

A color-based objective quality metric for point cloud contents

Irene Viola*, Shishir Subramanyam*[†], and Pablo Cesar*[†]

*Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

[†]TU Delft, Delft, The Netherlands

Abstract—In recent years, point clouds have gained popularity as a promising representation for volumetric contents in immersive scenarios. Standardization bodies such as MPEG have been developing new compression standards for point cloud contents to reduce the volume of data, while maintaining an acceptable level of visual quality. To do so, reliable metrics are needed in order to automatically estimate the perceptual quality of degraded point cloud contents. Whereas several objective metrics have been developed to assess the geometrical impairment of degraded point cloud contents, fewer publications have been devoted to evaluating color artifacts. In this paper, we propose new color-based objective metrics for quality evaluation of point cloud contents. Our work extracts color statistics from both reference and degraded point cloud contents, in order to assess the level of impairment. Using publicly available ground-truth data, we compare the performance of our proposed work with state-of-the-art metrics, and we demonstrate how the color metrics are able to achieve comparable results with respect to widely adopted solutions. Moreover, we combine color- and geometry-based metrics in order to provide a global quality score. The novelty of our works resides in simultaneously taking both degradation types into account, while being independent of the rendering process. Results show that our solution is able to overcome the limitations of focusing on only one type of degradation, achieving better performance with respect to current metrics.

Index Terms—point cloud, objective quality metrics, compression

I. INTRODUCTION

In recent years the availability of low cost sensors, affordable head mounted displays and the computational power of commodity hardware have allowed content providers to serve a broad range of users with immersive virtual reality experiences. In order to allow free movement within the virtual world these immersive experiences require volumetric photorealistic reconstructions of real world objects. Point clouds have emerged as a popular format to represent such reconstructions. A point cloud is a collection of points sampled from the surface of the object with a specified geometry in the form X, Y and Z coordinates along with some associated attributes at each point. Point clouds are typically captured by depth sensors or with multi-camera capture using photogrammetry. They allow a simple representation of geometry using a set of un-ordered points that require no pre-processing to reconstruct

the surface. Point clouds are also versatile as there are no restrictions imposed on the attributes that can be recorded at each point such as colour, surface normals, transparency and material properties. Point clouds can contain millions of points and require compression in order to store and transmit them efficiently. Uncompressed point cloud sequences can require several Gigabytes per second to transmit. Point cloud compression has received significant research interest in recent years including an MPEG standardisation activity with a new compression standard expected to be released later this year [1]. The upcoming MPEG VPCC codec has been shown to reduce the bandwidth requirements to the range of 3 to 55 Mbps [1].

In order to evaluate the effectiveness of lossy compression and to measure the errors introduced, a quality degradation metric is required. In previous research, several objective metrics to measure geometrical distortions have been proposed. However, such metrics are not able to capture the impact of color distortions on the visual quality of point clouds. In this paper, we propose the use of global color statistics, such as color histograms and correlograms, to compare the level of distortion of a degraded point cloud with respect to its unaltered reference. Moreover, we introduce a new metric, obtained through a linear regression between geometry-only and color-only metrics, in order to better approximate the perceived visual quality.

The contributions of this paper are two-fold. We present a thorough investigation of color-based statistic as a means to extract color features from point cloud data, and we employ several distance metrics in order to quantify the discrepancies between an altered point cloud and its original reference. To test the validity of color-based metrics, we use a publicly available dataset of objective and subjective quality scores for point cloud contents under compression distortions. Then, we analyze the performance improvements brought by combining geometry- and color-based approaches in a single, rendering-independent metric.

II. RELATED WORK

Full reference point cloud distortion metrics can be broadly categorized as point based or projection based. Point based metrics rely on finding correspondences for all points in the distorted cloud with points from the original cloud. Projection based metrics rely on the fact that point clouds are typically rendered before being consumed. Therefore, distortions can be

This paper was partly funded by the European Commission as part of the H2020 program, under the grant agreement 762111, “VRTTogether” (<http://vrtogether.eu/>).

measured by rendering the original and distorted point clouds from the the same viewpoint, and using image distortion metrics on the resulting 2D images.

A. Point Based Metrics

Point based metrics rely on identifying correspondences between points in the distorted decompressed cloud and the original uncompressed cloud. The euclidean distance between corresponding points can then be used to measure the distortion [2] introduced by compression. Cignoni et al proposed a mesh based approach where a mesh is created from the original point cloud in order to recreate the surfaces from which the points were sampled. The distance between the surfaces and the points in the distorted point cloud can then be measured [3]. However this method relies heavily on the mesh construction process and these are often tuned differently for specific reconstructed objects. Tian et al [4] propose a point to plane metric. The authors compute the normal to the surface at every point in the original uncompressed point cloud as an indication of the local surface. The displacement of every corresponding point in the distorted point cloud is then projected onto the normal to calculate the point to plane distance. Linear approximations of surfaces, as well as more complex curvature approximations, have also been used to quantify the angular distortions between two sets of points [5], [6]. Once the point distances have been calculated the overall geometric distortion can then be measured using mean squared error (MSE) or Hausdorf Distance. The quality can then be expressed using a Peak Signal to Noise Ratio (PSNR). Once point correspondences between the original and distorted cloud have been identified the changes to the color attribute introduced during compression can also be measured at each point.

B. Projection Based Metrics

Point clouds are typically rendered using graphics engines, by placing a primitive such as a triangle of fixed size at each point location. The primitives are then projected onto a 2D surface during rasterization, before being viewed by the user. The final image seen by the user can be used to measure the distortion, by comparing the images for the original and the distorted point cloud from the same viewport. Existing image distortion metrics such as SSIM, VIF and PSNR can be used to measure the distortion. This approach has been used by Quiroz et al [7] to drive rate distortion optimization in their codec design. Alexiou et al. [8] investigated the impact of the number of viewports on the performance of the metric, and proposed a weighting system based on user interaction. However, this approach is not suitable to be used in rate-distortion evaluation for lossy compression, as a large number of views would be needed in order to evaluate the geometry of the entire object.

III. COLOR-BASED OBJECTIVE QUALITY METRICS

In order to measure distortions introduced by compression on the colour domain, 2D image based metrics, such as PSNR, have been applied on the color information of couples of points

taken from the reference and distorted point cloud. The metrics thus capture variations in color on a point-to-point basis. However, these metrics fail to capture the broad variations in texture, often resulting in poor prediction capabilities [9].

The intuition behind this paper resides in trying to capture the difference between colored point clouds from a global perspective. We use color statistics, such as color histogram, to analyze how the general distribution of the colors changes when artifacts are introduced. We also incorporate spatial information using the concept of color correlogram [10].

A. Histogram-based objective quality metric

In image processing, color histograms represent the probability distribution of pixel values for the entire image. Distortions applied on the color channel, such as compression artifacts, tend to modify the statistical distribution of the colors, for example by applying quantization. Thus, computing the distance between the color histogram of a distorted model with respect to the one obtained from its unaltered reference, will give us an idea of the level of distortion.

We propose the use of histogram distance as a measure of distortion of a test point cloud with respect to a reference. We compute the color histogram on the luminance channel, which has been shown to better correlate with human perception of color [11]. Moreover, following the literature [12], we apply a weighted average between the distances associated with the $YCbCr$ channels, in order to have a distance metric that takes chrominance into account:

$$dist_{YCbCr} = \frac{6 \cdot dist_Y + dist_{Cb} + dist_{Cr}}{8}. \quad (1)$$

Several distances can be defined between two histograms. In our work, we compare the L_1 (Cityblock), L_2 (Euclidean) and L_∞ (Chebyshev) distances, as well as the χ^2 distance and the Earth Moving Distance (EMD) [13].

B. Correlogram-based objective quality metric

The concept of color correlogram was first introduced by Huang et al. [10], as a way to improve image indexing and comparison by introducing spatial correlation information to the color histogram. For 2D images, it is defined as the probability, for any pixel with color c_i , that a pixel at k -distance from it has color c_j .

As point cloud color information does not generally lie on a regular, fully occupied grid, the concept of k -distance becomes less meaningful, as it is not certain that, for a given point p_l , there will be a point p_m at distance k . In order to apply the correlogram concept to our data, we defined the color correlogram $\gamma^{(k)}$ as the probability, for any point p_l in point cloud \mathcal{P} with color $c_i \in [0, N - 1]$, that the k -nearest neighboring point p_m has color $c_j \in [0, N - 1]$:

$$\gamma^{(k)}(c_i, c_j) := \Pr_{p_l \in \mathcal{P}_{c_i}, p_m \in \mathcal{P}} [p_m \in \mathcal{P}_{c_j} \mid p_m \in \mathcal{N}_k(p_l)] \quad (2)$$

in which $\mathcal{N}(p_l)$ is the set of K -nearest neighbours of p_l , and \mathcal{P}_{c_i} is the set of points belonging to point cloud \mathcal{P} with

TABLE I: Performance indexes for histogram-based objective quality metrics, separately for luminance channel and weighted luminance and chroma channels (best performance in bold).

		Linear				Cubic			Logistic		
		SRCC	PLCC	RMSE	OR	PLCC	RMSE	OR	PLCC	RMSE	OR
Y	L_1	0.8679	0.5641	1.1233	0.8664	0.8176	0.7833	0.8578	0.8351	0.7485	0.7845
	L_2	0.8841	0.5318	1.1521	0.8707	0.8195	0.7796	0.8362	0.8532	0.7096	0.7802
	L_∞	0.8492	0.4673	1.2027	0.8621	0.7459	0.9061	0.8707	0.8157	0.7893	0.8103
	χ^2	0.8750	0.3718	1.2628	0.8707	0.6487	1.0353	0.8966	0.8321	0.7596	0.8319
	EMD	0.7110	0.4960	1.1812	0.8664	0.6452	1.0393	0.8534	0.6623	1.0273	0.8448
YCbCr	L_1	0.7879	0.5720	1.1158	0.9009	0.7521	0.8966	0.8793	0.7538	0.8939	0.8578
	L_2	0.7020	0.5161	1.1651	0.8836	0.6698	1.0101	0.8491	0.6763	1.0021	0.8319
	L_∞	0.6373	0.4615	1.2068	0.8793	0.6362	1.0705	0.8879	0.6327	1.0537	0.8319
	χ^2	0.7412	0.4199	1.2346	0.8836	0.6494	1.0344	0.8491	0.6873	0.9890	0.8621
	EMD	0.7036	0.5019	1.1766	0.8707	0.6438	1.0409	0.8448	0.6544	1.0346	0.8578

color c_i . The resulting correlogram matrix is $\gamma^{(k)} \in \mathbb{R}^{N \times N}$; in the case of 8-bit color information, $N = 256$. By definition, the correlogram is a symmetric matrix. It is worth noting that for $k = 1$, the nearest neighbor corresponds to the point under consideration. In this case, the resulting correlogram is a diagonal matrix, whose diagonal corresponds to the color histogram. Conversely, for $k = \infty$, the matrix contains the color histogram in every row.

Several norms and distances can be defined between two matrices. In this work, we use the L_1 , L_2 , L_{Fro} and L_∞ norms to get a global measure of difference between the correlograms associated with reference and distorted point cloud, along with the pairwise χ^2 distance and EMD. Moreover, we use the d_1 distance as defined in [10], as well as the sets of features described in [14]. For the latter, we create vector f_k , for both reference and distorted contents, by concatenating several features, namely diagonal energy, entropy, contrast, homogeneity and energy ratio, as defined in the aforementioned paper. We then compute the distance between feature vectors using the Euclidean distance. As with the color histogram, we compute the color correlogram in the $YCbCr$ space, and we report results for the Y channel, as well as the results obtained with the weighted average in (1).

IV. RESULTS

A. Experimental data

To test the validity of our metric, we used a publicly available dataset [9]. The dataset consists of subjective and objective quality scores given to 8 point cloud contents, of which 4 depicting human bodies, under compression distortions, for a total of 232 stimuli. We recreated the distorted contents following the description given by the authors, and we applied our metric on the resulting point clouds. We then computed several performance indexes on the data, to assess the prediction power of our metrics with respect to the subjective ground truth, expressed as Mean Opinion Score (MOS). In particular, Pearson Linear Correlation Coefficient (PLCC), Spearman Rank Correlation Coefficient (SRCC), Root Mean Square Error (RMSE) and Outlier Ratio (OR) were chosen to account for linearity, monotonicity, accuracy and consistency, respectively, following ITU-T Recommendations P.1401 [15].

Before computing the indexes, linear, cubic and logistic fitting was applied on the results of the objective metrics, similarly to what was done in [9].

B. Histogram-based objective quality metric

Table I depicts the results of the performance indexes for the histogram-based objective quality metrics. Results are shown separately for the luminance channel and for the $YCbCr$ weighted average.

Among the metrics under consideration, the L_2 distance on the Y channel achieves the best results for both cubic and logistic fitting, according to all the performances indexes under consideration. For linear fitting, the L_1 distance achieves the best performance in terms of PLCC, whereas for monotonicity and accuracy, the L_2 distance is the best performing one. In terms of consistency, the L_∞ distance achieves the best result with linear fitting.

Results obtained by performing a weighted average on the $YCbCr$ color space show that aberrations in the chrominance space do not correlate with human perception of distortions. In fact, incorporating chrominance distance information leads to generally poorer results in terms of performance indexes, with the sole exception of the L_1 metric with linear fitting, for which slightly better linearity and accuracy is achieved.

C. Correlogram-based objective quality metric

Figure 1 depicts the PLCC index obtained with the color correlogram comparison for values of k ranging from 2 to 20, for Y and weighted $YCbCr$ color spaces. Due to space restrictions, only the results from the best performing metric d_1 are displayed. A slight loss in performance can be seen, for all metrics, for $k = 3$, followed by a slow increase as values of k get larger. The best performance is achieved for $k = 20$ with cubic fitting, closely followed by logistic fitting. Figure 2 displays the distance d_1 results against the MOS, for all stimuli, separately for the cases $k = 3$ (worst performing case) and $k = 20$ (best performing case). Logistic fitting is also displayed. From the comparison of scatterplots, it is clearly visible how the metric for $k = 3$ is not able to fully capture the perceptual color loss. For example, a value of $d_1 = 1$ can be associated to stimuli that were rated as high as 4.5 or as low as 1, showing that the metric is not able to accurately

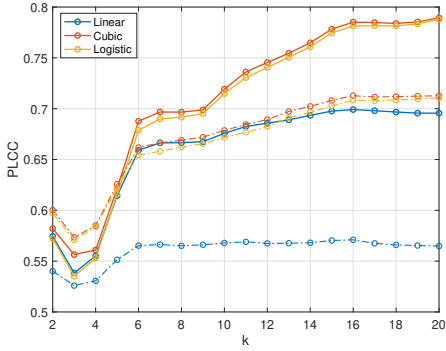


Fig. 1: PLCC index for correlogram comparison using d_1 distance, for different values of k (solid line for Y channel, dashed line for $YCbCr$).

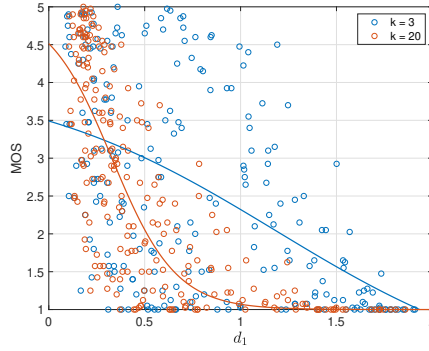


Fig. 2: Results of comparison between correlograms using d_1 distance, for $k = 3$ and $k = 20$, with relative logistic fitting.

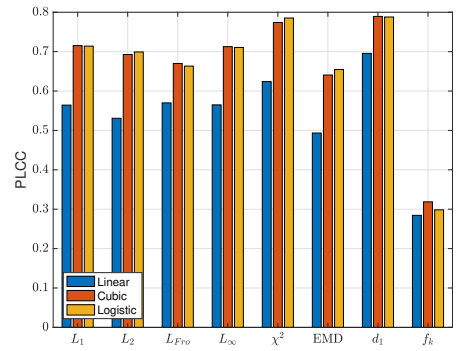


Fig. 3: PLCC indexes obtained with all distance metrics for correlogram comparison, performed on the Y channel, for $k = 20$.

predict the visual quality of the stimuli. On the other hand, the scatterplot for $k = 20$ appears more compact, following a monotonically decreasing trend. A comparison between the SRCC values obtained in the two cases (0.5553 for $k = 3$, 0.8075 for $k = 20$) confirms the observed trends.

Figure 3 depicts the values of the PLCC index achieved by every color correlogram distance under exam, for $k = 20$. Results are shown for linear, logistic and cubic fittings. Among the metrics, χ^2 and d_1 are the best performing ones, followed by the L -norms. EMD shows a slightly poorer performance, achieving a PLCC of 0.6595 with logistic fitting. Extracting features from the correlogram leads to a significantly poorer performance with respect to all the other metrics under consideration. The features to be extracted from the correlograms were proposed in relation to 2D image contents under transmission artifacts; however, results indicate that the same features do not correlate well with human perception of 3D contents.

An increase in performance is observed with higher values of k , for all metrics. However, even the best-performing metric fails to reach the same performance values as the ones observed with the histogram-based metrics. In fact, with the exception of the EMD metric, all metrics considered for the histogram case achieve a higher PLCC than the best-performing metric in the correlogram case. There are a few possible explanations for the phenomenon. The correlogram works by introducing spatial correlation as a way to describe changes in color statistics; as such, it does not describe the distribution of colors in the content, but rather the color distribution of points in relation with each other. Correlation with close neighbors leads to poorer results, as visible in Fig. 1. This is justifiable considering that local variations of colors may not be perceivable by the users, especially if the content is observed from a distance. Moreover, it is also possible that the size of the splats used to render the content masks some of the variations. As k increases, larger and larger neighborhoods of points are considered, thus moving the analysis to a more global scale.

D. Comparison with existing metrics

In order to assess the validity of our color-based metrics, we compare its performance with most commonly used objective quality metrics for point cloud contents. A comprehensive evaluation was performed in [9], along with the dataset used in this work. In Table II we report some of the results, along with our two proposed metrics for comparison. We refer the reader to the original work for a more exhaustive benchmark.

It can be observed that histogram-based metric with L_2 distance is able to outperform the majority of the metrics under consideration, and has comparable performance with respect to the best-performing metric $p2plane_{MSE}$ for logistic fitting. In particular, the $p2plane_{MSE}$ metric has slightly higher values of PLCC (0.858 against 0.853), as well as lower values of RMSE (0.700 against 0.710). On the other hand, our metric achieves better OR (0.780 for our metric, as opposed to 0.832 for the $p2plane_{MSE}$), whereas for SRCC, the two metrics are equivalent.

A notable increase in performance can be seen when using our histogram-based metric with respect to the traditional point cloud color-based metrics MSE_{YCbCr} and $PSNR_{YCbCr}$. The latter metrics use a point-to-point approach to compute the error between reference and distorted point clouds. However, as already observed in the correlogram-based metrics, a local approach does not seem to adequately approximate the perception of point cloud contents.

E. Combining color and geometry

Point-based metrics are traditionally used to assess the quality of geometry-only or color-only distortions. The histogram-based metric, as it is, can reflect geometry distortions, if they result in a loss of points. In that case, the statistics of the color distribution would likely change between the reference and distorted contents, even if the color information itself is not altered. On the other hand, geometry-only metrics are not able to assess the perceived quality of point cloud contents, if the distortions are introduced only on the color values without altering the topology.

TABLE II: Performance indexes for traditional objective quality metrics for point cloud contents [9]. Last two rows represent the color-only metric $H_{L_2}^Y$ and the combined metric d_{gc} presented in this paper.

	Linear				Cubic			Logistic		
	SRCC	PLCC	RMSE	OR	PLCC	RMSE	OR	PLCC	RMSE	OR
p2point _{MSE}	0.868	0.484	1.193	0.858	0.691	0.985	0.841	0.845	0.728	0.841
p2plane _{MSE}	0.884	0.448	1.219	0.862	0.663	1.021	0.841	0.858	0.700	0.832
PSNR - p2point _{MSE}	0.759	0.679	0.935	0.833	0.723	0.880	0.801	0.720	0.885	0.819
PSNR - p2plane _{MSE}	0.807	0.711	0.896	0.833	0.757	0.833	0.833	0.756	0.834	0.852
MSE _{YCbCr}	0.663	0.410	1.244	0.884	0.528	1.158	0.888	0.653	1.033	0.849
PSNR _{YCbCr}	0.660	0.646	1.040	0.879	0.654	1.032	0.866	0.653	1.033	0.849
PSNR	0.628	0.597	1.093	0.871	0.611	1.079	0.858	0.667	1.015	0.802
VIFp	0.742	0.697	0.978	0.853	0.716	0.951	0.823	0.698	0.977	0.858
$H_{L_2}^Y$	0.884	0.532	1.152	0.871	0.820	0.780	0.836	0.853	0.710	0.780
d_{gc}	0.920	0.490	1.186	0.866	0.806	0.804	0.875	0.904	0.585	0.724

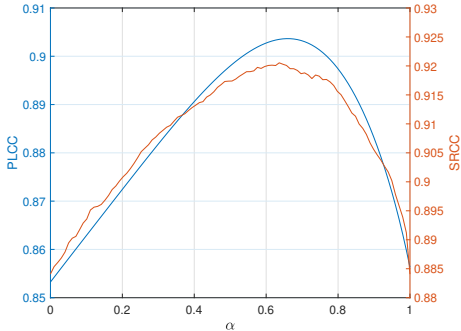


Fig. 4: Performance indexes PLCC and SRCC for values of α . For $\alpha = 0$, only the color-based metric is used; for $\alpha = 1$, only the geometry metric is selected.

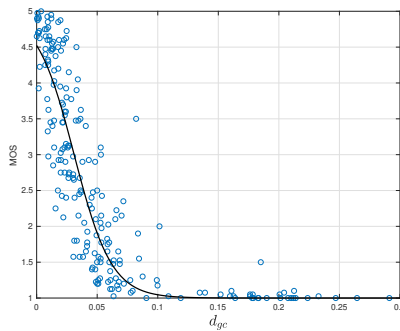


Fig. 5: Proposed metric d_{gc} against MOS values. To improve readability, results are shown for lower levels of d_{gc} only.

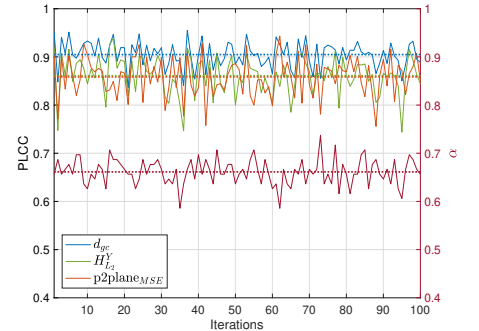


Fig. 6: PLCC index for every validation set for the proposed metric d_{gc} , along with the geometry- and color-based metrics, and chosen α .

To overcome the limitations of both approaches, we propose to combine geometry- and color-based metrics in an individual metric which can capture both deviations. In particular, we define a new metric d_{gc} as a linear combination of the geometry-only metric d_{geom} and the color-only metric d_{color} :

$$d_{gc} = \alpha \cdot d_{geom} + (1 - \alpha) \cdot d_{color}, \quad (3)$$

in which α is a real number in $[0, 1]$. In our case, the p2plane metric with MSE error $p2plane_{MSE}$ is selected as d_{geom} , whereas the L_2 -distance color histogram metric on the Y channel $H_{L_2}^Y$ is selected as d_{color} .

We performed a grid search on values of α , to understand whether an improvement in performance could be found by linearly combining the two metrics. Figure 4 displays the performances indexes PLCC and SRCC obtained for different values of α . Logistic fitting was applied on the metric d_{gc} before computing the indexes. As can be seen, combining the two metrics always leads to an increase in performance. In particular, the lowest indexes are obtained for $\alpha = 0$, which corresponds to using only the histogram-based color metric, and $\alpha = 1$, for which only the geometry metric is used. The highest value of linear correlation is reached for $\alpha = 0.6597$, for which $PLCC=0.9037$. In terms of SRCC, the highest value of 0.9205 is reached for $\alpha = 0.6364$. Figure 5

shows the scatterplot of our proposed metric against the MOS, for $\alpha = 0.6597$. For the best-performing α , the null hypothesis of equivalence between the PLCC associated with d_{gc} , and the PLCC obtained with $p2plane_{MSE}$ and $H_{L_2}^Y$, is rejected, according to the procedure described in [15].

In order to validate that the improvement in performance was not caused by overfitting, we performed 100 random splits on the data (80% training, 20% test set), and we computed the optimal level of α in terms of PLCC. We then used the α value we defined on the training set to compute d_{gc} on the test data. We also computed the PLCC obtained on the test set for the geometry- and color-based metrics alone ($p2plane_{MSE}$ and $H_{L_2}^Y$, respectively), to assess whether an increase in performance could be observed. Results of the iterations are shown in Figure 6. The mean value over the iterations is shown with a dotted line. Small variations can be observed in the performance index, depending on the composition of the test set. Nonetheless, the combined metric d_{gc} achieves better results in term of PLCC with respect to the color-only metric $H_{L_2}^Y$ and geometry-only metric $p2plane_{MSE}$ on 95 of the 100 iterations under consideration. On average, d_{gc} achieves a PLCC of 0.9048, $\sigma = 0.0269$, whereas $H_{L_2}^Y$ and $p2plane_{MSE}$ achieve a PLCC of 0.8610 ± 0.0395 and 0.8584 ± 0.0410 , respectively. In terms of SRCC, d_{gc} has an

average value of 0.9152 ± 0.0262 , whereas $H_{L_2}^Y$ achieves a value of 0.8826 ± 0.0380 , and $p2plane_{MSE}$ reaches an average value of 0.8777 ± 0.0352 .

V. DISCUSSION

The dataset we use to assess the validity of our metric has a comprehensive collection of compression artifacts, obtained using the upcoming MPEG standard for point cloud coding. In particular, the distortion points were based on the Common Test Conditions, which were defined by the experts in the MPEG community [16]. However, as it is, distortions to color and geometry are applied at the same time, i.e., there is no such case in which the geometry quantization parameter was left constant while the color parameter varied, or vice versa. This explains, for example, why a geometry-only metric is able to achieve a good performance in predicting the quality of compressed point cloud content, despite being blind to color artifacts. Our metric showed superior performance in the dataset, which in all respects represents the benchmarking standard for point cloud assessment. However, further tests with larger datasets, including geometry-only and color-only distortions, are needed in order to fully understand the generalization capabilities of our metric.

Objective quality assessment for image and video contents has been an active area of research in the past decades, and several models have been proposed in order to mimic the human perception of visual quality. In order to directly exploit the vast body of research that has been done on image and video objective quality assessment, it has been proposed to project the 3D model onto a planar grid. In this case, a 2D representation of the volumetric content from a certain viewpoint would be obtained, and its visual quality can then be assessed using popular metrics, such as PSNR or VIF [8]. The method is usually praised for its ability to simultaneously capture both geometry and color degradations. However, one main limitation resides in the fact that rendering parameters, such as splat size and orientation, distance from the camera, and viewpoint, are needed in order to compute the metric. On the other hand, the metric we propose has the advantage of combining color and geometry distortions, while being renderer-agnostic, i.e., no rendering parameter needs to be taken into consideration when computing the score. This represents an advantage when the metric needs to be used in contexts where multiple rendering parameters could be applied, or where the application is unknown, e.g., a streaming scenario in which different rendering devices could be used. Moreover, it combines local (in the form of geometric distortions) and global statistics (variations in color histogram), which are able to better capture how users perceive the contents.

VI. CONCLUSIONS

In this paper, we present an analysis of color-based metrics for visual quality assessment of point cloud contents. Specifically, we employ color statistics such as histograms and correlograms in order to assess the level of impairment of a

degraded point cloud with respect to its unaltered reference. We also combine color-only and geometry-only approaches to create a rendering-independent objective metric that is able to simultaneously capture the distortions brought in the topology and texture of the point cloud content. Results show that our approach is able to outperform current metrics in terms of performance indexes. Future work will focus on testing the metric on other types of distortions, as well as exploring different kinds of color statistics.

REFERENCES

- [1] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li, J. Llach, K. Mammou, R. Mekuria, O. Nakagami, E. Siahaan, A. Tabatabai, A. M. Tourapis, and V. 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] D. Girardeau-Montaut, M. Roux, R. Marc, and G. 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, p. W19, 2005.
- [3] 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.
- [4] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Geometric distortion metrics for point cloud compression," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3460–3464.
- [5] E. Alexiou and T. 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] G. Meynet, J. Digne, and G. 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.
- [7] R. De Queiroz and P. Chou, "Motion-compensated compression of dynamic voxelized point clouds," *IEEE Transactions on Image Processing*, vol. PP, pp. 1–1, 05 2017.
- [8] E. Alexiou and T. 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.
- [9] 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 Transactions on Signal and Information Processing*, vol. 8, p. e27, 2019.
- [10] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proceedings of IEEE computer society conference on Computer Vision and Pattern Recognition*. IEEE, 1997, pp. 762–768.
- [11] S. Winkler, M. Kunt, and C. J. van den Branden Lambrecht, "Vision and video: models and applications," in *Vision Models and Applications to Image and Video Processing*. Springer, 2001, pp. 201–229.
- [12] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC)," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1669–1684, 2012.
- [13] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *International journal of computer vision*, vol. 40, no. 2, pp. 99–121, 2000.
- [14] J. Redi, P. Gastaldo, I. Heynderickx, and R. 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.
- [15] ITU-T P.1401, "Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models," International Telecommunication Union, July 2012.
- [16] S. Schwarz, G. Martin-Cocher, D. Flynn, and M. Budagavi, "Common test conditions for point cloud compression," *Document ISO/IEC JTC1/SC29/WG11 w17766, Ljubljana, Slovenia*, 2018.