CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/168856/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Zhou, Wei, Yang, Qi, Chen, Wu, Jiang, Qiuping, Zhai, Guangtao and Lin, Weisi 2024. Blind quality assessment of dense 3D point clouds with structure guided resampling. ACM Transactions on Multimedia Computing, Communications and Applications 20 (8) 10.1145/3664199

Publishers page: http://dx.doi.org/10.1145/3664199

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Blind Quality Assessment of Dense 3D Point Clouds with Structure Guided Resampling

WEI ZHOU, School of Computer Science and Informatics, Cardiff University
QI YANG, Tencent MediaLab
WU CHEN AND QIUPING JIANG, School of Information Science and Engineering, Ningbo University
GUANGTAO ZHAI, Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University
WEISI LIN, School of Computer Science and Engineering, Nanyang Technological University

Objective quality assessment of 3D point clouds is essential for the development of immersive multimedia systems in real-world applications. Despite the success of perceptual quality evaluation for 2D images and videos, blind/no-reference metrics are still scarce for 3D point clouds with large-scale irregularly distributed 3D points. Therefore, in this paper, we propose an objective point cloud quality index with Structure Guided Resampling (SGR) to automatically evaluate the perceptually visual quality of dense 3D point clouds. The proposed SGR is a general-purpose blind quality assessment method without the assistance of any reference information. Specifically, considering that the human visual system (HVS) is highly sensitive to structure information, we first exploit the unique normal vectors of point clouds to execute regional pre-processing which consists of keypoint resampling and local region construction. Then, we extract three groups of quality-related features, including: 1) geometry density features; 2) color naturalness features; 3) angular consistency features. Both the cognitive peculiarities of the human brain and naturalness regularity are involved in the designed quality-aware features that can capture the most vital aspects of distorted 3D point clouds. Extensive experiments on several publicly available subjective point cloud quality databases validate that our proposed SGR can compete with state-of-the-art full-reference, reduced-reference, and no-reference quality assessment algorithms.

CCS Concepts: • **Computing methodologies → Image processing**.

Additional Key Words and Phrases: 3D point clouds, blind/no-reference, perceptual quality assessment, structure information, naturalness regularity, human visual system

ACM Reference Format:
Wei Zhou, Qi Yang, Wu Chen and Qiuping Jiang, Guangtao Zhai, and Weisi Lin. 2024. Blind Quality Assessment of Dense 3D Point Clouds with Structure Guided Resampling. In . ACM, New York, NY, USA, 21 pages.

This work was supported in part by the Natural Science Foundation of China (62271277), the Natural Science Foundation of Zhejiang (LR22F020002), the Natural Science Foundation of Ningbo (2022J081), the "Leading Goose" RD Program of Zhejiang Province under Grant No. 2024C01132, and Cardiff-DUT Collaboration Fund (AJ28101015). Corresponding author: Qiuping Jiang.
Authors' addresses: Wei Zhou is with the School of Computer Science and Informatics, Cardiff University, CF24 4AG, Cardiff, UK (email: zhouw26@cardiff.ac.uk); Qi Yang is with the Tencent MediaLab, Shanghai 200030, China (e-mail: chinoyang@tencent.com); Wu Chen and Qiuping Jiang are with the School of Information Science and Engineering, Ningbo University, Ningbo 315211, China (e-mail: jiangqiuping@nbu.edu.cn); Guangtao Zhai is with the Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zhaiguangtao@sjtu.edu.cn); Weisi Lin is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: wslin@ntu.edu.sg).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
Manuscript submitted to ACM

1 INTRODUCTION

With the modern advances in 3D data capture devices and rendering technologies, popular 3D point clouds have become one of the most important multimedia representations for providing 6 degrees of freedom [20, 48, 55]. As defined, a 3D point cloud contains a huge number of scattered points with certain attributes. Each 3D point owns geometry and attribute information, which can represent actual objects or environments more visually compared with traditional images and videos. To be specific, geometry stands for the 3D coordinates of the point, while the attribute is composed of additional descriptors such as RGB color, surface normal, reflectance and opacity, etc. Due to such an abundant pattern, there have emerged lots of real-world applications for 3D point clouds, including VR/AR/XR [51], automatic driving [59], scene understanding [8], among many others. Typically, 3D point clouds can be divided as two types: relatively smaller objects and large-scale scenes. In this work, we focus on the first type of dense 3D point clouds.

Although the data formats of 3D point clouds are very different, similar to conventional images and videos, a variety of distortions would be inevitably introduced during the processing chain of multimedia communication systems [21, 46, 50]. For example, the acquisition sensors may produce noise artifacts in captured 3D point clouds. Moreover, because of the large-scale volumetric content of 3D point clouds, executive downsampling and compression operations [39] can also lead to perceptual quality degradation. Consequently, how to effectively assess the perceptually visual quality of 3D point clouds is a significant and challenging problem, which can benefit other relevant processing stages, e.g., tuning the acquisition parameters for the optimal capture of 3D point clouds. Regarding to the quality assessment of 3D point clouds, a subjective test usually serves as the most accurate and reliable method [3, 18, 61]. However, since organizing subjective experiments are time-consuming and labor-intensive, developing efficient objective quality metrics is a promising alternative [60].

In general, the characteristics of point clouds make the design of objective quality assessment models more challenging than traditional 2D images or videos. Ideally, both the 2D and point cloud formats can represent a scene or an object in real-world. However, a 2D image is usually displayed in a regular shape, while a 3D point cloud is unstructured and composed of many distributed points. Thus, new challenges would arise for the objective quality assessment of 3D point clouds.

According to the availability of original reference 3D point clouds, the objective quality assessment methods of 3D point clouds can be divided into three categories, consisting of full-reference (FR) [27], reduced-reference (RR) [67], and no-reference (NR) [25] models. That is, the FR and RR models can access full and part data of reference point clouds, respectively. In contrast, the blind/NR models evaluate the perceptual quality of distorted 3D point clouds without any reference information. Among the NR models, general-purpose ones can evaluate distortion-generic 3D point clouds, which do not need to know about the distortion type and are more practical in real-world scenarios. Existing general-purpose NR quality assessment methods for 3D point clouds have been proposed mainly on the basis of deep neural networks (DNNs) [24, 25, 57]. By leveraging the remarkable learning ability of DNNs, perceptual quality can be directly predicted from the point data. However, the learning process often lacks the design philosophies of domain knowledge that reflect the human visual system (HVS) characteristics, causing unsatisfactory interpretability. In [63], a statistics-based quality assessment model for point clouds and meshes was proposed. To the best of our knowledge, so far there has been no exploration of explainable general-purpose NR quality assessment methods specifically designed for 3D point clouds based on domain knowledge.

As illustrated in Fig. 1, our main motivation comes from the observations of typical artifact appearances in various distorted 3D point clouds. From (b) to (d), we can find that the perceptual quality degradations are relevant to geometry

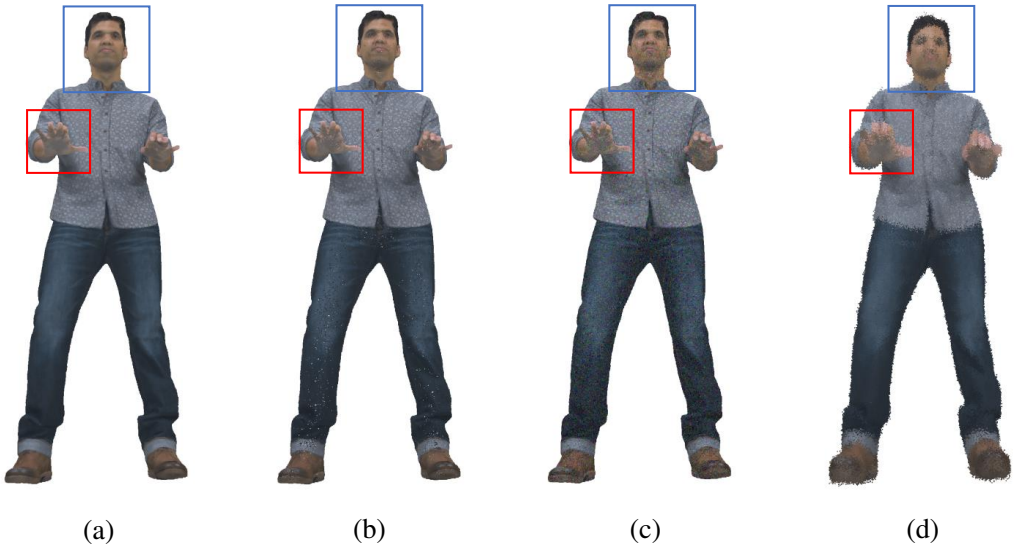


Fig. 1. Typical distortions in 3D point clouds. (a) A reference point cloud, (b)-(d) examples of distorted point clouds.

density, color difference and point orientation, respectively. For example, compared with the original reference point cloud in (a), the points are incomplete and have incorrect color in (b) and (c), respectively. Moreover, in (d), the orientations of these points are destroyed, as clearly observed by the object parts in the bounding boxes. Therefore, we aim to expose the black box of point cloud quality prediction by quantifying the distortions from the aspects of density, color and orientation, leading to the general-purpose NR point cloud quality index with Structure Guided Resampling (SGR).

In our proposed SGR, to extract efficient HVS-related features based on scattered points, regional pre-processing is first conducted, where keypoint resampling and local region construction are involved. In this pre-processing step, considering the huge amount of point clouds data and the crucial structure information, we first resample the test point cloud via normal vectors to obtain a serial of keypoints, and then construct local regions centered keypoints. Afterwards, inspired by the combined impacts of 3D geometry and associated attributes, 3 groups of quality-aware features including 1) geometry density features, 2) color naturalness features, and 3) angular consistency features, are employed to predict the perceptual quality of distorted 3D point clouds. Besides, we also consider the fundamental theory of natural scene statistics (NSS) [37, 43] in the extracted features. Finally, the quality-related features are fused into the overall SGR via a quality regression module. The effectiveness of our proposed SGR is verified on four subjective point cloud quality databases. The experiments demonstrate that the proposed SGR presents significant performance improvement compared with other NR metrics, and even better than most FR metrics. Besides, the proposed SGR shows stable performance with typical point cloud distortion types, such as Downsampling, Gaussian Noise, geometry-based point cloud compression (G-PCC), and video-based point cloud compression (V-PCC). The source codes are publicly available at https://github.com/weizhou-geek/SGR.

The main contributions of this paper are summarized as follows:

- We propose a general-purpose blind objective quality assessment algorithm for 3D point clouds based on the cognitive characteristics of the human brain and the regularity of NSS.
- According to the structural dependence of the HVS, the regional pre-processing is proposed in the SGR framework, which exploits the unique normal vectors of 3D point clouds.
- The integrated influence of 3D geometry and attributes information can be reflected by the proposed quality-aware features, which is different from conventional 2D quality assessment. Experimental results on existing subjective point cloud quality databases demonstrate the competitive performance of the proposed model.

The rest of this paper is organized as follows. Section II reviews related work on objective quality assessment of 3D point clouds. In Section III, we introduce the technical details of our proposed NR quality assessment method. Section IV presents the validation results of the proposed model. Finally, we conclude the paper in Section V.

2 RELATED WORK

To assess the perceptual quality of distorted 3D point clouds, comparing with the corresponding original reference point clouds is the most intuitive way if the pristine one is available. Therefore, there have emerged many FR quality assessment methods for 3D point clouds, which can be generally classified into point-based and projection-based metrics.

The point-based models directly use the 3D points in the reference and distorted content. Among them, the earliest point-based metric is point-to-point (p2po) [29], where representative mean squared error (MSE) or Hausdorff (HF) distance is often adopted for geometric peak signal-to-noise ratio (PSNR) computation. An alternative method is point-to-plane (p2pl) [47] that computes the PSNR between a test point and a corresponding normal vector in the reference point cloud. Except for the point-based geometry distortion measures, color information can also be used to compute the PSNR, such as the YUV channels or the luminance Y component [28]. But both p2po and p2pl quality metrics cannot precisely evaluate the perceptual quality of 3D point clouds under structural loss. Thus, the angular similarity (AS), also known as plane-to-plane, was proposed [1]. Moreover, on the basis of graph signal processing, Yang et al. [58] proposed graph similarity index (GraphSIM) and the extended multiscale GraphSIM (MS-GraphSIM) [62]. In addition, based on the designed mesh structural distortion measure (MSDM) [19], Meynet et al. [30] exploited local curvature statistics to develop the FR quality assessment method called PC-MSDM for the perceptual quality evaluation of 3D point clouds. They further proposed the point cloud quality metric (PCQM) [31] which utilizes the optimally weighted linear combination of geometric curvature and color features. Furthermore, similar to the well-known structural similarity index (SSIM) [52], Alexiou et al. [2] tried to capture the local changes of test point clouds, leading to the quality degradation measurement named PointSSIM. Diniz et al. [13] proposed the BitDance by using local and global statistics to evaluate the visual quality of point clouds. Lu et al. [26] presented the 3D edge similarity-based model (TDESM) for point cloud quality assessment.

Apart from the previously mentioned point-based FR methods, another kind of mainstream FR framework is to project 3D point clouds into multiple 2D images from various views and then conduct the 2D image quality assessment, namely projection-based models. For example, the SSIM [52] and its variant multiscale SSIM (MS-SSIM) [54] as well as the visual information fidelity in the pixel domain (VIFP) [41] are often employed to predict the visual quality of converted 2D images. Several projection methods can be adopted, where 6 perpendicular projection of a cube is one of typical ways [56]. Usually, different weights are allocated for extracted features from all projection planes.

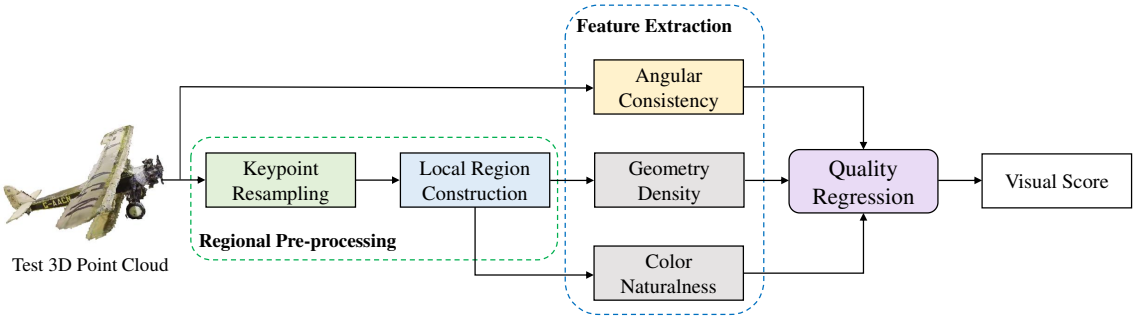


Fig. 2. Framework of the proposed SGR method which consists of regional pre-processing, quality-aware feature extraction, and quality regression. The regional pre-processing involves keypoint resampling and local region construction. For final visual score prediction, geometry density, color naturalness and angular consistency features are used to quantify the perceptual quality of 3D point clouds.

Since the full information of associated reference point clouds do not always exist, some RR point cloud quality assessment methods have emerged. For example, Viola et al. [49] proposed a RR quality metric for point cloud content (i.e., PCMRR), where a small set of features from the references are extracted and then delivered to the receiver side for evaluating the quality degradation of distorted point clouds. In [23], a sets of quantization parameters was utilized to estimate the visual quality of V-PCC compressed point clouds. In addition, Zhou et al. [67] presented a projection-based RR model by content-oriented saliency detection.

Generally, in practical scenarios, we cannot access any reference data. Thus, researchers resort to explore the challenging NR quality assessment methods for 3D point clouds. To the best of our knowledge, existing NR quality metrics specifically designed for point clouds are developed mainly with the assistance of DNNs. Liu et al. [24] designed a point cloud quality assessment network (PQA-Net) which consists of multi-view-based joint feature extraction and fusion, distortion type identification, and final quality prediction. In [57], the unsupervised NR quality framework called image transferred point cloud quality assessment (IT-PCQA) was proposed, where the adversarial domain adaptation was realized. By focusing on geometric distortion, Su et al. [45] presented the PRL-GQA according to pairwise rank learning. In [25], the ResSCNN was developed on the basis of sparse convolutional neural network. More recently, Zhang et al. [64] developed the NR point cloud quality assessment model called MM-PCQA, which is based on multi-modal learning. Bourbia et al. [5] proposed to use multi-view projection and DNNs for NR point cloud quality prediction.

However, the existing NR models are generally data-driven and uninterpretable, little consideration is given to design domain knowledge oriented perceptual features that reveal the characteristics of the HVS. Therefore, in this work, we aim to propose an efficient NR point cloud quality assessment method according to the HVS perception. Considering the structure information of scattered points and the great success of NSS regularity in 2D image quality prediction, we are the first to propose a blind/NR quality assessment method specifically designed for 3D point clouds based on well-defined quality-aware features upon regional pre-processing and the cognitive properties of the human brain.

3 PROPOSED POINT CLOUD QUALITY INDEX

The framework of our proposed SGR method is shown in Fig. 2, which is a general-purpose blind quality assessment model. Since the HVS is more sensitive to structures, we first resample keypoints from the input test 3D point cloud

based on generated normal vectors. Then, we construct local regions for the resampled points and extract distortion-related features. Finally, the quality regression module is used to map the extracted features onto visual scores. We will introduce the technical details of SGR in the following subsections.

3.1 Regional Pre-processing

Keypoint resampling. Suppose that a distorted 3D point cloud $\mathbf{P}$ contains M points. Usually, each point has three geometry coordinates (e.g., x, y, z) as well as three color attributes (e.g., r, g, b). We can denote such kind of 3D point cloud data as follows:

$$\mathbf{P} = \left[\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{\mathrm{M}}\right]^T \in \mathbb{R}^{\mathrm{M}\times 6}, \tag{1}$$

where $\mathbf{p}_m = (x_m, y_m, z_m, r_m, g_m, b_m)$ is a point in the point cloud. In this representation, the geometry coordinates and color attributes can be separated as $\mathbf{g}_m = (x_m, y_m, z_m)$ and $\mathbf{c}_m = (r_m, g_m, b_m)$, respectively. In addition, another attribute called normal vectors can be estimated from the point cloud.

Generally, each 3D point cloud data has numerous points. Moreover, these 3D points are irregular and unstructured. Nevertheless, the HVS inclines to perceive the structures of objects or environments. For traditional images and videos, such structure information has been widely used to design many objective quality measures [52–54]. Inspired by GraphSIM [58], we opt to extract the keypoints by graph-based resampling to realize the structure extraction of 3D point clouds. Specifically, considering that the resampled points should be high-frequency parts such as edges, contours, etc., GraphSIM uses a Haar-like high pass filter to extract point cloud skeleton based on original geometry coordinates. However, this method is too sensitive to the outlier, which may lead to unstable results. Therefore, we first compute the normal vectors and then treat them as the graph signal to guide the keypoint resampling process.

As defined, a graph filter represents a signal system that inputs a graph signal and outputs a tuned graph signal [9]. Mathematically, a linear and shift-invariant graph filter is a polynomial conversion of the graph shift operator $\mathcal{A} \in \mathbb{R}^{\mathrm{M}\times\mathrm{M}}$ computed by:

$$\mathrm{h}(\mathcal{A}) = \sum_{\mathrm{k}=0}^{\mathrm{K}-1} \mathrm{h}_{\mathrm{k}}\mathcal{A}^{\mathrm{k}} = \mathrm{h}_0\mathbf{I} + \mathrm{h}_1\mathcal{A} + \ldots + \mathrm{h}_{\mathrm{K}-1}\mathcal{A}^{\mathrm{K}-1}, \tag{2}$$

where h_{k} denotes the k − th filter coefficients and K is the length of the graph filter. $\mathbf{I}$ represents the identity matrix. With the predefined graph filter, for an input graph signal $\mathbf{\Phi} \in \mathbb{R}^{\mathrm{M}\times 1}$, the filtered output graph signal is a matrix-vector product which can be converted by eigendecomposition as:

$$\mathrm{h}(\mathcal{A})\Phi = \mathbf{V}\mathrm{h}(\mathbf{\Lambda})\mathbf{V}^{-1}\mathbf{\Phi}, \tag{3}$$

where the eigenvectors of $\mathcal{A}$ constitute the columns of matrix $\mathbf{V}$. The eigenvalue matrix $\mathbf{\Lambda} \in \mathbb{R}^{\mathrm{M}\times\mathrm{M}}$ is the diagonal matrix containing the corresponding eigenvalues of $\mathcal{A}$. Here, the eigenvalues are frequencies on the formed graph, which can be used to sort the 3D points.

Given a distorted 3D point cloud $\mathbf{P}$, from the above analysis, we resample it to extract graph keypoints by:

$$\hat{\mathbf{P}} = \mathrm{h}_{\mathrm{H}}(\mathbf{P}, \mathbf{\Phi}, \mathrm{K})_{\Theta}, \Theta << \mathrm{M}, \tag{4}$$

where $\hat{\mathbf{P}} \in \mathbb{R}^{\Theta\times 6}$, and $\mathrm{h}_{\mathrm{H}}(\cdot)$ is a high-pass graph filter with length equaling to K, and the graph signal $\mathbf{\Phi}$ is normal vector. Moreover, Θ represents the number of keypoints after the resampling operation. Similar to [62], we set K and Θ as 4 and $\mathrm{M}/10,000$, respectively.

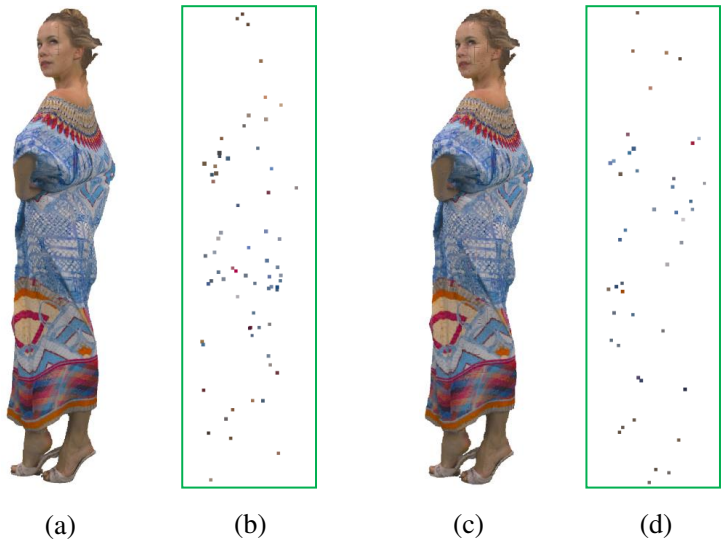


Fig. 3. Illustration of keypoint resampling for 3D point clouds. (a) and (c) are with different Octree-based compression levels. (b) and (d) are the corresponding keypoints resampled from (a) and (c), respectively. (a) MOS=8.5013, (b) the resampled keypoint number of (a) is 68, (c) MOS=6.3047, (d) the resampled keypoint number of (c) is 45.

As shown in Fig. 3, we give examples of the keypoint resampling results for 3D point clouds which are distorted by different Octree-based compression levels. Compared (c) to (a), the distortion level increases, which can be easily observed by the appearances of distorted point clouds. Moreover, the MOS values are 8.5013 and 6.3047 for (a) and (c), which means the visual quality of (c) is worse than that of (a). Meanwhile, we can see that the keypoints generally represent the shapes/structures of point clouds. The location of resampled keypoints is influenced by the distortion. In addition, with the increase of distortion levels, the number of resampled keypoints decreases from 68 to 45. Therefore, the keypoint resampling can be a promising quality indicator for subsequent local region construction and quality-aware feature extraction. This mainly benefit from the efficient structure information of normal attribute.

Local region construction. After the process of keypoint resampling, we exploit the generated keypoints to construct local regions. For each keypoint $\hat{\mathbf{p}_\theta}$ in $\hat{\mathbf{P}}$, we cluster its neighbors based on the Euclidean distance of the corresponding geometry coordinates as follows:

$$\tilde{\mathbf{P}} \subset \hat{\mathbf{P}}, \|\tilde{\mathbf{g}} - \hat{\mathbf{g}_\theta}\|_2^2 \leq \alpha, \tag{5}$$

where $\tilde{\mathbf{P}}$ denotes the groups of clustered neighbors of the keypoint $\hat{\mathbf{p}_\theta}$. $\tilde{\mathbf{g}}$ and $\hat{\mathbf{g}_\theta}$ are the geometry components of $\tilde{\mathbf{P}}$ and $\hat{\mathbf{p}_\theta}$, respectively. In addition, α is used to cluster neighbors, which is $1/20$ of minimum range among x, y, and z coordinates.

In Fig. 4, we show the structure guided local region construction of distorted 3D point cloud after keypoint resampling. Here, we take (a) in Fig. 3 as an example, where three constructed regions are given. It can be found that these local regions have different shapes, which are used for the quality-aware feature extraction.

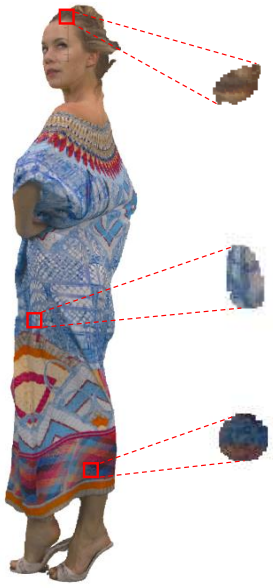


Fig. 4. Structure guided local region construction of a distorted 3D point cloud after keypoint resampling.

3.2 Geometry Density

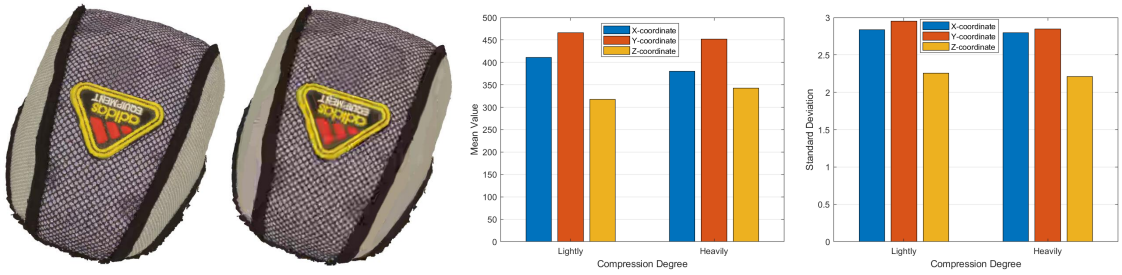
Based on the constructed local region, we exploit the geometry components to derivate geometry density features. To be specific, since the geometry information can be directly obtained from the local region, we separately compute the mean and standard deviation values of all three coordinates to serve as the geometry density features. For example, the mean and standard deviation of geometry information along x-coordinate are calculated by:

$$\tau_{\mathbf{g}}^{\mathbf{x}} = \frac{\sum_{i=1}^{N} \tilde{\mathbf{g}}_i^{\mathbf{x}}}{N}, \tag{6}$$

$$\xi_{\mathbf{g}}^{\mathbf{x}} = \sqrt{\frac{\sum_{i=1}^{N} \left(\tilde{\mathbf{g}}_i^{\mathbf{x}} - \tau_{\mathbf{g}}^{\mathbf{x}}\right)^2}{I}}, \tag{7}$$

where $\tilde{\mathbf{g}}_i^{\mathbf{x}}$ and N represent the $i-th$ geometry information and total number of graph points in the local region along x-coordinate, respectively. Similar to x-coordinate, we can also compute the mean and standard deviation of geometry information along y-coordinate and z-coordinate. With all three coordinates, we can obtain 6 dimensional geometry density features from the average operation of extracted features across various local regions.

In Fig. 5, we provide the changes of mean value and standard deviation for two 3D point clouds at different distortion degrees. Here, the point clouds are distorted with light and heavy V-PCC compression artifacts. As can be seen from this figure, the histogram comparisons of the distorted 3D point clouds demonstrate that the mean value and standard deviation from all three coordinates can effectively reveal the quality variation. In addition, the effectiveness of these features is mainly because they can reflect the density of 3D point clouds.



(a) (b) (c) (d)

Fig. 5. The changes of mean value and standard deviation for 3D point clouds at different distortion degrees. (a) Lightly compressed 3D point cloud, (b) heavily compressed 3D point cloud, (c) the mean value changes of (a-b), (d) the standard deviation changes of (a-b). For better comparisons, the mean values are 411.3826 / 380.2105, 466.3556 / 452.1393 and 317.2707 / 342.3323 for x, y and z coordinates, respectively. Meanwhile, the standard deviation are 2.8366 / 2.7960, 2.9506 / 2.8458 and 2.2579 / 2.2111 for x, y and z coordinates, respectively.

3.3 Color Naturalness

As for 3D point clouds, color information is a significant attribute. Thus, we further extract color naturalness features from the constructed local region. By taking the human perception into account, we first convert the RGB components to YUV space. Then, to normalize the distributions, we apply the zero-phase component analysis followed by local mean subtraction and divisive normalization [37] to the separate YUV channels of distorted local regions as:

$$\overline{\rho}_z(y,u,v) = \frac{\tilde{\rho}_z(y,u,v) - \mu(y,u,v)}{\delta(y,u,v) + \mathrm{C}} \tag{8}$$

where $\tilde{\rho}_z(y,u,v)$ is the pixel values of the constructed local regions in YUV space after the whitening filter of zero-phase component analysis [10, 66], and $\overline{\rho}_z(y,u,v)$ represents the final normalized YUV channels. $\mathrm{C} = 1$ is a small constant that avoids instabilities when the denominator tends to 0. The local mean and standard deviation of YUV channels can be computed by:

$$\mu(y,u,v) = \sum_{l=-L}^{L}\sum_{t=-T}^{T} \omega_{l,t}\tilde{\rho}_z(y,u,v), \tag{9}$$

$$\delta(y,u,v) = \sqrt{\sum_{l=-L}^{L}\sum_{t=-T}^{T} \omega_{l,t}\left(\tilde{\rho}_z(y,u,v) - \mu(y,u,v)\right)^2}, \tag{10}$$

where $\omega = \left\{\omega_{l,t} \mid l = -L,\ldots,L, t = -T,\ldots,T\right\}$ indicates a circularly symmetric Gaussian weighting function. Motivated by [32], we set $L = T = 3$.

In Fig. 6, we show an example of statistical probability distribution for normalized coefficients, where the 3D point cloud is distorted by downscaling artifacts. On the basis of its symmetrically Gaussian-like appearance, we use the generalized Gaussian distribution (GGD) [40] to capture the regularity of NSS, which is given by:

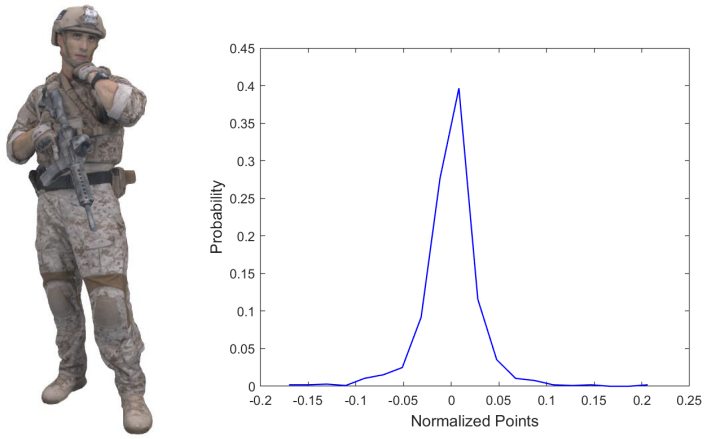
$$f\left(x;\lambda,\varepsilon^2\right) = \frac{\lambda}{2\eta\Gamma(1/\lambda)} e^{-\left(\frac{|x|}{\eta}\right)^{\lambda}}, \tag{11}$$

where

$$\eta = \varepsilon \sqrt{\frac{\Gamma(1/\lambda)}{\Gamma(3/\lambda)}}, \tag{12}$$

and $\Gamma(\cdot)$ is the gamma function as follows:

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx, a > 0. \tag{13}$$



(a) (b)

Fig. 6. An example of statistical probability distribution for normalized coefficients. (a) A distorted 3D point cloud with downscaling artifacts, (b) the corresponding statistical probability distribution for normalized coefficients of (a).

In the GGD, λ controls the shape of statistical probability distribution and ε^2 reflects the variance. For each color channel of the input distorted point cloud, we estimate 2 parameters (λ, ε^2) from the GGD fit by the moment matching-based approach [40]. Additionally, we use four scales containing the original scale as well as the reduction by the factors of 2, 4, and 8. Totally, with YUV channels, we have 24 dimensional color naturalness features which are calculated by the average operation among all the local regions.

3.4 Angular Consistency

Except for the geometry and color information, another important attribute is normal that can indicate the orientations of points. To fully use the unique normal vectors, apart from regional pre-processing, they have also been verified to be correlated to visual quality [1, 47, 49]. Therefore, we propose angular consistency features based on the normal information. Specifically, for each 3D point in the distorted point clouds, we first estimate the normal vector $\vec{\chi}_i$ of the point by using multiple neighboring points to fit a local plane for determining the normal vector. It should be noted that the number of neighboring points for normal vector estimation is validated in the experiments. Then, we adopt the k-nearest algorithm to choose the set of neighbors for each normal vector. In such case, assume that the normal vector of each neighbor point is denoted by $\vec{\chi}_j$, we compute the cosine similarity between $\vec{\chi}_i$ and $\vec{\chi}_j$ by:

$$\Omega = \cos(\zeta) = \frac{\vec{\chi}_i \cdot \vec{\chi}_j}{\|\vec{\chi}_i\| * \left\|\vec{\chi}_j\right\|}, \tag{14}$$

where $\Omega \in [-1, 1]$ and ζ is the angular value of the two normal vectors. Based on this, the inverse cosine of the obtained Ω is calculated as:

$$\zeta' = \arccos(|\Omega|), \tag{15}$$

where $\zeta' \in [0, \pi/2]$ and thus the ultimate angular similarity can be computed in the range of $[0, 1]$ as follows:

$$\psi = 1 - \frac{2\zeta'}{\pi}. \tag{16}$$

From the above-mentioned computation, we obtain the angular similarity matrix for every distorted point cloud. On one hand, we conduct an average operation for the point number dimension. On the other hand, 5 kinds of statistics are calculated for the k-nearest neighbor dimension, where the first two statistical features are the mean ι and standard deviation σ. Furthermore, the skewness and kurtosis are computed as:

$$S = E\left[\left(\frac{\psi_k - \iota}{\sigma}\right)^3\right], \tag{17}$$

$$K = E\left[\left(\frac{\psi_k - \iota}{\sigma}\right)^4\right], \tag{18}$$

where $E[\cdot]$ denotes the expectation operator, and ψ_k is the angular similarity value. Except for these moments, as suggested in [14], the entropy is also obtained by:

$$H = -\sum_k p(\psi_k) log p(\psi_k), \tag{19}$$

where $p(\psi_k)$ is the probability of ψ_k. These statistics of angular similarity matrix can be regarded as the angular consistency of normal vectors.

To extract the angular consistency features, inspired by the hierarchical multi-scale perception of the HVS [16, 54], here we employ more than one scale for better feature representations. As illustrated in Fig. 7, two scales of PCL compressed point clouds are shown, where (a) and (b) have 132, 518 points and 66, 259 points, respectively. From the two scales, we have 10 dimensional angular consistency features, which are used for quantifying visual distortions of 3D point clouds.

3.5 Quality Regression

After the regional pre-processing and quality-aware feature extraction, a quality regression module is applied to map the relevant features onto final quality scores. Many regressors can be used for this purpose, among whom we select the well-known support vector regression (SVR), random forest regression (RFR), and Extra Trees due to their capacity and popularity in quality assessment problems [12, 17, 34, 42, 65]. Especially, they have also been widely used for the spatial-domain and transform-domain NSS in the 2D image quality evaluation task [15, 33, 35, 38].

In our framework, given a test 3D point cloud, all features from geometry density, color naturalness and angular consistency aspects are extracted and concatenated to a feature vector. Before predicting the quality result of the test 3D point cloud, we use the labeled point clouds to train the regressor. With the trained regressor, we can predict the

(a) (b)

Fig. 7. Demonstration of two scales for distorted 3D point clouds. (a) A distorted 3D point cloud with PCL compression, (b) the downsampled scale of (a).

perceptual quality of any input test 3D point cloud. The Library for Support Vector Machines (LIBSVM) [7] is utilized to implement SVR with a radial basis function kernel, together with the RFR and Extra Trees proposed in [6, 11, 34] are used for our experiments. Additionally, we follow the common regressor parameter settings that have been used in the training of mainstream quality assessment models.

4 VALIDATION OF PROPOSED METHOD

4.1 Evaluation Databases and Criteria

We conduct experiments to validate our proposed SGR on four publicly available subjective point cloud quality databases, including Waterloo [22, 44], M-PCCD [4], SJTU [56] and LS-PCQA [25] databases. Specifically, followed by the recent NR quality assessment work [24], we first use the selected Waterloo database to compare the proposed SGR with a variety of state-of-the-art FR, RR and NR quality assessment methods for 3D point clouds. This database contains 20 original reference point cloud content with different geometry and texture complexities. Four distortion types are considered to generate the corresponding quality-degraded point clouds, which include 60 downsampling distorted point clouds, 180 distorted point clouds with Gaussian noise, 320 G-PCC (T) compressed point clouds, and 180 V-PCC compressed point clouds. Moreover, the adopted SJTU database consists of 6 pristine point cloud content, i.e., longdress, loot, redandblack, shiva, soldier, and statue. To produce distorted point clouds, the Octree-based compression, color noise, downsampling, and geometry Gaussian noise are introduced to the original point clouds. As for M-PCCD and LS-PCQA databases, the numbers of original point cloud content are 8 and 85, respectively. By involving a variety of distortion types, 232 and 930 distorted 3D point clouds are generated in the two databases.

Note that all the 3D point clouds in the adopted four databases have no pre-defined normal vectors. Apart from the reference and distorted point clouds, each database also provides subjective quality labels in the form of mean opinion score (MOS), which are rated by viewers with several display modes, such as the direct point cloud format [56] and converted video format [22].

To validate the proposed SGR and compare it with other state-of-the-arts, we adopt three commonly-used evaluation criteria in quality assessment field, including Spearman Rank-Order Correlation Coefficient (SROCC), Kendall Rank-Order Correlation Coefficient (KROCC), and Pearson Linear Correlation Coefficient (PLCC). Among these evaluation criteria, the SROCC is usually used to measure prediction monotonicity, while KROCC can be applied to evaluate the ordinal association between two measured quantities. Moreover, the PLCC can be used to evaluate prediction accuracy. It should be noted that higher correlation coefficients indicate better performance for objective quality models.

Besides, before computing the PLCC for different objective quality assessment approaches, a five-parameter logistic nonlinear fitting function [36] is adopted to map the predicted quality scores into a common scale as:

$$g(o) = \beta_1\left(\frac{1}{2} - \frac{1}{1+e^{(\beta_2(o-\beta_3))}}\right) + \beta_4 o + \beta_5, \tag{20}$$

where $(\beta_1 ... \beta_5)$ are five parameters to be fitted. o represents the raw objective score produced by objective quality models and $g(o)$ denotes the regressed score after the nonlinear mapping.

Table 1. Performance evaluation of different point cloud content, where the best NR is in red and the best FR is in blue.

Method Type	Method Name	Banana			Cauliflower			Mushroom			Pineapple			Overall		
		SROCC	KROCC	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC
FR	$PSNR_{MSE,p2po}$ [29]	0.64	0.47	0.73	0.49	0.35	0.55	0.63	0.48	0.66	0.41	0.30	0.52	0.48	0.33	0.50
	$PSNR_{MSE,p2pl}$ [47]	0.54	0.40	0.57	0.30	0.23	0.40	0.47	0.37	0.55	0.29	0.21	0.26	0.36	0.25	0.37
	$PSNR_Y$ [28]	0.62	0.47	0.72	0.54	0.40	0.56	0.60	0.44	0.79	0.74	0.57	0.77	0.53	0.37	0.56
	AS_{MSE} [1]	0.34	0.26	0.39	0.24	0.18	0.37	0.26	0.18	0.39	0.25	0.20	0.23	0.21	0.14	0.23
	$SSIM_{projected}$ [52]	0.77	0.60	0.72	0.77	0.61	0.80	0.73	0.57	0.85	0.82	0.66	0.82	0.58	0.42	0.59
	GraphSIM [58]	0.46	0.37	0.53	0.59	0.43	0.61	0.65	0.50	0.68	0.15	0.10	0.37	0.46	0.32	0.47
	PCQM [31]	0.74	0.56	0.64	0.69	0.52	0.67	0.76	0.60	0.71	0.79	0.68	0.74	0.71	0.52	0.65
	PointSSIM [2]	0.18	0.12	0.10	0.21	0.19	0.36	0.33	0.26	0.36	0.27	0.22	0.33	0.18	0.14	0.14
RR	*PCMRR* [49]	0.47	0.38	0.42	0.29	0.20	0.30	0.18	0.14	0.19	0.35	0.26	0.31	0.26	0.19	0.29
NR	*NR-3DQA* [63]	0.67	0.50	0.66	0.65	0.47	0.64	0.79	0.63	0.75	0.70	0.53	0.73	0.59	0.42	0.64
	PQA-Net [24]	0.52	0.39	0.53	0.69	0.52	0.70	0.71	0.56	0.77	0.89	0.71	0.87	0.69	0.51	0.70
	PRLGQA [45]	0.17	0.12	0.36	0.22	0.13	0.14	0.18	0.13	0.19	0.19	0.14	0.27	0.11	0.08	0.11
	IT-PCQA [57]	0.64	0.48	0.68	0.46	0.32	0.55	0.55	0.39	0.56	0.55	0.40	0.61	0.42	0.29	0.45
	MM-PCQA [64]	0.58	0.43	0.61	0.38	0.28	0.53	0.11	0.07	0.34	0.63	0.46	0.64	0.45	0.32	0.51
	ResSCNN [25]	0.28	0.20	0.49	0.23	0.18	0.40	0.16	0.10	0.24	0.36	0.26	0.53	0.42	0.29	0.47
	Proposed SGR (SVR)	0.68	0.51	0.70	0.68	0.50	0.70	0.80	0.66	0.85	0.73	0.58	0.79	0.73	0.55	0.74
	Proposed SGR (RFR)	0.86	0.68	0.83	0.70	0.52	0.72	0.82	0.65	0.86	0.73	0.55	0.78	0.73	0.55	0.75
	Proposed SGR (Extra Trees)	0.77	0.61	0.78	0.74	0.56	0.77	0.74	0.58	0.85	0.55	0.41	0.68	0.65	0.48	0.68

Table 2. Performance evaluation of various distortion types, where the best NR is in red and the best FR is in blue.

Method Type	Method Name	Downsampling			Gaussian Noise			G-PCC			V-PCC		
		SROCC	KROCC	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC
FR	$PSNR_{MSE,p2po}$ [29]	0.86	0.67	0.96	0.63	0.44	0.67	0.39	0.29	0.41	0.42	0.30	0.48
	$PSNR_{MSE,p2pl}$ [47]	0.76	0.52	0.85	0.62	0.43	0.67	0.42	0.30	0.42	0.46	0.32	0.48
	$PSNR_Y$ [28]	0.69	0.55	0.70	0.84	0.68	0.90	0.73	0.55	0.75	0.32	0.25	0.48
	AS_{MSE} [1]	0.77	0.52	0.96	0.67	0.49	0.70	0.03	0.03	0.13	0.49	0.32	0.60
	$SSIM_{projected}$ [52]	0.87	0.73	0.91	0.66	0.48	0.84	0.63	0.48	0.63	0.35	0.29	0.44
	GraphSIM [58]	0.79	0.64	0.97	0.71	0.57	0.72	0.61	0.44	0.70	0.32	0.27	0.61
	PCQM [31]	0.89	0.73	0.85	0.87	0.70	0.89	0.85	0.69	0.72	0.59	0.45	0.59
	PointSSIM [2]	0.91	0.76	0.97	0.63	0.46	0.70	0.77	0.60	0.78	0.48	0.34	0.51
RR	*PCMRR* [49]	0.88	0.70	0.89	0.88	0.73	0.89	0.19	0.13	0.20	0.55	0.42	0.50
NR	*NR-3DQA* [63]	0.90	0.75	0.81	0.86	0.69	0.89	0.64	0.49	0.56	0.54	0.37	0.51
	PQA-Net [24]	0.80	0.64	0.97	0.64	0.44	0.75	0.67	0.51	0.68	0.45	0.30	0.60
	PRLGQA [45]	0.86	0.63	0.65	0.21	0.14	0.24	0.33	0.23	0.35	0.42	0.26	0.41
	IT-PCQA [57]	0.81	0.63	0.62	0.56	0.40	0.54	0.44	0.31	0.44	0.66	0.50	0.73
	MM-PCQA [64]	0.95	0.84	0.96	0.87	0.68	0.90	0.62	0.45	0.69	0.51	0.36	0.58
	ResSCNN [25]	0.93	0.81	0.82	0.79	0.63	0.78	0.54	0.37	0.53	0.75	0.58	0.71
	Proposed SGR (SVR)	0.82	0.61	0.95	0.91	0.74	0.94	0.70	0.52	0.73	0.75	0.57	0.79
	Proposed SGR (RFR)	0.84	0.67	0.97	0.85	0.68	0.92	0.69	0.50	0.72	0.72	0.53	0.76
	Proposed SGR (Extra Trees)	0.80	0.61	0.91	0.85	0.67	0.90	0.50	0.36	0.57	0.66	0.48	0.72

4.2 Performance of Objective Models

In order to validate the performance of our proposed SGR, we compare the proposed SGR with state-of-the-art FR, RR and NR models on the largest subjective point cloud quality database, i.e., the Waterloo database [22, 44]. These include 8 FR methods, where both point-based and projection-based approaches are compared. Among them, classical point-based metrics contain $PSNR_{MSE,p2po}$, $PSNR_{MSE,p2pl}$, $PSNR_Y$, AS_{MSE}, *GraphSIM*, *PCQM*, *PointSSIM*, and the projection-based FR metric involves $SSIM_{projected}$. Representative RR model is also taken into consideration, such as *PCMRR*. In addition, we compare with 6 NR methods, consisting of *NR-3DQA*, *PQA-Net*, *PRLGQA*, *IT-PCQA*, *MM-PCQA*, and *ResSCNN*.

By following [24], we conduct the comparison for various visual content and distortion types. The evaluation results of different point cloud content are reported in TABLE 1. We can observe that our proposed SGR achieves competitive results for all the test visual content. Besides, by using different regressors, our framework delivers promising mapping results, demonstrating that the proposed SGR does not rely on specific regressor. More importantly, our proposed SGR is superior to many deep learning metrics regarding the overall performance. It should be noted that the *PQA-Net*, *PRLGQA*, *IT-PCQA*, *MM-PCQA*, and *ResSCNN* are all deep learning-based NR quality assessment models for 3D point clouds. In addition, beyond the performance results for individual visual content, we show the performance comparisons of various distortion types in TABLE 2. Again, our proposed SGR outperforms some deep learning models for all distortion types, especially for Gaussian noise, G-PCC and V-PCC compression. Therefore, the two tables demonstrate the superiority of our SGR algorithm.

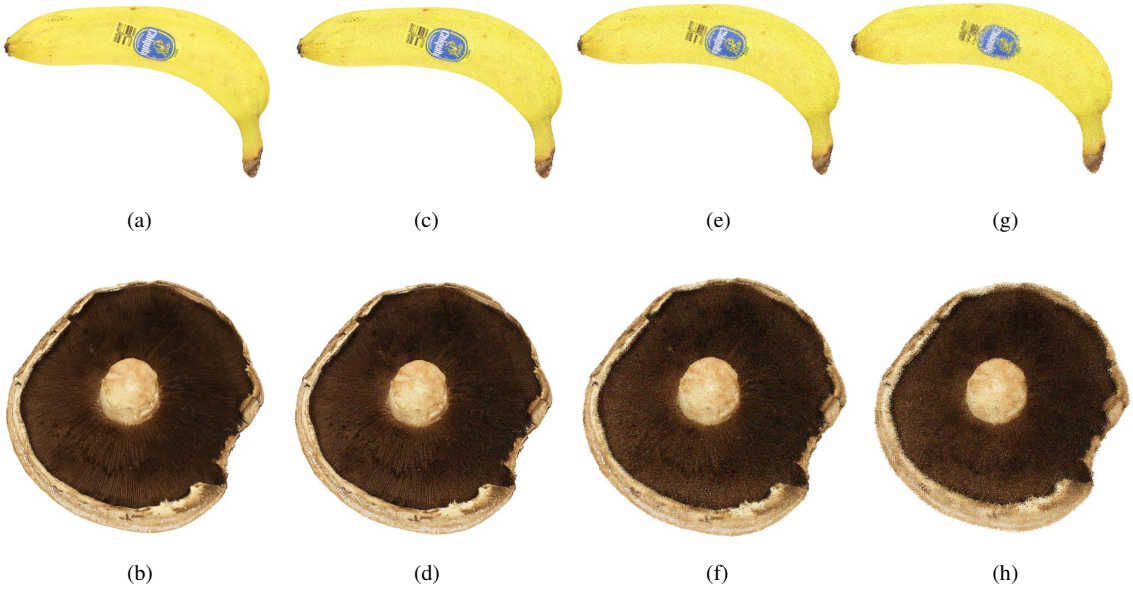


Fig. 8. Examples of perceptual quality prediction, where each row represent an original point cloud content (a-b) and the corresponding distortion degree increases from left to right. (c) SGR(SVR)=74.4780 / SGR(RFR)=61.9948 / MOS=60.4722, (d) SGR(SVR)=73.5027 / SGR(RFR)=68.9962 / MOS=75.0988, (e) SGR(SVR)=49.7316 / SGR(RFR)=46.3371 / MOS=49.6677, (f) SGR(SVR)=43.2950 / SGR(RFR)=44.6210 / MOS=46.5499, (g) SGR(SVR)=38.7809 / SGR(RFR)=36.8399 / MOS=27.0171, (h) SGR(SVR)=36.6015 / SGR(RFR)=35.2080 / MOS=37.9496. Higher SGR or MOS indicates better visual quality.

4.3 Visualization Comparison

In addition to quantitative performance results, we compare visualized point clouds with the predicted scores by our proposed SGR. Both SVR and RFR predicted quality results are computed, since they deliver the relatively good performance for Banana and Mushroom content from TABLE 1.

We show some examples of perceptual quality prediction in Fig. 8. Each row represents one visual content. That is, the original reference content are banana and mushroom, which contain 807,184 and 1144,603 points, respectively. By introducing the Gaussian noise, the reference point cloud would be distorted with various degrees. From the trends of SGR (SVR) and SGR (RFR) values, we can see that the proposed SGR method successfully distinguishes distorted point cloud data of different qualities. Therefore, the proposed model can effectively evaluate the perceptual quality of distorted 3D point clouds. In addition, for point clouds with relatively lower quality, i.e., (e-f) and (g-h), the prediction results from two models are more consistent. This may be because compared to low-quality point clouds, it is harder to predict the perceptual quality of high-quality point cloud data.

Table 3. Performance comparison on other subjective databases, where the best is in red and the second best is in blue.

Database	M-PCCD		SJTU		LS-PCQA	
Method	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC
NR-3DQA [63]	0.86	0.83	0.88	0.91	0.56	0.57
PQA-Net [24]	0.60	0.65	0.82	0.85	0.09	0.06
PRLGQA [45]	0.45	0.36	0.36	0.33	0.16	0.12
IT-PCQA [57]	0.75	0.82	0.85	0.93	0.45	0.55
MM-PCQA [64]	0.81	0.78	0.86	0.89	0.58	0.62
ResSCNN [25]	0.72	0.71	0.85	0.87	0.57	0.57
Proposed SGR (SVR)	0.80	0.87	0.89	0.93	0.47	0.52
Proposed SGR (RFR)	0.91	0.92	0.84	0.89	0.65	0.67
Proposed SGR (Extra Trees)	0.88	0.88	0.80	0.87	0.53	0.56

Table 4. Performance results of individual component on SJTU database, where the best is in bold.

Method	SROCC	KROCC	PLCC
only Geometry Density	0.58	0.41	0.70
only Color Naturalness	0.54	0.40	0.72
only Angular Consistency	0.72	0.55	0.84
without Geometry Density	0.83	0.66	0.89
without Color Naturalness	0.71	0.54	0.86
without Angular Consistency	0.78	0.61	0.85
Proposed SGR (SVR)	**0.89**	**0.72**	**0.93**

4.4 Validity on Other Subjective Databases

Apart from the largest Waterloo database, we further test the performance of the proposed SGR on other subjective databases, including M-PCCD [4], SJTU [56], and LS-PCQA [25]. Note that only the distorted 3D point clouds with individual/single distortion types are tested.

In the experiments, followed by [24], each database is randomly divided into training and testing sets with 80%-20% splitting. Our model is performed 1,000 times running. TABLE 3 provides the comparison results, where the state-of-the-art NR models are compared with our proposed SGR. From this table, we can find that the proposed SGR achieves the

best performance on all databases, which demonstrates the effectiveness of our SGR in general. In addition, regarding the adopted regression model, the performance of SVR is better for the SJTU database, while the results of RFR are higher for the other two databases. This may be because the distributions of these subjective databases are different.

Table 5. Performance results of keypoint resampling strategies on SJTU database, where the best is in bold.

Method	SROCC	KROCC	PLCC
Color	0.85	0.69	0.91
Geometry	0.86	0.69	0.91
Proposed SGR (SVR)	**0.89**	**0.72**	**0.93**

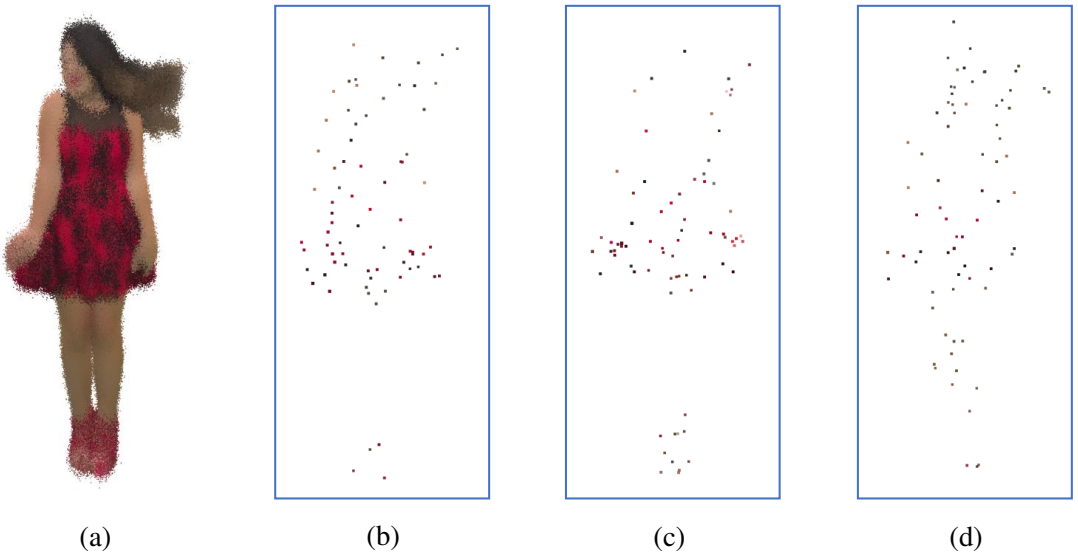


Fig. 9. Keypoint resampling comparisons. (a) An example of distorted point clouds, (b-d) the corresponding keypoints resampled from (a) by using geometry, color, and normal vectors, respectively.

4.5 Ablation Study

Since the proposed SGR is composed of three groups of quality-aware features including geometry density, color naturalness and angular consistency, it is worthwhile to validate the performance of each individual feature component. In TABLE 4, we conduct the ablation study and show the results. We can find that each component has a certain contribution to the final model. Moreover, integrating all the components leads to the best results of the proposed SGR framework.

Additionally, the keypoint resampling is based on the unique normal vectors, we test the cases by using other information, such as the color attributes and geometry information. The performance results are listed in TABLE 5, where the proposed method outperforms the others. One possible explanation may be that the normal vectors can

represent the structures of point clouds. Here, we also give the comparisons of keypoint resampling in Fig. 9. As can be seen in this figure, we take a point cloud distorted by severe Gaussian noise as an example. First, by observations, we can find that the resampled keypoints in both (b) and (c) cannot cover the woman's leg part. Instead, by using the normal vectors in (d), the keypoints cover more visual content. Second, since the extracted quality-aware features after regional pre-processing are based on geometry and color aspects, exploiting the normal vectors for keypoint resampling can make use of more known information of 3D point clouds (normal, geometry and color).

Table 6. Performance comparison of cross database validation, where the best is in bold. The training and testing databases are Waterloo and M-PCCD, respectively.

Method	SROCC	KROCC	PLCC
NR-3DQA [63]	0.10	0.07	0.13
PQA-Net [24]	0.30	0.20	0.42
PRLGQA [45]	0.30	0.20	0.23
IT-PCQA [57]	0.46	0.33	0.43
MM-PCQA [64]	0.52	0.37	0.52
ResSCNN [25]	0.32	0.22	0.32
Proposed SGR (SVR)	0.54	0.40	0.52
Proposed SGR (RFR)	**0.59**	**0.44**	**0.59**
Proposed SGR (Extra Trees)	0.44	0.32	0.47

Table 7. Evaluation on computational time versus correlation performance.

Method	Time (s)	SROCC
NR-3DQA [63]	6.91	0.88
PQA-Net [24]	6.09	0.82
PRLGQA [45]	60.95	0.36
IT-PCQA [57]	14.76	0.85
MM-PCQA [64]	16.39	0.86
ResSCNN [25]	5.38	0.85
Proposed SGR (SVR)	1.66	0.89

4.6 Model Generality and Computational Time

First, to verify the model generality of our proposed SGR, we train our method on the Waterloo database, and then test it on the M-PCCD database. The performance results are shown in Table 6. Here, we conduct the comparison with state-of-the-art NR metrics. From the table, we can observe that the proposed SGR method delivers better model generality when compared with the others.

Second, we provide evaluation on computation time of the NR quality assessment methods for 3D point clouds. It should be noted that we also list the correlation performance for fair comparison. The evaluation results are illustrated in Table 7. Here we choose SVR as it is the best regressor for SJTU database. Moreover, except for the *NR-3DQA*, the other NR metrics are all deep learning models. It can be seen that our proposed SGR is demonstrated to have lower computation with relatively competitive performance, which further prove the superiority of our proposed method regarding time complexity.

Table 8. Performance results of different keypoint numbers on SJTU database, where M is the total number of points before resampling.

Keypoint Number	SROCC	KROCC	PLCC
$M/1,000$	0.89	0.73	0.94
$M/2,000$	0.89	0.72	0.93
$M/10,000$	0.89	0.72	0.93
$M/20,000$	0.85	0.68	0.90

4.7 Parameter Sensitivity Test

It is interesting to explore the parameter sensitivity test in our proposed SGR framework. For example, the adopted keypoint number and the number of points used for local plane fitting when estimating normal vectors are explored in our experiment. Note that we exploit the Matlab function pcnormals for the normal computation. The performance results are shown in TABLE 8 and TABLE 9, respectively. First, we can observe that when the keypoint number decreases from $M/10,000$ to $M/20,000$, the performance drops a lot. Besides, if we increase the keypoint number, the performance values are similar but the computational complexity would increase. Thus, here we set the keypoint number equaling to $M/10,000$, which can balance the model performance and complexity. Second, 12 points used for local plane fitting can deliver the best performance. With more and less points, the results would decrease. Therefore, we choose this parameter equaling to 12 in the proposed SGR method.

Table 9. Performance results of various point numbers used for local plane fitting when computing normals on SJTU database, where the best is in bold.

Point Number	SROCC	KROCC	PLCC
6	0.79	0.61	0.85
11	0.87	0.71	0.91
12	**0.89**	**0.72**	**0.93**
13	0.87	0.70	0.91
15	0.85	0.68	0.90
18	0.83	0.65	0.89

Table 10. Performance comparison of color naturalness feature scales on SJTU database, where the best is in bold.

Feature Scale	SROCC	KROCC	PLCC
1	0.87	0.70	0.92
2	0.85	0.67	0.91
3	0.86	0.69	0.91
4	**0.89**	**0.72**	**0.93**
5	0.87	0.70	0.92
6	0.87	0.70	0.92

Furthermore, we change the color naturalness feature scales and report the performance comparisons in TABLE 10. From this table, we can see that when we choose four scales, our SGR has the best performance. Except for the feature scales, we also investigate different color spaces as illustrated in TABLE 11. Here we adopt five color spaces, including

Table 11. Performance evaluation of different color spaces on SJTU database, where the best is in bold.

Color Space	SROCC	KROCC	PLCC
RGB	0.84	0.66	0.90
HSV	0.79	0.62	0.87
LAB	0.87	0.70	0.92
YCbCr	0.86	0.68	0.91
YUV	**0.89**	**0.72**	**0.93**

RGB, HSV, LAB, YCbCr, and YUV. The results show that the YUV space used in our SGR method outperforms than the other color spaces. This further validates the effectiveness of the proposed SGR approach.

5 CONCLUSION

In this paper, we develop a general and efficient blind/no-reference quality assessment method for 3D point clouds, which is based on structure guided resampling. Inspired by the characteristics of the HVS and the perception of typical visual distortions, the proposed SGR model operates for distorted 3D geometry information and associated attributes without any special access to the reference, where regional pre-processing, quality-related feature extraction, and quality regression are involved in the proposed framework. In the experiments, we demonstrate that our proposed SGR algorithm can correlate well with human quality ratings on serveral subject-rated point cloud quality databases. Further, we also show the effectiveness of each constituted component of our proposed method.

In the future, except for dense 3D point clouds, we will explore subjective experiments for sparse point clouds. We also plan to explore more powerful resampling strategies and distortion-aware features to improve the quality assessment model. Moreover, the way to apply our proposed method to the automatic optimization of existing 3D point cloud processing algorithms could be another direction.

REFERENCES

[1] Evangelos Alexiou and Touradj Ebrahimi. 2018. Point cloud quality assessment metric based on angular similarity. In *International Conference on Multimedia and Expo*. IEEE, 1–6.

[2] Evangelos Alexiou and Touradj Ebrahimi. 2020. Towards a point cloud structural similarity metric. In *International Conference on Multimedia & Expo Workshops*. IEEE, 1–6.

[3] Evangelos Alexiou, Evgeniy Upenik, and Touradj Ebrahimi. 2017. Towards subjective quality assessment of point cloud imaging in augmented reality. In *2017 IEEE 19th International Workshop on Multimedia Signal Processing*. IEEE, 1–6.

[4] Evangelos Alexiou, Irene Viola, Tomás M Borges, Tiago A Fonseca, Ricardo L De Queiroz, and Touradj Ebrahimi. 2019. A comprehensive study of the rate-distortion performance in MPEG point cloud compression. *APSIPA Transactions on Signal and Information Processing* 8 (2019).

[5] Salima Bourbia, Ayoub Karine, Aladine Chetouani, Mohammed El Hassouni, and Maher Jridi. 2023. No-reference 3D Point Cloud Quality Assessment using Multi-View Projection and Deep Convolutional Neural Network. *IEEE Access* 11 (2023), 26759–26772.

[6] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[7] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011), 1–27.

[8] Jingdao Chen, Zsolt Kira, and Yong K Cho. 2019. Deep learning approach to point cloud scene understanding for automated scan to 3D reconstruction. *J. Comput. Civ. Eng* 33, 4 (2019), 04019027.

[9] Siheng Chen, Dong Tian, Chen Feng, Anthony Vetro, and Jelena Kovačević. 2017. Fast resampling of three-dimensional point clouds via graphs. *IEEE Transactions on Signal Processing* 66, 3 (2017), 666–681.

[10] Zhibo Chen, Wei Zhou, and Weiping Li. 2017. Blind stereoscopic video quality assessment: From depth perception to overall experience. *IEEE Transactions on Image Processing* 27, 2 (2017), 721–734.

[11] A Criminisi, J Shotton, and E Konukoglu. 2011. Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning [internet]. *Microsoft Research* (2011).

[12] Sathya Veera Reddy Dendi and Sumohana S Channappayya. 2020. No-reference video quality assessment using natural spatiotemporal scene statistics. *IEEE Transactions on Image Processing* 29 (2020), 5612–5624.

[13] Rafael Diniz, Pedro Garcia Freitas, and Mylene CQ Farias. 2021. Color and geometry texture descriptors for point-cloud quality assessment. *IEEE Signal Processing Letters* 28 (2021), 1150–1154.

[14] Yuming Fang, Kede Ma, Zhou Wang, Weisi Lin, Zhijun Fang, and Guangtao Zhai. 2014. No-reference quality assessment of contrast-distorted images based on natural scene statistics. *IEEE Signal Processing Letters* 22, 7 (2014), 838–842.

[15] Pedro Garcia Freitas, Welington YL Akamine, and Mylene CQ Farias. 2018. No-reference image quality assessment using orthogonal color planes patterns. *IEEE Transactions on Multimedia* 20, 12 (2018), 3353–3360.

[16] Kalanit Grill-Spector and Rafael Malach. 2004. The human visual cortex. *Annu. Rev. Neurosci.* 27 (2004), 649–677.

[17] Ke Gu, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang. 2014. Using free energy principle for blind image quality assessment. *IEEE Transactions on Multimedia* 17, 1 (2014), 50–63.

[18] Alireza Javaheri, Catarina Brites, Fernando Pereira, and Joao Ascenso. 2020. Point cloud rendering after coding: Impacts on subjective and objective quality. *IEEE Transactions on Multimedia* 23 (2020), 4049–4064.

[19] Guillaume Lavoué, Elisa Drelie Gelasca, Florent Dupont, Atilla Baskurt, and Touradj Ebrahimi. 2006. Perceptually driven 3D distance metrics with application to watermarking. In *Applications of Digital Image Processing XXIX*, Vol. 6312. SPIE, 150–161.

[20] Li Li, Zhu Li, Shan Liu, and Houqiang Li. 2020. Occupancy-map-based rate distortion optimization and partition for video-based point cloud compression. *IEEE Transactions on Circuits and Systems for Video Technology* 31, 1 (2020), 326–338.

[21] Hao Liu, Hui Yuan, Qi Liu, Junhui Hou, Huanqiang Zeng, and Sam Kwong. 2021. A hybrid compression framework for color attributes of static 3D point clouds. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 3 (2021), 1564–1577.

[22] Qi Liu, Honglei Su, Zhengfang Duanmu, Wentao Liu, and Zhou Wang. 2022. Perceptual Quality Assessment of Colored 3D Point Clouds. *IEEE Transactions on Visualization and Computer Graphics* (2022).

[23] Qi Liu, Hui Yuan, Raouf Hamzaoui, Honglei Su, Junhui Hou, and Huan Yang. 2021. Reduced reference perceptual quality model with application to rate control for video-based point cloud compression. *IEEE Transactions on Image Processing* 30 (2021), 6623–6636.

[24] Qi Liu, Hui Yuan, Honglei Su, Hao Liu, Yu Wang, Huan Yang, and Junhui Hou. 2021. PQA-Net: Deep No Reference Point Cloud Quality Assessment via Multi-View Projection. *IEEE Transactions on Circuits and Systems for Video Technology* 31, 12 (2021), 4645–4660.

[25] Yipeng Liu, Qi Yang, Yiling Xu, and Le Yang. 2023. Point cloud quality assessment: Dataset construction and learning-based no-reference metric. *ACM Transactions on Multimedia Computing, Communications and Applications* 19, 2s (2023), 1–26.

[26] Zian Lu, Hailiang Huang, Huanqiang Zeng, Junhui Hou, and Kai-Kuang Ma. 2022. Point cloud quality assessment via 3D edge similarity measurement. *IEEE Signal Processing Letters* 29 (2022), 1804–1808.

[27] Rufael Mekuria, Kees Blom, and Pablo Cesar. 2016. Design, implementation, and evaluation of a point cloud codec for tele-immersive video. *IEEE Transactions on Circuits and Systems for Video Technology* 27, 4 (2016), 828–842.

[28] Rufael Mekuria, Sebastien Laserre, and Christian Tulvan. 2017. Performance assessment of point cloud compression. In *Visual Communications and Image Processing*. IEEE, 1–4.

[29] RN Mekuria, Zhu Li, C Tulvan, and P Chou. 2016. Evaluation criteria for PCC (point cloud compression). (2016).

[30] Gabriel Meynet, Julie Digne, and Guillaume Lavoué. 2019. PC-MSDM: A quality metric for 3D point clouds. In *International Conference on Quality of Multimedia Experience*. IEEE, 1–3.

[31] Gabriel Meynet, Yana Nehmé, Julie Digne, and Guillaume Lavoué. 2020. PCQM: A full-reference quality metric for colored 3D point clouds. In *International Conference on Quality of Multimedia Experience*. IEEE, 1–6.

[32] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing* 21, 12 (2012), 4695–4708.

[33] Anush Krishna Moorthy and Alan Conrad Bovik. 2011. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing* 20, 12 (2011), 3350–3364.

[34] Hatef Otroshi-Shahreza, Arash Amini, and Hamid Behroozi. 2022. Feature-based no-reference video quality assessment using Extra Trees. *IET Image Processing* 16, 6 (2022), 1531–1543.

[35] Soo-Chang Pei and Li-Heng Chen. 2015. Image quality assessment using human visual DOG model fused with random forest. *IEEE Transactions on Image Processing* 24, 11 (2015), 3282–3292.

[36] Ann Marie Rohaly, Philip J Corriveau, John M Libert, Arthur A Webster, Vittorio Baroncini, John Beerends, Jean-Louis Blin, Laura Contin, Takahiro Hamada, David Harrison, et al. 2000. Video quality experts group: Current results and future directions. In *Visual Communications and Image Processing*, Vol. 4067. SPIE, 742–753.

[37] Daniel L Ruderman. 1994. The statistics of natural images. *Network: Computation in Neural Systems* 5, 4 (1994), 517.

[38] Michele A Saad, Alan C Bovik, and Christophe Charrier. 2012. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *IEEE Transactions on Image Processing* 21, 8 (2012), 3339–3352.

[39] Sebastian Schwarz, Marius Preda, Vittorio Baroncini, Madhukar Budagavi, Pablo Cesar, Philip A Chou, Robert A Cohen, Maja Krivokuća, Sébastien Lasserre, Zhu Li, et al. 2018. Emerging MPEG standards for point cloud compression. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 9, 1 (2018), 133–148.

[40] Karnran Sharifi and Alberto Leon-Garcia. 1995. Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. *IEEE Transactions on Circuits and Systems for Video Technology* 5, 1 (1995), 52–56.

[41] Hamid R Sheikh and Alan C Bovik. 2006. Image information and visual quality. *IEEE Transactions on Image Processing* 15, 2 (2006), 430–444.

[42] Likun Shi, Wei Zhou, Zhibo Chen, and Jinglin Zhang. 2019. No-reference light field image quality assessment based on spatial-angular measurement. *IEEE Transactions on Circuits and Systems for Video Technology* 30, 11 (2019), 4114–4128.

[43] Anuj Srivastava, Ann B Lee, Eero P Simoncelli, and S-C Zhu. 2003. On advances in statistical modeling of natural images. *Journal of Mathematical Imaging and Vision* 18, 1 (2003), 17–33.

[44] Honglei Su, Zhengfang Duanmu, Wentao Liu, Qi Liu, and Zhou Wang. 2019. Perceptual quality assessment of 3D point clouds. In *2019 IEEE International Conference on Image Processing*. IEEE, 3182–3186.

[45] Zhiyong Su, Chao Chu, Long Chen, Yong Li, and Weiqing Li. 2022. No-reference Point Cloud Geometry Quality Assessment Based on Pairwise Rank Learning. *arXiv preprint arXiv:2211.01205* (2022).

[46] Simeng Sun, Tao Yu, Jiahua Xu, Wei Zhou, and Zhibo Chen. 2022. GraphIQA: Learning distortion graph representations for blind image quality assessment. *IEEE Transactions on Multimedia* (2022).

[47] Dong Tian, Hideaki Ochimizu, Chen Feng, Robert Cohen, and Anthony Vetro. 2017. Geometric distortion metrics for point cloud compression. In *International Conference on Image Processing*. IEEE, 3460–3464.

[48] Jeroen van der Hooft, Tim Wauters, Filip De Turck, Christian Timmerer, and Hermann Hellwagner. 2019. Towards 6dof http adaptive streaming through point cloud compression. In *Proceedings of the ACM International Conference on Multimedia*. 2405–2413.

[49] Irene Viola and Pablo Cesar. 2020. A reduced reference metric for visual quality evaluation of point cloud contents. *IEEE Signal Processing Letters* 27 (2020), 1660–1664.

[50] Jianqiang Wang, Hao Zhu, Haojie Liu, and Zhan Ma. 2021. Lossy point cloud geometry compression via end-to-end learning. *IEEE Transactions on Circuits and Systems for Video Technology* 31, 12 (2021), 4909–4923.

[51] Yue Wang, Shusheng Zhang, Bile Wan, Weiping He, and Xiaoliang Bai. 2018. Point cloud and visual feature-based tracking method for an augmented reality-aided mechanical assembly system. *International Journal of Advanced Manufacturing Technology* 99, 9 (2018), 2341–2352.

[52] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612.

[53] Zhou Wang and Qiang Li. 2010. Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing* 20, 5 (2010), 1185–1198.

[54] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. 2003. Multiscale structural similarity for image quality assessment. In *Asilomar Conference on Signals, Systems & Computers*, Vol. 2. IEEE, 1398–1402.

[55] Xinju Wu, Yun Zhang, Chunling Fan, Junhui Hou, and Sam Kwong. 2021. Subjective quality database and objective study of compressed point clouds with 6DoF head-mounted display. *IEEE Transactions on Circuits and Systems for Video Technology* 31, 12 (2021), 4630–4644.

[56] Qi Yang, Hao Chen, Zhan Ma, Yiling Xu, Rongjun Tang, and Jun Sun. 2020. Predicting the perceptual quality of point cloud: A 3D-to-2D projection-based exploration. *IEEE Transactions on Multimedia* 23 (2020), 3877–3891.

[57] Qi Yang, Yipeng Liu, Siheng Chen, Yiling Xu, and Jun Sun. 2022. No-Reference Point Cloud Quality Assessment via Domain Adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 21179–21188.

[58] Qi Yang, Zhan Ma, Yiling Xu, Zhu Li, and Jun Sun. 2020. Inferring point cloud quality via graph similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020).

[59] Xiangyu Yue, Bichen Wu, Sanjit A Seshia, Kurt Keutzer, and Alberto L Sangiovanni-Vincentelli. 2018. A LiDAR point cloud generator: from a virtual world to autonomous driving. In *Proceedings of the ACM on International Conference on Multimedia Retrieval*. 458–464.

[60] Guangtao Zhai and Xiongkuo Min. 2020. Perceptual image quality assessment: a survey. *Science China Information Sciences* 63 (2020), 1–52.

[61] Juan Zhang, Wenbin Huang, Xiaoqiang Zhu, and Jenq-Neng Hwang. 2014. A subjective quality evaluation for 3D point cloud models. In *International Conference on Audio, Language and Image Processing*. IEEE, 827–831.

[62] Yujie Zhang, Qi Yang, and Yiling Xu. 2021. MS-GraphSIM: Inferring Point Cloud Quality via Multiscale Graph Similarity. In *Proceedings of the ACM International Conference on Multimedia*. 1230–1238.

[63] Zicheng Zhang, Wei Sun, Xiongkuo Min, Tao Wang, Wei Lu, and Guangtao Zhai. 2022. No-reference quality assessment for 3D colored point cloud and mesh models. *IEEE Transactions on Circuits and Systems for Video Technology* (2022).

[64] Zicheng Zhang, Wei Sun, Xiongkuo Min, Quan Zhou, Jun He, Qiyuan Wang, and Guangtao Zhai. 2022. MM-PCQA: Multi-Modal Learning for No-reference Point Cloud Quality Assessment. *arXiv preprint arXiv:2209.00244* (2022).

[65] Wei Zhou, Likun Shi, Zhibo Chen, and Jinglin Zhang. 2020. Tensor oriented no-reference light field image quality assessment. *IEEE Transactions on Image Processing* 29 (2020), 4070–4084.

[66] Wei Zhou, Jiahua Xu, Qiuping Jiang, and Zhibo Chen. 2021. No-reference quality assessment for 360-degree images by analysis of multifrequency information and local-global naturalness. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 4 (2021), 1778–1791.

[67] Wei Zhou, Guanghui Yue, Ruizeng Zhang, Yipeng Qin, and Hantao Liu. 2023. Reduced-reference quality assessment of point clouds via content-oriented saliency projection. *IEEE Signal Processing Letters* 30 (2023), 354–358.