R3-PCQA: Ray-Reprojection-Reinforcement for No-Reference 3D Point Cloud Quality Assessment

Supplementary Material

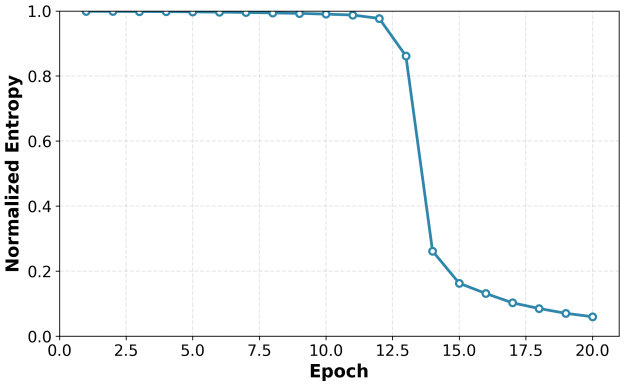

Figure 6. Convergence of the policy network's normalized entropy during the joint training phase

A. Implementation details

All experiments in this study were conducted on a single NVIDIA RTX A4500 GPU. Each point cloud was rendered as 2D projection images with a resolution of 1080×1080 from icosahedral viewpoints using Open3D [49]. These images were then resized to 448×448 before being fed into the model.

Geometric-aware reprojection. In the geometric-aware reprojection, the distance threshold ρ for ray-point intersection was set to 6. This threshold determines whether a 3D point is considered to lie on the ray cast from a 2D pixel location.

R3-PCQA architecture. For the RGB encoder $\mathcal{E}^{\text{rgb}}$, we employed a ResNet-18 [11] pretrained on ImageNet [7]. The context H_v for the QSS is extracted from Layer 1 of ResNet-18, which provides intermediate feature representations suitable for spatial localization. Each local patch $\mathbf{h}_{v,n}$ is embedded into a feature vector $\mathbf{h'}_{v,n} \in \mathbb{R}^D$ with $D = 128$ through a CNN block. The point cloud encoder $\mathcal{E}^{\text{pc}}$ is based on the PointNet++ [29] architecture. Both encoders output features with dimensionality $D = 128$, which are used throughout the local view encoder for feature fusion and token generation. The global view attention module uses multi-head attention with 3 heads. The multi-head self-attention used in the QSS is configured with 4 heads to capture inter-patch relationships.

R3-PCQA training. The R3-PCQA model was trained for a total of 40 epochs, consisting of a 20-epoch warm-up phase followed by a 20-epoch joint training phase. We used the AdamW optimizer [18] with a learning rate of 1×10^{-4} and a batch size of 4. To accelerate training, we applied mixed-precision training with GradScaler [22]. For training stability, MOS values were normalized to [0, 1] based on each dataset's maximum MOS value, and predictions were rescaled to the original MOS range of the respective dataset during inference. In the joint training phase, the temperature parameter τ in the policy network was gradually annealed from 1.0 to 0.5 over 20 epochs, enabling a smooth transition from exploration-focused to exploitation-focused policy learning.

Table 6. Ablation study on reward distribution strategies.

Reward Strategy	SRCC ↑	PLCC ↑
Uniform Reward	0.8863	0.8780
Attention-based Reward	**0.9017**	**0.8882**

B. Additional results

Entropy convergence. Figure 6 illustrates how the entropy of the action probabilities evolves for each view as the number of epochs increases during the joint training phase. The entropy decreases gradually until around 10 epochs and then drops more rapidly. By epoch 20, it has converged to approximately 0.1, indicating that the policy learned to make confident selections.

Ablation on ρ for reprojection. Point clouds are inherently sparse, so simple linear rays often fail to find valid anchors and pass through empty space between points. Therefore, cylindrical rays with a valid radius ρ are essential. To validate setting $\rho = 6$ in this experiment, as shown in Fig. 7, we experimented with varying ρ. For $\rho \leq 2$, rays often missed valid anchors due to sparsity. While $\rho = 4$ achieves a 100% detection rate, it can also select points that are not visible in the current projection, as observed for $\rho \in \{1, 2, 4\}$. In contrast, $\rho = 6$ robustly captured visible anchors while avoiding false selections; larger values provided no practical benefit, so we chose $\rho = 6$ as the best setting.

Ablation on reward. To validate the effectiveness of the proposed attention-based credit assignment mechanism, we conducted an ablation study on reward distribution strategies. We compared a uniform reward method, which equally assigns the global reward across all viewpoints, against our Attention-based reward approach, which weights rewards proportionally by α_v. Table 6 shows that attention-based reward achieves higher SRCC and PLCC,

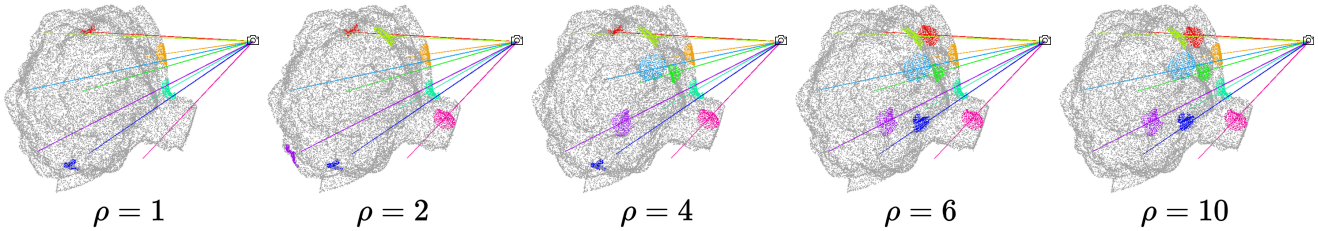

Figure 7. Ablation on ρ for reprojection.

Figure 8. Additional qualitative comparisons on the WPC dataset. For each distorted point cloud, we show the rendered multiview projections together with the ground-truth MOS, the predicted score of MM-PCQA, and the prediction of the proposed R3-PCQA. In all examples, R3-PCQA produces quality scores that are closer to the MOS than MM-PCQA, demonstrating better consistency with human subjective evaluation across various object categories.

Table 7. Ablation on number of viewpoints.

Views	SRCC ↑	PLCC ↑
4	0.8699	0.8543
6	0.8912	0.8816
12	0.8941	0.8863
20	**0.9017**	0.8882
24	0.8943	0.8851
40	0.8913	**0.8972**

demonstrating that credit assignment effectively focuses learning on selecting quality-salient subclouds at important viewpoints.

Ablation on number of viewpoints. We adopted an icosahedron-based multiview setup with 20 viewpoints. To validate this choice, we compared and analyzed the performance of R3-PCQA across viewpoint configurations of 4 (tetrahedron), 6 (octahedron), 12 (dodecahedron), and 20 (icosahedron) regular polyhedra, as well as 24 (tetrakis hexahedron) and 40 (Fibonacci sphere) viewpoints. As shown in Tab. 7, performance consistently improved with an increasing number of viewpoints, saturating at the 20-viewpoint icosahedron configuration. Increasing the viewpoints to 24 or 40 did not yield significant further gains.

Table 8. Ablation on candidate subclouds (N).

N (Subclouds)	2	4	9	16
SRCC ↑	0.8804	0.8798	**0.9017**	0.8883
PLCC ↑	0.8713	0.8701	**0.8882**	0.8815

Instead, such non-uniform configurations or excessive view counts caused view overlap and redundant directions, only increasing computational time. In conclusion, the 20-icosahedron configuration provided the most efficient performance while maintaining an optimal balance between spatial coverage and computational cost.

Ablation on the number of candidate subclouds. To validate our choice of $N = 9$ candidate subclouds, we evaluated the model's performance across varying values of $N \in \{2, 4, 9, 16\}$. As shown in Tab. 8, performance drops for $N \leq 4$, likely due to insufficient spatial coverage. Conversely, increasing N to 16 degrades performance despite higher computational cost, which can be attributed to information redundancy. Thus, $N = 9$ offers the optimal balance between efficiency and accuracy.

Ablation on warm-up phase. To validate the necessity of the warm-up phase in our curriculum learning strategy, we compared our approach, which includes a warm-up,

Table 9. Ablation study on the warm-up strategy.

Warm-up	SRCC ↑	PLCC ↑
No	0.8897	0.8793
Yes	**0.9017**	**0.8882**

against a no warm-up approach that performs joint training from the beginning. As shown in Tab. 9, the strategy incorporating the warm-up phase achieves higher performance. This validates the effectiveness of our curriculum strategy, which performs policy learning after stable feature representation learning.

Additional qualitative comparisons. Figure 8 presents additional qualitative comparison results on the WPC dataset. For each distorted point cloud, we visualize its multiview projections and report the ground-truth MOS, the predicted quality scores of the proposed R3-PCQA, and the state-of-the-art baseline MM-PCQA. Across diverse object categories and distortion levels, the predicted scores of R3-PCQA are consistently closer to the MOS than those of MM-PCQA, indicating that our model better aligns with human perceptual judgments.

Qualitative analysis of QSS. To further understand the behavior of the QSS, we visualize the dynamically selected regions in Fig. 9. As shown in the examples, the red bounding box explicitly highlights the 2D spatial region corresponding to the specific candidate subcloud selected by the QSS. The rightmost column presents a zoomed-in view of this bounded area to reveal its detailed local characteristics. As shown, rather than evaluating all regions of the view equally or naively aggregating features from all available candidates, the QSS actively directs its attention to a single visually prominent area exhibiting severe localized degradations. The zoomed-in views demonstrate that the learned policy isolates subclouds containing critical structural or textural artifacts, such as geometric noise, missing geometry, or heavy distortion. This visual evidence confirms that our reinforcement learning strategy learns a rational and efficient selection mechanism, successfully capturing the salient degradations that naturally draw human visual attention.

C. Discussion and limitations

R3-PCQA demonstrates robust performance across diverse scenarios. Notably, by sampling a fixed number of points via KNN, our method maintains a consistent structural input, making it inherently robust to varying point cloud densities. Furthermore, our reinforcement-learning-based single subcloud selection is grounded in visual perception theories [9, 10, 24, 26], which state that severe local degradations dominate overall perceived quality. While this selective approach is more resource efficient than processing the entire point cloud, we acknowledge that the overall framework remains computationally intensive due to multiview 2D rendering. Thus, R3-PCQA is currently positioned as a high-accuracy offline quality assessment metric, with optimizing inference speed left for future work.

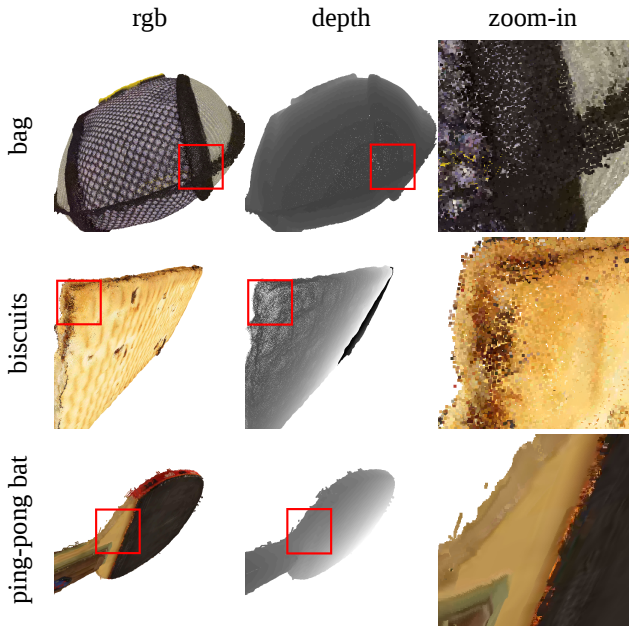

Figure 9. Qualitative results of the QSS. The red boxes indicate 2D regions corresponding to the selected 3D subclouds, while the zoomed-in views highlight severe localized degradations.