Abstract—Point cloud quality assessment (PCQA) has become an appealing research field in recent days. Considering the importance of saliency detection in quality assessment, we propose an effective full-reference PCQA metric which makes an attempt to utilize the saliency information to facilitate quality prediction, called point cloud quality assessment using 3D saliency maps (PQSM). Specifically, we first propose a projection-based point cloud saliency map generation method, in which depth information is introduced to better reflect the geometric characteristics of point clouds. Then, we construct point cloud local neighborhoods to derive three structural descriptors to indicate the geometry, color and saliency discrepancies. Finally, a saliency-based pooling strategy is proposed to generate the final quality score. Extensive experiments are performed on four independent PCQA databases. The results demonstrate that the proposed PQSM shows competitive performances compared to multiple state-of-the-art PCQA metrics.

Index Terms—objective quality assessment, point cloud, saliency map, structural similarity

I. INTRODUCTION

With the rapid development of 3D capturing technologies, point clouds (PCs) have emerged as a popular and prominent format to represent 3D photo-realistic content. A PC is composed of a great number of disordered points, and each point is attached with the information of its 3D coordinates and additional attributes such as RGB color and normals. In recent years, we have witnessed the occurrence of many applications based on PCs (e.g., navigation [1], industrial robotics[2] and mixed reality[3], etc) which require PC data with high quality. However, a variety of distortions could be induced during the relevant processing of PCs, such as acquisition, compression [4][5], transmission [6], and rendering, which may impair the perceptual quality of the human visual system (HVS). Therefore, point cloud quality assessment (PCQA) has become an appealing research field to ensure provide high-quality human perception service.

PCQA aims to directly reflect or approximate the perception quality of HVS for specific PCs. Subjective PCQA is the most straightforward and reliable method, but also time-consuming and expensive. Thus, effective objective PCQA metrics which correlate well with HVS for practical application are highly desired. In this paper, we focus on the full-reference (FR) objective PCQA metrics. Existing FR-PCQA metrics either project 3D PCs onto 2D planes, utilizing the well-developed image quality assessment (IQA) algorithms; or inferring the perceptual quality by modeling the 3D properties. In terms of the former, projection-based metrics like [7] project PCs onto several perpendicular image planes and aggregate image-based features to evaluate the PC quality. In terms of the latter, point-wise metrics, such as the point-to-point[8] and the point-to-plane[9], calculate the differences of the Euclidean distance related features between the corresponding point pairs. Meanwhile, some other metrics, e.g., PointSSIM[10], GraphSIM[11], MS-GraphSIM[12] and MPED[13], construct local neighborhoods to involve the structural features, which usually have better performances due to the fact that HVS is more sensitive to structural information.

One of the things that above methods overlooked is that they hardly consider the vital role of visual saliency in quality prediction. Specifically, the visual saliency is an important characteristic of HVS which can reflect the human attention distribution. In the domain of image quality assessment, some metrics integrated with saliency information have achieved great success, such as VSI [14]. One reason that there are few PCQA metrics involved with visual saliency is due to the lag great success, such as VSI [14]. One reason that there are few PCQA metrics involved with visual saliency is due to the lag PC saliency detection. For the PC saliency detection, most methods are directly conducted on uncolored 3D PCs[15][16], which is usually utilized in machine vision tasks rather than human vision tasks.

Inspired by the strategy proposed in [17], we use the projection to transfer the achievement of image saliency detection to PC and propose a novel saliency guided metric, called the point cloud quality assessment using 3D saliency maps (PQSM). Specifically, we first propose to incorporate depth-related maps into the projection-based saliency detection model to involve the intrinsic geometric information of PCs. We then design three descriptors to quantitively measure the distortions on geometry, color and saliency maps. After that, a saliency-based pooling strategy is adopted to generate the final score. Our method is examined by extensive experiments on four publicly accessible PCQA databases. The results show the competitive performance of our proposed model in inferring the perceived quality of distorted PCs.

This paper is supported in part by National Natural Science Foundation of China (61971282, U20A20185) and STCSM under Grant (22DZ2229005). The corresponding author is Yiling Xu(e-mail: yl.xu@sjtu.edu.cn).
II. METHOD

In this section, we first illustrate how to enhance the projection-based saliency maps via depth information. Then, we detail the construction of the proposed PQSM. The overall framework of the proposed metric is shown in Fig. I.

A. Projection-based Point Cloud Saliency Maps

The projection-based saliency generation algorithms for PCs first generate the 2D saliency maps based on projected images, then map these saliency values from pixels in the images to the corresponding points of the original PCs [17]. Although such methods have achieved considerable correlations with human attention, they ignore the intrinsic geometric characteristic of PCs. Intuitively, in 3D space, the closer an item is to us, the more likely it is to grab our attention [7]. Therefore, the depth information obtained from the projection process can be leveraged to involve geometry characteristics of PCs in the generation of saliency maps [18].

To simulate the above perceptual property, we utilize depth-related weights to enhance projection-based saliency maps.

Consider one point cloud $X$ as $\mathcal{P} = \{p_1, p_2, \ldots, p_N\} \in \mathbb{R}^{N \times 6}$, where each point $p_i = [p_{i}^{O}, p_{i}^{C}] = [x, y, z, R, G, B]$ is a 6-dimension vector containing 3D coordinates $p_{i}^{O}$ and RGB attributes $p_{i}^{C}$. Let $\mathcal{P}$ be orthogonally projected onto $L$ different 2D planes, on which we obtain $L$ projected texture maps and $L$ depth maps. Note that when deriving depth maps, the distance between each projected plane and the closest point in the point cloud is set as 10. We define the $l$-th projected texture map as $\mathcal{Y}_l = \{\mathbf{y}_j\}_{j=1}^{N_l} \in \mathbb{R}^{N_l \times 6}$ and its corresponding depth map as $\mathcal{Z}_l = \{d_{i,l}\}_{i=1}^{N_l} \in \mathbb{R}^{N_l}$, where $N_l$ denotes the pixel number of $\mathcal{Y}_l$ or $\mathcal{Z}_l$ and $d_{i,l}$ denotes the depth value of $i$-th pixel in $\mathcal{Z}_l$. Since one projected image may not contain all the points in $\mathcal{P}$, we have $N_l \leq N$. For each texture map, we first calculate its 2D saliency maps based on the established image saliency detection algorithm from [19]. For the texture map $\mathcal{Y}_l$, we represent its saliency map as $\mathbf{s}_l = \{s_{i,l}\}_{i=1}^{N_l} \in \mathbb{R}^{N_l}$, where $s_{i,l}$ represents the saliency value of $i$-th pixel.

Then, we compute the depth-related weights based on the depth map $\mathcal{Z}_l$ and then merge the weights into the 2D saliency map. Specifically, for $i$-th pixel located in the $\mathcal{Z}_l$, we can derive one new depth enhanced saliency map $\tilde{s}_l = \{\tilde{s}_{i,l}\}_{i=1}^{N_l}$ as follows,

$$
\tilde{s}_{i,l} = \frac{e^{-\frac{d_{i,l}}{\sigma_s}}}{\sum_{j=1}^{N_l} e^{-\frac{d_{j,l}}{\sigma_s}}} \cdot s_{i,l},
$$

where $\sigma_s$ is a constant to be determined. According to Eq. (1), we endow higher weights for points with small depths because these points are considered to be closer to human eyes and therefore are more salient.

Finally, for each point in point clouds, we average the depth-enhanced 2D saliency values of its corresponding pixels in $L$ projected images to derive the 3D saliency map $\mathbf{s}_l = \{\tilde{s}_{i,l}\}_{i=1}^{N}$, which is utilized for subsequent PCQA task.

B. Point Cloud Quality Assessment using 3D Saliency Maps

1) Neighborhood Construction: Consider the reference point cloud as $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^{N} \in \mathbb{R}^{N \times 6}$ and the distorted point cloud as $\mathcal{Y} = \{\mathbf{y}_j\}_{j=1}^{M} \in \mathbb{R}^{M \times 6}$. We first define a series of local patch pairs by ball query. Specifically, for a given point $\mathbf{x}_i$ in the reference PC, we have two local patches in $\mathcal{X}$ and $\mathcal{Y}$ as,

$$
\begin{align*}
N_{\mathbf{x}_i,\mathcal{X}} &= \{\mathbf{x}_j \mid ||\mathbf{x}_j - \mathbf{x}_i||_2 \leq r\}, \\
N_{\mathbf{x}_i,\mathcal{Y}} &= \{\mathbf{y}_j \mid ||\mathbf{y}_j - \mathbf{x}_i||_2 \leq r\},
\end{align*}
$$

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where \( r \) denotes the searching radius. We then compute the features over \( N^X_{x_i,r} \) and \( N^Y_{x_i,r} \).

2) Local Geometry Similarity: Geometric distortions, such as downsampling and geometry Gaussian noise, usually result in a change of the point density and point distribution, which is easily captured by the HVS. To quantitatively measure the point density and point distribution, we propose two statistical features within Euclidean distances between \( x_i \) and its neighbors in \( N^X_{x_i,r} \) and \( N^Y_{x_i,r} \) as follows:

\[
\mu^X_{i,geo} = \frac{\sum_{x_j \in N^X_{x_i,r}} \|x_j^O - x_i^O\|^2_2}{|N^X_{x_i,r}|}, \tag{3}
\]

\[
\sigma^X_{i,geo} = \frac{\sum_{x_j \in N^X_{x_i,r}} (\|x_j^O - x_i^O\|^2 - \mu^X_{i,geo})^2}{|N^X_{x_i,r}|}. \tag{4}
\]

\( \mu^Y_{i,geo} \) and \( \sigma^Y_{i,geo} \) can be obtained via the same paradigm as Eq. (3) and (4). Then we mimic the similarity measurement used in SSIM [20] to fuse the above statistical features as follows:

\[
F_1 = \frac{2\mu^X_{i,geo} \cdot \mu^Y_{i,geo} + T_1}{(\mu^X_{i,geo})^2 + (\mu^Y_{i,geo})^2 + T_1}, \quad \frac{2\sigma^X_{i,geo} \cdot \sigma^Y_{i,geo} + T_1}{(\sigma^X_{i,geo})^2 + (\sigma^Y_{i,geo})^2 + T_1}, \tag{5}
\]

where \( T_1 \) are small constants to avoid instability.

3) Local Color Similarity: It is widely believed that the HVS is more sensitive to contrast than absolute intensity [21]. Therefore, we try to calculate local color contrast in \( N^X_{x_i,r} \) and \( N^Y_{x_i,r} \) to reflect color distortions of PCs in our metric. Considering the sensitivity of the HVS to luminance channel [11], we first calculate the luminance information according to the ITU-R Recommendation BT.709 [22] as following,

\[
x_i^L = x_i^C \cdot [0.257, 0.504, 0.098]^T + 16. \tag{6}
\]

Similar to Eq. (3), we then derive the \( \mu^X_{i,lum} \) by merging the luminance difference between \( x_i \) and \( N^X_{x_i,r} \) to represent the local color variation:

\[
\mu^X_{i,lum} = \frac{\sum_{x_j \in N^X_{x_i,r}} \|x_j^L - x_i^L\|^2_2}{|N^X_{x_i,r}|}. \tag{7}
\]

\( \mu^Y_{i,lum} \) can be obtained in the same way. Next, the quality feature reflecting color distortion is defined as

\[
F_2 = \frac{2\mu^X_{i,lum} \cdot \mu^Y_{i,lum} + T_1}{(\mu^X_{i,lum})^2 + (\mu^Y_{i,lum})^2 + T_1}. \tag{8}
\]

4) Local Saliency Similarity: According to Section II-A, we respectively calculate the saliency maps of \( X \) and \( Y \) as \( S_X = \{s^O_i\} \in \mathbb{R}^N \) and \( S_Y = \{s^O_j\} \in \mathbb{R}^M \), respectively. Referring to some previous researches [14], human attention distribution changes along with different distortions, thus highly correlating with visual perception. Therefore, we can utilize the discrepancy between \( S_X \) and \( S_Y \) to assist in quality assessment. The mean of the saliency difference between \( x_i \) and \( N^X_{x_i,r} \) is first obtained by

\[
\mu^X_{i,sal} = \frac{\sum_{x_j \in N^X_{x_i,r}} \|s_j^O - s_i^O\|^2_2}{|N^X_{x_i,r}|}, \tag{9}
\]

\( \mu^Y_{i,sal} \) can be calculated similarly. Then we define the quality feature reflecting the distance between two saliency maps as follows:

\[
F_3 = \frac{2\mu^X_{i,sal} \cdot \mu^Y_{i,sal} + T_2}{(\mu^X_{i,sal})^2 + (\mu^Y_{i,sal})^2 + T_2}, \tag{10}
\]

where \( T_2 \) is another small constant.

5) Global Similarity: We have obtained three quality features for each point, which separately measure the local discrepancy of geometry structure, color contrast and saliency information. These features are then fused together to generate one local index:

\[
SIM_i = F_1 \cdot F_2 \cdot F_3. \tag{11}
\]

An appropriate pooling strategy is desired to fuse all local indices. It has been widely accepted that not all points in point clouds share equal importance, and these points with high saliency values usually perform a more vital role in quality assessment. Therefore, the saliency map is naturally considered as the weighting factor for our pooling. Specifically, the final quality score is defined as follows:

\[
Q = \frac{SIM_i \cdot x_i^S}{\sum_{i=1}^N x_i^S}. \tag{12}
\]

III. EXPERIMENTS

A. Database and Evaluation Criteria

- SJTU-PCQA: It has 9 reference PCs and 378 distorted samples.
- WPC: It contains 20 reference PCs and 740 distorted samples.
- ICIP2020: It contains 6 reference PCs and 90 distorted samples.
- M-PCCD: It contains 8 reference PCs and 232 distorted samples.

The Pearson linear correlation coefficient (PLCC), the Spearman rank order correlation coefficient (SROCC), and the root mean square error (RMSE) are utilized to evaluate the performance of different models. Higher PLCC or SROCC values indicate better model performance, whereas a lower RMSE value is indicative of better performance.

B. Selection of Model Parameters and 2D Saliency Method

Here we present the detailed parameter setting and the selection of the 2D saliency method used in Section II-A.

\( \sigma_s \) in the weight calculation. We determine \( \sigma_s \) in Eq. (1) as one tenth of the maximum side length of the bounding box, alleviating the impacts from different scales of PCs.

\( r \) in the neighborhood construction We set \( r \) as the average of distances between each point in the reference PC.
TABLE II
PERFORMANCE COMPARISON OF DIFFERENT PCQA METRICS ON THE SJTU-PCQA, WPC, ICIP2020, AND M-PCCD DATASETS

| Metrics                  | SJTU-PCQA[7] | WPC[27] | ICIP2020[28] | M-PCCD[29] | Weighted-AVE |
|--------------------------|--------------|---------|--------------|------------|--------------|
|                           | PLCC         | SROCC   | RMSE         | PLCC       | SROCC        | RMSE         | PLCC        | SROCC        | RMSE         | PLCC        | SROCC        | RMSE         |
| 2point-MSE[30]           | 0.837        | 0.566   | 18.70        | 0.888      | 0.878        | 1.224        | 0.778       | 0.957        | 0.515        | 0.562       | 0.639        | 10.26        |
| 2point-MSE-PSNR[30]      | 0.877        | 0.566   | 18.70        | 0.888      | 0.878        | 1.224        | 0.778       | 0.957        | 0.515        | 0.562       | 0.639        | 10.26        |
| 2point-Hausdoff[30]      | 0.223        | 0.524   | 2.366        | 0.469      | 0.446        | 20.24        | 0.383       | 0.515        | 0.486        | 0.523       | 0.457        | 11.54        |
| 2point-Hausdoff-PSNR[30] | 0.219        | 0.566   | 2.368        | 0.469      | 0.446        | 20.24        | 0.383       | 0.515        | 0.486        | 0.523       | 0.457        | 11.54        |
| pointSSIM-geo[10]        | 0.753        | 0.676   | 1.596        | 0.488      | 0.446        | 20.01        | 0.383       | 0.515        | 0.486        | 0.523       | 0.457        | 11.54        |
| pointSSIM-col[10]        | 0.725        | 0.704   | 1.672        | 0.510      | 0.454        | 19.71        | 0.383       | 0.515        | 0.486        | 0.523       | 0.457        | 11.54        |
| IW-SSIM\textunderscore projection-based[27] | 0.779        | 0.783  | 1.422        | 0.850      | 0.848        | 12.062       | 0.999       | 0.897        | 0.472        | 0.717       | 0.749        | 0.948        |
| M-PED-1-norm[13]         | 0.898        | 0.886   | 1.067        | 0.695      | 0.673        | 16.47        | 0.964       | 0.951        | 0.303        | 0.846       | 0.869        | 0.726        |
| M-PED-2-norm[13]         | 0.898        | 0.890   | 1.070        | 0.637      | 0.624        | 17.67        | 0.967       | 0.961        | 0.290        | 0.867       | 0.887        | 0.678        |
| PQSM (proposed)          | 0.894        | 0.884   | 1.086        | 0.753      | 0.737        | 15.08        | 0.909       | 0.902        | 0.474        | 0.899       | 0.911        | 0.597        |

and the 10-th nearest point in the distorted PC, which can also mitigate the impact of the different scales of PCs.

$T_1$ and $T_2$ in the similarity pooling, $T_1$ is set to be 0.001. $T_2$ is set to be $10^{-14}$ due to the small weight values after the process of normalization.

2D saliency method. For the 2D saliency detection method used in Section II-A, we test several eminent algorithms on SJTU-PCQA and show the performance in terms of SROCC in TABLE I. We can see that UHF provides the best performance, which is consequently utilized in our metric.

C. Overall Performance Comparison

We evaluate our proposed method and several other state-of-art PCQA metrics on four databases and list the performance in TABLE II. The top three performances of each database are indicated in bold.

According to TABLE II, we can see that the proposed PQSM achieves competitive performance across the four databases. Specifically, PQSM is among the top three performances on three databases, i.e., SJTU-PCQA, WPC and M-PCCD. This is because the saliency information is comprehensively taken into our consideration in our method. The visual saliency reflects human visual fixation, which could help define which points play a more significant role in the distortion.

To investigate the impact of saliency information in PQSM, we measure the performances resulting from different feature combinations with different pooling strategies on SJTU-PCQA, which are listed in TABLE III. Note that $AVE$ and $SAW$ represent average pooling and saliency-based weighting, respectively. From TABLE III, we can see that both saliency feature and saliency-based weighting improve the model performance, which states that introducing proper saliency information is beneficial for PCQA tasks. We also investigate the impact of $r$ (cf. Eq. (2)) in the neighborhood construction in TABLE IV, where $N_r$ is set as the average distance between each point in the reference PC and its $N_r$-th nearest point in the distorted PC. We can see that PQSM achieves the best performance when $N_r = 10$, which supports our parameter selection in Sec III-B.

D. Ablation Study

To investigate the impact of saliency information in PQSM, we measure the performances resulting from different feature combinations with different pooling strategies on SJTU-PCQA, which are listed in TABLE III. Note that $AVE$ and $SAW$ represent average pooling and saliency-based weighting, respectively. From TABLE III, we can see that both saliency feature and saliency-based weighting improve the model performance, which states that introducing proper saliency information is beneficial for PCQA tasks. We also investigate the impact of $r$ (cf. Eq. (2)) in the neighborhood construction in TABLE IV, where $N_r$ is set as the average distance between each point in the reference PC and its $N_r$-th nearest point in the distorted PC. We can see that PQSM achieves the best performance when $N_r = 10$, which supports our parameter selection in Sec III-B.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
REFERENCES

[1] M. Whitty, S. Cossell, K. S. Dang, J. Guivant, and J. Katupitiya, “Autonomous Navigation Using A Real-time 3D Point Cloud,” in Australas. Conf. Robot. Automat., 2010, pp. 1–3.

[2] M. Liu, “Robotic online path planning on point cloud,” IEEE Trans. Cybern., vol. 46, no. 5, pp. 1217–1228, 2016.

[3] F. Capraro and S. Milani, “Rendering-aware point cloud coding for mixed reality devices,” in 2019 IEEE Int. Conf. Image Process., 2019, pp. 3706–3710.

[4] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuca, S. Lasserre, Z. Li, J. Llach, K. Mammou, R. Mekuria, O. Nakagami, E. Siahana, A. Tabatabai, A. M. Tourapis, and V. F. Figueiredo, “Zakharchenko, Emerging MPEG Standards for Point Cloud Compression,” IEEE J. Emerg. Sel. Topics Circuits Syst., vol. 9, no. 1, pp. 133–138, Dec. 2019.

[5] R. Schnabel and R. Klein, “Octree-base-Point-Cloud Compression,” in Symp. on Point-Based Graph. The Eurographics Association, 2006, pp. 111–120.

[6] G.-R. Jang, Y.-D. Shin, J.-H. Park, and M.-H. Baeg, “Real-time Point-Cloud Data Transmission for Teleoperation Using H.264/AVC,” in IEEE Int. Symp. Secur. Rescue Robot., Oct. 2014, pp. 1–6.

[7] Q. Yang, H. Chen, Z. Ma, Y. Xu, R. Tang, and J. Sun, “Predicting the Perceptual Quality of Point Cloud: A 3D-to-2D Projection-Based Exploration,” IEEE Trans. Multimedia, vol. 23, pp. 3877–3891, Oct. 2021.

[8] R. Mekuria, K. Blom, and P. Cesar, “Design, Implementation, and Evaluation of a Point Cloud Codec for Tele-Immersive Video,” IEEE Trans. Circuits Syst. Video Technology, vol. 27, no. 4, pp. 828–842, Mar. 2017.

[9] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, “Geometric Distortion Metrics for Point Cloud Compression,” in IEEE Int. Conf. Image Process., Sept. 2017, pp. 3460–3464.

[10] E. Alexiou and T. Ebrahimi, “Towards a point cloud structural similarity metric,” in IEEE Int. Conf. Multimedia Expo. Workshops, 2020, pp. 1–6.

[11] 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, Dec. 2022.

[12] Y. Zhang, Q. Yang, and Y. Xu, MS-GraphSIM: Inferring Point Cloud Quality via Multiscale Graph Similarity. NY, USA: Association for Computing Machinery, Oct. 2021, p. 1230–1238.

[13] Q. Yang, Y. Zhang, S. Chen, Y. Xu, J. Sun, and Z. Ma, “MPED: Quantifying point cloud distortion based on multiscale potential energy discrepancy,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 5, pp. 6037–6054, 2023.

[14] L. Zhang, Y. Shen, and H. Li, “VSI: A Visual Salience-Induced Index for Perceptual Image Quality Assessment,” IEEE Trans. Image Process., vol. 23, no. 10, pp. 4270–4281, Aug. 2014.

[15] Y. Guo, F. Wang, and J. Xin, “Point-wise Salience Detection on 3D Point Clouds via Covariance Descriptors,” Vis. Comput., vol. 34, no. 10, pp. 1325–1338, 2018.

[16] X. Ding, W. Lin, Z. Chen, and X. Zhang, “Point Cloud Salience Detection by Local and Global Feature Fusion,” IEEE Trans. Image Process., vol. 28, no. 11, pp. 5379–5393, May 2019.

[17] V. F. Figueiredo, G. L. Sandri, R. L. de Queiroz, and P. A. Chou, “Salience Maps for Point Clouds,” in 2020 IEEE 22nd Int. Workshop on Multimedia Signal Process., Sept. 2020, pp. 1–5.

[18] C. Lang, T. V. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan, “Depth Matters: Influence of Depth Cues on Visual Salience,” in European Conf. on Comput. Vision. Springer, 2012, pp. 101–115.

[19] H. R. Tavakoli and J. Laaksonen, “Bottom-Up Fixation Prediction Using Unsupervised Hierarchical Models,” in Comput. Vision – ACCV 2016 Workshops, C.-S. Chen, J. Lu, and K.-K. Ma, Eds. Cham: Springer International Publishing, 2017, pp. 287–302.

[20] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, 2004.

[21] X. Qian, J. Han, G. Cheng, and L. Guo, “Optimal contrast based saliency detection,” Pattern Recognit. Lett., vol. 34, no. 11, pp. 1270–1278, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865513001591

[22] I. ITU, “Parameter Values for The HDTV Standards for Production and International Programme Exchange,” Recommendation ITU-R BT. pp. 709–5, 2002.

[23] A. Torralba, A. Oliva, M. S. Castelhano, and J. M. Henderson, “Contextual Guidance of Eye Movements and Attention in Real-world Scenes: The Role of Global Features in Object Search.” Psychological Rev., vol. 113, no. 4, p. 766, 2006.

[24] H. Rezazadegan Tavakoli, E. Rahtu, and J. Heikkilä, “Fast and Efficient Saliency Detection Using Sparse Sampling and Kernel Density Estimation,” in Scandinavian Conf. on Image Anal. Springer, 2011, pp. 666–675.

[25] N. Murray, M. Vanrell, X. Otazu, and C. A. Parraga, “Saliency Estimation Using A Non-parametric Low-level Vision Model,” in Conf. on Comput. Vision and Pattern Recognit., June 2011, pp. 433–440.

[26] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, “SUN: A Bayesian Framework for Saliency Using Natural Statistics,” J. Vision, vol. 8, no. 7, pp. 32–32, 12 2008.

[27] Q. Liu, H. Su, Z. Duannu, W. Liu, and Z. Wang, “Perceptual Quality Assessment of Colored 3D Point Clouds,” IEEE Trans. Vis. Comput. Graphics, pp. 1–1, Apr. 2022.

[28] E. Alexiou, I. Viola, T. M. Borges, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, “A comprehensive study of the rate-distortion performance in mpeg point cloud compression,” APSIPA Trans. Signal Inf. Process., vol. 8, p. e27, 2019.

[29] MPEG reference software. [Online]. Available: http://mpegx.int-evry.fr/software/MPEG/PCC/TF/mpeg-pcc-dmetric