IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 21 July 2022, accepted 5 August 2022, date of publication 16 August 2022, date of current version 1 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3198995

RESEARCH ARTICLE

Joint Geometry and Color Projection-Based Point Cloud Quality Metric

ALIREZA JAVAHERI[1], (Member, IEEE), CATARINA BRITES[2,3], (Member, IEEE), FERNANDO PEREIRA[2,3], (Fellow, IEEE), AND JOÃO ASCENSO[2,3], (Senior Member, IEEE)

[1]Ingenuity Labs Research Institute, Queen's University, Kingston, ON K7L 3N6, Canada
[2]Instituto de Telecomunicações, 1049-001 Lisbon, Portugal
[3]Instituto Superior Técnico, University of Lisbon, 1049-001 Lisbon, Portugal

Corresponding author: João Ascenso (joao.ascenso@lx.it.pt)

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) through the Project entitled Progressive Point Cloud Representation under Grant PTDC/EEI-PRO/7237/2014 and Grant UID/EEA/50008/2019.

ABSTRACT Point cloud coding solutions have been recently standardized to address the needs of multiple application scenarios. The design and assessment of point cloud coding methods require reliable objective quality metrics to evaluate the level of degradation introduced by compression or any other type of processing. Several point cloud objective quality metrics have been recently proposed to reliably estimate human perceived quality, including the so-called projection-based metrics. In this context, this paper proposes a joint geometry and color projection-based point cloud objective quality metric which solves the critical weakness of this type of quality metrics, i.e., the misalignment between the reference and degraded projected images. Moreover, the proposed point cloud quality metric exploits the best performing 2D quality metrics in the literature to assess the quality of the projected images. The experimental results show that the proposed projection-based quality metric offers the best subjective-objective correlation performance in comparison with other metrics in the literature. The Pearson correlation gains regarding D1-PSNR and D2-PSNR metrics range between ∼5% to ∼70% on three different datasets when data with all coding degradations is considered.

INDEX TERMS Coding, degradation, point cloud, projection, quality assessment, recoloring.

I. INTRODUCTION

Recent advances in 3D acquisition and reconstruction technologies have enabled many visual immersive applications, such as virtual and augmented reality, immersive communications, and video gaming. Point cloud (PC) is an emerging 3D visual representation format that is becoming rather popular due to its easy acquisition and capability to realistically represent 3D objects and visual scenes. However, since realistic PCs require a large number of points, a compact representation of PCs is essential for storage and transmission applications and services. PC coding is a rather new and challenging problem due to the unstructured nature of PCs where each point, i.e., filled voxel in voxelized PCs, is associated with a 3D coordinate; moreover, each point has often associated attributes such as color, transparency, reflectance, etc. In recent years, several efforts in PC coding were able to significantly reduce the bitrate while still maintaining the PC quality and fidelity. MPEG has already developed two PC coding standards [1], [2], [3], notably Geometry-based Point Cloud Compression (G-PCC) for static and progressive acquired PCs, and Video-based Point Cloud Compression (V-PCC) for dynamic PCs. In this context, the assessment of PC quality is very important as it plays a significant role in the design and optimization of coding solutions as well as in the validation of the quality of experience offered to the users.

The best way to reliably measure PC quality is through subjective quality assessment where a specially designed framework collects opinion scores from a minimum number of subjects. In the literature, there are many subjective PC quality studies available, considering different ways to visualize the PC [4], [5], render the PC [6], [7], and also

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Shu.

different types of degradation [8], [9]. However, since subjective quality assessment is expensive and time-consuming, reliable objective quality metrics are critical to facilitate the design of more efficient PC coding solutions and assess the quality of experience offered to the users. In the literature, the performance of multiple PC objective quality metrics has been assessed through the correlation between corresponding objective and subjective quality scores, notably for different codecs, with different types of degradation [6], [10].

In the literature, a few works have already exploited the idea of measuring the PC quality by projecting the 3D PC into one or more 2D images, i.e., by converting a 3D PC into several 2D images, a more traditional type of data. In the context of PC coding, this type of approach was successfully exploited to achieve efficient compression, as demonstrated by the MPEG V-PCC standard [2], [3]. In the context of PC quality assessment, these 2D projected images can be obtained by performing multiple projections into different viewpoints, i.e., using different projection centers. After projection, the most recent, and powerful 2D image quality metrics available can be exploited without any changes, to assess the entire PC quality through the projected images. However, the projection-based metrics available in the literature are not yet showing better subjective-objective correlation performance than the popular point-to-point quality metrics, where correspondences are established in the 3D space and errors/distances in position or color are accounted for.

The most critical weakness of projection-based metrics is caused by the inability of 2D quality metrics to efficiently handle local displacement errors since pixel-level (or region-level) comparisons are usually made. Due to lossy PC coding, geometry distortions (or degradations) cause many displacements and thus a lower correlation performance. Typically, 2D objective quality metrics consider that these displaced pixels/regions have high distortion when, in fact, the small or medium geometry degradations are perceptually well tolerated, especially when color is also available as some degree of masking may happen [6]. For example, small displacement errors in the projected images due to geometry distortions may not be perceived by humans but may lead to high objective distortions when 2D quality metrics are used to assess the quality of the PC projected images. Another critical issue for some PC codecs is related to the difference between the number of points in the reference and decoded PCs. This often occurs when the PC coding solution uses planar or triangular approximations of the PC surface, and more points may be recreated at the decoder side when these surfaces are sampled or when the PC coding solution reduces the number of coded points using octree pruning. This implies that one of the projected 2D images, the reference or the degraded one, may have, for some positions, pixels occupied while these pixels are not filled in the other projected 2D image, thus leading to large pixel-based mismatches. In some past works [11], [12], these positions are either ignored or an occupied position is compared with a non-occupied position (usually filled with some background color). However, both solutions negatively impact the final quality metric correlation performance since, for some cases, these pixels are visually important and should not be ignored; moreover, the quality score should not depend on an arbitrarily selected background color.

In this context, this paper proposes a novel joint geometry and color projection-based PC quality metric, which addresses the weaknesses and issues above, thus achieving a higher objective-subjective correlation performance. The key original ideas underpinning this novel quality metric are twofold:

- Reference and degraded projected images are compared for two fixed geometry conditions, notably the reference and degraded geometries. Later, these two quality scores are fused, thus implying that the proposed approach implicitly considers both geometry and color distortions/degradations. By comparing images created for the same geometry condition, reference or degraded, the aforementioned undesired misalignments are avoided and there is no difference between the number of points on the reference and degraded PCs for the same geometry condition.
- A padding operation is performed in the 2D domain to avoid assigning an arbitrary, uniform color to the background (i.e., not projected) pixels. Because of the aligned geometries, a pixel (in denser PC) would never be compared with a background pixel (in sparser PC). However, if a uniform background value is assigned, the 2D quality metric would be biased due to these regions for which no distortions would occur. The proposed padding operation mitigates the impact of these background pixels.

To achieve its objectives, the rest of the paper is organized as follows. Section II briefly reviews the state-of-the-art on PC objective quality metrics. Section III describes the proposed joint geometry and color projection-based PC quality metric. Experimental results are presented and analyzed in Section IV and, finally, conclusions are offered in Section V.

II. BACKGROUND WORK ON PC OBJECTIVE QUALITY ASSESSMENT

In this section, the state-of-the-art on PC objective quality metrics is briefly reviewed, by addressing first, point-based metrics, followed by feature-based metrics and, finally, projection-based metrics.

A. POINT-BASED PC QUALITY METRICS

A point-based quality metric compares the geometry or attributes of the reference and degraded PCs directly, point-by-point, after defining the necessary point correspondences. The most popular point-based geometry quality metrics are the Point-to-Point (Po2Po) [13] and Point-to-Plane (Po2Pl) [14] metrics. In a Po2Po metric, for every point in a degraded/reference PC, the nearest neighbor is obtained in the corresponding reference/degraded PC (thus a point correspondence is obtained); after, the Hausdorff distance or

the Mean Squared Error (MSE) distance are computed over all pairs of points. The main disadvantage of these types of metrics is that they do not consider that PC points represent the surface of an object(s) in the visual scene. To solve this issue, Point-to-Plane (Po2Pl) metrics have been proposed by Tian *et al.* [14], which model the underlying surface at each point as a plane perpendicular to the normal vector at that point. This type of metrics results in smaller errors for the points closer to the PC surface, associated with a plane. Currently, the MPEG-adopted PC geometry quality metrics are the Po2Po MSE (D1) and Po2Pl MSE (D2) distances/errors and their associated PSNR [15]. Moreover, a Plane-to-Plane (Pl2Pl) metric has been proposed by Alexiou *et al.* [16], which measures the similarity between the underlying surfaces associated with the corresponding points in the reference and degraded PCs. In this case, tangent planes are estimated for both the reference and degraded points and the associated angular similarity is assessed.

In [17], Javaheri *et al.* propose a geometry quality metric based on the Generalized Hausdorff distance, which corresponds to the maximum distance for a specific percentage of data rather than the whole data, thus filtering some outlier points. The Generalized Hausdorff distance between two PCs adopted in this quality metric may be computed for both Po2Po and Po2Pl metrics. In [18], Javaheri *et al.* also propose a so-called Point-to-Distribution (Po2D) metric based on the Mahalanobis distance between a point in a PC and its K nearest neighbors in the other PC. The mean and covariance matrix of the corresponding distribution is computed and used to measure the Mahalanobis distance between points in one PC and their corresponding set of nearest neighbors in the other PC. These distances are averaged to compute the final quality score. Later, they propose a joint color and geometry point-to-distribution quality metric in [19] based on the scale-invariance property of Mahalanobis distance. In [20], Javaheri *et al.* propose resolution adaptive metrics which improve the state-of-the-art D1-PSNR and D2-PSNR metrics by using as normalization factors the PC rendering and intrinsic resolutions.

There are not many point-based quality metrics for PC attributes and specifically for color. However, the Po2Po PSNR for color in the YCbCr color space is widely used by MPEG and in the literature to evaluate PC color degradations. This metric works like the Po2Po geometry metrics, with the error corresponding now to the difference between the colors of the points in some PC correspondence. This metric may either be computed only for the luminance (Y-PSNR) or chroma components (Cb/Cr-PSNR) individually or as a weighted average of all color components (YUV-PSNR).

B. FEATURE-BASED PC QUALITY METRICS

A feature-based PC quality metric computes a quality score based on the difference between some local or/and global features extracted from the reference and degraded PCs. Meynet *et al.* propose in [21] the so-called Point Cloud Multi-Scale Distortion metric (PC-MSDM), a structural similarity-based PC geometry quality metric based on local curvature statistics. This quality metric computes the surface curvature associated with each point and establishes after point-based correspondences. The metric score corresponds to the Gaussian weighted curvature statistics for a set of local neighborhoods.

In [22], Viola *et al.* design a PC quality metric based on the histogram and correlogram of the luminance component. After, the proposed color quality metric is fused with the Po2Pl MSE geometry metric (D2) using a linear model with a weighting parameter determined using a grid search method.

In [23], Diniz *et al.* propose the so-called Geotex metric, which is based on Local Binary Pattern (LBP) descriptors adapted to PCs and applied to the luminance component. To apply it on PCs, the LBP descriptor is computed on a local neighborhood corresponding to the K-nearest neighbors of each point in the other PC. Histograms of the extracted feature maps are obtained for both the reference and degraded PCs and used to compute the final quality score using a distance metric such as the f-divergence [24]. In [25], Diniz *et al.* extend the Geotex metric by considering multiple distances, notably Po2Pl MSE for geometry and the distance between LBP statistics [23] for color. In [26], Diniz *et al.* also propose another quality metric, which computes Local Luminance Patterns (LLP) on the K nearest neighbors of each point on the other PC.

In [27], Meynet *et al.* propose the Point Cloud Quality Metric (PCQM) metric, which combines the geometry features used in [21] with five color features related to lightness, chroma, and hue. PCQM corresponds to the weighted average of the differences for the geometry and color features between the reference and degraded PCs. In [28], Viola *et al.* propose the first reduced reference quality metric that jointly evaluates geometry and color. A set of seven statistical features such as the mean and standard deviation are extracted from the reference and degraded PCs in the geometry, texture, and normal vector domain, in a total of 21 features. The reduced quality score is computed as the weighted average of the differences for all these features between the reference and degraded PCs.

Inspired by the SSIM quality metric for 2D images, Alexiou *et al.* propose in [29] a quality metric using local statistical dispersion features. These statistical features are extracted in a local neighborhood around each point in the reference and degraded PCs considering four independent 'attributes', notably geometry, color (luminance), normal, and curvature information. The final quality metric is obtained by pooling the feature value differences between associated points in the reference and degraded PCs. In [30], Yang *et al.* propose a quality metric based on graph similarity, which extracts keypoints by resampling the reference PC and creates local graphs centered at these keypoints for both reference and degraded PCs. Some local similarity features are computed based on the graph topology with the quality metric value corresponding to the similarity between these

features. In [31], Diniz *et al.* extract local descriptors that represent geometry-aware texture information of the PCs, namely the Local Color Pattern (LCP) and different variations of the LBP descriptor. Then, statistics of the descriptors are obtained and used to compute the objective quality score.

C. PROJECTION-BASED PC QUALITY METRICS

A projection-based PC quality metric maps the 3D reference and degraded PCs onto some selected 2D planes and computes the quality score by comparing the projected images using some 2D image quality metrics. The first projection-based PC quality metric has been proposed by Queiroz *et al.* in [11]. This metric starts by projecting the reference and degraded PCs onto the six faces of a bounding cube around the PC, concatenates the corresponding projected images, and measures the 2D PSNR between the corresponding degraded and reference concatenated projected images.

In [12], Torlig *et al.* propose a rendering software for PC visualization on 2D screens, which performs the orthographic projection of a PC onto the six faces of the PC bounding box. A 2D quality metric is then applied to the reference and degraded projected images resulting from the rendering and the final quality score is obtained by computing the average over the six pairs of projected images. In [33], Alexiou *et al.* study the impact of the number of projected 2D images (each corresponding to a specific view) on the subjective-objective correlation performance of projection-based quality metrics. It is shown that even a single view may be enough to achieve a reasonable correlation performance. Moreover, a projection-based PC quality metric weighting the projected images according to the user interactions performed during the subjective test is proposed. In [37], the quality metric proposed in [12] is benchmarked considering a different number of views, pooling functions, etc. The best performance is achieved when 2D quality metrics are applied to the projections from 42 different views and pooled with an l1-norm.

In [32], Liu *et al.* propose a no-reference deep learning-based quality metric, called Point cloud Quality Assessment Network (PQA-Net). It starts by projecting the PC into six different images which are then processed by a convolutional neural network for feature extraction. A distortion type identification network and quality vector prediction network process these features to obtain the final quality score.

In [34], Wu *et al.* propose two projection-based objective quality metrics: a weighted view projection-based metric and a patch projection-based metric. In both cases, 2D quality metrics are applied to texture and geometry maps. The patch projection-based metric significantly outperforms the weighted view projection-based metric. In [35], Liu *et al.* propose a PC quality metric based on attention mechanisms and the principle of information content weighted pooling. The proposed quality metric translates, rotates, scales, and orthogonally projects PCs to 12 different views and evaluates the quality of the projected images using the IW-SSIM [36] 2D metric.

III. PROPOSED PROJECTION-BASED PC QUALITY METRIC

In this section, the architecture and walkthrough of the proposed joint geometry and color projection-based PC quality metric are presented; after the most relevant modules are explained in detail.

A. ARCHITECTURE AND WALKTHROUGH

Figure 1 shows the proposed Joint Geometry and Color Projection-based PC Quality Metric architecture, referred to from now on as *JGC-ProjQM*. This metric will have multiple variants depending on the 2D quality metric used in the 2D Quality Assessment module.

The key idea behind this metric is that the degraded and reference PCs are processed in two parallel branches, one associated with the reference geometry and the other associated with the degraded geometry, to obtain two intermediate quality scores which are fused at the end. To avoid misalignment errors, before applying the 3D to 2D projection, the reference and degraded PCs are processed to obtain two PCs to be compared with the same geometry:

- **Reference Geometry Branch** - In the architecture top branch, the geometry of the reference PC is used; the reference PC geometry is recolored with the color of the degraded PC and the resulting PC is compared with the reference PC (naturally, including the original color).
- **Degraded Geometry Branch** - In the architecture bottom branch, the geometry of the degraded PC is used; the degraded PC geometry is recolored with the reference PC color and the resulting PC is compared with the degraded PC (naturally, including the degraded color).

Naturally, the color attributes assigned to the points in the two PCs with the same geometry in the two branches will be different, notably using the color data with and without coding degradations. Although the geometry (which defines the shape of an object) is not explicitly considered in the proposed quality metric, it is considered in an implicit way. First, the proposed two-branch architecture exploits the geometry of both reference and degraded PCs, since the projection of 3D points to 2D planes only depends on the PC geometry. Second, the recoloring process which uses the underlying reference geometry to estimate the reference color for (new) degraded points and the degraded geometry to estimate the degraded color for reference points.

Before applying the *JGC-ProjQM* metric, the PCs must be voxelized to some fixed precision, if they are available in floating-point precision. This step is important to perform the 3D (voxels) to 2D (pixels) projection. Nowadays, most available PC data are already in fixed precision, i.e., the PC has been voxelized, and, thus, this step may not be needed; for this reason, it is not included in the architecture in Figure 1. Moreover, both V-PCC and G-PCC standard codecs code fixed precision PCs or perform voxelization as

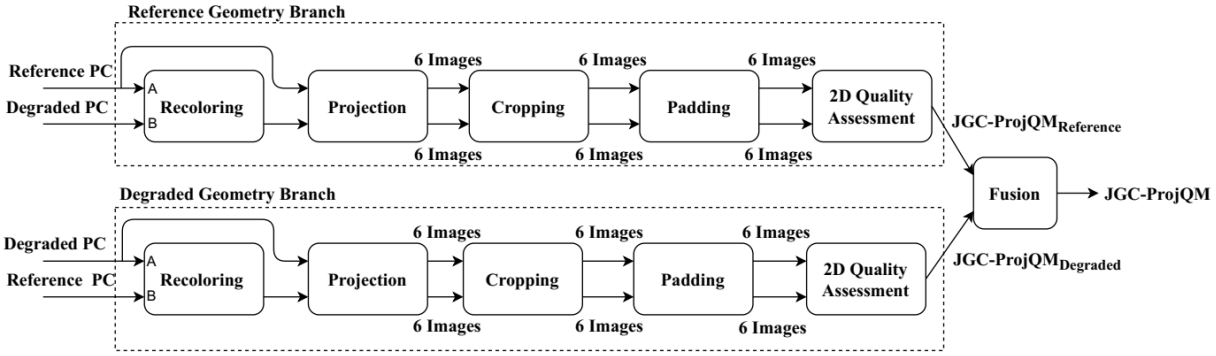

FIGURE 1. Proposed joint geometry and color projection-based PC quality metric (JGC-ProjQM) architecture.

a pre-processing step before coding. The main modules in the proposed *JGC-ProjQM* quality metric pipeline are briefly explained in the following walkthrough:

1. **Recoloring:** Due to the lossy coding of geometry data, the positions of the points in the reference and degraded PCs are not the same. Thus, when a projection is made, the resulting reference and degraded projected 2D images have regions that are not aligned, even when the degraded PC has only slight geometry degradations. This creates a problem for 2D quality metrics which are typically not robust to misalignments (or displacements) and often perform very poorly in this situation. Therefore, to align the reference and degraded projected images, a recoloring step is applied where the degraded (or reference) PC color is mapped onto the reference (or degraded) geometry. In this way, the color degradation is compared for the two geometry conditions (reference and degraded geometry, each at a time), without any misalignments (or a different number of points) while still considering both the geometry and color degradations. This is a key technical novelty of the proposed *JGC-ProjQM* quality metric. This solution also exploits the fact that color degradations typically have a higher perceptual impact than geometry degradations, notably due to masking effects [6]. More details about this module are presented in Section III.B.
2. **Projection:** The reference, degraded and the two recolored PCs obtained in the previous step are orthographically projected onto the six faces of a cube to obtain six projected images for each PC, this means six non-overlapping images for each of the four PCs, i.e., reference, degraded, reference recolored and degraded recolored. Although another type of projection could be used, the low-complexity orthographic projection is enough to assess the quality degradations, especially considering that the PC pairs to compare have now the same geometry. In this process, six binary occupancy maps are also obtained for each projected PC; these occupancy maps serve to signal if a 2D image pixel corresponds (or not) to a point (filled voxel) in the 3D PC. The size of the projected images and occupancy maps only depends on the precision p of the PC, thus commonly obtaining a $2^p \times 2^p$ size map. More details about this module are presented in Section III.C.
3. **Cropping:** After the projection, and depending on the PC size and position, the projected images may have a rather large background area (area without projected pixels around the PC object(s)), notably in comparison with the image area occupied with the PC points. These empty background areas can act as a distractor for the 2D quality metric, notably if the same (uniform) color value is assigned to all background pixels and thus they must be reduced as much as possible with a cropping procedure. More details about this module are presented in Section III.D.
4. **Padding:** If an arbitrary, uniform background value is used for background pixels, i.e., pixels positions on the projected images that have not been filled, the 2D quality metric may be biased, thus lowering the *JGC-ProjQM* quality prediction power. Computing the 2D quality metric only on the foreground pixels is also not a good solution. In the proposed method, after recoloring, the geometries are similar, and thus padding is not necessary to avoid comparing foreground and background pixels. However, some 2D quality metrics such as SSIM cannot be applied to arbitrarily shaped objects. In this context, the padding module targets the creation of a seamless image, where the background positions are filled with interpolated/padded values, thus obtaining an image that is more suitable for quality assessment. In this process, the background holes inside the PC foreground area are also padded in the same way as the empty areas around the projected PC. More details about this module are presented in Section III.E.
5. **2D Quality Assessment:** At this stage, a 2D image quality metric is computed between the six reference, padded images and the corresponding degraded, padded images for the same view/projection; this happens for the two architectural branches. The output of

this process is six objective quality scores, one for each pair of projected, padded images, corresponding to each projection plane, which must be pooled together. The final *JGC-ProjQM* metric performance has been studied for several 2D quality metrics, so-called metric variants, using common pooling functions, e.g., max, min, and weighted average. Since it was found that the final subjective-objective correlation performance is rather similar for the various pooling functions, it was decided to adopt average pooling to obtain a single quality score for each architectural branch of the proposed projection-based PC quality metric.

6. **Fusion:** All modules previously described are included in the two architectural branches of the proposed projection-based PC quality metric to obtain two intermediate quality scores, notably $JGC\text{-}ProjQM_{reference}$ and $JGC\text{-}ProjQM_{degraded}$. These two intermediate quality scores represent the quality associated with the projected images as measured by a 2D quality metric, for two different geometry conditions, i.e., reference and degraded geometries, and must be fused to obtain the final *JGC-ProjQM* quality metric. Although different fusion strategies are possible, even applying machine learning techniques, it was found that a linear regression was enough to obtain a high subjective-objective correlation performance, without the risk of overfitting. More details about this process are presented in Section III.F.

B. RECOLORING

The main challenge with projection-based PC quality metrics is that geometry distortions may cause misalignment errors between the reference and degraded projected 2D images. The 2D quality metrics do not typically handle well local displacement errors; as a consequence, when the same pixel location in images projected for reference and degraded geometry is compared, the measured error may not express well the user-perceived distortion since the color of different 3D positions in the reference and degraded PCs are used. Figure 2 shows the frontal projection for the *Egyptian Mask* PC before coding (reference geometry and color) and after G-PCC decoding and recoloring with the original/reference color, for the lowest geometry rate. Figure 2 also shows a residual image with the difference between the previous reference and decoded, recolored projected images (with some enhancement for better visualization). Although the color in both PCs is the same and the PCs are visually similar, the residual image shows large misalignment errors.

To overcome this key problem, this paper proposes an original solution involving computing the 2D quality metrics with different color data under two geometry conditions, notably reference and degraded geometries. The idea is to use the PC geometry, reference and degraded, and to perform recoloring to assign the degraded color to the reference geometry (top branch of the architecture) and the reference color to the degraded geometry (bottom branch of the architecture).

FIGURE 2. Egyptian Mask PC projected from front view: (a) Reference PC; (b) Decoded G-PCC geometry for lowest geometry rate, recolored with the original color; (c) Residual image between (a) and (b) after enhancement.

By using this approach, the projected images are always geometry-aligned within each branch.

To recolor PC A, with the color of PC B, each point in PC A will have a color assigned using the color of one or more corresponding points in PC B. In the proposed recoloring algorithm, the color for each point in PC A after recoloring is determined as follows:

1. For each point in PC A, the nearest neighbor in PC B is found (NN_A) and, for each point in PC B, the nearest neighbor in PC A is found (NN_B).
2. For each point a in PC A, perform: if point a is listed in the nearest neighbors of some points in PC B ($a \in NN_B$), its color is the average color of the points in PC B which have point a as their nearest neighbor, as defined in (1):

$$C_a = \sqrt{\frac{\sum_{b \in B, NN_B(b)=a} C_b^2}{\sum_{b \in B, NN_B(b)=a} 1}} \tag{1}$$

   Otherwise, its color is the color of its nearest neighbor listed in NN_A. In (1), C_a and C_b are the colors at points a and b. The denominator counts the number of points in PC B which have point a in PC A as its nearest neighbor.

Figure 3 illustrates the recoloring process for the *Amphoriskos* PC using a point-based rendering solution with cube primitives. Figure 3(a) shows the reference PC (reference geometry and color) on the left and the recolored PC (reference geometry and decoded color) on the right. Figure 3(b) shows the degraded PC (decoded geometry and color) on the left and the recolored PC (decoded geometry and reference color) on the right. Each pair of PCs have the same geometry and either the reference or degraded colors; because of the geometry alignment, it is now possible to compare these color values using a simple, direct pixel-to-pixel correspondence.

C. PROJECTION

Projection is the core module of the proposed *JGC-ProjQM* architecture, where 3D PCs are mapped onto six 2D planes from different perspectives, thus creating the projected images. The proposed projection procedure is based on the orthographic projection, an often-used parallel projection that renders objects with suitable shapes and sizes. In this

FIGURE 3. Amphoriskos PC decoded with G-PCC in Octree geometry coding mode and Lifting color coding mode at the lowest rate: (a) PCs obtained for the reference geometry with the reference (left) and decoded color, after recoloring (right); (b) PCs obtained for the decoded geometry with the decoded color (left) and reference color, after recoloring (right).

procedure, each PC point is projected to a pixel in a 2D image while considering its visibility (or occlusion) for each specific viewing perspective and, thus, each projection plane. The proposed projection procedure considers two main steps:

- **Mapping:** The planar (or 2D) images are generated by applying the orthographic projection for the six different sides (planes) of the precision box, i.e., the box surrounding the PC object with a size defined by the coordinates' precision. For each plane, a point is projected onto the plane as long as the point is not occluded by another point closer to the same plane. With this orthographic mapping to six faces of the surrounding cube, all visible points of the PC surface are covered (even for complex objects) and thus most coding artifacts are accounted for. It has been previously shown that there is no significant performance gain by using more than six projection planes [33].
- **Filtering:** Points that are projected onto a projection plane but do not belong to the PC surface closer to the plane must be removed. Considering a typical rendering process, these points will be occluded due to the use of primitives around each point or to surface reconstruction techniques. These points are unduly projected when there is some empty space between points in the surface closer to the plane and, therefore, points from the opposite side of the object are unduly projected onto this plane. Since these points are not visible after PC rendering is performed, a filtering technique is used to remove these points from the projected images.

1) MAPPING

The mapping algorithm projects every point visible from the perspective associated with each specific PC precision box plane as follows:

1. Six images, each corresponding to a projection plane are initialized with a uniform background color, e.g., white. In practice, the background color can be any color (in this case 255 is used) since the background pixels will be later padded and, thus, filled with non-uniform, interpolated values. Six planes are defined, notably $PL = \{xy, xz, yz, \acute{x}\acute{y}, \acute{x}\acute{z}, \acute{y}\acute{z}\}$, the first three with (0,0,0) origin and, the last three, including the opposite point, i.e. $(2^p, 2^p, 2^p)$, where p is the PC coordinates precision; these planes are shown in Figure 4. Initially, these images are set to 255 (as mentioned above) according to (2):

$$I_{PL} = \begin{bmatrix} 255 & \cdots & 255 \\ \vdots & \ddots & \vdots \\ 255 & \cdots & 255 \end{bmatrix}_{2^p \times 2^p \times 6} \tag{2}$$

where I_{PL} is the projected image associated with any of the planes (PL), e.g., I_{xy} or $I_{\acute{y}\acute{z}}$.

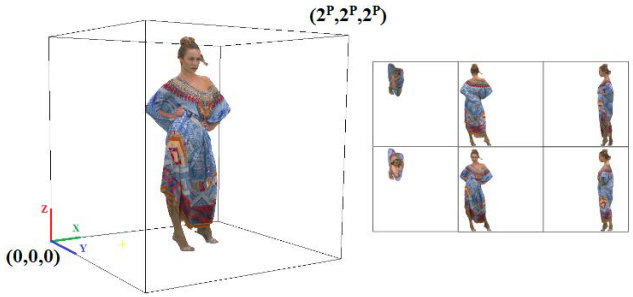

FIGURE 4. The six precision box planes surrounding a PC object: $\{xy, xz, yz, \acute{x}\acute{y}, \acute{x}\acute{z}, \acute{y}\acute{z}\}$ and their associated projected images.

2. Six binary images corresponding to the occupancy map (OM) for each projected image are initialized with '0', i.e., non-occupied pixel/voxel, as follows:

$$OM_{PL} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{2^p \times 2^p \times 6} \tag{3}$$

where OM_{PL} is the occupancy associated with any of the planes (PL), e.g., OM_{xy} or $OM_{\acute{y}\acute{z}}$.

3. To keep track of the occluded points, for the three coordinates, two depth maps are used to store the minimum projected depth ($NearD$) and maximum projected depth ($FarD$). These depth maps are initialized with 0 and 2^p, respectively, according to (4).

$$\begin{aligned} NearD &= \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{2^p \times 2^p} \\ FarD &= \begin{bmatrix} 2^p & \cdots & 2^p \\ \vdots & \ddots & \vdots \\ 2^p & \cdots & 2^p \end{bmatrix}_{2^p \times 2^p} \end{aligned} \tag{4}$$

While $NearD$ keeps record of the depth of the closest projected point to the $(\acute{x}\acute{y}, \acute{x}\acute{z}, \acute{y}\acute{z})$ planes, $FarD$ keeps record of the depth of the closest projected point to the (xy, xz, yz) planes, for every projected position.

4. For each point $P_i = (p_x, p_y, p_z)$ in the PC with color C_p, two parallel planes from the set PL will be jointly

processed, starting with $(xy, \acute{x}\acute{y})$ .The following steps are performed:

i. Define the depth D_i of point P_i as the value for the coordinate p_z since z is the perpendicular axis to both the xy and $\acute{x}\acute{y}$ planes.
ii. If D_i is less than or equal to the maximum depth at position $(p_{\acute{x}}, p_{\acute{y}})$ in *FarD* for this plane, then the point is projected onto position $(p_{\acute{x}}, p_{\acute{y}})$ of image $I_{\acute{x}\acute{y}}$. The corresponding pixel in the occupancy map is also set to '1' and the corresponding maximum depth is updated to the depth D_i as follows:

$$\begin{aligned} &\text{if } D_i \leq FarD(p_{\acute{x}}, p_{\acute{y}}) \text{ then} \\ &\quad I_{\acute{x}\acute{y}}\left(p_{\acute{x}}, p_{\acute{y}}\right) = C_p \text{ and} \\ &\quad OM_{\acute{x}\acute{y}}\left(p_{\acute{x}}, p_{\acute{y}}\right) = 1 \text{ and} \\ &\quad FarD\left(p_{\acute{x}}, p_{\acute{y}}\right) = D_i \end{aligned} \tag{5}$$

iii. If D_i is larger than or equal to the minimum depth at position (p_x, p_y) in *NearD* for this plane, then the point is projected onto position (p_x, p_y) of image $I_{x,y}$. The corresponding pixel in the occupancy map is also set to '1' and the corresponding minimum depth is updated to the depth D_i as follows:

$$\begin{aligned} &\text{if } D_i \geq NearD(p_x, p_y) : \\ &\quad I_{x,y}\left(p_x, p_y\right) = C_p \text{ and} \\ &\quad OM_{x,y}\left(p_x, p_y\right) = 1 \text{ and} \\ &\quad NearD\left(p_x, p_y\right) = D_i \end{aligned} \tag{6}$$

Finally, steps i-iii have to be repeated for the other two pairs of planes in PL, i.e., $(xz, \acute{x}\acute{z})$ and $(yz, \acute{y}\acute{z})$, thus obtaining the two pairs of images $(I_{\text{xz}}, I_{\acute{x}\acute{z}})$ and $(I_{\text{yz}}, I_{\acute{y}\acute{z}})$.

2) FILTERING

At this stage, every point that is not occluded should have been projected. In this context, it is possible that some points from the surface farther to the projection plane may be visible from it and, thus, projected onto it. This occurs because the surfaces close and farther away to the projected plane may not have the same density and may not be aligned. Thus, some pixel positions may be filled from surfaces that are not even visible after rendering (from the perspective of the projected plane).

These points should be filtered out by comparing their depth to the depth of their neighboring pixels in a window w since the depth of these faraway points are significantly different from their projected neighboring closest points. The algorithm for filtering the points from the 'back part' of the PC that are not seen by the users from that perspective, proceeds as follows:

1. For each occupied pixel (u, v) in the projected image for planes $PL \in \{xy, xz, yz\}$, compute the difference between *NearD(u,v)* and the average of *NearD* values for its neighbors in a 2D window with size $w \times w$. If this difference is smaller than a predefined threshold τ, then set that position to unoccupied in the associated occupancy map and set the corresponding projected image position to the background value (7), as shown at the bottom of the page, where $\tau > 0$ is found through experimentation.
2. For each occupied pixel (u, v) in the projected image for planes $\acute{PL} \in \{\acute{x}\acute{y}, \acute{x}\acute{z}, \acute{y}\acute{z}\}$, compute the difference between *FarD(u,v)* and the average of *FarD* values for its neighbors in a 2D window with size $w \times w$. If this difference is smaller than a predefined threshold, then set that position to unoccupied in the associated occupancy map and set the corresponding projected image position of the background value (8), as shown at the bottom of the page, where $\tau > 0$ is found through supervised experiments.

The algorithm above is a proximity filtering algorithm that depends on the threshold value τ that should be set according to the curvature of the object's surface, which is typically not very high. Even if a few PC front points are unduly 'filtered', the impact is small as they will be filled during the padding process performed next. After some experiments, it was found that a fixed filtering threshold τ of 20 was effective in the filtering of these already projected pixels.

D. CROPPING

The output of the previous step corresponds to six projected images with size $2^p \times 2^p$, one per projection plane. These images may contain a significant number of background pixels corresponding to empty areas around the PC object projection. To remove the undesired influence of this background data, the excessive background around the projected

$$\begin{aligned} &\text{if } \left| NearD_K(u,v) - \frac{\sum_{(i,j)\in w} NearD_{PL}(i,j) \times OM_{PL}(i,j)}{\sum_{(i,j)\in w} OM_{PL}(i,j)} \right| \leq \tau : \\ &I_{PL}(u,v) = 255 \text{ and } OM_K(u,v) = 0 \end{aligned} \tag{7}$$

$$\begin{aligned} &\text{if } \left| FarD_{\acute{K}}(u,v) - \frac{\sum_{(i,j)\in w} FarD_{\acute{PL}}(i,j) \times OM_{\acute{PL}}(i,j)}{\sum_{(i,j)\in w} OM_{\acute{PL}}(i,j)} \right| \geq \tau : \\ &I_{\acute{PL}}(u,v) = 255 \text{ and } OM_{\acute{K}}(u,v) = 0 \end{aligned} \tag{8}$$

PC is cropped out to the minimum size including the occupied pixels after projection. First, a bounding box surrounding the PC object(s) for each projected image is obtained for the reference and degraded PCs, using the occupancy map obtained during the projection step. More precisely, by scanning from top left to bottom right, the positions of the first and last occupied pixels in each occupancy map are used to define the bounding box for the object in the projected map. Note that the reference and degraded bounding boxes for each view are identical as using aligned geometries. Next, cropping is performed using the obtained bounding box for the reference/degraded projected images associated with each plane.

Figure 5 shows an example of the cropping operation for the *Longdress* PC. While the full projected image is shown in Figure 5(a) with the precision bounding box in green, Figure 5(b) shows the cropped image with a largely reduced background.

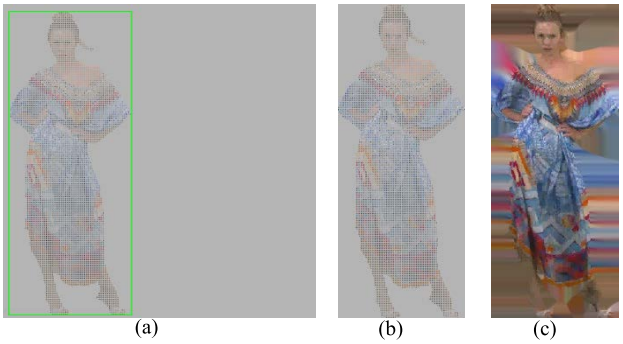

FIGURE 5. Frontal view projection for Longdress PC decoded with G-PCC in Octree coding mode and RAHT color coding mode at the medium rate: (a) projected image with the bounding box in green; (b) cropped image; (c) padded image. The padded image looks 'brighter' since the non-padded images still have many light grey background pixels inside the foreground area.

E. PADDING

After cropping, the background information is significantly removed. However, background pixels associated with unoccupied points inside the object and some background areas around the object still exist. To avoid the negative impact of the background color pixels when computing the 2D quality metric, the color values for the unoccupied pixels (and thus with uniform background values) should be set using some 2D interpolation technique, thus more appropriately filling all empty spaces. This operation aims to replicate the effect of PC rendering, which creates continuous surfaces without holes using appropriate rendering primitives, thus removing the bias on the quality score due to pixels which artificially have the same identical value for reference and degraded projected images, i.e., areas without any distortion. This padding approach allows applying any 2D quality metric, which uses as input a rectangular image without requiring any adaptations of the 2D quality metric to work with arbitrarily shaped 2D regions, i.e., to consider only foreground pixels. To fill the non-occupied pixels, it is proposed to use an image inpainting technique from the literature called Navier-Stokes [38], which has been selected due to its good performance. The occupancy maps created during the projection process are used as a padding mask to guarantee that only the unoccupied pixels are padded. Figure 5(c) shows an example of a padded image for the *Longdress* PC.

F. FUSION

Before obtaining the final *JGC-ProjQM* PC quality metric, two intermediate scores are computed, $JGC\text{-}ProjQM_{reference}$ and $JGC\text{-}ProjQM_{degraded}$, corresponding to the two parallel branches in the architecture, one corresponding to the reference geometry and another to the degraded geometry. To combine the two intermediate quality scores, the following linear model is proposed:

$$JGC\text{-}ProjQM = \alpha JGC\text{-}ProjQM_{reference} + \beta JGC\text{-}ProjQM_{degraded} \quad (9)$$

In (9), the α and β parameters are estimated using a least-squares linear regression procedure that minimizes the residual sum of the squared difference between the objective scores predicted by the linear approximation, *JGC-ProjQM*, and the Mean Opinion Scores (MOS) available in some selected dataset. For this paper, the used dataset was the MPEG Point Cloud Compression Dataset (M-PCCD) [39]. The α and β parameters obtained with this procedure are presented in Section IV.B. Although more complex models may be selected, they typically require more parameters, bringing the risk of overfitting to the selected dataset. This is rather critical as there are not that many PC datasets available with subjective scores and representative geometry and color degradations, especially compared to image and video datasets.

IV. PERFORMANCE ASSESSMENT

The main objective of this section is to evaluate the proposed *JGC-ProjQM* PC quality metric compared with the best performing PC quality metrics available in the literature, projection-based or not. Moreover, an ablation study is presented to assess the performance impact of each module in the *JGC-ProjQM* architecture.

A. SUBJECTIVE EVALUATION DATASETS

To evaluate the proposed *JGC-ProjQM* performance, three PC quality assessment datasets were used. These datasets were created under different conditions, namely using 2D and head-mounted (HMD) displays for visualization and allowing the subjects to interact with the PC (active method) or just rendering a 2D video (passive method). The datasets are described next:

1. **M-PCCD dataset** [37], publicly available in [39]. The MPEG Point Cloud Compression Dataset (M-PCCD) is rather popular in the PC community and includes stimuli where geometry and color are both

coded. The test material in this dataset corresponds to nine PCs, including four inanimate objects and five human figures. While *Longdress*, *Loot*, *Redandblack*, *Soldier*, *The20smaria* and *Head* are from the MPEG repository [40], *Romanoillamp* and *Biplane* are from the JPEG repository [41] and *Amphoriskos* from the Sketchfab dataset [42]. The *Redandback* PC has been used for training the subjects. The PCs are shown in Figure 6 and their characteristics listed in Table 1. The PCs have been coded in the following conditions: i) 24 rates for the G-PCC standard with six different rates for each combination of Octree and TriSoup geometry coding modes with the RAHT and Lifting color coding modes; and ii) five rates for the V-PCC standard. The rates were selected based on the MPEG Common Test Conditions (CTC) recommendations [15]. The subjective scores have been obtained in two separate laboratories, each with 20 subjects.

FIGURE 6. PCs in the M-PCCD dataset. From top left to bottom right: (a) Amphoriskos; (b) Biplane; (c) Head; (d) Romanoillamp; (e) Longdress; (f) Loot; (g) Redandblack; (h) Soldier; (i) The20smaria.

TABLE 1. M-PCCD PCs and associated characteristics.

Name	Type	Repository	Precision	No. Points
Amphoriskos	Object	*Sketchfab*	10-bit	814.474
Romanoillamp	Object	JPEG repository	10-bit	636.127
Biplane	Object	JPEG repository	10-bit	1.181.016
Head	Object	MPEG repository	9-bit	938.112
Longdress	People	MPEG repository	10-bit	857.966
Loot	People	MPEG repository	10-bit	805.285
Redandblack	People	MPEG repository	10-bit	757.691
Soldier	People	MPEG repository	10-bit	1.089.091
The20smaria	People	MPEG repository	10-bit	1.553.937

2. **ICIP2020 dataset** [9], publicly available at [44]: In this dataset, six different PCs are coded with V-PCC, G-PCC Octree and G-PCC TriSoup using five rates for both texture and geometry. The six (static) PCs are *Longdress*, *Loot*, *Soldier*, *Redandback, Ricardo10* and *Sarah9*, available in the JPEG repository [41], all representing human figures. For G-PCC, the PCs were coded with five rates (and thus quality levels) from those defined in the MPEG CTC recommendations; for V-PCC, they are coded with four rates as defined in MPEG CTC recommendations and an additional one to achieve even low bitrates. A total of 96 scores were obtained per evaluation session, with a total of 76 subjects from four laboratories.
3. **SIAT dataset** [34], publicly available at [45]: Twenty different PCs were selected, including six full-body human figures, four upper-body figures and 10 inanimate objects. Again, the PCs were obtained from the MPEG, JPEG and sketchfab repositories. All PCs were coded with V-PCC in 17 different rates. These rates are several (not all) combinations of geometry QPs from (0, 20, 28, 32, 36) and color QPs from (0, 27, 32, 37, 47); a losslessly compressed reference was also included. A mix of 38 experts and non-expert subjects from one lab performed the subjective test.

All datasets include both the MOS values as well as the reference and degraded/decoded PCs. For all datasets, the subjective quality assessment methodology was the Double Stimulus Impairment Scale (DSIS). Moreover, the outlier detection algorithm described in the ITU-R Recommendation BT.500-13 [43] was performed, separately for each laboratory, to exclude subjects whose ratings deviated drastically from the remaining quality scores. In all datasets, no outliers were found.

Moreover, a point-based rendering procedure was used [6] to obtain PCs for all datasets. For the M-PCCD dataset, an adaptive point size approach was selected while for the ICIP2020 and SIAT datasets, the point size was kept fixed. For M-PCCD and ICIP2020 datasets, PCs were shown side-by-side in a 2D display: For the M-PCCD dataset, an interactive navigation application allowed the subjects to select and modify their viewpoint while, for the ICIP2020 dataset, no interaction was possible, and thus PCs were converted to video sequences of 12s, according to a predefined trajectory. For the SIAT dataset, PCs were rendered with an immersive application that supports six degrees of freedom (6DoF). In this case, subjects were allowed to freely explore the environment and observe the original and distorted PCs side-by-side from any angle and without occlusion. An HMD and two hand controllers (from HTC Vive) were used to allow the subjects to interact with the PCs.

Regarding data processing, for the M-PCCD and ICIP2020 datasets, the MOS was computed by averaging all the subject scores for each stimulus, while for the SIAT dataset, DMOS values were computed as the difference between the scores of the degraded and original PCs. Note that original PCs were included as a hidden reference and thus also subjectively evaluated.

B. FUSION PARAMETERS OVERFITTING CHECKING

The final *JGC-ProjQM* metric is a linear combination of the intermediate quality metrics associated with the two

architectural branches as shown in Figure 1. The α and β parameters are estimated with a least-squares linear regression procedure which uses the objective and subjective scores for all PCs and therefore is important to confirm that no overfitting happens. For this purpose, the M-PCCD dataset is randomly split into 75% training data and 25% test data 100 times. For each iteration, the α and β parameter values are estimated from the training data split and used to compute the final PC quality metric score for the test data split. The subjective-objective performance is measured with the Pearson Linear Correlation Coefficient (PLCC) which has been computed for each iteration and is shown in Figure 7. The average performance for all iterations (i.e., for different splits) is computed as the average PLCC over all iterations and is also shown in Figure 7 as a red line. The PLCC performance computed using all data for both training and testing is also shown in Figure 7. The analysis of the results shows that the final *JGC-ProjQM* values are very close to the average *JGC-ProjQM* value, thus implying that the obtained performance is not due to overfitting. This overfitting analysis is made for four 2D quality metrics, notably SSIM, MS-SSIM, FSIM, and Y-PSNR.

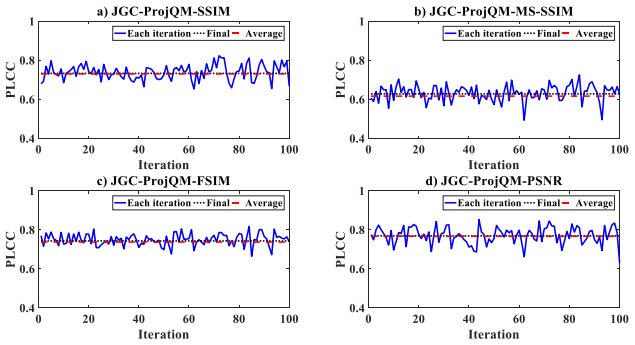

FIGURE 7. JGC-ProjQM PLCC overfitting analysis for the following 2D quality metrics: (a) SSIM; (b) MS-SSIM; (c) FSIM; (d) Y-PSNR.

The α and β parameters obtained by linear regression as well as the ratio between them are shown in Table 2 when several 2D quality metrics are used. The obtained fusion parameters values show that they depend significantly on the 2D quality metric. While for some 2D quality metrics, such as FSIM and VSI, more weight is given to the degraded geometry branch, for others, such as DISTS, LPIPS and MS-SSIM, the original geometry branch contributes more to the final score.

C. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the objective-subjective correlation performance of the proposed PC quality metric[1] is evaluated and compared with some relevant benchmarks. In the following, non-linear fitting is applied to all objective quality scores,

[1]A python implementation is made available online at https://github.com/AlirezaJav/Projection-based-PC-Quality-Metric

TABLE 2. *JGC-ProjQM* fusion parameters values for several 2D quality metrics.

2D quality metric	α	β	$\lvert\alpha/\beta\rvert$
DISTS	0.499	0.339	1.47
LPIPS	1.065	-0.191	5.58
FSIM	0.041	0.703	0.06
VSI	0.041	0.640	0.06
HaarPSI	0.239	0.553	0.43
VIPF	0.673	0.159	4.23
SSIM	0.283	0.455	0.62
MS-SSIM	0.874	-0.252	3.47
PSNR-HVS-M	-0.271	1.028	0.26
PSNR-HVS	-0.492	1.249	0.39
Y-PSNR	-0.519	1.275	0.41

in this case, using the following logistic function:

$$MOS_p = \beta_2 + \frac{\beta_1 - \beta_2}{1 + e^{-\left(\frac{Q_i - \beta_3}{\beta_4}\right)}} \tag{10}$$

where Q_i is the objective metric score and $\beta_1, \ldots, \beta_4$ are the regression model parameters. This approach allows to fit the objective metric scores to the perceptual (MOS) scale to obtain the fitted predicted MOS scores.

To assess the quality metrics performance, the PLCC, Spearman Rank Order Correlation Coefficient (SROCC), and Root Mean Squared Error (RMSE) are used. The correlation coefficients, which have floating-point values between 0 and 1, are rounded to two digits and multiplied by 100 to express their values in percentage. When the PLCC and SROCC scores are close to '100', the predicted objective quality scores are highly correlated and have a monotonic relationship with the ground-truth MOS scores. As a measure of monotonicity, SROCC does not depend on the selected fitting function.

A wide range of objective PC quality metrics available in the literature was selected as benchmarks to evaluate the proposed PC quality metric more effectively. These metrics are listed in Table 3.

1) PROJECTION-BASED METRICS COMPARISON

In this section, the proposed *JGC-ProjQM* metric performance is compared with the projection-based metrics proposed in [37] and [11] for the M-PCCD dataset. For a fair comparison, the same 2D quality metric is used in the *JGC-ProjQM* quality metric and the selected benchmarks. In [37], VIFP, SSIM, MS-SSIM and Y-PSNR are used, while, in [11], only Y-PSNR is used.

Table 4 clearly shows that the proposed *JGC-ProjQM* metric significantly outperforms the already available projection-based PC quality metrics, for the same 2D quality metric.

The minimum and maximum *JGC-ProjQM* gains are 18.3% for PLCC, 18.0% for SROCC and 0.3 for RMSE, compared to *Alexiou et al.* [37]. These gains are rather large and consistent across the selected quality metric performance measures, i.e., PLCC, SROCC and RMSE, especially

TABLE 3. PC quality metrics selected as benchmarks for performance comparison.

Metric Name	Short Description
D1and D1-PSNR [15]	Po2Po MSE error and associated PSNR.
D2 and D2-PSNR [15]	Po2Pl MSE error and associated PSNR.
Hausdorff distance and Hausdorff PSNR [13]	Po2Po and Po2Pl Hausdorff distance and associated PSNR.
Y-MSE and Y-PSNR [15]	MSE between luminance of points and their nearest neighbor and associated PSNR.
Pl2Pl Angle-MSE [16]	Mean squared of angular similarity between underlying surfaces at each point and its nearest neighbor.
PCQM [27]	Weighted average of geometry curvature features and lightness, chroma and hue features between reference and decoded PCs.
PointSSIM [29]	Color features are extracted in a local neighborhood around each point in the reference and degraded PCs using variance as statistical dispersion measure.
H^{Y}_{L2} *[22]*	Computed from the color histogram and correlogram of the luminance component.
d_{qc} *[22]*	Linear combination of Po2Pl MSE with H^{Y}_{L2}.
PCM_{RR} *[28]*	Reduced reference quality metric corresponding to the weighted average of 7 geometry, 7 color and 7 normal features.
Geotex [23]	Local Binary Pattern descriptors applied to the luminance are extracted and used for quality metric computation.
Proj-Y-PSNR [11]	Projection-based quality metric using six faces of a cube; the six images are concatenated and Y-PSNR are computed.
Proj-Y-MSSIM and Proj-Y-VIFP [37]	Projection-based metric using 42 views; the final metric is obtained using several 2D metrics and pooled using l1-norm.
GH-PSNR [17]	Generalized Hausdorff distance-based PSNR, considering 98% of the distances, using maximum pooling function.
RA-PSNR [20]	Resolution-Adaptive PSNR considering the rendering resolution computed over the 10 nearest neighbors.
Po2D PSNR [18][19]	PSNR based Point-to-Distribution metric where the mean Mahalanobis distances between a point and a distribution of points is computed.

considering that [37] uses 42 views, which is a much larger (and also more complex) number of views than the six views considered in *JGC-ProjQM*. Another interesting conclusion is that VIFP and Y-PSNR lead to the best and worst correlation performance, respectively.

2) 2D QUALITY ASSESSMENT METRICS IMPACT

As stated in Section III, the proposed *JGC-ProjQM* metric is flexible enough to accommodate any 2D quality metric. In this section, the performance of several variants of the proposed metric is evaluated for several 2D quality metrics (*JGC-ProjQM-2DMetricName*), again for the M-PCCD dataset. This will allow identifying the 2D quality metric that leads to the best correlation performance. In this case, the same 2D quality metric is used in both architectural branches of the proposed quality metric, i.e., reference and degraded, especially because the fusion module works best when the quality range and scale are similar for both the reference and degraded geometry branches. Since the used 2D quality metric can significantly influence the *JGC-ProjQM* performance,

TABLE 4. Objective-subjective correlation performance comparison of the proposed *JGC-ProjQM* metric with [37] and [11] for a set of 2D quality metrics.

Projection-based PC quality metric	2D quality metric	PLCC	SROCC	RMSE
JGC-ProjQM	VIFP	**83.0**	**85.5**	**0.760**
Alexiou et al. [37]		71.5	74.2	0.951
JGC-ProjQM Gain		11.5	11.3	0.191
JGC-ProjQM	SSIM	**80.9**	**81.3**	**0.800**
Alexiou et al. [37]		62.6	63.3	1.061
JGC-ProjQM Gain		18.3	18.0	0.261
JGC-ProjQM	MS-SSIM	**79.5**	**82.8**	**0.830**
Alexiou et al. [37]		70.9	75.2	0.959
JGC-ProjQM Gain		8.6	7.6	0.129
JGC-ProjQM	Y-PSNR	**77.1**	**79.1**	**0.870**
Alexiou et al. [37]		66.7	62.8	1.013
JGC-ProjQM Gain		10.4	16.3	0.143
de Queiroz et al. [11]		43.0	33.1	1.228
JGC-ProjQM Gain		34.1	46.0	0.358

TABLE 5. *JGC-ProjQM* objective-subjective correlation performance for a large set of 2D quality metrics, ordered from best to worse PLCC.

2D quality metric	*PLCC*	*SROCC*	*RMSE*
DISTS	**94.7**	**95.6**	**0.439**
LPIPS	92.3	93.2	0.523
FSIM	88.2	90.1	0.640
VSI	85.4	87.6	0.707
HaarPSI	84.8	87.7	0.721
VIPF	83.0	85.5	0.758
SSIM	79.5	82.8	0.800
MS-SSIM	80.9	81.3	0.825
PSNR-HVS-M	78.7	81.3	0.840
PSNR-HVS	78.4	80.5	0.845
Y-PSNR	77.1	79.1	0.866

a wide set of available quality 2D metrics are evaluated, notably: 1) Y-PSNR; 2) PSNR-HVS [46]; 3) PSNR-HVS-M [47]; 4) Structural Similarity Index Metric (SSIM) [48]; 5) Multi-Scale Structural Similarity Index Metric (MS-SSIM) [49]; 6) Visual Information Fidelity Measure (VIFP) [50]; 7) Feature Similarity Index (FSIM) [51]; 8) Visual Saliency Index (VSI) [52]; 9) Learned Perceptual Image Patch Similarity (LPIPS) [53]; 10) Deep Image Structure and Texture Similarity (DISTS) [54]; and 11) Haar Perceptual Similarity Index (HaarPSI) [55].

Table 5 shows the *JGC-ProjQM* correlation performance for a large set of 2D quality metrics, considering all possible coding degradations in the used dataset (all codecs data). The DISTS 2D quality metric has the best correlation performance among all the 2D quality metrics while LPIPS, FSIM, and HaarPSI come in the following positions. Both DISTS and LPIPS are very recent 2D quality metrics that use powerful deep-learning features to perform the quality assessment. More specifically, DISTS includes both color and structure similarity components, which are weighted to achieve a higher correlation with the perceived quality and to be invariant to small color changes. LPIPS computes

distances between features extracted from the reference and degraded projected images at different layers of a neural network. For both DISTS and LPIPS metrics, a perceptual feature space is used. Typically, these quality metrics weigh more general appearance changes than small color changes, where their elements may have different locations, sizes, colors and orientations. This fits rather well with the projected images obtained by the proposed *JGC-ProjQM* metric where small color changes may occur due to the recoloring process. The proposed *JGC-ProjQM* metric with the four 2D quality metrics with best correlation performance, i.e., *DISTS*, *LPIPS*, *FSIM*, *VSI*, will be used for the remaining experiments reported in this paper.

3) OVERALL POINT CLOUD QUALITY METRICS PERFORMANCE COMPARISON

In this section, the performance of the proposed *JGC-ProjQM* metric is compared with many state-of-the-art PC quality metrics available in the literature for the three selected datasets.

The correlation evaluation is first done for the M-PCCD dataset; since this is a very popular dataset, it is possible to compare results with a large number of benchmark quality metrics, nowadays available in the literature. A more extensive analysis is also performed for this dataset since: 1) allows to analyze the G-PCC and V-PCC codecs independently and jointly; 2) correlations can be obtained for a large number of PC quality metrics.

In this context, Table 6 shows the PLCC, SROCC, and RMSE scores for the selected benchmark quality metrics using the M-PCCD dataset, considering all the codecs scores together as well as the V-PCC and G-PCC scores individually. When a large set of quality metrics is proposed in a reference, only the best variants are included. This separation can be justified by the fact that the PC quality metrics performance may be significantly influenced by the different types of coding artifacts generated by different PC coding solutions (which are rather different). This split was performed as follows: i) G-PCC decoded PCs, including TriSoup and Octree geometry coding modes as well as RAHT and Lifting color coding modes; ii) V-PCC decoded PCs; and iii) all decoded PCs together. From Table 6, where the best correlation results are highlighted in bold and the second best are underlined, the following conclusions may be derived:

- **Overall correlation performance:** The proposed *JGC-ProjQM* metric using DISTS is the best performing metric for PC quality assessment, achieving the best PLCC, SROCC and RMSE scores. This result also confirms that by projecting a PC into several 2D images (which are close to what a user sees) and then exploiting the power of 2D quality metrics, a top correlation performance can be obtained. The proposed *JGC-ProjQM* with the LPIPS and LogP2D-JGY [19], [20] quality metrics also have a very high correlation performance.
- ***JGC-ProjQM* vs point-based PC quality metrics:** The proposed *JGC-ProjQM* metric significantly outperforms the point-based D1-PSNR and D2-PSNR and plane-to-plane quality metrics that are currently used by the MPEG and JPEG standardization groups. The correlation gains are rather significant, notably up to 32.3% for PLCC and 47.7% for SROCC for all data. Moreover, *JGC-ProjQM-DISTS* outperforms the best point-based metric in the literature (Po2D LogP2D-JGY) by 1.8% for both PLCC and SROCC for all data, and larger gains for all remaining point-based quality metrics.
- ***JGC-ProjQM*vs feature-based PC quality metrics:** The overall correlation performance of the best proposed projection-based metric, i.e., *JGC-ProjQM-DISTS*, is almost 2.1% and 3.8% higher in PLCC and SROCC, respectively, than the best feature-based quality metric, i.e., PointSSIM. The feature-based quality metrics often come in second place, achieving also rather good correlation performance, especially compared to point-based quality metrics. The only exception in point-based metrics is Po2D LogP2D which is the third-best metric. This metric also implicitly considers some local statistical features such as the mean and standard deviation.
- **V-PCC decoded data:** The *JGC-ProjQM-DISTS* is the best performing metric for V-PCC decoded data with 86.4% for PLCC and 85.3% for SROCC. The benchmark point-based and projection-based quality metrics fail to reliably assess the V-PCC decoded quality, except for LogP2D-JGY which shows high performance. The *JGC-ProjQM-DISTS* metric outperforms the best projection-based metric in the literature (Proj-Y-VIFP) by 42.7% for PLCC and 49.7% for SROCC. For the best point-based metric (LogP2D-JGY), the performance increase is 1.8% for both PLCC and SROCC. The gains against the best MPEG/JPEG adopted metrics (D2-PSNR) are also 26.1% in PLCC and 30% in SROCC. The gains against the best feature-based metric (PointSSIM) are 3.4% and 0.8% for PLCC and SROCC, respectively. In summary, *JGC-ProjQM-DISTS*, LogP2D-JYC and PointSSIM show relatively higher performance compared to the other quality metrics.
- **G-PCC decoded data**: The *JGC-ProjQM-DISTS* metric variant is the best performing metric for G-PCC decoded data with 95.8% and 96% for PLCC and SROCC, respectively. Feature-based and point-based metrics (except for Pl2Pl) also show acceptable performance for G-PCC decoded data. However, *JGC-ProjQM-DISTS* outperforms the best MPEG/JPEG adopted metrics (D2-PSNR) by 12.4% and 8.7% for PLCC and SROCC, respectively. The correlation gains against the best feature-based quality metric (PointSSIM) are 1.4% and 3.1% for PLCC and SROCC, respectively. The benchmark projection-based metrics do not show an

TABLE 6. Objective-subjective correlation performance comparison between the proposed *JGC-ProjQM* metric and the benchmarking metrics using the M-PCCD dataset. Cells with a dash (-) show missing performance since some of these results were obtained from the relevant paper.

Type		Metric Name	*All*			*V-PCC*			*G-PCC*		
			PLCC	*SROCC*	*RMSE*	*PLCC*	*SROCC*	*RMSE*	*PLCC*	*SROCC*	*RMSE*
Point-based	*Po2Po*	*D1* [15]	84.8	86.8	0.722	46.3	42.0	0.928	88.6	90.0	0.651
		D1-PSNR [15]	77.7	79.7	0.857	30.4	28.2	0.997	82.5	83.9	0.794
		Hausdorff [13]	1.4	37.0	1.360	14.7	-17.5	1.047	5.3	54.4	1.404
		Hausdorff PSNR [13]	66.1	36.6	1.021	27.1	-14.9	1.008	76.0	53.3	0.912
		GH-PSNR [17]	84.6	86.9	0.726	57.8	57.3	0.854	88.5	89.9	0.653
		RA-PSNR [20]	88.8	90.2	0.626	68.9	67.3	0.759	91.0	91.8	0.584
		Y-MSE [15]	66.3	66.2	1.018	37.9	33.3	0.969	70.3	70.3	0.998
		Y-PSNR [15]	67.1	66.2	1.009	37.6	33.3	0.970	71.4	70.3	0.984
	Po2Pl	*D2* [15]	85.9	88.4	0.695	73.5	68.8	0.710	87.9	90.6	0.669
		D2-PSNR [15]	80.5	83.8	0.808	60.3	55.3	0.835	83.4	87.3	0.774
		Hausdorff [13]	67.2	50.5	1.007	23.8	12.8	1.017	78.4	66.3	0.871
		Hausdorff PSNR [13]	56.3	49.3	1.124	28.6	13.5	1.003	68.7	65.3	1.020
		GH-PSNR [17]	84.3	87.9	0.731	75.1	71.2	0.691	87.5	91.0	0.680
		RA-PSNR [20]	88.9	89.9	0.622	79.9	76.9	0.629	90.3	91.5	0.604
	Po2D	*MMD* [18]	86.9	88.7	0.672	71.8	69.0	0.729	88.8	90.3	0.647
		MMD-PSNR [18]	86.9	88.7	0.673	71.9	69.0	0.728	88.7	90.3	0.648
		P2D- JGY (NN25) [19]	92.9	93.8	0.503	84.2	**85.5**	0.564	94.0	94.6	0.477
		LogP2D- JGY (NN25) [19]	92.9	93.8	0.502	84.3	**85.5**	0.563	94.1	94.6	0.477
	Pl2Pl	*Angle-MSE* [16]	62.4	47.7	1.063	51.6	34.1	0.897	69.0	55.0	1.016
Feature-based		*PCQM* [27]	89.9	91.6	0.597	-	-	-	-	-	-
		d_{gc} [22]	90.4	92.0	0.585	75.3	74.0	0.689	92.5	93.9	0.533
		H^{Y}_{L2} [22]	85.3	88.4	0.710	65.7	68.3	0.789	87.9	91.9	0.669
		PCM_{RR} [28]	90.2	90.7	0.573	71.6	64.8	0.731	89.2	91.0	0.636
		PointSSIM [29]	92.6	91.8	0.514	83.0	84.5	0.584	94.4	92.9	0.462
		Geotex [23]	-	87.9	-	-	-	-	-	-	-
		GraphSim [30]	94.3	95.3	0.453	-	-	-	-	-	-
		PQA-Net [32]	60.0	65.0	-	-	-	-	-	-	-
		LBP_N [31]	90.3	91.7	1.339	-	-	-	-	-	-
Projection-based		*Proj-Y-MS-SSIM* [37]	70.9	75.2	0.959	31.9	35.4	0.992	75.3	50.1	0.924
		Proj-Y-VIFP [37]	71.5	74.2	0.951	43.7	35.6	0.942	75.0	79.2	0.929
		Proj-Y-PSNR [11]	43.0	33.1	1.228	29.7	-10.1	1.000	47.7	35.7	1.234
		Proj-IW-SSIM [35]	71.7	74.9	0.948	-	-	-	-	-	-
Proposed Projection-based		*JGC-ProjQM-FSIM*	88.2	90.1	0.640	71.0	72.1	0.737	90.3	92.1	0.604
		JGC-ProjQM-VSI	85.4	87.6	0.707	64.5	63.9	0.800	88.1	90.1	0.664
		JGC-ProjQM-LPIPS	92.3	93.2	0.523	80.7	79.5	0.618	93.5	94.2	0.497
		JGC-ProjQM-DISTS	**94.7**	**95.6**	**0.439**	**86.4**	85.3	**0.526**	**95.8**	**96.0**	**0.402**

acceptable performance for the quality assessment of G-PCC decoded PCs.

For a more exhaustive performance assessment, the subjective-objective correlation results for the four best variants of the proposed *JGC-ProjQM* metric are also compared with other state-of-the-art quality metrics using the ICIP2020 and SIAT datasets. Table 7 and 8 show the PLCC, SROCC and RMSE scores for the selected benchmark quality metrics using the two additional datasets. The following conclusions can be derived:

- ICIP2020 dataset: The proposed *JGC-ProjQM-DISTS* metric variant is still the best performing metric as shown in Table 7. In this case, an even higher correlation performance could be obtained (96%), when compared to the M-PCCD dataset; this behavior is consistent for both PLCC, SROCC and RMSE. Interestingly, the GraphSim metric (the 2nd best for the M-PCCD dataset) performs much worse, thus highlighting it suffers from generalization problems.
- SIAT dataset: The proposed *JGC-ProjQM-DISTS* metric variant is the third best quality metric among the assessed quality metrics, as shown in Table 8. The overall subjective-objective correlation scores are lower for this dataset, and even the IW-SSIM projection-based metric proposed by the authors of the SIAT dataset [34] only achieves 81.8%. This may suggest that when PCs are visualized with HMD devices in 3D immersive environments, the proposed projection-based quality metrics may need to be enhanced to account for other factors, such as the intrinsic quality of the original PC.

Finally, the overall correlation performance evaluation has shown that the proposed *JGC-ProjQM-DISTS* metric variant outperforms many state-of-the-art PC quality metrics, notably being the best for the M-PCCD and ICIP2020 datasets and the third best for the SIAT dataset. These simultaneous top rankings for three different datasets are not repeated by any other quality metric. For example, well-known PCQM [27] and PointSSIM [29] are among the best with PLCC and

TABLE 7. Objective-subjective correlation performance comparison between the proposed *JGC-ProjQM* metric and the benchmarking metrics using the ICIP2020 [9] dataset.

Type		Metric Name	PLCC	SROCC	RMSE
Point-based	*Po2Po*	*D1* [15]	94.6	93.4	0.368
		D1-PSNR [15]	86.8	85.5	0.540
		Hausdorff [13]	40.1	53.1	1.045
		Hausdorff PSNR [13]	54.8	45.6	0.911
		Y-MSE [15]	87.6	89.2	0.551
		Y-PSNR [15]	88.7	89.2	0.525
	Po2Pl	*D2* [15]	95.9	95.0	**0.321**
		D2-PSNR [15]	91.3	91.0	0.443
		Hausdorff [13]	53.4	61.3	0.966
		Hausdorff PSNR [13]	58.0	54.7	0.887
	Pl2Pl	*Angle-MSE* [16]	92.5	91.2	0.432
Feature-based		*PCQM* [27]	26.3	41.1	1.136
		PCM_{RR} [28]	72.3	88.7	1.136
		PointSSIM [29]	67.6	56.4	0.837
		GraphSim [30]	86.0	84.5	0.579
Projection-based		*Proj-IW-SSIM* [35]	91.0	89.7	0.457
Proposed Projection-based		*JGC-ProjQM-FSIM*	88.0	88.0	0.530
		JGC-ProjQM-VSI	89.6	88.3	0.510
		JGC-ProjQM-LPIPS	94.1	93.0	0.390
		JGC-ProjQM-DISTS	**96.0**	**95.2**	0.330

TABLE 8. Objective-subjective correlation performance comparison between the proposed *JGC-ProjQM* metric and the benchmarking metrics using the SIAT [34] dataset.

Type		Metric Name	*PLCC*	*SROCC*	*RMSE*
Point-based	*Po2Po*	*D1* [15]	31.4	39.6	0.122
		D1-PSNR [15]	28.5	32.8	0.124
		Hausdorff [13]	29.8	37.9	0.123
		Hausdorff PSNR [13]	28.3	32.7	0.124
		Y-PSNR [15]	43.4	34.8	0.121
	Po2Pl	*D2* [15]	35.0	41.3	0.121
		D2-PSNR [15]	31.1	34.6	0.123
		Hausdorff [13]	32.2	38.6	0.122
		Hausdorff PSNR [13]	30.1	34.2	0.123
	Pl2Pl	*Angle-MSE* [16]	23.7	26.9	0.128
Feature-based		*PCQM* [27]	65.4	66.7	0.099
		PointSSIM [29]	78.1	69.6	0.082
Projection-based		*Proj-Y-MS-SSIM* [37]	33.0	25.9	0.122
		Proj-Y-VIFP [37]	52.8	55.6	0.110
		Patch-Proj-IW-SSIM [34]	**81.8**	**69.7**	**0.074**
Proposed Projection-based		*JGC-ProjQM-FSIM*	70.0	63.0	0.091
		JGC-ProjQM-VSI	56.0	52.0	0.110
		JGC-ProjQM-LPIPS	68.0	64.0	0.100
		JGC-ProjQM-DISTS	70.4	64.1	0.090

SROCC above 90% on M-PCCD but they perform rather poorly on SIAT and ICIP datasets. Moreover, this shows that the proposed *JGC-ProjQM-DISTS* metric variant generalizes well and, thus, the α and β parameters that were learned by linear regression on the M-PCCD dataset are suitable for other datasets and application scenarios, e.g. even when the PCs are visualized with HMD devices.

4) JGC-PROJQM METRIC ABLATION STUDY

The proposed *JGC-ProjQM* metric includes several modules that have a different impact on the overall correlation performance. To individually assess the impact of each *JGC-ProjQM* module, an ablation study is performed using the M-PCCD dataset. More precisely, the performance of the four best *JGC-ProjQM* variants is measured for all stimuli (all codecs scores) included in the dataset, each time turning off one of the architectural modules while keeping the others, notably recoloring, cropping, and padding. Table 9 shows the PLCC and SROCC results after non-linear logistic fitting for this ablation study. The following conclusions can be derived:

TABLE 9. Objective-subjective correlation performance for the ablation study of the proposed JGC-ProjQM metric.

Proposed JGC-ProjQM Variant	Correlat. Metric	Correlat. Performance	Removed Module		
			Recoloring	Cropping	Padding
JGC-ProjQM-DISTS	*PLCC*	94.7	93.7	83.7	91.4
	SROCC	95.6	94.9	86.7	91.1
JGC-ProjQM-LPIPS	*PLCC*	92.3	89.5	80.0	81.1
	SROCC	93.2	90.7	81.8	80.8
JGC-ProjQM-FSIM	*PLCC*	88.2	78.5	81.5	86.6
	SROCC	90.1	81.5	83.7	87.5
JGC-ProjQM-VSI	*PLCC*	85.4	74.6	81.6	82.7
	SROCC	87.6	77.3	83.9	83.4

- **Recoloring**: The correlation performance results show the importance of the recoloring module for the *JGC-ProjQM* performance. For example, for *JGC-ProjQM-VSI* and *JGC-ProjQM-FSIM*, the absence of recoloring leads to losses of 10.8% and 9.7% for PLCC and 10.3% to 8.6% for SROCC. The performance losses after removing recoloring are lower for *JGC-ProjQM-DISTS* and *JGC-ProjQM-LPIPS*, mainly because these two recent 2D quality metrics are robust to geometry distortions and transformations. However, the performance gains by using the recoloring module are up to 10.8% and 10.3% for PLCC and SROCC, respectively.
- **Cropping**: The cropping module significantly improves the *JGC-ProjQM* performance by removing background pixels around the projected PC that are common to the reference and degraded PCs. These background areas work as a distractor for the 2D quality assessment metric and lower the *JGC-ProjQM* prediction power, even when the background pixels are padded. The performance gains associated with the cropping module are up to 12.3% and 11.4% for PLCC and SROCC, respectively.
- **Padding**: The padding module also improves the *JGC-ProjQM* correlation performance. While most of the excessive background is removed by cropping, some background area around the object remains. Moreover, background pixels that are visible in the object surface, often due to sparse sampling during acquisition or the removal of points during coding, are filled by padding. The performance gains by using the padding module are up to 11.2% and 12.4% for PLCC and SROCC, respectively.

It is important to note that the correlation performance gains for these modules may be significantly higher when less powerful 2D image metrics, i.e., with lower overall correlation performance, are used, such as MS-SSIM.

V. CONCLUSION

This paper proposes a novel joint geometry and color projection-based PC quality metric, *JGC-ProjQM*, that compares PCs in the 2D domain for two geometry conditions, i.e., reference and degraded geometry. The projection-based PC quality metric applies a projection to obtain six projected images, corresponding to different views over the PC, which are cropped and padded before performing a 2D quality assessment. After, any 2D quality metric can be applied and the intermediate quality scores for the two geometry conditions are fused to obtain the final *JGC-ProjQM* quality score. The objective-subjective correlation results show that the proposed *JGC-ProjQM* metric shows promising performance generally and the *JGC-ProjQM-DISTS* variant outperforms all the state-of-the-art PC quality metrics and, thus, PC quality can be efficiently measured in the 2D domain, especially when powerful 2D quality metrics are also exploited. Future work may consider the integration of visual saliency information and attention models into the proposed PC quality assessment framework to further improve the correlation performance.

REFERENCES

[1] S. Schwarz *et al.*, "Emerging MPEG standards for point cloud compression," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 133–148, Mar. 2019.
[2] L. Cui, R. Mekuria, M. Preda, and E. S. Jang, "Point-cloud compression: Moving picture experts group's new standard in 2020," *IEEE Consum. Electron. Mag.*, vol. 8, no. 4, pp. 17–21, Jul. 2019.
[3] E. S. Jang, M. Preda, K. Mammou, A. M. Tourapis, J. Kim, D. B. Graziosi, S. Rhyu, and M. Budagavi, "Video-based point-cloud-compression standard in MPEG: From evidence collection to committee draft [standards in a nutshell]," *IEEE Signal Process. Mag.*, vol. 36, no. 3, pp. 118–123, May 2019.
[4] E. Alexiou and T. Ebrahimi, "Impact of visualisation strategy for subjective quality assessment of point clouds," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, San Diego, CA, USA, Jul. 2018, pp. 1–6.
[5] X. Wu, Y. Zhang, C. Fan, J. Hou, and S. Kwong, "Subjective quality database and objective study of compressed point clouds with 6DoF head-mounted display," 2020, *arXiv:2008.02501*.
[6] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Point cloud rendering after coding: Impacts on subjective and objective quality," *IEEE Trans. Multimedia*, vol. 23, pp. 4049–4064, 2021.
[7] E. Zerman, C. Ozcinar, P. Gao, and A. Smolic, "Textured mesh vs coloured point cloud: A subjective study for volumetric video compression," in *Proc. 12th Int. Conf. Quality Multimedia Exp. (QoMEX)*, Athlone, Ireland, May 2020, pp. 1–6.
[8] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Subjective and objective quality evaluation of 3D point cloud denoising algorithms," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Hong Kong, Jul. 2017, pp. 1–6.
[9] S. Perry, H. P. Cong, L. A. da Silva Cruz, J. Prazeres, M. Pereira, A. Pinheiro, E. Dumic, E. Alexiou, and T. Ebrahimi, "Quality evaluation of static point clouds encoded using MPEG codecs," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Abu Dhabi, UAE, Oct. 2020, pp. 3428–3432.
[10] 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 Trans. Signal Inf. Process.*, vol. 8, no. 1, pp. 1–27, 2019.
[11] R. L. de Queiroz and P. A. Chou, "Motion-compensated compression of dynamic voxelized point clouds," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3886–3895, Aug. 2017.
[12] 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," *Proc. SPIE*, vol. 10752, Sep. 2018, Art. no. 107520I.
[13] R. Mekuria, Z. Li, C. Tulvan, and P. Chou, *Evaluation Criteria for PCC (Point Cloud Compression)*, Standard ISO/IEC MPEG N16332, Geneva, Switzerland, Jun. 2016.
[14] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Geometric distortion metrics for point cloud compression," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Beijing, China, Sep. 2017, pp. 3460–3464.
[15] *Common Test Conditions for Point Cloud Compression*, Standard ISO/IEC JTC1/SC29/WG11 N18474, 3DG Group, Geneva, Switzerland, Mar. 2019.
[16] E. Alexiou and T. Ebrahimi, "Point cloud quality assessment metric based on angular similarity," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, San Diego, CA, USA, Jul. 2018, pp. 1–6.
[17] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "A generalized Hausdorff distance based quality metric for point cloud geometry," in *Proc. 12th Int. Conf. Quality Multimedia Exp. (QoMEX)*, Athlone, Ireland, May 2020, pp. 1–6.
[18] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Mahalanobis based point to distribution metric for point cloud geometry quality evaluation," *IEEE Signal Process. Lett.*, vol. 27, pp. 1350–1354, 2020.
[19] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "A point-to-distribution joint geometry and color metric for point cloud quality assessment," in *Proc. IEEE 23rd Int. Workshop Multimedia Signal Process. (MMSP)*, Tampere, Finland, Oct. 2021, pp. 1–6.
[20] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, "Improving PSNR-based quality metrics performance for point cloud geometry," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Abu Dhabi, UAE, Oct. 2020, pp. 3438–3442.
[21] G. Meynet, J. Digne, and G. Lavoué, "PC-MSDM: A quality metric for 3D point clouds," in *Proc. 11th Int. Conf. Quality Multimedia Exp. (QoMEX)*, Berlin, Germany, Jun. 2019, pp. 1–3.
[22] I. Viola, S. Subramanyam, and P. Cesar, "A color-based objective quality metric for point cloud contents," in *Proc. 12th Int. Conf. Quality Multimedia Exp. (QoMEX)*, Athlone, Ireland, May 2020, pp. 1–6.
[23] R. Diniz, P. G. Freitas, and M. C. Q. Farias, "Towards a point cloud quality assessment model using local binary patterns," in *Proc. IEEE Int. Conf. Quality Multimedia Exp. (QoMEX)*, Athlone, Ireland, May 2020, pp. 1–6.
[24] I. Vajda, "On the F-divergence and singularity of probability measures," *Periodica Math. Hungarica*, vol. 2, nos. 1–4, pp. 223–234, Mar. 1972.
[25] R. Diniz, P. G. Freitas, and M. C. Q. Farias, "Multi-distance point cloud quality assessment," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Abu Dhabi, UAE, Oct. 2020, pp. 3443–3447.
[26] R. Diniz, P. G. Freitas, and M. C. Q. Farias, "Local luminance patterns for point cloud quality assessment," in *Proc. IEEE 22nd Int. Workshop Multimedia Signal Process. (MMSP)*, Tampere, Finland, Sep. 2020, pp. 1–6.
[27] G. Meynet, Y. Nehmé, J. Digne, and G. Lavoué, "PCQM: A full-reference quality metric for colored 3D point clouds," in *Proc. 12th Int. Conf. Quality Multimedia Exp. (QoMEX)*, Athlone, Ireland, May 2020, pp. 1–6.
[28] I. Viola and P. Cesar, "A reduced reference metric for visual quality evaluation of point cloud contents," *IEEE Signal Process. Lett.*, vol. 27, pp. 1660–1664, 2020.
[29] E. Alexiou and T. Ebrahimi, "Towards a point cloud structural similarity metric," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, London, U.K., Jul. 2020, pp. 1–6.
[30] Q. Yang, Z. Ma, Y. Xu, Z. Li, and J. Sun, "Inferring point cloud quality via graph similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 6, pp. 3015–3029, Jun. 2022.
[31] R. Diniz, P. G. Freitas, and M. C. Q. Farias, "Point cloud quality assessment based on geometry-aware texture descriptors," *Comput. Graph.*, vol. 103, pp. 31–44, Apr. 2022.
[32] 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," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 12, pp. 4645–4660, Dec. 2021.
[33] E. Alexiou and T. Ebrahimi, "Exploiting user interactivity in quality assessment of point cloud imaging," in *Proc. IEEE Int. Conf. Quality Multimedia Exp. (QoMEX)*, Berlin, Germany, Jun. 2019, pp. 1–6.

[34] X. Wu, Y. Zhang, C. Fan, J. Hou, and S. Kwong, "Subjective quality database and objective study of compressed point clouds with 6DoF head-mounted display," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 12, pp. 4630–4644, Dec. 2021.
[35] Q. Liu, H. Su, Z. Duanmu, W. Liu, and Z. Wang, "Perceptual quality assessment of colored 3D point clouds," *IEEE Trans. Vis. Comput. Graphics*, early access, Apr. 13, 2022, doi: 10.1109/TVCG.2022.3167151.
[36] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.
[37] 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 Trans. Signal Inf. Process.*, vol. 8, no. 1, pp. 1–27, 2019.
[38] M. Bertalmio, A. L. Bertozzi, and G. Sapiro, "Navier–Stokes, fluid dynamics, and image and video inpainting," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Kauai, HI, USA, Dec. 2001, p. 1.
[39] MMSPG and EPFL. *M-PCCD: MPEG Point Cloud Compression Dataset*. Accessed: Aug. 2022. [Online]. Available: https://www.epfl.ch/labs/mmspg/downloads/quality-assessment-for-point-cloud-compression/
[40] 3DG Group. *MPEG Point Cloud Repository*. Accessed: Aug. 2022. [Online].Available:http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/
[41] JPEG Committee. *JPEG Pleno Database*. Accessed: Aug. 2022. [Online]. Available:http://plenodb.jpeg.org/
[42] *Sketchfab*. Accessed: Aug. 2022. [Online]. Available:https://sketchfab.com/
[43] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, document ITU-R BT.500-13, International Telecommunication Union, Geneva, Switzerland, 2012.
[44] UBI. *ICIP 2020 Dataset*. Accessed: Aug. 2022. [Online]. Available: http://emergimg.di.ubi.pt/icip2020PC.html
[45] SIAT Video Team. *SIAT-PCQD*. Accessed: Aug. 2022. [Online]. Available: http://codec.siat.ac.cn/video_download_siat-pcqd.html
[46] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli, "New full-reference quality metrics based on HVS," in *Proc. Workshop Video Process. Quality Metrics Consumer Electron.*, Scottsdale, AZ, USA, Jan. 2006, pp. 1–4.
[47] N. Ponomarenko, F. Silvestri, K. Egiaz, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions," in *Proc. Int. Workshop on Video Process. Quality Metrics*, Scottsdale, AZ, USA, Jan. 2007.
[48] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[49] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2003, pp. 1398–1402.
[50] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
[51] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
[52] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Aug. 2014.
[53] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 586–595.
[54] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Image quality assessment: Unifying structure and texture similarity," Apr. 2020, *arXiv:2004.07728*.
[55] R. Reisenhofer, S. Bosse, G. Kutyniok, and T. Wiegand, "A Haar wavelet-based perceptual similarity index for image quality assessment," *Signal Process., Image Commun.*, vol. 61, pp. 33–43, Feb. 2018.

ALIREZA JAVAHERI (Member, IEEE) received the B.S. degree in computer engineering from Bu-Ali Sina University, the M.Sc. degree in artificial intelligence from the Sharif University of Technology, Iran, and the Ph.D. degree in electrical and computer engineering from the Instituto Superior Teìcnico (IST), Universidade de Lisboa. He was also a Researcher with the Instituto de Telecomunicações. He is currently a Postdoctoral Research Fellow with the ECE Department, Queen's University and Ingenuity Laboratories, Kingston, Canada. His research interests include point cloud compression and quality assessment, and computer vision.

CATARINA BRITES (Member, IEEE) received the E.E., M.Sc., and Ph.D. degrees in electrical and computer engineering from the Instituto Superior Teìcnico (IST), Universidade Teìcnica de Lisboa, Lisbon, Portugal, in 2003, 2005, and 2011, respectively. She is currently a Postdoctoral Researcher with the Multimedia Signal Processing Group, Instituto de Telecomunicações, IST. She has contributed with more than 70 international journals and conference papers. She has been an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING and serves as a technical program committee member for several journals and conferences in the multimedia signal processing field. Her current research interests include 2D/3D video processing and coding, plenoptic imaging representations, and machine learning.

FERNANDO PEREIRA (Fellow, IEEE) received the B.S., M.Sc., and Ph.D. degrees in electrical and computer engineering from the Instituto Superior Teìcnico (IST), Universidade de Lisboa, Lisbon, Portugal, in 1985, 1988, and 1991, respectively. He is currently with the Department of Electrical and Computer Engineering, IST, and Instituto de Telecomunicações. He is also the JPEG Requirements Chair and has contributed more than 300 papers in international journals, conferences, and workshops. His research interests include video analysis, coding, description and adaptation, and advanced multimedia services.

JOÃO ASCENSO (Senior Member, IEEE) received the E.E., M.Sc., and Ph.D. degrees from the Instituto Superior Teìcnico (IST), Universidade Teìcnica de Lisboa, Lisbon, Portugal, in 1999, 2003, and 2010, respectively, all in electrical and computer engineering. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, IST, and a member of the Instituto de Telecomunicações. He has authored more than 100 papers in international conferences. His current research interests include visual coding, quality assessment, light-field and point cloud processing, and indexing and searching of multimedia content. He is an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING and IEEE TRANSACTIONS ON MULTIMEDIA.