

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-13

论文引用格式: Jiang Zhaoyi, Shen Yuhui, Zhu Jiajun, Yang Bailin. Point cloud geometric compression integrating global frequency domain and local spatial domain features[J/OL]. Journal of Image and Graphics, XXXX: 1-13. DOI: 10.11834/jig.260535. (江照意, 沈宇辉, 朱佳骏, 杨柏林. 融合频域全局与空域局部特征的点云几何压缩[J/OL]. 中国图象图形学报, XXXX: 1-13. DOI: 10.11834/jig.260535.) [DOI: 10.11834/jig.260535]

融合频域全局与空域局部特征的点云几何压缩

江照意, 沈宇辉, 朱佳骏, 杨柏林*

浙江工商大学计算机科学与技术学院, 杭州 310018

摘要: 目的 针对现有点云几何压缩方法存在忽略频域信息、难以平衡感受野和计算复杂度的问题, 提出了一种融合频域全局与空域局部特征的点云几何压缩方法, 旨在提升压缩性能。方法 首先在点云压缩网络中引入了快速傅里叶变换, 利用点云的频域信息来构建点云的全局特征, 有效地扩大了感受野, 避免了注意力机制的高复杂度。同时, 设计了一种基于八叉树的卷积, 利用八叉树的层次化结构和紧凑表示, 高效提取点云的多尺度局部几何特征, 克服了基于张量的稀疏卷积计算与内存开销大的缺陷。此外, 还提出了一种通道位移卷积, 利用通道位移操作, 在不改变八叉树卷积核大小的条件下进一步扩大感受野, 有效缓解八叉树中感受野随层级增加而受限的问题。结果 在8iVFB(8i voxelized full bodies)、OwlII(OwlII dynamic human textured mesh sequence dataset)和MVUB(Microsoft voxelized upper bodies)等公开点云数据集上进行实验验证, 以每点比特数(bits per point, bpp)、编解码时间和内存占用作为评价指标。实验结果表明, 所提方法在跨数据集和同源数据集测试场景下均取得优于现有方法的压缩性能, 在跨数据集场景下平均bpp较基于几何的点云压缩方法G-PCC(geometry-based point cloud compression)和SparsePCGC分别降低43.00%和5.91%, 同时相比SparsePCGC显著降低编解码时间和内存占用。结论 所提出的方法通过融合频域全局特征建模与空域局部结构增强, 有效提升了点云几何结构的表示能力和占用概率预测精度, 在保证较低计算复杂度的同时实现了更加高效的点云几何压缩。

关键词: 点云几何压缩; 八叉树卷积; 频域感知; 上下文熵编码; 占用概率预测

Point cloud geometric compression integrating global frequency domain and local spatial domain features

Jiang Zhaoyi, Shen Yuhui, Zhu Jiajun, Yang Bailin*

School of Computer Science and Technology, Zhejiang Gongshang University, Hangzhou 310018, China

Abstract: Objective Existing point cloud geometry compression methods generally suffer from two major limitations. On the one hand, most of them rely heavily on spatial-domain feature extraction and pay insufficient attention to frequency-domain information, so the potential of frequency-domain cues in describing global geometric structures has not been fully exploited. On the other hand, they find it difficult to strike a good balance between receptive field size and computational complexity. Convolution-based methods are often constrained by limited receptive fields, especially in hierarchical octree structures, where the effective receptive field decreases significantly as the resolution increases. Transformer-based methods can capture long-range dependencies, but they introduce considerable computational and memory costs, making them

收稿日期: 2026-08-26; 修回日期: 2026-10-09

* 通信作者: 杨柏林 ybl@zjgsu.edu.cn

基金项目: 国家自然科学基金(62172366)

Supported by: National Natural Science Foundation of China(62172366)

© 中国图象图形学报版权所有

hard to apply to large-scale point cloud compression. To address these problems, this paper proposes a point cloud geometry compression method that integrates frequency-domain global features with spatial-domain local features, aiming to improve compression performance while keeping computational complexity low. **Method** First, a fast Fourier transform is introduced into the point cloud compression network. The input node features are transformed into the frequency domain based on the Morton order of octree nodes, and the resulting frequency components are analyzed and reconstructed. Low-frequency components provide global structural information of point cloud regions, while high-frequency components preserve local geometric details and boundary variations. In this way, the frequency-domain information of the point cloud is exploited to construct global features, which effectively enlarges the receptive field and avoids the high complexity of attention mechanisms. Second, an octree-based convolution is designed. By leveraging the hierarchical structure and compact representation of the octree, this convolution efficiently extracts multi-scale local geometric features of the point cloud, overcomes the high computational and memory overhead of tensor-based sparse convolution, and avoids redundant computation on empty regions. Third, a channel-shift convolution is proposed. Using channel shift operations along the octree node sequence, this design further enlarges the receptive field without changing the size of the octree convolution kernel or introducing significant additional computational overhead, thereby effectively alleviating the problem that the receptive field in an octree becomes increasingly limited as the hierarchy grows deeper. Finally, a multi-stage occupancy prediction strategy is adopted to progressively estimate child-node occupancy probabilities, and the predicted probabilities are combined with context-based entropy coding to generate compact bitstreams. Within each octree level, the prediction is refined stage by stage, so that the network can make full use of the available geometric context and produce more reliable occupancy estimates. Through the collaboration of frequency-domain global modeling and spatial-domain local enhancement, the proposed framework achieves more accurate occupancy probability estimation and more efficient geometric compression. **Result** Experiments were conducted on several public point cloud datasets, including 8iVFB (8i voxelized full bodies), OwlII (OwlII dynamic human textured mesh sequence dataset), and MVUB (Microsoft voxelized upper bodies). Bits per point (bpp), encoding and decoding time, and memory consumption were used as evaluation metrics. The proposed method was compared with the traditional geometry-based point cloud compression method G-PCC and the learning-based SparsePCGC method under both cross-dataset and within-dataset evaluation scenarios. Experimental results show that the proposed method achieves better compression performance than existing methods in both scenarios. In the cross-dataset scenario, the average bpp of the proposed method reaches 0.5625, which is reduced by 43.00% and 5.91% compared with G-PCC and SparsePCGC, respectively. Meanwhile, compared with SparsePCGC, the proposed method significantly reduces encoding and decoding time, achieving approximately 30% acceleration in compression efficiency, and reduces memory consumption by more than 70% on both the 8iVFB and OwlII datasets. The model size is also reduced, demonstrating the compactness and efficiency of the proposed architecture. Ablation studies further verify the effectiveness of each proposed component. The frequency-domain feature aggregation module improves global structural modeling, while the spatial feature aggregation module enhances local geometric representation. Combining these two modules achieves the best compression performance, confirming their complementary advantages. These results indicate that frequency-domain and spatial-domain features play complementary roles. **Conclusion** By integrating frequency-domain global feature modeling with spatial-domain local structure enhancement, the proposed method effectively improves the representation capability of point cloud geometric structures and the accuracy of occupancy probability prediction. The designed octree convolution and channel-shift convolution expand the feature modeling capability while maintaining low computational complexity. Extensive experiments demonstrate that the proposed method achieves superior compression performance, lower computational cost, and reduced memory consumption compared with existing approaches. The proposed framework provides an effective solution for efficient point cloud geometry compression and has potential application value in large-scale three-dimensional data storage and transmission scenarios.

Key words: point cloud geometry compression; octree convolution; frequency perception; context-based entropy coding; occupancy probability prediction

0 引言

点云作为一种直接、灵活且高保真的三维数据表示形式(Schwarz等, 2018),已成为当前三维视觉领域最重要的数据载体之一(Nguyen等, 2021c),广泛应用于诸多重要领域,如自动驾驶(Martins等, 2024),增强现实(Hu等, 2025)和三维地图(Lyu等, 2024)。随着点云数据的爆发式增长,每帧点云可包含高达数百万个点,给数据存储和网络传输带来了巨大挑战。因此,高效的点云压缩技术已成为支撑点云应用落地的关键基础技术(元辉等, 2026)。

点云是一种无序、稀疏且非结构化的三维数据形式,其数学建模和信号处理理论与传统的图像有着本质区别,这也给点云压缩带来了巨大的挑战。为此,国际运动图像专家组(Moving Picture Experts Group, MPEG)于2017年启动了点云压缩标准化工作,并于2020年正式发布了两种点云压缩国际标准:基于视频的点云压缩方法(video-based point cloud compression, V-PCC)(Graziosi等, 2020)和基于几何的点云压缩方法(geometry-based point cloud compression, G-PCC)(Li等, 2024)。G-PCC在稀疏点云压缩方面表现优异,已成为当前点云压缩的基准参考。然而,G-PCC采用手工设计的上下文模型和熵编码策略,难以充分建模点云数据的复杂统计特性,压缩性能仍有较大提升空间。

随着深度学习技术的快速发展,基于深度学习的点云压缩方法在压缩性能上已经超越传统的手工设计方法(Nguyen等, 2023)。现有的点云几何压缩技术主要分为基于卷积神经网络(convolutional neural networks, CNN)和基于Transformer(Han等, 2021)两大类。稀疏卷积作为CNN的典型代表,是一种专为高维稀疏数据设计的卷积操作,能够高效地处理点云的稀疏性特征(Choy等, 2019),避免对空体素的无效计算,具有计算效率高、内存占用低的显著优势(Yu等, 2022)。CNN擅长提取局部特征,捕获复杂的数据模式,但受其感受野的限制,无法捕捉全局依赖。在层级八叉树结构中,相同尺寸的卷积核在不同的层上拥有不同的感受野,虽然在粗糙层卷积可以具有较大的感受野,但是在精细层卷积的感受野呈指数级减小。直接增大卷积核会导致模型复杂度大大增加,与压缩率的增益不成正比,这揭

示了层级八叉树结构中感受野与模型复杂度之间的固有矛盾。Transformer通过自注意力机制直接关注输入序列中任意位置的信息,能够有效捕捉全局依赖关系,可以有效解决CNN无法建模长距离依赖的局限性。但同样Transformer对局部信息的敏感度较差,并且计算和内存开销非常大,时间复杂度是 $O(n^2)$,尤其在处理长序列时,需要大量的内存和计算资源。此外,现有方法仅从空域分析点云的特征,忽略了频域信息对点云压缩的影响。从信号处理的角度来看,点云作为一种三维离散信号,在不同尺度下呈现出差异化的频率特性。低频分量通常刻画整体结构与平滑区域,而高频分量通常对应局部细节与突变信息。若缺乏对频率分量的调控,可能会限制模型对结构信息的充分利用。

为此,本文提出了一种融合频域全局与空域局部特征的点云几何压缩方法,旨在提高点云几何压缩性能的同时保持较高的压缩效率。该方法在八叉树卷积框架下设计了两个核心模块:频域特征聚合模块与空间特征聚合模块,分别用于建模全局结构与局部几何细节。频域特征聚合模块通过对节点特征进行频率域变换与特征重构,实现对不同频率分量的分离与建模,从而刻画点云整体几何结构及潜在的长程依赖关系,在避免高复杂度注意力机制的同时有效扩大感受野。空间特征聚合模块在八叉树邻域建模基础上,引入通道位移卷积以增强特征交互能力。该卷积通过通道维度的特征位移实现跨邻域信息传递,在不增加卷积核大小的前提下扩展感受野,缓解感受野受限问题,从而更有效地刻画局部几何结构。两个模块在统一框架下进行融合,共同服务于子节点占用概率预测,使模型能够同时兼顾全局结构表达与局部细节建模,从而提升几何压缩表示的紧致性与编码效率。

1 相关工作

1.1 基于CNN的点云几何压缩

基于CNN的点云几何压缩方法已取得显著进展,尤其在压缩效率和性能方面展现了强大优势。

学习型点云几何压缩最早受到图像压缩中自回归模型的启发。VoxelDNN是较早提出的基于深度神经网络的点云几何压缩方法之一(Nguyen等, 2021a)。该方法将点云体素化,在八叉树结构下逐

层编码体素占用信息,通过3D卷积神经网络对体素占用的条件概率进行建模,再结合算术编码实现无损压缩,取得了明显的性能提升,验证了深度概率模型在点云几何压缩中的有效性。在此基础上,MSVoxelDNN引入多尺度体素上下文建模策略(Nguyen等,2021b),在不同分辨率层级上联合预测体素占用概率,增强跨尺度空间相关性建模,提升了压缩性能。但这类体素表示的方法在高分辨率点云上计算和存储开销较大。SparseVoxelDNN通过引入稀疏卷积机制(Nguyen等,2022),仅对非空体素区域进行建模与计算,在保证压缩性能的同时显著降低了模型复杂度,提高了方法在高分辨率点云场景下的可扩展性。这一思路也逐渐成为后续学习型点云压缩方法的重要设计原则。

VoxelContext-Net通过显式构建节点上下文关系(Que等,2021),对当前节点的占用情况进行条件概率预测,相较于固定卷积邻域的方法具有更强的上下文表达能力。PCGCv2采用分层式稀疏卷积网络替代传统稠密卷积(Wang等,2021),通过几何细化策略与自适应熵编码机制,在保持较低计算复杂度的同时显著提升了率失真性能。SparsePCGC将稀疏卷积引入八叉树概率模型(Wang等,2022),仅对非空节点进行建模,使方法在保证压缩性能的同时具备更好的计算效率。

针对现有点云压缩方法推理时间太慢而无法满足激光雷达实时采集需求的问题,RENO提出了一种基于稀疏卷积的实时压缩框架(You等,2025),通过稀疏占用码直接推导占用信息,并设计了一种新的单阶段占用预测,从而达到实时的编解码速度且模型非常小。

总体而言,基于CNN的点云压缩方法已形成较为成熟的技术路线,在压缩效率、计算复杂度及内存占用等维度持续取得突破性进展。然而,这类方法主要依赖空间域的局部上下文信息,对点云结构中的全局相关性和频域特性利用仍较为有限。本文利用快速傅里叶变换提取点云的频域信息来构建全局特征,有效地获取了全局感受野。提出了一种通道位移卷积,在不引入额外计算开销的前提下扩大了局部的感受野,增强对点云结构特性的表达能力。

1.2 基于Transformer的点云几何压缩

为弥补卷积神经网络在建模长程依赖方面的不足,近年来研究者开始将注意力机制和Transformer

结构引入点云压缩领域。

OctAttention在八叉树节点建模过程中引入自注意力机制(Fu等,2022),捕捉不同空间区域之间的非局部依赖,辅助节点占用概率预测,在复杂几何结构区域具有一定优势。进一步地,OctFormer将八叉树节点序列化为token(Cui等,2023),并利用多层Transformer对节点特征进行全局建模,在不同空间位置 and 不同层级之间建立长程依赖关系,在高复杂度点云场景下展现出较强的上下文建模能力,但其计算复杂度也相对较高。

在熵建模方面,EHEM提出了一种熵感知的层级概率建模方法(Song等,2023),通过结合注意力机制对点云的概率分布进行建模,使编码过程能够更加自适应地分配比特资源。ECM-OPCC则进一步提出端到端的八叉树点云压缩框架(Jin等,2024),将上下文建模与熵编码过程进行统一优化,在复杂点云场景中表现出较好的泛化性能。此外,TopNet提出的自顶向下树结构建模思想(Wang等,2025),虽然最初用于点云生成任务,但其层级化建模方式为后续点云压缩方法提供了有益的参考。

总体来看,基于Transformer的点云压缩方法在捕获全局依赖关系方面具有明显优势,但其较高的计算和存储开销在一定程度上限制了其在高分辨率点云压缩场景下的应用。为实现更加高效的点云压缩,有必要探索计算复杂度更低、同时具备有效上下文建模能力的替代方案。

1.3 基于频域的点云特征学习方法

随着深度学习在点云任务中的广泛应用,研究者逐渐意识到仅依赖空间域特征难以充分刻画复杂几何结构,因而开始探索将频域或谱域分析引入点云表示框架中,以增强几何结构表达。

Wen等人(2024)提出在点云图结构上构建可学习的谱域小波变换PointWavelet,通过在频域设计卷积算子捕捉多尺度几何特征,从而在分类和分割任务中取得了优于传统空间域方法的性能。

Kim等人(2025)提出基于三维傅里叶分解的点云频域建模方法,在连续三维空间中引入可学习的傅里叶基函数,将点云属性显式分解为振幅与相位分量,分别表征外观属性与几何结构信息,实现了几何与属性特征的有效解耦,并在分类、分割及风格迁移等任务中验证了其表达能力。但该方法依赖点级的连续傅里叶映射与三角函数计算,计算复杂度和

显存开销随点数与频率维度线性增长,难以高效扩展至大规模点云或层级化几何表示场景。

Jin 等人(2026)则将频域感知机制直接融入点云网络结构设计中,在特征空间引入频率分解与多尺度融合策略,在多尺度特征提取过程中显式区分不同频率成分,并通过频域信息的跨尺度融合,有效提升了模型对长程依赖和全局上下文的感知能力。但其频域操作仍主要作用于中间特征空间,频率分量与真实空间尺度之间的对应关系相对间接。

基于学习的频域点云方法将频域信息作为可学习组件嵌入端到端网络,使模型能够根据数据分布自适应地利用不同频率成分,为点云特征建模提供了新的视角。相比 Transformer,一维快速傅里叶变换具有更优的计算复杂度,为高效点云压缩提供了可行的路径。

2 方 法

提出了一种面向点云几何压缩的层级化特征建模与预测框架,如图 1 所示。该方法以八叉树结构为基础,对输入点云进行逐层空间划分,并在树形层级中对节点占用状态进行自顶向下的预测与编码。为增强特征表达能力,该框架在节点特征建模阶段引入频域与空域相结合的特征聚合策略。输入特征经嵌入映射后被划分为频域分支与空域分支,其中频域分支通过快速傅里叶变换(fast Fourier transform, FFT)对特征进行频率分解,以捕获跨节点的全局结构信息;空域分支通过轻量级卷积与通道位移操作对局部几何模式进行建模。两类特征在融合后再通过层级上采样映射到子节点空间,用于预测下一层节点的占用概率。

2.1 八叉树结构

现有基于 CNN 的点云几何压缩方法大多采用体素化表示的稀疏卷积。体素化将空间划分为规则的网格,便于快速访问每个体素的特征。但体素化会在大部分空白区域生成无效体素,浪费大量的内存和计算资源。也有基于八叉树表示的稀疏卷积方法,如 OCNN(Wang 等, 2017)。该方法在低分辨率点云上具有较高的计算效率,但由于其结构设计的限制,当应用于高分辨率点云时,时间复杂度显著增加。

因此,本文采用了英伟达提出的结构化点云

(structured point cloud, SPC)(Tsang, 2024)作为八叉树结构。如图 1 所示,根节点位于第 0 层,绿色的节点表示占用节点,保存了占用值和坐标,占用值表示其孩子节点的占用情况,坐标表示其在点云中的具体位置。叶子节点没有孩子节点,故不再保存占用值。虚线框表示包含整层上下文信息的输入特征。该特征经过特征建模与上采样后生成对应子节点的特征表示,并基于此进一步预测占用概率。

该八叉树结构使点云数据的存储与处理更加高效。节点以紧凑的线性数组形式存储,并通过金字塔索引与点层级实现不同层级节点之间的快速访问,避免了传统指针式树结构的额外内存开销,减少了不必要的内存访存,从理论上保证了较低的内存占用和较高的计算效率。在计算效率方面,该结构将空间邻接关系隐式编码于索引映射中,使八叉树卷积能够在连续内存上以批量方式执行,显著降低了计算复杂度并提高了 GPU(graphics processing unit)并行度,使其在大规模点云场景下具备更高的卷积执行效率。

2.2 频域特征聚合

点云数据通常在空域进行处理,但其高度稀疏性和复杂结构使得预测变得困难。尤其在大型场景或复杂环境中,空域信息可能包含冗余或噪声,影响占用概率的准确预测。

尽管快速傅里叶变换已被广泛应用于规则栅格图像的频率特征分析,但 3D 点云的无序性和非规则度量空间特性限制了标准频域工具的直接应用,导致其在点云频域特征提取中的潜力尚未得到充分挖掘。频域不仅可以在低计算复杂度下高效捕捉全局长程依赖,突破自注意力机制的计算瓶颈,还可以通过紧凑的频域表示学习,建立通用的数据先验信息(Jiang 等, 2024)。从频率视角审视点云,频域能有效地捕捉全局特征,并且频域特征本质上包含点云结构的一些先验信息,这些先验可能在网络设计中被有效利用。

在点云处理中,直接对三维点云数据做傅里叶变换具有一定的挑战性,为了更高效地处理点云数据,可将点云数据转换成一维的形式进行傅里叶变换。这样做虽然失去了完整的结构信息,但处理更加便捷,能够简化计算、提高效率,以平衡计算效率与数据表示精度。基于此提出了频域特征聚合模块。

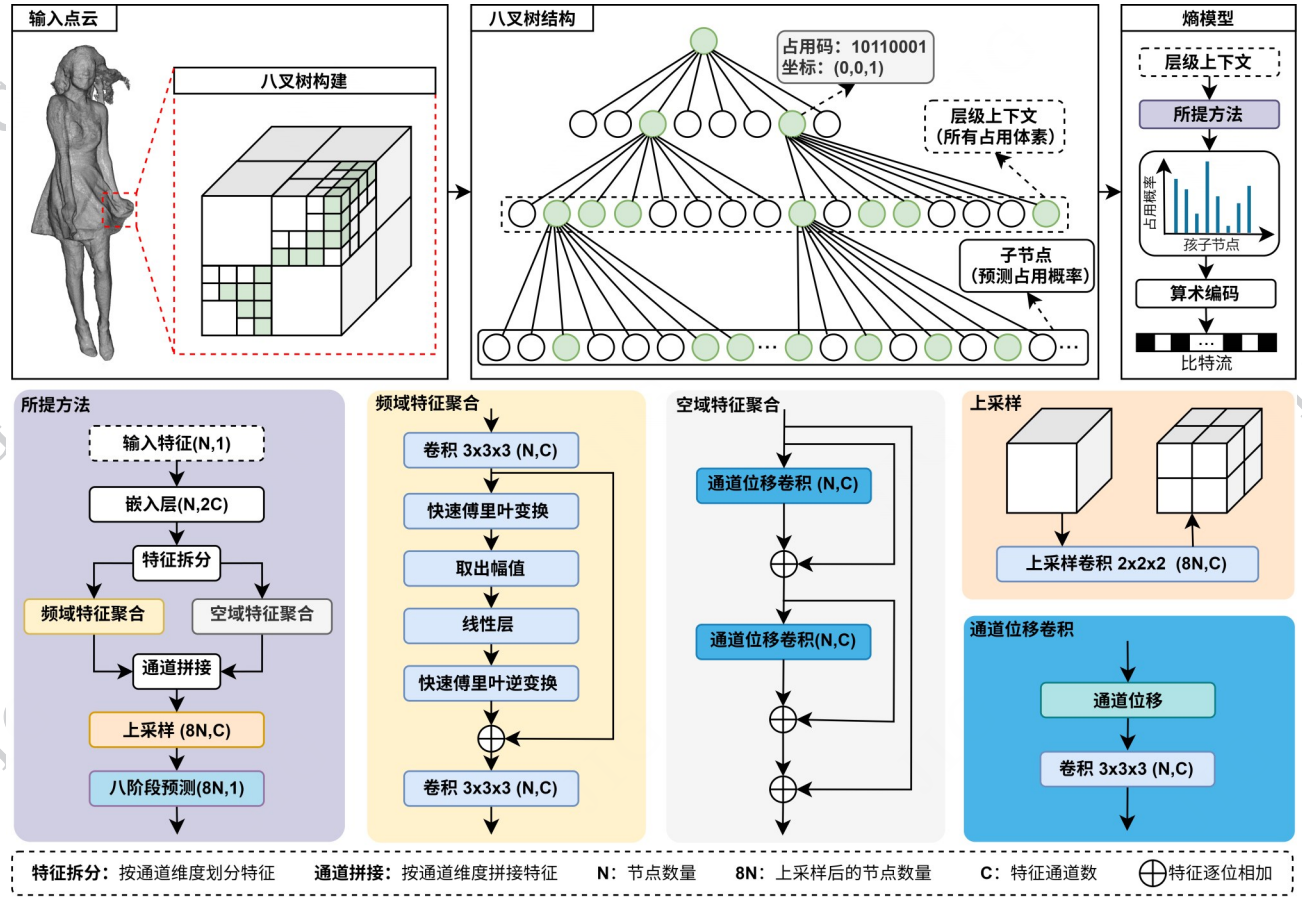


图1 整体框架

Fig. 1 Overall framework

输入特征 $F_f \in \mathbb{R}^{N \times C}$, 首先通过一个卷积提取点云的空间特征 $F_0 \in \mathbb{R}^{N \times C}$, 捕捉点云的局部特征, 为快速傅里叶变换做好准备。接下来, 按八叉树节点组织的 Morton 顺序进行一维快速傅里叶变换将空域的特征转换到频域。在频域中, 低频部分代表了点云的整体结构, 而高频部分通常对应于点云的细节部分。在占用概率预测中, 低频信息有助于判断哪些区域整体被占用, 高频信息有助于处理占用区域的边界细节。通过适当的滤波和处理, 频域特征聚合模块能够有效提升模型在这些方面的表现。取出幅值 $F_i \in \mathbb{R}^{N \times C}$ 经过一个线性变换, 然后再通过快速傅里叶逆变换将其转换回空域。最后通过残差连接和卷积进行信息整合得到输出 $F_a \in \mathbb{R}^{N \times C}$ 。完整的过程可以形式化为:

$$F_0 = \text{Conv}_3(F_f) \quad (1)$$

$$F_i = \text{FFT}(F_0) \quad (2)$$

$$F_a = \text{Conv}_3(\text{IFFT}(\text{Linear}(F_i)) + F_0) \quad (3)$$

式中, $\text{Linear}(\cdot)$ 表示一维线性变换, $\text{Conv}_k(\cdot)$ 表示

卷积核大小为 k 的稀疏卷积, $\text{FFT}(\cdot)$ 和 $\text{IFFT}(\cdot)$ 分别表示一维傅里叶变换和一维傅里叶逆变换。

2.3 空间特征聚合

在使用普通的点云稀疏卷积时, 网络的感受野不足, 模型只能理解非常局部区域的信息, 无法捕获较大范围的结构特征。如果网络层数太少或卷积核过小, 信息传递就会受到限制, 增大感受野通常需要更多的卷积层或更大的卷积核, 这会直接增加计算复杂度(Li等, 2025)。

为了提高特征处理的效率、改善压缩性能, 并使模型能够更好地适应不同类型的特征图, 提出了通道位移卷积, 通过调整通道方式, 可以在不同的通道范围内增加信息的传播路径, 从而实现感受野的扩展。通道位移的方式如图2所示。因为点云是稀疏的, 标准化通道位移操作的实现是困难的, 所以使用结构化点云的 Morton 顺序进行通道位移, 这种方法可以在较低的计算成本下达到较大的感受野。通过设计这个模块, 可以有效地提高网络对大范围上下

文信息的捕捉能力。

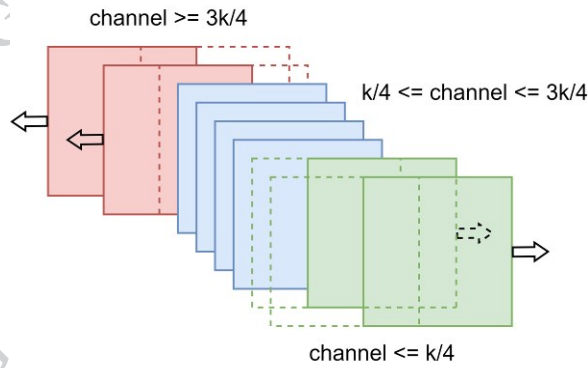


图2 通道位移

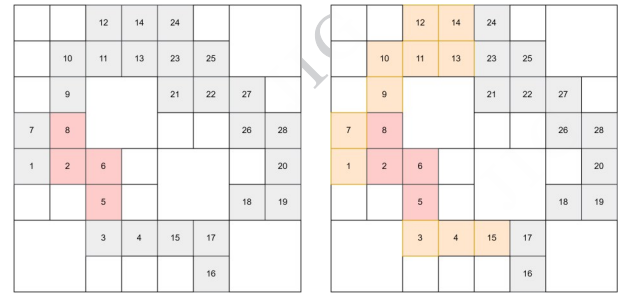
Fig. 2 Channel shift

不同颜色的框表示不同的通道数范围,绿色框是前 1/4 的通道,蓝色框是中间 1/2 的通道,红色框是后 1/4 的通道。虚线框表示通道位移之前的位置,实线框表示通道经过位移之后所在的位置。箭头方向则表明了通道位移的方向。每一个通道位移的距离不同,以获取不同距离的特征,以绿色框的通道为例,假设总通道数 k 为 32。那么 $k/4$ 就是 8,这 8 个通道分别向右移动 1-8 个距离。

图 3 和图 4 展示了感受野的变化。图中颜色标记的节点是当前要进行卷积的节点,红色标记的是 kernel 的感受野范围,橙色标记的是通道位移卷积扩展的感受野范围。以中心节点 6 为例,图 3-a 表示标准 $3 \times 3 \times 3$ 卷积的感受野范围。图 3-b 表示改进后的通道位移卷积的感受野范围,相比标准 $3 \times 3 \times 3$ 卷积拥有更大的感受野。

在预测过程中可能出现节点邻域为空的情况,此时标准 $3 \times 3 \times 3$ 卷积无法获取局部信息。以中心节点 6 为例,如图 4-a 所示,其邻域没有节点,只能利用自身信息进行占用概率的预测。使用通道位移卷积后感受野如图 4-b 所示,得到了有效扩大。

基于上述通道位移卷积,本文设计了空间特征聚合模块。具体而言,给定输入空间特征 $F_s \in \mathbb{R}^{N \times c}$,首先通过一个通道位移卷积模块对局部特征进行建模,并通过残差连接得到中间结果 $F_0 \in \mathbb{R}^{N \times c}$ 。然后将 F_0 再次输入通道位移卷积模块以进一步增强局部特征表达能力,并与原始输入 F_s 和中间特征 F_0 残差相加,从而得到最终的空间聚合特征 F_c 。上述描述可表达为:



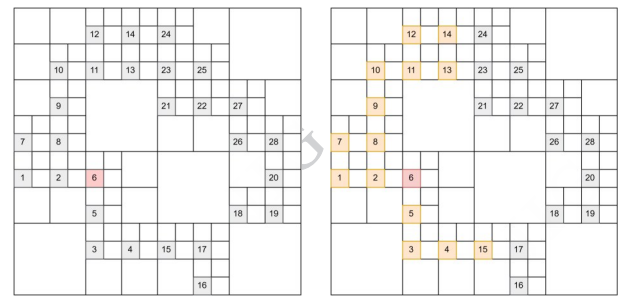
(a) 卷积感受野

(b) 通道位移卷积感受野

((a) convolutional receptive field; (b) channel shift convolutional receptive field)

图3 d层的感受野

Fig. 3 The receptive field of layer d



(a) 卷积感受野

(b) 通道位移卷积感受野

((a) convolutional receptive field; (b) channel shift convolutional receptive field)

图4 d+1层阶段1的感受野

Fig. 4 The receptive field of layer d+1

$$F_0 = \text{CSCNN}(F_s) + F_s \quad (4)$$

$$F_c = \text{CSCNN}(F_0) + F_s + F_0 \quad (5)$$

式中, $\text{CSCNN}(\cdot)$ 表示通道位移卷积。

2.4 多阶段占用预测与八叉树更新策略

为充分利用空间上下文信息并提升占用预测的建模能力,本文采用多阶段占用预测与八叉树更新策略。首先对每个待分裂节点上采样,将其映射为 8 个具有明确空间位置的子节点。随后将其按照在三维空间中的相对位置划分为 8 个空间分组。在每个预测阶段中,仅对当前分组内的子节点进行占用概率预测,而其余子节点保持未预测状态。随着预测阶段的推进,已预测子节点的结构信息被逐步引入,为后续分组提供更充分的空间上下文,从而降低占用预测的不确定性。

图 5 给出了一个二维四叉树的示例。在每个阶段,占用预测模块以当前阶段要预测节点和已预测

占用节点的特征为输入,输出当前阶段预测节点的占用概率。为了保持整体编码效率,占用预测模块仅通过一个通道位移卷积模块与一个标准的八叉树卷积实现不同空间位置之间的隐式信息交互;随后,采用三个连续的八叉树点卷积对展开后的子节点进行占用概率预测。在每个阶段中,模型输出的占用概率将与对应的真实占用标签共同用于算术编码,

以生成码流;在解码端,根据预测概率与码流重建子节点的占用状态,并据此更新八叉树结构。重建为占用的子节点将参与后续阶段的上下文建模,而重建为空的子节点则被剪枝,不再参与后续计算。通过子节点展开、分组预测、结构更新的多阶段迭代过程,八叉树结构得以逐步细化,并始终与当前阶段的预测结果保持一致。

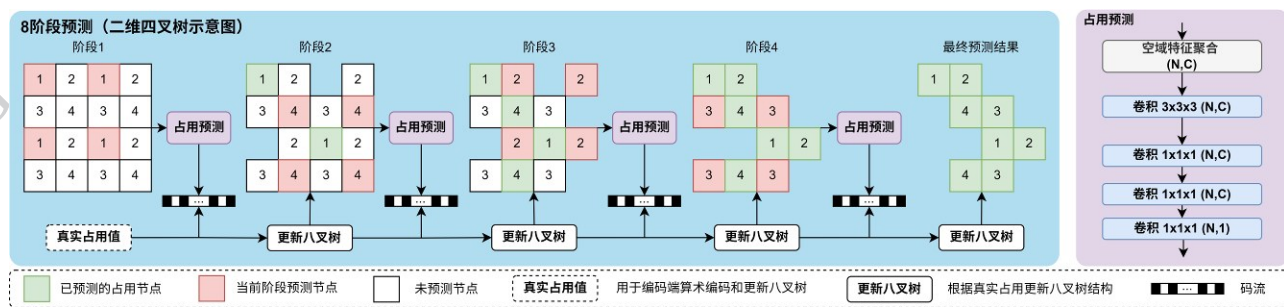


图5 基于空间分组的多阶段占用预测

Fig. 5 Multi-Stage Occupancy Prediction Based on Spatial Grouping

3 实验结果与分析

在本节中,首先介绍实验设置,包括数据集、实现细节和评估指标。随后展示了所提方法的点云压缩性能。此外,还对所提方法设计因素进行了消融实验,以分析它们对压缩性能的影响。

3.1 数据集

为全面评估所提出方法的几何无损压缩性能及其泛化特性,本文采用了两种训练与测试场景。第一种为跨数据集场景,通过在不同数据分布之间进行训练与测试,重点考查模型在点云密度和几何结构发生变化时的泛化能力;第二种为同源数据集场景,在保证训练集与测试集相互独立的前提下,评估模型在一致数据分布条件下的压缩性能。

跨数据集场景里使用 SparsePCGC 提供的经过预处理和增强的 ShapeNet (an information-rich 3D model repository) 数据集 (Chang 等, 2015) 作为训练集。测试时使用 8iVFB (8i voxelized full bodies) 数据集 (d'Eon 等, 2017) 与 OwlII (OwlII dynamic human textured mesh sequence dataset) 数据集 (Xu 等, 2017)。为了方便与 SparsePCGC 进行比较,选取 8iVFB 的 longdress_vox10_1300、redandblack_vox10_1550、soldier_vox10_0690、loot_vox10_1200 和 OwlII

的 basketball_player_vox11_0200、dancer_vox11_0001 进行测试。它们具有不同的几何精度和结构特性,能够代表多样化的物体,用于评估模型的泛化性能。

同源数据集场景分别使用 8iVFB 和 MVUB (Microsoft voxelized upper bodies) 数据集 (Loop 等, 2016) 进行训练和测试。在 8iVFB 数据集中,选取 longdress_vox10 和 soldier_vox10 序列作为训练数据, loot_vox10 和 redandblack_vox10 序列作为测试数据。在 MVUB 数据集中,选取 Andrew_vox10、David_vox10、Sarah_vox10 序列作为训练数据, Phil_vox9/10 和 Ricardo_vox9/10 序列作为测试数据。所有序列均来源于同一数据集,但在动作姿态及局部细节方面存在差异。该设置主要用于评估模型在相同数据分布条件下的压缩性能与稳定性。

3.2 实现细节

为了公平对比,按照上述场景设置分别训练了跨数据集场景和同源数据集场景的模型。在跨数据集实验中,ShapeNet 数据集的模型训练覆盖八叉树第 4 至第 8 层,而 8iVFB 与 MVUB 数据集则训练第 6 至第 10 层,各数据集的最大层数由其点云精度决定。测试阶段,模型对第 4 层至最深层 (即该数据集所能达到的最大八叉树层级) 范围内的所有层级进行压缩。对于最低的 4 层 (即第 1 至第 4 层),其坐标信息单独采用 G-PCC 进行编码。

在本实验中,所提出的几何无损压缩方法基于深度学习框架实现。为了保证实验的可重复性和高效性,所提模型在 Pytorch 中实现,所有模型均在 NVIDIA RTX 6000 GPU(48GB 内存)上训练和测试。模型的训练过程采用了标准的 Adamw 优化器,权重衰减设置为 0.01,初始学习率设置为 0.001,批量大小设置为 1。训练的最大迭代次数为 30 次。损失函数是每个八叉树节点估计占用概率与真实占用情况之间的二元交叉熵损失。

3.3 评估指标和比较方法

对于点云几何无损压缩,度量每点的平均比特流大小,表示为每点比特(bits per point, bpp)。此外,评估编码和解码时间以评估压缩过程的计算效率。为了全面评估所提方法的压缩性能,将其与已有点云几何压缩方法进行了对比。对于跨数据集场景,与基于学习的方法 SparsePCGC 和传统方法 G-PCC 进行了对比。对于同源数据集场景,与基于学习的方法 OctAttention、VoxelDNN、MSVoxelDNN、ECM-OPCC、TopNet 进行了对比,同时也与传统方法 G-PCC 进行了比较。

3.4 实验结果

表 1 给出了所提方法与 G-PCC 和 SparsePCGC 在多个动态点云数据集上的无损压缩对比结果。所提方法在全部序列上均取得最低 bpp,平均码率为 0.5625,较 G-PCC 和 SparsePCGC 分别降低 43.00% 和 5.91%。在 OwlII 的 player 与 dancer 序列上进一步取得超过 11% 的码率下降,表明模型在高分辨率复杂场景下具备更强的空间建模能力;在 8iVFB 的 vox10 序列上保持约 3% 的稳定增益,体现出良好的泛化性。结果验证了所提结构在占用概率建模与熵

编码效率方面的有效性。

表 2 给出了在同源数据集场景下不同无损点云几何压缩方法的性能对比。所提方法在所有测试序列上均取得了最低的比特率,表现出最优的压缩性能。具体而言,所提方法相比 G-PCC 在平均 bpp 上实现了 42.94% 的比特率节省。该结果表明,在同源训练与测试设置下,所提方法能够更有效地建模点云几何结构,从而显著提升无损几何压缩效率。

表 3 展示了编解码时间的对比结果。所提方法在编码与解码阶段均显著降低了运行时间,相比 SparsePCGC 实现了约 30% 的时间节省。该效率提升

主要得益于模型结构设计的优化。一方面,通过减少常规卷积层的数量,有效降低了整体计算开销;另一方面,采用基于傅里叶变换的全局频域特征提取模块,在较低计算成本下建模长程依赖关系。此外,引入不增加额外计算复杂度的通道位移卷积,在保持高效性的同时进一步增强了特征表达能力。

表 4 对比了所提方法与 SparsePCGC 在不同数据集上的内存占用情况。在 8iVFB 数据集上,所提方法的内存占用比 SparsePCGC 减少了约 72.81%;在 OwlII 数据集上,减少了约 72.74%。该结果表明,所提出的方法在不同规模和复杂度的点云数据集上均具有良好的内存效率,能够在保持压缩性能的同时显著降低内存需求,有助于提升方法在实际工程应用和资源受限环境中的可行性。

表 5 对比了所提方法与 SparsePCGC 在模型参数规模方面的差异。结果表明,所提方法的模型参数量更低,相较于 SparsePCGC 减少约 17.19%。这说明所提方法在网络结构设计上更加紧凑,在不依赖

表 1 跨数据集场景模型实现无损压缩所需 bpp 对比

Table 1 Comparison of bpp required for lossless compression by cross-dataset scenario models

点云	G-PCC	SparsePCGC	所提方法	相比 G-PCC	相比 SparsePCGC
longdress_vox10	1.015	0.625	0.6040	-40.49%	-3.36%
loot_vox10	0.97	0.596	0.5754	-40.68%	-3.46%
redandblack_vox10	1.1	0.69	0.6717	-38.94%	-2.65%
soldier_vox10	1.03	0.628	0.6060	-41.17%	-3.50%
player_vox11	0.898	0.52	0.4609	-48.67%	-11.37%
dancer_vox11	0.88	0.514	0.4569	-48.08%	-11.11%
Average	0.9822	0.5955	0.5625	-43.00%	-5.91%

注:粗体表示最优结果。

表 2 同源数据集场景模型实现无损压缩所需 bpp 对比

Table 2 Comparison of bpp required for lossless compression of same-source dataset scene models								
点云	G-PCC	OctAttention	VoxelDNN	MSVoxelDNN	ECM-OPCC	TopNet	所提方法	相比 G-PCC
loot_vox10	0.95	0.62	0.64	0.73	0.55	0.53	0.47	-50.03%
redandblack_vox10	1.09	0.73	0.73	0.87	0.66	0.66	0.57	-47.45%
phil_vox9	1.23	0.83	0.92	-	0.79	0.79	0.72	-41.46%
phil_vox10	1.07	0.79	0.83	1.02	0.76	0.76	0.69	-35.51%
ricardo_vox9	1.04	0.72	0.72	-	0.76	0.76	0.62	-40.38%
ricardo_vox10	1.07	0.72	0.75	0.95	0.69	0.69	0.62	-42.06%
Average	1.08	0.74	0.77	0.89	0.70	0.70	0.62	-42.94%

注:粗体表示最优结果。

表 3 时间性能对比

Table 3 Time Performance Comparison				
方法	8iVFB(s)		Owlii(s)	
	编码	解码	编码	解码
G-PCC	5.788	4.76	20.365	14.79
SparsePCGC	0.965	0.85	2.235	2.05
所提方法	0.633	0.55	1.57	1.4
相比 G-PCC	-89.07%	-88.35%	-92.29%	-90.52%
相比 SparsePCGC	-34.46%	-35.04%	-29.75%	-31.71%

注:粗体表示最优结果。

表 4 内存占用对比

Table 4 Memory Usage Comparison		
方法	8iVFB(GB)	Owlii(GB)
SparsePCGC	5.03	15.52
所提方法	1.37	4.23
相比 SparsePCGC	-72.81%	-72.74%

注:粗体表示最优结果。

大规模参数的情况下即可实现有效的上下文建模。这一特性有助于降低模型的存储开销,为模型的部署提供了有利条件。

表 5 模型大小对比

Table 5 Model Size Comparison			
方法	SparsePCGC	所提方法	相比 SparsePCGC
模型大小(MB)	1.251	1.036	-17.19%

注:粗体表示最优结果。

3.5 可视化分析

为了直观验证本文方法的重建性能,图 6 给出了在 redandblack_vox10_1550 上的可视化对比结果,从左到右依次为原始点云、本文方法重建点云以及逐点重建误差热力图。可以看出,重建点云的全部点均呈现深蓝色,且经统计所有点的逐点几何误差均为 0mm。这表明本文方法在该序列上实现了无损重建,验证了其在点云压缩任务中的有效性。

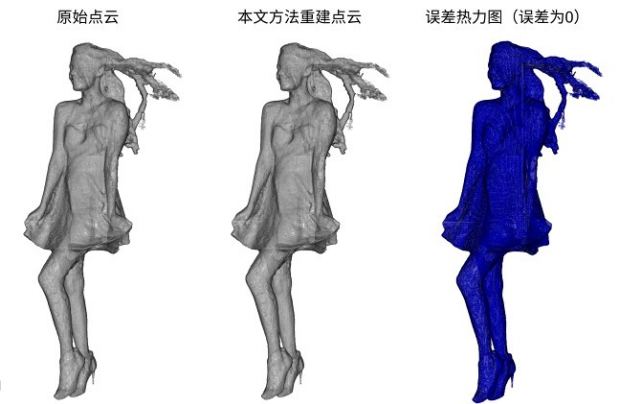


图 6 重建结果可视化对比
Fig. 6 Reconstruction results visualization comparison

3.6 消融实验

为了进一步验证所提方法的有效性,还需评估表 6 中各个模块的贡献。该实验在跨数据集场景下进行。模型 A 作为基线模型,仅使用普通的八叉树卷积进行特征聚合,其 BPP 为 0.6512。在此基础上分别引入不同模块进行对比分析。

仅引入空域特征聚合模块(模型 B,以该模块替代原八叉树卷积)时,BPP 降至 0.6328,较基线下降

表6 不同模块的消融实验

Table 6 Ablation experiment for different modules

模型	频域特征 聚合模块	空域特征 聚合模块	BPP(↓)
A			0.6512
B		√	0.6328(-2.83%)
C	√		0.6292(-3.39%)
D	√	√	0.6143(-5.67%)

注:粗体表示最优值,√表示采用对应模块。

2.83%,说明引入通道位移模块的空域上下文建模能够有效提升点云几何结构的表达能力,从而改善压缩效率。当仅引入频域特征聚合模块(模型C,以该模块替代八叉树卷积)时,BPP降至0.6292,较基线下降3.39%。表明频域特征能够捕获点云中更具全局性的结构信息,在几何冗余建模方面具有更强的表征能力。当同时引入频域特征聚合模块和空域特征聚合模块(模型D,双分支融合)时性能最优,BPP进一步降至0.6143,较基线下降5.67%。表明两种特征聚合方式具有良好互补性,联合使用能更充分挖掘点云在空间域与频率域中的冗余信息,显著提升压缩性能。综上所述,频域与空域特征聚合模块均对点云几何无损压缩有积极作用,且二者联合使用增益最显著。

4 结 论

针对点云几何无损压缩中八叉树逐层编码阶段上下文信息利用不充分、局部特征感受野受限以及全局结构信息建模不足等问题,提出了一种融合频域全局与空域局部特征的点云几何压缩方法。通过传统八叉树编码框架内引入频域特征聚合模块和空域特征聚合模块的协同建模策略,在不显著增加计算与内存开销的前提下,有效提升了节点占用概率预测的准确性,从而实现了更优的几何压缩性能与更高的编解码效率。

在多个公开点云数据集上的实验表明,所提方法在跨数据集和同源数据集测试场景下均能够取得较好的压缩性能。与G-PCC、SparsePCGC以及其他基于学习的点云几何压缩方法相比,所提方法能够更充分地利用点云的空间结构信息,在降低压缩率的同时保持较好的编解码效率。消融实验进一步

表明,频域特征聚合模块和空域特征聚合模块均能够带来独立的性能增益,二者联合使用能够进一步提升压缩性能,验证了频域与空域特征协同建模的有效性。

尽管本文方法在压缩性能与效率方面取得了显著提升,但仍存在一定局限性。当前框架主要针对八叉树结构下的几何占用建模,在极稀疏或超高分辨率点云场景下仍有进一步优化空间,且尚未涉及几何与属性信息的联合压缩。未来将围绕模型的鲁棒性与可扩展性展开深入研究,针对复杂点云场景进一步优化层级特征与上下文建模方法,并探索几何与属性信息的协同编码机制,并拓展至动态及实时压缩场景,以进一步提升方法的工程实用性与应用价值。

参考文献(References)

- Chang A X, Funkhouser T, Guibas L, Hanrahan P, Huang Q, Li Z, et al. 2015. Shapenet: An information-rich 3d model repository. arXiv preprint. [DOI:10.48550/arXiv.1512.03012]
- Choy C, Gwak J and Savarese S. 2019. 4d spatio-temporal convnets: Minkowski convolutional neural networks//2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR). California, USA: IEEE: 3070 - 3079 [DOI:10.1109/CVPR.2019.00319]
- Cui M, Long J, Feng M, Li B and Kai H. 2023. OctFormer: Efficient octree-based transformer for point cloud compression with local enhancement//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA: AAAI Press: 470 - 478 [DOI:10.1609/aaai.v37i1.25121]
- d'Eon E, Harrison B, Myers T and Chou P A. 2017. 8i voxelized full bodies-a voxelized point cloud dataset. ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document WG11M40059/WG1M74006, 7(8), 11. [DOI:10.57702/698qsiw2]
- Fu C, Li G, Song R, Gao W and Liu S. 2022. Octattention: Octree-based large-scale contexts model for point cloud compression//Proceedings of the AAAI conference on artificial intelligence. Virtual Event: AAAI Press: 625 - 633 [DOI:10.1609/aaai.v36i1.19942]
- Graziosi D, Nakagami O, Kuma S, Zaghetto A, Suzuki T and Tabatabai A. 2020. An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and geometry-based (G-PCC). APSIPA Transactions on Signal and Information Processing, 9, e13. [DOI:10.1017/ATSIP.2020.12]
- Han K, Xiao A, Wu E, Guo J, Xu C and Wang Y. 2021. Transformer in transformer//Advances in neural information processing systems. Virtual Event: Curran Associates, Inc: 15908 - 15919 [DOI:10.48550/arXiv.2103.00112]

- Hu Y, Gong R and Wang Y. 2025. Bits-to-photon: End-to-end learned scalable point cloud compression for direct rendering//2025 IEEE International Conference on Image Processing (ICIP). AK, USA: IEEE: 953 – 958 [DOI:10.1109/ICIP55913.2025.11084608]
- Jiang X, Zhang X, Gao N and Deng Y. 2024. When fast fourier transform meets transformer for image restoration//European conference on computer vision. Milano, Italy: Springer: 381 – 402 [DOI: 10.1007/978-3-031-72995-9_22]
- Jin W, Zhou J, Li N and Liu X. 2026. Global frequency-aware multi-scale feature learning for point cloud normal estimation. *Engineering Applications of Artificial Intelligence*, 164, 113227. [DOI: 10.1016/j.engappai.2025.113227]
- Jin Y, Zhu Z, Xu T, Lin Y and Wang Y. 2024. Ecm-opcc: Efficient context model for octree-based point cloud compression//ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Seoul, Republic of Korea: IEEE: 7985 – 7989 [DOI: 10.1109/ICASSP48485.2024.10446374]
- Kim D, Kim C, Ko H and Hwang S J. 2025. Fourier decomposition for explicit representation of 3d point cloud attributes. *arXiv preprint arXiv:2503.10055*. [DOI:10.48550/arXiv.2503.10055]
- Li D, Li L, Chen Z and Li J. 2025. Shiftwiseconv: Small convolutional kernel with large kernel effect//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE: 25281 – 25291 [DOI: 10.1109/CVPR52734.2025.02354]
- Li G, Gao W and Gao W. 2024. Mpeg geometry-based point cloud compression (g-pcc) standard. In *Point Cloud Compression: Technologies and Standardization* (pp. 135 – 165). Springer. [DOI:10.1007/978-981-97-1957-0_7]
- Loop C, Cai Q, Escolano S O and Chou P A. 2016. Microsoft voxelized upper bodies-a voxelized point cloud dataset. ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document m38673 M, 72012(42), 2016. [DOI:10.57702/11beg7xg]
- Lyu W, Ke W, Sheng H, Ma X and Zhang H. 2024. Dynamic downsampling algorithm for 3D point cloud map based on voxel filtering. *Applied Sciences*, 14(8), 3160. [DOI:10.3390/app14083160]
- Martins N A, Cruz L A d S and Lopes F. 2024. Impact of LiDAR point cloud compression on 3D object detection evaluated on the KITTI dataset. *EURASIP Journal on Image and Video Processing*, 2024 (1), 15.
- Nguyen D T and Kaup A. 2022. Learning-based lossless point cloud geometry coding using sparse tensors//2022 IEEE International Conference on Image Processing (ICIP). Bordeaux, France: IEEE: 2341 – 2345 [DOI:10.1109/ICIP46576.2022.9897827]
- Nguyen D T and Kaup A. 2023. Lossless point cloud geometry and attribute compression using a learned conditional probability model. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(8), 4337 – 4348. [DOI:10.1109/TCSVT.2023.3239321]
- Nguyen D T, Quach M, Valenzise G and Duhamel P. 2021a. Learning-based lossless compression of 3d point cloud geometry//ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto, ON, Canada: IEEE: 4220 – 4224 [DOI:10.1109/ICASSP39728.2021.9414763]
- Nguyen D T, Quach M, Valenzise G and Duhamel P. 2021b. Multiscale deep context modeling for lossless point cloud geometry compression//2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). Shenzhen, China: IEEE: 1 – 6 [DOI: 10.1109/ICMEW53276.2021.9455990]
- Nguyen H X, Trestian R, To D and Tatipamula M. 2021c. Digital twin for 5G and beyond. *IEEE Communications Magazine*, 59(2), 10 – 15. [DOI:10.1109/MCOM.001.2000343]
- Que Z, Lu G and Xu D. 2021. Voxelcontext-net: An octree based framework for point cloud compression//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE: 6038 – 6047 [DOI:10.1109/CVPR46437.2021.00598]
- Schwarz S, Preda M, Baroncini V, Budagavi M, Cesar P, Chou P A, et al. 2018. Emerging MPEG standards for point cloud compression. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 9(1), 133 – 148. [DOI:10.1109/JETCAS.2018.2885981]
- Song R, Fu C, Liu S and Li G. 2023. Efficient hierarchical entropy model for learned point cloud compression//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE: 14368 – 14377 [DOI: 10.1109/CVPR52729.2023.01381]
- Tsang C F a S, Maria and Lalleche, Jean-Francois and Perel, Or and Loop, Charles and Takikawa, et al. 2024. Kaolin: A PyTorch Library for Accelerating 3D Deep Learning Research [EB/OL]. <https://github.com/NVIDIAGameWorks/kaolin>
- Wang J, Ding D, Li Z, Feng X, Cao C and Ma Z. 2022. Sparse tensor-based multiscale representation for point cloud geometry compression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45 (7), 9055 – 9071. [DOI: 10.1109/TPAMI. 2022. 3225816]
- Wang J, Ding D, Li Z and Ma Z. 2021. Multiscale point cloud geometry compression//2021 Data Compression Conference (DCC). Snowbird, UT, USA: IEEE: 73 – 82 [DOI:10.1109/DCC50243.2021.00013]
- Wang P-S, Liu Y, Guo Y-X, Sun C-Y and Tong X. 2017. O-cnn: Octree-based convolutional neural networks for 3d shape analysis. *ACM Transactions On Graphics (TOG)*, 36(4), 1 – 11. [DOI:10.1145/3072959.3073608]
- Wang X, Zhang Y, Liu T, Liu X, Xu K, Wan J, et al. 2025. TopNet: Transformer-efficient occupancy prediction network for octree-structured point cloud geometry compression//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE: 27305 – 27314 [DOI: 10.1109/

CVPR52734.2025.02543]

Wen C, Long J, Yu B and Tao D. 2024. PointWavelet: Learning in spectral domain for 3-D point cloud analysis. IEEE Transactions on Neural Networks and Learning Systems, 36 (3), 4400 - 4412. [DOI:10.1109/TNNLS.2024.3363244]

Xu Y, Lu Y and Wen Z. 2017. OwlII dynamic human mesh sequence dataset//ISO/IEC JTC1/SC29/WG11 m41658, 120th MPEG Meeting. 8 [DOI:10.57702/2bp0n9tx]

You K, Chen T, Ding D, Asif M S and Ma Z. 2025. Reno: Real-time neural compression for 3d lidar point clouds//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE: 22172 - 22181 [DOI: 10.1109/CVPR52734.2025.02065]

Yu J, Wang J, Sun L, Wu M-E and Zhu Q. 2022. Point cloud geometry compression based on multi-layer residual structure. Entropy, 24 (11), 1677. [DOI:10.3390/e24111677]

Yuan H, Ding D D, Zhang W, Gao W, Xu Y L, Liu Q, et al.2026.

Research progress of 3D point cloud coding and communication. Journal of Image and Graphics, 31 (6): 1619-1670 (元辉, 丁丹丹, 张伟, 高伟, 徐昇凌, 刘祺, et al. 2026. 三维点云编码通信进展. 中国图象图形学报, 31 (6), 1619 - 1670) [DOI: 10.11834/jig.250625]

作者简介

江照意,男,副教授,主要研究方向为计算机图形学,虚拟现实。E-mail:zyjiang@zjgsu.edu.cn

杨柏林,通信作者,男,教授,主要研究方向为数字媒体、大数据可视分析、虚拟现实、人工智能。E-mail:ybl@zjgsu.edu.cn

沈宇辉,男,硕士研究生,主要研究方向为计算机视觉。E-mail:1258321999@qq.com

朱佳俊,男,硕士研究生,主要研究方向为计算机视觉。E-mail: 24020100097@pop.zjgsu.edu.cn