PCQD-AR: Subjective Quality Assessment of Compressed Point Clouds with Head-mounted Augmented Reality

Chunling Fan[1], Yun Zhang[2], linwei Zhu[3], and Xinju Wu[4]

[1]Shenzhen Polytechnic
[2]Sun Yat-Sen University
[3]SIAT
[4]City University of Hong Kong

July 19, 2023

Abstract

In this letter, the colored point cloud quality assessment in Augmented Reality (AR) environment was fully studied through subjective test. Firstly, we present a point cloud dataset, named Point Cloud Quality Dataset-AR (PCQD-AR), including ten reference point clouds and their 90 distorted versions, which were encoded by the reference software of Video-based Point Cloud Compression (V-PCC) under different pairs of geometry and texture quantization parameters. Then, the impact of geometry and texture distortions on perceived quality of point clouds in the AR environment was discussed in detail. Moreover, we evaluate the performance of existing objective point cloud quality assessment metrics on the proposed dataset. The subjective dataset including the values of Mean Opinion Score (MOS) will be released after acceptance.

PCQD-AR: Subjective Quality Assessment of Compressed Point Clouds with Head-mounted Augmented Reality

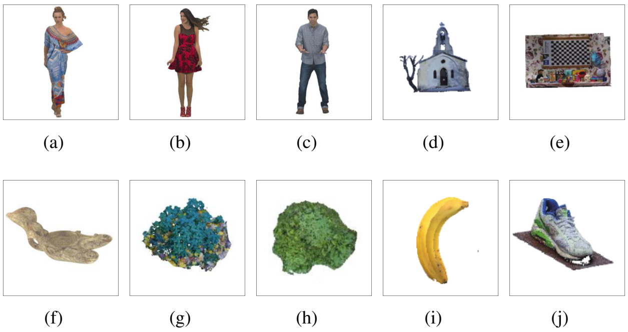
Chunling Fan,[1] Yun Zhang,[2] Linwei Zhu,[3] and Xinju Wu[4]
[1] *Shenzhen Polytechnic*
[2] *Sun Yat-sen University*
[3] *Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences*
[4] *City University of Hong Kong*
Email: zhangyun2@mail.sysu.edu.cn

In this letter, the colored point cloud quality assessment in Augmented Reality (AR) environment was fully studied through subjective test. Firstly, we present a point cloud dataset, named Point Cloud Quality Dataset-AR (PCQD-AR), including ten reference point clouds and their 90 distorted versions, which were encoded by the reference software of Video-based Point Cloud Compression (V-PCC) under different pairs of geometry and texture quantization parameters. Then, the impact of geometry and texture distortions on perceived quality of point clouds in the AR environment was discussed in detail. Moreover, we evaluate the performance of existing objective point cloud quality assessment metrics on the proposed dataset. The subjective dataset including the values of Mean Opinion Score (MOS) will be released after acceptance.

Introduction: Augmented Reality (AR) is gaining popularity as an immersive scenario that integrates computer-generated Three Dimensional (3D) models into the real world. Point clouds are often used to represent realistic models in the AR environment. A point cloud is usually composed of millions of 3D points where each point includes geometric information (e.g., X, Y, Z), texture information (e.g., R, G, B), reflectance, and so on. The required bit rate for storage and transmission is huge, so it is often necessary to compress point clouds in practical applications. However, due to lossy compression and other inevitable noises, the visual perception of point cloud models may be compromised.

Recently, Point Cloud Quality Assessment (PCQA) has been a research hotspot. The metrics can be roughly divided into two groups: objective and subjective metrics. The objective PCQA metric tries to design a mathematical model to assess the point cloud quality automatically, which can be further classified into two groups: point based and projection based methods. The former computes point-wise distortion in 3D space between the reference and distorted point clouds, e.g., point-to-point [1] and point-to-plane [2]. These methods hardly consider the structural information which is important for the Human Visual System (HVS). Alexiou *et al.* [3] calculated two indices in terms of Structure Similarity Index Measure (SSIM) for geometry and color information, and then combined them as the final quality score of point clouds. Yang *et al.* [4] applied graph signal processing to extract local features of key points in the point clouds. Liu *et al.* [5] used sparse convolutional neural network to extract features from the whole point cloud without downsampling. The latter projects the 3D points onto 2D planes and then applies existing 2D image quality assessment metrics. Wu *et al.* [6] proposed a patch projection based method, in which neighboring points are clustered and projected as 2D patches in an image. However, due to the limited research on perceptual quality of the HVS and insufficient point cloud dataset, the existing objective PCQA methods have not reached a high degree of correlation with the HVS.

The subjective PCQA is the most straightforward and reliable approach to evaluate the perceived quality, which also provides benchmark for objective PCQA metrics. As listed in Table 1, according to the equipment and the degree of interactivity in subjective test, existing works can be roughly divided into three categories: 1) 2D/3D monitor and non-interactive, 2) 2D/3D monitor and interactive, 3) Head Mounted Display (HMD) and Six Degrees of Freedom (6DoF). In the first group, point clouds are displayed on the 2D/3D monitor according to the set trajectory which does not support interaction [7–10]. In the second group, the point clouds are displayed on a 2D/3D display, the viewers can interact with the models, such as translate, rotate [5], and zoom [11]. In the third group, the point clouds are displayed in the HMD and the viewers can interact with the models in 6DoF. Wu *et al.* [6] explored the effect of geometry and texture attributes in compression distortion with 6DoF HMD in Virtual Reality (VR) environment. Alexiou *et al.* [12] first introduced AR device into subjective point cloud test, but they only explored geometry distortion without considering color information. AR is different from VR in two aspects: on the one hand, the background of AR is real environment, while the background of VR is virtual environment; on the other hand, the principles of display devices are different since optical see-through combiners are required for AR. These will affect the visual perception of the 3D point cloud models by the HVS.



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

(f) (g) (h) (i) (j)

Fig 1 *Snapshots of ten reference point clouds used in our subjective test. (a)Longdress. (b)Redandblack. (c)Loot. (d)House without roof. (e)ULB Unicorn. (f)Romanoillamp. (g)Newgrass. (h)Grass. (i)Bananamesh.(j)Nike.*

To study the impact of colored point cloud quality degradation in AR environment, we conduct a subjective test and construct a point cloud dataset. The main contributions of our work are listed as follows:

1. We conduct a subjective test of point clouds in the AR environment. The subjects are allowed to view the point clouds wearing Hololens and walk freely to interact in 6DoF.
2. We establish a point cloud dataset in 6DoF AR environment which is composed of ten static reference colored point clouds and their 90 distorted versions who are encoded with V-PCC, each reference is compressed with nine distortion levels by changing the geometry and texture Quantization Parameters (QPs).
3. We further discuss the impact of point clouds geometry and texture distortions in the AR environment.

Subjective Quality Assessment: To generate a diverse dataset, ten popular colored point cloud models were selected as reference shown in Fig. 1, which can be divided into two categories, i.e., human figures (Figs.1(a)-(c)) and inanimate objects (Figs.1(d)-(j)). They were selected from the MPEG point cloud dataset [13], the JPEG Pleno dataset [14], and the online platform Sketchfab. The V-PCC was applied to produce distorted versions as it shows a high compression performance. For each reference point cloud, many distorted versions were first generated by modifying the geometry and texture QPs. Nine distortion levels were then selected for each reference according to significant differences among different distortion levels.

To evaluate the point cloud quality exhibited in the real-world scene, the subjective test was conducted in a controlled laboratory. The AR display device used in our experiment is Microsoft Hololens (1st gen). The Hololens is equipped with a transparent holographic lens with a holographic resolution of 2 HD 16:9 light engines, which can generate 2.3 million total light points. Our experiment platform was constructed in Unity and deployed to the Hololens. The experimental environment and rendered results are shown in Fig. 2. Complex texture background and strong light will affect the perception of the point cloud models by human eyes. In order to reduce the influence of the background, a normal room with simple background was chosen. Specifically, the laboratory is equipped with LED lamps of 6400K color temperature and the color of the wall is cream-white.

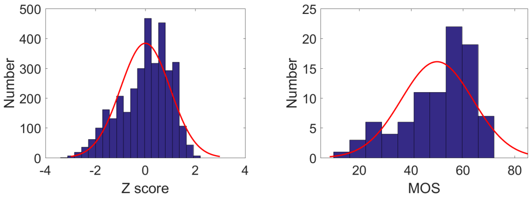
To ensure that each point cloud is suitable to be exhibited in a real-world scene, in the pre-processing stage, the CloudCompare software was used to rotate, translate and scale, so that each point cloud is within a

Table 1. Summary of existing point cloud datasets.

Category	Dataset	Ref. No.	Distortion type	Attributes	Display	Interaction
group 1	[7]	20	Downsampling, Gaussion Noise, S/V/L-PCC	Colored	2D	passive
	[8]	6	V-PCC, G-PCC	Colored	2D	passive
	[9]	4	FFmpeg, TFAN	Colored	2D	passive
	[10]	6	PCL, G-PCC,V-PCC	Colored	2D	passive
group 2	[11]	9	V-PCC, G-PCC	Colored	2D	rotation, translation, and zoom
	[5]	104	V-PCC, G-PCC, Downsampling, etc	Colored	2D	rotation
group 3	[6]	20	V-PCC	Colored	VR HMD	6DoF
	[12]	5	Gaussion Noise, Octree-pruning	Colorless	AR HMD	6DoF
	PCQD-AR	10	V-PCC	Colored	AR HMD	6DoF

(a) Environment (b) rendered results

Fig 2 *Experimental environment and rendered results.*



(a) Z scores (b) MOS

Fig 3 *Distribution of scores.*

similar bounding box (600,1000,400). Based on the fact that the selected codec only processes point clouds with integer coordinates, the positions of points were rounded and the duplicate points were removed. It is worth mentioning that due to the limitation of the viewing field of Hololens, in order to better observe and compare the reference and the distorted point cloud, the initial distance between subject and the center of two point clouds is set to 3.5 meters. It means that subjects can observe two point clouds at the same time, thus avoiding dizziness caused by frequent switching perspectives. As aforementioned, the subjects are required to stand in a fixed location before the experiment and walk freely to view and perceive during the experiment. Finally, the subjects go back to the initial location and rate. The details of pre-processed point clouds are shown in Table 2.

In this study, in order to avoid fatigue, 90 distorted point clouds were randomly and non-overlappingly divided into two sessions. Two sessions took about one hour in total and subjects were asked to have a rest for ten minutes between two sessions. In each session, the reference and distorted point clouds were displayed side-by-side. Note that subjects were told which was the reference and asked to evaluate the overall quality of each distorted version. As recommended by ITU-R BT.500-15 [15] and [16], the Double-Stimulus Impairment Scale (DSIS) evaluation methodology was selected in our experiment. The quality of the point cloud is defined by five discrete levels [17] (Excellent: imperceptible, Good: perceptible but not annoying, Fair: slightly annoying, Poor: annoying, and Bad: very annoying), which is in the range [0.5, 5], where 0.5 denotes the worst quality and 5 means that there is merely no difference between the distorted and the reference point clouds.

38 subjects (20 males and 18 females) participated in our subjective test aged from 22 to 35 years old. Specifically, seven of them are familiar with image and video coding or quality assessment and other 31 subjects have no experience in quality assessment. At the beginning, a color blindness and color weakness test was performed. In addition, a training session was performed to ensure that each subject is familiar with the AR equipment and the artifacts caused by the corresponding distortion type.

In subjective test, there may be outliers in the collected samples, which may be caused by the fatigue or other reasons. A procedure from Recommendation ITU-R BT 500.13 [15] for outlier detection was per-

Table 2. Reference point clouds after pre-processing.

Content	Source	Pre.	Points	Bounding Box	QP (geometry, texture)
Fig.1(a)	MPEG JPEG	×	797,178	(348,995,381)	(10,27),(10,37),(10,47), (16,10),(22,10),(32,10), (16,27),(22,37),(32,47)
Fig.1(b)	MPEG JPEG	×	729,133	(393,977,232)	(10,27),(10,37),(10,47), (27,10),(37,10),(47,10), (27,27),(37,37),(47,47)
Fig.1(c)	MPEG JPEG	×	797,178	(348,995,381)	(05,22),(05,32),(05,42), (16,05),(28,05),(36,05), (16,22),(28,32),(36,42)
Fig.1(d)	MPEG	✓	452,106	(477,423,468)	(10,27),(10,37),(10,47), (16,10),(36,10),(46,10), (16,27),(36,37),(46,47)
Fig.1(e)	MPEG	✓	949,797	(565,301,332)	(05,27),(05,37),(05,47), (16,05),(22,05),(32,05), (16,27),(22,37),(32,47)
Fig.1(f)	JPEG	✓	330,655	(517,355,352)	(10,16),(10,32),(10,42), (27,10),(37,10),(47,10), (27,16),(37,32),(47,42)
Fig.1(g)	Sketchfab	✓	452,106	(477,423,468)	(10,27),(10,37),(10,47), (27,10),(37,10),(47,10), (27,27),(37,37),(47,47)
Fig.1(h)	Sketchfab	✓	421,643	(265,347,294)	(10,16),(10,32),(10,42), (27,10),(37,10),(47,10), (27,16),(37,32),(47,42)
Fig.1(i)	Sketchfab	✓	202,770	(219,363,112)	(10,27),(10,37),(10,47), (16,10),(22,10),(32,10), (16,27),(22,37),(32,47)
Fig.1(j)	Sketchfab	✓	195,438	(302,213,300)	(10,16),(10,32),(10,42), (27,10),(37,10),(47,10), (27,16),(37,32),(47,42)

formed. No outlier was detected which indicates that the collected scores are all reliable. To give a uniform score across all subjects, we first transformed the raw subjective scores to Z scores. Fig. 3(a) shows the distribution of the Z scores, which indicates that about 99% of the Z scores lie in [-3, 3], then we mapped the Z scores to the range [0,100] by a linear mapping. Finally, we calculate the Mean Opinion Scores (MOS) for each sequence [15]. The histogram of the MOS of all distorted point clouds is shown in Fig. 3(b). It indicates that the perceptual quality has spanned a wide range of visual quality from severely annoying to imperceptible with a good separation.

Discussions: Although the four distortion levels correspond to different texture and geometry QPs for each reference point cloud, respectively, we have L1<L2<L3<L4. E.g. for sequence “Longdress”, textures L1, L2, L3, L4 are 10, 16, 22, 32. Table 3 lists the averaged MOS comparison for each distortion level across all the 3D point cloud models in our subjective test. We find that the averaged MOS decreases as the increase of distortion levels which indicates that the perceived quality of the point cloud degrades as the increase of distortion levels.

Fig. 4 presents the MOS for all the distorted point cloud in our subjective test. The X axis lists the geometry and texture QP for each distortion level, e.g., 10,27 denotes the geometry QP is 10 and texture QP is 27. It can be observed that: 1) When the texture QP increases while the geometry QP remains stable, the visual quality decreases for 90% of the test sequences. Only the visual quality for the “Grass” does not decrease. The texture in this sequence is rich and messy, so there exists strong masking effect. 2) When the geometry QP increases while the texture QP remains

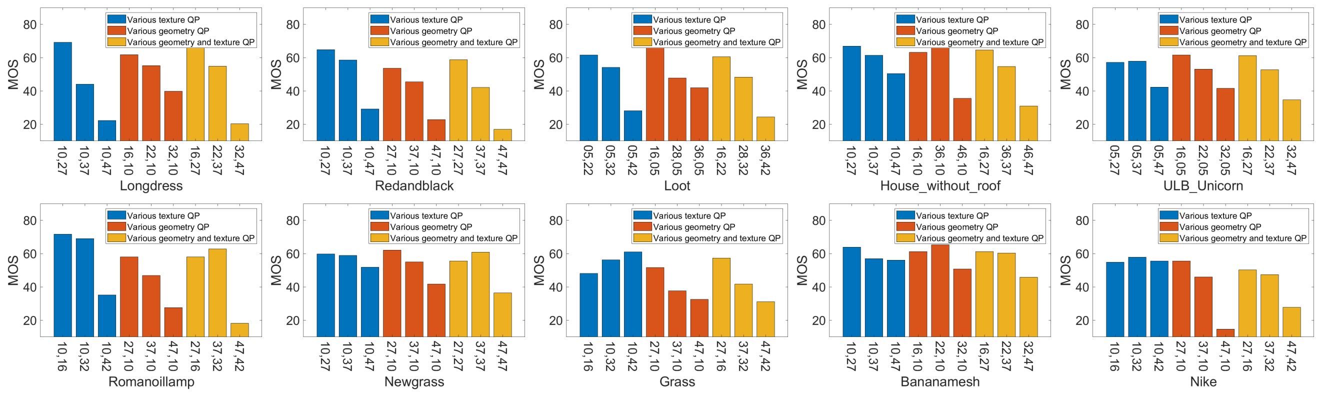


Fig 4 *The MOS for all the distorted 3D point cloud models in our subjective test.*

Table 3. The averaged MOS comparison for each distortion level across all the 3D point cloud models in our subjective test.

MOS		Texture QP L1	L2	L3	L4
Geometry QP	L1	61.26	57.53	44.76	-
	L2	59.48	59.17	-	-
	L3	52.51	-	53.06	-
	L4	37.67	-	-	32.04

Table 4. Performance comparison among objective PCQA metrics on the proposed PCQD-AR dataset.

Metrics	PLCC	SRCC	KRCC
MSE(p2point)	0.5775	0.5170	0.3998
PSNR(p2point)	0.3888	0.3089	0.2242
MSE(p2plane)	0.6502	0.5311	0.4144
PSNR(p2plane)	0.4103	0.3389	0.2453
PC-MSDM [18]	0.4068	0.4342	0.3218
PC-SSIM [3]	0.6210	0.4869	0.3416
Wu *et al.* [6]	0.6894	0.5798	0.4203
Graph-SSIM [4]	**0.7085**	**0.6254**	**0.4482**

stable, the visual quality decreases for all the test sequences. Furthermore, when the geometry QP of the sequence “House_without_roof” increases from 16 to 36, the visual quality keeps the same, but when the geometry QP drops from 36 to 46, the visual quality decreases significantly. When the geometry QP of the sequence “Bananamesh” increases from 16 to 22, the visual quality keeps the same, but when the geometry QP increases from 22 to 32, the visual quality decreases significantly. It is found that there is large flat area in “House_without_roof” and “Bananamesh”. When the geometry QP increases slightly, the geometric distortion is masked by the texture, and the human eye can perceive the geometric distortion only when the geometric distortion is large. 3) Compared with the inanimate object sequences, the MOS drops more for the human figure sequences, indicating that the human eye is more sensitive to the distortion in these sequences.

Performance comparison of objective methods: To evaluate the proposed dataset, some existing objective PCQA metrics are adopted, such as MSE, PSNR, PC-MSDM [18], PC-SSIM [3], Wu *et al.* [6], and GraphSIM [4]. Table 4 presents the performance comparison among objective point clouds quality assessment metrics on the proposed dataset. The indices Pearson Linear Correlation Coefficients (PLCC), Spearman's Rank Order Correlation Coefficients (SROCC), and Kendall Rank-order Correlation Coefficient (KRCC) are selected to present the performance. The best performing metric was found to be GraphSIM, but the metric is not highly consistent with HVS. One of the possible reasons may be that these methods do not consider the context in the AR environment, which can be further explored in the future.

Conclusions: We conduct a subjective test to study the perceived quality of point clouds in the 6DoF AR environment. Ten colored point clouds are encoded with different pairs of geometry and texture QPs to generate distorted versions. We find that the subjects are sensitive to the distortion in the human figure sequences. The context in the AR environment needs to be explored in the future PCQA work. The proposed point cloud dataset will be a benchmark for PCQA in the 6DoF AR environment.

References

1. Mekuria, R., Li, H., Tulvan, C., Chou, P. A.: Evaluation criteria for PCC (Point Cloud Compression). ISO/IEC MPEG n163(2016)
2. Tian, D., Ochimizu, H., Feng, C., Cohen, R., Vetro, A.: Geometric distortion metrics for point cloud compression. IEEE Int. Conf. Image Process. Beijing, China, 3460–3464 (2017)
3. Alexiou, E., Ebrahimi, T.: Towards a point cloud structural similarity metric. IEEE Int. Conf. Multimedia Expo Workshops 1–6 (2020)
4. Yang, Q., Ma, Z., Xu, Y., Li, Z., Sun, J.: Inferring point cloud quality via graph similarity. IEEE Trans. Pattern Anal. Machine Intell. 44(6), 3015–3029 (2020)
5. Liu, Y., Yang, Q., Xu, Y., Yang, L.: Point cloud quality assessment: Dataset construction and learning-based no-reference metric. ACM Trans. Multimedia Comput., Commun. and Applications 19(2), 1–26 (2023)
6. Wu, X., Zhang, Y., Fan, C., Hou, J., Kwong, S.: Subjective quality dataset and objective study of compressed point clouds with 6DoF head-mounted display. IEEE Trans. Circuits Syst. Video Technol. 19(2), 1–26 (2021)
7. Liu, Q., Su, H., Duanmu, Z., Liu, W., Wang, Z.: Perceptual quality assessment of colored 3D point clouds. IEEE Trans. Vis. Comput. Graph. 29(8), 3642–3655 (2023)
8. Perry, H. P., Cong, H., da Silva Cruz, L. A., Prazeres, J., Pinheiro, A., Dumic, E., Alexiou, E., Ebrahimi, T.: Quality evaluation of static point clouds encoded using MPEG codecs. IEEE Int. Conf. Image Process. 3428–3432 (2020)
9. Cao, K., Xu, Y., Cosman, P.: Visual quality of compressed mesh and point cloud sequences. IEEE Access 8, 171203–171217 (2020)
10. Alireza, J., Catarina, B., Fernando, P., Joao, A.: Point cloud rendering after coding: impacts on subjective and objective quality. IEEE Trans. Multimedia 23, 4049–4064 (2021)
11. Alexiou, E., Viola, I., Borges, T. M., Fonseca, T. A., de Queiroz, R. L., Ebrahimi, T.: A comprehensive study of the rate-distortion performance in MPEG point cloud compression. APSIPA Trans. Signal Inf. Process. 8, 1–27 (2019)
12. Alexiou, E., Upenik, E., Ebrahimi, T.: Towards subjective quality assessment of point cloud imaging in augmented reality. IEEE 19th Int. Workshop Multimedia Signal Process. London-Luton, UK, 1–6 (2017)
13. MPEG: MPEG point cloud dataset. Avaliable: http://mpegfs.intevry.fr/MPEG/PCC/DataSets/pointCloud/CfP/(2017)
14. JPEG: JPEG pleno dataset. Avaliable: https://jpeg.org/plenodb/(2016)
15. ITU-R BT.500-15: Methodologies for the subjective assessment of the quality of television images. Int. Telecommun. Union (2023)
16. Nehmé, Y., Farrugia, J., Dupont, F., Callet, P. L., Lavoué, G.: Comparison of subjective methods for quality assessment of 3D graphics in virtual reality. ACM Trans. Appl. Percept. 18(2), 1–23(2021)
17. Yang, H., Fang, Y., Lin, W.: Perceptual quality assessment of screen content images. IEEE Trans. Image Process. 24(11), 4408–4421(2015)
18. Meynet, G., Digne, J., Lavoué, G.: PC-MSDM: a quality metric for 3D point clouds. IEEE Int. Conf. Quality Multimedia Experience 24(11), Berlin, Germany, 1–3(2019)