A reduced reference metric for visual quality evaluation of point cloud contents

Viola, Irene; Cesar, Pablo

DOI
10.1109/LSP.2020.3024065

Publication date
2020

Document Version
Accepted author manuscript

Published in
IEEE Signal Processing Letters

A reduced reference metric for visual quality evaluation of point cloud contents

Irene Viola, *Member, IEEE*, and Pablo Cesar, *Senior Member, IEEE*

***Abstract*—Point cloud representation has seen a surge of popularity in recent years, thanks to its capability to reproduce volumetric scenes in immersive scenarios. New compression solutions for streaming of point cloud contents have been proposed, which require objective quality metrics to reliably assess the level of degradation introduced by coding and transmission distortions. In this context, reduced reference metrics aim to predict the visual quality of the transmitted contents, while requiring only a small set of features to be sent in addition to the streamed media. In this paper, we propose a reduced reference metric to predict the quality of point cloud contents under compression distortions. To do so, we extract a small set of statistical features from the reference point cloud in the geometry, color and normal vector domain, which can be used at the receiver side to assess the visual degradation of the content. Using publicly available ground-truth datasets, we compare the performance of our metric to widely-used full reference metrics. Results demonstrate that our metric is able to effectively predict the level of distortion in the degraded point cloud contents, achieving high correlation values with respect to subjective scores.**

***Index Terms*—objective quality metric, point cloud, compression, reduced reference metric**

I. Introduction

Recent advances in 3D acquisition and rendering technologies, such as low-cost sensors and cross reality (XR) devices, as well as commodity hardware with sufficient computational power, have led to a renewed interest in photo-realistic immersive virtual reality experiences. In order to enable free movement in the virtual world in 6 degrees of freedom, a three-dimensional representation model is needed. Among others, point cloud represents a popular format to acquire, transmit and render volumetric content. However, the large amount of data comprising a point cloud content can easily become a bottleneck in current storage and delivery systems. To alleviate the problem, point cloud compression has been extensively examined in recent years, and a new compression standard is expected to be released by the MPEG standardisation body [1].

In order to design and evaluate new compression solutions that effectively remove redundancy in the data, without compromising on its visual quality, subjective or objective measures of quality distortions are usually employed. While the former is commonly considered as ground-truth information regarding the perceptual merit of distorted contents, it is cumbersome and expensive to execute. Thus, great effort has been spent in the literature in order to create algorithmic solutions that can mimic users' perception. Objective metrics for visual quality assessment are commonly classified as Full Reference (FR), Reduced Reference (RR) and No Reference (NR), depending on the availability, at computation time, of undistorted reference information. FR metrics are undoubtedly the most popular for point cloud contents, as they can leverage information from the entire uncompressed content to estimate the distortion. On the other hand, RR and NR metrics can be usefully employed when little to no information is known about the original content, which is a common occurrence at the receiver's side in broadcast and streaming scenarios. However, to the best of our knowledge, no RR or NR metric has been proposed yet for point cloud contents.

In this paper, we propose a new RR metric for visual quality assessment of point cloud contents. In particular, we extract a small set of features from a given reference content, based on both structure and attribute domains. Such features are then transmitted alongside the content, and are used at the receiver side in order to predict the visual quality of the content under exam. Moreover, we find the best combination of the proposed features through a linear optimization algorithm. We test the validity of our metric on four publicly available point cloud datasets with ground-truth subjective scores. Our results show the informative value of our features, demonstrating that our metric is capable of accurately predicting the visual quality of point cloud contents and achieving better performance with respect to well-established FR point cloud metrics. An implementation of the proposed metric is available here: https://github.com/cwi-dis/PCM_RR.

II. Related work

FR objective quality metrics for point cloud contents can be broadly classified as a) point-based or b) projection-based. In point-based metrics, correspondences between the points in the reference and distorted contents are used as the basis for the computation. Several point-based approaches have been proposed in the literature to assess distortions in the geometry and color domain. Point-to-point metrics are computed using the Euclidean distances between pairs of associated points that belong to the reference and the content under assessment [2]. Point-to-plane metrics, on the other hand, rely on computing the projected error between a displaced point and a corresponding normal vector in the associated reference point cloud [3]. A point-to-mesh approach has also been proposed, in which the distance between a displaced point and its reference reconstructed surface is computed [4]. However,

This paragraph of the first footnote will contain the date on which you submitted your paper for review. This paper was partly funded by the European Commission as part of the H2020 program, under the grant agreement 762111, "VRTogether" (http://vrtogether.eu/)

I. Viola is with the Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands (e-mail: irene@cwi.nl).

P. Cesar is with the Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands, and with TU Delft, Delft, The Netherlands. (e-mail: p.s.cesar@cwi.nl).

TABLE I
FEATURE VECTOR $\Phi_{\mathcal{S}}$ FOR GEOMETRY AND LUMINANCE PROPERTIES.

Feature Name	Definition
Mean	$f_1 = \frac{1}{P}\sum_i s_i$
Std	$f_2 = \sqrt{\frac{1}{P}\sum_i (s_i - f_1)^2}$
Median	$f_3 = \frac{1}{2}(\hat{\mathcal{S}}_{\lfloor \frac{P+1}{2} \rfloor} + \hat{\mathcal{S}}_{\lceil \frac{P+1}{2} \rceil})$
Mode	$f_4 = \{s_i \mid \mathbb{P}(s_i) \geq \mathbb{P}(s_j), j \neq i\}$
Entropy	$f_5 = -\sum_i \mathbb{P}(s_i) \log_2 \mathbb{P}(s_i)$
Energy	$f_6 = \sum_i \mathbb{P}(s_i)^2$
Sparsity	$f_7 = \frac{\lvert Z \rvert}{N^B}, Z = \{s_i \mid \mathbb{P}(s_i) \neq 0\}$

TABLE II
FEATURE SET Φ^N FOR NORMAL CONSISTENCY PROPERTIES.

Feature Name	Definition
Mean of Means	$f_1 = \frac{1}{kP}\sum_i \sum_j \theta(i,j)$
Mean of Stds	$f_2 = \frac{1}{k}\sum_j \sqrt{\frac{1}{P}\sum_i (\theta(i,j) - \frac{1}{P}\sum_l \theta(l,j))^2}$
Mean of Medians	$f_3 = \frac{1}{k}\sum_j \frac{1}{2}(\hat{\Theta}^{\{j\}}_{\lfloor \frac{P+1}{2} \rfloor} + \hat{\Theta}^{\{j\}}_{\lceil \frac{P+1}{2} \rceil})$
Std of Means	$f_4 = \sqrt{\frac{1}{P}\sum_i (\frac{1}{k}\sum_j \theta(i,j) - f_1)^2}$
Entropy	$f_5 = -\sum_i \mathbb{P}(\tilde{\theta_i}) \log_2 \mathbb{P}(\tilde{\theta_i})$
Energy	$f_6 = \sum_i \mathbb{P}(\tilde{\theta_i})^2$
Sparsity	$f_7 = \frac{\lvert \tilde{Z} \rvert}{N^B}, Z = \{\tilde{\theta_i} \mid \mathbb{P}(\tilde{\theta_i}) \neq 0\}$

as the method relies heavily on the mesh construction process, it is considered suboptimal. Plane-to-plane metrics have been suggested, using angular similarity among normal vectors in reference and distorted point cloud objects to assess the level of impairment [5]. Once the point distances have been calculated, the overall geometric distortion is usually measured using Mean Squared Error (MSE) or Hausdorff Distance, and a measure of quality can be expressed using Peak Signal to Noise Ratio (PSNR). Using similar approaches, the changes to the color attribute introduced during compression can also be measured at each point. A generalized Hausdorff distance has been proposed to improve the performance of geometry-based metrics [6], as well as a scale-invariant, point-to-distribution geometry metric based on Mahalanobis distance [7]. Recently, curvature statistics have also been proposed in order to estimate the distortion of a point cloud with respect to its reference [8], and they have been extended to include color information [9]. Viola et al. incorporate color distortion in geometry-based metrics, using luminance histogram information [10], whereas Diniz et al. use local binary pattern descriptors to estimate texture distortion [11]. In [12], Alexiou et al. propose the usage of local statistical features in order to obtain a global measure of degradation, similarly to the Structural Similarity Index (SSIM) in the image domain.

Projection-based metrics rely on mapping the original and distorted point clouds on planar surfaces, and then using popular image quality assessment metrics on the resulting projected images. The approach has the advantage of naturally combining geometry and color distortions; moreover, it can leverage existing image distortion metrics, such as PSNR and SSIM. The approach has been pioneered by Queiroz et al [13] to drive the rate-distortion optimization in their codec, and its performance has been analyzed in [14]. Alexiou et al. [15] investigated the impact of the number of viewports on the performance of the metric, and proposed a weighting system based on user interaction.

III. PROPOSED METRIC

RR metrics need to rely on extracting a set of features from a reference content in order to predict the level of distortion in the content under assessment. As the set of features needs to be transmitted alongside the content, it needs to be as informative as possible while maintaining a low cardinality. RR metrics have been adopted in the image and video community in order to produce a real-time estimation of visual quality at the receiving side of the transmission [16], [17]. However, adapting a RR framework to point cloud contents requires rethinking in terms of what dimensions will be affected by compression and transmission distortions. Traditional static 2D contents lie on a regular grid, which is unlikely to be tampered with. Hence, distortions will likely be present in the luminance or color domain. On the other hand, static point cloud contents can be distorted in the geometrical domain, along with the point attributes domain. In our work, we propose to use statistical features computed on the geometry information, luminance channel, and normal vectors, in order to measure the level of distortion of a degraded point cloud content.

A. Geometry-based features

Distortions in the geometrical composition of a point cloud content include reduction in the number of points, such as compression-based artifacts, or their displacement, such as additive Gaussian noise [18]. The intuition behind this work resides in the fact that both types of distortions will likely result in changes in the statistical distribution of the points along the three axes (x, y, z).

Given a point cloud $\mathcal{P}$, comprised of P points $p_i(x_i, y_i, z_i)$, we define the set $\mathcal{X}$ as the coordinates of all the points along the x-axis:

$$\mathcal{X} := \{x_i \mid p_i(x_i, y_i, z_i) \in \mathcal{P}\} \tag{1}$$

Sets $\mathcal{Y}, \mathcal{Z}$ are similarly defined. For each set $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, we also compute the relative histograms $H^{\mathcal{X}}, H^{\mathcal{Y}}, H^{\mathcal{Z}}$, defined as the probability $\mathbb{P}(x_i)$ that a point $p_i \in \mathcal{P}$ would have coordinate $x_i \in \mathcal{X}$ (respectively, $\mathbb{P}(y_i)$ for $y_i \in \mathcal{Y}$, and $\mathbb{P}(z_i)$ for $z_i \in \mathcal{Z}$). The number of bins N^B in the histogram is based on the maximum range of the set. For each of the three sets, a vector of features $\Phi^G_{\mathcal{S}}$ is then computed. To minimize the impact of the point cloud orientation in 3D space, the feature set Φ^G is obtained through max pooling, resulting in 7 features. Table I gives a definition of the feature vector Φ for a given set $\mathcal{S}$. Note that $\hat{\mathcal{S}}$ denotes the ordered list of $\mathcal{S}$.

B. Luminance-based features

When it comes to measure distortion on the color attributes of a point cloud content, it has been shown that computing global characteristics, such as color histograms, allows to better capture the perceptual level of degradation with respect to point-based solutions [10]. Following the recent literature, we

compute our features in the luminance channel, which shows better correlation with human perception of color [19]. We convert the color attributes R, G, B at each point $p_i \in \mathcal{P}$ using the matrix defined in ITU-R Recommendation BT.709 [20], in order to obtain the set L, comprised of all the luminance-channel attributes of point cloud $\mathcal{P}$. We also compute the luminance histogram H^L, defined as the probability $\mathbb{P}(\tilde{y_j})$ that a point $p_i \in \mathcal{P}$ has luminance value $\tilde{y_j}$. In this case, $N^B = 256$. The chosen set of features is defined similarly to geometry-based features in section III-A, and is summarized in Table I. Thus, the set Φ^L is comprised of 7 features.

C. Normal-based features

Normal attributes indicate the orientation of a given point in 3D space; as such, they are informative on the underlying planar surface of which the points can be considered as a sample. Thus, they have been used in the literature as an indicator for visual distortions in the geometrical domain of a given point cloud [3], [5]. In order to extract meaningful features from our point cloud contents, we first introduce the notion of *normal consistency* as a measure of the similarity between the normal vector of a point, with respect to the normal vectors of its neighbors. In particular, for each point $p_i \in \mathcal{P}$ with normal vector $\boldsymbol{n}_i$, we select the set $\mathcal{K}$ of k-nearest neighbors of p_i, sorted according to distance. Then, for each point $p_j \in \mathcal{K}$, we compute the angular similarity $\theta(i,j) \in \Theta$ between $\boldsymbol{n}_i$ and $\boldsymbol{n}_j$, following [5]. The result is matrix $\Theta \in \mathbb{R}^{P,k}$.

As there are two dimensions to Θ, the feature set Φ^N cannot be obtained by directly applying what seen in Table I. Moreover, the values under exam more closely resemble a continuous distribution, whereas sets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ and L are discrete. Thus, we redefine the feature vector to be more informative of the normal consistency attributes. In particular, we do not use the concept of mode, as the probability of a single value for continue distributions is equal to 0. Moreover, we define the histogram H^N as the probability $\mathbb{P}(\tilde{\theta_i})$ that the value $\tilde{\theta_i}$, obtained by averaging the values of θ across k, lies in a predetermined interval, induced by the number of bins N^B. The set Φ^N is comprised of 7 features, summarized in Table II.

D. Unified perceptual quality score

The feature sets Φ^G, Φ^Y, Φ^N, comprised of 21 features $\hat{f}_i$ extracted from a given distorted point cloud content, are compared to the features extracted from the corresponding reference content. In particular, for each pair of features $(f_i, \hat{f}_i)$, we compute the absolute difference $d_i = |f_i - \hat{f}_i|$. We then obtain our perceptual quality score as a linear combination of such differences:

$$PCM_{RR} = \sum_i w_i d_i. \tag{2}$$

The weights $w_i \in [0, 1]$ are obtained and validated via training on a point cloud dataset, as described in the following section.

TABLE III
PERFORMANCE RESULTS OF THE PROPOSED METRIC IN THE CROSS-VALIDATION ON M-PCCD [21].

	SRCC ↑	**PLCC** ↑
PCM_{RR} (LpOCV)	0.826 (σ = 0.102)	0.798 (σ = 0.111)
PCM_{RR} (MCCV)	**0.907** (σ = 0.028)	**0.901** (σ = 0.029)
D1	0.759	0.720
D2	0.807	0.756

IV. RESULTS

A. Experimental setup

To train and evaluate our metric, we use the publicly available dataset M-PCCD [21], consisting of subjective and objective quality scores assigned to 8 point cloud contents (4 human bodies, 4 inanimate objects) under compression distortions, resulting in 232 stimuli. We extract the features described in Section III from the reference and distorted point clouds. As normal vectors were not given along with the dataset, we estimate them using the built-in MATLAB function. For the computation of Θ, we set $k = 9$, while the corresponding $N^B = 300$. Features are computed and stored in single float precision, requiring 84 bytes to be transmitted.

To obtain the weights w_i, we run a linear optimization algorithm, which aims at maximizing the Pearson Linear Correlation Coefficient (PLCC) between our metric PCM_{RR} and the corresponding subjective scores, after logistic fitting [22]. To see how the metric generalizes to previously unseen contents, we perform Leave p Out cross-validation (LpOCV) by selecting 4 contents out of the 8 provided to be used for testing, and training on the remaining 4. We repeat the procedure for all $\binom{8}{4} = 70$ pairs, and we report the average performance. Additionally, we perform Monte Carlo cross-validation (MCCV) with 100 random splits on our dataset (80% training, 20% test). Finally, we perform cross-dataset validation on 3 additional point cloud datasets: PointXR [23], IRPC [24], and SJTU-PCQA [25], using the optimal weights defined in the LpOCV step. The PointXR dataset includes 5 static point cloud contents depicting cultural heritage, encoded using an octree-based geometry module and 2 different color compression schemes. The IRPC dataset includes 6 point cloud contents under 3 types of geometric compression distortion, which are evaluated in 3 different settings, only one of which includes undistorted color information. The SJTU-PCQA dataset, finally, includes 9 contents under 7 types of distortions, both on the color and geometry domain. Following ITU-T Recommendations P.1401 [22], the performance of our metric is assessed using the Spearman Rank Correlation Coefficient (SRCC), along with the aforementioned PLCC, to account for monotonicity and linearity, respectively, after logistic fitting.

B. Results

Table III reports the mean correlation coefficient, along with the corresponding standard deviation σ, obtained through cross-validation in dataset M-PCCD [21]. To offer a comparison with widely-used metrics in the state of the art, we also report the results of metrics D1 and D2, as defined

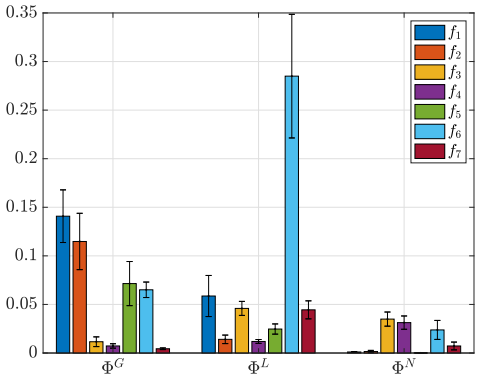

Fig. 1. Optimal weights for each feature, averaged across the LpOCV splits, with relative 95% confidence intervals.

TABLE IV
CROSS-DATASET VALIDATION ON POINTXR [23], IRPC [24] AND SJTU-PCQA [25]. WEIGHTS FROM LpOCV ON M-PCCD [21].

	PointXR	**IRPC**	**SJTU-PCQA-1**	**SJTU-PCQA-2**
SRCC ↑	0.928	0.302	0.820	0.830
PLCC ↑	0.956	0.434	0.821	0.821

and employed in the MPEG standardization efforts [1]. As no training of parameters is involved, correlation results are reported for the entire dataset. However, it should be noted that those metrics are full reference, thus including information from the full point cloud content, and only assess distortion in the geometrical domain. We refer the reader to the supplemental material, as well as [21] and [10], for a more complete benchmarking on the same dataset. It can be observed that our metric is outperforming the aforementioned FR solutions both in terms of PLCC and SRCC, for both cross-validation methods. Figure 1 depicts the optimal weight for each feature, averaged across the 70 pairs in the LpOCV, with relative confidence intervals. Features are grouped per feature set to facilitate comprehension. It can be observed that the two largest weights (0.285 and 0.141) are assigned to features f_6 in set Φ^L and f_1 in set Φ^G, which corresponds to the energy of the luminance histogram, and the mean in the geometry domain, respectively. Generally, the weights appear to be balanced between structure and color information, although less weight is given to normal vector features: set Φ^L accounts for 48.47% of the total weights, whereas sets Φ^G and Φ^N comprise 51.53% (41.54% and 9.99%, respectively).

Table IV reports the results of cross-dataset validation on the three datasets PointXR [23], IRPC [24] and SJTU-PCQA [25], using the weights illustrated in Fig. 1. For the PointXR dataset we select the alternating variant, as it was associated with better accuracy. To ensure a fair evaluation, for the IRPC dataset we select the experiment in which both geometry and color were rated. Similarly, for the SJTU-PCQA dataset we select distortions Downsampling + Color Noise (SJTU-PCQA-1) and Geometry Gaussian Noise + Color Noise (SJTU-PCQA-2). The best performance is obtained on the PointXR dataset, followed by the SJTU-PCQA-2 and SJTU-PCQA-1 datasets. This might be explained by the fact that the type of distortions in the two datasets more closely resemble the training dataset M-PCCD, as they apply geometrical and color distortions simultaneously. In particular, despite the fact that different types of geometric distortions are present in SJTU-PCQA-2 and SJTU-PCQA-1, our metric is able to capture random displacements as well as variations in the number of points, as shown by the similar performance in the two datasets. On the other hand, the worst results are obtained for dataset IRPC. The dataset was created by applying distortions uniquely on the geometry domain, while the color information was uncompressed, and obtained through recoloring. Thus, the color information may act as a distractor [12], hiding impairments in the geometry domain. As our proposed weights heavily include a measure of color distortion, a less than optimal performance in this dataset is to be expected.

TABLE V
ABLATION STUDIES ON M-PCCD [21].

	Φ^G	Φ^L	Φ^N	$\Phi^{\{G,L\}}$	$\Phi^{\{G,N\}}$	$\Phi^{\{L,N\}}$	PCM_{RR}
SRCC ↑	0.694	0.841	0.655	0.850	0.761	0.817	0.826
PLCC ↑	0.672	0.816	0.667	0.822	0.754	0.790	0.798

We refer the readers to the supplemental material for the exact value of the optimal weights for each feature. Using the optimal weights, the performance on the entire M-PCCD dataset is PLCC = 0.868 and SRCC = 0.889.

C. Ablation studies

In order to understand the prediction power of our features, we run ablation studies on the dataset M-PCCD [21]. In particular, we run our linear optimization algorithm on the feature sets Φ^G, Φ^L, Φ^N (7 features each), and pairwise combinations $\Phi^{\{G,L\}}$, $\Phi^{\{G,N\}}$ and $\Phi^{\{L,N\}}$. Results are summarized in Table V, where they are compared to the full feature set $PCM_{RR} := \Phi^{\{G,L,N\}}$. Results are shown on average over LpOCV splits. Among the single feature sets, Φ^L achieves the best performance, showing that luminance distortion is the best indicator of global visual quality on the dataset under exam. However, clear gains can be observed when geometry information is added to the luminance features, as shown by the increase in performance for $\Phi^{\{G,L\}}$. Moreover, it is shown that combining geometry and normal features greatly improves the performance over the single sets, achieving closer performance to Φ^L. The best performance is achieved when only geometry and luminance features are used, which might lead to the assumption that normal features are not necessary. However, when using only the features in set $\Phi^{\{G,L\}}$, a poorer performance was obtained across other datasets (e.g., PointXR, SRCC = 0.899; SJTU-PCQA-1, SRCC = 0.794) indicating worse generalization capabilities.

V. CONCLUSION

In this paper, we propose a reduced reference metric for visual quality assessment of point cloud contents. Our set of features, extracted from reference point cloud contents, requires few bytes to be transmitted alongside the content. Results on 4 publicly-available datasets demonstrate the informative value of our proposed features, and confirm the high performance of our metric. Future work will focus on testing the metric on a larger array of degradations, and test whether the feature space can be further reduced. An implementation of the metric can be found here: https://github.com/cwi-dis/PCM_RR.

REFERENCES

[1] Sebastian Schwarz, Marius Preda, Vittorio Baroncini, Madhukar Budagavi, Pablo Cesar, Philip A Chou, Robert A Cohen, Maja Krivokuća, Sébastien Lasserre, Zhu Li, Joan Llach, Khaled Mammou, Rufael Mekuria, Ohji Nakagami, Ernestasia Siahaan, Ali Tabatabai, Alexis M. Tourapis, and Vladyslav Zakharchenko, "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, March 2019.

[2] Daniel Girardeau-Montaut, Michel Roux, Raphaël Marc, and Guillaume Thibault, "Change detection on points cloud data acquired with a ground laser scanner," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, no. part 3, pp. W19, 2005.

[3] Dong Tian, Hideaki Ochimizu, Chen Feng, Robert Cohen, and Anthony Vetro, "Geometric distortion metrics for point cloud compression," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3460–3464.

[4] Paolo Cignoni, Claudio Rocchini, and Roberto Scopigno, "Metro: measuring error on simplified surfaces," in *Computer graphics forum*. Wiley Online Library, 1998, vol. 17, pp. 167–174.

[5] Evangelos Alexiou and Touradj Ebrahimi, "Point cloud quality assessment metric based on angular similarity," in *2018 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2018, pp. 1–6.

[6] Alireza Javaheri, Catarina Brites, Fernando Pereira, and Joao Ascenso, "A generalized Hausdorff distance based quality metric for point cloud geometry," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 2020, pp. 1–6.

[7] Alireza Javaheri, Catarina Brites, Fernando Pereira, and Joao M Ascenso, "Mahalanobis based point to distribution metric for point cloud geometry quality evaluation," *IEEE Signal Processing Letters*, 2020.

[8] Gabriel Meynet, Julie Digne, and Guillaume Lavoué, "PC-MSDM: A quality metric for 3D point clouds," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2019, pp. 1–3.

[9] Gabriel Meynet, Yana Nehmé, Julie Digne, and Guillaume Lavoué, "PCQM: A full-reference quality metric for colored 3D point clouds," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 2020, pp. 1–6.

[10] Irene Viola, Shishir Subramanyam, and Pablo César, "A color-based objective quality metric for point cloud contents," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6.

[11] Rafael Diniz, Pedro Garcia Freitas, and Mylène CQ Farias, "Towards a point cloud quality assessment model using local binary patterns," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 2020, pp. 1–6.

[12] Evangelos Alexiou and Touradj Ebrahimi, "Towards a point cloud structural similarity metric," in *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2020, pp. 1–6.

[13] Ricardo De Queiroz and Philip Chou, "Motion-compensated compression of dynamic voxelized point clouds," *IEEE Transactions on Image Processing*, vol. PP, pp. 1–1, 05 2017.

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

[15] Evangelos Alexiou and Touradj Ebrahimi, "Exploiting user interactivity in quality assessment of point cloud imaging," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2019, pp. 1–6.

[16] Zhou Wang, Guixing Wu, Hamid R Sheikh, Eero P Simoncelli, En-Hui Yang, and Alan C Bovik, "Quality-aware images," *IEEE transactions on image processing*, vol. 15, no. 6, pp. 1680–1689, 2006.

[17] Judith Redi, Paolo Gastaldo, Ingrid Heynderickx, and Rodolfo Zunino, "Color distribution information for the reduced-reference assessment of perceived image quality," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 20, pp. 1757 – 1769, 01 2011.

[18] Evangelos Alexiou and Touradj Ebrahimi, "On subjective and objective quality evaluation of point cloud geometry," in *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2017, pp. 1–3.

[19] Stefan Winkler, Murat Kunt, and Christian J. van den Branden Lambrecht, "Vision and video: models and applications," in *Vision Models and Applications to Image and Video Processing*, pp. 201–229. Springer, 2001.

[20] ITU-R BT.709, "Parameter values for the HDTV standards for production and international programme exchange," International Telecommunication Union, June 2015.

[21] Evangelos Alexiou, Irene Viola, Tomás M. Borges, Tiago A. Fonseca, Ricardo L. de Queiroz, and Touradj Ebrahimi, "A comprehensive study of the rate-distortion performance in MPEG point cloud compression," *APSIPA Transactions on Signal and Information Processing*, vol. 8, pp. e27, 2019.

[22] ITU-T P.1401, "Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models," International Telecommunication Union, July 2012.

[23] Evangelos Alexiou, Nanyang Yang, and Touradj Ebrahimi, "PointXR: A toolbox for visualization and subjective evaluation of point clouds in virtual reality," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6, Dataset and software release information available at: https://www.epfl.ch/labs/mmspg/pointxr/.

[24] Alireza Javaheri, Catarina Brites, Fernando Pereira, and Joao Ascenso, "Point cloud rendering after coding: Impacts on subjective and objective quality," 2019.

[25] Q. Yang, Z. Ma, Y. Xu, R. Tang, and J. Sun, "Predicting the perceptual quality of point cloud: A 3D-to-2D projection-based exploration," *Submitted to IEEE Transactions on Multimedia*, 2020.