Signal Processing: Image Communication 135 (2025) 117262

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

PointPCA+: A full-reference Point Cloud Quality Assessment metric with PCA-based features

Xuemei Zhou [a,c,*], Evangelos Alexiou [b], Irene Viola [a], Pablo Cesar [a,c]

[a] Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[b] TNO Netherlands Organization for Applied Scientific Research, The Hague, The Netherlands
[c] TU Delft, Delft, The Netherlands

ARTICLE INFO

MSC:
41A05
41A10
65D05
65D17

Keywords:
Point cloud
Perceptual Quality Assessment
PCA
Full-reference
Feature selection
Random forest

ABSTRACT

This paper introduces an enhanced Point Cloud Quality Assessment (PCQA) metric, termed PointPCA+, as an extension of PointPCA, with a focus on computational simplicity and feature richness. PointPCA+ refines the original PCA-based descriptors by employing Principal Component Analysis (PCA) solely on geometry data; additionally, the texture descriptors are refined through a direct application of the function on YCbCr values, enhancing the efficiency of computation. The metric combines geometry and texture features, capturing local shape and appearance properties, through a learning-based fusion to generate a total quality score. Prior to fusion, a feature selection module is incorporated to identify the most effective features from a proposed super-set. Experimental results demonstrate the high predictive performance of PointPCA+ against subjective ground truth scores obtained from four publicly available datasets. The metric consistently outperforms state-of-the-art solutions, offering valuable insights into the design of similarity measurements and the effectiveness of handcrafted features across various distortion types. The code of the proposed metric is available at https://github.com/cwi-dis/pointpca_suite/.

1. Introduction

Point cloud is prevailing among the available 3D imaging formats in recent years. It is essentially a collection of points, where each point has attributes of geometry, color, reflectance, etc. However, through acquisition, compression, transmission, and rendering, the quality of a point cloud can be degraded, which necessitates effective and efficient Point Cloud Quality Assessment (PCQA) metrics. These metrics provide a guide on the design, optimization, and parameter tuning of point cloud processing pipelines. PCQA metrics have been extensively utilized in various applications, including benchmarking visual tasks such as restoration [1,2], compression [3–5], super-resolution [6,7], as well as for quality monitoring in various systems [8–11].

Based on the availability of reference point cloud data, PCQA can be generally divided into three categories, namely Full Reference (FR), Reduced Reference (RR) and No Reference (NR). Compared with RR and NR, FR requires the fully available reference during the execution, which mimics the way the Human Vision System (HVS) perceives quality as a comparison of differences between a degraded and a pristine version [12]. From a methodology standpoint, PCQA is composed of two main components: feature extraction and feature regression [13–16]. Researchers have utilized both hand-crafted and learning-based features, usually combined with a non-linear function or learning-based regressor [17,18]. End-to-end schemes can significantly improve the prediction performance by fitting the ground truth well. However, the HVS mechanisms behind them are difficult to explain [19] . In addition, one of the most challenging issues of end-to-end learning-based approaches is their requirement for large amounts of labeled data for training [20–22]. The learning model may have difficulties in handling various contents and distortions if the training set is not sufficiently large or fails to adequately represent real-world contents [23]. Besides, the shortage of data may probably cause serious over-fitting problems. In fact, only very size-limited, distortion-type unbalanced point cloud datasets are available in the PCQA field.

In previous research, the perceptual quality of point clouds under the FR framework has been explored using multiple geometric features, textural features, or a combination of both. However, these studies typically compute geometric or textural similarity measurements from the referenced and distorted point clouds, independently [24–26]. This is achieved by identifying correspondences among points based on geometry. Considering the aforementioned limitations and drawing inspiration from PointPCA [25], our emphasis in this paper is on non-deep-learning-based FR PCQA metrics. To this aim, we describe

* Corresponding author.
E-mail address: xuemei.zhou@cwi.nl (X. Zhou).

https://doi.org/10.1016/j.image.2025.117262
Received 30 December 2023; Received in revised form 16 September 2024; Accepted 9 January 2025
Available online 17 January 2025

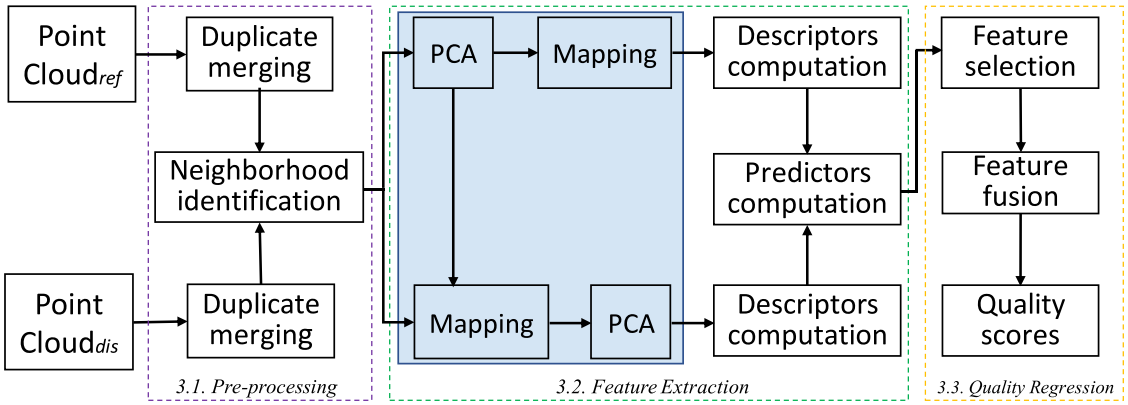


Fig. 1. PointPCA+ architecture: both the reference and the distorted point cloud are passing from every stage to compute a quality score. Operations in the blue box are applied only to the geometry data of point clouds. Only the reference point cloud serves as a reference to identify neighborhoods.

PointPCA+ and extend our prior analysis [27], incorporating results and discussions that take into account distortion types and feature importance. The contributions of the paper are threefold:

- We extend the PointPCA framework by performing PCA on the geometry data of the reference point cloud and transforming both the reference and distorted point clouds onto the new basis. This way we can capture differences in their shape properties effectively.
- We utilize *knn* algorithm to determine the neighborhood, which is faster and returns a consistent number of points, therefore further decreasing the computational cost of subsequent processing steps.
- We perform extensive experimentation on four publicly available datasets, demonstrating that PointPCA+ consistently achieves superior performance across four distinct datasets. A thorough analysis investigates the effectiveness of different handcrafted features concerning specific distortion types, aiming to discern which features are more impactful for different distortion type categories.

2. Related works

2.1. Ground-truth labeling and subjective dataset construction

In the development of objective PCQA metrics, subjective quality assessment datasets are the basis for their design and validation. Ground-truth ratings for visual impairments in stimuli, namely Mean Opinion Score (MOS) or Differential MOS (DMOS), are obtained through subjective quality assessment experiments [28,29]. These scores primarily indicate the degree of degradation affecting the contents with little information about the representations of distortion. Existing synthesized PCQA datasets, generated either by performing a test in a controlled laboratory environment or by mimicking it with point cloud processing algorithms, are size-limited and distortion-type and distortion-level unbalanced, often by design [30–33]. For example, when using test methods such as absolute category rating, the distortion level for a certain distortion type must be visually distinguishable to obtain meaningful MOS values and to avoid fatiguing subjects. Regarding the compression distortion of PCQA, recent research indicates a monotonic relationship with bit rate when evaluating compression quality [11,34]. However, it is crucial to emphasize that subjective quality evaluation sometimes does not follow a monotonic behavior. This is particularly evident in cases where the balance between geometry and color quality in point cloud content is challenging to establish [35]. This complexity adds difficulty to discerning between different point cloud coding modules. Therefore, it becomes imperative to integrate both geometry and texture distortions in assessing the perceptual quality of point clouds for PCQA. The utilization of effective geometry- and texture-related features can enhance the optimization of various algorithms associated with perception.

2.2. Objective quality assessment of point cloud

Objective PCQA metrics can be divided into point-based, projection-based, and feature-based models based on the way to process the point clouds. Point-based metrics such as point-to-point, point-to-plane, and their variants [36–39] measure degradations between the original and distorted point clouds per point, mainly based on Euclidean or color space distances [40]. Alexiou et al. [41] propose the angular similarity of tangent planes among corresponding points, which considers neighborhood information. These metrics are computationally efficient but suffer from a crude correspondence of matching between points.

Projection-based approaches adapt existing Image Quality Assessment (IQA) metrics to PCQA. The first attempt is reported in [42], which describes a framework for predicting point cloud quality by employing 2D IQA metrics on 6 orthographic projected views. This methodology was further extended in [43] by assigning non-uniform weighting on each view, the weights of a stimulus are computed as the ratio of the duration of inspection on corresponding views, divided by the total time. Liu et al. [44] provide a PCQA model based on the principle of information content weighted structural similarity (IW-SSIM) [45]. Icosphere and a series of transformations are employed to generate viewpoints. PQA-Net [46] takes 6 orthographic projections of point clouds as inputs, features are extracted after convolution neural network blocks, and they share a distortion identification and a quality prediction module that assist in obtaining final quality scores. PQA-Net adopts a two-step strategy to train the multi-task neural network. However, the projection process and the number of viewpoints have a non-negligible impact on the final prediction accuracy; besides, how to combine the quality score on each viewpoint into a score is also not straightforward.

Feature-based metrics consider perceptual loss from both geometry and texture properties. Viola et al. [47] extract the color statistics, histogram, and correlogram to assess the level of impairment and combine the color-based metrics with geometry-based metrics to form a global quality score. Alexiou et al. [26] employ the local distributions of point clouds to predict perceptual degradations from topology and color. Yang et al. [48] construct a local graph centered at resampled key points for both reference and distorted point clouds, with the color gradient on the local graph being used to measure distortions. Diniz et al. [49] introduced a texture descriptor based on perceptual color distance patterns, which is scale and rotation invariant [50]. Meynet et al. [24] utilize an optimally-weighted linear combination of curvature-based and color-based features to evaluate visual quality. Diniz et al. [51] adopt the statistical information of the extracted geometry/color features and feed them into a regression model.

Deep learning-based modules have also been used to extract perceptual features. An extension using coarse-to-fine progressive knowledge transfer based on HVS is given in [52]. Zhang et al. [53] make use

of multi-modal information to address the PCQA problem; the quality-aware encoder features are optimized with the assistance of symmetric cross-modality attention. They first split point clouds into two sub-models, with a point cloud encoder to extract the geometry feature, and an image encoder to obtain the texture feature. Subsequently, the quality-aware encoder features are optimized with the assistance of symmetric cross-modality attention. Zhang et al. [54] utilize natural scene statistics and entropy on the quality-related geometry and color feature domains, which are projected from 3D space, and support vector regression is employed to obtain quality scores. IT-PCQA [55] utilizes the rich prior knowledge in images and builds a bridge between 2D and 3D perception in the field of quality assessment. Specifically, a hierarchical feature encoder and a conditional discriminative network is proposed to extract effective latent features and minimize the domain discrepancy. pmBQA [56] proposes a projection-based blind quality indicator via multimodal learning by using four homogeneous modalities (i.e., texture, normal, depth and roughness). Interested readers may refer to [40] for a more comprehensive review of the literature.

In summary, point-based schemes may neglect the high-dimensional properties of point clouds and the interplay among these dimensions, thereby limiting their effectiveness. Projection-based methods often rely on 2D IQA, which may not adequately capture the intrinsic characteristics of point clouds. Feature-based schemes tend to have a high level of complexity, while the interpretability of deep learning-based methods is a drawback with training requiring a huge amount of data.

3. Proposed method

In Fig. 1, the PointPCA+ framework is illustrated, which is split into three modules, namely, (a) pre-processing, (b) feature extraction, and (c) quality regression which are introduced in the following subsections. Note that a FR metric typically uses either the pristine or the impaired content as a reference, or both. In our metric design, only the pristine point cloud serves as a reference.

3.1. Pre-processing

To ensure coherent geometry and color information without redundancies, points with identical coordinates that belong to the same point cloud are merged [37]. The color of a merged point is obtained by averaging the color of corresponding points sharing the same coordinates. For an FR PCQA metric, identifying matches between reference and distorted point clouds is crucial for comparing corresponding local properties. In our method, we use the *knn* algorithm to identify neighborhood pairs between two point clouds. In particular, for each point that belongs to a reference point cloud $\mathcal{A}$, we find its N nearest reference points, and its N nearest distorted points from the distorted point cloud $\mathcal{B}$, in terms of Euclidean distance.

3.2. Feature extraction

To capture local perceptual quality degradations of a distorted point cloud, we compute geometry and texture descriptors based on the identified neighborhoods. Statistics based on these descriptors are subsequently calculated and serve as predictors of visual quality. Features are finally obtained via pooling over these predictors. As mentioned earlier, our method uses only the pristine point cloud as a reference to find the matches in the distorted point cloud.

Table 1
Definition of descriptors.

	Descriptor	Definition	Distance
Geometric	Error vector	$\boldsymbol{e} = (\boldsymbol{\omega}_i^{\mathcal{B}} - \boldsymbol{\omega}_i^{\mathcal{A}})$	r_α
	Error along axes	$\epsilon_m = (\boldsymbol{\omega}_i^{\mathcal{B}} - \boldsymbol{\omega}_i^{\mathcal{A}})^T \cdot \mathbf{u}_m$	r_β
	Error from origin	$\boldsymbol{\epsilon}^{\mathcal{F}} = \boldsymbol{\omega}_i^{\mathcal{F}}$	r_α, r_β
	Mean	$\boldsymbol{\mu}^{\mathcal{B}} = \frac{1}{N}\sum_n \boldsymbol{\omega}_n^{\mathcal{B}}$	r_α, r_β
	Variance	$\lambda^{\mathcal{F}} = \frac{1}{N}\sum_n \left(\boldsymbol{\omega}_n^{\mathcal{F}} - \boldsymbol{\mu}^{\mathcal{F}}\right)^2$	r_δ
	Sum of variance	$\Sigma^{\mathcal{F}} = \sum_m \lambda_m^{\mathcal{F}}$	r_δ
	Covariance	$\boldsymbol{\Sigma} = \frac{1}{N}\sum_n \left(\boldsymbol{\omega}_n^{\mathcal{A}} - \boldsymbol{\mu}^{\mathcal{A}}\right) \cdot \left(\boldsymbol{\omega}_n^{\mathcal{B}} - \boldsymbol{\mu}^{\mathcal{B}}\right)^{\mathrm{T}}$	r_γ
	Omnivariance	$\mathrm{O}^{\mathcal{F}} = \sqrt[3]{\prod_m \lambda_m^{\mathcal{F}}}$	r_λ
	Eigenentropy	$\mathrm{E}^{\mathcal{F}} = -\sum_m \lambda_m^{\mathcal{F}} \cdot \log \lambda_m^{\mathcal{F}}$	r_δ
	Anisotropy	$\mathrm{A}^{\mathcal{F}} = (\lambda_1^{\mathcal{F}} - \lambda_3^{\mathcal{F}})/\lambda_1^{\mathcal{F}}$	r_δ
	Planarity	$\mathrm{P}^{\mathcal{F}} = (\lambda_2^{\mathcal{F}} - \lambda_3^{\mathcal{F}})/\lambda_1^{\mathcal{F}}$	r_δ
	Linearity	$\mathrm{L}^{\mathcal{F}} = (\lambda_1^{\mathcal{F}} - \lambda_2^{\mathcal{F}})/\lambda_1^{\mathcal{F}}$	r_δ
	Scattering	$\mathrm{S}^{\mathcal{F}} = \lambda_3^{\mathcal{F}}/\lambda_1^{\mathcal{F}}$	r_δ
	Change of curvature	$\mathrm{C}^{\mathcal{F}} = \lambda_3^{\mathcal{F}}/\sum_m \lambda_m^{\mathcal{F}}$	r_δ
	Parallelity	$\mathscr{P}_m = 1 - \mathbf{u}_m \cdot \mathbf{v}_m^{\mathcal{B}}$	–
	Angular similarity	$\theta = 1 - \frac{2\cdot\arccos(\cos(\mathbf{u}_m, \mathbf{v}_m^{\mathcal{B}}))}{\pi}$	–
Textural	Mean	$\tilde{\boldsymbol{\mu}}^{\mathcal{F}} = \frac{1}{N}\sum_n \boldsymbol{p}_n^{t,\mathcal{F}}$	r_δ
	Variance	$\tilde{s}^{\mathcal{F}} = \frac{1}{N}\sum_n \left(\boldsymbol{p}_n^{t,\mathcal{F}} - \tilde{\boldsymbol{\mu}}^{\mathcal{F}}\right)^2$	r_δ
	Sum of variance	$\tilde{\Sigma}^{\mathcal{F}} = \sum_m \tilde{s}_m^{\mathcal{F}}$	r_δ
	Covariance	$\tilde{\boldsymbol{\Sigma}} = \frac{1}{N}\sum_n \left(\boldsymbol{p}_n^{t,\mathcal{A}} - \tilde{\boldsymbol{\mu}}^{\mathcal{A}}\right) \cdot \left(\boldsymbol{p}_n^{t,\mathcal{B}} - \tilde{\boldsymbol{\mu}}^{\mathcal{B}}\right)^{\mathrm{T}}$	r_γ
	Omnivariance	$\tilde{\mathrm{O}}^{\mathcal{F}} = \sqrt[3]{\prod_m \tilde{s}_m^{\mathcal{F}}}$	r_δ
	Entropy	$\tilde{\mathrm{H}}^{\mathcal{F}} = -\sum_m \tilde{s}_m^{\mathcal{F}} \cdot \log \tilde{s}_m^{\mathcal{F}}$	r_δ

3.2.1. Geometry descriptors

Given a query point $\mathbf{p}_i$ of $\mathcal{A}$, the subscript i denotes the point index, $1 \le i \le |\mathcal{A}|$, and $|\mathcal{A}|$ is the cardinality. The coordinates of $\mathbf{p}_i$'s N nearest neighbors in $\mathcal{F}$ are indicated as $\boldsymbol{p}_n^{g,\mathcal{F}} = (x_n, y_n, z_n)^{\mathrm{T}}$, with $1 \le n \le N$ and $\mathcal{F} \in \{\mathcal{A}, \mathcal{B}\}$. The geometry of $\mathbf{p}_i$ is denoted as $\boldsymbol{p}_i^{g,\mathcal{A}}$, and the geometry of its closest neighbor in $\mathcal{B}$ is denoted as $\boldsymbol{p}_i^{g,\mathcal{B}}$. Initially, the covariance matrix $\boldsymbol{\Sigma}_i^{\mathcal{A}}$ is computed as

$$\boldsymbol{\Sigma}_i^{\mathcal{A}} = \frac{1}{N}\sum_{n=1}^{N} \left(\boldsymbol{p}_n^{g,\mathcal{A}} - \bar{\boldsymbol{p}}_i^{g,\mathcal{A}}\right) \cdot \left(\boldsymbol{p}_n^{g,\mathcal{A}} - \bar{\boldsymbol{p}}_i^{g,\mathcal{A}}\right)^{\mathrm{T}}, \tag{1}$$

where $\bar{\boldsymbol{p}}_i^{g,\mathcal{A}}$ indicates the centroid, given as

$$\bar{\boldsymbol{p}}_i^{g,\mathcal{A}} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{p}_n^{g,\mathcal{A}}. \tag{2}$$

Then, eigen-decomposition is applied to $\boldsymbol{\Sigma}_i^{\mathcal{A}}$, to obtain the eigenvectors which form an orthonormal basis $\mathbf{V}^{\mathcal{A}}$ composed of eigenvectors $\mathbf{v}_m^{\mathcal{A}}$, where $m = 1, 2, 3$. Next, we map the reference and distorted neighborhoods to the new orthonormal basis, denoted as $\boldsymbol{\omega}_n^{\mathcal{F}} = (\boldsymbol{p}_n^{g,\mathcal{F}} - \bar{\boldsymbol{p}}_i^{g,\mathcal{A}}) \cdot \mathbf{V}^{\mathcal{A}}$. Finally, we apply PCA to the covariance matrix of $\boldsymbol{\omega}_n^{\mathcal{B}}$ and compute the eigenvectors $\mathbf{v}_m^{\mathcal{B}}$. This process is visually demonstrated in Fig. 2, showcasing the distinction between the two bases.

The merit inherent in projecting the geometry of both the reference and distorted point clouds onto a shared orthonormal basis, established by the reference point cloud, lies in the capacity to unify the representation of geometry degradation within a common space. This enables a more precise measurement of geometry similarity within the framework of the FR PCQA paradigm.

The mapped coordinates of the reference and distorted points $\boldsymbol{\omega}_n^{\mathcal{F}}$, the eigenvectors $\mathbf{v}_m^{\mathcal{F}}$ and the unit vectors $\mathbf{u}_m$, with $\mathbf{u}_1 = [1,0,0]^{\mathrm{T}}$, $\mathbf{u}_2 = [0,1,0]^{\mathrm{T}}$ and $\mathbf{u}_3 = [0,0,1]^{\mathrm{T}}$, are used to construct the geometric descriptors defined in Table 1.

3.2.2. Texture descriptors

The color space is first converted from RGB to YCbCr [57]. This conversion is motivated by the fact that the human eye is more sensitive to changes in brightness than changes in color according to HVS. We denote the texture information of $\mathbf{p}_i$'s N nearest neighbors in $\mathcal{F}$ as $\boldsymbol{p}_n^{t,\mathcal{F}} = (Y_n, Cb_n, Cr_n)^{\mathrm{T}}$. The proposed 6 texture descriptors are defined in Table 1.

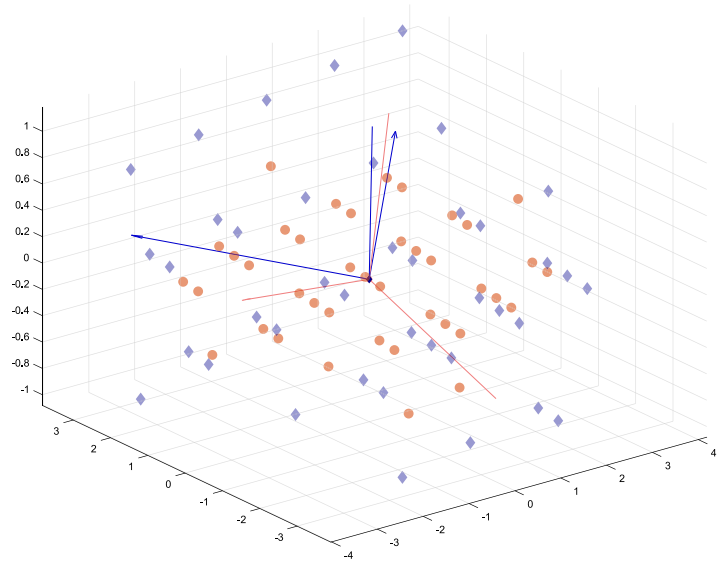


Fig. 2. The orthonormal basis formed by both reference (*longdress*) and distorted (*longdress_Octree-Lifting_R04*) point clouds. The orange color represents the geometry around one point (2130th point) of the reference point cloud and the corresponding basis after PCA operation; the purple color represents the geometry around the matched point of the distorted point cloud and the corresponding basis after PCA operation.

(a) *longdress* (b) $\mu(\mathrm{L}^{\mathcal{A}})$: PointPCA (c) $\mu(\mathrm{L}^{\mathcal{A}})$: PointPCA+ (d) $\mu(\mathrm{P}^{\mathcal{A}})$: PointPCA (e) $\mu(\mathrm{P}^{\mathcal{A}})$: PointPCA+

Fig. 3. The point cloud *longdress* and statistical features using mean of linearity (Figs. 3(b)–3(c)), planarity (Figs. 3(d)–3(e)). The amplitudes of statistical features are color-mapped, with red indicating higher and blue lower values. It can be noticed that the mean of linearity 3(b) and planarity 3(d) of PointPCA/PointPCA+ capture high- and low-frequency geometric regions, respectively. Additionally, PointPCA+ has lower complexity as the neighborhood is determined through *knn*.

3.2.3. Explanation of descriptors

Each geometry descriptor represents an interpretable shape property inside the neighborhood. Specifically, $\boldsymbol{e}$ denotes the error vector between the mapped coordinates of the reference query point and its nearest neighbor, and ϵ_m is the projected distance of the error vector across the mth axis. The $\boldsymbol{\epsilon}$ is used to capture the Euclidean and projected distances of the mapped reference query point or its nearest distorted neighbor from the centroid and principal axes, respectively. $\boldsymbol{\mu}^{\mathcal{B}}$, $\lambda^{\mathcal{F}}$, $\Sigma^{\mathcal{F}}$ and $\boldsymbol{\Sigma}$ reveal local statistics. $\mathrm{E}^{\mathcal{F}}$ provides an estimation of the space uncertainty on the projected surfaces. Additionally, $\mathscr{P}_m$ and θ_m assess the parallelity and the angular dispersion of the distorted plane. The remaining geometry descriptors explore the topology of a local region from different aspects, relying on the spatial dispersion along different principal axes. $\tilde{\boldsymbol{\mu}}^{\mathcal{F}}$, $\tilde{\mathbf{s}}^{\mathcal{F}}$ and $\tilde{\Sigma}^{\mathcal{F}}$ of the YCbCr channel express the intrinsic distribution of luminance and chromatic components. $\tilde{\boldsymbol{\Sigma}}$ and $\tilde{\mathrm{O}}^{\mathcal{F}}$ show the variability of color information. $\tilde{\mathbf{H}}^{\mathcal{F}}$ provides an estimation of color uncertainty of the local region. Every descriptor is computed per point $\mathbf{p}_i$.

(a) *longdress-Octree-Lifting-R02* (b) $\mathrm{L}^{\mathcal{B}}$ (c) $\mathrm{P}^{\mathcal{B}}$

Fig. 4. The predictors of linearity (Fig. 4(b)) and planarity (Fig. 4(c)) of point cloud *longdress-Octree-Lifting-R02* (*Fig. 4(a)*). The amplitudes of predictors are color-mapped, with red indicating higher and blue lower values. Notably, the relative differences in linearity and planarity of PointPCA+ effectively highlight regions that can form linear structures and detect bulges, such as wrinkles in clothing, in the *longdress*.

3.2.4. Predictors

Predictors are defined as the error samples obtained by computing a distance over descriptor values. We define different distance functions for different descriptors. We use the Euclidean distance to measure the point-to-point distances between query point pairs under the new basis

$$r_\alpha = \sqrt{\sum_m \boldsymbol{d_1}^2}, \tag{3}$$

where $\boldsymbol{d_1}$ is the difference between two points. We use the absolute value to measure the point-to-plane distance, as

$$r_\beta = |\boldsymbol{d_2}|, \tag{4}$$

where $\boldsymbol{d_2}$ indicates the projected distance between a point and the reference axes. We use the following definition of relative difference for the covariance features

$$r_\gamma = \frac{|\boldsymbol{q}^{\mathcal{A}} \odot \boldsymbol{q}^{\mathcal{B}} - \boldsymbol{Q}|}{\boldsymbol{q}^{\mathcal{A}} \odot \boldsymbol{q}^{\mathcal{B}}}, \tag{5}$$

where $\{\boldsymbol{q}^T = \lambda^T, \mathbf{Q} = \boldsymbol{\Sigma}\}$ and $\{\boldsymbol{q}^T = \tilde{s}^T, \mathbf{Q} = \tilde{\boldsymbol{\Sigma}}\}$, for geometry and texture attributes, respectively, $\odot$ is for element-wise product. We use the relative difference formula [26], for the remaining descriptors

$$r_\delta = \frac{|\phi^{\mathcal{A}} - \phi^{\mathcal{B}}|}{|\phi^{\mathcal{A}}| + |\phi^{\mathcal{B}}| + \varepsilon}, \tag{6}$$

where ε is a small constant to avoid undefined operations. Finally, the definitions of parallelity and angular similarity descriptors incorporate a distance function. For notational purposes only, we define distances r_ρ and r_θ to be identical to the definitions of $\mathscr{P}_m$ and θ_m, respectively. Table 1 enlists distance function(s) used per descriptor.

3.2.5. Features

Features are defined by pooling over predictor values. Specifically, predictors $\psi_{i,j,k}$ are obtained per point $\mathbf{p}_i$, descriptor j, and distance function r_k, $k \in \{\alpha, \beta, \gamma, \delta, \rho, \theta\}$. This is done for all descriptors j in Table 1, using the corresponding distances r_k. Through pooling, we obtain a feature $f_{j,k}$ for every predictor:

$$f_{j,k} = \frac{1}{|\mathcal{A}|} \sum_{i=1}^{|\mathcal{A}|} \psi_{i,j,k}. \tag{7}$$

3.3. Quality regression

To obtain a quality score that is well-aligned with the HVS, the Recursive Feature Elimination (RFE) algorithm is used to select the most relevant predictor set among all the proposed predictors. RFE [58] improves model accuracy, and efficiency, and reduces overfitting. Machine learning-based regression models have been extensively used to tackle the quality regression problem in the domain of quality assessment, we then use the random forest algorithm to regress the selected predictors to the final quality score.

3.4. Differences with PointPCA

In comparison to PointPCA, our pre-processing methodology involves solely utilizing the pristine point cloud as a reference, dispensing with the need for both pristine and distorted point clouds. Regarding geometric features, we transform the xyz values of both the pristine and distorted point clouds into the basis formed by the pristine point cloud, eliminating the necessity for separate bases for distinct point cloud sets. For textural features, we directly apply statistical functions to the color values in the YCbCr space without resorting to PCA decomposition. The decision to forego PCA on texture stems from the lack of physical significance post-decomposition of YCbCr channel values, as observed in geometry. Simultaneously, this approach aids in reducing computational complexity. Additionally, PointPCA+ adopts distinct distance functions for various descriptors, eliminating constraints on the computation of mean and standard deviation values. We illustrate the disparities between PointPCA and PointPCA+ for a specific point cloud, *longdress*, in Fig. 3. Notably, we visualize the linearity and planarity geometric descriptors before comparison. From Fig. 3, the geometric dissimilarity measurements differ when using *knn* and *r-search* methods. For PointPCA, we employed the *r-search* method with $r = 0.008 \times B$, where B represents the maximum length of the bounding box of the reference point cloud. In contrast, for PointPCA+, we used the *knn* algorithm with $k = 81$. While for the statistical features μ, the *knn* algorithm with $k = 9$ was applied. Both the linearity and planarity are decreased compared with PointPCA. We also visualize the predictions corresponding to linearity and planarity in Fig. 4. We can see that the predictors exactly describe the line and the uneven region of point cloud *longdress*.

4. Experimental results

In this Section, we report the evaluation results of the proposed PointPCA+ metric under three public datasets in Section 4.2 with other 8 state-of-the-art metrics. Moreover, we report the performance achieved in the ICIP 2023 Point Cloud Visual Quality Assessment (PCVQA) grand challenge[1] in Section 4.3. Specifically, the challenge consists of 5 tracks, which correspond to different use cases in which quality metrics are typically used. The first two tracks aim to assess the perceptual fidelity of distorted contents with/without respect to the originals for any level of distortion, respectively. This is the most generic and traditional set-up for quality metrics. The next two tracks focus on metrics for high-end quality with/without access to the original content. These are desirable in applications such as content production, high-quality streaming, digital twins, etc. The last track should be sensible to quality differences within different processed versions of the same point cloud content, which is suitable for optimization scenarios. We participated in Track#1 FR broad-range quality estimation, Track#3 FR high-range quality estimation, and Track#5 FR intra-reference quality estimation. Additional analysis related to cross-dataset validation and feature importance is carried out across all the aforementioned datasets, as detailed in Section 4.4 and Section 4.5. These sections aim to demonstrate the generalizability of the proposed PointPCA+.

4.1. Setup

4.1.1. Datasets

Three publicly available datasets were recruited for performance evaluation, namely, M-PCCD, SJTU, and WPC. The M-PCCD [60] consists of 8 point clouds whose geometry and color are encoded using V-PCC and G-PCC variants, resulting in 232 distorted stimuli. Detailed distortion types include Octree-Lifting, Octree-RAHT, TriSoup-Lifting, TriSoup-RAHT and V-PCC. The contents in M-PCCD depict either human figures or objects. The SJTU [61] includes 9 reference point clouds with each point cloud corrupted by seven types of distortions under six levels, generating 378 distorted stimuli. Detailed distortion types include Octree-based compression, Color Noise (CN), Geometric Gaussian Noise (GGN), downsampling, and combinations of the CN, GGN and downsampling. SJTU includes 5 human body models and 4 inanimate objects. The WPC [44] contains 20 reference point clouds with each point cloud degraded under five types of distortions and different levels, leading to 740 distorted stimuli. Detailed distortion types include Octree-LPCC, TriSoup-SPCC, V-PCC, Gaussian noise and downsampling. WPC dataset only collects objects including snacks, fruits and vegetables, etc. The Broad Quality Assessment of Static Point Clouds (BASICS) [34] is used in the ICIP 2023 PCVQA grand challenge, and comprises 75 point clouds from 3 different semantic categories: (i) Humans & Animals, (ii) Inanimate Objects, and (iii) Buildings & Landscapes. Each point cloud is compressed with 3 compression methods from the MPEG standardization field, i.e., Octree-RAHT, Octree-Predlift and V-PCC; 1 learning-based algorithm, i.e., GeoCNN, at varying compression levels, resulting in 1494 processed point clouds. BASICS dataset is aimed to provide a foundation for research that supports the telepresence applications, in terms of compression and quality assessment. Table 3 summarized the characteristics of the four datasets.

[1] https://sites.google.com/view/icip2023-pcvqa-grand-challenge

Table 2
SROCC performance on M-PCCD, SJTU and WPC datasets.

Metric	PointPCA+	PointPCA [25]	PCQM [24]	PointSSIM [26]	BitDance [59]	Plane2Plane [41]	P2Plane_MSE [40]	P2P_MSE [40]	PSNR Y [40]
M-PCCD	**0.943**±0.022	0.941±0.032	0.940 ± 0.032	0.925 ± 0.024	0.859 ± 0.061	0.847 ± 0.076	0.901 ± 0.025	0.896 ± 0.042	0.798 ± 0.162
SJTU	0.865±0.064	**0.890**±0.056	0.862 ± 0.030	0.708 ± 0.070	0.748 ± 0.077	0.761 ± 0.039	0.578 ± 0.155	0.612 ± 0.157	0.743 ± 0.083
WPC	0.857±0.040	**0.866**±0.036	0.749 ± 0.036	0.465 ± 0.059	0.451 ± 0.054	0.454 ± 0.069	0.452 ± 0.065	0.563 ± 0.071	0.614 ± 0.061

Table 3
Overview of M-PCCD, SJTU, WPC and BASICS datasets.

Datasets	Contents		Distortion types		Distortion levels	Total
M-PCCD	Humans & Inanimate Objects	8	Octree-Lifting, Octree-RAHT, TriSoup-Lifting, TriSoup-RAHT, V-PCC	5	G-PCC: 6 V-PCC: 5	232
SJTU	Humans & Inanimate Objects	9	Octree-based compression, CN, GGN, downsampling, CN + GGN, CN + downsamplin, GGN + downsampling	7	6	378
WPC	Inanimate Objects	20	Octree-LPCC, TriSoup-SPCC, V-PCC, Gaussian noise, downsampling	5	Geometry: 3 Texture:1/3/4	740
BASICS	1. Humans & Animals 2. Inanimate Objects 3. Buildings & Landscapes	75	Octree-RAHT, Octree-Predlift, V-PCC, GeoCNN	4	G-PCC: 5 V-PCC: 6	1494

4.1.2. Evaluation metrics

For the PCVQA Grand Challenge, all submissions undergo testing on a designated test set curated by the organizers. The evaluation of performance relies on five standard criteria provided by the organizers: including Pearson Linear Correlation Coefficient (PLCC), Spearman Rank Order Correlation Coefficient (SROCC), Difference/Similar Analysis quantified by Area Under the Curve (D/S_{AUC}), Better/Worse Analysis quantified by Correct Classification percentage (B/W_{CC}) [62], and the Runtime Complexity (RC). Notably, no function is employed for score mapping. Additionally, for the assessment of performance metrics on three other commonly used datasets, the criteria include SROCC, PLCC, Kendall Rank Order Correlation Coefficient (KROCC), and Mean Square Error (MSE) are chosen. Higher values of PLCC, SROCC, and KROCC indicate better performance in terms of correlation with human opinion while lower RMSE indicates better consistency.

4.1.3. Implementation details

We use RFE to select the best feature set among all the predictors, with the best SROCC performance on the PCVQA grand challenge test set. In the inference stage, the default configuration of scikit-learn (version 1.2.2) in Python is used. Regarding the neighborhood size for the computation of descriptors, $K = 81$ is chosen considering complexity and performance, after experimenting with $K \in \{9, 25, 49, 81, 121\}$.

4.2. Performance evaluation on M-PCCD, SJTU and WPC

We compare PointPCA+ with existing FR point-based quality metrics, the results are shown in Table 2. The best performance among these metrics is highlighted in boldface, with the second best underlined. Specifically, each dataset is split into two partitions that contain 80% and 20% of the contents for training and testing, respectively, with all the distorted versions of a specific content placed in one partition. For M-PCCD, SJTU, and WPC, we use 6/2, 7/2, and 16/4 contents for training/testing, respectively. Then, a quality prediction model is trained on the training data and tested on the corresponding testing data of the same dataset, for within-dataset validation. this process is repeated for all possible 80%–20% splits of each dataset, leading to 28, 36, and 4845 testing partitions and an equal number of corresponding quality prediction models for M-PCCD, SJTU, and WPC respectively. Finally, the average and the standard deviation of SROCC index computed across all testing splits of each dataset, are reported. From Table 2 we can see that PCA-based metrics are competitive with the highest SROCC on the three datasets, especially the performance of PointPCA on WPC is increased by 15.62% in terms of SROCC though the performance of PointPCA+ is a slightly lower than PointPCA (0.866 VS 0.857).

Table 4
Track#1 (FR broad range): top 4 performance comparison on the official PCVQA grand challenge test set, evaluated by the challenge organizers. Best in bold and second best underlined. Our submission is ranked in 2nd place.

Submission	SROCC	PLCC	D/S_{AUC}	B/W_{CC}	RC(s)
KDDIUSCJoint	**0.875**	**0.917**	**0.888**	**0.970**	42.80
PointPCA+	0.874	0.909	0.871	0.961	1000.00
SJTU MMLAB	0.871	0.896	0.832	0.955	**8.60**
SlowHand	0.791	0.825	0.805	0.924	130.47

Table 5
Track#3 (FR high range): top 4 performance comparison on the official PCVQA grand challenge test set, evaluated by the challenge organizers. Best in bold and second best underlined. Our submission is ranked in 3rd place.

Submission	SROCC	PLCC	D/S_{AUC}	B/W_{CC}	RC(s)
SJTU MMLAB	**0.630**	**0.592**	**0.665**	**0.909**	**8.60**
KDDIUSCJoint	0.551	0.516	0.642	0.872	42.80
PointPCA+	0.603	0.479	0.625	0.886	1000.00
SlowHand	0.377	0.423	0.565	0.780	130.47

Table 6
Track#5 (FR intra-reference): top 4 performance comparison on the official PCVQA grand challenge test set, evaluated by the challenge organizers. Best in bold and second best underlined. Our submission is ranked in 3rd place.

Submission	D/S_{AUC}	B/W_{CC}	RC(s)
SJTU MMLAB	0.808	**0.947**	**8.60**
KDDIUSCJoint	**0.822**	0.933	42.80
PointPCA+	0.811	0.938	1000.00
SlowHand	0.753	0.854	130.47

4.3. Performance evaluation on BASICS

We split BASICS into training–validation-test with 60%-20%–20% following the rules from the PCVQA grand challenge [63]. Table 4 to Table 6 show the official evaluation results of Track#1, Track#3 and Track#5, respectively.

Referencing Tables 5–6, several notable observations emerge from the competition results across all three FR tracks. (1) Despite strong performances in Track 1 and Track 5, none of the teams attained satisfactory results in Track 3 of in PCVQA, highlighting the challenges associated with fine-grained PCQA. (2) Examining PointPCA+, it is evident that the extracted features within a neighborhood size of 81, combined with the application of statistical functions (e.g., mean and variance) on geometry, may fail to capture subtle differences between two point clouds. This highlights the limitations of using statistical features to capture subtle differences.

Table 7
Cross-dataset validation among M-PCCD, SJTU, WPC and BASICS datasets. Both the training and testing used all the content among the datasets. Best in bold.

	Test															
	M-PCCD				SJTU				WPC				BASICS			
Train	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
M-PCCD	–	–	–	–	**0.726**	**0.725**	**0.541**	3.148	0.500	0.469	0.328	4.003	0.847	**0.777**	**0.589**	**0.899**
SJTU	**0.855**	**0.881**	**0.708**	2.578	–	–	–	–	**0.578**	**0.608**	**0.442**	**2.575**	0.732	0.717	0.523	1.946
WPC	0.731	0.878	0.703	4.767	0.610	0.604	0.435	**2.049**	–	–	–	–	**0.848**	0.726	0.536	3.675
BASICS	0.832	0.880	0.692	**1.057**	0.579	0.645	0.472	2.810	0.500	0.490	0.347	3.202	–	–	–	–

4.4. Cross-dataset validation

To verify the generalization and robustness of the proposed PointPCA+, we conduct cross-dataset experiments among all 4 datasets. We train the model using the entire content of one dataset and then test it separately using the entire content of the other three datasets. The experimental results are shown in Table 7. From Table 7, we can draw the following observations:

1. PointPCA+ performs well in generalization and robustness, particularly when trained on the small M-PCCD dataset and tested on the large BASICS dataset. Combined with Table 2, the cross-dataset evaluation performance is even higher than certain FR PCQA metrics, for example BitDance [59], PSNR_Y [40], etc.
2. Compared with M-PCCD, SJTU and BASICS datasets, the WPC dataset has the worst performance among all the evaluation metrics. SROCC and PLCC of PointPCA+ on WPC trained on M-PCCD and BASICS achieve the same accuracy as random guessing, this may be because the contents in WPC only contain objects, and the distortion types of WPC are more complex compared with the other three datasets.
3. PointPCA+ shows better SROCC performance on the SJTU dataset compared to the WPC dataset. This is likely because SJTU shares specific human figures with the M-PCCD dataset and includes both human and object categories for M-PCCD and BASICS, while SJTU and WPC have no overlapping content.

In conclusion, the cross-dataset performance of PointPCA+ is promising but its generalization depends on dataset composition. Training on large, diverse datasets and testing on smaller ones normally yields better results, while the reverse leads to poor generalization. However, this cannot hold if there exists a domain shift (i.e., in our case, contents and distortion type) between the testing set and the training set.

4.5. Performance on individual distortion type

To further explore the effectiveness of the designed geometry and texture features for a specific distortion type, we test the performance per distortion type per dataset, with the results listed in Table 8. We can draw the following observations.

1. When assessing compression distortion, V-PCC emerges as the most challenging distortion type of which to evaluate perceptual quality, aligning with findings from a prior study [46]. Predicting the perceptual quality of compression distortions from G-PCC and learning-based methods proves to be more manageable, with Octree-Lifting exhibiting a slight advantage over Octree-RAHT on M-PCCD and BASICS datasets.
2. For CN and Gaussian noise distortions, CN has the poorest performance. However, prediction accuracy improves by 16.06% for CN+GGN in terms of SROCC on the SJTU dataset, indicating that geometry-related features help capture these distortions. Gaussian noise performs worst on WPC, as it affects both geometry and texture, creating a compounded distortion.
3. For downsampling distortion types, the performance of PointPCA+ is notably high on the SJTU dataset but relatively lower on the WPC dataset. This implies that downsampling on objects presents a challenge for the HVS to discern, as it may exert a masking effect on objects more prominently than on humans.

Table 8
Performance comparison of PointPCA+ metrics for different distortion types per dataset. DT refers to distortion type and NC denotes the number of contents.

Dataset	DT	NC	PLCC	SROCC	KROCC	RMSE
M-PCCD	Octree-Lifting	48	0.938	0.991	0.962	0.826
	Octree-RAHT	48	0.962	0.984	0.931	0.704
	TriSoup-Lifting	48	0.951	0.965	0.879	0.597
	TriSoup_RAHT	48	0.970	0.963	0.870	0.561
	V-PCC	40	0.823	0.815	0.674	1.266
SJTU	CN	54	0.621	0.741	0.545	2.594
	CN + GGN	54	0.894	0.860	0.697	0.825
	Downsampling	54	0.969	0.944	0.848	0.562
	Downsampling+CN	54	0.946	0.937	0.788	1.371
	Downsampling+GGN	54	0.981	0.965	0.879	0.736
	GGN	54	0.955	0.958	0.848	0.810
	Octree (PCL)	54	0.975	0.965	0.879	0.797
WPC	Downsampling	60	0.802	0.795	0.594	1.177
	Octree (LPCC)	80	0.916	0.881	0.708	0.850
	Trisoup (SPCC)	240	0.987	0.923	0.818	0.594
	Gaussian noise	180	0.628	0.650	0.417	1.826
	V-PCC	180	0.737	0.756	0.566	2.054
BASICS	GeoCNN	294	0.971	0.941	0.787	0.305
	Octree-Predlift	375	0.975	0.937	0.792	0.289
	Octree-RAHT	375	0.895	0.876	0.693	0.441
	V-PCC	450	0.787	0.690	0.514	0.371

In summary, PointPCA+ exhibits proficiency in predicting compression distortions. However, its effectiveness diminishes when confronted with distortion instances primarily manifesting in color values, as observed for CN distortions on the SJTU dataset. Additionally, when the equilibrium between geometric and color is disrupted, as exemplified by Gaussian noise on the WPC dataset, PointPCA+ struggles to accurately gauge the extent of degradation.

4.6. Feature importance of PointPCA+ for different distortion types per dataset

We evaluated the effectiveness of the 40 selected features after RFE. The model was trained on 80% of the dataset, and feature importance was calculated on the remaining 20% for each distortion type. This was done using the permutation importance technique [64] based on MSE. Permutation importance shows how crucial a feature is for a specific model, rather than its standalone predictive value, helping to assess the generalization ability of the features across different distortion types. The feature importance of PointPCA+ across the four datasets is shown in Fig. 5. The feature importance for each distortion type per dataset is illustrated in Fig. 6. Combining Fig. 5 and Fig. 6, we can draw the following conclusions:

1. The feature importance ranking varies across different datasets, primarily due to distinctions in content and distortion types among the four datasets. The variance on the z axis ($f^{z}_{\lambda,\delta}$) within the geometric features obtained the lowest importance ranking on both M-PCCD and BASICS, ranked 29th and 38th on SJTU and WPC datasets, respectively.
2. Geometric features consistently exhibit superior feature importance rankings when compared to textural features, as observed in the top-5 features across all four datasets. In M-PCCD, SJTU,

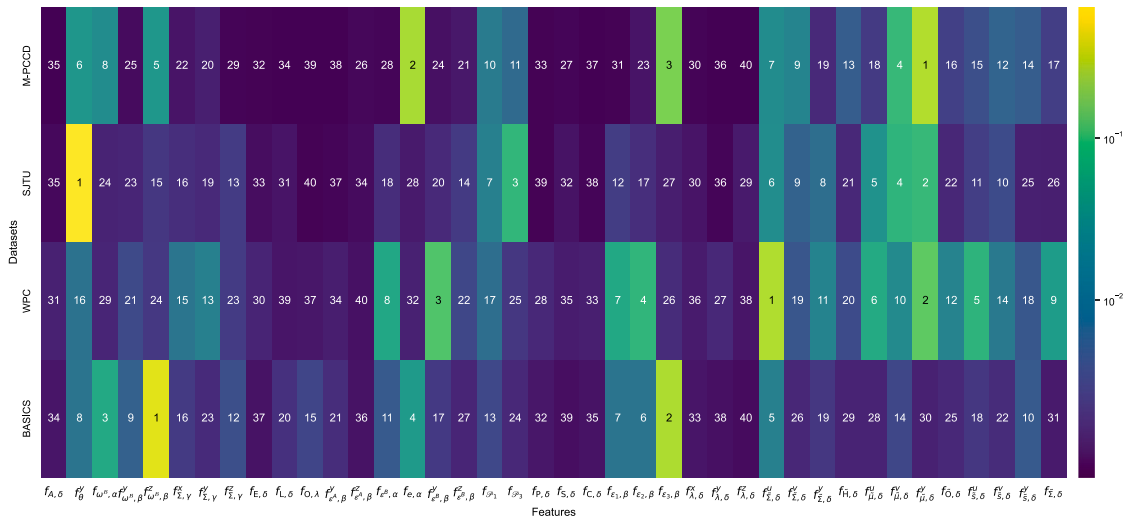
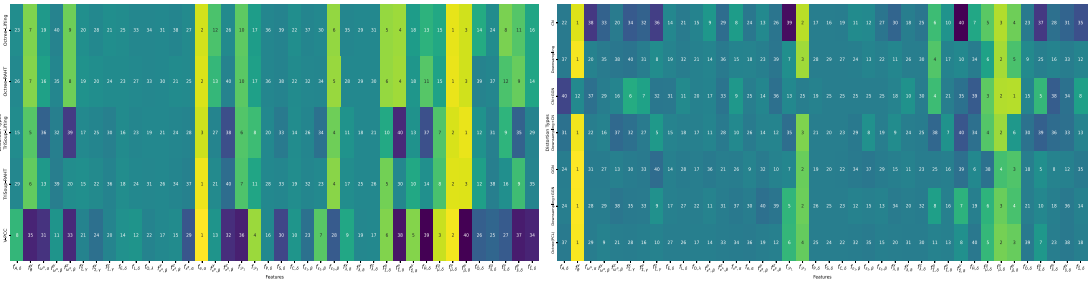
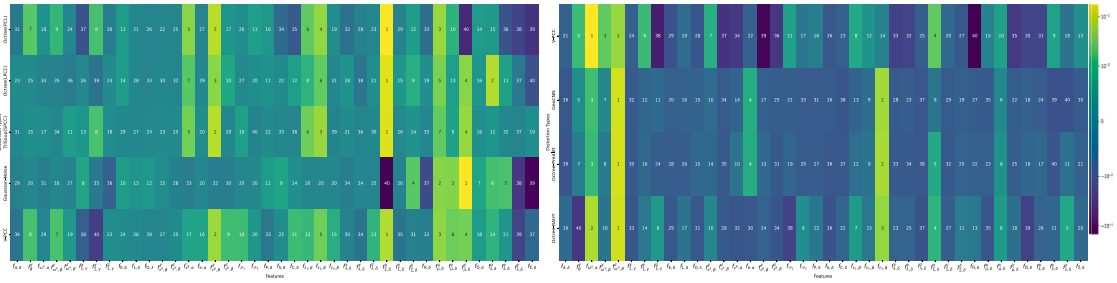


Fig. 5. The feature importance of all the extracted geometry and texture features within pointPCA+ metric for M-PCCD, SJTU, WPC and BASICS datasets. The numbers are obtained by ranking the 40 features based on the importance score per dataset.



(a) M-PCCD (b) SJTU



(c) WPC (d) BASICS

Fig. 6. Feature importance of PointPCA+ per distortion type for M-PCCD, SJTU, WPC, and BASICS datasets, separately. The numbers are obtained by ranking the 40 features based on the permutation feature importance score per distortion type.

and WPC, the top-3 rankings comprise a combination of texture and geometry features. However, in BASICS, only geometry features are represented in the top-3 ranking.

3. The point-to-point distance ($f_{\boldsymbol{e},\alpha}$) and the mean value of v channel ($f^{v}_{\tilde{\boldsymbol{\mu}},\delta}$) demonstrate strong performance across various distortion types in M-PCCD dataset. Notably, in TriSoup-Lifting, the mean value of y channel ($f^{y}_{\tilde{\boldsymbol{\mu}},\delta}$) excels but performs less optimally in the case of V-PCC distortion. In SJTU dataset, the cosine similarity of the y axis (f^{y}_{θ}) outperforms other features for all distortion types, except for CN+GGN distortion. Similarly, the covariance of the u channel ($f^{u}_{\tilde{\boldsymbol{\Sigma}}}$) excels for all distortion types, with the exception of Gaussian noise in WPC. Meanwhile, in BASICS, the projected distances of the distorted centroid from reference planes on the z axis ($f^{z}_{\boldsymbol{\omega}^{B},\beta}$) consistently yield high rankings across all distortion types.

4.7. Discussion

At the dataset level, performance is significantly influenced by the total number of point clouds. In extensive datasets encompassing diverse content, prediction accuracy tends to be lower compared to smaller datasets with fewer variations, even when their distortion types are similar. Furthermore, distortion levels play a crucial role in impacting prediction accuracy, with fine-grained distortion proving more challenging than coarse division. In the context of point clouds, compound distortion does not necessarily result from the mixture of multiple distortion types, even a single distortion type can concurrently compromise texture and geometry. Hand-crafted geometry and texture features exhibit distinct strengths and weaknesses across various distortion types. Combining both types of features adaptively with distortions may enhance prediction accuracy.

5. Conclusion and future work

This paper proposes a PCA-based FR PCQA metric, namely PointPCA+, which relies on an enriched set of lower complexity descriptors with respect to its PointPCA predecessor. After a pre-processing step, features are extracted from both geometric and textural domains. A subset of features is selected to enhance the stability of the model, and a learning-based feature fusion based on ensemble learning is applied to the feature subset, to provide a quality score for a distorted point cloud. Our experimental results demonstrate that PointPCA+ outperforms the majority of existing PCQA metrics, reaching second place in Track#1 and third place in Track#3 and Track#5 of the ICIP 2023 PCVQA grand challenge.

Compared to other teams in the PCVQA grand challenge, PointPCA+ has the highest computational complexity, with a processing time of 1000 ms, while the lowest reported is 8.6 ms. This inefficiency stems from the point-based FR PCQA framework, especially the PCA decomposition applied to each point. One solution to reduce the computational load is to downsample the point cloud before processing, while maintaining its overall structure. Additionally, both PointPCA and PointPCA+ mainly focus on local features, neglecting the broader geometric and texture context of the point cloud. To improve, incorporating global features that capture more comprehensive information is essential. This will lead to a more efficient and holistic characterization of point clouds, meeting the demands of the PCQA field. These additions are poised to contribute to a more holistic and efficient characterization of the point cloud, aligning with the rigorous demands of the PCQA field.

CRediT authorship contribution statement

Xuemei Zhou: Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Evangelos Alexiou:** Writing – review & editing, Validation, Supervision, Methodology. **Irene Viola:** Writing – review & editing, Visualization, Validation, Supervision, Funding acquisition. **Pablo Cesar:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported through the NWO WISE grant and the European Commission Horizon Europe program, under the grant agreement 101070109, *TRANSMIXR* https://transmixr.eu/. Funded by the European Union.

Data availability

Data will be made available on request.

References

[1] I. Sipiran, A. Mendoza, A. Apaza, C. Lopez, Data-driven restoration of digital archaeological pottery with point cloud analysis, Int. J. Comput. Vis. 130 (9) (2022) 2149–2165.

[2] W. Cao, J. Wu, Y. Shi, D. Chen, Restoration of individual tree missing point cloud based on local features of point cloud, Remote Sens. 14 (6) (2022) 1346.

[3] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, H. Yang, Reduced reference perceptual quality model with application to rate control for video-based point cloud compression, IEEE Trans. Image Process. 30 (2021) 6623–6636.

[4] H. Su, Q. Liu, Y. Liu, H. Yuan, H. Yang, Z. Pan, Z. Wang, Bitstream-based perceptual quality assessment of compressed 3D point clouds, IEEE Trans. Image Process. 32 (2023) 1815–1828.

[5] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P.A. Chou, R.A. Cohen, M. Krivokuća, S. Lasserre, Z. Li, J. Llach, K. Mammou, R. Mekuria, O. Nakagami, E. Siahaan, A. Tabatabai, A.M. Tourapis, V. Zakharchenko, Emerging MPEG standards for point cloud compression, IEEE J. Emerg. Sel. Top. Circuits Syst. 9 (1) (2019) 133–148.

[6] X. Xu, B. Cizmeci, A. Al-Nuaimi, E. Steinbach, Point cloud-based model-mediated teleoperation with dynamic and perception-based model updating, IEEE Trans. Instrum. Meas. 63 (11) (2014) 2558–2569.

[7] T.M. Borges, D.C. Garcia, R.L. de Queiroz, Fractional super-resolution of voxelized point clouds, IEEE Trans. Image Process. 31 (2022) 1380–1390.

[8] Z. Liu, Q. Li, X. Chen, C. Wu, S. Ishihara, J. Li, Y. Ji, Point cloud video streaming: Challenges and solutions, IEEE Netw. 35 (5) (2021) 202–209.

[9] J. Van Der Hooft, T. Wauters, F. De Turck, C. Timmerer, H. Hellwagner, Towards 6dof http adaptive streaming through point cloud compression, in: ACM MM, 2019, pp. 2405–2413.

[10] M. Bassier, S. Vincke, H. De Winter, M. Vergauwen, Drift invariant metric quality control of construction sites using BIM and point cloud data, ISPRS Int. J. Geo-Inf. 9 (9) (2020) 545.

[11] S. Subramanyam, I. Viola, J. Jansen, E. Alexiou, A. Hanjalic, P. Cesar, Subjective QoE evaluation of user-centered adaptive streaming of dynamic point clouds, in: QoMEX, IEEE, 2022, pp. 1–6.

[12] W. Lin, S. Lee, et al., Visual saliency and quality evaluation for 3D point clouds and meshes: An overview, APSIPA Trans. Signal Inf. Process. 11 (1) (2022).

[13] P. Ye, J. Kumar, L. Kang, D. Doermann, Unsupervised feature learning framework for no-reference image quality assessment, in: CVPR, 2012, pp. 1098–1105, http://dx.doi.org/10.1109/CVPR.2012.6247789.

[14] L. Zhang, L. Zhang, A.C. Bovik, A feature-enriched completely blind image quality evaluator, IEEE Trans. Image Process. 24 (8) (2015) 2579–2591.

[15] J. Xu, P. Ye, Q. Li, H. Du, Y. Liu, D. Doermann, Blind image quality assessment based on high order statistics aggregation, IEEE Trans. Image Process. 25 (9) (2016) 4444–4457.

[16] W. Xue, L. Zhang, X. Mou, Learning without human scores for blind image quality assessment, in: CVPR, 2013.

[17] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, M. Manohara, et al., Toward a practical perceptual video quality metric, Netflix Tech. Blog 6 (2) (2016) 2.

[18] X. Zhou, Y. Zhang, N. Li, X. Wang, Y. Zhou, Y.-S. Ho, Projection invariant feature and visual saliency-based stereoscopic omnidirectional image quality assessment, IEEE Trans. Broadcast. 67 (2) (2021) 512–523.

[19] Y. Zhang, S. Kwong, S. Wang, Machine learning based video coding optimizations: A survey, Inform. Sci. 506 (2020) 395–423.

[20] Y. Liu, Q. Yang, Y. Xu, L. Yang, Point cloud quality assessment: Dataset construction and learning-based no-reference metric, ACM Trans. Multimed. Comput. Commun. Appl. 19 (2s) (2023) 1–26.

[21] Y. Zhang, X. Gao, L. He, W. Lu, R. He, Objective video quality assessment combining transfer learning with CNN, IEEE Trans. Neural Netw. Learn. Syst. 31 (8) (2019) 2716–2730.

[22] J. Kim, A.D. Nguyen, S. Lee, Deep CNN-based blind image quality predictor, IEEE Trans. Neural Netw. Learn. Syst. 30 (1) (2018) 11–24.

[23] X. Wang, J. Xiong, H. Gao, W. Lin, Regression-free blind image quality assessment, 2023, arXiv preprint arXiv:2307.09279.

[24] G. Meynet, Y. Nehmé, J. Digne, G. Lavoué, PCQM: A full-reference quality metric for colored 3D point clouds, in: 2020 Twelfth International Conference on Quality of Multimedia Experience, QoMEX, 2020, pp. 1–6, http://dx.doi.org/10.1109/QoMEX48832.2020.9123147.

[25] E. Alexiou, X. Zhou, I. Viola, P. Cesar, PointPCA: Point cloud objective quality assessment using PCA-based descriptors, EURASIP J. Image Video Process. 2024 (1) (2024) 1–27.

[26] E. Alexiou, T. Ebrahimi, Towards a point cloud structural similarity metric, in: 2020 IEEE International Conference on Multimedia Expo Workshops, ICMEW, 2020, pp. 1–6, http://dx.doi.org/10.1109/ICMEW46912.2020.9106005.

[27] X. Zhou, E. Alexiou, I. Viola, P. Cesar, PointPCA+: Extending PointPCA objective quality assessment metric, in: 2023 IEEE International Conference on Image Processing Challenges and Workshops, ICIPCW, IEEE, 2023, pp. 1–5.

[28] ITU-R BT.500-13, Methodology for the Subjective Assessment of the Quality of Television Pictures, International Telecommunications Union, 2012.

[29] ITU-T P.1401, Methods, Metrics and Procedures for Statistical Evaluation, Qualification and Comparison of Objective Quality Prediction Models, International Telecommunication Unionn, 2012.

[30] K. Ma, Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, L. Zhang, Waterloo exploration database: New challenges for image quality assessment models, IEEE Trans. Image Process. 26 (2) (2016) 1004–1016.

[31] V. Hosu, H. Lin, T. Sziranyi, D. Saupe, KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment, IEEE Trans. Image Process. 29 (2020) 4041–4056.

[32] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, TID2008-a database for evaluation of full-reference visual quality assessment metrics, Adv. Mod. Radioelectron. 10 (4) (2009) 30–45.

[33] X. Zhou, I. Viola, E. Alexiou, J. Jansen, P. Cesar, QAVA-DPC: Eye-tracking based quality assessment and visual attention dataset for dynamic point cloud in 6 DoF, in: 2023 IEEE International Symposium on Mixed and Augmented Reality, ISMAR, IEEE Computer Society, Los Alamitos, CA, USA, 2023, pp. 69–78, http://dx.doi.org/10.1109/ISMAR59233.2023.00021.

[34] A. Ak, E. Zerman, M. Quach, A. Chetouani, A. Smolic, G. Valenzise, P.L. Callet, BASICS: Broad quality assessment of static point clouds in compression scenarios, 2023, arXiv preprint arXiv:2302.04796.

[35] J. Prazeres, M. Pereira, A.M. Pinheiro, Quality evaluation of point cloud compression techniques, 2023, Available At SSRN 4549486.

[36] R. Mekuria, K. Blom, P. Cesar, Design, implementation, and evaluation of a point cloud codec for tele-immersive video, IEEE Trans. Circuits Syst. Video Technol. 27 (4) (2017) 828–842.

[37] D. Tian, H. Ochimizu, C. Feng, R. Cohen, A. Vetro, Geometric distortion metrics for point cloud compression, in: 2017 IEEE International Conference on Image Processing, ICIP, 2017, pp. 3460–3464, http://dx.doi.org/10.1109/ICIP.2017.8296925.

[38] A. Javaheri, C. Brites, F. Pereira, J. Ascenso, Improving PSNR-based quality metrics performance for point cloud geometry, in: 2020 IEEE International Conference on Image Processing, ICIP, 2020, pp. 3438–3442, http://dx.doi.org/10.1109/ICIP40778.2020.9191233.

[39] A. Javaheri, C. Brites, F. Pereira, J. Ascenso, A generalized hausdorff distance based quality metric for point cloud geometry, in: 2020 Twelfth International Conference on Quality of Multimedia Experience, QoMEX, 2020, pp. 1–6, http://dx.doi.org/10.1109/QoMEX48832.2020.9123087.

[40] E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, P. Cesar, Chapter 18 - Subjective and objective quality assessment for volumetric video, in: G. Valenzise, M. Alain, E. Zerman, C. Ozcinar (Eds.), Immersive Video Technologies, Academic Press, ISBN: 978-0-323-91755-1, 2023, pp. 501–552, http://dx.doi.org/10.1016/B978-0-32-391755-1.00024-9.

[41] E. Alexiou, T. Ebrahimi, Point cloud quality assessment metric based on angular similarity, in: 2018 IEEE International Conference on Multimedia and Expo, ICME, 2018, pp. 1–6, http://dx.doi.org/10.1109/ICME.2018.8486512.

[42] E. Alexiou, T. Ebrahimi, Exploiting user interactivity in quality assessment of point cloud imaging, in: 2019 Eleventh International Conference on Quality of Multimedia Experience, QoMEX, 2019, pp. 1–6, http://dx.doi.org/10.1109/QoMEX.2019.8743277.

[43] E. Alexiou, T. Ebrahimi, Exploiting user interactivity in quality assessment of point cloud imaging, in: 2019 Eleventh International Conference on Quality of Multimedia Experience, QoMEX, IEEE, 2019, pp. 1–6.

[44] Q. Liu, H. Su, Z. Duanmu, W. Liu, Z. Wang, Perceptual quality assessment of colored 3D point clouds, IEEE Trans. Vis. Comput. Graphics (2022) 1.

[45] Z. Wang, Q. Li, Information content weighting for perceptual image quality assessment, IEEE Trans. Image Process. 20 (5) (2011) 1185–1198.

[46] Q. Liu, H. Yuan, H. Su, H. Liu, Y. Wang, H. Yang, J. Hou, PQA-Net: Deep no reference point cloud quality assessment via multi-view projection, IEEE Trans. Circuits Syst. Video Technol. 31 (12) (2021) 4645–4660.

[47] I. Viola, S. Subramanyam, P. Cesar, A color-based objective quality metric for point cloud contents, in: 2020 Twelfth International Conference on Quality of Multimedia Experience, QoMEX, 2020, pp. 1–6, http://dx.doi.org/10.1109/QoMEX48832.2020.9123089.

[48] Q. Yang, Z. Ma, Y. Xu, Z. Li, J. Sun, Inferring point cloud quality via graph similarity, IEEE Trans. Pattern Anal. Mach. Intell. 44 (6) (2022) 3015–3029.

[49] R. Diniz, P.G. Freitas, M. Farias, A novel point cloud quality assessment metric based on perceptual color distance patterns, Electron. Imaging 33 (2021) 1–11.

[50] R. Diniz, P.G. Freitas, M.C. Farias, Multi-distance point cloud quality assessment, in: 2020 IEEE International Conference on Image Processing, ICIP, 2020, pp. 3443–3447, http://dx.doi.org/10.1109/ICIP40778.2020.9190956.

[51] R. Diniz, P.G. Freitas, M.C. Farias, Point cloud quality assessment based on geometry-aware texture descriptors, Comput. Graph. (2022).

[52] Q. Liu, Y. Liu, H. Su, H. Yuan, R. Hamzaoui, Progressive knowledge transfer based on human visual perception mechanism for perceptual quality assessment of point clouds, 2022, arXiv preprint arXiv:2211.16646.

[53] Z. Zhang, W. Sun, X. Min, Q. Zhou, J. He, Q. Wang, G. Zhai, MM-PCQA: Multi-modal learning for no-reference point cloud quality assessment, 2022, arXiv preprint arXiv:2209.00244.

[54] Z. Zhang, W. Sun, X. Min, T. Wang, W. Lu, G. Zhai, No-reference quality assessment for 3D colored point cloud and mesh models, 2021, arXiv preprint arXiv:2107.02041.

[55] Q. Yang, Y. Liu, S. Chen, Y. Xu, J. Sun, No-reference point cloud quality assessment via domain adaptation, in: CVPR, 2022, pp. 21179–21188.

[56] W. Xie, K. Wang, Y. Ju, M. Wang, pmBQA: Projection-based blind point cloud quality assessment via multimodal learning, in: ACM MM, 2023, pp. 3250–3258.

[57] ITU-R BT.709-6, Parameter Values for the HDTV Standards for Production and International Programme Exchange, International Telecommunication Unionn, 2015.

[58] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Mach. Learn. 46 (2002) 389–422.

[59] R. Diniz, P.G. Freitas, M.C.Q. Farias, Color and geometry texture descriptors for point-cloud quality assessment, IEEE Signal Process. Lett. 28 (2021) 1150–1154.

[60] E. Alexiou, I. Viola, T.M. Borges, T.A. Fonseca, R.L. De Queiroz, T. Ebrahimi, A comprehensive study of the rate-distortion performance in MPEG point cloud compression, APSIPA Trans. Signal Inf. Process. 8 (2019) e27.

[61] Q. Yang, H. Chen, Z. Ma, Y. Xu, R. Tang, J. Sun, Predicting the perceptual quality of point cloud: A 3d-to-2d projection-based exploration, IEEE Trans. Multimed. 23 (2020) 3877–3891.

[62] L. Krasula, K. Fliegel, P. Le Callet, M. Klíma, On the accuracy of objective image and video quality models: New methodology for performance evaluation, in: 2016 Eighth International Conference on Quality of Multimedia Experience, QoMEX, 2016, pp. 1–6, http://dx.doi.org/10.1109/QoMEX.2016.7498936.

[63] A. Chetouani, G. Valenzise, A. Ak, E. Zerman, M. Quach, M. Tliba, M.A. Kerkouri, P.L. Callet, ICIP 2023 - Point Cloud Visual Quality Assessment Grand Challenge, IEEE, 2023.

[64] L. Breiman, R. Cutler, Random forests machine learning, J. Clin. Microbiol. 2 (2001) 199–228.