IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 10, 2021, accepted May 31, 2021, date of publication June 7, 2021, date of current version June 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3087183

Layered Projection-Based Quality Assessment of 3D Point Clouds

TIANXIN CHEN[1], CHUNYI LONG[1], HONGLEI SU[1], LIJUN CHEN[1], JIERU CHI[1], ZHENKUAN PAN[2], HUAN YANG[2], AND YUXIN LIU[2]

[1]School of Electronic Information, Qingdao University, Qingdao 266071, China
[2]School of Computer Science and Technology, Qingdao University, Qingdao 266071, China

Corresponding author: Honglei Su (suhonglei@qdu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61772294, in part by the Natural Science Foundation of Shandong Province under Grant ZR2018PF002, in part by the Key Research and Development Program of Shandong Province (Special Project for Public Welfare) under Grant 2019GGX101021, in part by the Joint funding for Smart Computing of Shandong Natural Science Foundation of China under Grant ZR2019LZH002, and in part by the State Key Laboratory of High Performance Server and Storage Technology, Inspur Group, Jinan, China.

ABSTRACT Point clouds are subject to various distortions during point cloud processing missions, any of which may lead to quality degradation. Consequently, predicting point cloud quality has attracted a lot of attention. In this paper, a layered projection-based point cloud quality metric (LP-PCQM) is proposed. We layer the distorted point cloud and its original version firstly and then extract the geometry and color features of layers. The geometry feature is obtained using the projection-based method and the color features are extracted upon RGB by using the point-based method. Finally, the LP-PCQM is a weighted linear combination of an optimal subset of these pooled geometry and color features of layers. To verify the performance of LP-PCQM, we compare it with other eight metrics including both point-based metrics and projection-based metrics on the WPC, SJTU-PCQA, and ICIP2020 database respectively. Experimental results show that the proposed metric exhibits better and more robust performance.

INDEX TERMS Layered projection, image quality assessment, 3D point cloud, point cloud quality assessment.

I. INTRODUCTION

Thanks to the advancement of imaging technologies, more and more formats are used to present visual content in the immersive media field. Depending on the capturing device, those formats can correspond to holograms [1], light fields [2], or point cloud (PC) [3], etc. Among these, PCs denote a practical content representation that allows users to visualize static or dynamic scenes in a more immersive way and can be directly exploited in systems aiming to provide immersive experiences with higher degrees of freedom [4], so PC recently becomes one of the first choices to represent 3D visual contents and has a wider variety of applications with the development of VR/AR [5], [6].

Compared to 2D images, 3D PCs have their geometric coordinates and other associated attributes that describe surface properties. To describe accurately a 3D scene, PCs require a large number of points, which limits their use in actual applications. Most PC processing missions, e.g., compression [7]–[9], transmission [10], display [11]–[13] and other technologies [14], [15] will bring in various distortions and lead to quality degradation reducing the satisfaction of the quality of experience (QoE) of end users. For applications in which PCs are ultimate to be viewed by human beings, the only ''correct'' method of quantifying visual quality is the subjective evaluation. However, the subjective evaluation is time-consuming and expensive, and cannot be directly embedded into a practical system as the optimization metric [16]. Naturally, the majority of the efforts on quality assessment have been devoted to the development of objective metrics [17], which is to develop quantitative measures that can automatically predict perceived PC quality. It is worth mentioning that objective PC quality assessment (PCQA) methods can be used not only to dynamically monitor and adjust PC quality to achieve a good perceived quality for end-users but also to optimize algorithms and parameter settings of PC processing methods.

The associate editor coordinating the review of this manuscript and approving it for publication was Jinjia Zhou.

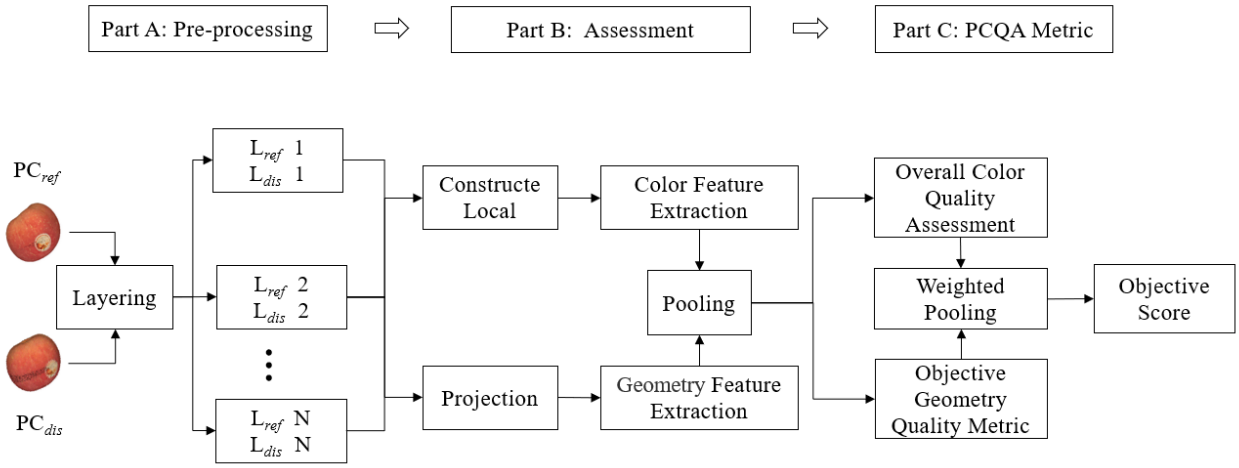


FIGURE 1. Framework of layered projection-based point cloud quality metric (LP-PCQM). *N* represents the total number of layers.

Objective PCQA metrics are commonly classified as Full Reference (FR) [17]–[28], Reduced Reference (RR) [29], [30], and No Reference (NR), depending on the availability of reference information. Compared with RR and NR metrics, people have more research on FR metrics which can be classified as (a) point-based [18]–[25] and (b) projection-based [17], [26]–[28]. Most studies in the past had emphasized on the geometric distortion measurement of PC object, such as the point-to-point (po2point), point-to-plane (po2plane), point-to-mesh (po2mesh) [18] and plane-to-plane (pl2plane) [19]. Recently, curvature statistics which were initially introduced and applied on polygonal meshes [31] have also been proposed to estimate the distortion of a PC concerning its reference [20] and they have been extended to include color information [21]. Different from the above methods, Yang *et al.* [22], Liu *et al.* [23], Viola *et al.* [24], and Alexiou and Ebrahimi [25] developed new research directions for PCQA from breakthroughs, such as deep learning, histogram, or SSIM [32] and so on.

We propose a layered projection-based point cloud quality metric (LP-PCQM). Fig. 1 shows the framework of the proposed LP-PCQM. We firstly layer the reference PC (PC$_{ref}$) and the distorted PC (PC$_{dis}$) and then extract the layers of features related to geometry and color respectively. Geometry features are based on layered projection and image metrics are employed. This process adopts the layered projection method, which is to project the distorted PC layer (L$_{dis}$) and its corresponding reference PC layer (L$_{ref}$) onto the plane fitted by the L$_{ref}$ to obtain the information images of layers, instead of mapping onto six different planes. Extending the definition in [21], color features of layers are extracted from the RGB of points and their local neighborhoods. The proposed metric is then computed as a linear combination (computed through multiple linear regression analysis) of an optimal subset of these pooled geometry and color features of layers.

This paper is organized as follows: Section II is the related work, which summarizes the existing PCQA metrics and analyzes their advantages and disadvantages. We describe the proposed method in section III. Section IV are experimental results and analysis, including the determination of model parameters and the performance comparison of the proposed metric and others. Section V is the conclusion of this paper.

II. RELATED WORK

Inspired by the vast amount of previous works on image/video quality assessment, these PCQA metrics are mostly FR [17]–[28] meaning that a complete reference content is assumed to be known and NR/RR [29], [30] PCQA metrics are introduced recently.

RR and NR metrics can be usefully employed when little to no information is known about the original content [30]. Liu *et al.* [29] proposed a RR model to accurately predict the mean opinion score (MOS) of V-PCC compressed 3D PCs from the quantization parameters of the geometry and color encoders. Viola and Cesar [30] extracted a small set of features from a given reference content based on both structure and attribute domains. Such features were then transmitted alongside the content, and were used at the receiver side to predict the visual quality of the content under exam. Moreover, they combined the proposed features through a linear optimization algorithm. In the following, we mainly study two categories of FR PCQA metrics: point-based [18]–[25] and projection-based [17], [26]–[28].

In the first category, the earliest researches, only considering geometric information, calculated the distance error, normal vector, and tangent plane between the PC$_{ref}$ and the PC$_{dis}$, which are classified as po2point, po2plane, po2mesh, and pl2plane. The po2mesh metric is based on projected distances between the PC under evaluation and the reference object. Considering that the objective scores heavily depend on the selected surface reconstruction algorithm.

Thus, the po2mesh metric is considered as a sub-optimal solution for the quality assessment of PCs [4]. The po2point metric is based on geometric distances of associated points between the PC_{ref} and PC_{dis}, but it does not consider the fact that points in a PC usually represent surfaces on the object. The po2plane metric is based on the projected error along the normal of a reference point and, essentially, larger costs are assigned to points that deviate from the underlying surface [18]. The pl2plane metric is based on the angular similarity of tangent planes. PC-MSDM [20] is based on local curvature statistics and can be viewed as an extension for PCs of the MSDM [33] metric suited for 3D meshes. However, both of them are limited by the high complexity of searching for the neighboring points to construct the normal or curvature. The state-of-the-art point-based methods that assess the color of a PC_{dis} are based on conventional formulas used in 2D content representations. In particular, points that belong to the content under assessment are associated with points that belong to the reference model, using typically the nearest neighbor algorithm. And the total color degradation value is based either on the color MSE, or the PSNR [17]. However, the metric only considers color information. Recently, Meynet *et al.* introduced PCQM, which is an optimally weighted linear combination of geometry-based and color-based features [21]. Viola *et al.* incorporated color distortion in geometry-based metrics, using luminance histogram information [24]. However, the histogram works by introducing spatial correlation as a way to describe changes in color statistics. As such, it does not describe the distribution of colors in the content, but rather the color distribution of points concerning each other, which leads to a poorer correlation with close neighbors results. Whereas Alexiou *et al.* proposed the usage of local statistical features in order to obtain a global measure of degradation, similarly to the Structural Similarity (SSIM) in the image domain [25]. In [22], Yang *et al.* filtered the PC firstly and then evaluated the quality prediction considering the impact of color transformation on quality. However, this model presents relatively high computation complexity due to some operations, such as graph construction [27]. With the proposal of PC deep learning networks [34] such as PointNet [35], Liu *et al.* [23] used a mathematical model of Artificial Neural Networks to evaluate the quality of PC. It is based on the stack of sparse convolutional layers and residuals, extracts hierarchical features, and pools them globally to obtain feature vectors. Finally, the subsequent network sends the feature vectors to the regression module to predict the final quality fraction. Although the metric has better robustness to different distorted PCs, the performance may be reduced due to insufficient training set on a small data set. At present, the amount of data on the public PC databases is not large enough, so this method has certain limitations.

The above point-based methods, considering geometry, color, or geometry-plus-color, depend on individual errors that are assigned to pairs of associated points. Conversely, the projection-based methods [17], [26]–[28] are to transform the 3D PC into 2D images through projecting and then evaluate the PC quality, which not only are able to capture both geometry and color distortions but also evaluate the quality of PCs from the human eye perception. In the projection-based approaches, Torlig *et al.* projected voxelized PCs onto 2D planes and conventional 2D imaging metrics are employed. Moreover, the number of projections to compute the objective scores was set to six, while each view was treated as of equal importance [28]. In principle, different perspectives of a 3D model might be of different importance, as they could be more or less representative or informative regarding the presented content. So the projection-based PCQA was extended by investigating the impact of applying different camera layouts to capture views of the models, as well as exploiting user interactivity data in [26]. Alexiou *et al.* [17] proposed that snapshots of the models are typically acquired from the software used for consumption in order to reflect the views observed by the users. However, this method needs to consider additional influencing factors, such as the distance between the content and the cameras, the direction of the cameras, the lighting conditions, and the type of projection (e.g., orthographic, perspective). So we do not consider this method of capturing views. Yang *et al.* chose to project the 3D PC onto six perpendicular image planes of a cube for the color texture image and corresponding depth image and aggregated image-based global and local features (e.g., edge, depth, pixel-wise similarity, complexity) among all projected planes for a final objective metric [27]. Compared with the previous one, this projection method does not have to consider many additional influencing factors, but it will lose detailed information (see Section III-A2). In short, how to derive an effective PC quality assessment model over 2D planes requires further exploration.

III. A LAYERED PROJECTION-BASED POINT CLOUD QUALITY METRIC

Considering that PCs have at least two types of information per point (color and geometry), the main idea of the proposed method is to extract color and geometry features from the PC_{ref} and PC_{dis}. The proposed method is divided into the following stages: (1) Preprocessing of PC_{ref} and PC_{dis} respectively. We layer PCs firstly and then proceed in two steps. One step is to project PC layers for extracting geometric features, and another step is to establish local neighborhoods among PC layers for extracting color features. (2) Geometry and color quality assessment. (3) Overall quality assessment.

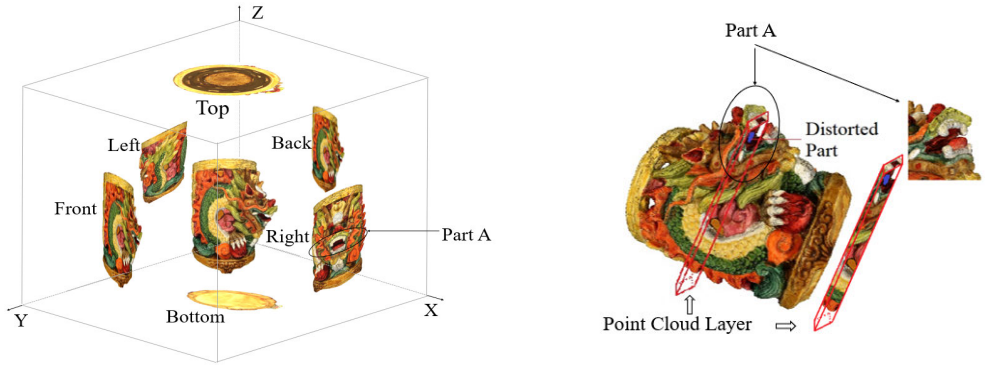
A. POINT CLOUD PRE-PROCESSING

1) LAYERING

We propose the concept of PC layering for two reasons. One reason is to reduce time complexity. The PC data is scattered, disorderly and huge. We have to traverse the entire PC for each search when searching for the nearest neighbors of all points. After layering, each point narrows the search range to reduce time complexity when we extract color features. The other reason is that it is beneficial for overcoming the

(a) Point cloud (b) Reference point cloud layer (c) Distorted point cloud layer

FIGURE 2. Schematic diagram of layering effect.



(a) Point cloud is overall mapped to six planes (b) Layer superiority diagram

FIGURE 3. The multi-sides projection and layered projection.

shortcomings of multi-sides overall projection as seen in Section III-B1.

The method of layering is as follows. A point in a PC has its coordinates (x, y, z) and attributes, such as color values, reflectivities and normal vectors. According to the coordinates, the PC is divided into N layers and Z_{max} is the maximum value of z of PC_{ref} or PC_{dis}:

$$d = \frac{Z_{max}}{N}. \tag{1}$$

where d is the thickness of each layer. Then the layer number of each point can be record as t,

$$t = \lfloor \frac{z}{d} \rfloor. \tag{2}$$

Fig. 2 shows the effect of dividing certain layer of the apple, where L_{ref} means the PC_{ref} layer, and L_{dis} means the PC_{ref} layer. After layering, we project the PCs' layers and construct local neighborhoods.

2) LAYERED-PROJECTION

The multi-sides overall projection method is to project the entire PC on multiple different planes. The projection method we mainly study in this paper is that each 3D PC is mapped to six perpendicular 2D image planes via the orthographic projection. These image planes, e.g., referred to as the ''front'', ''back'', ''left'', ''right'', ''top'' and ''bottom'' planes, correspond to six faces of a cube [27], as shown in Fig. 3 (a). The color value of every point with spatial coordinates is associated with an image pixel in the respective projected image. During the projection, if another point with identical coordinates and a smaller distance from the projection plane is identified, the first point is ignored, and the color value of the pixel is given by the second point [28]. This may lose some detailed information for the part A appearing to be surrounded as shown in Fig. 3 (b) because the part A is not the closest to the projected planes. So the multi-sides projection has some limitations to a certain extent. In order to solve this problem, we propose the layered-based projection. Only observing the PC layer as shown in Fig. 3 (b), the details of part A are more obvious. Considering that only a certain number of projection planes are used in practice and one projection view could be enough to achieve high performance [26], we project the PC on the fitting plane obtained by the least-squares method [36] which is an optimization process. The purpose of establishing the fitting plane is to reduce the points that are projected on the same position repeatedly and to better reflect the geometric structure on the plane for considering all points. We perform PCs projection via MATLAB and implement it as follows: Firstly, L_{ref} and L_{dis} project on the same fitting plane F obtained by

$$F = Ax + By + Cz + D, \tag{3}$$

where A, B, C and D are the coefficients of the fitting equation, and x, y and z are the 3D coordinates of points of the corresponding PC layer. Secondly, for projected points of sample, we plot them on the plane as the information images

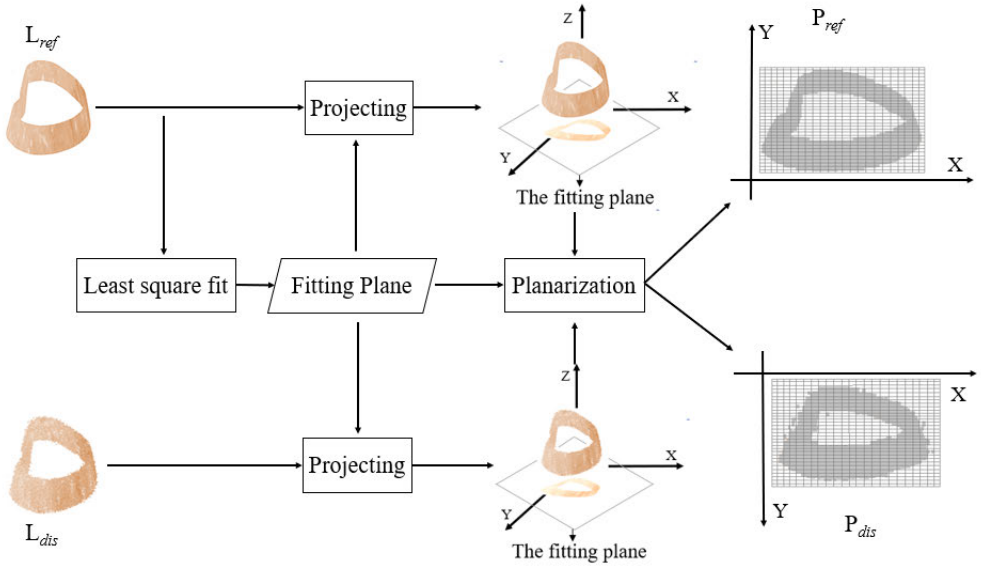


FIGURE 4. The formation of layered projection.

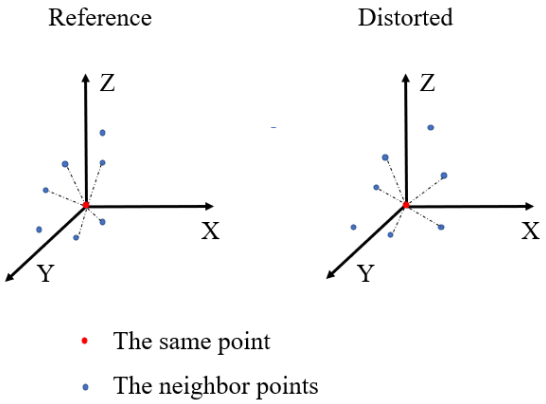


FIGURE 5. Two local neighborhoods of the L_{ref} and the L_{dis}.

storing the corresponding gray value of the PCs via its x coordinates and y coordinates. If multiple points share same x and y while different z, the associated information image value is derived as the average of the gray values of the points. This is the planarization process, where the L_{ref} of 2D image is P_{ref} and the L_{dis} of 2D image is P_{dis}. The layered projection process of PC is shown in Fig. 4.

3) NEIGHBORHOOD DETERMINATION

To establish a correspondence preparing for color features extraction, for each point belonging to PC_{ref}, the existing approach in [19] simply considers the nearest neighbor of p belonging to PC_{dis}. However, this will ignore the surface structures leading to inaccurate correspondence. In [21], Meynet *et al.* searched, for each p, the projection $\hat{p}$ on the 3D surface subtended by PC_{dis} and established two local neighborhoods with p and $\hat{p}$ as the center, but it heavily tends to rely on specific surface reconstructions. In order to establish a more effective and simple correspondence, we propose a new method combining the above two ideas. The same points are found in the L_{ref} and L_{dis} as the center to create local neighborhoods using the K-nearest neighbor method [37] in preparation for the establishment of color-related features as shown in Fig. 5. To choose the best neighborhood computation parameters for the proposed metric, we tested different neighborhood values (K=15 and k=20), but found no major performance advantages. Therefore, K is set as 20 in LP-PCQM.

B. GEOMETRIC AND COLOR QUALITY ASSESSMENT

1) GEOMETRIC QUALITY ASSESSMENT

In the field of 2D images, image quality metrics include the mean squared error (MSE) [38], Peak Signal-to-Noise Ratio (PSNR) [39], Structural Similarity (SSIM) [32], and so on. MSE is computed by averaging the squared intensity differences of distorted and reference image pixels, which objectively quantifies the strength of the error signal. However, two distorted images with the same MSE may have different types of errors, some of which are much more visible than others. And people often found it hard to understand the MSE between multiple PCs. So PSNR appeared. Let x_v and y_v be the vth pixel in the original image x and the distorted image y, respectively. The MSE and PSNR between the two images are given by

$$\mathrm{MSE} = \frac{1}{V}\sum_{v=1}^{V}(x_v - y_v)^2, \tag{4}$$

$$\mathrm{PSNR} = 10log_{10}(\frac{L^2}{\mathrm{MSE}}), \tag{5}$$

where V is the total number of pixels in the image and L is the maximum dynamic range. For 8 bits gray-scale images,

L is 255. The PSNR value approaches infinity as the MSE approaches zero. This shows that a higher PSNR value provides a higher image quality. At the other end of the scale, a small value of the PSNR implies high numerical differences between images [40]. However, the PSNR is based on the error between the corresponding pixels, and the difference value is not proportional to the subjective feelings of people. Instead, SSIM considers to be correlated with the quality perception of the human visual system (HVS) [41] and is based upon separated comparisons of local luminance, contrast, and structure between original and distorted images. Given two images x_v and y_v, the SSIM can be calculated in the following way:

$$\mathrm{SSIM}(x_v, y_v) = l(x_v, y_v)c(x_v, y_v)s(x_v, y_v), \tag{6}$$

$$l(x_v, y_v) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \tag{7}$$

$$c(x_v, y_v) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \tag{8}$$

$$s(x_v, y_v) = \frac{2\sigma_{xy} + C_3}{\sigma_x + \sigma_y + C_3} \tag{9}$$

where μ_x is the mean of x_v, μ_y is the mean of y_v, σ_x is the standard deviation of x_v, σ_y is the standard deviation of y_v, σ_{xy} is the covariance of x_v and y_v. C_1, C_2 and C_3 are the constant used to maintain stability. Instead of using traditional error summation methods, such as MSE and PSNR, the SSIM improves the disadvantages of PSNR and has better performance in image quality assessment (IQA) [42]. Besides, the evaluation results of SSIM is more consistent with the human eyes. This is more helpful for PCQA of the HVS oriented. In addition, the perception of 3D geometric objects by human eyes has the same structural characteristics as that of 2D images. The pixel points of 2D image belong to plane distribution and regular arrangement. However, 3D geometric objects belong to spatial distribution and are irregular arrangement [43]. So the PC layer is converted into a 2D image through layered projection, and SSIM is used to evaluate the quality of geometry between the L_{ref} and L_{dis}. In practice, we use IW-SSIM due to its better overall performance than SSIM [44]. Let $x_{e,v}$ and $y_{e,v}$ be the vth local image patches (extracted from the vth evaluation window) at the eth scale. E_e is the number of evaluation windows in the scale, and V is the number of scales. $g_{e,v}$ is the information content weight computed at the vth spatial location in the eth scale [44]. Then the eth scale IW-SSIM measure is computed as

$$\text{IW-SSIM}_\mathrm{e} = \frac{\sum_v g_{e,v}c(x_{e,v}, y_{e,v})s(x_{e,v}, y_{e,v})}{\sum g_{e,v}}, \tag{10}$$

where e=1,…,V-1.

$$\text{IW-SSIM}_\mathrm{e} = \frac{1}{\mathrm{E}_e}\sum_v l(x_{e,v},y_{e,v})c(x_{e,v},y_{e,v})s(x_{e,v},y_{e,v}), \tag{11}$$

for e=V. The final overall IW-SSIM measure is then computed as

$$\text{IW-SSIM} = \prod_{e=1}^{V} (\text{IW-SSIM}_e)^{\beta_e}, \tag{12}$$

where β_e is obtained in [44]. So the geometry feature of layer f_{1n} is:

$$f_{1n} = \text{IW-SSIM}(P_{ref}, P_{dis}), \tag{13}$$

where f_{1n} represents the geometry feature of the nth layer.

2) COLOR QUALITY ASSESSMENT

Only the luminance components of the images are used for IQA in IW-SSIM [44] and the human eyes perceive color more sensitively than the brightness of the image. So we need to extract color features to evaluate color quality. Meynet *et al.* [21] converted RGB to the perceived color space Lab2000HL [45], and obtained the direct relationship between each point of brightness and two chromaticity values to extract color features, but their color features were not good. Based on this, we directly use the raw information RGB of the PC without any special processing to avoid damaging data conversion so as to more accurately reflect the color quality. Similar to [21], the color value c is:

$$c = \sqrt{r^2 + g^2 + b^2}, \tag{14}$$

where r, g, and b are components of RGB. In order to compare the color values of two PCs, we define the color distance Δh as:

$$\Delta h = \sqrt{(r_{ref} - r_{dis})^2 + (g_{ref} - g_{dis})^2 - (c_{ref} - c_{dis})^2}, \tag{15}$$

where r_{ref}, g_{ref} are the two components of the RGB and c_{ref} is color values about the PC_{ref}; r_{dis}, g_{dis} are the two components of the RGB and c_{dis} is color values about the PC_{dis}. In HVS, perceiving local changes is significantly more sensitive than individual point intensity changes. Therefore, we use the method described in Section III-A3 to construct corresponding local information neighborhoods in the L_{ref} and L_{dis}. Extending the definition in [21], the color features are:

$$f_{2i} = \frac{1}{(\mu_p^L - \mu_{\hat{p}}^L)^2 + C_4}, \tag{16}$$

$$f_{3i} = \frac{2\sigma_p^L\sigma_{\hat{p}}^L + C_5}{(\sigma_p^L)^2 + (\sigma_{\hat{p}}^L)^2 + C_5}, \tag{17}$$

$$f_{4i} = \frac{\sigma_{p\hat{p}}^L + C_6}{\sigma_p^L\sigma_{\hat{p}}^L + C_6}, \tag{18}$$

$$f_{5i} = \frac{1}{C_7(\mu_p^c - \mu_{\hat{p}}^c)^2 + 1}, \tag{19}$$

$$f_{6i} = \frac{1}{C_8\bar{\Delta H}^2 + 1}, \tag{20}$$

where μ_p^L, $\mu_{\hat{p}}^L$ are the mean neighborhood gray values of point p_i, $\hat{p}_i$ respectively; σ_p^L, $\sigma_{\hat{p}}^L$ are the standard deviations of neighborhood gray values of point p_i, $\hat{p}_i$ respectively; $\sigma_{p\hat{p}}^L$ is the covariance of σ_p^L and $\sigma_{\hat{p}}^L$; μ_p^c, $\mu_{\hat{p}}^c$ are the mean neighborhood color value of point p_i and $\hat{p}_i$; $\Delta\bar{H}$ is the Gaussian weighted average of the corresponding neighborhood Δh; C_1-C_5 are the constants that guarantee the stability of the expression and are fixed as 1.

The above features $f_{2i} - f_{6i}$ are computed for each point and are in [0, 1]. We pool all points' feature values to obtain color features of layers $f_{2n} - f_{6n}$,

$$f_{jn} = \frac{\sum_{i \epsilon Q_s} f_{ji}}{S}, \quad j = 2, 3, 4, 5, 6. \tag{21}$$

where Q_s is the corresponding point set of the PC layer; s is the number of the corresponding point set; f_{ji} is the feature value of points representing the jth color feature of the ith points.

C. THE PROPOSED METRIC

The features $f_{1n} - f_{6n}$ presented above are computed for each PC layer. Therefore, we pool them to get the overall features:

$$f_j = \frac{\sum_{n=1}^{N} f_{jn}}{N}, \quad j = 1, 2, 3, 4, 5, 6, \tag{22}$$

where N is the number of layers.

It is a difficult task to assess computationally the visual quality of a 3D PC composed of both geometry and color information. In order to combine them, we use the method of weighted pooling all features for obtaining the objective score W inspired by [46]:

$$W = \sum_{j=1}^{6} w_j f_j, \tag{23}$$

where w_j is the weight, obtained according to the subsequent experiments.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

We used the following data sets for experiment [27], [47], [48].

WPC Database[1] [47]: The database consists of 20 contents, all of which are inanimate objects. It includes 740 test PCs and each PC_{ref} is processed into 5 distortion types causing by downsampling (DS), Gaussian noise contamination (GN), and three state-of-the-art PCC algorithms (S-PCC/G-PCC (Trisoup), V-PCC, and L-PCC/G-PCC (Octree)) with different levels [49].

SJTU-PCQA Database [27]: This database has 10 contents, including both human bodies and inanimate objects, and we chose 6 of them. It includes 252 test PCs, where each PC_{ref} is processed into 7 different distortion types in six levels. Specifically, the distortions are obtained with Octree-based compression (OT), Color noise (CN), Geometry Gaussian noise (GGN), Downscaling (DS), Downscaling and Color noise (D+C), Downscaling and deometry Gaussian noise (D+G), Color noise and Geometry Gaussian noise (C+G). They adopted the single stimulus method for subjective rating. MOSs are given in the range of [1,10].

ICIP2020 Database [48]: This database has six contents, including both human bodies and inanimate objects. It includes 112 test PCs, where each PC_{ref} is processed into three different distortion types at six levels. The distortions are generated by the MPEG-3DGC codecs, namely the video-based PC codec (V-PCC) and two variants of the geometric-based PC codec (G-PCC). The MOSs we used is the summary result of all labs provided by the author.

The following evaluation metrics were used to compare the performance of features or PCQA metrics. Pearson Linear correlation coefficient (PLCC) is used to assess the prediction accuracy after a nonlinear mapping between the subjective and calculated scores. The video quality experts group (VQEG) [50] recommended mapping the dynamic range of the calculated scores into a common scale using

$$q_u = k_1\left(\frac{1}{2} - \frac{1}{1 + e^{k_2(s_u - k_3)}}\right) + k_4 s_u + k_5, \tag{24}$$

where s_u is the calculated score of the uth PC_{dis}, q_u is the corresponding mapped score. k_1-k_5 are the regression model parameters to be fitted by minimizing the sum of squared differences between MOSs and calculated scores. The PLCC value can then be computed as

$$\text{PLCC} = \frac{\sum_{u=1}^{U}(q_u - \bar{q})(o_u - \bar{o})}{\sqrt{\sum_{u=1}^{U}(q_u - \bar{q})^2 (o_u - \bar{o})^2}}, \tag{25}$$

where o_u denotes the MOS of the uth PC_{dis}. $\bar{o}$ and $\bar{q}$ denote the mean values of o_u and q_u respectively. Similarly, Spearman rank-order correlation coefficient (SROCC) for prediction monotonicity is

$$\text{SROCC} = 1 - \frac{6\sum_{u=1}^{J} d_u^2}{J(J^2 - 1)}, \tag{26}$$

where d_u as difference between uth the calculated score and MOS, and J for the total number of samples. And Root mean squared error (RMSE) for prediction consistency is

$$\text{RMSE} = \sqrt{\frac{1}{J}\sum_{u=1}^{J}(q_u - \bar{q})^2}. \tag{27}$$

The higher the values of the PLCC and SROCC, the better the performance of models. On the contrary, the lower RMSE, the better performance of models.

A. PARAMETER DETERMINATION

In order to obtain the best combination of parameters, we chose the WPC database to train for its larger data set. In our training, the default number of layers N was 128 (an exponent of 2) and C_1-C_5 were set as 1 to maintain the stability of the features. A linear model was considered to

[1] https://github.com/qdushl/Waterloo-Point-Cloud-Database

TABLE 1. Performance comparison between each feature and MOS.

Feature	PLCC	SROCC	RMSE
f_1	0.62	0.61	18.05
f_2	0.60	0.59	18.28
f_3	0.49	0.46	20.03
f_4	0.30	0.10	21.88
f_5	0.58	0.56	18.74
f_6	0.01	0.23	22.92

TABLE 2. Recommended weight.

w_1	w_2	w_3
044	0.36	0.20

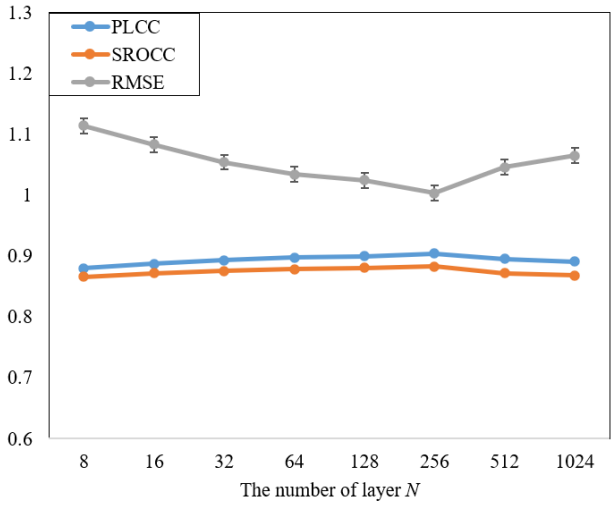


FIGURE 6. The effect of the number of layer on correlation.

determine the proportion of f_1-f_6. However, in order to prevent overfitting, it is necessary to evaluate the performance of each feature and choose the optimal features subset. Table 1 illustrates the prediction performance of each feature separately. f_1, f_2, and f_5 present the better performance compared to f_6 which seems to be less relevant as seen from Table 1. For a given subset of features f_1-f_5, the weights were optimized as follows: Features were normalized first due to their non-uniform. And then the optimal combination of f_1, f_2, and f_3 was obtained through $\alpha = C_4^2 + C_4^3 + C_4^4$ experiments using multiple linear regression analysis, and the effective weights are shown in Table 2.

The objective scores obtained by LP-PCQM are between [0,1]. The closer to 1, the less distortion of the PC. After determining the proposed metric, we tested the optimal number of layers N on the STJU-PCQA database for its various distortion types and a moderate amount of data so as to obtain test results more efficiently. Similarly, we tested the superiority of layered projection on the same database.

We increased N by an exponential power of 2 to test the optimal number of layering. In the course of the experiment, we found that the smaller N is, the higher the time complexity is. When N is less than 8, evaluating quality of a PC will need about twenty minutes, or even forty minutes. Moreover, the N is too small, and the perceived detail are less. Therefore, we did not study the influence of N less than 8 on our metric. The experimental results are shown in Fig. 6. There is not much difference between the results of N=128 and 256, but when N is the former, less time is needed, so we recommend N=128. In the following test, N was set as 128.

For the layered projection performance analysis, we validated the proposed layered-based method and compared with IW-SSIM$_{tp}$, IW-SSIM$_{lp}$ and OP-PCQM. IW-SSIM$_{tp}$ combines with IW-SSIM to evaluate PCs of quality based on six-sides overall projection. All points are orthographically projected on six planes. When multiple points are projected at the same position, the average gray value of these points is taken as the pixel value. IW-SSIM$_{lp}$ bases on the proposed layered projection. OP-PCQM is the proposed metric without layering. The results are shown in Table 3. It's clear that IW-SSIM$_{lp}$ is better than IW-SSIM$_{tp}$, indicating that the currently proposed projection has an advantage over the six-sides overall projection as seen from Fig. 3. In particular, LP-PCQM attains higher predictive power than OP-PCQM reflecting the advantages of layering. In addition, as introduced in Section III-A, we look for the maximum value of the Z-axis

TABLE 3. Performance comparison between multi-sides overall projection and layered projection.

Projection Type	PLCC	SROCC	RMSE
IW-SSIM$_{tp}$	0.79	0.78	1.43
IW-SSIM$_{lp}$	0.85	0.79	1.25
OP-PCQM	0.85	0.84	0.97
LP-PCQM	0.90	0.88	1.02

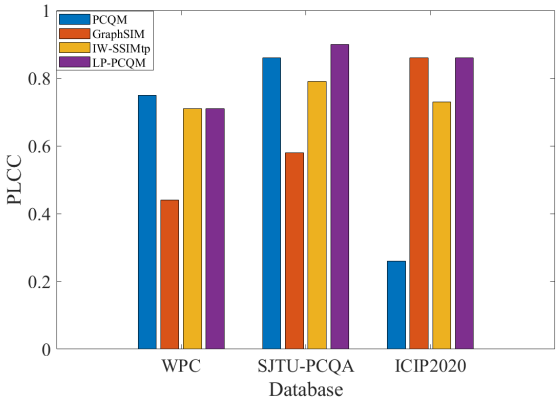
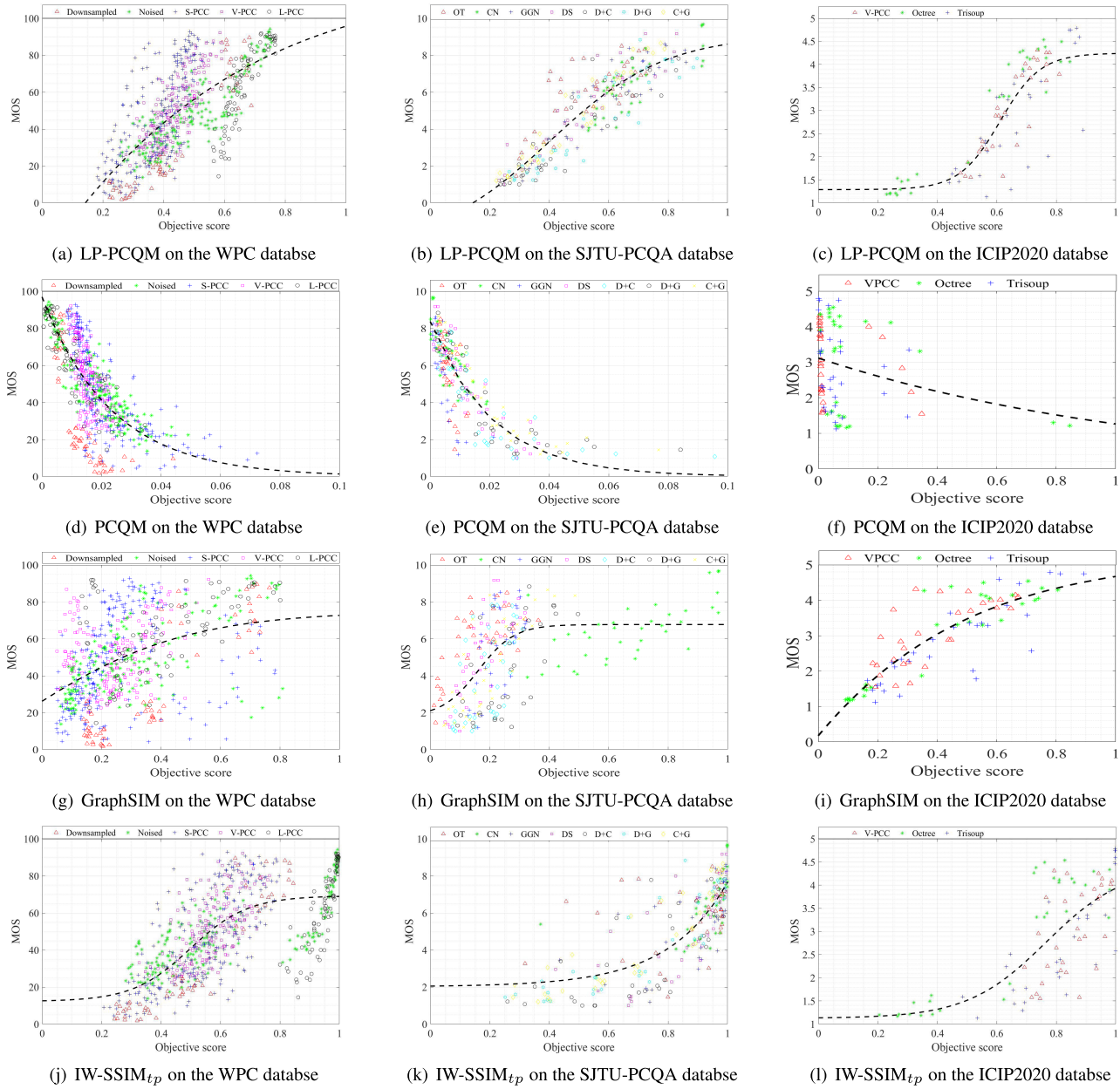


FIGURE 7. PLCC of LP-PCQM, PCQM, GraphSIM and IW-SSIM$_{tp}$.



(a) LP-PCQM on the WPC databse (b) LP-PCQM on the SJTU-PCQA databse (c) LP-PCQM on the ICIP2020 databse

(d) PCQM on the WPC databse (e) PCQM on the SJTU-PCQA databse (f) PCQM on the ICIP2020 databse

(g) GraphSIM on the WPC databse (h) GraphSIM on the SJTU-PCQA databse (i) GraphSIM on the ICIP2020 databse

(j) IW-SSIM$_{tp}$ on the WPC databse (k) IW-SSIM$_{tp}$ on the SJTU-PCQA databse (l) IW-SSIM$_{tp}$ on the ICIP2020 databse

FIGURE 8. Scatter plots of objective score vs. MOS on the WPC, SJTU-PCQA, ICIP2020 database. Also the best fitting logistic fuctions curves are shown.

to divide PCs. In fact, we can choose any perspective to divide the PC, such as X-axis, Y-axis because projection-based PCQA has rotation invariance [26].

B. PERFORMANCE EVALUATION

We validated the proposed LP-PCQM and compared them with MPEG point-based metrics, PCQM [21], PointSSIM [25], GraphSIM [22] and IW-SSIM$_{tp}$ on the WPC [47], SJTU-PCQA [27] and ICIP2020 [48] database, of which MPEG point-based metrics include PSNR$_{po2point}$MSE, PSNR$_{po2plane}$MSE, PSNR$_{po2point}$Hausdorff, PSNR$_{po2plane}$-Hausdorff obtained by two different pooling strategies, e.g., mean square error (MSE) or Hausdorff distance. The parameters of PCQM, PointSSIM and GraphSIM were set as the default value. In order to provide quantitative measures on the performance of the objective quality assessment models, we compared the predicted scores with the MOSs provided in the datasets using the PLCC, SROCC, and RMSE. The evaluation results are given in Table 4.

As reported in Table 4, LP-PCQM consistently offers the leading performance (mostly ranked at the top) in three databases, while po2point, po2plane and PointSSIM could not offer competitive performance with our LP-PCQM in predicting the subjective MOS since the performance of PointSSIM is inherently unstable [25] and MPEG metrics only considering geometrical distortion. In the following, we focus on comparing and analyzing LP-PCQM with the PCQM, GraphSIM and IW-SSIM$_{tp}$.

TABLE 4. Performance comparison of 9 metrics on 3 publicly available databases.

	WPC [47]			SJTU-PCQA [27]			ICIP2020 [48]		
Metric	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
$PSNR_{po2point}$,MSE	0.43	0.41	20.66	0.65	0.62	1.79	0.70	0.70	0.82
$PSNR_{po2plane}$,MSE	0.40	0.37	21.06	0.66	0.59	1.79	-	-	-
$PSNR_{po2point}$,Hausdorff	0.34	0.26	21.54	0.58	0.55	1.92	0.61	0.59	0.91
$PSNR_{po2plane}$,Hausdorff	0.34	0.29	21.55	0.58	0.55	1.89	-	-	-
PCQM	**0.75**	**0.75**	**15.20**	0.86	0.84	1.22	0.26	0.41	1.14
PointSSIM	0.34	0.31	21.53	0.75	0.74	1.55	0.68	0.56	0.84
GraphSIM	0.44	0.44	20.56	0.59	0.57	1.89	0.86	0.84	0.58
IW-$SSIM_{tp}$	0.71	0.72	16.20	0.79	0.78	1.43	0.73	0.69	0.78
LP-PCQM	0.71	0.72	16.12	**0.90**	**0.88**	**1.02**	**0.86**	**0.85**	**0.58**

On the SJTU-PCQA database for all impairments, LP-PCQM presents the best model performance with evaluated metrics (PLCC, SROCC, RMSE) as (0.90, 0.88, 1.02). In comparison, the remaining metrics are worse than LP-PCQM, especially GraphSIM with (0.59, 0.57, 1.89). For the WPC database, the PCQM is the best performing metric, followed by the LP-PCQM and IW-$SSIM_{tp}$ and PLCC and SROCC for these three metrics are very similar. Although LP-PCQM and IW-$SSIM_{tp}$ have the same PLCC and SROCC, the RMSE of the former is lower. Among the four metrics, GraphSIM's performance is the worst. And on the ICIP2020 database, LP-PCQM also has a good consistency, with the same PLCC and RMSE as GraphSIM, while has a better SROCC. PCQM has the lowest correlation. For a closer look at the performance of these metrics, Fig. 7 shows a graph of the PLCC for all databases. From this figure, it is clear that the performance of GraphSIM and PCQM is unstable in different databases, with drawbacks such as large performance differences and low robustness. The IW-$SSIM_{tp}$ performs consistently, but its performance is consistently not among the best. In comparison, LP-PCQM shows a more robust accuracy performance and demonstrates reliable correlations with subjective MOSs on distinct distortion types and various contents.

For better illustration, we have also provided the scatter plots along with the best fitting logistic function shown in Fig. 8 for four metrics on the WPC, SJTU-PCQA, ICIP2020 database respectively. GraphSIM, LP-PCQM and LP-PCQM are positively correlated with MOS, while PCQM is negatively correlated with MOS. We can see from Fig. 8 that scattered points predicted by LP-PCQM are more concentrated and regularly distribute around the fitting curve than other metrics, especially on the SJTU-PCQA database. Though PCQM provides better performance on the WPC database, they are not reliable and consistent, especially on the ICIP2020 database, where most predictions are away from the prediction line. For GraphSIM, it's worth noting that in Fig. 8 (g)-(i), predicted values mostly distribute disorderly around the fitting curve. As shown in Fig. 8 (j)-(l), the predicted values of IW-$SSIM_{tp}$ are less concentrated around the fitting curve and have more outliers than LP-PCQM on the WPC and ICIP2020 database respectively, providing worse performance in certain types of impairments (e.g., WPC database for ''Noised'' and 'L-PCC' distortion). In summary, for the SJTU-PCQA and ICIP2020 database, the proposed metric has better correlation values than the MPEG metrics, while for the WPC database, our metric has a (close) second-best performance.

V. CONCLUSION

In this work, we present a layered projection-based PCQA metric and use local and global statistics of the points to assess the perceived quality from both geometry and color. Our metric is obtained by the linear combination of the geometric and color features which are derived from PCs' projection information and neighborhood information. We use eight PCQA metrics to compare the performance with LP-PCQM on three public available PCQA databases. Our experimental results show that layering has a positive impact on the projection-based PCQA model, which has improved performance compared to the multi-sides overall projection-based PCQA model. Compared with the state-of-the art metrics, experimental results show that LP-PCQM has accuracy and robust performance. In summary, this paper is important for the study of projection-based PCQA models. There are still limitations in this paper. LP-PCQM has worse performance in certain types of impairments. Moreover, the point clouds evaluated by the proposed model are all distorted point clouds after voxelization, which still need to be voxelized first for 3D mesh models, which limits the application of LP-PCQM.

In future research, we will try to explore more effective and innovative point cloud quality features. In the experiments, we used PLCC, SROCC, and RMSE. But when data highly skewed or disturbance no longer be normal, we can use interquartile range(IQR). Compared with the evaluated metrics used in this paper, IQR is more suitable for the robust statistic when evaluating heavily distorted point clouds, IQR can better reflect the performance of the PCQA model.

APPENDIX

po2point	point-to-point
po2plane	point-to-plane
pl2mesh	plane-to-mesh
pl2plane	plane-to-plane
PC_{ref}	reference point cloud
PC_{dis}	distorted point cloud
L_{ref}	reference point cloud layer
L_{dis}	distorted point cloud layer
Z_{max}	the maximum value of z
d	the thickness of each layer
t	the number of layer each point belongs to
F	the fitting plane
P_{ref}	the reference point cloud layer of image
P_{dis}	the distorted point cloud layer of image
p	each point of reference point cloud
$\hat{p}$	each point of distorted point cloud
x_v	the vth pixel in the original image x
y_v	the vth pixel in the distorted image y
V	the total number of pixels in the image
L	the maximum dynamic range
μ_x	the mean of x_v
μ_y	the mean of y_v
σ_x	the standard deviation of x_v
σ_y	the standard deviation of y_v
σ_{xy}	the covariance of x_v and y_v
C_1-C_8	the constant used to maintain stability
$x_{e,v}$, $y_{e,v}$	the vth local image patches at the eth scale
E_e	the number of evaluation windows in the scale
V	the number of scales
$g_{e,v}$	the information content weight computed at the vth spatial location in the eth scale
f_{1n}-f_{6n}	the features used to evaluate the quality of point cloud layer
f_1-f_6	the features used to evaluate the quality of point cloud
f_{2i}-f_{6i}	the features value of each point
c	the color value
c_{ref}	the color value of reference point cloud
c_{dis}	the color value of distorted point cloud
Δh	the color distance
r_{ref}, g_{ref}	the color distance about reference point cloud
r_{dis}, g_{dis}	the color distance about distorted point cloud
p_i	the ith point in the reference point cloud
$\hat{p}_i$	the ith point in the distorted point cloud
μ_p^L, $\mu_{\hat{p}}^L$	the mean neighborhood gray values of point p_i, $\hat{p}_i$
σ_p^L, $\sigma_{\hat{p}}^L$	the standard deviations of neighborhood gray values of point p_i, $\hat{p}_i$
$\sigma_{p\hat{p}}^L$	the covariance of σ_p^L and $\sigma_{\hat{p}}^L$
μ_p^c, $\mu_{\hat{p}}^c$	the mean neighborhood color values of point p_i, $\hat{p}_i$
$\bar{\Delta H}$	the Gaussian weighted average of the corresponding neighborhood Δh
W	the objective score
w_j	the jth weight
q_u	the corresponding mapped score
s_u	the calculated score of the uth PC_{dis}
k_1-k_5	the regression model parameters to be fitted by minimizing the sum of squared differences between MOSs and calculated scores
o_u	the MOS of the uth PC_{dis}
$\bar{o}$, $\bar{q}$	the mean values of o_u and q_u respectively.
d_u	the difference between uth the calculated score and MOS
J	the total number of samples

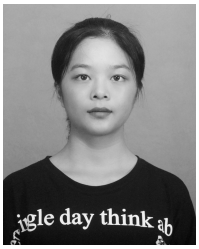
REFERENCES

[1] P. Luo, X. Lu, and L. Zhong, "Research progress and application of digital holography," *Laser Mag.*, no. 6, pp. 8–10, 2006.
[2] L. Si, G. Jiang, X. Hu, and B. Liu, "Retina 3D perception reconstruction algorithm based on visual light field image," *IEEE Access*, vol. 8, pp. 196804–196812, 2020.
[3] Y. Peng, M. Chang, Q. Wang, Y. Qian, Y. Zhang, M. Wei, and X. Liao, "Sparse-to-dense multi-encoder shape completion of unstructured point cloud," *IEEE Access*, vol. 8, pp. 30969–30978, 2020.
[4] E. Alexiou and T. Ebrahimi, "Benchmarking of objective quality metrics for colorless point clouds," in *Proc. Picture Coding Symp. (PCS)*, Jun. 2018, pp. 51–55.
[5] J. Yang, J. Luo, D. Meng, and J.-N. Hwang, "QoE-driven resource allocation optimized for uplink delivery of delay-sensitive VR video over cellular network," *IEEE Access*, vol. 7, pp. 60672–60683, 2019.
[6] A. Rohan, M. Rabah, and S.-H. Kim, "Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2," *IEEE Access*, vol. 7, pp. 69575–69584, 2019.
[7] Q. Liu, H. Yuan, J. Hou, H. Liu, and R. Hamzaoui, "Model-based encoding parameter optimization for 3D point cloud compression," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2018, pp. 1981–1986.
[8] Q. Liu, H. Yuan, R. Hamzaoui, and H. Su, "Coarse to fine rate control for region-based 3D point cloud compression," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2020, pp. 1–6.
[9] Q. Liu, H. Yuan, J. Hou, R. Hamzaoui, and H. Su, "Model-based joint bit allocation between geometry and color for video-based 3D point cloud compression," *IEEE Trans. Multimedia*, early access, Sep. 10, 2020, doi: 10.1109/TMM.2020.3023294.
[10] K. Sato, R. Shinkuma, T. Sato, E. Oki, T. Iwai, D. Kanetomo, and K. Satoda, "Prioritized transmission control of point cloud data obtained by LIDAR devices," *IEEE Access*, vol. 8, pp. 113779–113789, 2020.
[11] B. Li, Y. Zhang, B. Zhao, and H. Shao, "3D-ReConstnet: A single-view 3D-object point cloud reconstruction network," *IEEE Access*, vol. 8, pp. 83782–83790, 2020.
[12] C. Suo, Z. Liu, L. Mo, and Y. Liu, "LPD-AE: Latent space representation of large-scale 3D point cloud," *IEEE Access*, vol. 8, pp. 108402–108417, 2020.
[13] M.-J. Chen, D.-K. Kwon, L. K. Cormack, and A. C. Bovik, "Optimizing 3D image display using the stereoacuity function," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep. 2012, pp. 617–620.
[14] W. Hu, X. Gao, G. Cheung, and Z. Guo, "Feature graph learning for 3D point cloud denoising," *IEEE Trans. Signal Process.*, vol. 68, pp. 2841–2856, 2020.
[15] S. Shi, X. Wang, and H. Li, "PointRCNN: 3D object proposal generation and detection from point cloud," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 770–779.

[16] G. Zhai and X. Min, "Perceptual image quality assessment: A survey," *Sci. China Inf. Sci.*, vol. 63, no. 11, pp. 1–52, Nov. 2020.
[17] 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, pp. 1–27, Nov. 2019.
[18] 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.
[19] 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.
[20] G. Meynet, J. Digne, and G. Lavoué, "PC-MSDM: A quality metric for 3D point clouds," in *Proc. 11th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, Jun. 2019, pp. 1–3.
[21] 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. Qual. Multimedia Exper. (QoMEX)*, May 2020, pp. 1–6.
[22] Q. Yang, Z. Ma, Y. Xu, Z. Li, and J. Sun, "Inferring point cloud quality via graph similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Dec. 24, 2021, doi: 10.1109/TPAMI.2020.3047083.
[23] Y. Liu, Q. Yang, Y. Xu, and L. Yang, "Point cloud quality assessment: Large-scale dataset construction and learning-based no-reference approach," Dec. 2020, *arXiv:2012.11895*. [Online]. Available: https://arxiv.org/abs/2012.11895
[24] I. Viola, S. Subramanyam, and P. Cesar, "A color-based objective quality metric for point cloud contents," in *Proc. 12th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, May 2020, pp. 1–6.
[25] 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.
[26] E. Alexiou and T. Ebrahimi, "Exploiting user interactivity in quality assessment of point cloud imaging," in *Proc. 11th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, Jun. 2019, pp. 1–6.
[27] 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*, early access, Oct. 23, 2020, doi: 10.1109/TMM.2020.3033117.
[28] E. M. Torlig, E. Alexiou, T. A. D. Fonseca, R. L. D. Queiroz, and T. Ebrahimi, "A novel methodology for quality assessment of voxelized point clouds," *Proc. SPIE Opt. Eng. Appl.*, vol. 10752, Sep. 2018, Art. no. 107520I.
[29] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, and H. Yang, "Reduced reference perceptual quality model and application to rate control for 3D point cloud compression," 2020, *arXiv:2011.12688*. [Online]. Available: http://arxiv.org/abs/2011.12688
[30] 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.
[31] G. Lavoué, "A multiscale metric for 3D mesh visual quality assessment," in *Computer Graphics Forum*. Oxford, U.K.: Blackwell, 2011, pp. 1427–1437.
[32] 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.
[33] G. Lavoué, E. D. Gelasca, F. Dupont, A. Baskurt, and T. Ebrahimi, "Perceptually driven 3D distance metrics with application to watermarking," in *Applications of Digital Image Processing XXIX*, vol. 6312. San Diego, CA, USA: International Society for Optics and Photonics, 2006, Art. no. 63120L.
[34] M. Zhang, H. You, P. Kadam, S. Liu, and C.-C.-J. Kuo, "PointHop: An explainable machine learning method for point cloud classification," *IEEE Trans. Multimedia*, vol. 22, no. 7, pp. 1744–1755, Jul. 2020.
[35] 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. 77–85.
[36] X. Tong, Y. Xu, Z. Ye, S. Liu, L. Li, H. Xie, F. Wang, S. Gao, and U. Stilla, "An improved phase correlation method based on 2-D plane fitting and the maximum kernel density estimator," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1953–1957, Sep. 2015.
[37] Y. Cheng, X. Li, and P. Li, "Extraction of point cloud data features of different vegetation morphology by K-neighborhood method," *Jiangxi Sci.*, vol. 34, no. 6, pp. 875–878, 2016.
[38] W. Xue, X. Mou, L. Zhang, and X. Feng, "Perceptual fidelity aware mean squared error," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 705–712.
[39] M. A. Baig, A. A. Moinuddin, and E. Khan, "PSNR of highest distortion region: An effective image quality assessment method," in *Proc. Int. Conf. Electr., Electron. Comput. Eng. (UPCON)*, Nov. 2019, pp. 1–4.
[40] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 2366–2369.
[41] P. Gupta, P. Srivastava, S. Bhardwaj, and V. Bhateja, "A modified PSNR metric based on HVS for quality assessment of color images," in *Proc. Int. Conf. Commun. Ind. Appl.*, Dec. 2011, pp. 1–4.
[42] Y. Niu, Y. Zhong, W. Guo, Y. Shi, and P. Chen, "2D and 3D image quality assessment: A survey of metrics and challenges," *IEEE Access*, vol. 7, pp. 782–801, 2019.
[43] X. Zhang, "Efficient compression algorithm and quality evaluation of 3D color point cloud," Ph.D. dissertation, College Commun. Inf. Eng., Shanghai Univ., Shanghai, China, 2017.
[44] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.
[45] I. Lissner and P. Urban, "Toward a unified color space for perception-based image processing," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1153–1168, Mar. 2012.
[46] W. Yuanji, L. Jianhua, L. Yi, F. Yao, and J. Qinzhong, "Image quality evaluation based on image weighted separating block peak signal to noise ratio," in *Proc. Int. Conf. Neural Netw. Signal Process.*, vol. 2, 2003, pp. 994–997.
[47] H. Su, Z. Duanmu, W. Liu, Q. Liu, and Z. Wang, "Perceptual quality assessment of 3D point clouds," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 3182–3186.
[48] S. Perry, H. P. Cong, L. A. da Silva Cruz, J. Prazeres, M. Pereira, A. Pinheiro, E. Dumic, E. Alexiou, and T. Ebrahimi, "Quality evaluation of static point clouds encoded using MPEG codecs," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 3428–3432.
[49] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuca, S. Lasserre, Z. Li, J. Llach, K. Mammou, R. Mekuria, O. Nakagami, E. Siahaan, A. Tabatabai, A. M. Tourapis, and V. Zakharchenko, "Emerging MPEG standards for point cloud compression," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 133–148, Mar. 2019.
[50] 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.

TIANXIN CHEN was born in Binzhou, Shandong, China, in 1999. He is currently pursuing the B.S. degree in electronic engineering with the University of Qingdao, Qingdao.

His research interest includes immersive media quality assessment.



CHUNYI LONG was born in Yuli, Guangxi, China, in 1998. She is currently pursuing the B.S. degree in electronic engineering with the University of Qingdao, Qingdao.

Her research interest includes immersive media quality assessment.

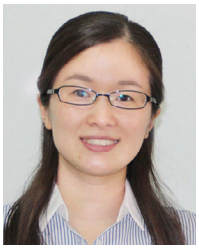
HONGLEI SU received the B.A.Sc. degree from the Shandong University of Science and Technology, Qingdao, China, in 2008, and the Ph.D. degree from Xidian University, Xi'an, China, in 2014. Since September 2014, he has been working as an Assistant Professor with the School of Electronic Information, Qingdao University, Qingdao. From March 2018 to March 2019, he worked as a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests include perceptual image processing, immersive media processing, and computer vision.

LIJUN CHEN was born in Qujing, Yunnan, China, in 2000. She is currently pursuing the bachelor's degree.

Her research interest includes immersive media quality assessment.

JIERU CHI received the B.Sc. degree in automation and the M.Sc. degree in control science and engineering from Shandong University, Jinan, China, in 1992 and 1995, respectively, and the Ph.D. degree from the Institute of Complexity Science, Qingdao University, Qingdao, China. She is currently an Associate Professor with the School of Automation and Electrical Engineering, Qingdao University. Her current research interests include motor control and robot control, applied nonlinear control, and intelligent systems.

ZHENKUAN PAN received the Ph.D. degree from the Shanghai Jiao Tong University of Science and Technology, in 1992. Since 1996, he has been a Full Professor with Qingdao University. He is currently working on the application of multibody system dynamics and control. He is the author of more than 300 articles. His interests include computer vision, image processing, and pattern recognition. He is a member of the Virtual Reality Professional Committee of China Graphic Image Association.



HUAN YANG received the B.S. degree in computer science from the Heilongjiang Institute of Technology, China, in 2007, the M.S. degree in computer science from Shandong University, China, in 2010, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore, in 2015. She is currently working with the College of Computer Science and Technology, Qingdao University, Qingdao, China. Her research interests include image/video processing and analysis, perception-based modeling and quality assessment, object detection/recognition, and machine learning.

YUXIN LIU was born in Zhoukou, Henan, China, in 1997. He is currently pursuing the master's degree in electronic engineering with the University of Qingdao, Qingdao.

His research interest includes immersive media quality assessment.