

# No-Reference Quality Assessment for 3D Colored Point Cloud and Mesh Models

Zicheng Zhang<sup>1</sup>, Student Member, IEEE, Wei Sun<sup>1</sup>, Xiongkuo Min<sup>1</sup>, Member, IEEE,  
Tao Wang<sup>1</sup>, Wei Lu<sup>1</sup>, and Guangtao Zhai<sup>1</sup>, Senior Member, IEEE

**Abstract**—To improve the viewer's Quality of Experience (QoE) and optimize computer graphics applications, 3D model quality assessment (3D-QA) has become an important task in the multimedia area. Point cloud and mesh are the two most widely used digital representation formats of 3D models, the visual quality of which is quite sensitive to lossy operations like simplification and compression. Therefore, many related studies such as point cloud quality assessment (PCQA) and mesh quality assessment (MQA) have been carried out to measure the visual quality of distorted 3D models. However, most previous studies utilize full-reference (FR) metrics, which indicates they can not predict the quality level in the absence of the reference 3D model. Furthermore, few 3D-QA metrics consider color information, which significantly restricts their effectiveness and scope of application. In this paper, we propose a no-reference (NR) quality assessment metric for colored 3D models represented by both point cloud and mesh. First, we project the 3D models from 3D space into quality-related geometry and color feature domains. Then, the 3D natural scene statistics (3D-NSS) and entropy are utilized to extract quality-aware features. Finally, a support vector regression (SVR) model is employed to regress the quality-aware features into visual quality scores. Our method is validated on the colored point cloud quality assessment database (SJTU-PCQA), the Waterloo point cloud assessment database (WPC), and the colored mesh quality assessment database (CMDM). The experimental results show that the proposed method outperforms most compared NR 3D-QA metrics with competitive computational resources and greatly reduces the performance gap with the state-of-the-art FR 3D-QA metrics. The code of the proposed model is publicly available now at <https://github.com/zzc-1998/NR-3DQA>.

**Index Terms**—3D model quality assessment, colored point cloud, colored mesh, no-reference quality assessment, natural scene statistics.

## I. INTRODUCTION

NOWADAYS, with the rapid development of computer graphics, the digital representation of 3D models has been widely studied and used in a wide range of application

scenarios such as virtual reality (VR), medical 3D reconstruction, and video post-production [1], etc. Among the digital representation forms of 3D models, point cloud and mesh are the most widely used formats in practical. A point cloud based object is a set of points in space, in which each point is described with geometry coordinates and sometimes with other attributes such as color and surface normals. Mesh is more complicated because it is a collection of vertices, edges, and faces which together define the shape of a 3D model. Except for geometry information, the 3D mesh may also contain other appearance attributes, such as color and material. Both point cloud and mesh are able to vividly display exquisite models and complex scenes. However, since point cloud and mesh record the omnidirectional details of objects and scenes, lossless 3D models usually need large storage space and very high transmission bandwidth in practical applications. Hence, a variety of 3D processing algorithms such as simplification and compression, etc. have been proposed to satisfy the specific needs, which inevitably cause damage to the visual quality of 3D models [2], [3]. Additionally, in some 3D scanning model APIs like Apple 3D object capture [4] and Intel Lidar Camera Realsense [5], some slight disturbance such as blur, noise, etc. may be introduced to the constructed 3D models.

Therefore, to improve users' Quality of Experience (QoE) in 3D fields and optimize 3D compression and reconstruction systems, it is of great significance to develop quality metrics for point cloud quality assessment (PCQA) and mesh quality assessment (MQA). However, different from the format of 2D media like images and videos where pixels are distributed in the fixed grid, the points in 3D models are distributed irregularly in the space, which brings a huge challenge for 3D-QA tasks. For example, the neighborhood of pixels in 2D media can be easily obtained, thus the local features are available by analyzing the relationship between the pixel and its neighborhood. However, the neighborhood for points in 3D models is very ambiguous, which makes it difficult to conduct local feature analysis. Although some 3D-QA databases [6]–[10] have been proposed to push forward the development of 3D-QA algorithms, the difficulty of collecting 3D models and conducting subjective experiments greatly limit the size of the 3D-QA databases, which may also restrict relevant research, especially for NR 3D-QA metrics.

## A. Previous 3D-QA Works

Quality assessment can be divided into subjective quality assessment and objective quality assessment according to whether human observers are involved. It is known that,

Manuscript received 21 February 2022; revised 2 May 2022; accepted 22 June 2022. Date of publication 28 June 2022; date of current version 28 October 2022. This work was supported in part by the National Science Foundation of China under Grant 61831015 and in part by the National Key Research and Development Program of China under Grant 2021YFE0206700. This article was recommended by Associate Editor L. Agostini. (Corresponding author: Guangtao Zhai.)

The authors are with the Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zzc1998@sjtu.edu.cn; sunguwei@sjtu.edu.cn; minxiongkuo@sjtu.edu.cn; fl603011.wangtao@sjtu.edu.cn; SJTU-Luwei@sjtu.edu.cn; zhaiguangtao@sjtu.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCSVT.2022.3186894>.

Digital Object Identifier 10.1109/TCSVT.2022.3186894

1051-8215 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

in subjective quality assessment, large numbers of people are required to assess the visual quality of 3D models and give their subjective scores. Although the mean opinion scores (MOSs) collected from human observers are straightforward and precise, they cannot be used in practical applications due to the huge time consumption and high expense [11]. Therefore, objective quality metrics are urgently needed to predict the visual quality of 3D models automatically. The 3D objective quality assessment can then be divided into full-reference (FR), reduced-reference (RR), and no-reference (NR) 3D-QA methods. FR 3D-QA methods work by comparing the difference between the reference 3D models and distorted 3D models, RR 3D-QA methods employ part of the reference models' information for comparison, and NR 3D-QA methods only analyze the distorted 3D models to give the perceptual quality scores. Considering the complexity of 3D models, in the literature, a large part of 3D-QA metrics are full-reference and they only take geometry features into consideration [12]–[19]. When it comes to 3D models with color information, limited works have been proposed [20]–[25]. Clearly, the NR 3D-QA for colored models has fallen behind. In this section, we briefly review the development of 3D-QA and introduce the mainstream methods designed for 3D-QA tasks.

1) *The Development of PCQA*: The earliest FR-PCQA metrics usually focus on the geometry aspect at the point level, such as p2point [12], p2plane [13], and p2mesh [16]. The p2point estimates the levels of distortion by computing the distance vector between the corresponding points. The p2plane further projects the distance vector on the normal orientation for evaluating the quality loss. The p2mesh first reconstructs the point cloud to mesh and then measures the distance from points to the reconstructed surface to predict the quality level, however, it greatly depends on the reconstruction algorithms and lacks stability. Since the point-level difference is difficult to reflect the complex structural distortions, some studies further consider other structural characteristics for PCQA. For example, Alexiou and Ebrahimi [17] adopt the angular difference of point normals to estimate the degradations. Javaheri *et al.* [18] utilize the generalized Hausdorff distance to reflect the distortions caused by compression operations.

In some situations, color information can not be ignored, which challenges the PCQA methods considering only geometry information. To incorporate the color information into PCQA models, Meynet *et al.* [23] propose a metric to use the weighted linear combination of curvature and color information to evaluate the visual quality of distorted point clouds. Inspired by SSIM [26] and Alexiou and Ebrahimi [20] compute the similarity of four types of features, including geometry, normal vectors, curvature, and color information. What's more, some studies [6], [9] try to predict the quality level by evaluating the projected 2D images from 3D models. The advantage is that image quality assessment (IQA) metrics have been well developed while the disadvantage is that there is inevitable information loss during the projection and the projected images are easily affected by the projected angles and viewpoints.

Few NR-PCQA metrics have been developed so far, which include a learning-based approach [27] using two modified PointNet [28] as feature extraction backbone and also a learning-based method [29] using multi-view projection.

2) *The Development of MQA*: The MQA metrics can be categorized into two types: model-based metrics [14], [15], [19], [30], [31] which operate directly on the 3D models, and IQA-based metrics [21], [22], [32]–[34] which operate on the rendering snapshots of 3D models. Supported by the vast amount of previous research on FR-IQA methods [35]–[37], many FR-MQA models have been proposed using similar manners, which usually compute local features at the vertex level and then pool the features into a quality value. For example, MSDM2 [14] uses the differences of structure (captured via curvature statistics) computed on local neighborhoods to predict the quality level. DAME [19] measures the differences in dihedral angles between the reference and the distorted meshes to evaluate the quality loss. FMPD [15] estimates the local roughness difference derived from Gaussian curvature to assess the quality of the distorted mesh. However, these metrics only take geometry information into consideration.

Furthermore, to analyze the influence of color information, some color-involved FR-MQA metrics have been carried out. Tian and AlRegib [21] introduce a global distance over texture image using Mean Squared Error (MSE) to quantify the effect of color information. Guo *et al.* [22] exploit SSIM to calculate the texture image distance as the color information features. Nehmé *et al.* [7] introduce a metric to incorporate perceptually relevant curvature-based and color-based features to evaluate the visual quality of colored meshes.

Recently, thanks to the effectiveness of machine learning technologies, some learning-based NR-MQA metrics have been proposed. Abouelaziz *et al.* [30] extract features using dihedral angle models and train a support vector machine for feature regression. Later, Abouelaziz *et al.* [31] scale the curvature and dihedral angle into 2D patches and utilize the convolution neural network (CNN) for training. They further introduce a CNN framework with saliency views rendered from 3D meshes [32]. However, the NR methods mentioned above only extract geometry features and may fail to accurately predict the scores of colored meshes.

## B. Our Approach

Generally speaking, a large part of the metrics mentioned above are full-reference metrics, which can make full use of the relationship between the reference and distorted 3D models and give relatively accurate results. However, the disadvantage of FR 3D-QA metrics is also obvious. They are not able to work in the absence of reference 3D models. Unfortunately, in many application scenarios like 3D reconstruction, pristine reference is not always available. Therefore, inspired by the huge success of natural scene statistics (NSS) in IQA tasks [9], [38], [39], some researchers [30], [40], [41] use similar ideas of extracting handcrafted features and estimating the statistical parameters with certain NSS distributions as quality-aware information. In this paper, we further push forward the application of NSS for both colored point cloud and mesh models, and we summarize such a concept as 3D-NSS. Specifically, NSS is a discipline within the field of perception, which is dependent on the premise that the perceptual system is designed to interpret natural scenes [39]. Through observations of feature distributions, we find that the reference features obey certain NSS distributions and different types of distortions change the appearance of such feature distributions. Therefore, we believe

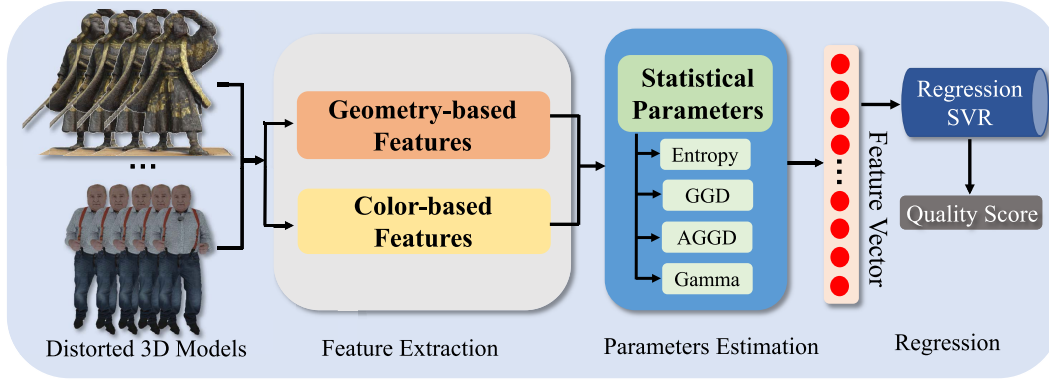


Fig. 1. The framework of the proposed method. Geometry-based and color-based features are first extracted from the distorted 3D models. Then various statistical parameters are estimated from the extracted features to form the feature vector. Finally, the quality scores are given through the SVR regression module.

that 3D-NSS is effective to quantify the visual quality of 3D models in the presence of distortion. More details of 3D-NSS are discussed in Section III.

Specifically, we first project the 3D models into quality-related geometry and color feature domains. Then, 3D-NSS and entropy are utilized to extract these color and geometry characteristics. Finally, the obtained features are integrated into a quality value through the support vector regression (SVR) model. In order to test the effectiveness of different types of features and different kinds of distribution models, we test the performance of various combinations of the statistical parameters to find the optimal combination. Further, we conduct the data-sensitivity experiment, the ablation study, and the computational efficiency experiment to demonstrate the effectiveness of our method. In-depth discussions are given as well.

### C. Contributions of This Paper

We summarize our contributions as follows:

- We push forward the development of NSS in the 3D-QA fields based on previous research [30], [40], [41] and we systematically summarize the concept as 3D-NSS.
- We propose a no-reference quality assessment metric for both colored point cloud and mesh. We extract features not only from the geometry aspect, but also from the color aspect. Furthermore, it is the first method that can deal with both NR-PCQA and NR-MQA with color information. The proposed NR 3D-QA model follows a common NR framework, which means it is easy to modify our metric for performance improvement or meeting other specific needs.
- We deeply investigate the effectiveness of different types of features and different kinds of NSS models, which can provide useful guidelines for future research.
- Compared with the state-of-the-art methods, our method is more efficient in the computation process, which means the proposed method is potentially more capable of handling practical situations. The code of the proposed model is also released for promoting the development of NR 3D-QA.

The paper is organized as follows. Section II describes the feature projection processes. Section III describes the process of quantifying the distortions to statistical parameters.

Section IV presents the experiment setup and the experimental results. Section V summarizes of this paper.

## II. FEATURE PROJECTION

The framework of the proposed method is illustrated in Fig. 1, which includes a feature extraction module, a parameters estimation module, and a regression module. Point cloud and mesh have similar quality-aware geometric properties as well as the attached color attributes. Therefore, we uniformly design the point cloud and mesh processing framework, and determine the specific projection domains based on their characteristics. Before introducing the proposed model, we first define a distorted 3D object  $O$  as:

$$\begin{aligned} O &\in \{P, M\}, \\ P &= \{Points\}, \\ M &= \{Vertices, Edges, Faces\}, \end{aligned} \quad (1)$$

where  $P$  and  $M$  mean that the 3D object is represented by point cloud and mesh respectively, and the color information is attached to *Points* in point cloud and *Vertices* in mesh respectively.

### A. Geometry Feature Projection

Geometry features usually have a strong correlation with human perception, which has been firmly proved in 3D-QA studies [12]–[19]. Although the geometry features are computed in different ways for point cloud and mesh, they share similar characteristics for the visual quality of 3D models. In this section, we project the given 3D model into several quality-aware geometry feature domains:

$$\begin{aligned} F_{geo} &= \mathbf{Projection}_{geo}(O), \\ O &\in \{P, M\}, \end{aligned} \quad (2)$$

where  $F_{geo}$  indicates the set of geometry feature domains of the 3D model,  $\mathbf{Projection}_{geo}(\cdot)$  denotes the geometry projection function.

1) *Point Cloud Geometry Feature Domains*: Considering that the point cloud lacks surfaces, we first need to get the neighborhood set for each point so we can further extract the geometry features. Given the point cloud  $P = \{p_i\}_{i=1}^N$ ,



the neighborhood  $P_{Nbi}$  of each vertex  $p_i$  can be obtained utilizing the k-nearest neighbors (k-NN) algorithm:

$$P_{Nb} = \underset{k=10}{\text{KNN}}(P),$$

$$\text{Dist}(p, q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_z - q_z)^2}, \quad (3)$$

where  $N$  indicates the number of points in the point cloud,  $P_{Nb}$  is the set of all neighborhoods,  $\text{KNN}(\cdot)$  denotes the k-NN algorithm function, the  $k$  level is set as 10, and Euclidean distance  $\text{Dist}(p, q)$  is exploited as the distance function. With the computed neighborhood  $P_{Nbi}$ , the corresponding covariance matrix  $C_i$  for each point  $p_i$  (represented by 3D geometry coordinates) can be denoted as:

$$C_i = \frac{1}{K} \sum_{j=1}^K (p_{n_j} - \hat{p})(p_{n_j} - \hat{p})^\top,$$

$$\times \{p_{n_1}, \dots, p_{n_K}\} \in P_{Nbi}, \quad (4)$$

where  $p_{n_j}$  and  $\hat{p}$  are vectors of dimension  $3 \times 1$ ,  $C_i$  is a matrix of dimension  $3 \times 3$ ,  $K$  stands for the size of neighborhood  $P_{Nbi}$ , and  $\hat{p}$  represents the centroid of neighborhood  $P_{Nbi}$ . Therefore, the eigenvector problem for the covariance matrix  $C_i$  can be described as:

$$C_i \cdot v_l = \lambda_l \cdot v_l, l \in \{1, 2, 3\}, \quad (5)$$

where  $(\lambda_1, \lambda_2, \lambda_3)$  indicate the eigenvalues,  $(v_1, v_2, v_3)$  represent the corresponding eigenvectors, and we assume that  $\lambda_1 > \lambda_2 > \lambda_3$ . Thus, a total of three eigenvalues are obtained for each point  $p_i$  in the point cloud  $P$ .

In previous studies [42], [43], the eigenvalues mentioned above are used to calculate various geometry features for tasks like classification, simplification, and segmentation. These features have shown great ability to describe the complicated geometry structure information and achieved outstanding performance. Hence, we selected several geometry features that may be helpful for point cloud quality assessment:

- **Curvature:** Curvature is the amount by which a curve deviates from being a straight line, which is frequently used to describe roughness or smoothness.
- **Anisotropy:** Anisotropy is used to exhibit variations in geometrical properties for different directions.
- **Linearity:** Linearity is the property for estimating the similarity to a straight line.
- **Planarity:** Planarity is utilized to measure the similarity to a planar surface.
- **Sphericity:** Sphericity is the measure of how closely the shape of an object resembles that of a perfect sphere.

Moreover, all the geometry features described above represent the local distribution patterns and are calculated at the point level, which indicates that each  $p_i$  has its own curvature, anisotropy, linearity, planarity, and sphericity. The formulations [44] for geometry features are denoted as:

$$\text{Cur}(p_i) = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}, \quad (6)$$

$$\text{Ani}(p_i) = \frac{\lambda_1 - \lambda_3}{\lambda_1}, \quad (7)$$

$$\text{Lin}(p_i) = \frac{\lambda_1 - \lambda_2}{\lambda_1}, \quad (8)$$

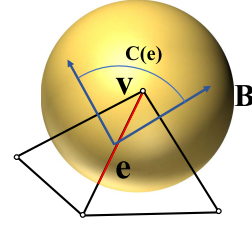


Fig. 2. Example of the variables corresponding to the curvature calculation, where the red line indicates  $|e \cap B|$  and  $C(e)$  indicates the curvature contribution of  $e$ .

$$\text{Pla}(p_i) = \frac{\lambda_2 - \lambda_3}{\lambda_1}, \quad (9)$$

$$\text{Sph}(p_i) = \frac{\lambda_3}{\lambda_1}, \quad (10)$$

where  $\lambda_1, \lambda_2$ , and  $\lambda_3$  refer to the corresponding eigenvalues respectively. Through the operations mentioned above, the point cloud is transformed into 5 geometry feature domains.

2) **Mesh Geometry Feature Domains:** Compared with point cloud, the 3D mesh is more complicated because it is a collection of vertices, edges, and faces which together define the shape of a 3D model. Therefore, the feature extraction process is a bit different from that in the point cloud.

a) **Curvature:** Thanks to the edges and faces information contained in mesh, the computation of curvature is more precise than the point cloud. So far, various curvature definitions have been introduced for 3D meshes [45], [46], among which the average curvature is proven to be able to stably describe the local structural information of 3D meshes. Hence, a weighted average curvature method introduced in [46] is adopted to measure the roughness of embedded surfaces in 3D meshes. Given the mesh  $M = \{V, E, F\}$ , the weighted average curvature is calculated by averaging the curvature contributions over a certain region  $B$  around each vertex  $v_i$ :

$$\text{Cur}(v_i) = \frac{1}{|B|} \sum_{e \in E} C(e) |e \cap B|, \quad (11)$$

where  $\text{Cur}(v_i)$  represents the weighted average curvature for vertex  $v$ , region  $B$  is a sphere region centered at  $v$  and its radius is  $1/100$  of the bounding box of the 3D mesh,  $|B|$  stands for the surface area of region  $B$ .  $E$  is the set of edges that are linked to the vertex  $v$ ,  $C(e)$  indicates the curvature contribution of  $e$  (the signed angle between the normals of the two oriented triangles incident to edge  $e$ ),  $|e \cap B|$  denotes the weight which is defined as the length of  $e \cap B$  (always between 0 and  $|e|$ ). An example of the variables corresponding to the curvature calculation is shown in Fig. 2.

b) **Dihedral angle:** For a 3D mesh, the dihedral angle is the angle between the normals of two adjacent faces, which has been utilized in [47] as an effective indicator for measuring the loss of the mesh simplification. Furthermore, [19] exploits oriented dihedral angles instead of the simple dot product of normals to better distinguish the convex and concave angles. Inspired by similar ideas, we adopt the dihedral angles set to describe the visual quality of 3D meshes. Assuming that the vertices for two adjacent faces  $f_1$  and  $f_2$  are  $\{v_1, v_3, v_4\}$  and  $\{v_2, v_3, v_4\}$ , we have:

$$\text{Dih}_{f_1, f_2} = \arccos(n_1 \cdot n_2) \cdot \text{sgn}(n_1 \cdot (v_2 - v_1)), \quad (12)$$

where  $Dih_{f_1, f_2}$  is the oriented dihedral angle between two adjacent faces  $f_1$  and  $f_2$ ,  $n_1$  and  $n_2$  are the normals of  $f_1$  and  $f_2$ ,  $\text{sgn}(\cdot)$  denotes the signum function which is used to decide the orientation of the dihedral angle.

c) *Face area and angle*: Area and angle are two simple attributes of 3D mesh faces, which can be easily computed by making use of the coordinates of the face's vertices. In the mesh smoothing algorithm proposed in [48], attributes including face angle are used to predict the new location of the smoothed nodes. While in the 3D mesh encoding method introduced in [49], face angle is utilized to instruct the compression of 3D meshes, which indicates that face angle is related to the quality of 3D meshes. Therefore, to further measure the visual quality degradation of 3D meshes, the face area and angle are collected as feature sets.

Finally, the mesh is projected into 4 geometry feature domains:

- **Curvature**: The weighted average curvature is used to describe the geometry characteristics like roughness or smoothness.
- **Dihedral Angle**: Dihedral angle is employed as a useful descriptor for measuring the caused degradations.
- **Face Area & Angle**: These two attributes are highly correlated with the effectiveness of lossy operations like compression.

### B. Color Feature Projection

Color is a significant aspect of visual quality assessment. For a colored point cloud, the color is directly determined by the color information of the point, while for a colored mesh, the color of the surface is generally rendered by the color information of the contained vertices. Besides, the color information in the 3D models is usually stored in the form of RGB channels. However, the RGB color space has been proven to have a poor correlation with human perception. Therefore, we adopt the LAB color transformation as the color feature projection:

$$\begin{aligned} F_{col} &= \mathbf{Projection}_{col}(O), \\ O &\in \{P, M\}, \end{aligned} \quad (13)$$

where  $F_{col}$  indicates the set of color feature domains of the 3D model,  $\mathbf{Projection}_{col}(\cdot)$  stands for the color projection function, and  $P$  and  $M$  represent the distorted point cloud and mesh respectively. The detailed color transformation is formulated as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.7688 & 1.7517 & 1.1301 \\ 1.0000 & 4.5906 & 0.0601 \\ 0 & 0.0565 & 5.5942 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (14)$$

$$\begin{cases} L = 116f\left(\frac{Y}{Y_n}\right) - 16, \\ A = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right), \\ B = 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right), \end{cases} \quad (15)$$

where  $R, G, B$  represent the corresponding RGB color channels,  $X, Y, Z$  stand for the corresponding XYZ color channels,  $L, A, B$  denote the corresponding RGB color channels, and  $X_n, Y_n, Z_n$  describe the specified white achromatic reference

illuminant. The  $f(\cdot)$  function is described as:

$$f(t) = \begin{cases} \sqrt[3]{t}, & \text{if } t > \delta^3, \\ \frac{t}{3\delta^2} + \frac{4}{29}, & \text{otherwise,} \end{cases} \quad (16)$$

where  $\delta$  is set as  $\frac{6}{29}$ . Finally, the LAB color channels are computed as the color feature domains.

## III. ESTIMATING STATISTICAL PARAMETERS

Through prior knowledge and observations of the corresponding feature distributions, we find that the characteristic statistical parameters of NSS models can be changed by the presence of distortion. Therefore, we choose entropy and several NSS models including the generalized Gaussian distribution (GGD), the general asymmetric generalized Gaussian distribution (AGGD), and the shape-rate Gamma distribution for parameters estimation to quantify the perceptual quality of 3D models.

### A. Basic Statistical Parameters

For each set of features, we exploit the normalization operation as the pre-processing:

$$\begin{aligned} \hat{F} &= \frac{F - \text{mean}(F)}{\text{std}(F) + C}, \\ F &\in \{F_{geo}, F_{col}\}, \end{aligned} \quad (17)$$

where  $F$  represents the feature domain,  $\text{mean}(\cdot)$  is the average function,  $\text{std}(\cdot)$  denotes the standard deviation function, and  $C$  is a small constant to avoid instability. Entropy is believed to be highly correlated with the quantization distortion. Fig. 3 presents the normalized probability distributions of LAB channels for color quantization distortion. It can be observed that with the increasing quantization levels, the corresponding distributions become significantly sparse, meaning the number of distinct colors is reduced. Considering that some simplification and compression algorithms usually introduce quantization operations to the 3D models, we decide to use entropy as one of the quality-aware features. Then the entropy can be derived as:

$$\begin{aligned} E &= \mathbf{Entropy}(\hat{F}), \\ \hat{F} &\in \{F_{geo}, F_{col}\}. \end{aligned} \quad (18)$$

where  $\mathbf{Entropy}(\cdot)$  indicates the entropy function,  $F_{geo}$  and  $F_{col}$  are the normalized feature distributions for geometry and color respectively.

### B. GGD Parameters

Fig. 4 shows one reference model and three distorted point cloud samples with different distortion types (compression, downsampling, geometry Gaussian noise). It can be observed that the reference curvature and anisotropy distributions shown in Fig. 4 (e, i) exhibit Gamma-like appearance while the reference linearity and planarity distributions shown in Fig. 4 (m, q) tend to be Gaussian-like. Obviously, with different distortion types, the shapes of the distributions are changed. For example, with closer inspections, we can see that compression adds more weight to the tail of curvature and anisotropy distributions shown in Fig. 4 (f, j) while downsampling makes

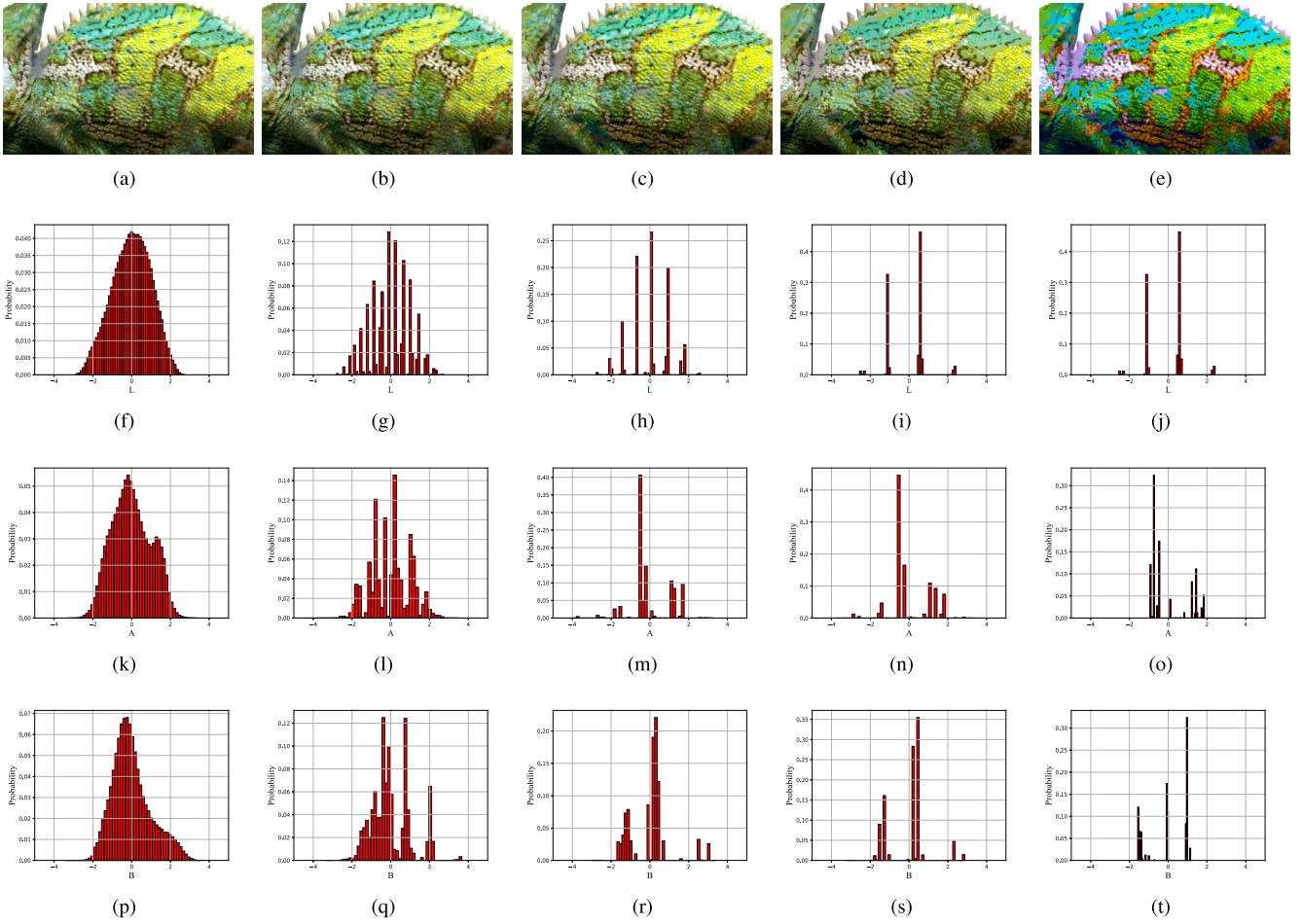


Fig. 3. A comparison example for color distortion in the CMDM database [7]. (a) represents the snapshots of the reference 3D mesh while (b)-(e) stand for the snapshots of the meshes with 4 increasing levels of color information quantization. (f), (k), (p) represent the normalized probability distributions of LAB channels for (a) model, (g), (l), (q) represent the normalized probability distributions of LAB channels for (b) model, (h), (m), (r) represent the normalized probability distributions of LAB channels for (c) model, (i), (n), (s) represent the normalized probability distributions of LAB channels for (d) model, and (j), (o), (t) represented the normalized probability distributions of LAB channels for (e) model respectively.

curvature and anisotropy distributions shown in Fig. 4 (g, k) more centered. What's more, geometry Gaussian noise makes all the geometry distributions more smooth and adds more weight to the tail of curvature and anisotropy distributions shown in Fig. 4 (h, l) on both sides.

The GGD distribution is effective for capturing a broad spectrum of statistics and reflecting the changes in distribution shapes, especially for the tail part [50]. To capture the corresponding characteristics of 3D model distributions, we propose to use the generalized Gaussian distribution (GGD):

$$GGD(x; \alpha, \beta^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right), \quad x \in \{F_{geo}, F_{col}\}, \quad (19)$$

where  $\beta = \sigma\sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}$ ,  $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$ ,  $\alpha > 0$  is the gamma function, and the two estimated parameters  $(\alpha, \beta^2)$  indicate the shape and variance of the distribution. What's more, considering that the variance for normalized distribution is fixed as 1, we estimate the GGD parameters from distributions before normalization.

### C. AGGD Parameters

However, in some situations, the shape of the distribution is not symmetrical. We can see clearly from Fig. 4 that the reference planarity distribution represented by Fig. 4 (q) has a longer tail on the left while the planarity distribution represented by Fig. 4 (t) has a longer tail on the right. What's more, all types of distortions change weight of the tails for curvature and anisotropy distributions represented by Fig. 4 (f, g, h, j, k, l). Therefore, we utilize the general asymmetric generalized Gaussian distribution (AGGD) model to extract more detailed parameters from normalized feature distributions to describe the different spread extent in two directions:

$$AGGD(x; v, \sigma_l^2, \sigma_r^2) = \begin{cases} \frac{v}{(\beta_l + \beta_r)\Gamma(\frac{1}{v})} \exp\left(-\left(\frac{-x}{\beta_l}\right)^v\right) & x < 0, \\ \frac{v}{(\beta_l + \beta_r)\Gamma(\frac{1}{v})} \exp\left(-\left(\frac{x}{\beta_r}\right)^v\right) & x \geq 0, \end{cases} \quad x \in \{F_{geo}, F_{col}\}, \quad (20)$$

where  $\beta_l = \sigma_l\sqrt{\frac{\Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})}}$ ,  $\beta_r = \sigma_r\sqrt{\frac{\Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})}}$ ,  $\eta = (\beta_r - \beta_l)$ , the parameter  $v$  decides the shape of the distribution,  $\sigma_l^2$  and  $\sigma_r^2$

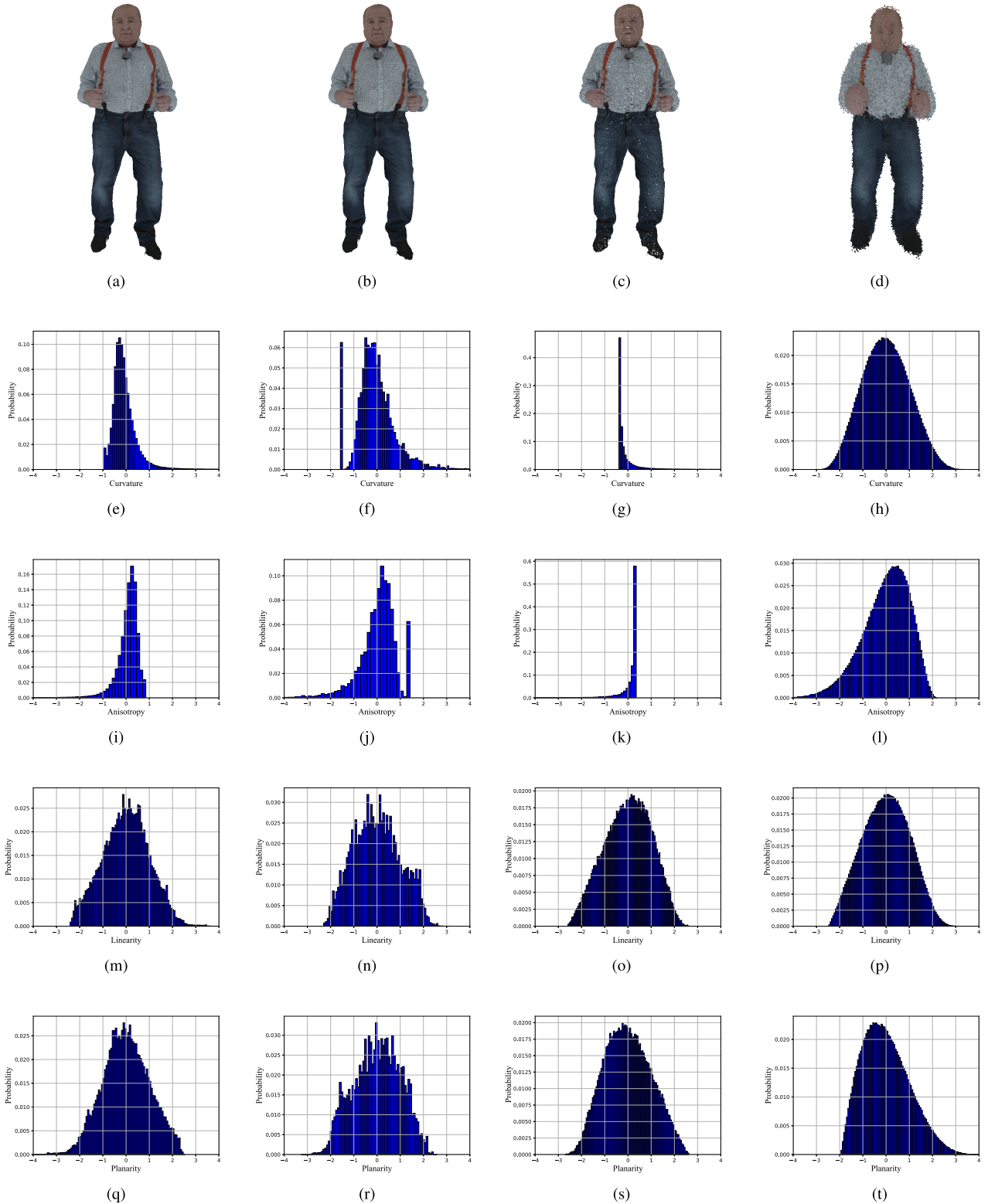


Fig. 4. Examples of point cloud samples from SJTU-PCQA database [6]. (a) is the reference point cloud sample, while (b), (c), and (d) are 3 distorted point cloud samples with different distortion types (compression, downsampling, geometry Gaussian noise). Some features' (*Cur*, *Ani*, *Lin*, *Pla*) normalized probability distributions are selected as examples. More specifically, (e), (i), (m), and (q) are the corresponding features' normalized probability distributions of (a) model, (f), (j), (n), and (r) are the corresponding features' normalized probability distributions of (b) model, (g), (k), (o), and (s) are the corresponding features' normalized probability distributions of (c) model, (h), (l), (p), and (t) are the corresponding features' normalized probability distributions of (d) model respectively.

are scale parameters that refer to the spread extent on the left and right sides of the distribution respectively,  $\eta$  is a parameter that can better fit the AGGD model stated in [38]. Additionally, the AGGD can be recognized as the extension



TABLE I  
SUMMARY OF FEATURES EXTRACTED IN THE PROPOSED METHOD

Format	Feature Domains	Feature ID	Feature Description	Computation
Point cloud	$(Cur, Ani, Lin, Pla, Sph, L, A, B)$	$f_{p1} - f_{p16}$	Mean and standard deviation for each feature domain	Eq.(17)
	$(Cur, Ani, Lin, Pla, Sph, L, A, B)$	$f_{p17} - f_{p24}$	Entropy for each feature domain	Eq.(18)
	$(Cur, Ani, Lin, Pla, Sph, L, A, B)$	$f_{p25} - f_{p40}$	GGD $(\alpha, \beta^2)$ for each feature domain	Eq.(19)
	$(Cur, Ani, Lin, Pla, Sph, L, A, B)$	$f_{p41} - f_{p72}$	AGGD $(\eta, v, \sigma_l^2, \sigma_l^2)$ for each normalized feature domain	Eq.(20)
	$(Cur, Ani, Lin, Pla, Sph, L, A, B)$	$f_{p73} - f_{p88}$	Gamma $(\alpha, \beta)$ for each normalized feature domain	Eq.(21)
Mesh	$(Cur, Dih, Far, Fan, L, A, B)$	$f_{m1} - f_{m14}$	Mean and standard deviation for each feature domain	Eq.(17)
	$(Cur, Dih, Far, Fan, L, A, B)$	$f_{m15} - f_{m21}$	Entropy for each feature domain	Eq.(18)
	$(Cur, Dih, Far, Fan, L, A, B)$	$f_{m22} - f_{m35}$	GGD $(\alpha, \beta^2)$ for each feature domain	Eq.(19)
	$(Cur, Dih, Far, Fan, L, A, B)$	$f_{m36} - f_{m63}$	AGGD $(\eta, v, \sigma_l^2, \sigma_l^2)$ for each normalized feature domain	Eq.(20)
	$(Cur, Dih, Far, Fan, L, A, B)$	$f_{m64} - f_{m77}$	Gamma $(\alpha, \beta)$ for each normalized feature domain	Eq.(21)

of GGD. When  $\sigma_l^2 = \sigma_r^2$ , the AGGD turns into GGD. Finally, the four parameters  $(\eta, v, \sigma_l^2, \sigma_l^2)$  are estimated to describe the characteristics of asymmetric distributions.

#### D. Gamma Parameters

The reference curvature and anisotropy distributions shown in Fig. 4 (e, i) are similar to the shape of Gamma distribution. Compression, downsampling, and Gaussian noise all change the shape and scale of the corresponding distributions and such distributions still exhibit Gamma-like appearance. Therefore, we propose to use Gamma distribution parameters to quantify the distortions. The shape-rate Gamma model is formulated as:

$$Gamma(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} x > 0, \\ x \in \{\hat{F}_{geo}, \hat{F}_{col}\}, \quad (21)$$

where  $\alpha$  and  $\beta$  stands for the shape and rate parameters and  $\alpha, \beta > 0$ .

#### E. Parameters Summary

In summary, we collect the average, standard deviation, and entropy values in the 3D model feature domains as the fundamental features. Then we use 3 representative distribution models to estimate 8 statistical parameters, including GGD  $(\alpha, \beta^2)$ , AGGD  $(\eta, v, \sigma_l^2, \sigma_l^2)$ , Gamma  $(\alpha, \beta)$ , for all feature domains. Finally, considering that the colored point cloud has 8 feature domains  $(Cur, Ani, Lin, Pla, Sph, L, A, B)$  and the colored mesh has 7 feature domains  $(Cur, Dih, Far, Fan, L, A, B)$ , 88 ( $8 \times 11$ ) features are computed for a single colored point cloud and 77 ( $7 \times 11$ ) features are computed for a single colored mesh respectively. The summary of features extracted in the proposed method is listed in Table I.

### IV. EXPERIMENT EVALUATION

#### A. Regression Model

After the feature extraction process, a feature vector is obtained to describe the characteristics of the 3D model.

In our experiment, we propose to use the support vector machine regressor (SVR) as the regression model, which is a common and effective choice to handle high dimensional data in previous quality assessment research [30], [38]. We employ the min-max normalization as the pre-processing to scale the features. Then the feature vector can be integrated into a quality score for evaluation. We employ the Python sklearn package [51] to implement the radial basis function (RBF) kernel SVR model with default settings.

#### B. Experiment Setup

1) *Experiment Setup for PCQA*: To test the performance of the proposed method, we employ the subjective point cloud assessment database (SJTU-PCQA) [6] and the Waterloo point cloud assessment database (WPC) proposed in [10].

The SJTU-PCQA database provides 420 point cloud samples distorted from 10 reference point clouds. Each reference point cloud is distorted with seven types of common distortions with six levels. Unfortunately, only 9 reference point clouds and their corresponding distorted point cloud samples are now available to the public, thus we can obtain 378 ( $9 \times 6 \times 7$ ) point cloud samples for the experiment. Since the proposed approach requires a training procedure to calibrate the SVR model and to avoid the influence of content overlap, we select 8 of the 9 groups of point clouds as training set and leave the rest 1 group as testing set. In order to ensure the validity of the results, we exhaustively list all the  $C_9^8 = 9$  database separations for the experiment and use the average performance as the final experimental results. In addition, the MOSs collected in the SJTU-PCQA database are divided by 10 in the training process to scale the MOSs to [0,1]. The predicted scores are re-scaled for validation in the testing process.

The WPC database includes 20 high-quality source point clouds and creates 740 distorted point clouds using downsampling, Gaussian noise, and three types of compression. Specifically, we maintain the same training set and testing set split as stated in [29]. Similarly, the MOSs provided in the WPC database are divided by 100 in the training process and the predicted scores are re-scaled for validation in the testing process. The comparison PCQA metrics can be categorized into two types:



TABLE II  
PERFORMANCE COMPARISON WITH COMPETITORS ON THE SJTU-PCQA AND WPC DATABASES

Ref	Type	Index	Metric	SJTU-PCQA				WPC			
				PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE
FR	Image-based	A	PSNR	0.2317	0.2422	0.1077	2.3124	0.4872	0.4235	0.3080	15.8133
		B	SSIM	0.3476	0.2987	0.1919	2.1770	0.4944	0.3878	0.3234	15.7749
		C	PB-PCQA	0.6076	0.6020	-	1.8635	-	-	-	-
	Model-based	D	GraphSIM	0.8449	0.8483	0.6448	1.5721	0.6163	0.5831	0.4194	17.1939
		E	PointSSIM	0.7136	0.6867	0.4964	1.7001	0.4667	0.4542	0.3278	20.2733
		F	PCQM	<b>0.8653</b>	<b>0.8544</b>	<b>0.6586</b>	<b>1.2162</b>	<b>0.7499</b>	<b>0.7434</b>	<b>0.5601</b>	<b>15.1639</b>
RR	Model-based	G	PCMRR	0.6101	0.4816	0.3362	1.9342	0.3433	0.3097	0.2082	21.5302
NR	Image-based	H	NIQE	0.3764	0.2214	0.1512	2.2671	0.3957	0.3887	0.2551	22.5502
		I	BRISQUE	0.2241	0.2051	0.1121	2.2428	0.4176	0.3781	0.2444	22.5414
		J	PQA-Net	-	-	-	-	<b>0.7000</b>	<b>0.6900</b>	<b>0.5100</b>	<b>15.1800</b>
	Model-based	K	ResCNN	0.5975	0.6187	-	-	-	-	-	-
		L	Proposed	<b>0.7382</b>	<b>0.7144</b>	<b>0.5174</b>	<b>1.7686</b>	0.6514	0.6479	0.4417	16.5716

- Image-based metrics: These metrics evaluate the quality of 3D models by assessing the quality of the corresponding 2D projections. Please refer to [6] for detailed projection process and we use only RGB projections. The FR image-based metrics include PSNR, SSIM [26], and PB-PCQA [6]. The NR image-based metrics include NIQE [52], BRISQUE [38], and PQA-net [29].
- Model-based metrics: Full-reference metrics operate directly from the 3D model, which include GraphSIM [53], PointSSIM [20], PCQM [23], and ResCNN [27]. Reduced-reference metric includes PCMRR [54]. Specifically, ResCNN is a model-based deep-learning approach.

2) *Experiment Setup for MQA*: The MQA method proposed in this paper is validated on the color mesh distortion measure (CMDM) database [7]. The database is generated from 5 source models subjected to geometry and color distortions. Then the source models are corrupted with 4 types of distortions based on color and geometry and each type of distortion is adjusted with 4 different strengths. The *Aix*, *Chamleon*, *Fish*, and *Samurai* models are selected for training while the *Ari* model is chosen for testing. Specifically, each distorted model is provided with 5 subjective scores according to its viewpoints and animation types. For simplification, we use the average of the 5 subjective scores as the final quality score for the distorted model.

In the literature, few metrics are proposed to deal with the 3D-QA tasks of color meshes. In order to evaluate the performance of the proposed method, some image-based metrics that might be able to predict the visual quality of 3D meshes are utilized as competitors. Unfortunately, few no-reference quality assessment metrics for 3D meshes are open-sourced, thus we try to reproduce some of the metrics. Specifically, the metrics used for comparison can be divided into two types:

- Image-based metrics: These metrics operate by evaluating the quality of 2D images rendered from the 3D colored meshes. The FR image-based metrics include PSNR and SSIM [26]. The NR image-based metrics include NIQE [52] and BRISQUE [38].
- Model-based metrics: The FR-MQA metric includes CMDM [7]. The NR-MQA metrics designed especially

for geometry-only 3D meshes include NR-SVR [30], NR-GRNN [55], NR-CNN [31].

### C. Evaluation Criterion

Four mainstream consistency evaluation criteria are utilized to compare the correlation between the predicted scores and MOSs, which include Spearman Rank Correlation Coefficient (SRCC), Kendall's Rank Correlation Coefficient (KRCC), Pearson Linear Correlation Coefficient (PLCC), and Root Mean Squared Error (RMSE). An excellent model should obtain values of SRCC, KRCC, and PLCC close to 1, and the value of RMSE near 0.

### D. Performance Discussion

1) *The Performance of PCQA Validated on the SJTU-PCQA and WPC Databases*: The experimental results for PCQA on the SJTU-PCQA database and the WPC database are shown in Table II. The best performance results of the NR-PCQA methods along with the best performance results of all methods are marked in bold. We can see clearly that the FR-PCQA metric PCQM outperforms other PCQA metrics. With the assistance of reference information, FR-PCQA methods tend to be more effective at predicting the quality levels of colored point clouds than NR-PCQA methods. The proposed method achieves first place on the SJTU-PCQA database and second place on the WPC database among the compared NR-PCQA metrics respectively. It's worth mentioning that the PQA-Net extracts features from the rendering views with deep-learning networks. The proposed method depends on handcrafted features and remains competitive considering the computational resource consumption of deep-learning networks.

With closer observations, we can make several more detailed analyses: 1-1) Image-based metrics using simple IQA models (such as PSNR and SSIM) are less effective. This is because the rendering views captured from various viewpoints are quite different in content. To get more quality information, more snapshots have to be sampled, which may confuse such IQA models because they are not capable of dealing with multiple pairs of images with dissimilar contents. Thanks to the strong learning ability of CNN, PQA-Net overcomes the difficulty

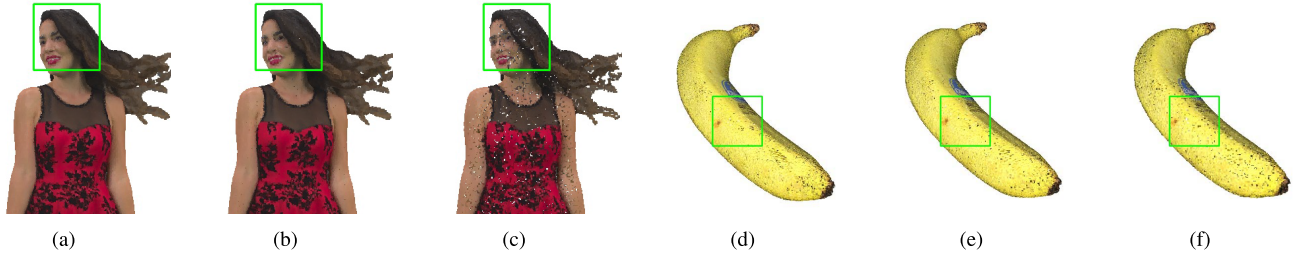


Fig. 5. Illustration of point clouds with consistently increasing downsampling distortion levels from the SJTU-PCQA [6] and the WPC [10] databases. (a), (b), and (c) are the snapshots of *rednandblack\_15*, *rednandblack\_16*, and *rednandblack\_17* from the SJTU-PCQA database. (d), (e), and (f) are the snapshots of *banana\_level\_9*, *banana\_level\_8*, and *banana\_level\_7* from the WPC database respectively.

TABLE III  
PERFORMANCE COMPARISON WITH COMPETITORS  
ON THE CMDM DATABASE

Ref	Type	Index	Metric	PLCC	SRCC	KRCC	RMSE
FR	Image-based	A	PSNR	0.7672	0.7735	0.7280	0.8832
		B	SSIM	0.7944	0.7817	0.7000	0.9656
	Model-based	C	CMDM	<b>0.9130</b>	<b>0.9000</b>	-	-
NR	Image-based	D	NIQE	0.4059	0.4768	0.3000	1.3352
		E	BRISQUE	0.5786	0.4882	0.3598	1.2237
	Model-based	F	NR-SVR	0.6082	0.4489	0.3420	1.3147
		G	NR-GRNN	0.6599	0.6948	0.5130	1.1121
		H	NR-CNN	0.5204	0.5022	0.3420	1.2804
		I	Proposed	<b>0.8626</b>	<b>0.8754</b>	<b>0.7222</b>	<b>0.6062</b>

and achieves high performance by utilizing the correlation of the rendering views. To conclude, we think that increasing the sample number of snapshots and optimizing learning models are the focus of image-based methods in the future. However, such methods inevitably require large amounts of computation resources and highly depend on the scale of the database. 1-2) Most model-based methods achieve lower performance on the WPC database than on the SJTU-PCQA database. We attempt to give the reasons. First, although the two databases manually introduce similar types of distortions, the WPC database employs some point cloud samples that are less sensitive to distortions in perceived visual quality. For example, as can be seen in Fig. 5, the *rednandblack* sample from the SJTU-PCQA database appears to be more distinct under the influence of downsampling distortions. However, the *banana* sample from the WPC database is less complex in geometry structure and limited in colorfulness, which makes the distortion difference less obvious. Second, the SJTU-PCQA database includes distorted point clouds with mixed distortions while the WPC database only adds one type of distortion to a single point cloud. It seems that point clouds with mixed distortions are more distinguishable in quality under close distortion levels. What's more, the WPC database has twice as many reference point clouds as the SJTU-PCQA database. The expanded scale and content diversity challenge the generalization ability of such methods as well.

2) *The Performance of MQA Validated on the CMDM Database:* Table III presents the experimental results on the CMDM database, in which the top performance results of the NR-PCQA methods along with the best performance results

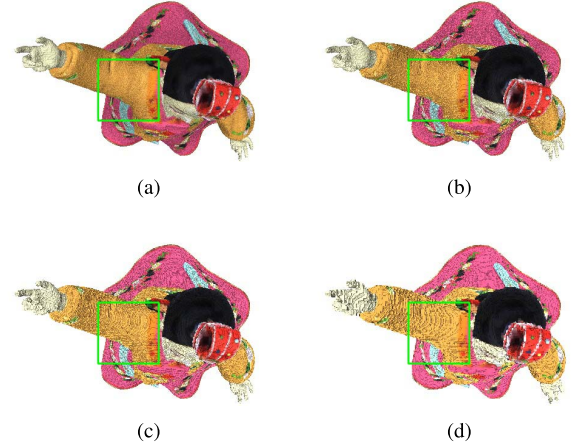


Fig. 6. Examples of meshes with increasing geometry quantization distortion levels from the CMDM database [7]. (a), (b), (c), and (d) are the snapshots of *Ari\_QuantGeom\_1*, *Ari\_QuantGeom\_2*, *Ari\_QuantGeom\_3*, and *Ari\_QuantGeom\_4* respectively.

of all methods are marked in bold. Surprisingly, the FR image-based metrics get remarkable performance results. It is because the CMDM database manually adjusts the employed 4 strengths of distortions to cover most range of visual quality levels (from imperceptible levels to high levels of impairment), which makes the distortion levels more coarse-grained. As illustrated in Fig. 6, the degradations caused by the 4 distortion strengths are quite obvious, thus enabling such simple methods to gain competitive performance.

The NR model-based MQA methods such as NR-SVR, NR-GRNN, and NR-CNN have a stronger ability to predict the visual quality than the NR image-based MQA methods, but they do not take color information into consideration, thus resulting in lower performance than the proposed method. The FR-MQA metric CMDM achieves first place in PLCC and SRCC and the proposed method obtains second place in PLCC and SRCC with a slight gap. The reason is that the coarse-grained levels of distortions make it easier for the proposed method to learn the relationship between the corresponding 3D-NSS parameters and quality scores. Therefore, even with relatively little training data, the proposed method still obtains good performance.

#### E. Ablation Study

To further test the effectiveness and contributions of different types of features (color and geometry) and different

TABLE IV  
PERFORMANCE OF THE ABLATION STUDY. THE BEST PERFORMANCE FOR EACH DATABASE IS MARKED IN BOLD

Type	Feature	SJTU-PCQA				WPC				CMDM			
		PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE
Geometry	F1	0.6406	0.6171	0.5283	1.8825	0.4280	0.4219	0.2801	19.5704	0.6214	0.4162	0.2852	0.9288
	F2	0.6082	0.5837	0.3963	1.8989	0.5042	0.4653	0.2992	18.7064	0.5872	0.4355	0.3022	1.0861
	F3	0.6259	0.5997	0.4166	1.9205	0.3874	0.3714	0.2350	20.4081	0.5656	0.3333	0.2481	1.0053
	F4	0.5187	0.4216	0.2713	1.9892	0.3889	0.3838	0.2761	21.2468	0.5358	0.4163	0.3096	1.0692
	(F1~F4)	0.7314	0.6937	0.5067	1.8096	0.5889	0.5827	0.4186	17.8330	0.6561	0.5076	0.3954	0.9267
Color	F5	0.1897	0.1250	0.0840	2.5758	0.1436	0.1327	0.0900	26.3404	0.1245	0.1101	0.0807	1.2254
	F6	0.2755	0.2170	0.1548	2.3927	0.1548	0.1009	0.0708	25.1412	0.0074	0.0291	0.0171	1.2582
	F7	0.0097	0.0069	0.0042	2.6471	0.1557	0.1836	0.1242	27.9339	0.1090	0.0793	0.0574	1.2556
	F8	0.0004	0.0012	0.0001	2.5288	0.1839	0.1488	0.0906	21.9701	0.1281	0.1860	0.0064	1.2617
	(F5~F8)	0.2351	0.2408	0.1685	2.4511	0.2220	0.1857	0.1599	24.3489	0.2985	0.1820	0.2283	1.2469
Basic	(F1+F5)	<b>0.7522</b>	<b>0.7449</b>	<b>0.5454</b>	<b>1.7004</b>	0.5335	0.5121	0.3647	21.4647	0.8481	0.8530	0.6955	0.6079
GGD	(F2+F6)	0.7173	0.6973	0.5035	1.7781	0.6172	0.5968	0.4353	17.1555	0.7819	0.7392	0.5821	0.7646
AGGD	(F3+F7)	0.6106	0.6122	0.4337	2.0208	0.4268	0.4243	0.2921	20.9246	0.8342	0.7974	0.6456	0.6989
Gamma	(F4+F8)	0.5841	0.4999	0.3504	2.0469	0.4584	0.4516	0.3177	19.8646	0.5982	0.5262	0.3951	0.9956
NSS	2D	0.3604	0.2123	0.1613	2.3125	0.4019	0.3919	0.2337	22.5500	0.5616	0.4772	0.2988	1.2347
	2D+3D	0.4122	0.3344	0.2951	2.2005	0.4658	0.4452	0.2500	21.2028	0.6658	0.6652	0.3700	0.9981
All	F1~F8	0.7382	0.7144	0.5174	1.7686	<b>0.6514</b>	<b>0.6479</b>	<b>0.4417</b>	<b>16.5716</b>	<b>0.8626</b>	<b>0.8754</b>	<b>0.7222</b>	<b>0.6062</b>

statistical parameters, we split the features into 8 groups: (1) F1: mean, standard deviation, and entropy values of geometry feature domains; (2) F2: GGD parameters of geometry feature domains; (3) F3: AGGD parameters of geometry feature domains; (4) F4: Gamma distribution parameters of geometry feature domains; (5) F5: mean, standard deviation, and entropy values of color feature domains; (6) F6: GGD parameters of color feature domains; (7) F7: AGGD parameters of color feature domains; (8) F8: Gamma distribution parameters of color feature domains. Then we can analyze the features' contributions by obtaining the performance of different combinations of feature groups. For example, the F2+F6 model only contains AGGD parameters of both geometry and color feature domains while the F1~F4 model only consists of statistical parameters of geometry feature domains.

Besides, most 3D models are perceived by humans only through the rendered 2D views. To find out whether the information of 2D views would help improve the performance of 3D-NSS, we test the effectiveness of the combined 2D-NSS and 3D-NSS features as well. The 2D-NSS features are extracted by the 2D views of objects using BRISQUE [38] and NIQE [52] (The redundant features of BRISQUE and NIQE are removed). The features extracted from different viewpoints are averaged to get the final 2D-NSS features. The proposed all groups' features in this paper are employed as the 3D-NSS features. We simply concatenate the 2D-NSS and 3D-NSS features as the quality-aware feature vector.

The performance results of the ablation study are shown in Table IV, where the mean, standard deviation, and entropy values of geometry and color feature domains are represented as the basic type of features (F1 + F5). It can be obviously observed that the geometry features contribute more to the final

quality score on all three databases. It might be because all three databases introduce more types of geometry distortions than color distortions and the geometry information weighs more in the human perception of 3D models.

By analyzing the different performances of different models, we can make several observations. 1) On the SJTU-PCQA database, we can see that the mean, standard deviation, and entropy values obtain the best performance, even higher than the proposed all groups. While on the WPC database, the all groups model is significantly superior to the separate distribution models. We attempt to give our explanation as follows. The ability of simple parameters to describe the shape of distributions is limited. For example, some distributions with very different shapes may have the same mean and variance. When the diversity and number of samples are not saturated, simpler parameters may achieve better results (more complex parameters are troubled by the redundant information), but when the diversity and number of samples increase, these simpler parameters are less effective, which requires the help of more complex parameters to describe the distributions. Therefore, we suggest using the proposed all groups model for practical application since such model is more capable of describing and distinguishing different distributions. 2) It seems that the 2D-NSS is not effective for 3D-QA tasks. We try to give the reason for such phenomenon as well. Single rendering 2D views contain limited information, therefore, we have to use multiple viewpoints to cover more quality information. However, the contents of snapshots from various viewpoints are quite different, such as the top and bottom views of the 3D models. The contents difference may have a much greater influence on the distributions rather than the distortions.

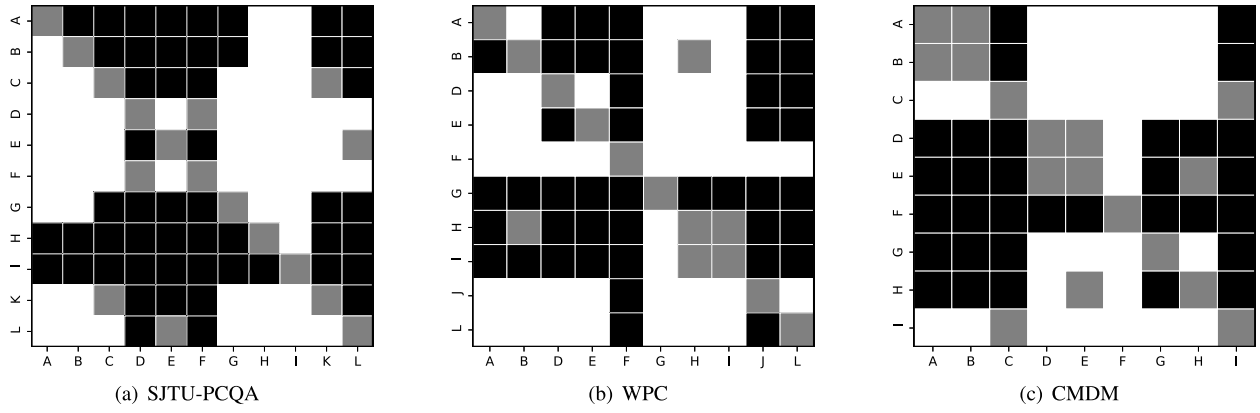


Fig. 7. Statistical test results of the proposed method and compared metrics on the SJTU-PCQA, WPC, and CMDM databases. A black/white block means the row method is statistically worse/better than the column one. A gray block means the row method and the column method are statistically indistinguishable. The metrics are denoted by the same index as in Table II and Table III respectively.

TABLE V  
PERFORMANCE OF THE DATA-SENSITIVITY EXPERIMENT. THE BEST PERFORMANCE FOR EACH DATABASE IS MARKED IN BOLD

Criteria	SJTU-PCQA				WPC				CMDM			
	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%
PLCC	0.6009	0.6458	0.6655	<b>0.7382</b>	0.5303	0.5575	0.5702	<b>0.6514</b>	0.6935	0.7464	0.8178	<b>0.8626</b>
SRCC	0.5851	0.6260	0.6462	<b>0.7144</b>	0.5208	0.5187	0.5542	<b>0.6479</b>	0.5807	0.6375	0.7422	<b>0.8754</b>
KRCC	0.4083	0.4402	0.4574	<b>0.5174</b>	0.3611	0.3655	0.3938	<b>0.4417</b>	0.4152	0.4682	0.5680	<b>0.7222</b>
RMSE	2.0708	1.9482	1.9018	<b>1.7686</b>	19.2921	18.9968	18.1487	<b>16.5176</b>	0.9064	0.8157	0.7208	<b>0.6062</b>

#### F. Statistical Test

To further analyze the performance of the proposed method, we conduct the statistical test in this section. We follow the same experiment setup as in [56] and compare the difference between the predicted quality scores with the subjective ratings. All possible pairs of models are tested and the results are listed in Fig. 7. It can be seen that our method is significantly superior to 7 compared PCQA metrics on the SJTU-PCQA and the WPC databases while our method also significantly outperforms 7 compared MQA metrics on the CMDM database. Specifically, the FR metric PCQM achieves significantly better performance than our method on both SJTU-PCQA and WPC databases, the FR metric CMDM is insignificantly distinguishable from our method on the CMDM database.

#### G. Data-Sensitivity Experiment

To find out the influence of the number of training samples, we conduct the data-sensitivity experiment by changing the proportion of training set as about 20%, 40%, 60%, and 80% (2, 4, 6, 8 sample groups for training on the SJTU-PCQA database, 4, 8, 12, 16 for training on the WPC database, and 1, 2, 3, 4 samples for training on the CMDM database correspondingly). The experimental results are exhibited in Table V, from which we can find that increasing the number of training samples is helpful to improve the performance of the proposed method. The CMDM database employs coarse-grained distortion levels and includes relatively small amounts of mesh samples, thus the increase of training data improves

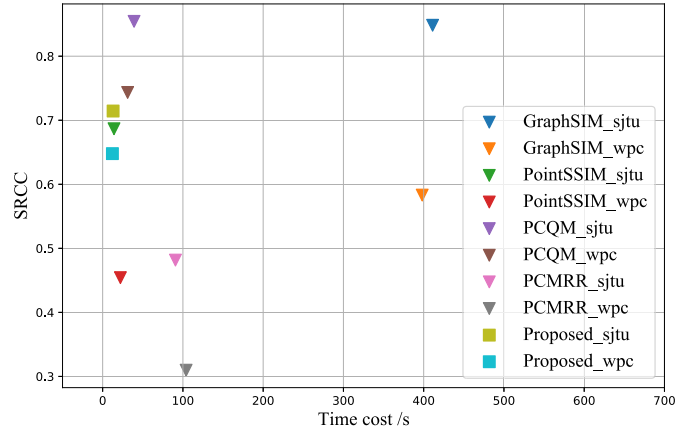


Fig. 8. The results of time cost comparison on the SJTU-PCQA and the WPC databases. The time cost refers to the average time consumption per point cloud for each database.

the performance more significantly. For the PCQA database, we think using fewer training samples may exaggerate the impact of noisy labels and result in the over-fit of the model, which causes the performance to drop. In all, increasing the diversity and number of the training samples are beneficial for obtaining more robustness and higher performance.

#### H. Computational Efficiency

Considering that the proposed method operates directly from the 3D model, we also focus on computational efficiency. The image-based metrics usually operate quite fast due to



the mature development of IQA metrics. Therefore, to make the comparison meaningful, we select model-based methods GraphSIM, PointSSIM, PCQM, and PCMR as competitors. We conduct the test on a computer with Intel (R) Core (TM) i5-3470 CPU @ 3.20 GHz and 8 GB RAM on the Windows platform. The corresponding time cost results are shown in Fig. 8, and we can see clearly that the average time cost of the proposed method is smaller than all the compared model-based methods, which indicates that our method achieves relatively considerable computational efficiency.

## V. CONCLUSION

This paper proposes a no-reference colored 3D model quality assessment metric based on entropy and 3D natural scene statistics. The proposed method deals with the quality assessment problems for both colored point cloud and mesh models. We first project the 3D models into corresponding quality-related geometry and color feature domains. Then entropy and various 3D-NSS parameters are estimated to better capture the representative characteristics and quantify the distortions that are more in line with human perception. The proposed method is validated on the colored point cloud quality assessment database (SJTU-PCQA), the Waterloo point cloud assessment database (WPC), and the colored mesh quality assessment database (CMDM). The experimental results show that our method outperforms most compared NR 3D-QA metrics with competitive computational resources and reduces the performance gap with the state-of-the-art FR 3D-QA metrics. The proposed method follows a common NR framework and can be easily modified and expanded to satisfy specific needs, which has great application potential. The code of the proposed method is released for promoting the development of NR 3D-QA.

## REFERENCES

- [1] H. Graf, S. P. Serna, and A. Stork, "Adaptive quality meshing for 'on-the-fly' volumetric mesh manipulations within virtual environments," in *Proc. IEEE Symp. Virtual Environ., Hum.-Comput. Interfaces Meas. Syst.*, Jul. 2006, pp. 178–183.
- [2] L. Li, Z. Li, S. Liu, and H. Li, "Occupancy-map-based rate distortion optimization and partition for video-based point cloud compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 1, pp. 326–338, Jan. 2021.
- [3] R. Mekuria, K. Blom, and P. Cesar, "Design, implementation, and evaluation of a point cloud codec for tele-immersive video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 4, pp. 828–842, Apr. 2017.
- [4] (May 1, 2022). *Apple Developer Document: Overview of Scanning and Detecting 3-Dimension Objects*. [Online]. Available: <https://wrywhisker.pulpfriktion.net/wallcrust/linear-colinear-felinear.html>
- [5] (May 1, 2022). *Intel Realsense*. [Online]. Available: <https://www.intelrealsense.com/>
- [6] Q. Yang, H. Chen, Z. Ma, Y. Xu, R. Tang, and J. Sun, "Predicting the perceptual quality of point cloud: A 3D-to-2D projection-based exploration," *IEEE Trans. Multimedia*, vol. 23, pp. 3877–3891, 2021.
- [7] Y. Nehmé, F. Dupont, J. P. Farrugia, P. Le Callet, and G. Lavoué, "Visual quality of 3D meshes with diffuse colors in virtual reality: Subjective and objective evaluation," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 3, pp. 2202–2219, Mar. 2021.
- [8] E. Alexiou, I. Viola, T. M. Borges, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, "A comprehensive study of the rate-distortion performance in MPEG point cloud compression," *APSIPA Trans. Signal Inf. Process.*, vol. 8, no. 1, p. e27, 2019.
- [9] E. M. Torlig, E. Alexiou, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, "A novel methodology for quality assessment of voxelized point clouds," *Proc. SPIE*, vol. 10752, pp. 174–190, Sep. 2018.
- [10] H. Su, Q. Liu, Z. Duanmu, W. Liu, and Z. Wang, "Perceptual quality assessment of colored 3D point clouds," 2021, *arXiv:2111.05474*.
- [11] W. Sun, X. Min, G. Zhai, and S. Ma, "Blind quality assessment for in-the-wild images via hierarchical feature fusion and iterative mixed database training," 2021, *arXiv:2105.14550*.
- [12] P. Cignoni, C. Rocchini, and R. Scopigno, "METRO: Measuring error on simplified surfaces," *Comput. Graph. Forum*, vol. 17, no. 2, pp. 167–174, Sep. 1998.
- [13] R. Mekuria and P. Cesar, "MP3DG-PCC, open source software framework for implementation and evaluation of point cloud compression," in *Proc. 24th ACM Int. Conf. Multimedia*, 2016, pp. 1222–1226.
- [14] G. Lavoué, "A multiscale metric for 3D mesh visual quality assessment," *Comput. Graph. Forum*, vol. 30, no. 5, pp. 1427–1437, 2011.
- [15] K. Wang, F. Torkhani, and A. Montanvert, "A fast roughness-based approach to the assessment of 3D mesh visual quality," *Comput. Graph.*, vol. 36, no. 7, pp. 808–818, Nov. 2012.
- [16] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Geometric distortion metrics for point cloud compression," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3460–3464.
- [17] E. Alexiou and T. Ebrahimi, "Point cloud quality assessment metric based on angular similarity," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2018, pp. 1–6.
- [18] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "A generalized Hausdorff distance based quality metric for point cloud geometry," in *Proc. 12th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2020, pp. 1–6.
- [19] L. Váša and J. Rus, "Dihedral angle mesh error: A fast perception correlated distortion measure for fixed connectivity triangle meshes," *Comput. Graph. Forum*, vol. 31, no. 5, pp. 1715–1724, Aug. 2012.
- [20] E. Alexiou and T. Ebrahimi, "Towards a point cloud structural similarity metric," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2020, pp. 1–6.
- [21] D. Tian and G. AlRegib, "BaTex3: Bit allocation for progressive transmission of textured 3-D models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 23–35, Jan. 2008.
- [22] J. Guo, V. Vidal, I. Cheng, A. Basu, A. Baskurt, and G. Lavoué, "Subjective and objective visual quality assessment of textured 3D meshes," *ACM Trans. Appl. Perception*, vol. 14, no. 2, pp. 1–20, Feb. 2017.
- [23] G. Meynet, Y. Nehmé, J. Digne, and G. Lavoué, "PCQM: A full-reference quality metric for colored 3D point clouds," in *Proc. 12th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2020, pp. 1–6.
- [24] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, and H. Yang, "Reduced reference perceptual quality model with application to rate control for video-based point cloud compression," *IEEE Trans. Image Process.*, vol. 30, pp. 6623–6636, 2021.
- [25] Z. Zhang *et al.*, "A no-reference visual quality metric for 3D color meshes," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2021, pp. 1–6.
- [26] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [27] Y. Liu, Q. Yang, Y. Xu, and L. Yang, "Point cloud quality assessment: Dataset construction and learning-based no-reference approach," 2020, *arXiv:2012.11895*.
- [28] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 652–660.
- [29] Q. Liu *et al.*, "PQA-Net: Deep, no., reference point cloud quality assessment via multi-view projection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 12, pp. 4645–4660, Dec. 2021.
- [30] I. Abouelaziz, M. El Hassouni, and H. Cherifi, "No-reference 3D mesh quality assessment based on dihedral angles model and support vector regression," in *Proc. Int. Conf. Image Signal Process.* Cham, Switzerland: Springer, May 2016, pp. 369–377.
- [31] I. Abouelaziz, M. E. Hassouni, and H. Cherifi, "A convolutional neural network framework for blind mesh visual quality assessment," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 755–759.
- [32] I. Abouelaziz, A. Chetouani, M. El Hassouni, L. J. Latecki, and H. Cherifi, "No-reference mesh visual quality assessment via ensemble of convolutional neural networks and compact multi-linear pooling," *Pattern Recognit.*, vol. 100, Apr. 2020, Art. no. 107174.

- [33] S. Yang, C.-H. Lee, and C.-C.-J. Kuo, "Optimized mesh and texture multiplexing for progressive textured model transmission," in *Proc. 12th Annu. ACM Int. Conf. Multimedia*, 2004, pp. 676–683.
- [34] F. Caillaud, V. Vidal, F. Dupont, and G. Lavoué, "Progressive compression of arbitrary textured meshes," *Comput. Graph. Forum*, vol. 35, no. 7, pp. 475–484, Oct. 2016.
- [35] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, vol. 2, 2003, pp. 1398–1402.
- [36] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [37] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [38] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
- [39] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
- [40] Y. Lin, M. Yu, K. Chen, G. Jiang, Z. Peng, and F. Chen, "Blind mesh quality assessment method based on concave, convex and structural features analyses," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2019, pp. 282–287.
- [41] I. Abouelaziz, M. El Hassouni, and H. Cherifi, "Blind 3D mesh visual quality assessment using support vector regression," *Multimedia Tools Appl.*, vol. 77, no. 18, pp. 24365–24386, Sep. 2018.
- [42] Q. Mérigot, M. Ovsjanikov, and L. J. Guibas, "Voronoi-based curvature and feature estimation from point clouds," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 6, pp. 743–756, Jun. 2011.
- [43] H. Thomas, J. Deschaud, B. Marcotegui, F. Goulette, and Y. L. Gall, "Semantic classification of 3D point clouds with multiscale spherical neighborhoods," 2018, *arXiv:1808.00495*.
- [44] T. Hackel, J. D. Wegner, and K. Schindler, "Fast semantic segmentation of 3D point clouds with strongly varying density," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 3 pp. 177–184, Jul. 2016.
- [45] T. Surazhsky, E. Magid, O. Soldea, G. Elber, and E. Rivlin, "A comparison of Gaussian and mean curvatures estimation methods on triangular meshes," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 1, Sep. 2003, pp. 1021–1026.
- [46] P. Alliez, D. Cohen-Steiner, O. Devillers, B. Lévy, and M. Desbrun, "Anisotropic polygonal remeshing," in *Proc. ACM SIGGRAPH Papers*, 2003, pp. 485–493.
- [47] M. Corsini, E. D. Gelasca, T. Ebrahimi, and M. Barni, "Watermarked 3-D mesh quality assessment," *IEEE Trans. Multimedia*, vol. 9, no. 2, pp. 247–256, Feb. 2007.
- [48] N. Mukherjee, "A hybrid, variational 3D smoother for orphaned shell meshes," in *Proc. IMR*, 2002, pp. 1–12.
- [49] H. Lee, P. Alliez, and M. Desbrun, "Angle-analyzer: A triangle-quad mesh codec," *Comput. Graph. Forum*, vol. 21, no. 3, pp. 383–392, Sep. 2002.
- [50] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 1, pp. 52–56, Feb. 1995.
- [51] (May 1, 2022). *Scikit-Learn for Python*. [Online]. Available: <https://scikit-learn.org/stable/>
- [52] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Nov. 2013.
- [53] Q. Yang, Z. Ma, Y. Xu, Z. Li, and J. Sun, "Inferring point cloud quality via graph similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 6, pp. 3015–3029, Jun. 2022.
- [54] I. Viola and P. Cesar, "A reduced reference metric for visual quality evaluation of point cloud contents," *IEEE Signal Process. Lett.*, vol. 27, pp. 1660–1664, 2020.
- [55] I. Abouelaziz, M. El Hassouni, and H. Cherifi, "A curvature based method for blind mesh visual quality assessment using a general regression neural network," in *Proc. 12th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, 2016, pp. 793–797.
- [56] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.



**Zicheng Zhang** (Student Member, IEEE) received the B.E. degree from Shanghai Jiao Tong University, Shanghai, China, in 2020, where he is currently pursuing the Ph.D. degree with the School of Electronic Information and Electrical Engineering. His research interests include image quality assessment, video quality assessment, and 3D visual quality assessment.



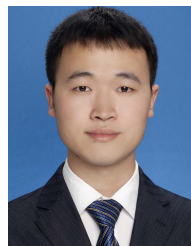
**Wei Sun** received the B.E. degree from the East China University of Science and Technology, Shanghai, China, in 2016. He is currently pursuing the Ph.D. degree with the Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University. His research interests include image quality assessment, perceptual signal processing, and mobile video processing.



**Xionghuo Min** (Member, IEEE) received the B.E. degree from Wuhan University, Wuhan, China, in 2013, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2018. From January 2016 to January 2017, he was a Visiting Student at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. He is currently a Post-Doctoral Fellow with Shanghai Jiao Tong University. His research interests include visual quality assessment, visual attention modeling, and perceptual signal processing. He received the Best Student Paper Award from IEEE ICME in 2016.



**Tao Wang** received the B.E. degree from Shanghai Jiao Tong University, Shanghai, China, in 2020, where he is currently pursuing the M.S. degree with the School of Electronic Information and Electrical Engineering. His research interests include image/video quality assessment and image aesthetic quality assessment.



**Wei Lu** received the B.E. degree from Shanghai Jiao Tong University, Shanghai, China, in 2020, where he is currently pursuing the master's degree with the Department of Electronic Engineering. His research interests include image quality assessment and video quality assessment.



**Guangtao Zhai** (Senior Member, IEEE) received the B.E. and M.E. degrees from Shandong University, Shandong, China, in 2001 and 2004, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2009. From 2008 to 2009, he was a Visiting Student with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he was a Post-Doctoral Fellow from 2010 to 2012. From 2012 to 2013, he was a Humboldt Research Fellow with the Institute of Multimedia Communication and Signal Processing, Friedrich Alexander University of Erlangen-Nuremberg, Germany. He is currently a Research Professor with the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University. His research interests include multimedia signal processing and perceptual signal processing. He received the Award of National Excellent Ph.D. Thesis from the Ministry of Education of China in 2012.