

Evaluation of point cloud features for no-reference visual quality assessment

Smitskamp, Gwennan; Viola, Irene; Cesar, Pablo

DOI

[10.1109/QoMEX58391.2023.10178459](https://doi.org/10.1109/QoMEX58391.2023.10178459)

Publication date

2023

Document Version

Final published version

Published in

2023 15th International Conference on Quality of Multimedia Experience, QoMEX 2023

Citation (APA)

Smitskamp, G., Viola, I., & Cesar, P. (2023). Evaluation of point cloud features for no-reference visual quality assessment. In *2023 15th International Conference on Quality of Multimedia Experience, QoMEX 2023* (pp. 147-152). (2023 15th International Conference on Quality of Multimedia Experience, QoMEX 2023). Institute of Electrical and Electronics Engineers (IEEE).
<https://doi.org/10.1109/QoMEX58391.2023.10178459>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Evaluation of point cloud features for no-reference visual quality assessment

Gwennan Smitskamp^{*†}, Irene Viola^{*}, and Pablo Cesar^{*†}
^{*}Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[†]TU Delft, Delft, The Netherlands

Abstract—The development and widespread adoption of immersive XR applications has led to a renewed interest in representations that are capable of reproducing real-world objects and scenes with high fidelity. Among such representations, point clouds have attracted the interest of industry and academia alike, and new compression solutions have been developed to facilitate their adoption in mainstream applications. To ensure the best quality of experience for the end-user in limited bandwidth scenarios, new full-reference objective quality metrics have been proposed, promoting features designed specifically for point cloud contents. However, the performance of such features to predict the quality of point cloud contents when the reference is not available is largely unexplored. In this paper, we evaluate the performance of features commonly used to model point cloud distortions in a no-reference framework. The obtained features are integrated into a quality value through a support vector regression model. Results demonstrate the potential of full-reference features for no-reference assessment.

Index Terms—3D model quality assessment, colored point cloud, no-reference quality assessment

I. INTRODUCTION

Recent years have witnessed the rise of eXtended Reality (XR) technologies, fuelled by developments in devices for capturing, delivering and rendering, as well as the demand for applications to use XR for use cases such as education, remote communication, gaming, and cultural heritage [1]. Such applications need to be populated with volumetric contents in order to enable 6 degrees of freedom navigation. Among others, point clouds have received significant interest from industry and academia as a relevant format for volumetric representation. A point cloud is a collection of points, defined by their coordinates in 3D space, as well as additional attributes that define certain properties of each point, such as texture or normal vector. In order to deliver a high-fidelity representation, millions of points might be needed for a given object; thus, significant effort has been spent in crafting efficient compression solutions to alleviate the storage and transmission requirements [2]–[4]. Such compression solutions

This work was supported through the European Commission Horizon Europe program, under the grant agreement 101070109, *TRANSMIXR* <https://transmixr.eu/>. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Directorate-General for Communications Networks, Content and Technology. Neither the European Union nor the granting authority can be held responsible for them.

979-8-3503-1173-0/23 \$31.00 ©IEEE

need objective quality metrics to automatically predict the visual quality of the point cloud against encoding degradations.

A way to split visual quality metrics in approach and application is whether the metric requires undistorted reference content to provide a measure of visual quality. Full-reference (FR) quality metrics compare distorted content with its original, while no-reference (NR) or blind metrics estimate the distortion from a single content. Using some information on the original content is referred to as reduced reference (RR). RR and NR metrics are fundamental to assess distortions when the pristine content is not available (for example at the receiver side of a transmission pipeline), or when it is not known in the first place (for example, to assess acquisition noise or enhancement processing such as super-resolution). However, most of the objective quality metrics that have been proposed to estimate and model visual distortions in point cloud contents, have been in the realm of FR assessment. Whereas early approaches for FR metrics focused on directly measuring the distance between attributes on a point basis and aggregating them into a global score, more recently, methods have been proposed to compute features based on point neighborhoods, which would describe certain properties of the point cloud content. Such features are generally defined and compared locally between the distorted and reference point clouds. The intuition behind this paper is to try to understand whether such locally defined features can be generalized and used for NR point cloud assessment.

In this paper, we investigate the effectiveness of adapting multiple types of features that have been defined for FR point cloud quality assessment in NR point cloud quality assessment methods. To do so, we transform known local descriptors to global descriptors, using different kinds of statistical distribution models, and we test their correlation with the point cloud subjective visual quality, using the framework defined in [5]. A set of hand-crafted local features are calculated per point or neighborhood. For each feature, descriptors based on the distribution of the values are extracted to create global properties of the point cloud. These properties are tested in a quality value prediction model to assess their potential to use in a NR point cloud quality metric.

II. RELATED WORK

FR metrics for point cloud contents can be divided into point-based or image-based, depending on whether the metric computation is performed on the original 3D domain, or

if a projection to 2D surfaces is applied before computing traditional image metrics [6], [7]. The latter have the advantage of offering a global view of the point cloud content, implicitly considering occlusions and visual masking in their computation; however, they are rendering-dependent, such that the performance is affected by the rendering parameters and environment, as well as the camera distance. Conversely, point-based metrics compute distortions by comparing attributes or features locally, either on a point basis or using neighborhoods. Notable exceptions are the histogram metrics proposed in [8] and [9], which compute global measures of color and geometry distortions without resorting to 2D projection.

Early work on point-based approaches focused on direct measurements of geometric error on a point basis, either by measuring the distance between matching points in reference and distorted contents (point-to-point [10]), by projecting the distance along the normal direction (point-to-plane [11]) or with respect to a local point distribution (point-to-distribution [12]), or by performing a comparison of normal vectors (plane-to-plane [13]). More recent methods have integrated color distortions, along with descriptors and measurements that better approximate the human visual system. PCQM [14] defines 8 features based on comparing curvature and luminance values between reference and distorted contents to obtain a global measure of distortion. PointSSIM [15] uses dispersion statistics of geometry, normal, curvature and color attributes to produce a structural similarity score. Yang et al. [16] define a measure of point cloud distortion based on multiscale gravitational potential energy, whereas Xu et al. [17] model the visual distortion of point cloud contents as the difference in elastic potential energy for each point. PointPCA [18] defines several geometric and textural descriptors based on PCA decomposition and uses statistical features to capture variations within neighborhoods. GraphSIM [19] constructs a graph from a sampled version of reference and distorted point clouds, and uses statistical moments defined on the color gradient to compute a measure of similarity.

Outside of the domain of FR metrics, few methods have been proposed in the realm of RR metrics [20], [21]. NR metrics have recently attracted some attention. A set of hand-crafted features is used in a deep-learning model [22] (geometric distance, mean curvature and luminance) and in a machine learning framework [5] (eigenvalue and color-based features) to learn a mapping to the ground-truth mean opinion score. PQA-net [23] maps the point cloud to 2D images using a multi-view projection strategy to extract a 384-dimensional feature vector, which is fed into two learning modules that combined calculate the quality of the degraded point cloud. Liu et al. [24] designed a sparse convolutional neural network to extract the hierarchical features, and predict quality scores with a regression module of a grand new dataset they created for these learning-based approaches. Tliba et al. [25] employ a shallow network to extract features from sparse patches of data, using both supervised and unsupervised learning to reach a global quality score.

TABLE I: Definition of FR features under use.

Feature	Definition
Curvature	$\lambda_3/(\lambda_1 + \lambda_2 + \lambda_3)$
Anisotropy	$(\lambda_1 - \lambda_3)/\lambda_1$
Linearity	$(\lambda_3 - \lambda_2)/\lambda_1$
Planarity	$(\lambda_2 - \lambda_3)/\lambda_1$
Sphericity	λ_3/λ_1
Variation	$\lambda_1/(\lambda_1 + \lambda_3 + \lambda_3)$
Omnivariance	$\sqrt[3]{\lambda_1 + \lambda_3 + \lambda_3}$
Eigenentropy	$\sum_{i \in \{1,2,3\}} \lambda_i \cdot \ln(\lambda_i)$
Sum of eigenvalues	$\lambda_1 + \lambda_2 + \lambda_3$
Density	$\sum_{n=1}^{ \mathbf{P}_i } \ x_n - x_i\ _2^2$
CV-PED	$g_{p_i} h_{p_i}$
PED	$m_{p_i} g_{p_i} h_{p_i}$
EPES-CO	$m_{p_i} h_{p_i}$
EPES	$1/2(m_{p_i} h_{p_i}) h_{p_i}^2$
CFGD	m_{p_i}/h_{p_i}

III. PROPOSED GLOBAL FEATURES

To transform FR local features to global descriptors of point clouds, we build upon the framework of Zhang et al. [5], where the authors reviewed the transformation of 5 geometric and 3 color features to a global descriptor using different statistical parameters. The proposed features were evaluated on their combined performance in a regression model to predict the quality score of a point cloud. Our extension to their models involves including new features adapted from the literature on FR point cloud quality metrics, as well as using a new statistical parameter that can aid in modeling the distribution of the features.

A. Features

The set of geometric descriptors is based on Principal Component Analysis to estimate the shape properties of a point or area of the point cloud. The process of estimating these features is explained in [5]: a neighborhood \mathbf{P}_i of k nearest points is defined for each point \mathbf{p}_i in point cloud \mathbf{P} :

$$\mathbf{P}_i = \text{knn}_k(\mathbf{p}_i) \quad (1)$$

The covariance matrix of this set is computed with :

$$C_i = \frac{1}{|\mathbf{P}_i|} \sum_{n=1}^{|\mathbf{P}_i|} (\mathbf{p}_n - \bar{\mathbf{p}}_i)(\mathbf{p}_n - \bar{\mathbf{p}}_i)^T \quad (2)$$

and three eigenvalues can be found through eigenvalue decomposition, where $\lambda_1 > \lambda_2 > \lambda_3$: $C_i \cdot v_j = \lambda_j \cdot v_j, j \in 1, 2, 3$.

With these eigenvalues, various geometry features can be calculated. In [5], the set of curvature, anisotropy, linearity, planarity, and sphericity is used. We extend it with additional features, which have been successfully used in FR metrics [18], namely Variation, Omnivariance, Eigensum and Eigenentropy. In addition to the PCA-based features, we define a geometric measure describing the density of a given neighborhood. We define it as the average Euclidean distance among the neighbors of a point.

In terms of color features, the framework of Zhang et al. [5] considers 3 color features l, a^*, b^* based on the CIELAB color space, and then aggregates them in a single color measure. We additionally consider the RGB, XYZ , and

HCL color spaces, and analyse the performance of each color attribute separately. In addition, we consider features that were proposed in the literature which combine both geometric and color properties, namely the Potential Energy Discrepancy (PED) [16], the Elastic Potential Energy Similarity (EPES) [17], and the Color Fluctuation over Geometric Distance (CFGD) [21]. We also test the geometry-only version of PED, namely CV-PED, and the elastic coefficient defined for EPES, namely EPES-CO. For all these features, three main concepts are introduced, namely mass, spatial field, and distance:

$$m_{p_i} = \begin{cases} \sum_{j=1}^3 w_j |(c_i)_j - (c_0)_j| + 1 & \text{if } \mathbf{c}_i, \mathbf{c}_0 \neq \phi \\ 1 & \text{otherwise} \end{cases} \quad (3)$$

$$g_{p_i} = (1 + e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_0\|_2^2}{\sigma^2}})^{-1} \quad (4)$$

$$h_{p_i} = \|\mathbf{x}_i - \mathbf{x}_0\|_2^2, \quad (5)$$

in which \mathbf{x} and \mathbf{c} are the position and color attributes of the point \mathbf{p}_i under exam, and w_j refers to the weight associated with each color attribute $j = \{R, G, B\}$ ($w_R = w_B = 1/4, w_G = 1/2$). For PED, the value is computed as $\text{PED}_{p_i} = m_{p_i} g_{p_i} h_{p_i}$; for CV-PED, we omit the color-based m value ($\text{CV-PED}_{p_i} = g_{p_i} h_{p_i}$). The elastic coefficient for EPES is computed as follows: $\text{EPES-CO}_{p_i} = m_{p_i} h_{p_i}$, and the computation of EPES integrates elastic energy ($\text{EPES}_{p_i} = 1/2(m_{p_i} h_{p_i})(h_{p_i}^2)$). Finally, CFGD is defined as $\text{CFGD}_{p_i} = m_{p_i}/h_{p_i}$. A summary of all features under consideration can be found in Table I.

B. Transforming local features to a global descriptor

The features we have extracted are computed on a point or neighborhood basis, forming a high-dimensional vector. To translate it to a global measure, the properties of the vector need to be extracted to model the distortion with a smaller number of features.

Zhang et al. [5] proposes extracting the mean μ , standard deviation σ , entropy H , generalized Gaussian distribution parameters (GGD), asymmetric generalized Gaussian distribution parameters (AGGD), and Gamma distribution parameters. In particular, they found that for feature vectors where higher or lower values are directly correlated with quality, mean and standard deviation are sufficient to extract enough information, whereas for features where the sparsity indicates quality, the entropy can be used to represent information loss. The feature can be described to fit a probability distribution when shown in a histogram. If the distribution is Gaussian-like, symmetric or asymmetric, the parameters of this distribution can be estimated. Moreover, the estimated parameter of the gamma distribution can be employed, which have been shown to be useful descriptors in mesh quality estimation [26].

We expand this set to include the estimated parameters of a beta distribution. The beta distribution is a generalization of the uniform distribution, and is bounded to a $[0, 1]$ support. It can be interpreted as a ratio of gamma distributions. As it is bounded, the feature vector needs to be normalized in order

to estimate its parameters. The c and d parameters of the beta distributions are estimated with $\mu = \text{mean}(f), \sigma^2 = \text{var}(f)$

$$c = \mu^2 \left(\frac{1 - \mu}{\sigma^2} \right) - \mu, d = c \left(\frac{1 - \mu}{\mu} \right). \quad (6)$$

Moreover, we include some additional statistical measures which were successfully used in RR point cloud assessment [20], namely median Md , mode Mo , energy En , and sparsity Sp . Thus, our global descriptor vector F comprises 17 values: $\mu, Md, Mo, \sigma, H, En, Sp, \text{GGD}(\alpha, \beta^2), \text{AGGD}(\eta, \nu, \sigma_l^2, \sigma_r^2), \text{Gamma}(a, b)$ and $\text{Beta}(c, d)$.

IV. EXPERIMENT

The SJTU-PCQA database provides 378 public point cloud objects with a subjective quality score [7]. It consists of 9 objects, each distorted with 7 types of distortion on 6 levels. The 7 distortions are Octree-based compression (OT), Color Noise (CN), Downscaling (DS), Downscaling and Color noise (D+C), Downscaling and Geometry Gaussian noise (D+G), Geometry Gaussian noise (GGN), and Color and Geometry Gaussian noise (C+G).

For each point cloud, the local features are calculated per point. After extensive testing, we found that a neighborhood size of 20 gave the best performance; hence, unless specified, we report results for $k = 20$. For each feature vector, distribution statistics are calculated or estimated to describe its behavior in the whole point cloud. For each feature under inspection, we end up with 17 global descriptor values, as described in the previous section.

To find the optimal version of a global description for each feature, a support vector machine regressor is modeled to fit the feature vector data to the desired quality value. We follow the framework defined in [5]; to split the data in training and testing sets, of the 9 point cloud objects, 7 in the training set and 2 in the test set. All combinations of this split are scored and averaged to a final performance score of the model.

We analyse the performance of each individual feature, as well as the combination of all features, according to ITU-T recommendation P.1401 [27]. In particular, we compute the Spearman Rank Correlation Coefficient (SRCC), Pearson Linear Correlation Coefficient (PLCC) and Root Mean Square Error (RMSE) to account for monotonicity, linearity and accuracy, respectively. We test our metric on the SJTU dataset, and additionally report the performance for the M-PCCD [28] and the ICIP2020 [29] datasets.

V. FEATURE AND DESCRIPTOR PERFORMANCE

A. PCA-based descriptors

Figure 1 depicts the SRCC obtained for each PCA-based feature under consideration and for each global descriptor value. We can see that the correlation value of each feature varies depending on the global descriptor under use: in particular, the GGD α and AGGD η appear to have a consistently low correlation for all the features, whereas the mode Mo has correlation results that are consistently in the 0.2 – 0.3 range. In general, considering the features singularly leads to poor

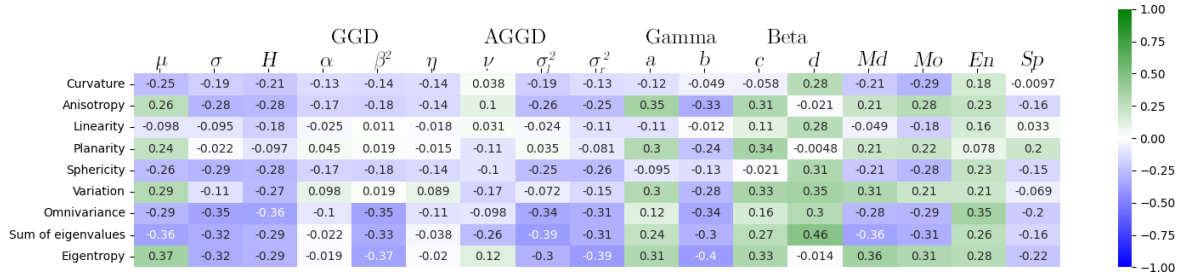


Fig. 1: SRCC values for each PCA-based feature and each global descriptor.

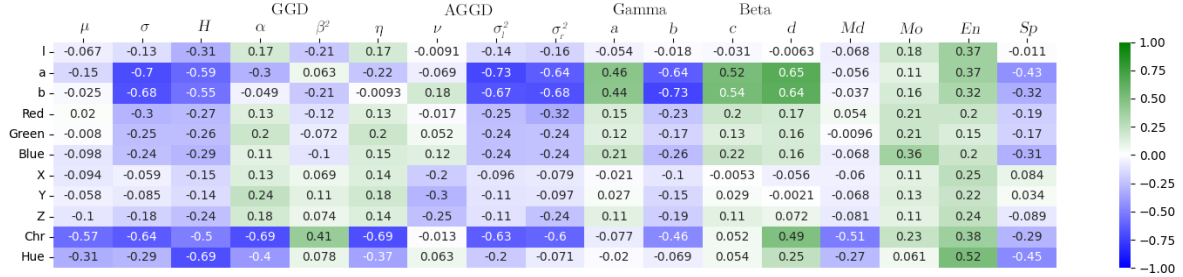


Fig. 2: SRCC values for each color-based feature and each global descriptor.

correlation, as the maximum absolute value we observe is 0.4. As the features all capture different aspects of the data, it is to be expected that no single feature would perform well. When we combine all features and descriptors using a linear SVR, we obtain a SRCC value of 0.84, demonstrating the predictive power of the features and descriptors.

Nonetheless, to avoid overfitting with a large number of features, we select a subset by considering feature/descriptor combinations that have $|SRCC| > 0.35$ or $|PLCC| > 0.3$ (that is, absolute correlation is greater than a threshold). To further analyse the predictive power of each feature and descriptor, we compute the SRCC between the feature/descriptors, to appropriately see whether they capture the same distortions or are complementary. Due to space limitations, we cannot show the full results of our analysis; the plots will be made available along with the code. We observe that certain features and descriptors exhibit high correlations, thus providing the same information to the model. In particular, Curvature $\{\mu, Md\}$, Anisotropy $\{\mu, Md\}$, and Sphericity $\{\mu, Md\}$ have an SRCC above 0.99; similarly for Eigensum $\{\mu, \sigma_r^2, Md\}$ and Eigentropy $\{\mu, Md, \sigma_r^2, b\}$. Thus, from each of the two sets, we only keep the feature/descriptor pair that exhibit the highest correlation with MOS scores (respectively, Anisotropy $\{\mu\}$ and Eigentropy $\{b\}$). The total set of features that we consider for the final model comprises 14 values: Curvature $\{a, Mo\}$, Planarity $\{\mu, Md, Mo, c\}$, Eigentropy $\{H, b, c\}$, Omnivariance $\{H\}$, Eigensum $\{c, d\}$, Anisotropy $\{\mu\}$, Sphericity $\{d\}$.

B. Color features

Figure 2 reports the SRCC values obtained for every color attribute and every global descriptor under consideration. Please note that to compute the correlation values, we only consider contents in the dataset which have color distortions. We can observe that, among the color attributes, chroma

attributes perform better than luminance attributes: the result is surprising, considering that many metrics put an emphasis on the luminance value as the carrier of the most important information for the human visual system. The results might be due to the fact that our descriptors are not really able to capture the underlying distribution of the luminance data, thus leading to subpar performance. Among all attributes, the a^* and b^* are the most promising, followed by the Hue and Chroma values. Considering all attributes together, we achieve a SRCC = 0.53 when considering only color distortions in the dataset, and SRCC = 0.26 in the full dataset.

Similar to what we did for the PCA-based descriptors, we only keep the feature/descriptor pairs whose $|SRCC| > 0.5$, and we compute the correlation between each feature/descriptor pair to further reduce our pool by only maintaining feature/descriptors which are not highly correlated with each other. Several sets exhibit high SRCC values, as can be expected considering that color spaces are a linear transformation of each other. We end up selecting the ones that have the best correlation with MOS scores, resulting in 12 values: $a^* \{\sigma, H, \sigma_r^2\}$, $b^* \{H, \sigma_r^2, b, d\}$, Chroma $\{\sigma, H, \nu, Md\}$, and Hue $\{H\}$.

C. Potential energy features

Figure 3 reports the results of using potential-energy-based features with each global descriptor. We can observe that, similarly to what we saw for other features, some descriptors seem to have a better correlation with subjective scores with respect to others; however, none of the feature/descriptor pairs achieves a very high correlation. The best performance is found for the d parameter of the beta distribution, which achieves an SRCC of 0.48 for the CV-PED feature.

To further reduce the number of feature/descriptor pairs to be used in our final model, the correlation analysis we performed thus far is inconclusive, as there is no clear trend

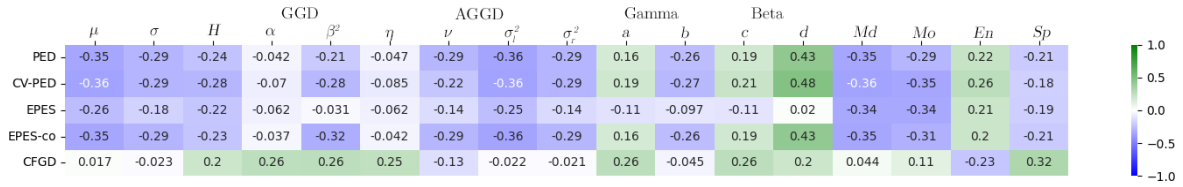


Fig. 3: SRCC values for each potential-based feature and each global descriptor.

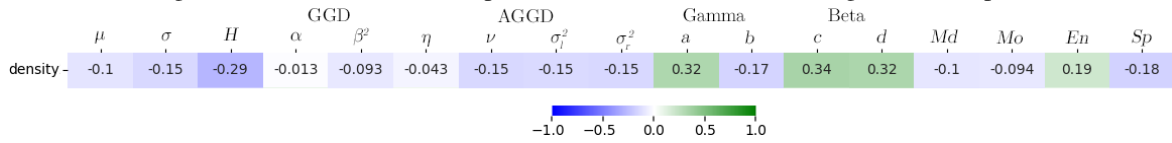


Fig. 4: SRCC values for each density-based feature and each global descriptor.

that would allow us to remove feature/descriptor pairs. So, we use recursive feature elimination to progressively remove features that would not contribute to the performance. We thus end up with the following set of 11 values: CV-PED $\{\mu, H, d\}$, EPES $\{Md\}$, EPES-CO $\{\mu\}$, CFGD $\{H, \nu, a, c, d, Sp\}$.

D. Density

Figure 4 shows the SRCC values for the density feature, for all descriptors under exam. We can observe that some distribution parameters better correlate with the quality than others; similarly to what we saw with PCA-based descriptors, we can also see that none of the descriptors taken singularly has high correlation values. When considering all the descriptors in a linear SVR, we obtain a SRCC = 0.47, only marginally better than single descriptor performance. Since we only have a set of 17 features, we do recursive feature elimination to only maintain a subset of descriptors that gives the best performance. From the elimination, we are left with 4 features: Density $\{b, H, d, \eta\}$.

VI. OVERALL PERFORMANCE

Table II reports the performance of notable metrics in the literature. In particular, we report results for some FR and NR metrics, and we perform a comparison with our full model considering all features and global descriptors, the subset we selected following Section V, compared with the best-performing descriptor in [5]. It can be observed that both our full model and the subset we selected outperform the baseline in terms of PLCC, SRCC and RMSE for all datasets under test. We can also see that the full set generally performs better with respect to the selected subset in terms of PLCC and SRCC; however, in terms of RMSE, the performance is worse for two out of three datasets, indicating that the full set of features might lead to more error in estimating the true quality, probably due to overfitting.

With respect to FR metrics, we can see that the proposed features reach quite a good performance. In particular, in the SJTU-PCQA dataset, the NR features are able to achieve similar performance with respect to FR metrics such as PointSSIM and PCQM, and outperform the RR metric; in fact, they are only outperformed by PointPCA. For the M-PCCD dataset, similarly, good performance is observed, comparable to most of the FR and RR metrics. In the case of the ICIP2020 dataset,

we also observe a good performance, that however does not reach the same heights as the FR and RR metrics under consideration.

For the sake of completeness, we also report the SRCC and PLCC values achieved by PQA-net [23] and GQI-VGG [22], in both variants. Please note that we did not run the metrics ourselves, instead opting to report the values as they are stated in the relative papers. We can see that with respect to PQA-Net, we achieve a slightly better performance in the SJTU dataset (Proposed: PLCC = 0.879, SRCC = 0.866, PQA-Net: PLCC = 0.85, SRCC = 0.82), while in the M-PCCD dataset, we are considerably better (Proposed: PLCC = 0.849, SRCC = 0.861, PQA-Net: PLCC = 0.60, SRCC = 0.65). These results indicate the potential of point-based features in predicting the quality of point cloud contents, especially with respect to projection-based approaches. However, we can also observe that deep learning approaches still outperform our combination of features. This might be due to a combination of the features that are extracted from the point cloud, and the machine learning method that is used to form a prediction score. In future work, we aim at considering several machine learning algorithms and different feature combinations to see whether better performance can be achieved.

VII. CONCLUSION

In this paper, we analyse how features commonly employed in full-reference quality assessment of point cloud contents can be used for no-reference quality assessment. We select a wide range of features, and we employ 17 global descriptors to model their distribution, using an SVR to combine the features into a single quality score. Results show that full reference features can achieve quite a good performance in predicting the quality of point clouds even in the absence of a reference, encouraging knowledge transfer and informing future work in the field. Future work will focus on testing other features and other machine learning methods to combine them in a single quality score.

REFERENCES

- [1] S. Gunkel, H. Stokking, M. Prins, O. Niamut, E. Siahann, and P. Cesar, "Experiencing virtual reality together: Social vr use case study," in *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, 2018, pp. 233–238.

TABLE II: Performance evaluation of state-of-the-art quality metrics.

Metric	SJTU-PCQA			ICIP2020			M-PCCD		
	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE
NR-3DQA [5]	0.781	0.760	1.644	0.741	0.704	0.825	0.533	0.553	1.220
Proposed (all)	0.879	0.866	1.350	0.883	0.884	0.667	0.849	0.861	0.871
Proposed (subset)	0.864	0.843	1.279	0.799	0.810	0.823	0.728	0.749	1.016
Point-to-point_MSE [10]	0.237	0.463	2.361	0.946	0.934	0.368	0.845	0.868	0.728
Plane-to-plane [13]	0.796	0.736	1.470	0.925	0.912	0.432	0.624	0.477	1.066
PointSSIM(curv) [15]	0.828	0.762	1.383	0.937	0.939	0.398	0.829	0.844	0.763
PCQM [14]	0.864	0.854	1.225	0.967	0.970	0.291	0.899	0.916	0.597
PCM_RR [20]	0.590	0.553	1.974	0.895	0.893	0.510	0.864	0.886	0.686
PointPCA(v1) [18]	0.935	0.946	0.483	0.931	0.920	0.417	0.880	0.853	1.154
<i>PQA-Ner*</i> [23]	0.85	0.82	-	-	-	-	0.60	0.65	-
<i>GQI-VGG16*</i> [22]	0.923	0.907	-	0.956	0.966	-	-	-	-
<i>GQI-VGG19*</i> [22]	0.925	0.912	-	0.952	0.966	-	-	-	-

- [2] T. Ebrahimi, S. Foessel, F. Pereira, and P. Schelkens, "JPEG Pleno: Toward an efficient representation of visual reality," *IEEE Multimedia*, vol. 23, no. 4, pp. 14–20, 2016.
- [3] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li *et al.*, "Emerging MPEG Standards for Point Cloud Compression," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 133–148, March 2019.
- [4] ISO/IEC DIS 23090-5, "Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC)," International Organization for Standardization, Jun. 2021.
- [5] Z. Zhang, W. Sun, X. Min, T. Wang, W. Lu, and G. Zhai, "No-reference quality assessment for 3d colored point cloud and mesh models," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 11, pp. 7618–7631, 2022.
- [6] E. M. Torlig, E. Alexiou, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, "A novel methodology for quality assessment of voxelized point clouds," in *Applications of Digital Image Processing XLI*, vol. 10752. International Society for Optics and Photonics, 2018, p. 107520I.
- [7] Q. Yang, H. Chen, Z. Ma, Y. Xu, R. Tang, and J. Sun, "Predicting the perceptual quality of point cloud: A 3d-to-2d projection-based exploration," *IEEE Transactions on Multimedia*, vol. 23, pp. 3877–3891, 2020.
- [8] I. Viola, S. Subramanyam, and P. Cesar, "A color-based objective quality metric for point cloud contents," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, ISSN: 2472-7814.
- [9] R. Diniz, P. G. Freitas, and M. C. Q. Farias, "Color and geometry texture descriptors for point-cloud quality assessment," vol. 28, pp. 1150–1154, conference Name: IEEE Signal Processing Letters.
- [10] 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.
- [11] 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)*, pp. 3460–3464, ISSN: 2381-8549.
- [12] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Mahalanobis based point to distribution metric for point cloud geometry quality evaluation," vol. 27, pp. 1350–1354, conference Name: IEEE Signal Processing Letters.
- [13] E. Alexiou and T. Ebrahimi, "Point cloud quality assessment metric based on angular similarity," in *2018 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, ISSN: 1945-788X.
- [14] G. Meynet, Y. Nehmé, J. Digne, and G. Lavoué, "PCQM: A full-reference quality metric for colored 3d point clouds," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, ISSN: 2472-7814.
- [15] E. Alexiou and T. Ebrahimi, "Towards a point cloud structural similarity metric," in *2020 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, pp. 1–6.
- [16] Q. Yang, S. Chen, Y. Xu, J. Sun, M. S. Asif, and Z. Ma, "Point cloud distortion quantification based on potential energy for human and machine perception." [Online]. Available: <http://arxiv.org/abs/2103.02850>
- [17] Y. Xu, Q. Yang, L. Yang, and J.-N. Hwang, "Epes: Point cloud quality modeling using elastic potential energy similarity," *IEEE Transactions on Broadcasting*, 2021.
- [18] E. Alexiou, I. Viola, and P. Cesar, "Pointpca: Point cloud objective quality assessment using pca-based descriptors," *arXiv preprint arXiv:2111.12663*, 2021.
- [19] Q. Yang, Z. Ma, Y. Xu, Z. Li, and J. Sun, "Inferring point cloud quality via graph similarity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [20] I. Viola and P. Cesar, "A reduced reference metric for visual quality evaluation of point cloud contents," vol. 27, pp. 1660–1664, conference Name: IEEE Signal Processing Letters.
- [21] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, and H. Yang, "Reduced reference perceptual quality model and application to rate control for 3d point cloud compression," vol. 30, pp. 6623–6636. [Online]. Available: <http://arxiv.org/abs/2011.12688>
- [22] A. Chetouani, M. Quach, G. Valenzise, and F. Dufaux, "Deep learning-based quality assessment of 3d point clouds without reference," in *2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2021, pp. 1–6.
- [23] Q. Liu, H. Yuan, H. Su, H. Liu, Y. Wang, H. Yang, and J. Hou, "PQA-net: Deep no reference point cloud quality assessment via multi-view projection," pp. 1–1, conference Name: IEEE Transactions on Circuits and Systems for Video Technology.
- [24] Y. Liu, Q. Yang, Y. Xu, and L. Yang, "Point cloud quality assessment: Dataset construction and learning-based no-reference approach." [Online]. Available: <http://arxiv.org/abs/2012.11895>
- [25] M. Tliba, A. Chetouani, G. Valenzise, and F. Dufaux, "Representation learning optimization for 3d point cloud quality assessment without reference," in *2022 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2022, pp. 3702–3706.
- [26] I. Abouelaziz, M. El Hassouni, and H. Cherifi, "No-reference 3d mesh quality assessment based on dihedral angles model and support vector regression," in *Image and Signal Processing*, ser. Lecture Notes in Computer Science, A. Mansouri, F. Nouboud, A. Chalifour, D. Mammas, J. Meunier, and A. Elmoataz, Eds. Springer International Publishing, pp. 369–377.
- [27] ITU-T P.1401, "Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models," International Telecommunication Union, July 2012.
- [28] 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," vol. 8, p. e27.
- [29] S. Perry, H. P. Cong, L. A. da Silva Cruz, J. Prazeres, M. Pereira, A. Pinheiro, E. Dunic, E. Alexiou, and T. Ebrahimi, "Quality evaluation of static point clouds encoded using mpeg codecs," in *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2020, pp. 3428–3432.