10.2352/EI.2022.34.9.IQSP-385

LEARNING-BASED 3D POINT CLOUD QUALITY ASSESSMENT USING A SUPPORT VECTOR REGRESSOR

Aladine Chetouani[1] *Maurice Quach*[2] *Giuseppe Valenzise*[2] *Frédéric Dufaux*[2]

[1]Laboratoire PRISME, Université d'Orléans, Orléans, France
[2]L2S, Centrale Supélec, Université Paris-Saclay, Gif-sur-Yvette, France

ABSTRACT

Recent advances in capture technologies have increased the production of 3D content in the form of Point Clouds (PCs). The perceived quality of such data can be impacted by typical processing including acquisition, compression, transmission, visualization, etc. In this paper, we propose a learning-based method that efficiently predicts the quality of distorted PCs through a set of features extracted from the reference PC and its degraded version. The quality index is obtained here by combining the considered features using a Support Vector Regression (SVR) model. The performance contribution of each considered feature and their combination are compared. We then discuss the experimental results obtained in the context of state-of-the-art methods using 2 publicly available datasets. We also evaluate the ability of our method to predict unknown PCs through a cross-dataset evaluation. The results show the relevance of introducing a learning step to merge features for the quality assessment of such data.

***Index Terms*—** 3D Point Cloud, Quality Assessment, Feature Fusion, Support Vector Regression

1. INTRODUCTION

Recent advances in capture technologies have increased the production of 3D content in the form of Point Clouds (PCs). As most multimedia contents, PCs may undergo different types of distortion introduced by several basic processing (acquisition, compression [1, 2], transmission, visualization, etc.), usually applied to transmit or visualize such data. To estimate the perceptual impact of these distortions on the perceived quality, subjective and objective evaluations are usually conducted. Subjective evaluation gives scores that reflect the perception of human observers through psycho-visual tests, while objective evaluation aims to automatically predict the subjective scores. As for 2D images and videos [3, 4], objective methods can be classified according to the availability of the reference PC: Full Reference (FR) approaches that need the reference PC, Reduced Reference (RR) approaches that exploit only partial information from the reference PC and No Reference (NR) approaches that predict the quality from only the distorted version of the reference PC. It is worth noting that such quality prediction methods can be useful in several applications [5, 6, 7].

Different point cloud objective metrics have been proposed in the literature, e.g., the point-to-point or the point-to-plane metrics [8]. The latter consists in projecting the point-to-point error vector along the local normal. In [9], the authors proposed a method based on the angular similarity between tangent planes. In [10], the authors proposed a metric called PC-MSDM by extending the well-known SSIM metric [11], widely used for 2D images, to PC, by considering features including local mean curvature. The authors extended that work later in [12] with a metric called PCQM, which takes into consideration also the color information. In [13], the authors proposed a new approach that focuses more on the distribution of the data. They introduced a new type of correspondence from point to distribution characterized by its mean and covariance using the well-known Mahalanobis distance. In [14], the authors proposed a color-focused metric that integrates geometry information. In [15], the authors adapted also the SSIM metric for point clouds using a number of features. In [16], the authors improved the point cloud PSNR metrics. Interesting learning-based methods were also proposed for 3D meshes [17, 18, 19, 20].

Contrary to the above-described metrics, in this paper we introduce a *learning-based* full-reference method to predict the quality of PCs. The proposed method is based on the extraction and the combination of a set of features. More precisely, we consider geometrical and color attributes extracted from the pristine PC and its degraded version. We then compute the distance between the considered features and use a Support Vector Regressor (SVR) to predict the quality of such distorted PC. Our method is compared with state-of-the-art metrics and the performance contribution of each considered feature is evaluated. We also evaluated quality prediction performance on unknown PC through a cross-dataset evaluation. Our results show the relevance of introducing a learning step to merge features for PC quality assessment.

Funded by the TIC-ART project, Region Centre-Val de Loire, France

The remainder of this paper is structured as follows: the proposed method is described in Section 2. Experiment results including feature analysis are discussed in Section 3, followed by the conclusion in Section 4.

2. PROPOSED METHOD

In this paper, we design a learning-based framework to predict the quality of PCs with reference. As illustrated by Fig. 1, our method is based on two main steps: feature extraction and feature fusion. The former aims to characterize the difference between the reference PC $\boldsymbol{X}$ and its degraded version $\boldsymbol{Y}$ through the comparison of some attributes, while the latter aims to derive a quality index from the features thus extracted using an SVR model.

2.1. Feature Extraction

Several features can be considered to characterize the difference between two PCs. In this study, geometrical and color attributes are considered since both have an impact on our perception. Geometry-based attributes aim to capture the structural deformations and the color-based attributes aim to capture the color deformations. More precisely, the mean curvature and roughness have been employed as geometrical attributes. The latter have been computed using CloudCompare software [21] with an automatic radius adjustment based on the density of each PC. These features are employed as follows.

Let us first define $\boldsymbol{F_k}$ as the $\boldsymbol{k^{th}}$ considered feature with **{$\boldsymbol{k=1}$: Color; $\boldsymbol{k=2}$: Mean curvature; $\boldsymbol{k=3}$: Roughness}**. After decomposing each PC into patches of size $32 \times 32 \times 1$, we compute for each considered feature $\boldsymbol{F_k}$ the distance $\boldsymbol{D_{F_k}(P_X \rightarrow P_Y)}$ between each patch $\boldsymbol{P_X}$ of $\boldsymbol{X}$ and its corresponding patch $\boldsymbol{P_Y}$ in $\boldsymbol{Y}$. Several distance criteria can be used. Here, the Mean Absolute Error (MAE) has been employed. We then form a feature vector $\boldsymbol{V_{F_k}}$ as follows:

$$\boldsymbol{V_{F_k}(X,Y) = \{\mu_{D_{F_k}(P_X \rightarrow P_Y)}; \sigma_{D_{F_k}(P_X \rightarrow P_Y)}\}}, \quad (1)$$

where $\boldsymbol{\mu_{D_{F_k}(P_X \rightarrow P_Y)}}$ and $\boldsymbol{\sigma_{D_{F_k}(P_X \rightarrow P_Y)}}$ represent the mean and the standard deviation of the patch-based distances $\boldsymbol{D_{F_k}(P_X \rightarrow P_Y)}$ of each feature $\boldsymbol{F_k}$, respectively.

In addition to the above feature vectors, the mean and standard deviation of the geometric distances between each point of $\boldsymbol{X}$ and its corresponding point in $\boldsymbol{Y}$ are also considered. The latter are also regrouped in a feature vector noted $\boldsymbol{V_G(X,Y)}$ (see Eq. 1).

2.2. Quality index

An SVR model is then used to predict the quality of the distorted PC. More precisely, we concatenate the above feature vectors to form a global feature vector $\boldsymbol{V(X,Y)}$ as follows:

$$\boldsymbol{V(X,Y)} = \begin{bmatrix} \boldsymbol{V_{F_1}(X,Y)} \\ \boldsymbol{V_{F_2}(X,Y)} \\ \boldsymbol{V_{F_3}(X,Y)} \\ \boldsymbol{V_G(X,Y)} \end{bmatrix}, \quad (2)$$

The input of our model is the global feature vector $\boldsymbol{V(X,Y)}$, while its output is the predicted quality score. We tested different kernel functions and the best result was achieved with the Gaussian kernel.

It is worth noting that existing PC quality metrics usually employ a function f that aims to symmetrize those metrics (i.e. $\boldsymbol{f(X,Y) = f(Y,X)}$). For a given metric M, the symmetrization function is often applied as follows:

$$\mathbf{Q_M(X,Y) = f(X \rightarrow Y, Y \rightarrow X)}, \quad (3)$$

where $\boldsymbol{Q_M(X,Y)}$ is the quality index computed between the PC $\boldsymbol{X}$ and its distorted version $\boldsymbol{Y}$ through the metric M.

For instance, the symmetrization function $\boldsymbol{f}$ is usually $\mathbf{min}$ and $\mathbf{max}$ for MSE and PSNR, respectively. In this study, instead of applying the symmetrization function that is not necessary the optimal solution, we rather consider three different configurations:

1. $\mathbf{V(X,Y)}$: The global feature vector is obtained by comparing points and patches of X and their corresponding points and patches in Y.
2. $\mathbf{V(Y,X)}$: The global feature vector is obtained by comparing points and patches of Y and their corresponding points and patches in X.
3. $\mathbf{V(X,Y)}$ and $\mathbf{V(Y,X)}$: The features are obtained by concatenating the global feature vectors of the two above-configurations.

The quality prediction capacity of our method is evaluated through each of these configurations in Section 3.2.

3. EXPERIMENTAL RESULTS

The efficiency of our method to predict the quality of PCs is evaluated in this section. To do so, we first describe the datasets used and the evaluation protocol applied. Then, we discuss the performance reached by each considered features as well as their combination. The results thus obtained are compared to state-of-the-art 3D PC metrics. Finally, a cross-dataset evaluation is carried-out to show the generalization ability of our method to predict the quality of unknown PCs.

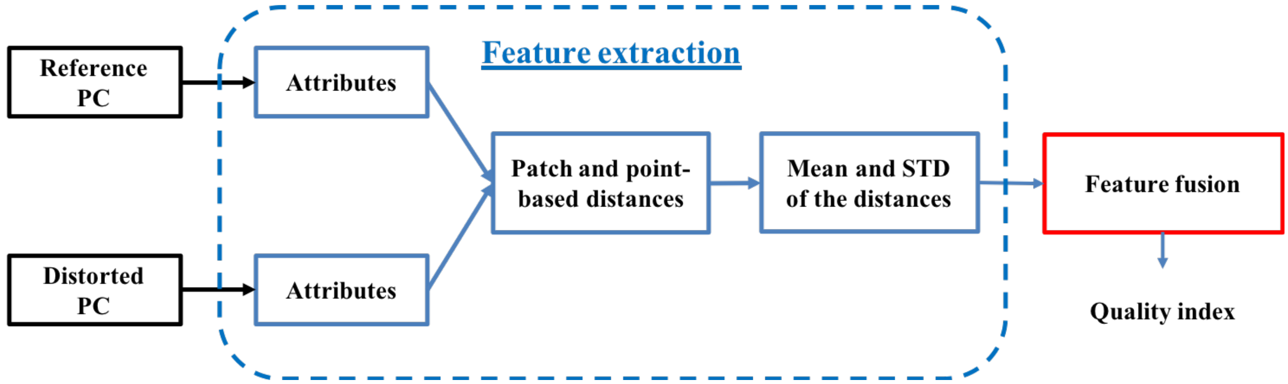

Fig. 1: Flowchart of the proposed 3D point cloud quality measure.

3.1. Datasets and Evaluation Protocol

In this study, two recent PC datasets are used: ICIP20 [22] and PointXR [23].

- **ICIP20** is composed of 6 commonly used point clouds from which 90 degraded versions were derived through 3 types of compression: V-PCC, G-PCC with triangle soup coding and G-PCC with octree coding. Each reference point cloud was compressed using five different levels.
- **PointXR** is composed of 5 point clouds from which 45 degraded versions were derived through G-PCC with octree coding for geometry compression and, Lifting and RAHT for color compression.

The above-described datasets are split into training and test sets N times (i.e. N folds). N corresponds to the number of reference point clouds of each dataset and it is thus equal to 6 for ICIP20 and 5 for PointXR. At each time, $N-1$ reference PCs and their degraded versions are used to train the model and, the rest (i.e. one reference PC and its degraded versions) are used to test the model.

Two commonly used measures are adopted as evaluation criteria: 1) Pearson Correlation Coefficient (PCC) and 2) Spearman Rank-Order Coefficient Correlation (SROCC). Both take values on $[0, 1]$ (absolute values), with 1 meaning the best correlation. These correlations are computed over each fold and the mean correlations are reported.

3.2. Feature Evaluation

Table 1 shows the mean correlations obtained for each of the three considered configurations on ICIP20 dataset. The best mean correlations are highlighted in bold. As can be seen, the first two configurations obtain close correlations with a slight difference in terms of mean SROCC. The third configuration that corresponds to the concatenation of the two global feature vectors, achieves the best performance with a mean PCC and SROCC equal to 0.973 and 0.969, respectively. The concatenation of all the feature vectors improves the global performance and, it is thus retained and compared with state-of-the art methods in Section 3.3.

Features	**PCC**	**SROCC**
1) $\mathbf{V}(\mathbf{X}, \mathbf{Y})$	0.967	0.958
2) $\mathbf{V}(\mathbf{Y}, \mathbf{X})$	0.967	0.965
3) $\mathbf{V}(\mathbf{X}, \mathbf{Y})$ **and** $\mathbf{V}(\mathbf{Y}, \mathbf{X})$	**0.973**	**0.969**

Table 1: Performance comparison in terms of mean correlations on ICIP20 dataset. The top mean PCC and SROCC values are highlighted in bold.

We then compare the performance contribution of each attribute and their combination only for the retained configuration. Table 2 shows the results obtained on ICIP20 dataset. The top two mean PCC and SROCC values are highlighted in bold. As can be seen, all the features obtain high correlations. Geometry and Color achieve the higher correlations, while Roughness obtains the lower performance. This result was expected since the former are usually used to estimate the quality of such data as well as for 3D meshes [24, 19]. The overall best performance is reached after combining all the considered features with an improvement gain in terms of mean PCC between 1.8 to 12%.

3.3. Comparison with the state-of-the-art

Our method is here compared with state-of-the-art metrics. Tables 3 and 4 show the results obtained for ICIP20 and PointXR datasets, respectively.

On ICIP20 (see Table 3), the proposed method performs the best with a mean PCC gain varying between 2.96

Features	PCC	SROCC
Geometry	**0.956**	0.945
Color	0.945	**0.955**
Curvature	0.910	0.898
Roughness	0.863	0.828
Combination of all features	**0.973**	**0.969**

Table 2: Performance comparison in terms of mean correlations for each considered features and their combination on ICIP20 dataset. The top two mean PCC and SROCC values are highlighted in bold.

and 10.57%. po2pointMSE and po2planeMSE obtain also high correlations outperforming PSNRpo2pointMSE and PSNRpo2planeMSE.

Method	PCC	SROCC
po2pointMSE	0.945	0.950
po2planeMSE	0.945	0.959
PSNRpo2pointMSE	0.880	0.934
PSNRpo2planeMSE	0.916	0.953
Proposed method	**0.973**	**0.969**

Table 3: Comparison with state-of-the-art methods on ICIP20. Best result is highlighted in bold.

On PointXR (see Table 4), our method performs also the best, followed by PSNRpo2pointMSE and PSNRpo2planeMSE. Both po2pointMSE and po2planeMSE obtain lower correlations.

Method	PCC	SROCC
po2pointMSE	0.887	0.978
po2planeMSE	0.855	0.942
PSNRpo2pointMSE	0.983	0.978
PSNRpo2planeMSE	0.972	0.950
Proposed method	**0.986**	**0.983**

Table 4: Comparison with state-of-the-art methods on PointXR. Best result is highlighted in bold.

3.4. Cross Dataset Evaluation

In this section, we evaluate the generalization ability of our method to predict the quality of unknown PCs by using ICIP20 (PointXR) as training set and PointXR (ICIP20) as test set. Table 5 shows the results obtained. As can be seen, high correlations are obtained when ICIP20 is used as training set. Whereas the correlations are not high as obtained in Section 3.3 when PointXR is used as training set. These results can be explained by the fact that ICIP20 is composed of PCs compressed using G-PCC and V-PCC, while PointXR contains only PCs compressed through G-PCC. Hence, the use of PointXR as training is more challenging since it allows to evaluate the capacity of our method to predict the quality of unknown PCs with unknown distortions.

Training set -> Test set	PCC	SROCC
ICIP20 -> PointXR	0.964	0.972
PointXR -> ICIP20	0.892	0.930

Table 5: Cross database evaluation.

4. CONCLUSION

In this paper, we proposed a learning-based method that efficiently predicts the quality of distorted PCs with reference. After splitting the reference PC $\boldsymbol{X}$ and its degraded version $\boldsymbol{Y}$ into patches, the mean absolute value between a set of features (i.e. curvature, roughness and color) extracted from each patches of both PCs are computed. Geometry distance between points of the two PCs was also considered as feature. The resulted feature vectors are then fed as input to an SVR model to predict the quality. The performance contribution of each considered feature and their fusion were also analyzed through difference configurations. The best configuration was retained and compared with state-of-the-art metrics using 2 publicly available datasets. The proposed learning-based method obtained high correlations on both datasets and showed a good ability to predict unknown PCs.

5. REFERENCES

[1] M. Quach, G. Valenzise, and F. Dufaux, "Learning Convolutional Transforms for Lossy Point Cloud Geometry Compression," in *2019 IEEE Intl. Conf. on Image Process. (ICIP)*, Sept. 2019, pp. 4320–4324, ISSN: 1522-4880.

[2] M. Quach, G. Valenzise, and F. Dufaux, "Folding-Based Compression Of Point Cloud Attributes," in *2020 IEEE Intl. Conf. on Image Process. (ICIP)*, Oct. 2020, pp. 3309–3313, ISSN: 2381-8549.

[3] A. Chetouani and L. Li, "On the use of a scanpath predictor and convolutional neural network for blind image quality assessment," *Signal Processing: Image Communication*, vol. 89, pp. 115963, 2020.

[4] A. Chetouani, "Image quality assessment without reference by mixing deep learning-based features," in *2020 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2020, pp. 1–6.

[5] A. Chetouani, S. Treuillet, M. Exbrayat, and S. Jesset, "Classification of engraved pottery sherds mixing deep-learning features by compact bilinear pooling," *Pattern Recognition Letters*, vol. 131, pp. 1 – 7, 2020.

[6] M. Hamidi, A. Chetouani, M. El Haziti, M. El Hassouni, and H. Cherifi, "Blind Robust 3D Mesh Watermarking Based on Mesh Saliency and Wavelet Transform for Copyright Protection," *Information*, vol. 10, no. 2, pp. 67, Feb. 2019.

[7] E. Fourati, W. Elloumi, and A. Chetouani, "Face anti-spoofing with image quality assessment," in *2017 2nd International Conference on Bio-engineering for Smart Technologies (BioSMART)*. IEEE, 2017, pp. 1–4.

[8] D. Tian et al., "Geometric distortion metrics for point cloud compression," in *2017 IEEE Intl. Conf. on Image Process. (ICIP)*, Beijing, Sept. 2017, pp. 3460–3464, IEEE.

[9] E. Alexiou and T. Ebrahimi, "Point Cloud Quality Assessment Metric Based on Angular Similarity," in *2018 IEEE Intl. Conf. on Multimedia and Expo (ICME)*, July 2018, pp. 1–6, ISSN: 1945-788X.

[10] G. Meynet, J. Digne, and G. Lavoué, "PC-MSDM: A quality metric for 3D point clouds," in *2019 11th Intl. Conf. on Quality of Multimedia Experience (QoMEX)*, June 2019, pp. 1–3, ISSN: 2472-7814, 2372-7179.

[11] Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Tran. on Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[12] G. Meynet et al., "PCQM: A Full-Reference Quality Metric for Colored 3D Point Clouds," in *2020 12th Intl. Conf. on Quality of Multimedia Experience (QoMEX 2020)*, Athlone, Ireland, May 2020.

[13] A. Javaheri et al., "Mahalanobis Based Point to Distribution Metric for Point Cloud Geometry Quality Evaluation," *IEEE Signal Process. Lett.*, vol. 27, pp. 1350–1354, 2020.

[14] I. Viola, S. Subramanyam, and P. Cesar, "A Color-Based Objective Quality Metric for Point Cloud Contents," in *2020 12th Intl. Conf. on Quality of Multimedia Experience (QoMEX)*, May 2020, pp. 1–6, ISSN: 2472-7814.

[15] E. Alexiou and T. Ebrahimi, "Towards a Point Cloud Structural Similarity Metric," in *2020 IEEE Intl. Conf. on Multimedia Expo Workshops (ICMEW)*, July 2020, pp. 1–6.

[16] A. Javaheri et al., "Improving PSNR-based Quality Metrics Performance For Point Cloud Geometry," in *2020 IEEE Intl. Conf. on Image Process. (ICIP)*, Oct. 2020, pp. 3438–3442, ISSN: 2381-8549.

[17] I. Abouelaziz, A. Chetouani, M. E. Hassouni, L. Latecki, and H. Cherifi, "Convolutional neural network for blind mesh visual quality assessment using 3d visual saliency," in *2018 IEEE International Conference on Image Processing, ICIP 2018, Athens, Greece, October 7-10, 2018*, 2018, pp. 3533–3537.

[18] I. Abouelaziz, A. Chetouani, M. El Hassouni, L. Latecki, and H. Cherifi, "3D visual saliency and convolutional neural network for blind mesh quality assessment," *Neural Computing and Applications*, 2019.

[19] A. Chetouani, "Three-dimensional mesh quality metric with reference based on a support vector regression model," *Journal of Electronic Imaging*, vol. 27, no. 4, pp. 1 – 9, 2018.

[20] I. Abouelaziz, A. Chetouani, M. El Hassouni, L. J. Latecki, and H. Cherifi, "No-reference mesh visual quality assessment via ensemble of convolutional neural networks and compact multi-linear pooling," *Pattern Recognition*, vol. 100, pp. 107174, 2020.

[21] "Cloudcompare," in *(version 2.11.1 Anoia) [GPL software]. Retrieved from http://www.cloudcompare.org/*, Athlone, Ireland.

[22] S. Perry et al., "Quality Evaluation Of Static Point Clouds Encoded Using MPEG Codecs," in *2020 IEEE Intl. Conf. on Image Process. (ICIP)*, Oct. 2020, pp. 3428–3432, ISSN: 2381-8549.

[23] E. Alexiou, N. Yang, and T. Ebrahimi, "PointXR: A Toolbox for Visualization and Subjective Evaluation of Point Clouds in Virtual Reality," in *2020 12th Intl. Conf. on Quality of Multimedia Experience (QoMEX)*, May 2020, pp. 1–6, ISSN: 2472-7814.

[24] I. Abouelaziz, A. Chetouani, M. E. Hassouni, L. J. Latecki, and H. Cherifi, "Combination of handcrafted and deep learning-based features for 3d mesh quality assessment," in *2020 IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 171–175.