Perceptual Quality Assessment of Colored 3D Point Clouds

Qi Liu, Honglei Su, *Member, IEEE,* Zhengfang Duanmu, *Student Member, IEEE,* Wentao Liu, *Student Member, IEEE,* and Zhou Wang, *Fellow, IEEE*

Abstract—3D point clouds have found a wide variety of applications in multimedia processing, remote sensing, and scientific computing. Although most point cloud processing systems are developed to improve viewer experiences, little work has been dedicated to perceptual quality assessment of 3D point clouds. In this work, we build a new 3D point cloud database, namely the Waterloo Point Cloud (WPC) database. In contrast to existing datasets consisting of small-scale and low-quality source content of constrained viewing angles, the WPC database contains 20 high quality, realistic, and omni-directional source point clouds and 740 diversely distorted point clouds. We carry out a subjective quality assessment experiment over the database in a controlled lab environment. Our statistical analysis suggests that existing objective point cloud quality assessment (PCQA) models only achieve limited success in predicting subjective quality ratings. We propose a novel objective PCQA model based on an attention mechanism and a variant of information content-weighted structural similarity, which significantly outperforms existing PCQA models. The database has been made publicly available at https://github.com/qdushl/Waterloo-Point-Cloud-Database.

Index Terms—Point cloud, subjective quality assessment, attention model, objective quality assessment.

1 Introduction

A 3D point cloud is a collection of points representing a 3D shape, object or environment. Each point can be described by its geometric coordinates and optional associated attributes. 3D point clouds [1]–[4] have found broad applications in manufacturing, construction, environmental monitoring, navigation, and animation. Many of these applications require high quality point clouds that faithfully reflect the geometry and perceptual attributes of the physical world. However, various distortions may be introduced during the acquisition, compression [5], transmission, storage, and rendering processes, leading to degraded perceptual quality by end users. Over the past decade, point cloud quality assessment (PCQA) has become an active field of research [6]–[62].

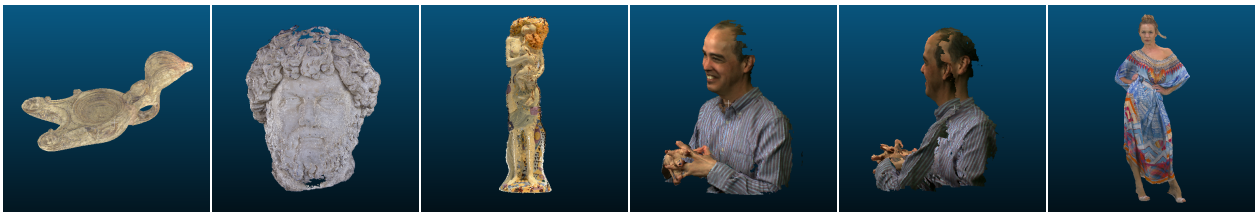
Since the human visual system (HVS) is the ultimate receiver of 3D point clouds in most applications, subjective quality assessment is the most straightforward and reliable approach to evaluate the quality of point clouds. A comprehensive subjective user study on a large-scale point cloud database brings several benefits. First, it advances our understanding about the HVS in evaluating the perceived quality of point clouds. Second, a diverse set of high quality source stimuli supply a fertile playground for point cloud processing algorithms, such as denoising [63], super-resolution [64], and compression [65]. Third, a subject-rated dataset provides a valuable source to train, validate, and test existing objective PCQA models.

Despite its importance, subjective quality evaluation is inconvenient, time-consuming, and expensive. To enable quality-centric point cloud systems in practice, objective PCQA models that can accurately predict subjective quality are highly desired. Although substantial effort has been made to develop objective PCQA models [12], [16], [21], [23], [25], [28]–[30], [32], [33], [35], [37]–[40], [44]–[47], [49]–[55], [55]–[62], they often fail to draw a connection to the HVS [66], [67] and/or struggle in handling the irregular representation of point clouds [21], [23]. Most importantly, none of these models is validated on large-scale subject-rated PCQA databases with diverse and high quality original point clouds, making their generalizability questionable.

In this work, we first introduce so-far the largest high quality point cloud dataset. By degrading the reference point clouds with diverse distortion types and levels, we create 740 distorted point clouds. A subjective experiment is then carried out in a controlled environment to evaluate the perceptual quality of these point clouds. The new database, named the Waterloo Point Cloud database (WPC), together with subjective labels are made publicly available to facilitate reproducible research. Using the WPC database, we conduct a comprehensive evaluation of existing objective PCQA models, which suggests that state-of-the-art models only achieve a moderate correlation with human visual perception. To overcome the problem, we develop an attention guided objective PCQA model, inspired by the information content weighted structural similarity measure (IW-SSIM) [69]. Experimental results demonstrate that the proposed model well correlates with subjective quality evaluations and significantly outperforms all existing PCQA models.

- *Qi Liu and Honglei Su are with the College of Electronic Information, Qingdao University, Qingdao, 266071, China (e-mail: sdqi.liu@gmail.com, suhonglei@qdu.edu.cn).*
- *Zhengfang Duanmu, Wentao Liu and Zhou Wang are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (e-mail: {zduanmu, w238liu, zhou.wang}@uwaterloo.ca).*

Corresponding author: Honglei Su.



(a) (b) (c) (d) (e) (f)

Fig. 1. Sample point clouds from existing datasets. (a) RomanOilLight. (b) Head. (c) Statue_Klimt. (d) Phil. (e) Phil2. (f) Longdress.

TABLE 1
COMPARISON OF EXISTING PUBLICLY AVAILABLE SUBJECT-RATED PCQA DATABASES

Database	Attribute	Source contents	Distortion type	Subject-rated point clouds
IRPC [24]	None, Color	6	PCL, G-PCC, V-PCC	54
vsenseVVDB [13]	Color	2	V-PCC	32
vsenseVVDB2 [34]	Color	8	Draco+JPEG, G-PCC, V-PCC	164
G-PCD [14], [15]	None	5	Octree-puring, Gassian noise	40
RG-PCD [9]	None	6	Octree-puring	24
M-PCCD [23]	Color	8	G-PCC, V-PCC	244
PointXR [36]	Color	5	G-PCC	100
NBU-PCD 1.0 [55]	Color	10	Octree	160
CPCD 2.0 [49]	Color	10	G-PCC, V-PCC, Gassian noise	360
SJTU-PCQA [45]	Color	10	Octree, downsampling, color and geometry noise	420
ICIP2020 [48]	Color	6	G-PCC, V-PCC	90
3DMDC [68]	Color	5	QGeo, QCol, SGeo, SCol	80
SIAT-PCQD [52]	Color	20	V-PCC	340
WPC (ours)	Color	**20**	Gassian noise, dowsampling, G-PCC, V-PCC	**740**

2 RELATED WORK

2.1 Existing PCQA Databases

The history of point cloud generation dated back at least to 1990's, when Turk and Levoy investigated computational methods for 3D surface reconstruction [71]. The resulting Stanford 3D scanning dataset is still in use in recent PCQA research [9], [14], [15]. The MPEG point cloud database [72] and the JPEG Pleno database [73] introduced more content types, such as cultural heritages, computer-generated objects, and human figures. These early databases provide a solid foundation for a series of subjective PCQA studies. Most later subject-rate point cloud databases were derived from these datasets [9], [13]–[15], [23], [24], [34], [36], [45], [48], [49], [52], [55], [68], as summarized in Table 1. Due to the simplified data collection process, these PCQA databases inherently suffer from several limitations. First, most point clouds in the Stanford 3D scanning repository [71] are colorless. Second, the scanning process fails to capture the aesthetic aspect of the objects, especially for those of cultural heritages. Typical examples include the "RomanOilLight" [73] as shown in Fig. 1 (a) and the "Head" [72] as shown in Fig. 1 (b). Third, some point clouds are of inferior perceptual quality, containing scanning noise ("Statue_Klimt" [72] in Fig. 1 (c)) or irregular edges ("Phil" [74] in Fig. 1 (d)). Fourth, many point clouds were scanned from a limited number of directions, as exemplified in Fig. 1 (e). However, real-world applications often require point clouds that allow for omni-directional presentation [9], [13]–[15], [23], [24], [52]. Fifth, existing PCQA databases are often of low content diversity. We wish to address these limitations in this study.

Besides the difficulty in dataset construction, the 3D nature and the irregular data representation of point clouds also create complications in the design of subjective experiment. Existing subjective PCQA tests [7]–[24], [34], [36], [38], [39], [45], [48], [52], [55], [56], [68], [70] are summarized in Table 2, where we include both publicly-available and privately-held studies. These studies vary significantly in the scoring methodology, the viewing display, the interaction method, and the rendering technology. There are several important observations. First, Double Stimulus Impairment Scale (DSIS) [75] is used more frequently as the testing protocol than Absolute Category Rating (ACR) and Pairwise Comparison (PC). One possible reason may be to make it easier for subjects to detect all artifacts and distortions such as color impairment [20], which is often difficult in reference-free tests such as ACR. Second, most subjective tests prefer 2D monitors as opposed to more advanced 3D monitors and HMDs. This might be because the latter often causes visual discomfort and fatigue that may affect the reliability of the subjective scores collected [76], [77]. Third, the subjective viewing experience may be either interactive where the subject controls the viewpoints and/or distances freely using a mouse or HMD, or passive where a predetermined virtual path of viewpoints and distances is used to create a video presented to the subject. Both methods work similarly well in existing subjective tests, and the passive method has a slight edge in terms of reproducibility. Fourth, point-based rendering is overwhelmingly popular in existing subjective tests due to its simplicity, stability and low computational complexity. While finding the best subjective testing methodology remains an open question, we opt to the most common settings in constructing the WPC dataset, as shown in Table 2.

2.2 Objective Quality Assessment of 3D Point Clouds

Depending on the application scope, objective PCQA models may be categorized into geometry-only models [16], [25], [26], [28]–[30], [33], [39], [40], [78] and general-purpose models [12], [21], [23], [28], [29], [32], [35], [37], [38], [44]–[47], [49]–[62]. Geometry-only PCQA models are dedicated to assess the perceptual quality of point clouds with only ge-

TABLE 2
EXPERIMENT SETTINGS IN EXISTING SUBJECTIVE TESTS

Literature	Methodology	Display device	Interaction method	Rendering mode
Javaheri *et al.* [11]	DSIS	2D monitor	Passive	Point
Javaheri *et al.* [7]	DSIS	2D monitor	Passive	Mesh
Alexiou *et al.* [8]	DSIS	2D monitor	Interactive	Point
Alexiou *et al.* [14]	DSIS, ACR	2D monitor	Interactive	Point
Alexiou *et al.* [15]	DSIS	HMD (AR)	Interactive	Point
Alexiou *et al.* [16]	DSIS, ACR	2D monitor	Interactive	Point
Alexiou *et al.* [17]	DSIS	2D monitor, HMD (AR)	Interactive	Point
Alexiou *et al.* [18]	DSIS, ACR	2D monitor	Interactive	Point
Alexiou *et al.* [9]	DSIS	2D monitor	Passive	Point, Mesh
Torlig *et al.* [12]	DSIS	2D monitor	Passive	Point
Alexiou *et al.* [10]	DSIS	3D monitor	Passive, Interactive	Mesh
Zhang *et al.* [19]	–	2D monitor	–	Point
Nehmé *et al.* [20]	DSIS, ACR	HMD (VR)	Passive	Mesh
Zerman *et al.* [13]	DSIS, PC	2D monitor	Interactive	Point
Alexiou *et al.* [21]	DSIS	2D monitor	Interactive	Point
da Silva Cruz *et al.* [22]	DSIS	2D monitor	Passive	Point
Alexiou *et al.* [23]	DSIS	2D monitor	Interactive	Point
Javaheri *et al.* [24]	DSIS	2D monitor	Passive	Point, Mesh
Jesús Gutiérrez *et al.* [70]	ACR-(HR)	HMD (MR)	Interactive	Mesh
Viola *et al.* [38]	DSIS	2D monitor	-	Point
Alexiou *et al.* [36]	DSIS, ACR	2D monitor, HMD (VR)	Passive, Interactive	Point
Zerman *et al.* [34]	ACR	2D monitor	Passive	Point
Javaheri *et al.* [39]	DSIS	2D monitor	Passive	Point
Hua *et al.* [55]	-	2D monitor	-	Point
Yang *et al.* [45]	ACR-(HR)	2D monitor	Interactive	Point
Perry *et al.* [48]	DSIS	2D monitor	Passive	Point
Liu *et al.* [56]	DSIS	2D monitor	Interactive	Point
Nehmé *et al.* [68]	DSIS	HMD (VR)	Interactive	Mesh
Wu *et al.* [52]	DSIS	HMD	Interactive	Point
WPC (ours)	DSIS	2D monitor	Passive	Point

ometric information, while general-purpose PCQA models take all quality-related attributes including color and surface normal into consideration. From the perspective of feature extraction, objective PCQA models may also be classified into point-based [16], [25], [28]–[30], [33], [35], [37]–[40], [44], [46], [47], [49]–[51], [54]–[57], [59], [60], [62] and projection-based models [12], [21], [23], [32], [45], [52], [53], [58], [61].

Both point-to-point and point-to-plane models employ variants of the Euclidean distance to quantify geometric distortions [25], [28], [29]. In [16], the cosine similarity measure is applied to the local surface normal. Similarly, the PC-MSDM model [30] computes the similarity between the curvature of the original and distorted signals. However, surface normal-based approaches are often susceptible to random noise in the acquisition process [16], [25], [28]–[30]. By incorporating machine learning techniques, [33] developed a generalized Hausdorff distance measure with enhanced robustness. In [39], a PSNR-based metric [25], [28], [29] is proposed by including a normalization factor that accounts for changes in the intrinsic point cloud resolution after rendering. In [29], PSNR-based methods are modified by a density coefficient determined by the peak of coordinate and the rendering resolution. In [40], a point-to-distribution quality assessment model is proposed by exploiting the correspondence between a point and a distribution of points in a small point cloud region. As such, the point cloud surface is characterized through the covariance of points within the local region, which is not overly influenced by the number of reconstructed points after decoding, but rather by a statistical characterization of the point locations.

Compared with geometry-only point clouds, colored point clouds have a broad range of applications. Many quality or distortion metrics for colored PCs have emerged recently [6], [28], [29], [35], [37], [38], [42], [44], [46], [47], [49]–[51], [54], [55], [57], [59], [60]. In [28], [29], point-to-point PSNR on the Y component (MPEG PSNR_Y) is used to estimate texture distortion of colored PCs, though such a direct extension of PSNR inevitably inherits the widely-known disadvantages of PSNR [66], [67]. Similarity based measures [66] are extended to PCQA [35], [37], [55]. In these methods, geometry-based, color-based, normal-based and curvature-based features are extracted from both the reference and distorted PCs, then geometry and color feature similarities are evaluated and combined to produce the overall objective scores. In [38], color histograms and correlograms are used to estimate the impairment of a distorted point cloud with respect to its reference. Geometry-only and color-only approaches are then combined to a rendering-independent objective PCQA metric. More recently, statistics of a variant of the Local Binary Pattern (LBP) [46], [47], Perceptual Color Distance Pattern (PCDP) [50] and Local Luminance Pattern (LLP) [51] descriptors are introduced to the area. In [54], the BitDance metric uses color and geometry texture descriptors. The statistics of color and geometry information of the reference and test PCs are compared and combined to estimate the perceived quality of the test point cloud. The GraphSIM approach [44], [60] uses graph signal gradient as a quality index to evaluate point cloud distortions. Considering the visual masking effect of point cloud's geometric information and the color perception of human eyes, the CPC-GSCT metric [49] uses geometric segmentation and color transformation respectively to construct geometric and color features and then to estimate the point cloud quality. Inspired by the point cloud generation process, the elastic potential energy similarity (EPES) model [59] introduces elastic forces to record the shaping of the point set, and uses the elastic potential energy difference to quantify point cloud distortion.

In addition to the aforementioned point-based models, there are also many projection-based models [12], [21], [23], [32], [45], [52], [53], such as projection-based PSNR (PSNRp) [12], [32], projection-based structural similarity (SSIMp) [12], [32], [66], projection-based multi-scale struc-

tural similarity (MS-SSIMp) [12], [32], [79] and projection-based pixel-domain visual information fidelity (VIFPp) [12], [32], [80]. Yang *et al.* [45] choose to project the 3D point cloud onto six perpendicular image planes of a cube for the color texture image and corresponding depth image, and aggregate image-based global and local features from all projected planes to a final objective index. Wu *et al.* [52] propose two projection-based objective quality evaluation methods: a weighted view projection based model and a patch projection based model. He *et al.* [53] project the colored texture information and curvature of point cloud onto 2D planes and extract texture and geometric statistical features, respectively, so as to characterize the texture and geometric distortion. However, these methods treat background padding pixels on projected image planes the same way as the foreground ones, leading to inferior quality prediction accuracy [69]. Alexiou *et al.* [21], [23] develop a post-processing algorithm to remove the influence of background pixels, but model complexity increases and robustness declines.

Reduced reference and no reference PCQA models have also been developed. Reduced reference PCQA models [6], [42] only require partial information about the reference PCs. In [42], geometry-based, normal-based and luminance-based features are extracted from the reference point cloud, transmitted alongside the content, and employed at the receiver side to help assess the quality of the distorted point cloud.In [6], two color features are proposed to estimate three content dependent parameters for reduced reference PCQA.No reference PCQA models require no information about the reference PCs [43], [57], [58], [61], [62]. Cao et al. [43] define point cloud quality as a function of the bitrate and observation distance. Nevertheless, bitrate alone cannot accurately estimate the point cloud quality, and the observation distance, a parameter often used in the rendering algorithm, is often not available in practical systems. The BQE-CVP metric [57] uses geometric feature, color feature and joint feature to develop a blind quality evaluator. Zhang et al. [62] project the 3D models from 3D space into quality related geometry and color feature domains, extract natural scene statistics (NSS) and entropy for quality aware features, and employ a Support Vector Regressor (SVR) to regress the quality-aware features into quality scores. Liu et al. [61] propose a deep learning-based PQA-Net model, which consists of a multi-view-based joint feature extraction and fusion (MVFEF) module, a distortion type identification (DTI) module, and a quality vector prediction (QVP) module. By using the distortion type labels, the DTI and the MVFEF modules are pre-trained to initialize the network parameters, and the full network is then jointly trained for quality prediction. Tao et al. [58] propose a point cloud projection and multi-scale feature fusion network that includes a joint color-geometric feature extractor, a two-stage multi-scale feature fusion, and a spatial pooling module.

3 Point Cloud Database Construction

3.1 Point Cloud Construction

Motivated by the lack of source 3D point clouds, we gather a collection of objects with diverse geometric and textural complexity, including snacks, fruits, vegetables, office supplies, and containers, etc. The selected contents are moderate in size and are omni-directional in viewing angle. Fig. 2 shows a snapshot for the reference point clouds constructed. The construction process is as follows.

- *Image acquisition*: The image acquisition is performed in a laboratory environment of a normal lighting condition without reflecting ceiling, walls and floors. A single-lens-reflex camera and a turntable are employed to take photos of an object from a variety of perspectives. A graph illustration of the acquisition process is shown in Fig. 3, where each photo is placed at its capture position relative to the object in the center.
- *3D reconstruction*: A sequence of operations including image alignment, sparse point cloud reconstruction, dense point cloud reconstruction, and point cloud merging are applied to each sequence of images using Agisoft Photoscan [81]. The resulting point clouds are further refined by Screened Poisson Surface Reconstruction [82] and re-sampled using CloudCompare [83].
- *Normalization*: Each point cloud is normalized to be fully contained in a unit-cube with a step size of 0.001, where duplicated points are removed [83]. A total of 20 voxelized point clouds are generated. The number of points in each point cloud ranges between 400K and 3M, with an average of 1.35M and a standard deviation of 656K. The specifications are given in Table 3.

3.2 Distortion Generation

We distort the source PCs with the following processes to simulate real-world application scenarios.

- *Downsampling*: Octree-based downsampling [83] is applied to the source point clouds. Each dimension is uniformly divided into 2^N intervals, where N represents the octree level. Points located in the same cube are then merged into one. In this study, N is set to be 7, 8, and 9, to cover diverse spatial resolutions.
- *Gaussian noise contamination*: White Gaussian noise is added independently to both geometry and texture elements with standard deviations of $\{0, 2, 4\}$ and $\{8, 16, 32\}$, respectively. Then both geometry and texture elements are rounded to the nearest integer, followed by points removal by Meshlab [84].
- *MPEG-PCC*: Two technologies were chosen as test models following MPEG's call for proposals for International Organization for Standardization [65]: G-PCC for static content and dynamically capturing, and V-PCC for dynamic content. In this work, G-PCC (Trisoup) reference codec [85] is employed to encode the original point clouds with 'max_NodeSizeLog$_2$' of $\{10\}$, 'NodeSizeLog$_2$' of $\{2, 4, 6\}$ and 'rahtQuantizationStep' of $\{64, 128, 256, 512\}$, respectively. G-PCC (Octree) [86] employs a downsampling method to encode the geometry information, and is thus not performed redundantly. We set the 'quantizationSteps' of texture encoding as $\{16, 32, 48, 64\}$. V-PCC reference codec [87] is employed to encode the original point clouds at three 'geometryQP' values and three 'textureQP' values, ranging from 35-50 and 35-50, respectively, followed by duplicated points removal [84].

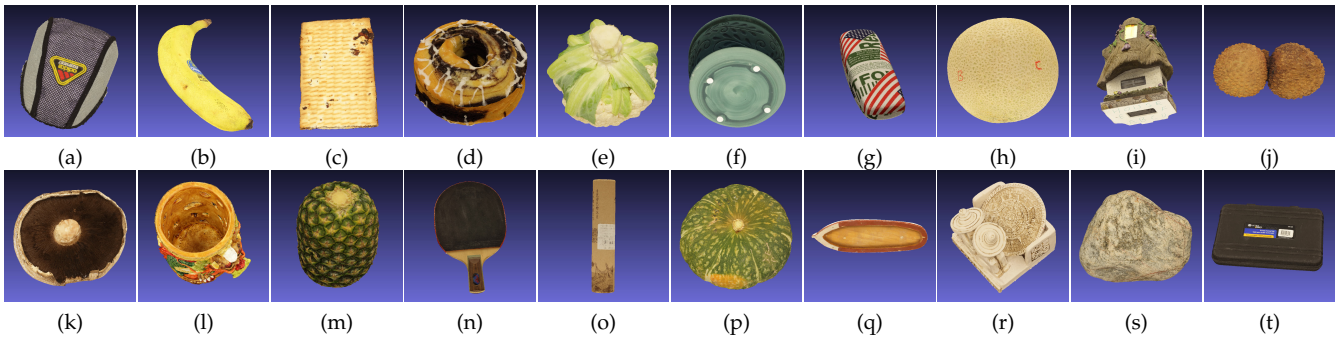


Fig. 2. Snapshots of acquired point clouds in the Waterloo Point Cloud database.

TABLE 3
CHARACTERISTICS OF ACQUIRED POINT CLOUDS IN THE WATERLOO POINT CLOUD DATABASE. X, Y, Z: COORDINATES IN THREE DIMENSIONS, DNN: DISTANCE BETWEEN NEAREST NEIGHBORS, GC: GEOMETRIC COMPLEXITY, AND TC: TEXTURAL COMPLEXITY.

Index	Name	Points	Xmin, Ymin, Zmin	Xmax	Ymax	Zmax	Min DNN	Max DNN	Description
a	Bag	1267845	0	879	1000	605	1	8.94	Daily supply, high TC
b	Banana	807184	0	828	1000	900	1	18.47	Fruit, low TC
c	Biscuits	952579	0	693	1000	631	1	10.05	Snack, thin, medium TC
d	Cake	2486566	0	1000	953	970	1	7.14	Snack, topological hole, medium TC
e	Cauliflower	1936627	0	1000	964	956	1	10.20	Vegetable, low TC
f	Flowerpot	2407154	0	896	1000	950	1	15.65	Container, thin wall, low TC
g	Glasses_case	716659	0	579	798	1000	1	24.37	Daily supply, high TC
h	Honeydew_melon	1431071	0	984	928	1000	1	29.97	Fruit, medium TC
i	House	1568490	0	717	1000	755	1	8.77	Crafts, high GC, high TC
j	Litchi	1039942	0	1000	510	550	1	2.45	Fruit, medium TC
k	Mushroom	1144603	0	1000	857	568	1	18.60	Vegetable, thin, different GC and TC on both sides
l	Pen_container	2878318	0	829	912	1000	1	23.58	Office supply, thin, high GC and different TC on both sides
m	Pineapple	1628910	0	733	949	1000	1	17.95	Fruit, high TC
n	Ping-pong_bat	703879	0	649	1000	400	1	5.10	Sports equipment, thin, different GC and TC on both sides
o	Puer_tea	412009	0	213	1000	230	1	6.71	Container, medium TC
p	Pumpkin	1340343	0	1000	934	756	1	3.74	Vegetable, high TC
q	Ship	684617	0	1000	288	375	1	3.61	Crafts, high GC, low TC
r	Statue	1637577	0	948	1000	819	1	52.20	Crafts, high GC, different TC on both sides
s	Stone	1086453	0	1000	815	586	1	75.77	Collection, high TC
t	Tool_box	1054211	0	1000	599	576	1	3.32	Container, low TC

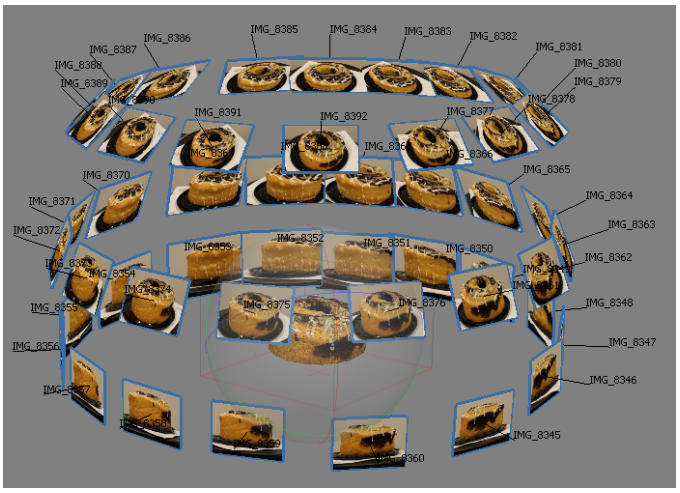


Fig. 3. Sample point cloud acquisition process.

In total, 760 point clouds with a wide range of visual quality levels are included in the WPC database.

Sampled distortion point clouds are shown in Fig. 4. It is interesting to observe that distorted point clouds not only exhibit loss of texture information similar to 2D images such as blockiness and blur, but also novel geometric distortion types. For example, hollow is caused by downsampling, where the point density may not be sufficient to cover the object surface. Holes and collapses arise from unsuccessful triangulations and inappropriate downsampling in G-PCC (Trisoup), respectively. Even when the triangulation is successful, geometric distortions may still appear as a consequence of ill-conditioned triangles. A sample case is given in the bottom right part of Fig. 4 (d). Moreover, a large 'geometryQP' in V-PCC may potentially introduce gaps and burrs. All these distortions are point cloud-specific, which create new challenges to objective PCQA models.

4 SUBJECTIVE EXPERIMENTS

4.1 Subjective User Study

We employ Technicolor Renderer [88] to render each point cloud to a video sequence. The rendering window, point size and point type parameters are set to 960×960, 1, and 'point', respectively. A horizontal and a vertical circles both with a radius of 5,000 are selected successively as the virtual camera path with the center of circles at the geometry center of an object. The remaining parameters are set as default. These settings preserve detail information as much as possible while maintaining the original point clouds to be watertight. One viewpoint is generated every two degrees on these circles, resulting in 360 image frames for each point cloud. Each distorted clip is then concatenated with its pristine counterpart into a 10-second video sequence for presentation. A screenshot is shown in Fig. 5.

Our subjective testing environment is the same as that for image acquisition. All video sequences are displayed on a 23.6" LCD monitor at a resolution of 1920×1080 with

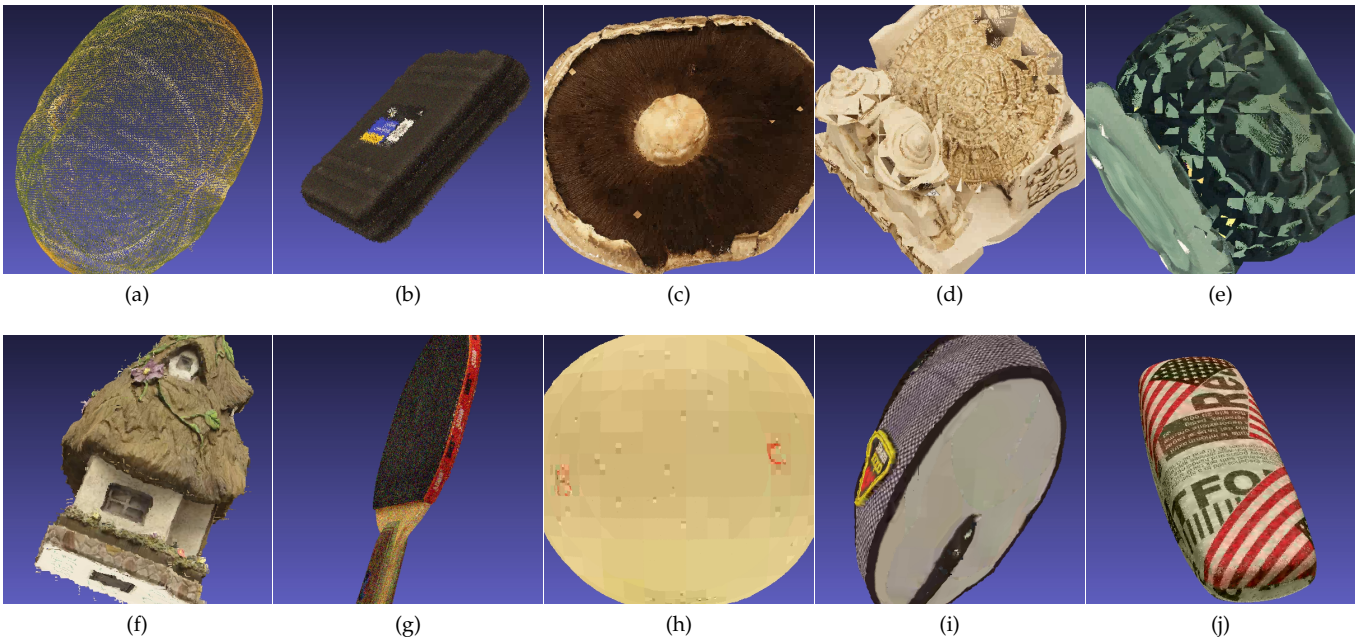


Fig. 4. Point cloud distortions. Geometry distortions: (a) Hollow. (b) Geometry noise. (c) Hole. (d) Shape distortion. (e) Collapse. (f) Gap and burr. Texture distortions: (g) Texture noise. (h) Blockiness. (i) Blur. (j) Color bleeding.

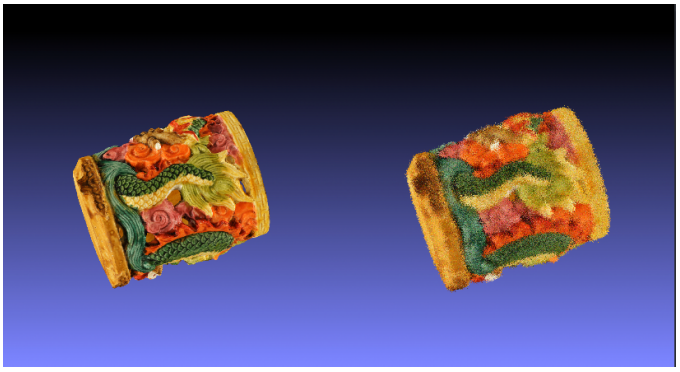
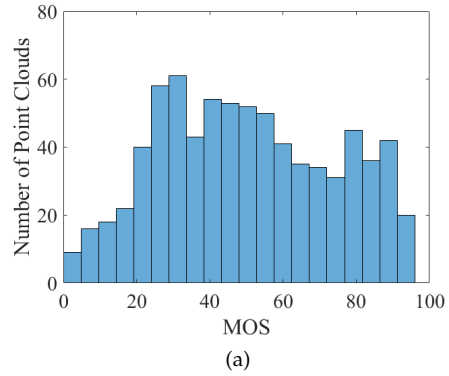


Fig. 5. Source and distorted point clouds of "PenContainer".



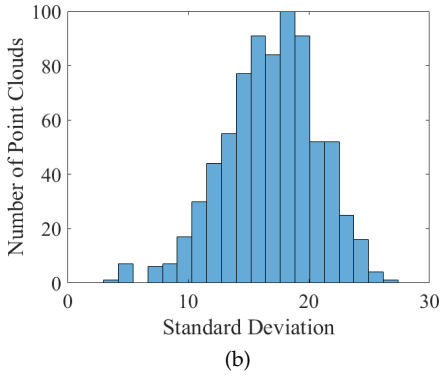


Fig. 6. MOS statistics of WPC database.

Truecolor (32bit) at 60 Hz. The monitor is calibrated in accordance with ITU-R Recommendation BT.500-13 [75]. DSIS methodology is applied in our subjective test [75]. Videos are displayed in random order using a customized graphical user interface, where subjective scores of individual viewers are recorded.

A total of 60 naïve subjects, including 32 males and 28 females aged between 21 and 40, participated in the subjective test. All subjects have normal or corrected-to-normal vision, and viewed videos from a distance of twice the screen height. Before the testing session, a training session is performed during which 18 videos different from those in the testing session are shown. The same methods are applied to generate videos used in both the training and testing sessions. Therefore, subjects knew what distortion types and levels to expect before the testing session, and thus the learning effects are kept minimal. Considering the limited testing capacity, each subject is assigned 10 objects in a circular fashion. Specifically, if subject i is assigned objects 1 to 10, then subject $i + 1$ watches objects 2 to 11. Each video is scored for 30 times, resulting in totally 22,800 subjective ratings, including 600 for source point clouds. For each subject, the whole study takes about 2 hours, which is divided into 4 sections with three 5-minute breaks in-between to minimize the impact of fatigue effect. For finer distinctions between ratings, 100-point continuous scale is utilized instead of a 5-point rating as in ITU-R ACR.

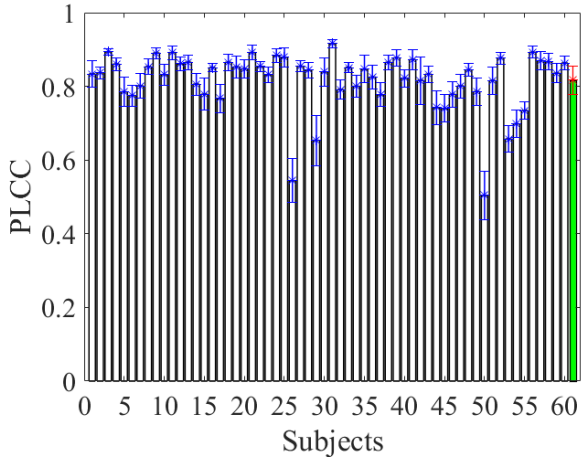


Fig. 7. PLCC and SRCC between individual subject rating and MOS. Rightmost column: performance of an average subject.

4.2 Subjective Data Analysis

After converting the raw subjective scores into Z-scores, an outlier removal scheme is applied [75]. No outlier detection is conducted participant-wise. Then the Z-scores are linearly rescaled to lie in the range of [0, 100]. The mean opinion score (MOS) for each distorted point cloud is calculated by averaging the re-scaled Z-scores from all valid subjects. The histograms for the MOS and the associated standard deviation are shown in Fig. 6, which demonstrates that the distorted point clouds span most of the quality range. Considering the MOS as the "ground truth", the performance of individual subjects can be evaluated by calculating the correlation coefficient between individual subject ratings and MOS values for each source point cloud, and then averaged over all source point clouds. Pearson linear correlation coefficient (PLCC) and Spearman rank-order correlation coefficient (SRCC) are employed as the evaluation criteria. Fig. 7 depicts the mean and standard deviation of each individual subject's performance, where most individual subjects perform quite consistently with relatively low variations across source point clouds. The average performance across all individual subjects is also given in the rightmost columns of Fig. 7.

4.3 Performance of Existing Objective PCQA Models

Using the aforementioned database, we test the performance of 13 PCQA models, which are selected to cover a wide range of design methodologies. Geometry distortion metrics except the MPEG metrics and algorithms unavailable to public are not included. The models include point-wise models: 1) point-to-point mean squared error-based PSNR ($\mathrm{PSNR}_{p2po,M}$) [26], [27], 2) point-to-point Hausdorff distance-based PSNR ($\mathrm{PSNR}_{p2po,H}$) [26], [27], 3) point-to-plane mean squared error-based PSNR ($\mathrm{PSNR}_{p2pl,M}$) [26], [27], 4) point-to-plane Hausdorff distance-based PSNR ($\mathrm{PSNR}_{p2pl,H}$) [26], [27], 5) point-to-point PSNR on color component (PSNR_{Y}) [28], [29], 6) PCM_{RR} [42], 7) PointSSIM [35], 8) PCQM [37], 9) GraghSIM [44]; and projection-based models: 10) projection-based PSNR (PSNR_{p}) [12], 11) projection-based structural similarity (SSIM_{p}) [12], [66], 12) projection-based multi-scale structural similarity ($\mathrm{MS\text{-}SSIM}_{p}$) [12], [79], and 13) projection-based pixel-domain visual information fidelity (VIFP_{p}) [12], [80]. The implementation of all models are obtained from the original authors or their public websites.

We use PLCC, SRCC and RMSE between MOSs and model predictions as the evaluation criteria, and the results are shown in Table 4, 5 and 6. First, it comes as no surprise that all geometry distortion models performs unfavorably to the geometry-plus-color PCQA models. Second, projection-based models, such as VIFP_{p}, provide the most promising results, but often fall short in making a distinction of the perceptual importance between the background and the regions corresponding to points in a 2D projection of a 3D point cloud.Third, even the best PCQA model only moderately correlates with human perception, leaving large space for improvement.

5 Objective Quality Assessment

5.1 Proposed PCQA Model

A point cloud can be omni-directionally inspected from a view-sphere at a given distance, while it is both cumbersome and unnecessary to use a large number of viewpoints when acquiring its 2D snapshots. Icosphere, a unit geodesic sphere created by subdividing a regular icosahedron with normalized vertices, are employed to generate viewpoints by uniformly distributed vertices [21], [89]. The number of vertices that may be generated is

$$N_v = 12 + 10\left(4^l - 1\right), \tag{1}$$

where l represents the subdivision level. For any point in a 3D point cloud, let $\mathbf{p} = (\mathbf{g}\ \mathbf{c})$ be a 6 dimensional row vector where $\mathbf{g}$ and $\mathbf{c}$ contain its 3D coordinates $(g_x\ g_y\ g_z)$ and attributed color information $(c_r\ c_g\ c_b)$, respectively. We use a series of transformations to obtain the projected images.

Firstly, we translate a point cloud to align its geometric center to the origin $(0\ 0\ 0)$. Specifically, for each $\mathbf{p}$

$$\mathbf{g}_t = \mathbf{g} - \mathbf{t}_r, \tag{2}$$

where $\mathbf{g}_t$ represents the 3D coordinates after translation, and $\mathbf{t}_r$ represents the translation vector equalling the geometric center coordinates of its corresponding reference point

TABLE 4
PLCC PERFORMANCE EVALUATION OF THE PROPOSED MODEL AGAINST EXISTING MODELS. ABSOLUTE PLCC ARE TAKEN FOR DISTORTION MEASURES FOR BETTER VISIBILITY.

Subset	Geometry distortion metric				Geometry-plus-color distortion metric									IW-SSIM$_p$
	PSNR$_{p2po,M}$	PSNR$_{p2po,H}$	PSNR$_{p2pl,M}$	PSNR$_{p2pl,H}$	PSNR$_Y$	PCM$_{RR}$	PointSSIM	PCQM	GraghSIM	PSNR$_p$	SSIM$_p$	MS-SSIM$_p$	VIFP$_p$	
Bag	0.7018	0.5116	0.6025	0.5487	0.8124	0.5947	0.5750	0.8658	0.7600	0.7849	0.8476	0.8641	0.8771	0.8480
Banana	0.7236	0.4086	0.5997	0.4086	0.7560	0.4797	0.1418	0.7145	0.5990	0.6337	0.7156	0.7812	0.7938	0.8724
Biscuits	0.5258	0.5197	0.5633	0.5203	0.7812	0.3520	0.4850	0.7798	0.7490	0.5318	0.6953	0.7563	0.7775	0.8904
Cake	0.4203	0.1327	0.2858	0.3577	0.5295	0.1209	0.1169	0.5832	0.4160	0.4848	0.6054	0.6096	0.6155	0.6743
Cauliflower	0.4555	0.2914	0.3483	0.2914	0.6332	0.4199	0.1865	0.7057	0.6010	0.4847	0.6515	0.6068	0.6581	0.8578
Flowerpot	0.7076	0.5271	0.6370	0.3816	0.6564	0.3224	0.3024	0.7030	0.6880	0.6779	0.8101	0.7963	0.8278	0.9368
GlassesCase	0.6028	0.5132	0.5141	0.4370	0.7861	0.3922	0.4921	0.8214	0.7100	0.7277	0.7963	0.8025	0.8089	0.8077
HoneydewMelon	0.4617	0.4337	0.4337	0.4337	0.7118	0.5654	0.5309	0.6539	0.7420	0.5291	0.7586	0.7543	0.8023	0.8989
House	0.6391	0.3956	0.4312	0.3792	0.7972	0.4016	0.4155	0.7537	0.7410	0.6668	0.8311	0.8241	0.8257	0.8347
Litchi	0.4291	0.3749	0.3472	0.3737	0.7201	0.4763	0.6147	0.7922	0.7260	0.6825	0.7685	0.8255	0.8545	0.9107
Mushroom	0.6406	0.4860	0.5456	0.4580	0.8022	0.2575	0.4410	0.8033	0.7120	0.5700	0.7994	0.8296	0.8450	0.8697
PenContainer	0.7782	0.5065	0.6688	0.5065	0.8132	0.5590	0.4916	0.8180	0.4080	0.8282	0.9183	0.9135	0.9153	0.9421
Pineapple	0.4678	0.2923	0.3719	0.2923	0.7466	0.3341	0.4556	0.7578	0.0410	0.5280	0.7214	0.7134	0.7456	0.7817
PingpongBat	0.7234	0.4191	0.6666	0.6320	0.8057	0.5132	0.5413	0.8600	-0.2480	0.5595	0.6601	0.7189	0.7922	0.9096
Pu'erTeaPot	0.3974	0.3688	0.3688	0.3688	0.8761	0.4041	0.4605	0.8400	-0.1240	0.8084	0.8155	0.8658	0.8897	0.9201
Pumpkin	0.5163	0.4919	0.4379	0.4919	0.6868	0.3401	0.3994	0.7462	0.2760	0.7250	0.8614	0.8606	0.8838	0.8976
Ship	0.7676	0.3848	0.6505	0.5535	0.7918	0.3976	0.4134	0.7578	-0.1990	0.7675	0.8201	0.8536	0.8791	0.9139
Statue	0.8208	0.4298	0.7011	0.4648	0.7579	0.2564	0.3585	0.7860	0.2970	0.8364	0.9228	0.9184	0.9244	0.9623
Stone	0.6140	0.5558	0.5161	0.5558	0.7882	0.3208	0.5102	0.8486	0.7850	0.7547	0.8282	0.8809	0.9233	0.8943
ToolBox	0.4485	0.2923	0.2846	0.2846	0.9039	0.4919	0.4776	0.7662	0.8050	0.4777	0.5469	0.5723	0.7141	0.8532
Downsampling	0.4247	0.5408	0.3323	0.4437	0.6368	0.6681	0.9545	0.8863	0.9330	0.6783	0.8529	0.9375	0.9700	0.9767
Gaussian noise	0.6867	0.6892	0.6867	0.6893	0.8706	0.7826	0.6743	0.9079	0.6200	0.8292	0.8213	0.8372	0.8467	0.9019
G-PCC (T)	0.4018	0.3029	0.4050	0.3405	0.6322	0.3553	0.5843	0.8075	0.5680	0.3291	0.6065	0.6545	0.8105	0.8154
V-PCC	0.1704	0.2156	0.2121	0.2866	0.3416	0.2805	0.3888	0.6563	0.4360	0.2903	0.3299	0.4397	0.7448	0.8419
G-PCC (O)	0	0	0	0	0.8067	0.6084	0.8183	0.8935	0.5720	0.7730	0.8258	0.8774	0.8950	0.8943
All	**0.4331**	**0.3425**	**0.3952**	**0.3412**	**0.6080**	**0.3775**	**0.3436**	**0.7486**	**0.4420**	**0.4989**	**0.6013**	**0.6701**	**0.7670**	**0.8504**

TABLE 5
SRCC PERFORMANCE EVALUATION OF THE PROPOSED MODEL AGAINST EXISTING MODELS. ABSOLUTE SRCC ARE TAKEN FOR DISTORTION MEASURES FOR BETTER VISIBILITY.

Subset	Geometry distortion metric				Geometry-plus-color distortion metric									IW-SSIM$_p$
	PSNR$_{p2po,M}$	PSNR$_{p2po,H}$	PSNR$_{p2pl,M}$	PSNR$_{p2pl,H}$	PSNR$_Y$	PCM$_{RR}$	PointSSIM	PCQM	GraghSIM	PSNR$_p$	SSIM$_p$	MS-SSIM$_p$	VIFP$_p$	
Bag	0.6669	0.4363	0.5751	0.4365	0.8051	0.6069	0.3236	0.8547	0.7320	0.7499	0.8438	0.8580	0.8725	0.8298
Banana	0.6471	0.1933	0.5691	0.2033	0.6785	0.5287	-0.0460	0.7686	0.5300	0.6759	0.7544	0.7968	0.7956	0.8627
Biscuits	0.5252	0.3085	0.4160	0.3368	0.7719	0.4130	0.4865	0.7945	0.7250	0.5299	0.6757	0.7195	0.7380	0.8900
Cake	0.3074	0.1724	0.1798	0.1796	0.5168	0.1503	0.0690	0.6043	0.4070	0.4365	0.5614	0.5602	0.5683	0.6598
Cauliflower	0.3501	0.0918	0.2058	0.1653	0.5927	0.4718	0.2224	0.6971	0.5520	0.4305	0.5967	0.5730	0.5820	0.8125
Flowerpot	0.6509	0.4348	0.5298	0.4515	0.6347	0.3058	0.3056	0.6984	0.6530	0.5963	0.7954	0.7776	0.8122	0.9211
GlassesCase	0.5845	0.2020	0.4390	0.3238	0.7826	0.3883	0.2288	0.8137	0.6830	0.7624	0.8269	0.8205	0.8201	0.7845
HoneydewMelon	0.4890	0.2768	0.3299	0.2300	0.6740	0.5742	0.4592	0.6439	0.7460	0.4512	0.7499	0.7207	0.7999	0.8954
House	0.5866	0.3429	0.4483	0.3434	0.7826	0.4905	0.2968	0.7845	0.7500	0.7119	0.8312	0.8246	0.8267	0.8196
Litchi	0.5109	0.3478	0.4291	0.3204	0.6496	0.4839	0.5026	0.7712	0.6840	0.6193	0.7231	0.8096	0.8556	0.8943
Mushroom	0.6396	0.3486	0.5156	0.3105	0.6550	0.2556	0.4113	0.7819	0.6730	0.5863	0.7297	0.8535	0.8658	0.8528
PenContainer	0.7720	0.2159	0.6688	0.3635	0.7963	0.6830	0.4059	0.8201	0.3760	0.8478	0.9372	0.9329	0.9334	0.9488
Pineapple	0.3777	0.1376	0.2785	0.1831	0.7217	0.4011	0.3267	0.7862	0.0310	0.5334	0.7193	0.7105	0.7285	0.7584
PingpongBat	0.5924	0.4958	0.4984	0.4357	0.7089	0.5526	0.4993	0.8224	-0.2280	0.5420	0.6785	0.7236	0.7947	0.8945
Pu'erTeaPot	0.6069	-0.1173	0.4746	-0.0384	0.8468	0.4308	0.3286	0.8528	-0.1210	0.7432	0.7636	0.8414	0.8637	0.9170
Pumpkin	0.4947	0.3092	0.3423	0.3068	0.6897	0.3241	0.2544	0.7802	0.1190	0.7347	0.8412	0.8497	0.8642	0.8831
Ship	0.7464	0.3404	0.6267	0.5158	0.7734	0.4400	0.3578	0.7793	-0.2050	0.7748	0.7847	0.8646	0.8829	0.9206
Statue	0.8040	0.2450	0.6707	0.4487	0.6968	0.1811	0.3390	0.7570	0.1730	0.7947	0.9030	0.9118	0.9059	0.9561
Stone	0.6219	0.3551	0.5129	0.3424	0.7115	0.3632	0.4924	0.8559	0.8040	0.6740	0.8303	0.8831	0.9203	0.8969
ToolBox	0.3937	0.1972	0.2969	0.1884	0.8760	0.5239	0.3364	0.8473	0.8170	0.4720	0.5889	0.6268	0.7093	0.8307
Downsampling	0.4815	0.5356	0.3251	0.4879	0.6172	0.7407	0.8478	0.8760	0.7650	0.5399	0.8039	0.8876	0.9212	0.9270
Gaussian noise	0.6155	0.6149	0.6194	0.6150	0.7895	0.7762	0.5931	0.8860	0.5760	0.6538	0.7509	0.7493	0.8067	0.8695
G-PCC (T)	0.3451	0.2811	0.3568	0.3085	0.6247	0.3044	0.5669	0.8212	0.4780	0.1968	0.6144	0.6572	0.8153	0.8203
V-PCC	0.1602	0.2051	0.1992	0.2370	0.3297	0.2966	0.3665	0.6431	0.2140	0.1998	0.3195	0.4213	0.7484	0.8458
G-PCC (O)	NaN	NaN	NaN	NaN	0.8100	0.6468	0.8234	0.8944	0.4260	0.7809	0.8391	0.8770	0.8976	0.8981
All	**0.4082**	**0.2578**	**0.3706**	**0.2883**	**0.5849**	**0.3603**	**0.3070**	**0.7471**	**0.4360**	**0.4601**	**0.6138**	**0.6656**	**0.7689**	**0.8481**

cloud. The reason $\mathbf{t}_r$ is used instead of the geometric center of a distorted point cloud is that geometric distortions may change the upper and lower bounds of the 3D coordinates, leading to misalignment of the projected images.

Secondly, we rotate the point cloud to obtain a number of viewpoints. More specifically, let $\mathbf{n}_v$, a 3 dimensional row vector, be the unit normal $\mathbf{n}_z$ be $(0\ 0\ 1)$, then the rotation vector $(\mathbf{r}\ \theta)$ can be calculated as

$$\mathbf{r} = \frac{\mathbf{n}_v \times \mathbf{n}_z}{\|\mathbf{n}_v \times \mathbf{n}_z\|} \tag{3}$$

and

$$\theta = \arccos(\mathbf{n}_v \cdot \mathbf{n}_z), \tag{4}$$

where $\|\cdot\|$ denotes the l^2 norm of a vector, $\mathbf{r}$ is the rotation axis, and θ is the axial angle. The rotation matrix $\mathbf{R}$ is obtained using $(\mathbf{r}\ \theta)$. Then we use $\mathbf{R}$ to calculate $\mathbf{g}_r$, the 3D coordinates after rotation, for each $\mathbf{p}$,

$$\mathbf{g}_r = \mathbf{g}_t\mathbf{R}. \tag{5}$$

Thirdly, a scaling transformation is applied to make 2D snapshots of all reference point clouds approximately watertight while keeping the details as much as possible. For each $\mathbf{p}$, this operation can be expressed as

$$\mathbf{g}_s = s \cdot \mathbf{g}_r, \tag{6}$$

where $\mathbf{g}_s$ represents 3D coordinates after scaling and s is a scaling factor. Since the values of the coordinates are rounded to integer numbers, for which the maximum rounding error is bounded by half of the pixel spacing, the default value of s is set to $1/2$. Empirically, we also find this leads to the best performance.

Fourthly, we use orthogonal projection [12] and rasterization to obtain a projected image. For each $\mathbf{p}$, the projected

TABLE 6
RMSE PERFORMANCE EVALUATION OF THE PROPOSED MODEL AGAINST EXISTING MODELS.

Subset	Geometry distortion metric				Geometry-plus-color distortion metric									IW-SSIM$_p$
	PSNR$_{p2po,M}$	PSNR$_{p2po,H}$	PSNR$_{p2pl,M}$	PSNR$_{p2pl,H}$	PSNR$_Y$	PCM$_{RR}$	PointSSIM	PCQM	GraghSIM	PSNR$_p$	SSIM$_p$	MS-SSIM$_p$	VIFP$_p$	
Bag	16.72	20.17	18.73	19.63	13.69	23.47	19.20	11.74	15.26	14.54	12.46	11.81	11.27	12.44
Banana	14.98	19.97	17.36	19.81	14.20	21.70	21.48	15.18	17.38	16.79	15.16	13.54	13.20	10.61
Biscuits	19.68	19.44	18.80	19.43	14.20	22.75	19.90	14.25	15.09	19.27	16.35	14.88	14.31	10.35
Cake	20.55	22.31	21.57	21.02	19.10	22.51	22.36	18.29	20.47	19.69	17.92	17.84	17.74	16.62
Cauliflower	19.95	21.44	21.01	21.44	17.35	22.41	22.02	15.88	17.91	19.60	17.00	17.82	16.88	11.52
Flowerpot	16.81	20.24	18.33	21.98	17.94	23.78	22.67	16.92	17.26	17.48	13.95	14.39	13.34	8.322
GlassesCase	18.06	19.68	19.41	20.36	13.99	22.63	19.72	12.91	15.94	15.52	13.69	13.50	13.31	13.34
HoneydewMelon	20.99	21.36	21.34	21.33	16.62	23.66	20.05	17.90	15.85	20.08	15.42	15.53	14.12	10.37
House	17.88	21.36	20.98	21.76	14.04	23.25	21.15	15.29	15.61	17.33	12.93	13.17	13.12	12.81
Litchi	20.98	21.53	21.78	21.54	16.11	23.22	18.32	14.17	15.96	16.97	14.86	13.11	12.06	9.595
Mushroom	17.02	19.37	18.58	19.71	13.23	22.17	19.89	13.20	15.56	18.21	13.32	12.38	11.85	10.94
PenContainer	14.76	20.26	17.47	20.26	13.68	23.50	20.46	13.53	21.46	13.17	9.306	9.563	9.465	7.879
Pineapple	18.04	19.52	18.95	19.53	13.58	20.41	18.17	13.32	20.20	17.34	14.14	14.30	13.61	12.73
PingpongBat	15.69	20.64	16.94	17.61	13.46	22.73	19.11	11.60	22.73	18.84	17.07	15.80	13.87	9.442
Pu'erTeaPot	21.75	22.15	22.05	22.11	11.43	23.71	21.04	12.86	23.71	13.96	13.72	11.86	10.82	9.288
Pumpkin	18.63	18.94	19.56	18.94	15.81	21.75	19.94	14.48	20.91	14.98	11.05	11.08	10.18	9.591
Ship	15.44	22.23	18.30	20.06	14.71	24.09	21.94	15.72	24.09	15.44	13.78	12.55	11.48	9.780
Statue	13.24	20.93	16.53	20.53	15.12	23.18	21.64	14.33	22.19	12.71	8.932	9.170	8.841	6.309
Stone	16.85	18.48	18.28	17.76	13.14	21.35	18.36	11.29	13.23	14.00	11.96	10.10	8.201	9.551
ToolBox	18.65	19.95	20.01	20.00	8.923	20.86	18.33	13.41	12.37	18.33	17.47	17.11	14.61	10.88
Downsampling	26.93	25.02	28.07	26.66	22.93	29.74	8.870	13.79	10.68	21.86	15.53	10.35	7.229	6.384
Gaussian noise	14.15	14.11	14.15	14.10	9.576	19.47	14.38	8.162	15.28	10.88	11.11	10.65	10.36	8.408
G-PCC (T)	22.59	23.51	22.56	23.20	19.11	24.67	20.02	14.64	20.30	23.30	19.61	18.65	14.45	14.28
V-PCC	16.76	16.61	16.62	16.29	15.99	17.01	15.67	12.84	15.30	16.28	16.06	15.28	11.35	9.179
G-PCC (O)	21.36	21.36	21.36	21.36	12.62	21.36	12.28	9.593	17.53	13.55	12.05	10.25	9.531	9.559
All	**20.66**	**21.54**	**21.06**	**21.55**	**18.20**	**22.92**	**21.53**	**15.20**	**20.56**	**19.87**	**18.32**	**17.02**	**14.71**	**12.06**

coordinates are given by

$$\widetilde{\mathbf{g}} = \mathbf{g}_s \mathbf{P}, \tag{7}$$

where

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{8}$$

represents an orthogonal projection matrix. If there are multiple points occupying the same location $(\widetilde{g}_x\ \widetilde{g}_y)$, the point with the largest value of g_z is maintained. To implement rasterization, we put **c** of all maintained points into their corresponding projected positions on a projection plane filled with **c** of $(127\ 127\ 127)$, and we obtain a projected image denoted as **I**.

Finally, we perform the above operations on each distorted point cloud PC_{dis} and its reference PC_{ref} to obtain $\mathbf{I}_{dis}(n)$ and $\mathbf{I}_{ref}(n)$ at the nth viewpoint of icosphere, respectively.

$\mathbf{I}_{dis}(n)$ and $\mathbf{I}_{ref}(n)$ have identical background pixels, and thus, the similarity between $\mathbf{I}_{dis}(n)$ and $\mathbf{I}_{ref}(n)$ is larger than that of PC_{dis} and PC_{ref}. Consequently, it is useful to remove the influence of background. Also note that in a 2D projection of a 3D point cloud, the perceptual importance of different regions changes significantly over space. In particular, the background region contains no information about the point clouds, and the importance of the other regions could also vary. The principle of information content weighted pooling [69] provides an excellent framework to account for such variations in importance, as exemplified by Fig. 8. Therefore, we propose to assess the perceptual quality of PC_{dis}, by

$$Q(PC_{dis}) = \sum_{n=1}^{N_v} \text{IW-SSIM}(\mathbf{I}_{ref}(n), \mathbf{I}_{dis}(n)), \tag{9}$$

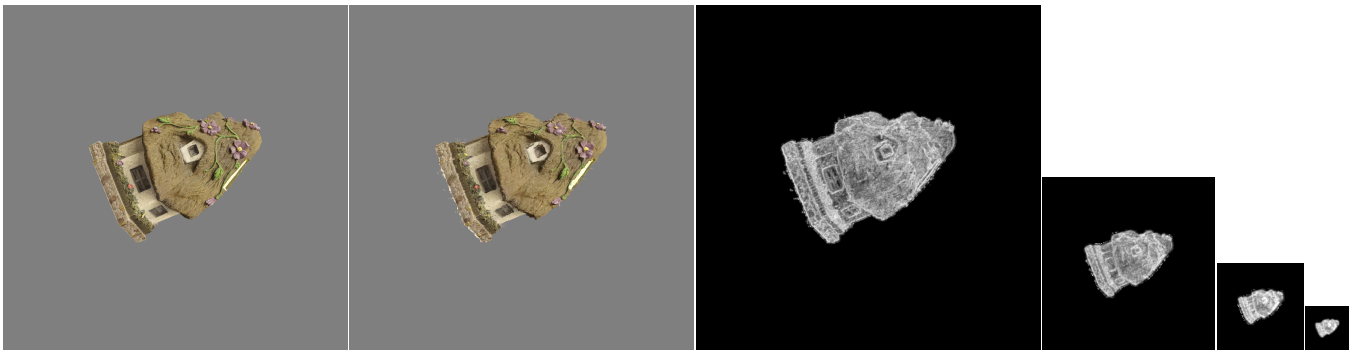
where the IW-SSIM evaluations [69] between all pairs of $\mathbf{I}_{ref}(n)$ and $\mathbf{I}_{dis}(n)$ are averaged to an overall quality measure of PC_{dis}, and we name the proposed method IW-SSIM$_p$, where all default parameters of IW-SSIM [69] in still image quality assessment are inherited.

5.2 Validation and Discussion

We validate the proposed IW-SSIM$_p$ model using the WPC database presented in Section 4 and compare its performance against existing objective PCQA models. Note that IW-SSIM$_p$ does not involve a training process and is independent of any existing PCQA databases including the WPC database. Tables 4, 5 and 6 summarize the PLCC, SRCC and RMSE evaluation results. We find that the performance of IW-SSIM$_p$ when $Nv = 12, 42, 162$ are very close to each other, while the computational complexity is proportional to Nv. Therefore we use $Nv = 12$ in all results reported here. It can be seen that the proposed model delivers the best performance in predicting subjective quality of 3D point cloud not only on the whole database but also on almost every subset. In addition, its PLCC and SRCC performance is at the same level as compared to an average human subject as given in Fig.7.

To ascertain that the improvement of the proposed model is statistically significant, we carry out a statistical significance analysis by following the approach introduced in [90]. First, a nonlinear regression function is applied to map the objective quality scores to predict the subjective scores. We observe that the prediction residuals all have zero-mean, and thus the model with lower variance is generally considered better. We conduct a hypothesis testing using F-statistics. Since the number of samples exceeds 50, the Gaussian assumption of the residuals approximately hold based on the central limit theorem [91]. The test statistic is the ratio of variances. The null hypothesis is that the prediction residuals from one quality model come from the same distribution and are statistically indistinguishable (with 95% confidence) from the residuals from another model. We compare every possible pair of objective models. The results are summarized in Table 7, where a symbol "1" means that the row model performs significantly better than the column model, a symbol "0" means the opposite, and a symbol "-" indicates that the row and column models are statistically indistinguishable.

There are several useful findings from the statistical sig-



(a) (b) (c) (d) (e) (f)

Fig. 8. Local information content maps. (a), (b) Snapshots of "House" and one of its distorted versions. (c), (d), (e), (f) Information content maps computed at four scales.

TABLE 7
STATISTICAL SIGNIFICANCE COMPARISON MATRIX BASED ON QUALITY PREDICTION RESIDUALS. A SYMBOL "1" MEANS THAT THE PERFORMANCE OF THE ROW MODEL IS STATISTICALLY BETTER THAN THAT OF THE COLUMN MODEL, A SYMBOL "0" MEANS THAT THE ROW MODEL IS STATISTICALLY WORSE, AND A SYMBOL "-" MEANS THAT THE ROW AND COLUMN MODELS ARE STATISTICALLY INDISTINGUISHABLE.

	$PSNR_{p2po,M}$	$PSNR_{p2po,H}$	$PSNR_{p2pl,M}$	$PSNR_{p2pl,H}$	$PSNR_Y$	PCM_{RR}	PointSSIM	PCQM	GraghSIM	$PSNR_p$	$SSIM_p$	MS-$SSIM_p$	$VIFP_p$	IW-$SSIM_p$
$PSNR_{p2po,M}$	-	-	-	-	0	-	-	0	-	-	0	0	0	0
$PSNR_{p2po,H}$	-	-	-	-	0	-	-	0	-	0	0	0	0	0
$PSNR_{p2pl,M}$	-	-	-	-	0	-	-	0	-	0	0	0	0	0
$PSNR_{p2pl,H}$	-	-	-	-	0	-	-	0	-	0	0	0	0	0
$PSNR_Y$	1	1	1	1	-	1	1	0	1	-	-	0	0	0
PCM_{RR}	-	-	-	-	0	-	-	0	-	0	0	0	0	0
PointSSIM	-	-	-	-	0	-	-	0	-	0	0	0	0	0
PCQM	1	1	1	1	1	1	1	-	1	1	1	1	-	0
GraphSIM	-	-	-	-	0	-	-	0	-	-	0	0	0	0
$PSNR_p$	-	1	1	1	-	1	1	0	-	-	-	0	0	0
$SSIM_p$	1	1	1	1	-	1	1	0	1	-	-	0	0	0
MS-$SSIM_p$	1	1	1	1	1	1	1	0	1	1	1	-	0	0
$VIFP_p$	1	1	1	1	1	1	1	-	1	1	1	1	-	0
IW-$SSIM_p$	1	1	1	1	1	1	1	1	1	1	1	1	1	-

nificance analysis. First, existing geometry distortion metrics are statistically indistinguishable from each other. Second, most geometry-plus-color metrics are statistically better than geometry distortion metrics. Third, the proposed IW-$SSIM_p$ model is statistically better than all existing models.

Finally, to investigate the generalization potential of the proposed model, we compared our model with the well-known state-of-the-art GraphSIM [44], PointSSIM [35], PCQM [37], PCM_{RR} [42] and NR-3DQA [62] on popular databases. Table 8, 9 and 10 depict the results of PLCC, SRCC and RMSE. The numbers in parentheses of Table 10 represent the score range of each database. We can draw the following conclusions. First, IW-$SSIM_p$ performs well on all databases except the IRPC database. One plausible explanation may be that IW-$SSIM_p$ is sensitive to the texture distortion, while the IRPC database only contains point clouds with geometry distortion. Second, PCQM and NR-3DQA are the best performers on the SJTU-PCQA database, meanwhile, GraghSIM has the best performance on the IRPC and M-PCCD database. Third, IW-$SSIM_p$ can achieve a higher prediction accuracy than the state-of-the-art GraphSIM model, with a much lower computational complexity.

TABLE 8
PLCC PERFORMANCE EVALUATION ON THE BENCHMARK DATABASES.

	GraghSIM	PointSSIM	PCQM	PCM_{RR}	NR-3DQA	IW-$SSIM_p$
SJTU-PCQA [45]	0.5910	0.7503	-0.8565	-0.5129	**0.8599**	0.7949
IRPC [24]	**0.8603**	0.5939	-0.1850	-0.0540	0.4080	0.0911
ICIP2020 [48]	0.8601	0.6758	-0.2634	-0.7233	0.4848	**0.9097**
M-PCCD [23]	**0.9428**	0.8519	-0.6070	-0.5535	0.5362	0.7172
WPC	0.4420	0.3436	-0.7486	-0.3775	0.3639	**0.8504**

TABLE 9
SRCC PERFORMANCE EVALUATION ON THE BENCHMARK DATABASES.

	GraghSIM	PointSSIM	PCQM	PCM_{RR}	NR-3DQA	IW-$SSIM_p$
SJTU-PCQA [45]	0.5710	0.7346	**-0.8439**	-0.5366	0.8336	0.7833
IRPC [24]	**0.7469**	0.5054	-0.4170	-0.2345	0.2849	0.1339
ICIP2020 [48]	0.8449	0.5638	-0.4108	-0.8868	0.4376	**0.8968**
M-PCCD [23]	**0.9535**	0.8328	-0.9155	-0.8885	0.5919	0.7487
WPC	0.4360	0.3070	-0.7471	-0.3603	0.3078	**0.8481**

6 CONCLUSION

In this work, we tackle the problem of 3D point cloud quality assessment. Our major contributions are fourfold. First, we construct 20 high quality, realistic and omni-directional dense point clouds with a wide range of geometric and textural complexity, which are voxelized with an average number of 1.35M points and a standard deviation of 656K, respectively. These point clouds provide fertile ground for PC processing and PCQA research. Second, we construct so far the largest point cloud database of diverse contents and distortion types and conduct a lab-controlled subjective user study. The new WPC database contains 740 distorted point clouds with MOS values approximately evenly distributed from poor to excellent perceived quality levels. Third, we conduct a comprehensive evaluation on existing objective PCQA models. Fourth, we propose a projection-based IW-$SSIM_p$ model that significantly outperforms existing objective PCQA methods.

There are several research directions that are worth future investigation. It should be noted that the all projection-

TABLE 10
RMSE PERFORMANCE EVALUATION ON THE BENCHMARK DATABASES.

	GraghSIM	PointSSIM	PCQM	PCM_{RR}	NR-3DQA	$IW\text{-}SSIM_p$
SJTU-PCQA [45] (10)	1.8910	1.5499	1.2169	2.3442	**1.1966**	1.4224
IRPC [24] (5)	**0.5207**	0.7944	0.9703	0.9874	0.9015	0.9833
ICIP2020 [48] (5)	0.5794	0.8373	1.1360	1.1360	0.9935	**0.4718**
M-PCCD [23] (5)	**0.4535**	0.7124	1.3603	1.3603	1.1482	0.9480
WPC (100)	20.5640	21.5273	15.1996	22.9234	21.3515	**12.0620**

based PCQA models so far including the proposed one is limited to outside views of the point clouds. One possible solution for PCQA of indoor scenes may be obtained by projecting the point cloud onto a panoramic image and then applying traditional image quality assessment models. Another promising direction is to leverage deep learning methods for PCQA research. State-of-the-art data-driven methods have demonstrated strong promises in full-reference [92] and no-reference image quality assessment [93], [94]. How to extend these works to PCQA is an interesting direction yet to be explored.

ACKNOWLEDGMENTS

This work was supported in part by Natural Sciences and Engineering Research Council of Canada, in part by the National Natural Science Foundation of China under Grants (61772294, 62172259), in part by Shandong Provincial Natural Science Foundation of China under Grant (ZR2018PF002, ZR2021MF025), in part by the open project program of state key laboratory of virtual reality technology and systems, Beihang University, under Grant VRLAB2021A01, in part by the Joint funding for smart computing of Shandong Natural Science Foundation of China under Grant ZR2019LZH002, and jointly completed by the State Key Laboratory of High Performance Server and Storage Technology, Inspur Group, Jinan, China.

REFERENCES

[1] S. Ye, D. Chen, S. Han, Z. Wan, and J. Liao, "Meta-pu: An arbitrary-scale upsampling network for point cloud," *IEEE Trans. Visualization and Computer Graphics*, 2021.

[2] H. Chen, M. Wei, Y. Sun, X. Xie, and J. Wang, "Multi-patch collaborative point cloud denoising via low-rank recovery with graph constraint," *IEEE Trans. Visualization and Computer Graphics*, vol. 26, no. 11, pp. 3255–3270, 2019.

[3] D. Zhang, X. Lu, H. Qin, and Y. He, "Pointfilter: Point cloud filtering via encoder-decoder modeling," *IEEE Trans. Visualization and Computer Graphics*, vol. 27, no. 3, pp. 2015–2027, 2020.

[4] X. Zhao, B. Zhang, J. Wu, R. Hu, and T. Komura, "Relationship-based point cloud completion," *IEEE Trans. Visualization and Computer Graphics*, 2021.

[5] 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*, 2020.

[6] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, and H. Yang, "Reduced reference perceptual quality model with application to rate control for video-based point cloud compression," *IEEE Trans. Image Processing*, vol. 30, pp. 6623–6636, 2021.

[7] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Subjective and objective quality evaluation of 3D point cloud denoising algorithms," in *Proc. IEEE Int. Conf. Multimedia and Expo Workshops*. IEEE, 2017, pp. 1–6.

[8] E. Alexiou and T. Ebrahimi, "On subjective and objective quality evaluation of point cloud geometry," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2017, pp. 1–3.

[9] E. Alexiou, T. Ebrahimi, M. Bernardo, M. Pereira, A. Pinheiro, L. da Silva Cruz, C. Duarte, L. Dmitrovic, E. Dumic, D. Matkovic, and A. Skodras, "Point cloud subjective evaluation methodology based on 2D rendering," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2018, pp. 1–6.

[10] E. Alexiou, A. Pinheiro, C. Duarte, D. Matkovic, E. Dumic, L. da Silva Cruz, L. Dmitrovic, M. Bernardo, M. Pereira, and T. Ebrahimi, "Point cloud subjective evaluation methodology based on reconstructed surfaces," in *Proc. SPIE Optical Engineering+Applications*. SPIE, 2018, pp. 107 520H.1–107 520H.14.

[11] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Subjective and objective quality evaluation of compressed point clouds," in *Proc. IEEE Int. Workshop on Multimedia Signal Processing*. IEEE, 2017, pp. 1–6.

[12] E. M. Torlig, E. Alexiou, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, "A novel methodology for quality assessment of voxelized point clouds," in *Applications of Digital Image Processing XLI*, vol. 10752. International Society for Optics and Photonics, 2018, p. 107520I.

[13] E. Zerman, P. Gao, C. Ozcinar, and A. Smolic, "Subjective and objective quality assessment for volumetric video compression," in *XVI Proc. IS&T Int. Symposium on Electronic Imaging: Image Quality and System Performance*. Society for Imaging Science and Technology IS&T, 2019, pp. 323:1–323:6.

[14] E. Alexiou and T. Ebrahimi, "On the performance of metrics to predict quality in point cloud representations," in *Applications of Digital Image Processing XL*, vol. 10396. Int. Society for Optics and Photonics, 2017, pp. 103 961H:301–103 961H:317.

[15] E. Alexiou, E. Upenik, and T. Ebrahimi, "Towards subjective quality assessment of point cloud imaging in augmented reality," in *Proc. IEEE Int. Workshop on Multimedia Signal Processing*. IEEE, 2017, pp. 1–6.

[16] E. Alexiou and T. Ebrahimi, "Point cloud quality assessment metric based on angular similarity," in *Proc. IEEE Int. Conf. Multimedia and Expo*. IEEE, 2018, pp. 1–6.

[17] ——, "Impact of visualization strategy for subjective quality assessment of point clouds," in *Proc. IEEE Int. Conf. Multimedia and Expo Workshops*. IEEE, 2018, pp. 1–6.

[18] ——, "Benchmarking of objective quality metrics for colorless point clouds," in *Proc. IEEE Picture Coding Symposium*. IEEE, 2018, pp. 51–55.

[19] J. Zhang, W. Huang, X. Zhu, and J.-N. Hwang, "A subjective quality evaluation for 3D point cloud models," in *Proc. IEEE Int. Conf. Audio, Language and Image Processing*. IEEE, 2014, pp. 827–831.

[20] Y. Nehmé, J.-P. Farrugia, F. Dupont, P. LeCallet, and G. Lavoué, "Comparison of subjective methods, with and without explicit reference, for quality assessment of 3D graphics," in *ACM Symposium on Applied Perception*. ACM, 2019, pp. 1–9.

[21] E. Alexiou and T. Ebrahimi, "Exploiting user interactivity in quality assessment of point cloud imaging," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2019, pp. 1–6.

[22] L. A. da Silva Cruz, E. Dumić, E. Alexiou, J. Prazeres, R. Duarte, M. Pereira, A. Pinheiro, and T. Ebrahimi, "Point cloud quality evaluation: Towards a definition for test conditions," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2019, pp. 1–6.

[23] 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. on Signal and Information Processing*, vol. 8, 2019.

[24] A. Javaheri, C. Brites, F. M. B. Pereira, and J. M. Ascenso, "Point cloud rendering after coding: Impacts on subjective and objective quality," *IEEE Trans. Multimedia*, 2020.

[25] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Geometric distortion metrics for point cloud compression," in *Proc. IEEE Int. Conf. Image Processing*, 2017, pp. 3460–3464.

[26] ——, "Evaluation metrics for point cloud compression," *ISO/IEC JTC1/SC29/WG11 MPEG, M39966*, 2017.

[27] ——, "Updates and integration of evaluation metric software for PCC," *ISO/IEC JTC1/SC29/WG11 MPEG, M40522*, 2017.

[28] R. Mekuria, K. Blom, and P. Cesar, "Design, implementation, and evaluation of a point cloud codec for tele-immersive video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 27, no. 4, pp. 828–842, April 2017.

[29] R. Mekuria, Z. Li, C. Tulvan, and P. Chou, "Evaluation criteria for PCC (point cloud compression)," *ISO/IEC JTC1/SC29/WG11 MPEG, N16332*, 2016.

[30] G. Meynet, J. Digne, and G. Lavoué, "PC-MSDM: A quality metric for 3D point clouds," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2019, pp. 1–3.

[31] E. Dumic, C. R. Duarte, and L. A. da Silva Cruz, "Subjective evaluation and objective measures for point clouds—state of the art," in *Int. Colloquium on Smart Grid Metrology (SmaGriMet)*. IEEE, 2018, pp. 1–5.

[32] H. Su, Z. Duanmu, W. Liu, Q. Liu, and Z. Wang, "Perceptual quality assessment of 3D point clouds," in *Proc. IEEE Int. Conf. Image Processing*. IEEE, 2019, pp. 3182–3186.

[33] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "A generalized hausdorff distance based quality metric for point cloud geometry," *arXiv preprint arXiv:2003.13669*, 2020.

[34] E. Zerman, C. Ozcinar, P. Gao, and A. Smolic, "Textured mesh vs coloured point cloud: A subjective study for volumetric video compression," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2020, pp. 1–6.

[35] E. Alexiou and T. Ebrahimi, "Towards a point cloud structural similarity metric," in *Proc. IEEE Int. Conf. Multimedia and Expo Workshops*. IEEE, 2020, pp. 1–6.

[36] E. Alexiou, N. Yang, and T. Ebrahimi, "PointXR: A toolbox for visualization and subjective evaluation of point clouds in virtual reality," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*, no. CONF, 2020.

[37] G. Meynet, Y. Nehmé, J. Digne, and G. Lavoué, "PCQM: A full-reference quality metric for colored 3D point clouds," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*, 2020.

[38] I. Viola, S. Subramanyam, and P. César, "A color-based objective quality metric for point cloud contents," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*, 2020.

[39] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Improving PSNR-based quality metrics performance for point cloud geometry," in *Proc. IEEE Int. Conf. Image Processing*. IEEE, 2020, pp. 3438–3442.

[40] ——, "Mahalanobis based point to distribution metric for point cloud geometry quality evaluation," *IEEE Signal Processing Letters*, vol. 27, pp. 1350–1354, 2020.

[41] E. Dumic, F. Battisti, M. Carli, and L. A. da Silva Cruz, "Point cloud visualization methods: a study on subjective preferences," in *28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 595–599.

[42] I. Viola and P. Cesar, "A reduced reference metric for visual quality evaluation of point cloud contents," *IEEE Signal Processing Letters*, 2020.

[43] K. Cao, Y. Xu, and P. Cosman, "Visual quality of compressed mesh and point cloud sequences," *IEEE Access*, vol. 8, pp. 171 203–171 217, 2020.

[44] Q. Yang, Z. Ma, Y. Xu, Z. Li, and J. Sun, "Inferring point cloud quality via graph similarity," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2020.

[45] 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*, 2020.

[46] R. Diniz, P. G. Freitas, and M. C. Farias, "Multi-distance point cloud quality assessment," in *Proc. IEEE Int. Conf. Image Processing*. IEEE, 2020, pp. 3443–3447.

[47] ——, "Towards a point cloud quality assessment model using local binary patterns," in *Proc. IEEE Int. Conf. Quality of Multimedia Experience*. IEEE, 2020, pp. 1–6.

[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 Processing*. IEEE, 2020, pp. 3428–3432.

[49] L. Hua, M. Yu, Z. He, R. Tu, and G. Jiang, "CPC-GSCT: Visual quality assessment for coloured point cloud based on geometric segmentation and colour transformation," *IET Image Processing*, 2021.

[50] R. Diniz, P. G. Freitas, and M. Farias, "A novel point cloud quality assessment metric based on perceptual color distance patterns," *Electronic Imaging*, vol. 2021, no. 9, pp. 256–1, 2021.

[51] R. Diniz, P. G. Freitas, and M. C. Farias, "Local luminance patterns for point cloud quality assessment," in *Proc. IEEE Int. Workshop on Multimedia Signal Processing*. IEEE, 2020, pp. 1–6.

[52] X. Wu, Y. Zhang, C. Fan, J. Hou, and S. Kwong, "Subjective quality database and objective study of compressed point clouds with 6dof head-mounted display," *IEEE Trans. Circuits and Systems for Video Technology*, 2021.

[53] Z. He, G. Jiang, Z. Jiang, and M. Yu, "Towards a colored point cloud quality assessment method using colored texture and curvature projection," in *Proc. IEEE Int. Conf. Image Processing*. IEEE, 2021, pp. 1444–1448.

[54] R. Diniz, M. Q. Farias, and P. Garcia-Freitas, "Color and geometry texture descriptors for point-cloud quality assessment," *IEEE Signal Processing Letters*, 2021.

[55] L. Hua, M. Yu, G. Jiang, Z. He, and Y. Lin, "VQA-CPC: a novel visual quality assessment metric of color point clouds," in *Optoelectronic Imaging and Multimedia Technology VII*, vol. 11550. International Society for Optics and Photonics, 2020, p. 1155012.

[56] Y. Liu, Q. Yang, Y. Xu, and L. Yang, "Point cloud quality assessment: Large-scale dataset construction and learning-based no-reference approach," *arXiv preprint arXiv:2012.11895*, 2020.

[57] L. Hua, G. Jiang, M. Yu, and Z. He, "BQE-CVP: Blind quality evaluator for colored point cloud based on visual perception," in *2021 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. IEEE, 2021, pp. 1–6.

[58] W.-x. Tao, G.-y. Jiang, Z.-d. Jiang, and M. Yu, "Point cloud projection and multi-scale feature fusion network based blind quality assessment for colored point clouds," in *Proc. ACM Int. Conf. Multimedia*, 2021, pp. 5266–5272.

[59] Y. Xu, Q. Yang, L. Yang, and J.-N. Hwang, "EPES: Point cloud quality modeling using elastic potential energy similarity," *IEEE Trans. Broadcasting*, 2021.

[60] Y. Zhang, Q. Yang, and Y. Xu, "Ms-graphsim: Inferring point cloud quality via multiscale graph similarity," in *Proc. ACM Int. Conf. Multimedia*, 2021, pp. 1230–1238.

[61] Q. Liu, H. Yuan, H. Su, H. Liu, Y. Wang, H. Yang, and J. Hou, "PQA-Net: Deep no reference point cloud quality assessment via multi-view projection," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 31, no. 12, pp. 4645–4660, 2021.

[62] Z. Zhang, W. Sun, X. Min, T. Wang, W. Lu, and G. Zhai, "No-reference quality assessment for 3d colored point cloud and mesh models," *arXiv preprint arXiv:2107.02041*, 2021.

[63] J. Zeng, G. Cheung, M. Ng, J. Pang, and C. Yang, "3D point cloud denoising using graph laplacian regularization of a low dimensional manifold model," *IEEE Trans. Image Processing*, December 2019.

[64] H. Wu, J. Zhang, and K. Huang, "Point cloud super resolution with adversarial residual graph networks," *arXiv preprint arXiv:1908.02111*, 2019.

[65] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li *et al.*, "Emerging MPEG standards for point cloud compression," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 133–148, 2018.

[66] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.

[67] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[68] Y. Nehmé, F. Dupont, J.-P. Farrugia, P. Le Callet, and G. Lavoué, "Visual quality of 3D meshes with diffuse colors in virtual reality: Subjective and objective evaluation," *IEEE Trans. Visualization and Computer Graphics*, 2020.

[69] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Processing*, vol. 20, no. 5, pp. 1185–1198, May 2011.

[70] J. Gutiérrez, T. Vigier, and P. L. Callet, "Quality evaluation of 3D objects in mixed reality for different lighting conditions," *Electronic Imaging*, vol. 2020, no. 11, pp. 128–1, 2020.

[71] G. Turk and M. Levoy, "Zippered polygon meshes from range images," in *Proc. ACM SIGGRAPH*. ACM, 1994, pp. 311–318.

[72] MPEG, "MPEG point cloud datasets," http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/datasets, 2017 (accessed May 28, 2018).

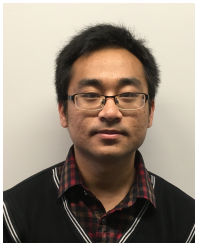
[73] JPEG, "JPEG Pleno database," https://jpeg.org/plenodb, 2016 (accessed May 19, 2018).

[74] C. Loop, Q. Cai, S. O. Escolano, and P. A. Chou, "Microsoft voxelized upper bodies-a voxelized point cloud dataset," *ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document m38673/M72012*, 2016.

[75] ITU, "Methodology for the subjective assessment of the quality of television pictures," *Recommendation BT. 500-13*, 2012.
[76] A. G. Solimini, "Are there side effects to watching 3D movies? a prospective crossover observational study on visually induced motion sickness," *PloS One*, vol. 8, no. 2, 2013.
[77] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems," *Displays*, vol. 29, no. 2, pp. 58–69, 2008.
[78] P. Cignoni, C. Rocchini, and R. Scopigno, "Metro: measuring error on simplified surfaces," in *Computer Graphics Forum*, vol. 17, no. 2. Wiley Online Library, 1998, pp. 167–174.
[79] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*. IEEE, 2003, pp. 1398–1402.
[80] H. Sheikh and A. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430–444, February 2006.
[81] Agisoft, "Agisoft Photoscan," http://www.agisoft.com, 2010 (accessed July 16, 2018).
[82] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM Trans. Graphics*, vol. 32, no. 3, pp. 29:1–29:13, June 2013.
[83] CloudCompare, "CloudCompare-3D point cloud and mesh processing software," https://www.danielgm.net/cc, 2003 (accessed April 15, 2018).
[84] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "Meshlab: an open-source mesh processing tool," in *Eurographics Italian Chapter Conference*, 2008, pp. 129–136.
[85] K. Mammou and P. Chou, "PCC test model category 13 v2," *ISO/IEC JTC1/SC29/WG11 MPEG, N17519*, 2018.
[86] K. Mammou, "PCC test model category 3 v1," *ISO/IEC JTC1/SC29/WG11 MPEG, N17349*, 2018.
[87] ——, "PCC test model category 2 v0," *ISO/IEC JTC1/SC29/WG11 MPEG, N17248*, 2017.
[88] C. Guede, J. Ricard, S. Lasserre, and J. Llach, "Technicolor point cloud renderer," *ISO/IEC JTC1/SC29/WG11 MPEG, N16902*, 2017.
[89] G. Lavoué, M. C. Larabi, and L. Váša, "On the efficiency of image metrics for evaluating the visual quality of 3D models," *IEEE Trans. Visualization and Computer Graphics*, vol. 22, no. 8, pp. 1987–1999, August 2015.
[90] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Processing*, vol. 15, no. 11, pp. 3440–3451, November 2006.
[91] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. John Wiley and Sons, 2014.
[92] K. Ding, K. Ma, S. Wang, and E. Simoncelli, "Image quality assessment: Unifying structure and texture similarity." *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2020.
[93] W. Zhang, K. Ma, J. Yan, D. Deng, and Z. Wang, "Blind image quality assessment using a deep bilinear convolutional neural network," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 30, no. 1, pp. 36–47, 2018.
[94] W. Zhang, K. Ma, G. Zhai, and X. Yang, "Uncertainty-aware blind image quality assessment in the laboratory and wild," *IEEE Trans. Image Processing*, vol. 30, pp. 3474–3486, 2021.

Qi Liu received the B.S. degree from Shandong Technology and Business University, Yantai, China in 2011, the M.S. degree from Xidian University, Xi'an, China in 2014 and the Ph.D. degree from Shandong University, Qingdao, China in 2021. From 2021.12 to now, she works as an Assistant Professor with the School of Electronic Information, Qingdao University, Qingdao, China. From 2018.09-2019.08, she also worked as an international visiting graduate student with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. Her research interests include point cloud compression and point cloud quality assessment.

Honglei Su (M'19) received the B.A.Sc degree from Shandong University of Science and Technology, Qingdao, China in 2008 and the Ph.D. degree from Xidian University, Xi'an, China in 2014. From 2014.09 to now, he works as an Assistant Professor with the School of Electronic Information, Qingdao University, Qingdao, China. From 2018.03-2019.03, he also 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, computer vision, etc.



Zhengfang Duanmu (S'15) received the B.A.Sc, M.A.Sc and Ph.D. degrees from the University of Waterloo, Waterloo, ON, Canada in 2015, 2017 and 2021 respectively. He is currently a Postdoctoral Fellow with the University of Waterloo. His research interests include perceptual image processing and quality-of-experience.

Wentao Liu (S'15) received the B.E. and M.E. degrees from Tsinghua University, Beijing, China in 2011 and 2014, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada in 2019. His research interests include perceptual quality assessment of images and videos.

Zhou Wang (S'99-M'02-SM'12-F'14) received the Ph.D. degree from The University of Texas at Austin, TX, USA, in 2001. He is currently a Professor and a Canada Research Chair with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests include image and video processing and coding; visual quality assessment and optimization; computational vision and pattern analysis; multimedia communications; and biomedical signal processing. He has more than 200 publications in these fields with over 70,000 citations (Google Scholar).

Dr. Wang serves as a member for the IEEE Multimedia Signal Processing Technical Committee from 2013 to 2015 and the IEEE Image, Video and Multidimensional Signal Processing Technical Committee from 2020 to 2022. He was elected as a Fellow of Canadian Academy of Engineering in 2016 and the Royal Society of Canada: Academy of Science in 2018. He was a recipient of the 2009 IEEE Signal Processing Society Best Paper Award, the 2013 *IEEE Signal Processing Magazine* Best Paper Award, the 2014 NSERC E.W.R. Steacie Memorial Fellowship Award, the 2015 Primetime Engineering Emmy Award, and the 2016 IEEE Signal Processing Society Sustained Impact Paper Award. He serves as a Senior Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING from 2022 to 2024, a Senior Area Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING from 2015 to 2019, an Associate Editor for IEEE SIGNAL PROCESSING LETTERS from 2006 to 2010, IEEE TRANSACTIONS ON IMAGE PROCESSING from 2009 to 2014, and IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2016 to 2018, and a Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING from 2013 to 2014 and from 2007 to 2009, among other journals.