

Omni-scene Perception-oriented Point Cloud Geometry Enhancement for Coordinate Quantization

Wang Liu¹, Wei Gao^{1,2*}

¹Guangdong Provincial Key Laboratory of Ultra High Definition Immersive Media Technology, Shenzhen Graduate School, Peking University, China, ²Peng Cheng Laboratory, China

lw.pkac@gmail.com, gaowei262@pku.edu.cn

Abstract

Information quantization has been widely adopted in multimedia content, such as images, videos, and point clouds. The goal of information quantization is to achieve efficient storage and transmission by reducing data precision or redundancy. However, the information distortion caused by quantization will lead to the degradation of signal fidelity and the performance of downstream tasks. This paper focuses on the geometry quantization distortion of point clouds and proposes a unified learning-based quality enhancement framework for omni-scene point clouds. Based on the characteristics of geometry quantization distortion, we analyze and find that existing upsampling methods are not competitive in dealing with point reduction and geometry displacement simultaneously caused by coordinate quantization. Therefore, we design a general rooting-growing-pruning paradigm to efficiently perceive the geometry feature of quantized point clouds and improve the quality significantly. In addition, a novel loss constraint term related to the quantization step parameter is proposed to further improve quality and accelerate model convergence. To the best of our knowledge, this is the first unified quality enhancement framework for object and scene point clouds with coordinate quantization. Extensive experiments verify the superiority of the proposed method on multi-scale point clouds with different levels of quantization distortion, including object (ModelNet40, 8iVFB) and scene (KITTI). In particular, the enhanced point clouds improve the performance of downstream analysis tasks, including classification and 3D object detection.

1. Introduction

As a preprocessing process of sensor data, information quantization can achieve data precision alignment and reduce information redundancy. The former realizes the inter-

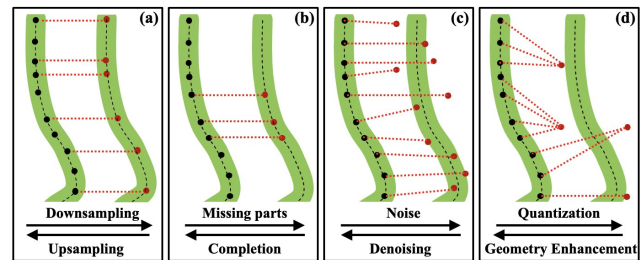


Figure 1. Simple visualization of common point cloud distortion types. It includes: (a) downsampling, (b) partial incompleteness, (c) noise, and (d) geometry quantization distortion (point reduction and coordinate offset). Their inverse processes lead to mainstream quality enhancement tasks, including upsampling, completion, denoising, and geometry enhancement. The black and red points represent the raw positions and the points with coordinate shifts caused by distortion, respectively. The dotted line depicts the latent surface, while the green band shows the area around it.

action of heterogeneous sensor data through precision conversion. The latter realizes efficient storage, transmission and calculation by reducing data precision. With the development of sensor technology, multimedia information is becoming more and more abundant, such as images, videos, audio, and point clouds. To reduce the burden of massive data calculation and interaction, information quantization converts continuous or high-precision data into discrete or low-precision representation. On the one hand, information quantization technology is concentrated in multimedia information compression algorithms. For example, the Motion Picture Experts Group (MPEG) [35] proposed a series of codec standards for images, videos and point clouds. On the other hand, the rapid development of large multimodal foundation models has posed great challenges to parameter storage and transmission. Quantization technology for model parameters alleviates this burden as much as possible while maintaining model performance, especially for lightweight deployment on mobile terminals.

As the basic and mature point cloud compression solution in MPEG, the geometry compression algorithm in the geometry-based point cloud compression standard (G-PCC) [34] directly adopts quantization technology to compress

*Corresponding author.

geometry information. However, the information distortion caused by quantization inevitably reduces the signal fidelity and the performance of downstream analysis tasks. This paper will study the problem of point clouds quantization distortion in G-PCC and improve the geometry quality in a post-processing manner.

To intuitively compare the similarities and differences between the quantization distortion and common geometry distortion types in point clouds, Fig. 1 briefly describes the above distortion process. As shown in sub-figure (d), geometry compression in G-PCC involves uniform quantization with a fixed quantization step parameter and a nonlinear mapping function. The quantization process causes both point reduction and geometry displacement, which shows a complex mapping relationship. In contrast, other common geometry distortions approximate a specific type of quality degradation. Therefore, existing common point cloud enhancement methods cannot be directly migrated to point clouds with the quantization distortion.

Point cloud enhancement methods for quantization distortion are mainly divided into three types: geometry distortion-oriented [10, 22], attribute-oriented distortion [23, 36], and hybrid distortion-oriented [47]. According to data organization manner, existing methods can also be divided into two types: point-based methods [49] and methods based on sparse convolution [2, 22, 47] or voxelization [10]. The former mainly works on small-scale object point clouds, while the latter serves large-scale object point clouds. However, existing methods are limited to independent processing strategies for point clouds with different scales. Moreover, these methods rarely consider scene point clouds with high sparsity.

Based on the defects in existing methods and the characteristics analysis for coordinate quantization distortion, we attempt to propose a unifying enhancement paradigm for coordinate quantization distortion. Specifically, we design a general rooting-growing-pruning mechanism to perceive the omni-scene point cloud features and semantic features. Our contributions are summarized as follows:

- We propose the first unified quality enhancement framework for point cloud with coordinate quantization distortion, which can deal with small-scale and large-scale point cloud simultaneously.
- Different from the loss function selected for common point cloud enhancement tasks, we design a novel loss function related to the quantization step parameter based on the characteristics of point cloud geometry distortion.
- Our proposed unified paradigm achieves state-of-the-art performance on a wide range of point cloud datasets and demonstrates excellent stability and generalization.
- The proposed method demonstrates omni-scene semantic feature perception capabilities, which ensures downstream tasks performance of the enhanced point cloud.

2. Related Work

Point Cloud Upsampling. Point cloud upsampling aims to increase the number of points in the sparse point cloud. Point cloud upsampling methods can be classified as optimization-based methods [3, 16] and deep learning-based methods. The latter can be summarized as point-based [12, 19, 21, 24, 28, 44], graph convolution-based [27], transformer-based [29, 32], sparse convolution-based [2], adversarial generative network-based [18], gradient field-based [15, 20] and diffusion-based [30]. Upsampling methods slightly improve the quality of point clouds with quantization distortion by increasing the number of points. However, it is subject to the assumption that the geometry structure is preserved, i.e. the exact coordinates.

Point Cloud Completion. The process of point cloud completion is to reconstruct a complete high-quality point cloud as much as possible based on part of the point cloud information. Recent years have witnessed the rapid development of point cloud completion methods based on deep learning [11, 38, 43, 45, 46, 48, 50]. Whether existing completion methods are capable of addressing quantization distortion will be discussed for the first time in this paper.

Point Cloud Denoising. Noise and outliers inevitably appear in the point cloud collection and transmission process, which will have a negative influence on downstream point cloud analysis tasks. Existing learning-based denoising methods can be roughly divided into offset estimation-based methods [7, 31], gradient field method [26], and diffusion method [40]. Due to the lack of prior information on the degree of quantization distortion, it is difficult for existing denoising methods to perform geometry quality enhancement.

Point cloud enhancement for quantization. There are two main types of point cloud geometry enhancement methods for coordinate quantization: traditional methods [4] and deep learning-based methods [1, 2, 10, 22, 47]. Borges et al. [4] first considered geometry quantization distortion removal as a special point cloud upsampling process, which effectively increases the number of points by building a lookup table. However, building a lookup table is an extremely time-consuming process. Deep learning-based methods have emerged currently and achieved remarkable performance and faster inference speed. Akhtar et al. [1] used a multi-scale sparse convolution framework to transform geometry enhancement into a probability prediction problem for neighborhood voxels. Fan et al. [10] adopted multi-scale 3D convolution with cross-entropy loss to adaptively generate enough effective points. However, it is also time-consuming for point-based methods to perform the spatial block partition process on sparse point clouds. Liu et al. [22] developed a coordinate offset prediction manner rather than point generation.

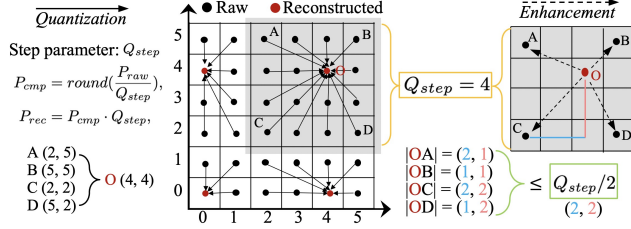


Figure 2. A toy example on the correlation between quantization step parameter Q_{step} and coordinate offset. $r2 : Q_{step} = 4$.

3. Motivation

3.1. Analysis of Geometry Quantization Distortion

Geometry information compression in the G-PCC framework [34] adopts the operator $\text{round}(\cdot)$ for coordinate quantization. The quantization and reconstruction process are defined as follows:

$$\begin{aligned} P_{cmp} &= \text{round}\left(\frac{P_{raw}}{Q_{step}}\right), \\ P_{rec} &= P_{cmp} \cdot Q_{step}, \end{aligned} \quad (1)$$

where P_{raw} , P_{cmp} , and P_{rec} represent the raw coordinate, quantized coordinate, and reconstructed coordinate, respectively. The degree of quantization distortion is controlled by the quantization step parameter Q_{step} . As shown in Fig. 2, it can be found that the absolute value of offset on each coordinate component from the reconstructed point (red point) to the raw point (black point) does not exceed half of the current quantization step parameter.

The above phenomenon describes the 2D coordinate offset, which can be extended to 3D point clouds. This prior information will be involved in the geometry quality enhancement process in the form of an additional loss constraint function in our work.

3.2. Rooting-Growing-Pruning Mechanism

Existing point cloud geometry quantization distortion removal methods [2, 47] generally follow the point cloud generation based on sparse convolution. However, the point cloud generation produced by the occupancy prediction process based on sparse convolution or voxelization often produces outliers, which ignore the continuous spatial structure. Recent work [22] has considered adding geometry offset learning to generation processes to alleviate this problem. However, the subjective results from Fig. 5 show that the existing enhancement paradigms still produce point clouds with geometry discontinuity. In other words, the discontinuities in the generated point cloud are propagated to the subsequent processes. Although some upsampling methods cover similar hierarchical prediction, they are still not compatible in removing geometry quantization distortion. The key is that these methods do not take into account the characteristics of quantization distortion: latent surface structure destruction, point reduction and coordinate offset.

Based on the above analysis, we decompose the quantization distortion removal task into three steps: rooting step, growing step, and pruning step. The reconstructed point cloud with geometry displacement deviates from the latent point cloud surface, which should be mitigated first in the rooting step. Subsequent experiments confirm that the rooting process is indeed the key operation to be considered. Then, the growing process is designed to eliminate the point number reduction.

A simplified explanation of our proposed rooting-growing-pruning mechanism is shown in Fig. 3. First, key points close to the underlying surface of the raw point cloud are generated, which is denoted as the rooting process. Second, the growing process refers to generating a large number of candidate points around the key points. Finally, candidate points are quantitatively selected and the corresponding geometry positions are corrected during the pruning process. The proposed mechanism is capable of perceiving the quantization distortion of omni-scene point clouds and improving the geometry quality.

4. Methods

Problem Definition. Given the raw point cloud $P_{raw}^{N \times 3} = \{p_i \mid 1 \leq i \leq N\}$ and the reconstructed point cloud with quantization distortion $P_{rec}^{S \times 3} = \{p_k \mid 1 \leq k \leq S, 1 \leq S \leq N\}$, where N and S represent the number of points in $P_{raw}^{N \times 3}$ and $P_{rec}^{S \times 3}$, respectively. In general, it satisfies $P_{rec}^{S \times 3} \not\subseteq P_{raw}^{N \times 3}$ and $P_{rec}^{S \times 3} \cap P_{raw}^{N \times 3} \neq \emptyset$. The goal of the geometry artifact removal task is to generate an enhanced point cloud $P_{pre}^{N \times 3}$ with high-quality to approximate the raw point cloud $P_{raw}^{N \times 3}$.

4.1. Framework

4.1.1. Small-scale point cloud

The number of points in the compressed small-scale point cloud is inconsistent and less. For convenience in training, the compressed point cloud is replicated to achieve the same input tensor size. In other words, existing points are copied directly to reach the target number before the rooting step ($S = N$). Therefore, the increase factor for the small-scale point cloud in the growing process is equal to 1, as shown in Fig. 3.

It should be noted that the highlight of our work is to propose a simple and unified enhancement paradigm rather than some exquisite and sophisticated feature extraction module design. Each stage consists of several feature extraction layers and geometry offset prediction layers. For the point cloud feature extractor, we directly refer to the global feature extraction module in [46] and the local feature extraction module in [14]. The local feature extraction module mainly introduces the point transformer unit, which is widely used in point cloud processing tasks.

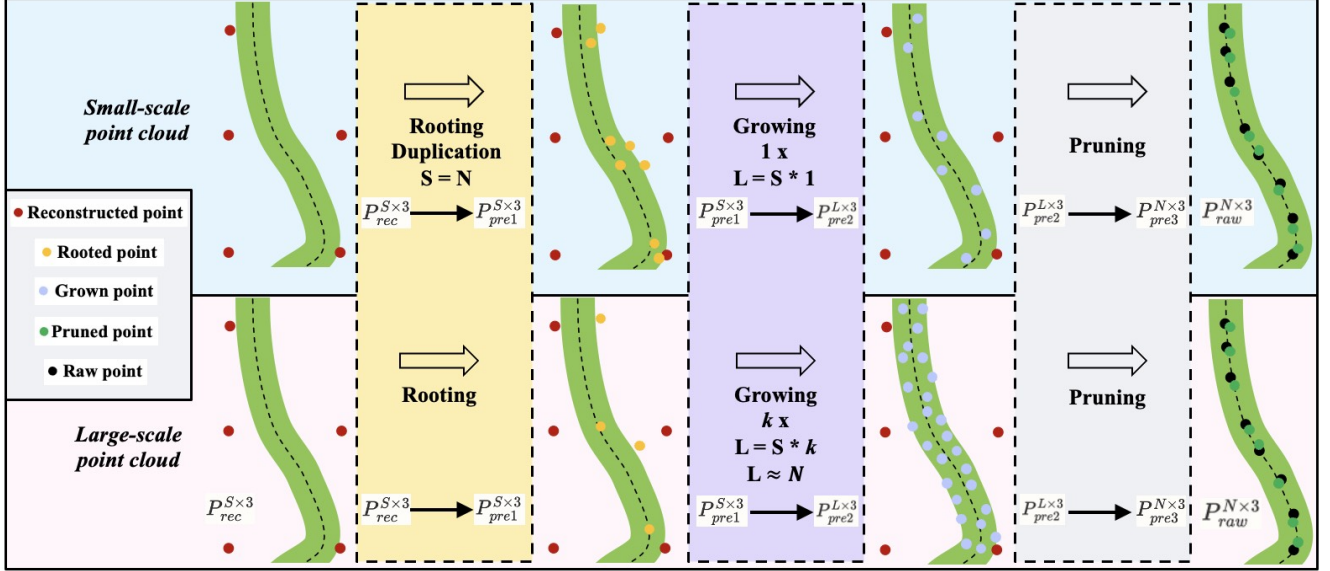


Figure 3. The pipeline of the proposed unified enhancement framework on small-scale and large-scale point clouds. In general, the rooting step forces the reconstructed point cloud to be close to the surface of the raw point cloud. The growing step can significantly make up for the number of points. The pruning step refines the rough results of the first two steps. All steps contribute to the quality enhancement of the compressed point cloud.

Our proposed geometry quality enhancement network for small-scale point clouds mainly includes three subnets: rooting subnet $r(\cdot)$, growing subnet $g(\cdot)$, and pruning subnet $p(\cdot)$. The rooting subnet $r(\cdot)$ consists of a global feature extraction (GFE) module and a multi-layer perceptron (MLP) layer. The relevant process is as follows:

$$\begin{aligned} F_g &= GFE(P_{rec}^{S \times 3}), \\ O_1^{S \times 3} &= MLP(F_g), \\ P_{pre1}^{S \times 3} &= O_1^{S \times 3} + P_{rec}^{S \times 3}. \end{aligned} \quad (2)$$

It should be noted that the rooting process does not involve local feature extraction because the geometry distribution of the reconstructed point cloud $P_{rec}^{S \times 3}$ has changed. The growing subnet $g(\cdot)$ takes the global feature F_g and $O_1^{S \times 3}$ as input and predicts $P_{pre2}^{S \times 3}$ through an MLP layer:

$$\begin{aligned} O_2^{S \times 3} &= MLP(F_g, O_1^{S \times 3}), \\ P_{pre2}^{S \times 3} &= O_1^{S \times 3} + O_2^{S \times 3} + P_{rec}^{S \times 3}. \end{aligned} \quad (3)$$

The pruning subnet $p(\cdot)$ is designed to further refine the $P_{pre2}^{S \times 3}$, and its structure is described as follows:

$$\begin{aligned} F_l^1 &= GLFE_1(P_{rec}^{S \times 3}, P_{pre2}^{S \times 3}, F_g), \\ F_l^2 &= GLFE_2(P_{pre2}^{S \times 3}, F_g), \\ F_l^3 &= GLFE_3(P_{pre2}^{S \times 3}, F_g, F_l^1, F_l^2), \\ O_3^{S \times 3} &= MLP(F_g, F_l^3, P_{pre2}^{S \times 3}), \\ P_{pre3}^{S \times 3} &= O_1^{S \times 3} + O_2^{S \times 3} + O_3^{S \times 3} + P_{rec}^{S \times 3}, \end{aligned} \quad (4)$$

where both $GLFE_1$ and $GLFE_2$ represent global and local feature extraction (GLFE) modules, while $GLFE_3$ adds an additional point cloud attention (GLTFE) module. $O_1^{S \times 3}, O_2^{S \times 3}, O_3^{S \times 3}$ record the geometry offset

predictions learned in the three stages, respectively. $P_{pre1}^{S \times 3}, P_{pre2}^{S \times 3}, P_{pre3}^{S \times 3}$ represent the enhanced point clouds in the three stages, respectively. The detailed structures of all GFE, GLFE and GLTFE modules are arranged in the supplementary material.

4.1.2. Large-scale point cloud

Different from the point-based organization method used in small-scale point clouds (1k–4k), the number of points in dense large-scale point clouds exceeds 1000k. Therefore, we adopt the efficient Minkowski sparse convolution [6] framework to process large-scale point clouds without patch partition. It should be noted that we still strictly follow the rooting-growing-pruning mechanism to build the network.

Our proposed geometry quality enhancement network for large-scale point clouds mainly includes three subnets: rooting subnet $r(\cdot)$, growing subnet $g(\cdot)$, and pruning subnet $p(\cdot)$. The rooting subnet $r(\cdot)$ adopts a U-Net structure with sparse convolutions similar to [23]. For large-scale point cloud, it is also necessary to extract global feature quickly and efficiently. Hence, we directly refer to our pre-trained model for the small-scale point cloud (=1024 points). To further extract the global feature of the large-scale point cloud, we randomly sample a fixed number of points (1024) on the reconstructed point cloud. Applying the pre-trained model on the small-scale point cloud to the sampled large-scale point cloud can make up for the shortcomings of sparse convolution in modeling spatial relationships. A sparse predicted point cloud $P_{pre1}^{S \times 3}$ is obtained in rooting process, which has the same number of points as the reconstructed point cloud $P_{rec}^{S \times 3}$. In second process, the growing subnet $g(\cdot)$ introduces a sparse convolution layer

with grid point generation to generate a target number of new points $P_{pre2}^{N \times 3}$ around the predicted point $P_{pre1}^{S \times 3}$. Finally, the pruning subnet $p(\cdot)$ with several sparse convolution layers is used to further refine $P_{pre2}^{N \times 3}$.

The relevant process is described as follows:

$$\begin{aligned} F_g &= GFE(P_{rec}^{S \times 3}), \\ O_1^{S \times 3} &= SUnet_1(P_{rec}^{S \times 3}, F_g), \\ P_{pre1}^{S \times 3} &= O_1^{S \times 3} + P_{rec}^{S \times 3}, \\ P_{pre2}^{N \times 3} &= SUnet_2(P_{pre1}^{S \times 3}), \\ P_{pre3}^{S \times 3} &= Snet(P_{pre2}^{N \times 3}) + P_{pre2}^{N \times 3}, \end{aligned} \quad (5)$$

where $SUnet_1$ represents a U-Net module with three sparse downsampling convolution layers and three sparse upsampling convolution layers, while $SUnet_2$ contains an additional expansion sparse convolution layer. The detailed structures of all modules are arranged in the supplementary material.

4.2. Loss Function

According to the point cloud data organization form, point cloud geometry quality enhancement methods based on deep learning need to choose different loss functions for training. For the voxel-based methods, binary cross-entropy (BCE) loss is adopted. For the point-based methods, Chamfer Distance (CD) loss [9] and Earth Mover's Distance (EMD) loss [33] are chosen. In addition, CD and EMD loss are also recommended as objective assessment metrics to measure the geometry difference. For voxel-based methods, the binary cross-entropy (BCE) loss is adopted as follows:

$$\mathcal{L}_p = -\frac{1}{N} \sum_{n=0}^N [c_n \log(q_n) + (1 - c_n) \log(1 - q_n)], \quad (6)$$

where N is the number of all voxels, c_n is the occupancy status of n -th voxel of raw point cloud P_r and q_n is the corresponding predicted occupancy probability of enhanced point cloud P_e . For the point-based methods, CD loss and EMD loss are defined as follows:

$$\mathcal{L}_c = \frac{1}{|P_r|} \sum_{x \in P_r} \min_{y \in P_e} \|x - y\|_2 + \frac{1}{|P_e|} \sum_{y \in P_e} \min_{x \in P_r} \|y - x\|_2, \quad (7)$$

$$\mathcal{L}_e = \min_{\phi: P_e \rightarrow P_r} \frac{1}{N} \sum_{x \in P_e} \|x - \phi(x)\|_2, \quad (8)$$

where $\phi: P_e \rightarrow P_r$ is a bijection. Moreover, CD and EMD are also chosen to measure the geometry difference.

As described in 3.1, prior information about the geometry is established as a loss function to constrain the offset prediction in our work. Given an input reconstructed point cloud $P_{rec}^{N \times 3}$ and the corresponding predicted coordinate offset $O_{rec}^{N \times 3}$, $O_{rec}^{N \times 3}$ is collapsed along the coordinate dimension into $O_{rec}^{3N \times 1}$, the loss function \mathcal{L}_q regarding the quantization step parameter Q_{step} is defined as follows:

$$\mathcal{L}_q = \max(\|O_{rec}^{3N \times 1}\|_2 - Q_{step}/2, 0). \quad (9)$$

In addition to the loss constraint function on the quantization step parameter, we also impose a repulsion loss on the enhanced point cloud. As we know, the geometry coordinates of the raw point cloud and the reconstructed point cloud are voxelized integer data without repeating points. Therefore, the absolute value of the minimum distance between the points of the enhanced point cloud exceeds one unit, which enables a uniform distribution of the point cloud. The repulsion loss \mathcal{L}_r is defined as follows:

$$\mathcal{L}_r = \max(1 - \left\| P_{pre}^{N \times 3} - \hat{P}_{pre, k=1}^{N \times 3} \right\|_2, 0), \quad (10)$$

where $P_{pre}^{N \times 3} = P_{rec}^{N \times 3} + O_{rec}^{N \times 3}$, $\hat{P}_{pre, k=1}^{N \times 3}$ represents the set of $K=1$ nearest neighbor points of $P_{pre}^{N \times 3}$.

Loss for small-scale point cloud. As previously analyzed, the goal of the first rooting process is to make the reconstructed point cloud close to the latent surface of the raw point cloud. Hence, the single CD loss $\mathcal{L}_{c'}$ from the generated point cloud to the raw point cloud followed by two additional loss functions \mathcal{L}_q and \mathcal{L}_r is adopted in this stage. The total loss function is defined as follows:

$$\mathcal{L} = \mathcal{L}_{c'} + \lambda_1 \cdot \mathcal{L}_q + \lambda_2 \cdot \mathcal{L}_r \quad (11)$$

where the hyperparameters $\lambda_1 = 10$ and $\lambda_2 = 1 \times 10^{-2}$ serve as the weight of the latter two loss functions.

Different from the CD loss in the first stage, we choose EMD loss as the optimization function of geometry signal fidelity in the latter two stages. All stages will be jointly optimized in an end-to-end manner, and the result of the last stage represents the final output.

Loss for large-scale point cloud. For large-scale point clouds, we adopt different loss functions in three stages. Due to the gap between the sparse convolution-based and point-based point cloud data organization manner, joint optimization cannot be achieved. Therefore, we train each stage separately.

The loss of the first stage (rooting) is defined as follows:

$$\mathcal{L} = \lambda_1 \cdot \mathcal{L}_e + \mathcal{L}_q + \mathcal{L}_r \quad (12)$$

where the hyperparameters $\lambda_1 = 1 \times 10^{-2}$ serve as the weight of the first item. The loss functions in the second and last stages adopt BCE loss and CD loss, respectively.

5. Experiments

Extensive experiments are performed on a wide range of datasets (ModelNet40 [42], ShapeNet [5], ScanObject [39], 8iVFB[8], and KITTI [13]), which serve as common datasets for point cloud analysis and processing. At the same time, various quantization distortion levels ($r1: Q_{step} = 8, r2: Q_{step} = 4, r3: Q_{step} = 2$) are also considered in the comparative experiments. In addition to existing quantization distortion removal methods (PU-Dense

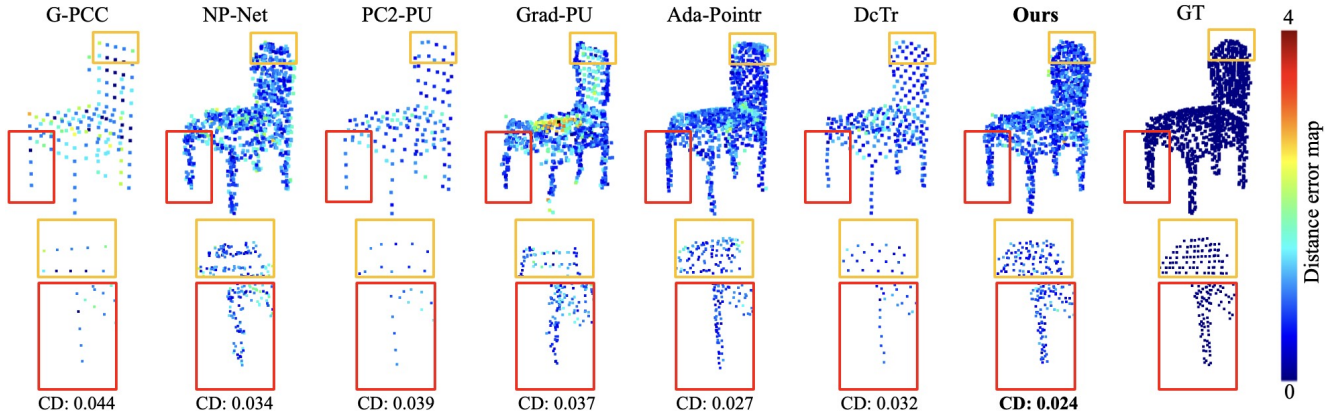


Figure 4. Subjective visualization results of different methods on small-scale object ModelNet40 [42] at $r_2 : Q_{step} = 4$. G-PCC: compressed point cloud with low quality. GT: raw point cloud with high quality.

[2], DGPP [10], Geo-Net [47], Grnet-ER[22]), we also extensively consider other mainstream point cloud enhancement methods, such as point cloud upsampling methods (PU-Net [44], NP-Net [12], PC2-PU [24], Grad-PU [15], RepKPU[32], PUDM [30]), point cloud completion methods (PCNet [46], Top-Net [38], GRNet [43], SeedFormer [50], Ada-Pointr [45], DcTr [11]) and point cloud denoising method (DMR-Net [25]). All the comparison baselines are retrained strictly according to the public official code and default settings.

The experiments are implemented on the PyTorch platform with Minkowski Engine [6]. Adam optimizer [17] is adopted to train on NVIDIA Tesla V100 GPU. For each set of experiments, all relevant **Datasets** and experimental **Settings** are explained in the supplementary materials.

Evaluation Metrics. According to the quantitative evaluation recommended by the common test conditions (CTC) provided by MPEG, the point-to-point (D1) distance based mean squared error peak signal-to-noise ratio (D1 PSNR) and point-to-plane (D2) distance based mean squared error peak signal-to-noise ratio (D2 PSNR) are adopted. Following the common point cloud enhancement tasks, Chamfer Distance (CD) is also considered in our experiments.

Results on small-scale point clouds. Table 1 and Table 2 show the quantitative results on ModelNet40 [42], ShapeNet [5], and ScanObject [39] with three most severe distortion levels. Overall, our method achieves the highest quality improvement on all objective evaluation metrics. It can be found that most point cloud upsampling methods and completion methods are not competent for point clouds with quantization distortion. The main reason is that these methods assume that the latent surface information of the input point cloud is not completely destroyed.

In addition to signal quality fidelity, we also considered the performance of the enhanced point cloud on downstream machine vision tasks, such as point cloud classification. The results in Table 3 show that the enhanced point

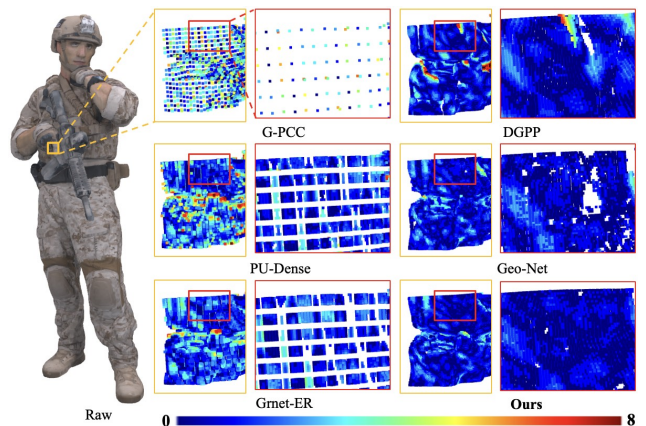


Figure 5. Subjective visualization results of different methods on large-scale object 8iVFB [8] at $r_1 : Q_{step} = 8$.

cloud obtained by our method achieves the highest classification accuracy. Fig. 4 presents the qualitative results of different methods on ModelNet40 [42] at $r_2 : Q_{step} = 4$. To reflect the superiority of our method more intuitively, the CD distance between the enhanced point cloud and the raw point cloud is converted into heat map. The closer to blue, the smaller the distance to the raw point cloud. The results show that the enhanced results of our method have better geometry integrity and distribution uniformity.

Results on large-scale point clouds. Table 4 shows the quantitative results on 8iVFB [8] with two most severe distortion levels ($r_1 : Q_{step} = 8, r_2 : Q_{step} = 4$). Overall, our method exhibits the highest quality improvement on all objective evaluation metrics. In particular, the quality improvement of our method is more significant under the most severe geometry distortion. Fig. 5 presents the qualitative results of different methods on 8iVFB [8]. The results show that our method is able to produce denser point clouds with better geometry integrity and distribution uniformity. The results of existing methods show obvious hole effects.

In addition to object point clouds, our proposed method is validated on LiDAR scene point clouds. To the best of

Table 1. Overall performance comparison for two objective evaluation metrics $\text{CD} \downarrow (\times 10^{-2}) / \text{D1 PSNR} \uparrow$ (dB) over three datasets ModelNet40 [42], ShapeNet [5], and ScanObject [39] at $r1 : Q_{step} = 8, r2 : Q_{step} = 4, r3 : Q_{step} = 2$. To verify the generalization, we choose the training set (9840 samples) of ModelNet40 [42] for training and evaluate on the test set (2468 [42], 2874 [5], and 518 [39] samples) of three datasets, respectively.

Methods & Datasets Quantization level	ModelNet40[42]			ShapeNet[5]			ScanObject[39]		
	r3	r2	r1	r3	r2	r1	r3	r2	r1
GPCC	3.22 / 39.16	5.24 / 34.38	8.96 / 28.91	3.05 / 39.17	5.08 / 34.45	8.78 / 29.09	3.14 / 39.15	5.20 / 34.38	9.18 / 28.82
PU-Net [44]	4.01 / 37.01	5.06 / 35.15	7.47 / 30.74	4.64 / 34.60	5.84 / 33.76	8.49 / 30.01	5.00 / 32.97	6.50 / 32.37	9.88 / 28.96
NP-Net [12]	2.98 / 40.56	4.11 / 37.57	5.91 / 33.78	2.73 / 41.32	3.81 / 38.04	5.68 / 33.56	2.87 / 40.66	4.03 / 37.94	6.14 / 32.95
PC2-PU [24]	3.23 / 39.11	4.38 / 36.37	6.70 / 31.82	2.83 / 40.65	4.51 / 36.00	6.71 / 31.95	2.98 / 39.88	4.55 / 35.73	6.86 / 31.82
Grad-PU [15]	3.85 / 36.69	4.74 / 34.96	7.48 / 30.91	3.25 / 38.46	4.41 / 35.78	7.54 / 30.79	3.45 / 37.48	4.45 / 35.55	7.74 / 30.94
RepKPU [32]	2.73 / 40.31	3.63 / 38.01	6.58 / 32.46	2.34 / 40.35	3.42 / 38.07	6.73 / 32.45	2.52 / 40.54	3.42 / 38.03	6.68 / 31.97
PUDM [30]	3.53 / 38.97	4.45 / 36.77	6.65 / 32.67	3.77 / 38.61	4.61 / 35.98	7.12 / 31.28	3.84 / 38.45	4.82 / 35.76	7.29 / 31.21
PCNet [46]	3.94 / 37.83	4.27 / 37.12	4.91 / 35.92	3.07 / 39.47	4.11 / 37.12	6.55 / 32.71	3.28 / 39.10	4.26 / 36.96	6.78 / 32.62
Top-Net [38]	4.16 / 37.32	4.41 / 36.75	5.01 / 35.59	4.46 / 36.62	4.71 / 35.87	5.51 / 34.34	5.78 / 33.89	6.04 / 33.35	6.96 / 32.00
GRNet [43]	3.03 / 39.09	4.00 / 37.33	4.77 / 35.55	2.73 / 40.43	4.16 / 37.05	5.24 / 34.41	2.81 / 40.13	3.62 / 38.44	5.08 / 35.18
SeedFormer [50]	3.40 / 39.60	4.30 / 36.80	6.32 / 33.09	3.15 / 39.73	4.05 / 37.61	6.78 / 32.86	3.23 / 39.86	4.44 / 36.55	7.87 / 31.60
Ada-Pointr [45]	3.54 / 39.05	3.95 / 37.98	4.74 / 36.11	3.39 / 39.25	3.95 / 37.68	5.13 / 34.82	3.70 / 38.29	4.35 / 36.75	6.13 / 33.05
DcTr [11]	2.84 / 40.31	3.98 / 37.62	4.77 / 35.81	2.56 / 41.07	3.90 / 37.70	5.03 / 35.18	2.52 / 41.92	3.96 / 37.69	6.10 / 33.20
DMR-Net [25]	4.19 / 36.63	5.58 / 33.12	8.75 / 27.78	4.37 / 35.63	9.18 / 27.06	6.24 / 31.19	5.02 / 33.95	6.59 / 31.57	9.36 / 27.64
PU-Dense [2]	2.55 / 40.21	--	--	2.25 / 41.21	--	--	2.45 / 40.96	--	--
DGPP [10]	2.80 / 38.23	5.09 / 33.05	6.71 / 30.82	2.81 / 39.55	4.28 / 35.26	6.12 / 32.28	2.70 / 39.24	4.78 / 33.92	7.52 / 30.46
Geo-Net [47]	2.59 / 39.46	--	--	2.27 / 40.93	--	--	2.44 / 40.25	--	--
Ours	2.32 / 43.02	3.47 / 39.40	4.47 / 36.86	2.25 / 42.82	3.41 / 39.13	4.76 / 35.99	2.36 / 42.83	3.54 / 39.03	5.43 / 34.58

Table 2. Overall performance comparison for objective evaluation metric $\text{D2 PSNR} \uparrow$ (dB) / over two datasets ModelNet40 [42] and ScanObject [39]. It also includes model parameter size: S(M) / computational complexity(MACs): M(G) / inference time: T (ms) per frame.

Dataset Level	ModelNet40[42]			ScanObject[39]			S / M / T
	r3	r2	r1	r3	r2	r1	
NP-Net	45.46	41.40	35.97	45.45	40.83	34.78	4.2 / 26.9 / 20
PC2-PU	45.68	41.95	38.54	45.75	41.69	36.82	0.4 / 3.5 / 27
RepKPU	45.06	41.25	35.15	44.84	40.86	34.22	1.0 / 1.2 / 75
PUDM	42.93	40.42	37.36	42.68	39.81	36.63	11.5 / 8.9 / 2319
GRNet	46.64	42.45	39.67	46.52	42.41	36.79	76.7 / 21.7 / 8
SeedFormer	44.32	42.26	38.72	43.84	40.29	34.63	2.6 / 7.0 / 31
Ada-Pointr	44.18	42.17	39.49	42.07	39.59	35.08	10.2 / 3.4 / 19
DcTr	46.97	43.01	40.17	47.19	42.26	35.90	54.1 / 13.8 / 46
DMR-Net	40.55	39.98	38.37	36.67	36.43	36.02	2.1 / 2.2 / 19
DGPP	46.19	42.09	39.12	46.41	41.25	36.08	5.6 / 19.4 / 12
Ours	47.81	43.57	40.34	47.27	42.53	36.95	11.1 / 7.6 / 23

our knowledge, this is the first attempt to directly perform quality enhancement on scene point clouds with quantization distortion. Table 5 provides the quality gains of the proposed method on KITTI [13]. Furthermore, it is also verified by the performance improvement on the 3D object detection task in Table 6.

6. Ablation study

Loss. Based on the inherent characteristics of point cloud geometry quantization distortion, we design a regularization loss function L_q regarding the quantization step parameter Q_{step} to constrain model convergence. Table 7 shows related ablation experiments on three types of datasets under the most severe geometry quantization distortion. Ex-

Table 3. The **accuracy** \uparrow (%) comparison of enhanced point clouds on classification task. The pre-trained DGCNN [41] is performed on raw, compressed (G-PCC) and enhanced point clouds in our experiments.

Dataset Level	ModelNet40[42]			ShapeNet[5]			ScanObject[39]		
	r3	r2	r1	r3	r2	r1	r3	r2	r1
Raw	91.69			98.71			74.52		
GPCC	89.87	55.11	9.44	98.57	92.03	55.29	72.81	55.59	25.47
PU-Net	66.73	37.16	4.25	85.11	69.69	13.98	55.76	42.85	11.88
NP-Net	91.04	82.53	24.15	98.36	98.19	96.03	71.77	62.48	10.84
PC2-PU	85.69	71.79	9.07	98.64	96.27	75.08	70.05	65.06	36.31
Grad-PU	62.93	48.98	18.31	97.84	90.40	60.16	66.95	63.85	41.13
RepKPU	82.69	76.82	13.53	98.19	96.38	76.20	59.72	55.08	34.58
PUDM	89.42	77.35	18.15	98.01	89.94	55.98	69.19	63.85	29.25
PCNet	78.36	75.32	68.56	96.59	92.69	90.88	46.29	38.9	34.42
Top-Net	74.47	74.35	70.87	97.11	95.48	94.25	46.64	34.93	34.94
GRNet	84.40	75.56	59.11	97.94	90.68	91.57	67.81	66.09	50.94
SeedFormer	90.15	78.77	35.77	98.50	97.77	82.11	70.57	46.99	14.11
Ada-Pointr	86.26	82.94	70.95	97.91	96.59	94.11	67.64	61.27	44.06
DcTr	89.71	75.28	62.43	96.93	87.2	91.34	72.46	55.77	43.55
DMR-Net	69.32	25.08	7.90	96.38	68.02	21.53	52.84	42.17	23.92
PU-Dense	89.26	--	--	98.56	--	--	72.28	--	--
DGPP	83.91	38.25	26.78	94.95	95.76	87.61	69.02	57.14	40.96
Geo-Net	87.43	--	--	98.22	--	--	71.42	--	--
Ours	91.13	89.58	78.57	98.68	98.36	96.69	74.01	69.02	50.77

perimental results show that the quality gain value of our method decreases without the supervision of L_q , which verifies the effectiveness of the loss function L_q we designed.

Paradigm. To establish a simple and effective point cloud geometry quantization distortion removal paradigm for both and small-scale object and large-scale object, we propose a unifying framework embedded with the rooting-growing-pruning process. Here, we perform ablation experiments

Table 4. Overall quality improvement performance comparison for three objective evaluation metrics $\Delta\mathbf{D1}$ PSNR \uparrow (dB) / $\Delta\mathbf{D2}$ PSNR \uparrow (dB) / $\Delta\mathbf{CD}\uparrow$ ($\times 10^{-2}$) over standard sequences 8iVFB [8] (*redandblack*, *soldier*) at $r1 : Q_{step} = 8, r2 : Q_{step} = 4$. It also includes running time (s) per frame and model parameter size (M) in the last three columns. Running time: data processing + inference.

Level Sequence	r1		r2		Time	Size
	<i>redandblack</i>	<i>soldier</i>	<i>redandblack</i>	<i>soldier</i>	<i>redandblack</i>	
DGPP [10]	8.18 / 6.98 / 4.47	8.70 / 7.05 / 4.52	9.95 / 7.67 / 2.57	10.09 / 7.83 / 2.60	18.75 + 2.90	5.59
PU-Dense [2]	6.26 / 4.03 / 3.72	6.21 / 3.98 / 3.72	10.98 / 9.65 / 2.68	11.56 / 9.89 / 2.71	3.88 + 1.06	13.17
Geo-Net [47]	7.80 / 8.46 / 4.55	7.34 / 8.64 / 4.55	8.13 / 8.82 / 2.52	7.79 / 9.12 / 2.55	3.68 + 1.83	0.72
Grnet-ER [22]	7.84 / 5.80 / 4.35	7.98 / 5.93 / 4.39	11.20 / 9.71 / 2.52	11.92 / 10.23 / 2.56	3.55 + 1.27	1.90
Ours	10.39 / 8.74 / 4.74	10.37 / 8.88 / 4.73	11.65 / 9.83 / 2.77	12.18 / 10.34 / 2.74	3.55 + 2.22	15.90

with each process, as shown in Table 8 and Table 9. Each stage in the rooting-growing-pruning process is abbreviated as R, G, and P, respectively. Experimental results show that compared to the removal of the pruning process, the reduction in performance gain caused by the removal of the rooting process is more serious (the first row). In other words, we validate the assertion that the rooting process we emphasize is the key to geometry quantization distortion removal.

Table 5. Quality improvement of our method for $\Delta\mathbf{D1}$ PSNR \uparrow (dB) / $\Delta\mathbf{D2}$ PSNR \uparrow (dB) / $\Delta\mathbf{CD}\uparrow$ at $r1$ over SemanticKITTI [13]. Trained on: *Sequence_00*.

<i>Sequence_01</i>	<i>Sequence_02</i>	<i>Sequence_03</i>	<i>Sequence_04</i>
2.82 / 1.14 / 82	3.26 / 3.41 / 109	2.67 / 0.82 / 74	3.01 / 2.23 / 96

Table 6. Overall performance on 3D object detection task. The pre-trained PointRCNN [37] is performed on raw, compressed and our enhanced LiDAR at $r1$ on KITTI(validation split set) [13]. The evaluation metric is Average Precision(AP: Easy / Moderate / Hard) with IoU threshold 0.7 for car and 0.5 for pedestrian.

Class	Raw	G-PCC	Ours
Car	88.9 / 78.6 / 77.4	86.7 / 75.5 / 70.1	88.2 / 76.9 / 72.7
Pedestrian	49.3 / 43.7 / 39.9	45.2 / 40.4 / 36.1	48.8 / 42.1 / 39.1

Table 7. Ablation experiments on loss function L_q . Quality improvement $\Delta\mathbf{CD}\uparrow$ ($\times 10^{-2}$) are calculated under the most severe distortion.

Dataset	ModelNet40	ShapeNet	ScanObject	<i>redandblack</i>	<i>soldier</i>
w/o L_q	4.38	3.85	3.49	4.62	4.59
Default	4.49	4.02	3.75	4.74	4.73

Table 8. Ablation experiments on rooting-growing-pruning paradigm. Overall quality improvement performance comparison for $\mathbf{CD}\downarrow$ ($\times 10^{-2}$) / $\mathbf{D1}$ PSNR \uparrow (dB) / $\mathbf{D2}$ PSNR \uparrow (dB) over ModelNet40 [42] at $r1 : Q_{step} = 8, r2 : Q_{step} = 4, r3 : Q_{step} = 2$.

Level	r1	r2	r3
w/o R	7.64 / 30.00 / 36.95	4.85 / 35.33 / 41.85	2.68 / 42.12 / 46.01
w/o G	4.54 / 36.71 / 40.02	3.89 / 38.31 / 42.18	2.58 / 42.38 / 46.94
w/o P	6.57 / 33.08 / 36.57	4.43 / 36.92 / 41.77	2.56 / 42.48 / 47.03
Default	4.47 / 36.86 / 40.34	3.47 / 39.40 / 43.57	2.32 / 43.02 / 47.81

Furthermore, we visualize the enhanced point clouds in the three stages and calculate the point-to-plane distance (D2 PSNR) in Fig. 6. We extracted the points of the wine glass on a plane to observe the geometry structure. It can be found that the point cloud obtained in the first rooting stage reaches the latent surface surrounding the raw point cloud.

Table 9. Ablation experiments on rooting-growing-pruning paradigm. Overall quality improvement performance comparison for $\Delta\mathbf{D1}$ PSNR \uparrow (dB) / $\Delta\mathbf{D2}$ PSNR \uparrow (dB) / $\Delta\mathbf{CD}\uparrow$ ($\times 10^{-2}$) over standard sequences 8iVFB [8] at $r1 : Q_{step} = 8$.

Sequence	<i>redandblack</i>	<i>soldier</i>
w/o R	8.07 / 5.96 / 4.43	8.24 / 6.31 / 4.48
w/o P	9.11 / 9.44 / 4.78	9.31 / 9.37 / 4.77
Default	10.39 / 8.74 / 4.74	10.37 / 8.88 / 4.73

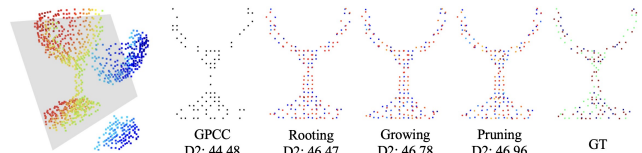


Figure 6. Subjective visualization results in three stages.

7. Conclusion

In this paper, we first analyze the characteristics of point cloud coordinate quantization distortion by comparing common distortion. After that, a simple unifying enhancement paradigm with rooting-growing-pruning mechanism is proposed to deal with point clouds with different scale and semantic scene. Extensive experiments verify the effectiveness and generalization of our method. In addition, the designed loss constraint term inspired by the geometry quantization distortion characteristics will become an effective tool for subsequent research. The performance of enhanced point clouds in downstream tasks is also discussed in depth.

Acknowledgments

This work was supported by The Major Key Project of PCL (PCL2024A02), Natural Science Foundation of China (62271013), Guangdong Provincial Key Laboratory of Ultra High Definition Immersive Media Technology (2024B1212010006), Guangdong Province Pearl River Talent Program (2021QN020708), Guangdong Basic and Applied Basic Research Foundation (2024A1515010155), Shenzhen Science and Technology Program (JCYJ20240813160202004, JCYJ20230807120808017), National Science and Technology Major Project (2020AAA0103501), Shenzhen Fundamental Research Program (GXWD20201231165807007-20200806163656003).

References

- [1] Anique Akhtar, Wen Gao, Xiang Zhang, Li Li, Zhu Li, and Shan Liu. Point cloud geometry prediction across spatial scale using deep learning. In *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, pages 70–73. IEEE, 2020. 2
- [2] Anique Akhtar, Zhu Li, Geert Van der Auwera, Li Li, and Jianle Chen. Pu-dense: Sparse tensor-based point cloud geometry upsampling. *IEEE Transactions on Image Processing*, 31:4133–4148, 2022. 2, 3, 6, 7, 8
- [3] Marc Alexa, Johannes Behr, Daniel Cohen-Or, Shachar Fleishman, David Levin, and Claudio T. Silva. Computing and rendering point set surfaces. *IEEE Transactions on visualization and computer graphics*, 9(1):3–15, 2003. 2
- [4] Tomás M Borges, Diogo C Garcia, and Ricardo L De Queiroz. Fractional super-resolution of voxelized point clouds. *IEEE Transactions on Image Processing*, 31:1380–1390, 2022. 2
- [5] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015. 5, 6, 7
- [6] Christopher Choy, JunYoung Gwak, and Silvio Savarese. 4d spatio-temporal convnets: Minkowski convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3075–3084, 2019. 4, 6
- [7] Dasith de Silva Edirimuni, Xuequan Lu, Zhiwen Shao, Gang Li, Antonio Robles-Kelly, and Ying He. Iterativepfn: True iterative point cloud filtering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 13530–13539, 2023. 2
- [8] E d’Eon, B Harrison, T Myers, and PA Chou. Iso/iec jtc1/sc29 joint wg11/wg1 (mpeg/jpeg) input document wg11m40059/wg1m74006; 8i voxelized full bodies-a voxelized point cloud dataset. *MPEG: Geneva, Switzerland*, 2017. 5, 6, 8
- [9] Haoqiang Fan, Hao Su, and Leonidas J Guibas. A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 605–613, 2017. 5
- [10] Xiaoqing Fan, Ge Li, Dingquan Li, Yurui Ren, Wei Gao, and Thomas H Li. Deep geometry post-processing for decompressed point clouds. In *2022 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2022. 2, 6, 7, 8
- [11] Ben Fei, Weidong Yang, Lipeng Ma, and Wen-Ming Chen. Dctr: Noise-robust point cloud completion by dual-channel transformer with cross-attention. *Pattern Recognition*, 133: 109051, 2023. 2, 6, 7
- [12] Wanquan Feng, Jin Li, Hongrui Cai, Xiaonan Luo, and Juyong Zhang. Neural points: Point cloud representation with neural fields for arbitrary upsampling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18633–18642, 2022. 2, 6, 7
- [13] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012. 5, 7, 8
- [14] Meng-Hao Guo, Jun-Xiong Cai, Zheng-Ning Liu, Tai-Jiang Mu, Ralph R Martin, and Shi-Min Hu. Pct: Point cloud transformer. *Computational Visual Media*, 7:187–199, 2021. 3
- [15] Yun He, Danhang Tang, Yinda Zhang, Xiangyang Xue, and Yanwei Fu. Grad-pu: Arbitrary-scale point cloud upsampling via gradient descent with learned distance functions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5354–5363, 2023. 2, 6, 7
- [16] Hui Huang, Shihao Wu, Minglun Gong, Daniel Cohen-Or, Uri Ascher, and Hao Zhang. Edge-aware point set resampling. *ACM transactions on graphics (TOG)*, 32(1):1–12, 2013. 2
- [17] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6
- [18] Ruihui Li, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. Pu-gan: a point cloud upsampling adversarial network. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7203–7212, 2019. 2
- [19] Ruihui Li, Xianzhi Li, Pheng-Ann Heng, and Chi-Wing Fu. Point cloud upsampling via disentangled refinement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 344–353, 2021. 2
- [20] Shujuan Li, Junsheng Zhou, Baorui Ma, Yu-Shen Liu, and Zhizhong Han. Learning continuous implicit field with local distance indicator for arbitrary-scale point cloud upsampling. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3181–3189, 2024. 2
- [21] Zhuangzi Li, Ge Li, Thomas H Li, Shan Liu, and Wei Gao. Semantic point cloud upsampling. *IEEE Transactions on Multimedia*, 2022. 2
- [22] Gexin Liu, Ruixiang Xue, Jiabin Li, Dandan Ding, and Zhan Ma. Grnet: Geometry restoration for g-pcc compressed point clouds using auxiliary density signaling. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 2, 3, 6, 8
- [23] Wang Liu, Wei Gao, and Xingming Mu. Fast inter-frame motion prediction for compressed dynamic point cloud attribute enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3720–3728, 2024. 2, 4
- [24] Chen Long, WenXiao Zhang, Ruihui Li, Hao Wang, Zhen Dong, and Bisheng Yang. Pc2-pu: Patch correlation and point correlation for effective point cloud upsampling. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 2191–2201, 2022. 2, 6, 7
- [25] Shitong Luo and Wei Hu. Differentiable manifold reconstruction for point cloud denoising. In *Proceedings of the 28th ACM international conference on multimedia*, pages 1330–1338, 2020. 6, 7
- [26] Shitong Luo and Wei Hu. Score-based point cloud denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4583–4592, 2021. 2

- [27] Guocheng Qian, Abdullellah Abualshour, Guohao Li, Ali Thabet, and Bernard Ghanem. Pu-gcn: Point cloud upsampling using graph convolutional networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11683–11692, 2021. 2
- [28] Yue Qian, Junhui Hou, Sam Kwong, and Ying He. Pugeonet: A geometry-centric network for 3d point cloud upsampling. In *European conference on computer vision*, pages 752–769. Springer, 2020. 2
- [29] Shi Qiu, Saeed Anwar, and Nick Barnes. Pu-transformer: Point cloud upsampling transformer. In *Proceedings of the Asian conference on computer vision*, pages 2475–2493, 2022. 2
- [30] Wentao Qu, Yuantian Shao, Lingwu Meng, Xiaoshui Huang, and Liang Xiao. A conditional denoising diffusion probabilistic model for point cloud upsampling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20786–20795, 2024. 2, 6, 7
- [31] Marie-Julie Rakotosaona, Vittorio La Barbera, Paul Guerrero, Niloy J Mitra, and Maks Ovsjanikov. Pointcleannet: Learning to denoise and remove outliers from dense point clouds. In *Computer graphics forum*, pages 185–203. Wiley Online Library, 2020. 2
- [32] Yi Rong, Haoran Zhou, Kang Xia, Cheng Mei, Jiahao Wang, and Tong Lu. Repkpu: Point cloud upsampling with kernel point representation and deformation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21050–21060, 2024. 2, 6, 7
- [33] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. The earth mover’s distance as a metric for image retrieval. *International journal of computer vision*, 40:99–121, 2000. 5
- [34] Ruwen Schnabel and Reinhard Klein. Octree-based point-cloud compression. *PBG@ SIGGRAPH*, 3, 2006. 1, 3
- [35] Sebastian Schwarz, Marius Preda, Vittorio Baroncini, Madhukar Budagavi, Pablo Cesar, Philip A Chou, Robert A Cohen, Maja Krivokuća, Sébastien Lasserre, Zhu Li, et al. Emerging mpeg standards for point cloud compression. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 9(1):133–148, 2018. 1
- [36] Xihua Sheng, Li Li, Dong Liu, and Zhiwei Xiong. Attribute artifacts removal for geometry-based point cloud compression. *IEEE Transactions on Image Processing*, 31:3399–3413, 2022. 2
- [37] Shaoshuai Shi, Xiaogang Wang, and Hongsheng Li. Pointcnn: 3d object proposal generation and detection from point cloud. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 770–779, 2019. 8
- [38] Lyne P. Tchapmi, Vineet Kosaraju, Hamid Rezatofighi, Ian D. Reid, and Silvio Savarese. TopNet: Structural point cloud decoder. *The IEEE Conference on Computer Vision and Pattern Recognition*, pages 383–392, 2019. 2, 6, 7
- [39] Mikaela Angelina Uy, Quang-Hieu Pham, Binh-Son Hua, Thanh Nguyen, and Sai-Kit Yeung. Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1588–1597, 2019. 5, 6, 7
- [40] Mathias Vogel, Keisuke Tateno, Marc Pollefeys, Federico Tombari, Marie-Julie Rakotosaona, and Francis Engelmann. P2p-bridge: Diffusion bridges for 3d point cloud denoising. In *European Conference on Computer Vision*, pages 184–201. Springer, 2024. 2
- [41] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, 38(5):1–12, 2019. 7
- [42] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1912–1920, 2015. 5, 6, 7, 8
- [43] Haozhe Xie, Hongxun Yao, Shangchen Zhou, Jiageng Mao, Shengping Zhang, and Wenxiu Sun. Grnet: Gridding residual network for dense point cloud completion. In *European Conference on Computer Vision*, pages 365–381. Springer, 2020. 2, 6, 7
- [44] Lequan Yu, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. Pu-net: Point cloud upsampling network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2790–2799, 2018. 2, 6, 7
- [45] Xumin Yu, Yongming Rao, Ziyi Wang, Jiwen Lu, and Jie Zhou. Adapointr: Diverse point cloud completion with adaptive geometry-aware transformers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 2, 6, 7
- [46] Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, and Martial Hebert. Pcn: Point completion network. In *2018 international conference on 3D vision (3DV)*, pages 728–737. IEEE, 2018. 2, 3, 6, 7
- [47] Junzhe Zhang, Tong Chen, Dandan Ding, and Zhan Ma. G-pcc++: Enhanced geometry-based point cloud compression. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 1352–1363, 2023. 2, 3, 6, 7, 8
- [48] Ruonan Zhang, Wei Gao, Ge Li, and Thomas H Li. Qinet: Decision surface learning and adversarial enhancement for quasi-immune completion of diverse corrupted point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–14, 2022. 2
- [49] Xiaoyu Zhang, Guibiao Liao, Wei Gao, and Ge Li. Tdrnet: Transformer-based dual-branch restoration network for geometry based point cloud compression artifacts. In *2022 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2022. 2
- [50] Haoran Zhou, Yun Cao, Wenqing Chu, Junwei Zhu, Tong Lu, Ying Tai, and Chengjie Wang. Seedformer: Patch seeds based point cloud completion with upsampling transformer. In *European conference on computer vision*, pages 416–432. Springer, 2022. 2, 6, 7