PERCEPTUAL QUALITY ASSESSMENT OF 3D POINT CLOUDS

Honglei Su[⋆†]*, Zhengfang Duanmu*[†]*, Wentao Liu*[†]*, Qi Liu*[‡†] *and Zhou Wang*[†]

[⋆] School of Electronic Information, Qingdao University
[†] Department of Electrical and Computer Engineering, University of Waterloo
[‡] School of Information Science and Engineering, Shandong University

ABSTRACT

The real-world applications of 3D point clouds have been growing rapidly in recent years, but effective approaches and datasets to assess the quality of 3D point clouds are largely lacking. In this work, we construct so far the largest 3D point cloud database with diverse source content and distortion patterns, and carry out a comprehensive subjective user study. We construct 20 high quality, realistic, and omni-directional point clouds of diverse contents. We then apply downsampling, Gaussian noise, and three types of compression algorithms to create 740 distorted point clouds. Based on the database, we carry out a subjective experiment to evaluate the quality of distorted point clouds, and perform a point cloud encoder comparison. Our statistical analysis find that existing point cloud quality assessment models are limited in predicting subjective quality ratings. The database will be made publicly available to facilitate future research.

Index Terms— point cloud, image quality assessment, subjective quality, point cloud compression, downsampling

1. INTRODUCTION

A 3D point cloud is a collection of points representing a 3D shape, object or environment. Each point has its own geometric coordinates and other associated attributes. 3D point clouds find a wide variety of applications in manufacturing, construction, environmental monitoring, navigation, animation, and rendering, among many others. Point clouds are subject to various distortions during acquisition, processing, compression, transmission, storage and rendering, any of which may lead to quality degradation. Subjective quality assessment is a straightforward and reliable approach to evaluate point cloud quality. Although expensive, inconvenient and time-consuming, a comprehensive subjective quality study has many benefits. First, it provides data to study user behaviors in evaluating perceived quality of point clouds. Second, it supplies a test set to evaluate and compare the relative performance of point cloud processing methods such as point cloud compression (PCC) algorithms. Third, it can be utilized as a benchmark to validate and compare the performance of objective point cloud quality assessment (PCQA) models.

There are a number of publicly available point cloud databases, including MPEG point cloud datasets [1], JPEG Pleno database [2], and Stanford 3D scanning repository [3]. These databases suffer from several problems that limit their usage in PCQA. These include inferior initial quality, limited meaningful viewpoints, and insufficient content types. As a result, subjective experiments derived from these databases [4, 5, 6, 7, 8, 9, 10] are inherently deficient in providing reliable evaluations of point cloud processing algorithms and objective PCQA models. The lack of realistic distortions such as PCC also reduces the relevancy of the databases in real-world applications. Moreover, most subject-rated databases are not publicly available, with the exception that the subjective data on a rather small database consisting of only two contents are released in [10].

In this work, we construct a subject-rated point cloud database covering diverse contents and distortion patterns. First, we generate so far the largest high quality point cloud dataset. The database consists of 20 contents of diverse geometric characteristics and textural patterns. Second, we derive 740 distorted point clouds of five distortion types from the dataset, where the distortion types include downsampling, Gaussian noise contamination, and three state-of-the-art PCC algorithms–S-PCC, V-PCC, and L-PCC [11]. Using the database, we carry out a subjective user study to evaluate perceived quality of distorted point clouds. We perform statistical analysis on the existing PCC algorithms, which leads to several interesting observations. Finally, we conduct a comprehensive evaluation on existing objective PCQA models.

2. SOURCE POINT CLOUD CONSTRUCTION

We gather a collection of objects with diverse geometric and textural complexity, including snacks, fruits, vegetables, office supplies, and etc. The selected contents are moderate in

This work was supported in part by Shandong Provincial Natural Science Foundation, China under Grants ZR2018PF002, and in part by Natural Sciences and Engineering Research Council of Canada. Correspondence author Honglei Su, Email: suhonglei@qdu.edu.cn.

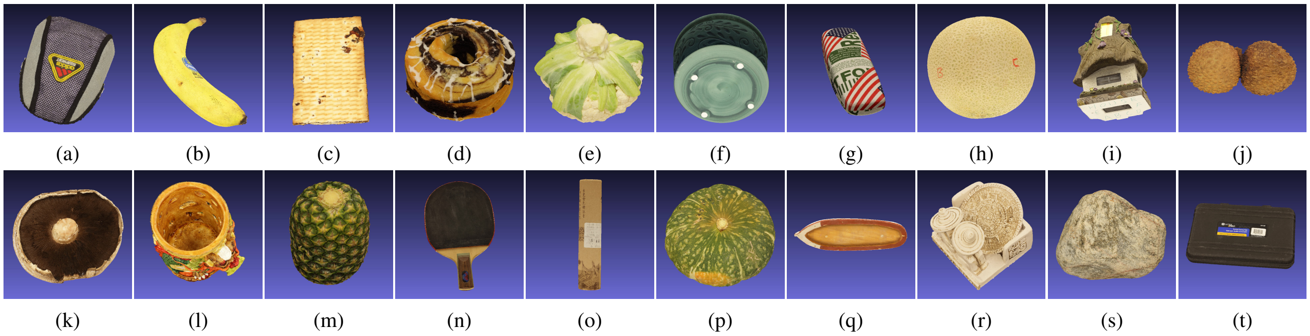

Fig. 1: Snapshots of acquired point clouds.

size and are omni-directionally acquirable. Fig. 1 shows snapshots of the objects in our point cloud dataset. The 3D point clouds are constructed using the following steps.

- *Image acquisition*: Image acquisition is conducted in standardized laboratory environment which has a normal lighting condition without reflecting ceiling walls and floor. A single-lens-reflex camera and a turntable are employed to take photos of an object from any angle. An example is shown in Fig. 2, where each photo is placed at its capture position relative to the object in the center.
- *3D reconstruction*: We apply image alignment, sparse point cloud reconstruction, dense point cloud reconstruction and point cloud merging to each sequence of images with Agisoft Photoscan [12]. The resulting point clouds are further refined by Screened Poisson Surface Reconstruction [13] and resampling using CloudCompare [14].
- *Standardization*: Each point cloud is normalized to a unit-cube with a step size of 0.001, where duplicated points are removed [14]. Finally, 20 voxelized point clouds are generated with an average number of 1.35M points and a standard deviation of 656K, respectively.

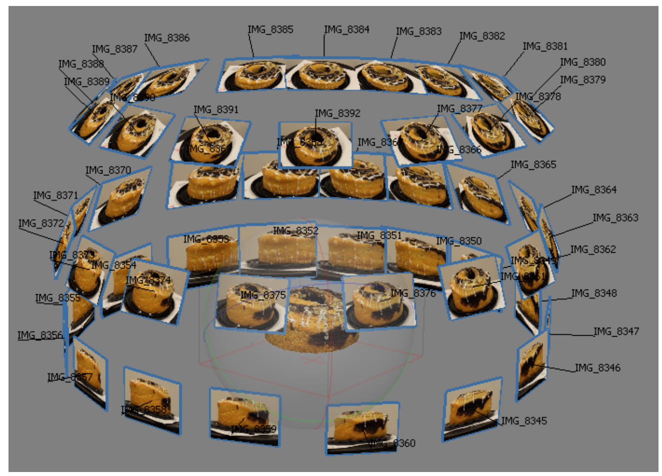

Fig. 2: Sample image acquisition process.

3. POINT CLOUD DATABASE CONSTRUCTION AND SUBJECTIVE QUALITY ASSESSMENT

3.1. Distortion Generation

Using the aforementioned point cloud as the source, we apply the following distortion processes.

- *Downsampling*: We apply octree-based downsampling [14] to the normalized point clouds. Each dimension is uniformly divided into 2^N intervals, where N represents the octree level. Then points located in the same cube are merged into one node. In this study, we set N to be 7, 8, and 9 respectively, to cover diverse spatial resolutions.
- *Gaussian noise contamination*: White Gaussian noise is added independently to both geometry and texture elements with standard deviation 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 [15].
- *S/V/L-PCC*: In 2017, MPEG issued a call for proposals on PCC methods for International Organization for Standardization [11]. Three technologies were chosen as test models in three categories: S-PCC for static content, V-PCC for dynamic content and L-PCC for dynamically capturing, respectively. In this work, S-PCC reference codec [16] is employed to encode the original point clouds with 'triSoupDepth' of $\{10\}$, 'triSoupLevel' of $\{4, 6, 8\}$ and 'rahtQuantizationStep' of $\{64, 128, 256, 512\}$, respectively. V-PCC reference codec [17] 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 [15]. L-PCC [18] employs downsampling method to encode

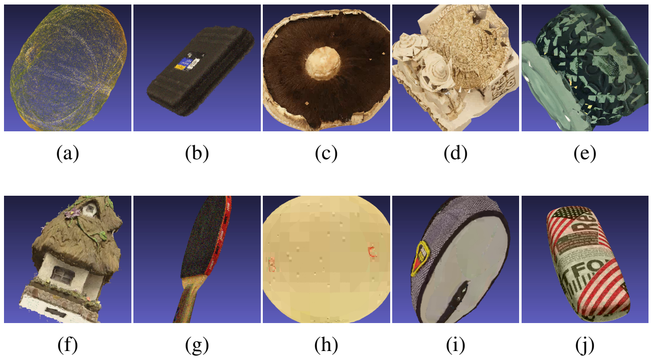

Fig. 3: Point cloud distortions. Geometry distortions: (a) Hollow. (b) Geometry noise. (c) Hole. (d) Shape distortion. (e) Collapse. (f) Gap and burr. Texture distortions: (g) Texture noise. (h) Blockiness. (i) Blur. (j) Color bleeding.

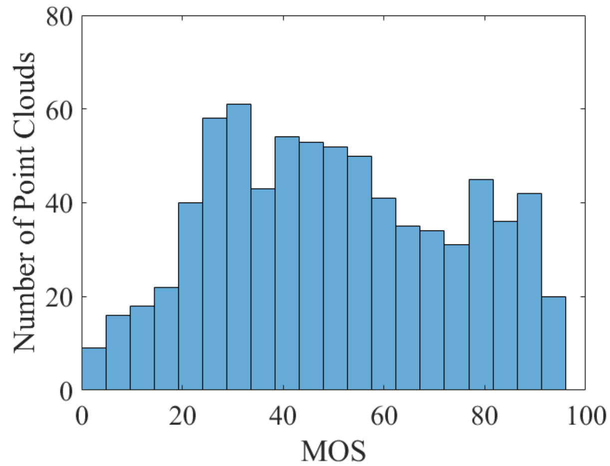

Fig. 4: MOS distribution.

the geometry information, and is thus not performed redundantly. We set the 'quantizationSteps' of texture encoding as $\{16, 32, 48, 64\}$.

Eventually, 740 distorted point clouds are generated in total by 5 distortion generators from 20 original point clouds.

Sampled distortion patterns in our database are shown in Fig. 3. It is interesting to observe that the distorted point cloud not only exhibits loss of texture information similar to 2D images such as blockiness and blur, but also novel geometric distortion types. Specifically, hollow is caused by point cloud downsampling, where the point density is not sufficient to cover the object surface. Holes and collapses arise from unsuccessful triangulations and inappropriate downsampling in S-PCC, respectively. Even when the triangulation is successful, geometry distortion may still appear as a consequence of ill-conditioned triangles. A sample case is given in the bottom right part of Fig. 3 (d). Moreover, a large 'geometryQP' in V-PCC potentially results in gaps and burrs. All these distortions are point cloud-specific, which create new challenges to objective PCQA models.

3.2. Subjective Quality Assessment

We employ Technicolor renderer [19] to render each point cloud to a video sequence. The rendering window, point size and point type are set to 960×960, 1 and 'point', respectively. A horizontal and a vertical circle 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 preserves 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 horizontally with its pristine counterpart into a 10-second video sequence for presentation.

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 Truecolor (32bit) at 60 Hz. The monitor is calibrated in accordance with ITU-R Recommendation BT.500-13 [20]. Double-stimulus impairment scale (DSIS) methodology is applied in our subjective test [20]. Videos are displayed in random order using a customized graphical user interface, where subjective scores of individual viewer are recorded.

A total of 60 naïve subjects, including 32 males and 28 females aged between 21 and 40, are recruited in the subjective test. All the 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 that are different from the videos in the testing session are shown to the subjects. The same methods are applied to generate videos used in both the training and testing sessions. Therefore, subjects knew what distortion types and levels would appear before the testing session, and thus the learning effects are kept minimal. Due to the limited subjective experiment capacity, we employed the following strategy. Each subject is assigned 10 objects in a circular fashion. Specifically, if subject i is assigned objects 1 to 10, then subject $i + 1$ watch objects 2 to 11. Each video is scored for 30 times and 22,800 subjective ratings, including 600 scores for reference point clouds, are collected in total. 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 influence of fatigue effect. Due to expanded range and finer distinctions between ratings, 100-point continuous scale is utilized instead of a discrete 5-point ITU-R Absolute Category Scale (ACR).

After converting subjective scores to Z-scores, we apply outlier removal scheme suggested in [20]. No outlier detection is conducted participant-wise. Then Z-scores are linearly

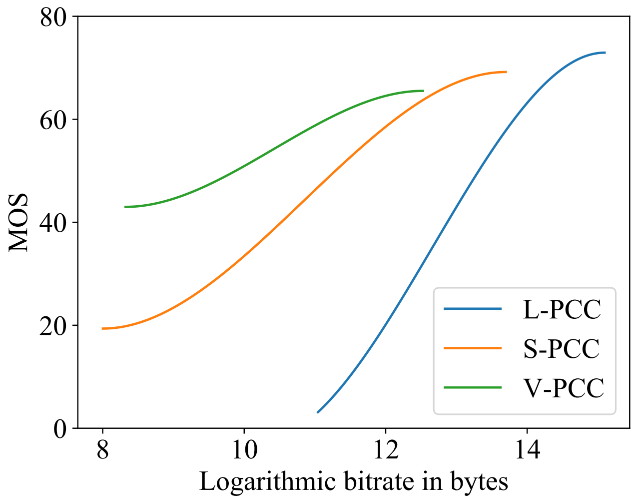

Fig. 5: Aggregated rate-quality curves of S/V/L-PCC reference codecs.

rescaled to lie in the range [0, 100]. Mean opinion score (MOS) for each distorted point cloud is calculated by averaging the re-scaled Z-scores from all valid subjects. The histogram for MOS distribution is shown in Fig. 4, which demonstrates the distorted point clouds span most of the quality range. Considering the MOS as the "ground truth", Pearson linear correlation coefficient (PLCC) and Spearman rank-order correlation coefficient (SRCC) between each viewer's scores and MOSs are calculated to estimate the performance of individual subjects. Both the mean PLCC and SRCC between each subject and MOS are as high as 0.81 with a relatively small standard deviation of 0.08, which indicates substantial agreement between individual subjects.

4. DATA ANALYSIS AND DISCUSSION

4.1. Performance Comparison of PCC Algorithms

Compression of point clouds has been an active research topic in the past few years. There is a significant interest in comparing the performance of state-of-the-art PCC. Here we compare the performance of the MPEG test models S/V/L-PCC [11]. It is worth mentioning that the default settings are selected for S/V/L-PCC reference codecs. Fine tuning of the codecs is not performed.

The aggregated rate-quality curves over all content of the PCC codecs are shown in Fig. 5. From the subjective test results and the rate-quality performance, we have several observations. First, V-PCC achieves the best performance on average at low bitrate range, while S-PCC reference codec being the second best. Second, at high bitrate, the performance gaps between PCC methods become narrower. Third, for S-PCC reference codec, not all rate-distortion surfaces are monotonous along the direction of increasing geometry bitrate due to unsuccessful triangulation. Fourth, in L-PCC reference codec, the perceived quality decreases rapidly when downsampling occurs with the decrease of geometry bitrate.

Table 1: Performance of Objective PCQA Models

Model	PLCC	SRCC	RMSE
PSNR$_{point2point,MSE}$	0.43	0.41	20.66
PSNR$_{point2point,Hausdorff}$	0.34	0.26	21.54
PSNR$_{point2plane,MSE}$	0.40	0.37	21.06
PSNR$_{point2plane,Hausdorff}$	0.34	0.29	21.55
PSNR$_Y$	0.61	0.58	18.20
AS$_{Mean}$	0.34	0.33	21.59
AS$_{RMS}$	0.33	0.32	21.66
AS$_{MSE}$	0.33	0.32	21.66
PSNR$_{projection}$	0.50	0.46	19.87
SSIM$_{projection}$	0.60	0.61	18.32
MS-SSIM$_{projection}$	0.67	0.67	17.02
VIFP$_{projection}$	**0.77**	**0.77**	**14.71**

4.2. Performance of Objective PCQA Models

The performance of both point-based [21, 22, 23, 24, 25, 26] and projection-based [9] objective PCQA models are tested and compared against our database. We employ PLCC, SRCC and root mean squared error (RMSE) as the evaluation criteria [27]. Table 1 shows the experimental results. We summarize the key observations as follows. Compared to PSNR$_Y$ [24, 25], geometry PSNR models [21, 24, 25] and angular similarity (AS) models [26] do not seem to provide adequate predictions of perceived quality of colored point clouds due to the lack of texture information. By contrast, projection-based models perform better, among which visual information fidelity in pixel domain (VIFP) [28] achieves the best performance compared to PSNR, structural similarity index (SSIM) [29] and multi-scale structural similarity (MS-SSIM) [30]. Nevertheless, the quality prediction accuracy is only moderate when compared with the performance of image quality models on 2D images.

5. CONCLUSION

We construct a point cloud database of diverse content and distortion variations and conduct a lab-controlled subjective user study. The database contains 760 point clouds with MOSs approximately evenly distributed from poor to excellent perceived quality levels. We find that V-PCC performs better than S-PCC and L-PCC at low bitrate. Our study also show that existing objective PCQA models are limited in providing accurate quality predictions, suggesting significant room for future improvement. The database will be made publicly available to facilitate future research.

6. REFERENCES

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

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

[3] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," in *Proc. ACM SIGGRAPH*, 1996, pp. 303–312.

[4] 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*, 2017, pp. 1–6.

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

[6] 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*, 2018, pp. 1–6.

[7] 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*, 2018, pp. 107520H.1–107520H.14.

[8] 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*, 2017, pp. 1–6.

[9] E. Torlig, E. Alexiou, T. Fonseca, R. de Queiroz, and T. Ebrahimi, "A novel methodology for quality assessment of voxelized point clouds," in *Proc. SPIE Optical Engineering+Applications*, 2018, pp. 107520I.1–107520I.17.

[10] E. Zerman, P. Gao, C. Ozcinar, and A. Smolic, "Subjective and objective quality assessment for volumetric video compression," in *IS&T Electronic Imaging, Image Quality and System Performance XVI*, in press.

[11] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. Chou, R. 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*, in press.

[12] Agisoft, "Agisoft Photoscan," http://www.agisoft.com, 2010 (accessed July 16, 2018).

[13] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM Transactions on Graphics*, vol. 32, no. 3, pp. 29:1–29:13, June 2013.

[14] CloudCompare, "CloudCompare-3D point cloud and mesh processing software," https://www.danielgm.net/cc, 2003 (accessed April 15, 2018).

[15] 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.

[16] K. Mammou and P. Chou, "PCC test model category 13 v2," *ISO/IEC JTC1/SC29/WG11 MPEG, N17519*, 2018.

[17] K Mammou, "PCC test model category 2 v0," *ISO/IEC JTC1/SC29/WG11 MPEG, N17248*, 2017.

[18] K Mammou, "PCC test model category 3 v1," *ISO/IEC JTC1/SC29/WG11 MPEG, N17349*, 2018.

[19] C. Guede, J. Ricard, S. Lasserre, and J. Llach, "Technicolor point cloud renderer," *ISO/IEC JTC1/SC29/WG11 MPEG, M40229*, 2017.

[20] ITU, "Methodology for the subjective assessment of the quality of television pictures," *Recommendation BT. 500-13*, 2012.

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

[22] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Evaluation metrics for point cloud compression," *ISO/IEC JTC1/SC29/WG11 MPEG, M39966*, 2017.

[23] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Updates and integration of evaluation metric software for pcc," *ISO/IEC JTC1/SC29/WG11 MPEG, M40522*, 2017.

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

[25] 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 Tech.*, vol. 27, no. 4, pp. 828–842, April 2017.

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

[27] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," http://www.vqeg.org, 2000 (accessed December 04, 2018).

[28] H. Sheikh and A. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430–444, February 2006.

[29] 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.

[30] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, 2003, pp. 1398–1402.