank methods were used on collected crowdsourcing data. Always based on saliency, Abouelaziz *et al.* [35, 36] proposed a no-reference deep learning approach where the convolutional network is fed with 2D saliency patches in order to estimate the visual quality of a 3D mesh.

3. THE PROPOSED METRIC

3.1. Visual attention and mesh visual quality assessment

To assess the perceptual quality of a 3D distorted mesh, we make the following assumption: the visual quality of a 3D mesh is more depreciated when the distortions are localized in its salient areas. On the contrary, the visual quality is less depreciated when distortions affect less or non salient areas. Figure 1 illustrates this case. We can remark that the proposed approach succeeds in assessing the visual quality in a way that it provides a lower quality score to the 3D mesh distorted in salient areas

(low visual quality) and a higher quality score to the 3D mesh distorted in less or non salient areas (better visual quality). Such a behavior correlated to human perception is not reached by the state-of-the-art full-reference algorithms on which we focus: TPDM [12], MSDM2 [9] and FMPD [11]. Indeed, they provide a higher score of quality to the 3D mesh distorted in visual salient areas and a lower score to the one affected in less or non salient areas, which is an unexpected result. The principal novelties of this paper rely on several key points:

1. The use of visual saliency for the full-reference (FR) quality assessment of 3D meshes.
2. A combination of two properties that the human visual system is sensitive to: visual saliency and roughness.
3. The use of a leave-one-out training based genetical algorithms for tuning the parameters associated to the combination of the two cited properties.
4. Performance comparison (3D mesh based performance and distortion based performance) study of six full-reference 3D mesh metrics based on the Spearman correlation coefficient index (SROOC), the Pearson Linear Correlation Coefficient (PLCC) and the statistical hypothesis test (test of Student).
5. The validation of the used saliency measure with eye-tracking experimentation.

3.2. Overview

The proposed metric is based on the idea that a modification of the saliency of a 3D mesh impairs its visual quality. Indeed, we suppose that the distortion of salient regions on the reference mesh which become less or more salient will decrease inevitably the visual quality. For this, the proposed approach uses a multi-scale saliency map as a basis to compute statistics on local corresponding neighborhoods that reflect the structures of the reference and the distorted meshes. Furthermore, in order to capture the visual masking effect which may happen when a rough region remains rough after being distorted, we consider a roughness map from which we compute the differences of mean roughness between corresponding neighborhoods. These statistics are used in four comparison functions optimized genetically for the prediction of the objective score of quality. Figure 2 illustrates the metric's pipeline.

4. PERCEPTUAL DISTANCE

4.1. Saliency comparison functions

To extract the structural information of a 3D mesh, we consider the multi-scale saliency map proposed in our previous work [9] from which we compute the mean, standard-deviation and covariance of saliency. Indeed, the saliency map associates a degree of saliency to each vertex of the mesh surface depending on its geometric attractiveness. For a local neighborhood $N(v_i)$

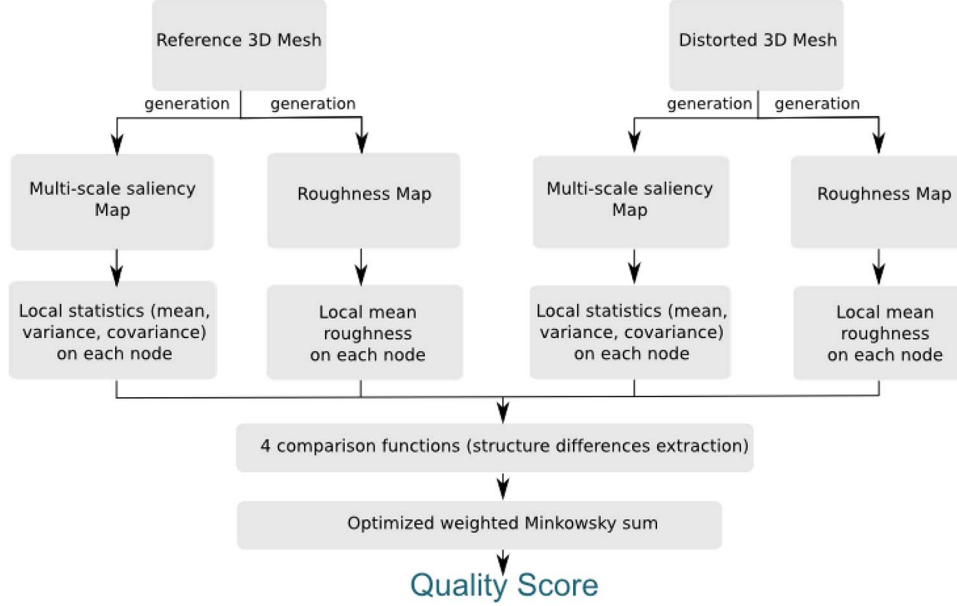


FIGURE 2. Block diagram illustrating the pipeline of the proposed perceptual metric.

representing the adjacent vertices of v_i on the mesh surface, we define the local mean saliency and the standard deviation respectively denoted $\mu_{N(v_i)}$ and $\sigma_{N(v_i)}$ as

$$\mu_{N(v_i)} = \frac{1}{|N(v_i)|} \sum_{v_j \in N(v_i)} MS(v_j) \quad (1)$$

$$\sigma_{N(v_i)} = \sqrt{\frac{1}{|N(v_i)|} \sum_{v_j \in N(v_i)} (MS(v_j) - \mu_{N(v_i)})^2}, \quad (2)$$

where $|N(v_i)|$ is the cardinality of the adjacent vertices of v_i .

For two corresponding local neighborhoods $N_1(v_i) = N_{\mathcal{M}_1}(v_i)$ and $N_2(v_i) = N_{\mathcal{M}_2}(v_i)$ from two 3D meshes \mathcal{M}_1 and \mathcal{M}_2 , we define the covariance σ_{xy} as follows:

$$\sigma_{N_1(v_i)N_2(v_i)} = \frac{1}{|N_1(v_i)|} \sum_{v_j \in N_1(v_i), N_2(v_i)} (MS_{\mathcal{M}_1}(v_j) - \mu_{N_1(v_i)}) (MS_{\mathcal{M}_2}(v_j) - \mu_{N_2(v_i)}), \quad (3)$$

where $MS_{\mathcal{M}_1}$ and $MS_{\mathcal{M}_2}$ represent the multi-scale saliency maps of the two compared meshes \mathcal{M}_1 and \mathcal{M}_2 , respectively.

Once the statistics computed, we define three comparison functions between corresponding neighborhoods $N_1(v_i)$ and $N_2(v_i)$ to quantify the deformation that affect the structural information of the distorted 3D mesh:

$$L(N_1(v_i), N_2(v_i)) = \frac{\|\mu_{N_1(v_i)} - \mu_{N_2(v_i)}\|_2}{\max(\mu_{N_1(v_i)}, \mu_{N_2(v_i)})} \quad (4)$$

$$C(N_1(v_i), N_2(v_i)) = \frac{\|\sigma_{N_1(v_i)} - \sigma_{N_2(v_i)}\|_2}{\max(\sigma_{N_1(v_i)}, \sigma_{N_2(v_i)})} \quad (5)$$

$$S(N_1(v_i), N_2(v_i)) = \frac{\|\sigma_{N_1(v_i)}\sigma_{N_2(v_i)} - \sigma_{N_1(v_i)N_2(v_i)}\|_2}{\sigma_{N_1(v_i)}\sigma_{N_2(v_i)}}, \quad (6)$$

where L , C and S refer to the saliency comparison, the contrast comparison and the structure comparison, respectively.

Our choice of using a saliency approach proposed in an anterior work [37] instead of a recent approach of the state-of-the-art is motivated by the accuracy of our saliency measure. Indeed, we have compared its saliency maps with saliency maps obtained from the *Eye Link* eye-tracker used in the psycho-visual experimentation proposed in [38] where 70 observers have been involved in order to detect the direction of their gaze when staring at 16 physical 3D objects. Figure 3 presents this comparison where the fixation results are encoded as vertex colors (red color on the shapes is associated to highly salient regions while dark blue color is associated to a weak saliency). In order to estimate the deviation between the saliency maps provided by our approach [37] and the ground truth saliency maps, we used the NMSE (Normalized Mean Squared Error). Indeed, the NMSE is an estimator of the overall deviations between predicted and measured values. NMSE interval is $[0, 1]$. Low NMSE is a metric for the validation of a predicting model. The smaller the NMSE, the more the predicting model is accurate. From Fig. 1, we can notice that the NMSE values are very low. This confirms the correlation between our predictive saliency model and the human visual perception.

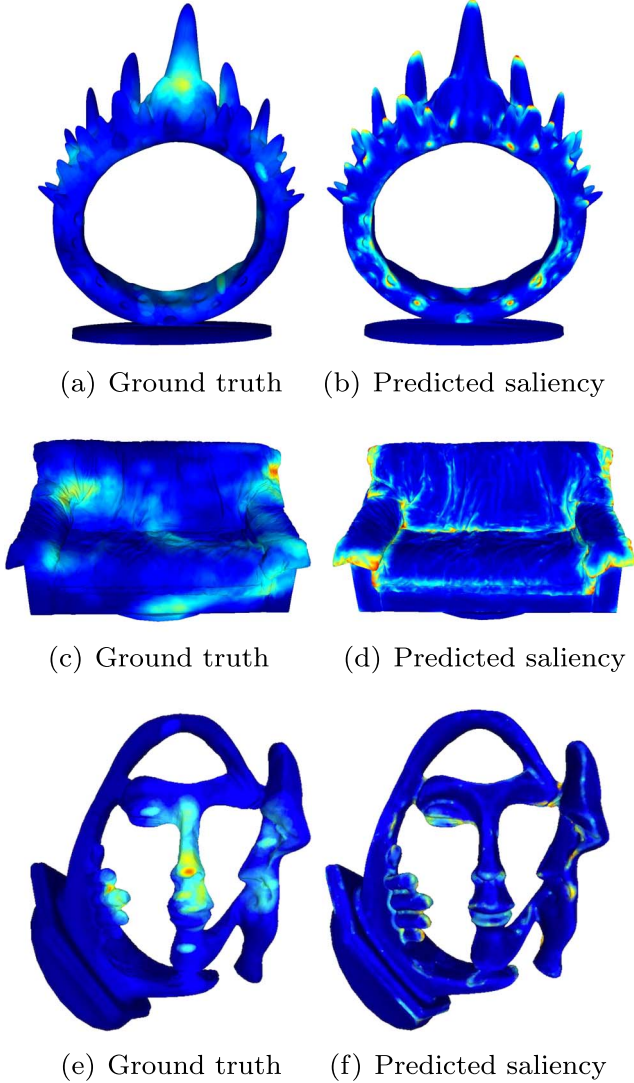


FIGURE 3. Comparison of the used saliency measure [37] with a ground truth (GT) based eye-tracking [38]. The ground truth results from the eye-tracker represent the aggregated human eye fixations regardless of the viewing directions. The NMSE between the ground truth saliency of the presented 3D meshes and their associated predicted saliency are $NMSE(a,b) = 0.04$, $NMSE(c,d) = 0.05$ and $NMSE(e,f) = 0.028$

4.2. Roughness comparison function

Once we have defined the saliency comparison functions, we noticed that the visual masking effect on the 3D meshes is not well captured by these functions when a rough region is present on the reference surface mesh. Indeed, given a rough and a smooth region, a distortion will be much more visible on the smooth region than on the rough one. To deal with this problem, we implemented the work described in [11] that provides a roughness map of a 3D mesh based on the Laplacian of the Gaussian curvature. The first stage of the roughness

computation pipeline begins by defining the Gaussian discrete curvature as:

$$CG_i = |2\pi - \sum_{j \in N_i^F} \alpha_j|, \quad (7)$$

where N_i^F represents the set of neighboring faces of the vertex v_i and α_j is the angle between the two incident edges of the face j in v_i . We then compute the mesh Laplacian matrix as:

$$\left\{ \begin{array}{l} D_{ij} = \frac{\cot(\hat{\beta}_{i,j}) + \cot(\hat{\beta}'_{i,j})}{2} \quad \text{for } j \in N(v_i) \\ D_{ii} = -\sum_j D_{ij} \end{array} \right\}, \quad (8)$$

where $\hat{\beta}_{i,j}$ and $\hat{\beta}'_{i,j}$ are the 2 opposite angles to the edge $e(v_i, v_j)$, respectively. The roughness is evaluated as a weighted difference between CG_i and the Gaussian curvatures of the neighbors, where the weights are determined according to the entries in the Laplacian matrix:

$$LR_i = \left| CG_i - \frac{\sum_{j \in N_i^V} D_{ij} \cdot CG_j}{\sum_{j \in N_i^V} D_{ij}} \right| \quad (9)$$

$$= \left| CG_i - \frac{\sum_{j \in N_i^V} D_{ij} \cdot CG_j}{D_{ii}} \right| \quad (10)$$

The second stage consists in modeling the local roughness by a power function to model the non-linearity of the masking effect. The basic idea is to induce a large visual difference result when a smooth region becomes a rough region (or vice versa) after modification and to induce a small difference when an originally rough region still remains rough after modification. The power function is defined as:

$$LRM_i = f(LR_i) = (LR_i)^a - (Th_l)^a, \quad (11)$$

where a is a parameter that controls the shape function. In order to account for the psychometric saturation effect, the range of the roughness is bounded to lie in the interval $[Th_l, Th_j]$.

The third stage permits to reduce the influence of the values that are greater than the mesh's average roughness. This average roughness (before any modulation) \overline{LR} of the mesh is defined as:

$$\overline{LR} = \frac{\sum_i LR_i \cdot s_i}{\sum_i s_i}, \quad (12)$$

where s_i is one third of the total area of the incident facets of v_i and $\sum_i s_i$ is the total area of the triangular manifold mesh. Then, this average is modulated as $\overline{LRM} = f(\overline{LR})$ to set a threshold $Th_{LRM} = \overline{LRM}$ for reducing any modulated roughness that is greater than this threshold according to the following equation:

$$LRF_i = Th_{LRM} + b(LRM_i - Th_{LRM}) \quad \text{for } LRM_i > Th_{LRM}, \quad (13)$$

where LRF_i is the final local roughness at v_i and b is a parameter that controls the magnitude of the reduction. As in [11], the parameter values used for all the experiments were chosen as follows: $Th_l = 5.0 \times 10^{-4}$, $Th_h = \max\{0.2, 5Th_{LRM}\}$, $a = 0.15$ and $b = 0.50$.

Consequently, we introduced a fourth function based on the comparison of the mean local roughness. The aim of this function is to induce a large difference when a smooth region becomes a rough region and is defined as follows:

$$R(N_1(v_i), N_2(v_i)) = \frac{\|\delta_{N_1(v_i)} - \delta_{N_2(v_i)}\|_2}{\max(\delta_{N_1(v_i)}, \delta_{N_2(v_i)})} \quad (14)$$

with $\delta_{N_1(v_i)} = \frac{1}{|N_1(v_i)|} \sum_{v_j \in N_1(v_i)} \text{RoughnessMap}(v_j)$. It is important to note that a saliency map is different from a roughness map, since only novel and non-redundant informations are pointed on a saliency map (see Fig. 4).

Finally our Saliency-based Mesh Quality Index (SMQI) between two 3D meshes \mathcal{M}_1 and \mathcal{M}_2 is defined by a weighted Minkowsky sum of their local distances:

$$\begin{aligned} SMQI(\mathcal{M}_1, \mathcal{M}_2) = & \left(\frac{1}{|V|} \sum_{|V|} L(N_1(v_i), N_2(v_i)) \right)^\alpha + \\ & \left(\frac{1}{|V|} \sum_{|V|} C(N_1(v_i), N_2(v_i)) \right)^\beta + \left(\frac{1}{|V|} \sum_{|V|} S(N_1(v_i), N_2(v_i)) \right)^\gamma \\ & + \left(\frac{1}{|V|} \sum_{|V|} R(N_1(v_i), N_2(v_i)) \right)^\delta, \end{aligned} \quad (15)$$

where α, β, γ and δ are obtained from an optimization based on genetic algorithms. Details on the genetic optimization will be given in the experiments section.

5. EXPERIMENTAL RESULTS

5.1. Datasets

In order to assess the effectiveness of the proposed approach and compare it against the state-of-the-art, we consider two subject-rated databases that serve as ground truth: (i) The Liris/Epfl General Purpose Database [8] and (ii) The Liris-Masking database [39]. The first database contains 88 3D meshes in total, obtained from four reference 3D meshes affected with two types of distortions: noise addition and smoothing. Different localisations have been considered to affect the reference 3D meshes: uniformly over the mesh surface, specifically on rough or smooth regions (visual masking simulation) and specifically on transitional areas between rough and smooth ones.

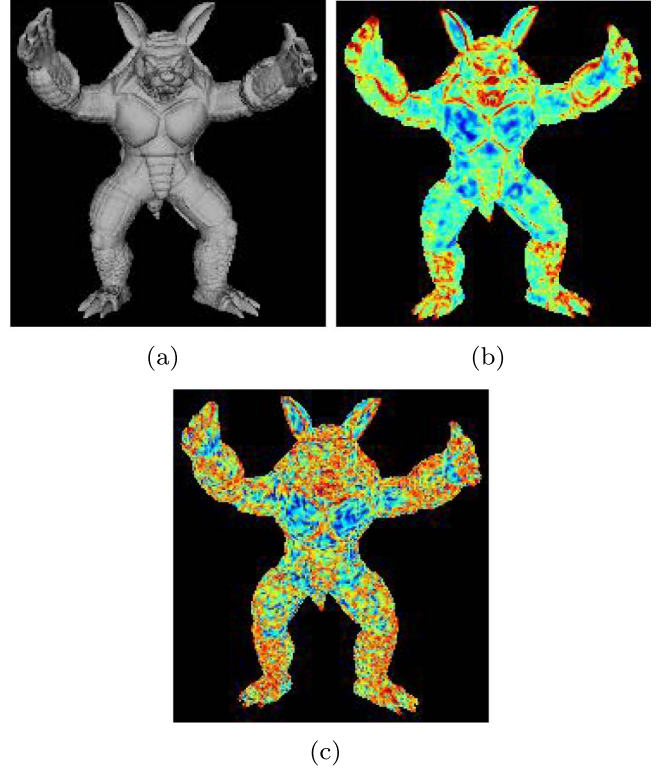


FIGURE 4. Difference between a multi-scale visual saliency map and a roughness map: (a) The original 3D mesh Armadillo, (b) its multi-scale saliency map and (c) its roughness map. Note that only novel and non-redundant informations are pointed on the saliency map.

The second database consists of 24 3D meshes obtained from four reference 3D meshes. The only considered distortion is the Noise Addition which is performed either on smooth areas or rough areas. The goal of such a database is to evaluate if a novel metric takes into account the visual masking effect enough in its pipeline. For these two datasets, the distortions have been performed according to three different strengths. 12 human observers have evaluated the quality of the 3D meshes to obtain the MOS (Mean Opinion score) values.

The performance of the proposed approach is measured by the Spearman Rank Ordered cORrelation Coefficient (SROOC) as well as the Pearson Linear Correlation Coefficient (PLCC) between the resulting objective scores of quality and the subjective ones (MOS) obtained from the subject-rated datasets. To compare the proposed approach, we chose seven full-reference well-known metrics of which the source code is available online. Two of which are purely geometric (no HVS features are considered in their pipeline), and the five remaining are perceptual-based metrics: HD [1], RMS [2], 3WDPM2 [6], MSDM2 [9], DAME [10], FMPD [11] and TPDM [40].

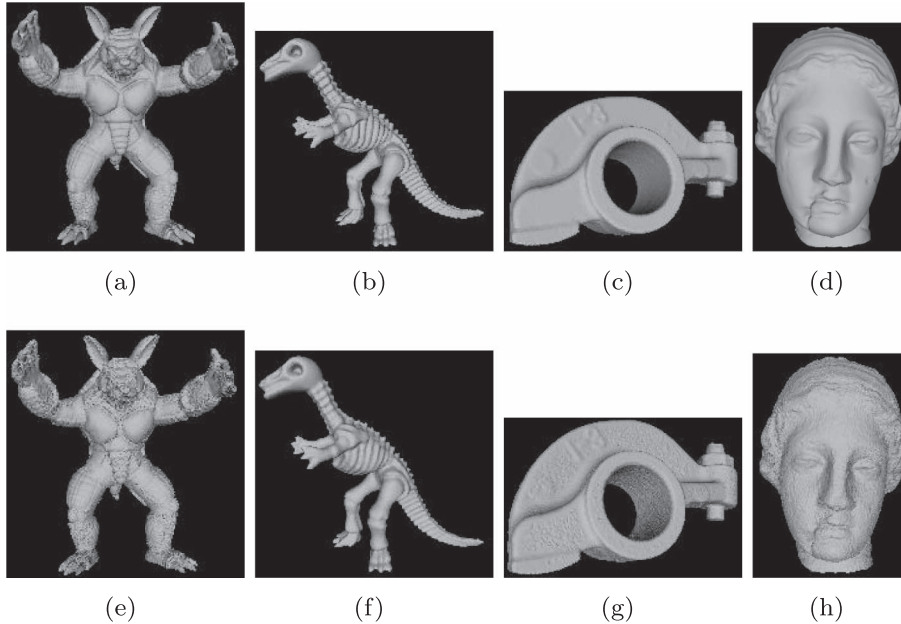


FIGURE 5. Example of 3D meshes from the Liris/Epfl General-Purpose database. In the top row are the four reference 3D meshes. In the bottom row are four examples of deformed 3D meshes, from left to right are respectively: Armadillo with noise on rough regions (MOS = 0.84), Dinosaur after a uniform smoothing (MOS = 0.43), RockerArm with noise on smooth regions (MOS = 0.75) and Venus with uniform noise (MOS = 1).

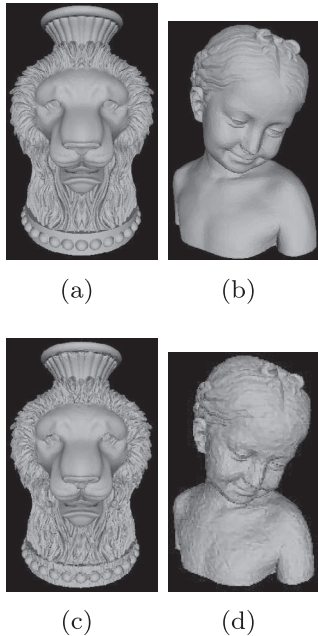


FIGURE 6. Example of 3D meshes from the Liris-Masking database. In the top row are two reference 3D meshes. In the bottom row are two examples of their distorted versions from left to right are respectively: Lion 3D mesh with noise on rough regions (MOS = 0.20), Bimba 3D mesh with noise in smooth regions (MOS = 1.0).

5.2. Results

The final perceptual distance equation includes four parameters (α , β , γ and δ). These are independent and may be difficult to tune manually. To remedy this, we chose a genetically-based optimization associated with a Leave-One-Out training since the two considered datasets (Liris/Epfl General Purpose and Liris-Masking) contains few 3D meshes. The goal of this approach is to perform the learning of the model on $k - 1$ observations and to validate it on the k^{th} one. This process is replicated $k \times 999$ times using with bootstrap on the Liris/Epfl General Purpose database. k refers to an observation. In our context, an observation is associated to MOS values associated to a reference 3D mesh and its 22 distorted versions. Since we have four reference 3D meshes, k is fixed to 4. The fitness function used to perform the genetic optimization is defined as:

$$f(\alpha, \beta, \gamma, \delta) = \sqrt{\sum_{i=0}^{k-1} (MOS_i - SMQI_i(\mathcal{M}_1, \mathcal{M}_2))^2}, \quad (16)$$

where MOS_i is the vector of MOS values of the observation i and $SMQI_i(\mathcal{M}_1, \mathcal{M}_2)$ is the perceptual distance computed with equation (15). After genetic optimization, we obtain $\alpha = 23.63$, $\beta = 3.26$, $\gamma = 5.04$ and $\delta = 0.77$. Note that in Fig. 1, the Gorilla 3D mesh was assessed with these parameters.

5.3. 3D mesh-based performance on the Liris/Epfl General-Purpose and Liris-Masking databases

Before computing the correlation coefficients between the obtained objective quality scores and the ground truth quality (MOS) scores provided by the subject-rated databases, we conduct a psychometric fitting between the previous [8] FMPDTPDM [13, 41]. Indeed, nonlinear quality rating may exist during the subjective rating experiments which induce a nonlinearity between the human scores and objective predictions quality. To correct this, we apply the cumulative Gaussian function to conduct the psychometric fitting [42]:

$$g(m, n, R) = \frac{1}{\sqrt{2\pi}} \int_{m+nR}^{\infty} e^{-t^2} dt, \quad (17)$$

where R is the objective distance. The psychometric fitting is conducted on the sample pairs which are constituted from the ground truth scores of quality and the calculated objective quality scores using the nonlinear least squares method in order to obtain the parameters m and n . Once these parameters are obtained, we transform objective quality values to a set of predicted MOS values in order to compute the correlation coefficients between the predicted MOS values and the ground truth MOS values.

Table 1 presents the SROOC values obtained with the proposed approach and the state-of-the-art metrics when the Liris/Epfl General-Purpose database is considered as a ground truth. We can notice that SMQI provides important correlation values for all the 3D meshes and particularly for the Venus 3D mesh where the PLCC value is the highest. Moreover, it appears that our proposed metric SMQI is the second best metric after TPDM over the entire database (the (SROOC, PLCC) values are (89.6%, 82.2%) for TPDM, (84.6%, 84.3%) for SMQI and (81.9%, 83.5%) for FMPD).

We have also tested and compared our proposed metric with the state-of-the-art metrics on the Liris-Masking database. The aim of such a comparison using the Liris-Masking database is to evaluate the capacity of quality assessment metrics to deal with the visual masking effect. Table 2 provides the Spearman correlation values of different metrics on this database. From these results, we can notice that SMQI is very competitive with the state-of-the-art approaches (between 83% and 100% correlation rate with human observers scores of quality). Otherwise, one correlation value associated to the Lion-vase 3D mesh is slightly lower in comparison with the correlation values provided by TPDM and MSDM2 metrics. We think that the computed multi-scale saliency map for this 3D mesh does not bring out the distorted regions.

We do not provide the correlation over the entire database. Indeed, in this context and on the contrary of Liris/Epfl General-Purpose database, the subjective evaluation protocol used while designing the Liris-Masking database have established the referential range for the rating separately for

TABLE 1. PLCC(r_p) and SROOC (r_s) values (%) of different viewpoint-independent metrics on the LIRIS/EPFL General Purpose database

Liris/Epfl G.P	HD [1, 2]		RMS [1, 2]		3DWPM2 [6]		MSDM2 [9]		DAME [10]		FMPD [11]		TPDM [40]		SMQI	
	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s
Armadillo	54.9	69.5	56.7	62.7	65.6	74.1	85.3	81.6	76.3	60.3	83.2	75.4	78.8	84.5	74.3	77.5
Venus	8.9	1.6	87.9	90.1	40.5	34.8	87.5	89.3	83.9	91	83.9	87.5	91.0	90.6	90.0	91.6
Dinosaur	47.5	30.9	0.0	0.3	44.6	52.4	85.7	85.9	88.9	92.8	88.9	89.6	89.0	92.2	80.3	84.8
RockerArm	23.4	18.1	17.3	7.3	54.7	37.8	87.2	89.6	80.1	85.8	84.7	88.8	91.4	92.2	92.6	91.8
Entire Database	11.4	13.8	28.1	26.8	49.6	49.0	81.4	80.4	75.2	76.6	83.5	81.9	86.2	89.6	84.3	84.6

each 3D mesh and therefore the correlation values over the whole set of 3D meshes are not meaningful [43].

From these results, it appears that SMQI succeeds in assessing the visual quality with the integration of visual saliency. The correlation rates associated to the Liris/Epfl General-Purposes database show that SMQI is the second best metric.

5.4. Distortion-based performance on the Liris/Epfl General-Purpose

In contrary to the precedent section, we present here the distortion-based correlation rates between the predicted quality scores provided by the FR-MVQA (Full-Reference Mesh Visual Quality Assessment) metrics and the human subjective quality scores from the LIRIS/EPFL General Purpose Database. The correlation is evaluated regarding each distortion type across different 3D meshes. The aim is to analyze the capability of the proposed approach and the state-of-the-art methods to assess the perceptual quality when the distortion is fixed and the 3D mesh is changed [44]. Table 3 presents the obtained SROOC and trials FR-MVQA algorithms when the noise distortion is considered. The noise is applied at different localizations on the surface mesh: uniformly, on rough regions, on intermediate regions (between smooth and rough areas) and smooth regions. From this table we can notice that the proposed approach is highly competitive with the state-of-the-art approaches in terms of correlation with human perception despite the different localizations of the distortion. Moreover, it obtains the highest correlation rate ((r_p, r_s) : (84%,79.6%)). Table 4 presents the obtained SROOC when the smooth distortion affects the 3D meshes. In comparison with the correlation values obtained when the noise distortion is considered, we can notice that the state-of-the-art metrics along with SMQI are less correlated to human scores of quality. The three most important correlation rates rise respectively to (r_p : 80.9, r_s : 72.8% - MSDM), (r_p : 63%, r_s : 69% - FMPD), (r_p : 77.3%, r_s : 77% - TPDM) and (r_p : 75.6, r_s : 67% - SMQI). All remaining metrics get SROOC values lesser than 60%. This can be explained by the fact that human beings tend to judge the smooth effect positively. Indeed, a smoothed surface will often be considered less distorted than a noisy one in comparison to their reference state.

5.5. Statistical significance and hypothesis testing

In order to establish which differences between FR-MVQA are statistically significant, we applied an hypothesis test using MOS values and the ratings provided by the trial FR-MVQA algorithms. This test is based on the t-test (Student test) that determines whether two populations means are equal or not. This yields us to take a statistically-based conclusion of superiority (or not) of an FR-MVQA metric. To the best of our knowledge, evaluating the statistical significance between two 3D quality metrics has never been done before. Four statistical

TABLE 2. PLCC(r_p) and SROOC (r_s) values (%) of different viewpoint-independent metrics on the LIRIS-Masking database

Liris Masking	HD [1, 2]		RMS [1, 2]		3DWPM2 [6]		MSDM2 [9]		DAME [10]		FMPD [11]		TPDM [40]		SMQI	
	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s
Armadillo	61.4	48.6	66.8	48.6	61.6	48.6	81.1	88.6	96.0	94.3	94.2	88.6	91.4	88.6	83.35	88.6
Lion	50.1	71.4	48.8	38.3	46.9	38.3	93.5	94.3	99.5	100	93.5	94.3	88.4	82.9	94.4	83.0
Dinosaur	55.8	48.6	70.9	71.4	70.8	71.4	95.6	100	89.4	82.9	96.9	94.3	97.1	100	95.8	100
Bimba	27.4	25.7	46.7	37.1	37.9	37.1	96.8	100	88.0	97.7	98.9	100	97.2	100	99.85	100

TABLE 3. PLCC(r_p) and SROOC (r_s) values (%) of different viewpoint-independent metrics on the Liris-General purpose database considering the Noise distortion localisations.

Liris/Epfl G.P	HD [1, 2]		RMS [1, 2]		3DWPM2 [6]		MSDM2 [9]		FMPD [11]		TPDM [40]		SMQI	
	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s
Noise unif.	57.3	63	62.4	79	87.2	88	80.17	78	84.6	84.6	82	84.6	60.4	63
Nose rough	61.4	71.3	54.1	58.7	84.9	87.9	74.9	64	70.4	61	82	79.7	72.1	75
Noise interm.	22.9	26	21.6	28	69.9	75	16.7	50	49.2	53	59.2	69.9	69.4	74.9
Noise smooth	12.9	2	45.7	54	87.2	85	65.6	62	90.3	87	80.4	76.9	87.9	87.5
All noise localisations	17.4	26.2	48.4	53	75	77	68.4	69.3	77.8	75.3	78	78.1	79.6	84

TABLE 4. PLCC(r_p) and SROOC (r_s) values (%) of different viewpoint-independent metrics on the Liris-General purpose database considering the Smooth distortion localisations.

Liris/Epfl G.P	HD [1, 2]		RMS [1, 2]		3DWPM2 [6]		MSDM2 [9]		FMPD [11]		TPDM [40]		SMQI	
	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s	r_p	r_s
Smooth unif.	52.7	59	10.7	6	90.1	93	85.4	80	51.8	54	71	69	79.1	69
Smooth rough	20.2	15	33.9	36	31	35	67.3	68	60.8	62	68	65.7	53.5	55
Smooth interm.	23	24	51.6	43	24.6	24	45	43	21.3	22	60.2	58.7	40.4	40.0
All smooth localisations	42.8	32	41.8	44	58.4	59	80.9	72.8	69.2	63	77.3	77	75.6	67

significance matrices are presented. The first and second matrices concern the superiority of the FR-MVQA metrics regarding to the 3D meshes of the Liris/Epfl General-Purpose and Liris-Masking databases. The third and fourth matrices are related to the superiority of the FR-MVQA metrics regarding the distortions types (noise and smooth) in the Liris/Epfl General-Purpose database.

Table 5 presents obtained results when a t-test is used to provide Statistical significance of FR-MVQA/MOS residuals. One-sided t-test is used between the quality values of the algorithms across the different 3D meshes. The null hypothesis is that the mean correlation of the two compared algorithms are equal at the 95% confidence level. The alternative hypothesis is that the mean correlation of the row-algorithm is higher (or lesser) than the mean correlation of the column algorithm. Each entry in Table 5 is coded using four or five symbols depending on the consideration of the correlation across the entire database. The position of each symbol corresponds to one subset of the Liris/Epfl General Purpose database. If the symbol equals 1, the FR- MVQA on the row is statistically better than the FR-MVQA on the column (0 means worse, - is used when two FR-MVQA are indistinguishable). The last column named P (Pertinence) refers to the relevance of each quality metric. The pertinence P_i of a FR-MVQA i on the row is defined as: $P_i = \sum_j c_{i,j}$ where $c_{i,j}$ is the minimum of the encoded symbols between the metric i and others metrics j . Hence, more is important he value P of a metric, more this metric is relevant. From the last column of the Table 5 presenting the statistical significance matrix considering the entire Liris/Epfl General Purpose Database in a 3D mesh-based performance, we can notice that the proposed approach (SMQI) performs better than all quality metrics with the exception of 3DWPM2. The Hausdorff metric (HD) seems to have the worst prediction performance in a way that all the state-of-the-art algorithms are better. The same remarks can be made for table 6 which represents the statistical significance matrix for the entire Liris-Masking database. Once again, this result attests the capability of the proposed approach in taking into account the visual masking effect.

Tables 7 and 8 present the statistical significance matrix regarding two types of distortions: noise and smooth. From these tables, it appears that SMQI is more correlated with human judgements than the remaining metrics when we consider the noise distortion. Considering the smooth degradation, SMQI is the second best metric in terms of statistical significance of correlation.

6. APPLICATION TO THE SMOOTHING PROCESSING ON INDEPENDENT 3D MESHES

In different 3D mesh (or 2D image) processing contexts, smoothing represents a crucial pretreatment step (noise removal, thresholding, etc.). Hence, the choice of which smoothing algo-

TABLE 5. Statistical significance matrix of FR-MVQA/MOS on Liris/Ep General Purpose database. Each symbol represents a subset of the database: Armadillo, Venus, Dinosaur, RockerArm and Entire Database. From the last column P we can deduce the following ranking: $HD < TPDM < RMS < FMPD = MSDM2 < SMQI < 3DWPM2$.

Liris Masking	HD [1, 2]	RMS [1, 2]	3DWPM2 [6]	MSDM2 [9]	FMPD [11]	TPDM [40]	SMQI	P
HD [1]	-----	00000	00000	00000	00000	00000	00000	0
RMS [2]	11111	-----	00000	00000	-0000	1-1-1	00000	2
3DWPM2 [6]	11111	11111	-----	11111	11111	11111	1---1	6
MSDM2 [9]	11111	11111	00000	-----	-----	11111	0-000	3
FMPD [11]	11111	-1111	00000	-----	-----	11111	00000	3
TPDM [40]	11111	0-0-0	00000	00000	00000	-----	00000	1
SMQI	11111	11111	0---0	1-111	11111	11111	-----	5

TABLE 6. Statistical significance matrix of FR-MVQA/MOS on Liris-Masking database. Each symbol represents a subset of the database: Armadillo, Lion Vase, Bimba and Dinosaur. From the last column P we can deduce the following ranking: $HD < TPDM = RMS < FMPD = MSDM2 < SMQI < 3DWPM2$.

Liris Masking	HD [1, 2]	RMS [1, 2]	3DWPM2 [6]	MSDM2 [9]	FMPD [11]	TPDM [40]	SMQI	P
HD [1]	----	0000	0000	0000	0000	0000	0000	0
RMS [2]	1111	----	0000	0000	00-0	01--	00000	1
3DWPM2 [6]	1111	1111	----	1-11	1-11	1111	1-11	6
MSDM2 [9]	1111	1111	0-00	----	----	0111	0--0	2
FMPD [11]	1111	11-1	0-00	----	----	01-1	0--0	2
TPDM [40]	1111	00--	0000	0000	00-0	----	0000	1
SMQI	1111	1111	0-00	1--1	1--1	1111	----	5

TABLE 7. Noise distortion-based statistical significance matrix of FR-MVQA/MOS on Liris/Epfl General Purpose database. Each symbol represents a localization (subset) of the noise distortion: Uniformly on the surface mesh, on rough regions, on intermediate regions, on smooth regions and all subsets. From the last column P we can deduce the following ranking: $HD < RMS < TPDM < MSDM2 < 3DWPM2 = FMPD < SMQI$.

Liris Masking	HD [1, 2]	RMS [1, 2]	3DWPM2 [6]	MSDM2 [9]	FMPD [11]	TPDM [40]	SMQI	P
HD [1]	-----	00000	00000	00000	00000	00000	00000	0
RMS [2]	11111	-----	00000	00000	00000	---0	00000	1
3DWPM2 [6]	11111	11111	----	-11-1	011-1	11111	01----	4
MSDM2 [9]	11111	11111	-00-0	-----	0-----	11111	000-0	3
FMPD [11]	11111	11111	100-0	1-----	-----	11111	-00-0	4
TPDM [40]	11111	---1-	00000	00000	00000	----	00000	2
SMQI	11111	11111	10----	111-1	-11-1	11111	----	5

TABLE 8. Smooth distortion-based statistical significance matrix of FR-MVQA/MOS on Liris/Epfl General Purpose database. Each symbol represents a localization (subset) of the smooth distortion: uniformly on the smooth surface, on rough regions, on intermediate regions and all subsets. From the last column P we can deduce the following ranking: $HD < FMPD = TPDM < RMS < MSDM2 < SMQI < 3DWPM2$.

Liris Masking	HD [1, 2]	RMS [1, 2]	3DWPM2 [6]	MSDM2 [9]	FMPD [11]	TPDM [40]	SMQI	P
HD [1]	----	0000	0000	0000	0000	0000	0000	0
RMS [2]	1111	----	0000	0-00	-111	1111	0000	4
3DWPM2 [6]	1111	1111	----	1111	1111	1111	1111	7
MSDM2 [9]	1111	1-11	0000	----	1111	1111	0000	5
FMPD [11]	1111	-000	0000	0000	----	1-0-	0000	1
TPDM [40]	1111	0000	0000	0000	0-1-	----	0000	1
SMQI	1111	1111	0000	1111	1111	1111	----	6

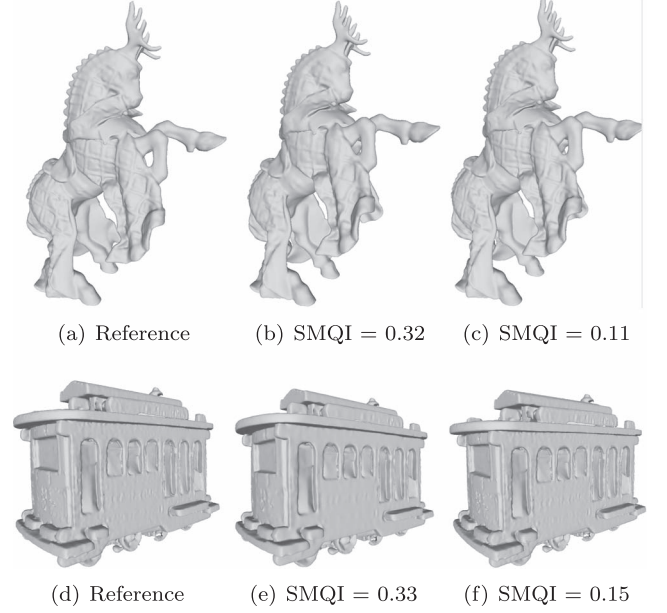


FIGURE 7. Results of the proposed metric (SMQI) on 3D meshes not be-longing to the database used in the learning step: (a) Reference 3D mesh representing a scan of a toy representing a horse (313346 faces), (b) Smoothed version of (a) using the Taubin approach, (c) Smoothed version of (a) using the structure-preserving Laplacian approach, (d) Reference 3D mesh representing a scan of a toy representing the San Francisco Cable Car (535490 faces), (e) Smoothed version of (d) using the Taubin approach and (f) Smoothed version of (d) using the structure-preserving Laplacian approach.

rithm to embed in a processing chain may be difficult since the smoothing results of different algorithms are difficult to assess visually. Quality assessment metrics can respond to a such problem. In order to attest the pertinence and the robustness of the proposed approach, we test it on two independent 3D meshes that were not involved in the genetically-based training. Figure 7 presents two 3D meshes from the GREYC 3D Colored Mesh Database [45] which have been smoothed (for 100 iterations) using two different approaches: Taubin smoothing and structure-preserving Laplacian provided in the open-source software Meshlab [46]. In this experiment we used the same values α , β , γ and δ associated to (15). From the top row of Fig. 7, we can notice that SMQI provides coherent scores of quality in accordance with human perception (we recall that a weak objective quality score refers to a good visual quality while an important quality score is associated to a bad visual quality). The smoothed version based on Taubin algorithm (Fig. 7b) obtains a perceived quality score of 0.32 which is greater to than the score obtained for the smoothed version based on the structure-preserving Laplacian algorithm (Fig. 7c). Indeed, one can observe that the rendering of Taubin algorithm (Fig. 7b) is more distorted visually in comparison to the one obtained with structure-preserving Laplacian algorithm

(Fig. 7c). The same analysis can be made for the second row of the Fig. 7. This experiments shows how the proposed approach can assess the visual quality of independent 3D meshes and how it can be applied in a real application related to the smoothing of 3D meshes.

7. CONCLUSION

We have proposed in this paper a new objective metric, called SMQI, for the assessment of 3D mesh visual quality. This perceptual metric compares structural informations of a reference 3D mesh and a distorted one. For this, we use a multi-scale visual saliency map as a basis for computing local statistics. In order to take into account the masking effect, a roughness map is used to measure the difference of local mean roughness between a pair of meshes. Four comparison functions (i.e., the three functions related to the comparison of the multi-scale saliency, and a roughness comparison function) are combined via a genetically optimized weighted Minkowski sum to provide an objective score of quality that quantifies the visual similarity between two meshes. Experimental results demonstrate the strong correlation of the proposed approach with the subjective results and its high competitiveness. We also conducted a statistical significance and hypothesis testing to compare the superiority of the proposed approach regarding the state-of-the-art methods. Finally a concrete application related to the assessment of smoothing algorithms showed the effectiveness of the proposed approach.

Funding

Agence Nationale de laRecherche (ANR-14-CE27-0001 GRAPH SIP); the European Union FEDER/FSE 2014/2020 (GRAPH SIP project); Saint-Lo Agglo under the project T2PSL.

REFERENCES

- [1] Aspert, N., Santa-Cruz, D. and Ebrahimi, T. (2002) Mesh: Measuring Errors Between Surfaces Using the Hausdorff Distance. *Proc. IEEE Int. Conf. on Multimedia and Expo*, Vol. 1, pp. 705–708.
- [2] Cignoni, P., Rocchini, C. and Scopigno, R. (1998) Metro: measuring error on simplified surfaces. *Comput. Graph. Forum*, 17, 167–174.
- [3] Karni, Z. and Gotsman, C. (2000) Spectral Compression of Mesh Geometry. In *Proc. 27th Annual Conf. on Computer Graphics and Interactive Techniques*, USA SIGGRAPH'00 279–286. ACM Press/Addison-Wesley Publishing Co.
- [4] Sorkine, O., Cohen-Or, D. and Toledo, S. (2003) High-Pass Quantization for Mesh Encoding. In *Proc. 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*. Goslar, DEU SGP'03 42–51. Eurographics Association.
- [5] Gelasca, E.D., Ebrahimi, T., Corsini, M. and Barni, M. (2005) Objective Evaluation of the Perceptual Quality of 3D Watermarking. In *IEEE Int. Conf. on Image Processing 2005*, pp. I–241.
- [6] Corsini, M., Gelasca, E., Ebrahimi, T. and Barni, M. (2007) Watermarked 3D mesh quality assessment. *IEEE Trans. Multimed.*, 9, 247–256.
- [7] Bian, Z., Hu, S. and Martin, R.R. (2009) Evaluation for small visual difference between conforming meshes on strain field. *J. Comput. Sci. Technol.*, 24, 65–75.
- [8] Lavoué, G., Drelie Gelasca, E., Dupont, F., Baskurt, A. and Ebrahimi, T. (2006) Perceptually driven 3D distance metrics with application to watermarking. *Proc. SPIE*, 63120L–63120L–12.
- [9] Lavoué, G. (2011) A multiscale metric for 3D mesh visual quality assessment. *Comput. Graph. Forum*, 30, 1427–1437.
- [10] Váša, L. and Rus, J. (2012) Dihedral angle mesh error: a fast perception correlated distortion measure for fixed connectivity triangle meshes. *Comput. Graph. Forum*, 31, 1715–1724.
- [11] Wang, K., Torkhani, F. and Montanvert, A. (2012) Technical section: A fast roughness-based approach to the assessment of 3D mesh visual quality. *Comput. Graph.*, 36, 808–818.
- [12] Torkhani, F., Wang, K., and Chassery, J. (2012) A curvature tensor distance for mesh visual quality assessment. In Bolc, L., Tadeusiewicz, R., Chmielewski, L. J., and Wojciechowski, K. (eds.), *Computer Vision and Graphics, Berlin, Heidelberg*, pp. 253–263. Springer, Berlin Heidelberg.
- [13] Dong, L., Fang, Y., Lin, W. and Soon, H.S. (2015) Perceptual quality assessment for 3D triangle mesh based on curvature. *IEEE Trans. Multimed.*, 17, 2174–2184.
- [14] Moorthy, A.K. and Bovik, A.C. (2009) Visual importance pooling for image quality assessment. *IEEE J. Sel. Top. Signal Process.*, 3, 193–201.
- [15] Liu, H. and Heynderickx, I. (2011) Visual attention in objective image quality assessment: Based on eye-tracking data. *IEEE Trans. Circuits Syst. Video Technol.*, 21, 971–982.
- [16] Farias, M.C.Q. and Akamine, W.Y.L. (2012) On performance of image quality metrics enhanced with visual attention computational models. *Electron. Lett.*, 48, 631–633.
- [17] Liu, H., Engelke, U., Wang, J. and P., L. C., and Heynderickx, I. (2013) How does image content affect the added value of visual attention in objective image quality assessment? *IEEE Signal Process. Lett.*, 20, 355–358.
- [18] Zhang, L., Shen, Y. and Li, H. (2014) Vsi: A visual saliency-induced index for perceptual image quality assessment. *IEEE Trans. Image Process.*, 23, 4270–4281.
- [19] Feng, X., Wan, W., Yi Da Xu, R., Perry, S., Zhu, S. and Liu, Z. (2018) A new mesh visual quality metric using saliency weighting-based pooling strategy. *Graph. Models*, 99, 1–12.
- [20] Lin, Y., Yu, M., Chen, K., Jiang, G., Chen, F. and Peng, Z. (2020) Blind mesh assessment based on graph spectral entropy and spatial features. *Entropy*, 22, 190.
- [21] Lee, C.H., Varshney, A. and Jacobs, D.W. (2005) Mesh Saliency. In *ACM SIGGRAPH 2005 Papers*. Association for Computing Machinery, New York, NY, USA SIGGRAPH'05 659–666.
- [22] Wu, J., Shen, X., Zhu, W. and Liu, L. (2013) Mesh saliency with global rarity. *Graph. Models*, 75, 255–264.
- [23] Maximo, A., Patro, R., Varshney, A. and Farias, R. (2011) A robust and rotationally invariant local surface descriptor with applications to non-local mesh processing. *Graph. Models*, 73, 231–242.

- [24] Leifman, G., Shtrom, E. and Tal, A. (2012) Surface Regions of Interest for Viewpoint Selection. *2012 IEEE Conf. on Computer Vision and Pattern Recognition*, June, pp. 414–421.
- [25] Johnson, A.E. and Hebert, M. (1999) Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21, 433–449.
- [26] Zhao, Y., Liu, Y., Song, R. and Zhang, M. (2012) A Saliency Detection Based Method for 3D Surface Simplification. *2012 IEEE Int. Conf. on Acoustics*, pp. 889–892. Speech and Signal Processing (ICASSP), March.
- [27] Song, R., Liu, Y., Martin, R.R. and Rosin, P.L. (2012) Conditional Random Field-Based Mesh Saliency. *2012 19th IEEE Int. Conf. on Image Processing*, Sept, pp. 637–640.
- [28] Yitian, Z., Yonghuai, L., and Ziming, Z. (2013) Using Region-Based Saliency for 3D Interest Points Detection. In Wilson, R., Hancock, E., Bors, A., and Smith, W. (eds.), *Computer Analysis of Images and Patterns*, pp. 108–116. Springer, Berlin Heidelberg.
- [29] Elad, M. (2005) Retinex by Two Bilateral Filters. In *Proc. 5th Int. Conf. on Scale Space and PDE Methods in Computer Vision*, Berlin, Heidelberg, Scale-Space’05, 217–229. Springer.
- [30] Song, R., Liu, Y., Martin, R.R. and Rosin, P.L. (2014) Mesh saliency via spectral processing. *ACM Trans. Graph.*, 33, 6:1–6:17.
- [31] Shengfa, W., Nannan, L., Shuai, L., Zhongxuan, L., Zhixun, S. and Hong, Q. (2015) Multi-scale mesh saliency based on low-rank and sparse analysis in shape feature space. *Comput. Aided Geom. Des.*, 35–36, 206–214. Geometric Modeling and Processing 2015.
- [32] Chen, X., Saporov, A., Pang, B. and Funkhouser, T. (2012) Schelling points on 3D surface meshes. *ACM Trans. Graph.*, 31.
- [33] Song, R., Liu, Y. and Rosin, P. (2019) Mesh saliency via weakly supervised classification-for-saliency cnn. *IEEE Trans. Vis. Comput. Graph.*
- [34] Lau, M., Dev, K., Shi, W., Dorsey, J. and Rushmeier, H. (2016) Tactile mesh saliency. *ACM Trans. Graph.*, 35.
- [35] Abouelaziz, I., Chetouani, A., El Hassouni, M., Latecki, L.J. and Cherifi, H. (2019) 3D visual saliency and convolutional neural network for blind mesh quality assessment. *Neural Comput. Appl.*, to appear.
- [36] Abouelaziz, I., Chetouani, A., Hassouni, M.E., Latecki, L.J. and Cherifi, H. (2018) Convolutional neural network for blind mesh visual quality assessment using 3D visual saliency. *25th IEEE Int. Conf. on Image Processing (ICIP)*, 3533–3537.
- [37] Nouri, A., Charrier, C. and Lézoray, O. (2015) Multi-scale mesh saliency with local adaptive patches for viewpoint selection. *Signal Process. Image Commun.*, 38, 151–166.
- [38] Wang, X., Koch, S., Holmqvist, K. and Alexa, M. (2018) Tracking the gaze on objects in 3D: How do people really look at the bunny? *ACM Trans. Graph.*, 37, 188:1–188:18.
- [39] Lavoué, G. (2009) A local roughness measure for 3D meshes and its application to visual masking. *ACM Trans. Appl. Percept.*, 5, 21:1–21:23.
- [40] Torkhani, F., Wang, K. and Chassery, J. (2014) A curvature-tensor-based perceptual quality metric for 3D triangular meshes. *Mach. Graph. Vis.*, 1–25.
- [41] Wu, J., Lin, W., Shi, G. and Liu, A. (2013) Perceptual quality metric with internal generative mechanism. *IEEE Trans. Image Process.*, 22, 43–54.
- [42] Engeldrum, P. (2000) *A Toolkit for Imaging Systems Development*. Imcotek Pr.
- [43] Lavoué, G. and Corsini, M. (2010) A comparison of perceptually-based metrics for objective evaluation of geometry processing. *IEEE Trans. Multimed.*, 12, 636–649.
- [44] Abouelaziz, I., Chetouani, A., El Hassouni, M. and Cherifi, H. (2017) Mesh Visual Quality Assessment Metrics: A Comparison Study. *13th Int. Conf. on Signal-Image Technology Internet-Based Systems (SITIS)*, Dec, 283–288.
- [45] Nouri, A., Charrier, C., and Lezoray, O. (2017) Greyc 3D Colored Mesh Database. Technical report. Normandie Univ, Uni-caen, Ensicaen, CNRS, GREYC.
- [46] Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. and Ranzuglia, G. (2008) Meshlab: an open-source mesh processing tool. 01, 129–136.
- [47] Nouri, A., Charrier, C. and Lézoray, O. (2017) 3D blind mesh quality assessment index. *Proc. IS&T Electronic Imaging, Three-Dimensional Image Processing, Measurement (3DIPM), and Applications*, 9–26.
- [48] Rogowitz, B.E. and Rushmeier, H.E. (2001) Are image quality metrics adequate to evaluate the quality of geometric objects? *Proc. SPIE*, 340–348.
- [49] Yildiz, Z., Oztireli, A., Capin, T. and Cipiloglu, Z. (2018) A machine learning framework for full-reference 3D shape quality assessment. *Vis. Comput.*, 36, 127–139.