DOI:10.19651/j.cnki.emt.2520074

基于特征域最近邻搜索和拼接的动态点云压缩算法

山 寿
(中国飞行试验研究院 西安 710089)

摘 要:动态点云在沉浸式通信与自动驾驶等前沿应用中具有重要价值,其高效压缩是实现实时传输与存储的关键。尽管已有基于规则与学习的方法在点云几何压缩方面取得进展,但现有方法在动态序列的帧间相关性利用上仍显不足。本文提出一种基于特征域最近邻搜索与拼接的动态点云几何压缩方法,将多尺度稀疏表示框架扩展至动态场景,并引入多尺度时间先验以增强帧间条件编码。具体而言,通过从已重建参考帧中提取分层特征,并与当前帧特征在特征域执行最近邻搜索与拼接,构建跨空间的时空上下文信息,从而更精确地估计体素占用概率。该方法在编码端仅传输部分特征,解码端结合参考帧信息重构时间先验,显著提升了压缩效率。实验在遵循 MPEG 通用测试条件的标准数据集上进行,结果表明,所提方法在多个测试序列上相较于现有基于规则与学习的压缩方法,在 D1-PSNR 与 D2-PSNR 下均取得大于 10%的显著 BD-Rate 增益,尤其在宽比特率范围内展现出优越的率失真性能,验证了其在动态点云几何压缩中的有效性与先进性。

关键词:动态点云压缩;特征域;最近邻搜索
中图分类号:TP2;TN919.8 **文献标识码**:A **国家标准学科分类代码**:510.1050

Research on dynamic point cloud compression algorithm based on feature domain nearest neighbor search and splicing

Shan Shou
(Chinese Flight Test Establishment, Xi'an 710089, China)

Abstract: Dynamic point clouds have significant value in cutting-edge applications such as immersive communication and autonomous driving, and their efficient compression is key to achieving real-time transmission and storage. Although rule-based and learning-based approaches have made progress in point cloud geometry compression, existing methods are still insufficient in leveraging inter-frame correlations in dynamic sequences. This paper proposes a dynamic point cloud geometry compression method based on feature-domain nearest neighbor search and concatenation, extending the multi-scale sparse representation framework to dynamic scenes, and introducing multi-scale temporal priors to enhance inter-frame conditional coding. Specifically, by extracting hierarchical features from the reconstructed reference frames and performing nearest-neighbor search and concatenation with the current frame features in the feature domain, spatiotemporal contextual information across spaces is constructed, thereby enabling more accurate estimation of voxel occupancy probabilities. This method transmits only partial features at the encoding end, and the decoding end reconstructs the temporal priors using reference frame information, significantly improving compression efficiency. The experiment was conducted on a standard dataset following the MPEG general test conditions. The results indicate that the method proposed in this paper achieves significant BD-Rate gains of over 10% in terms of D1-PSNR and D2-PSNR on multiple test sequences compared to existing rule-based and learning-based compression methods, particularly demonstrating superior rate-distortion performance across a wide range of bitrates. The test results validate the effectiveness and advancement of the algorithm proposed in this paper for dynamic point cloud geometry compression.

Keywords: dynamic point cloud compression; feature-domain; nearest-neighbor search

0 引 言

动态点云在全息通信、自动驾驶设备等应用中具有重要意义,而高效压缩动态点云几何(point cloud geometry, PCG)对于服务提供至关重要。除了国际标准化组织运动图像专家组(International Organization for Standardization,

收稿日期:2025-10-21

ISO/International Electrotechnical Commission Moving Picture Expert Group，ISO/IEC MPEG）所提出的基于规则的点云几何压缩（point cloud geometry compression，PCGC）技术，如基于视频的 PCC（video-based point cloud compression，V-PCC）和基于几何的 PCC（geometry-based point cloud compression，G-PCC）[1-3]之外，其展现出的高效压缩能力，为 SLAM 系统生成的大规模、动态点云数据的存储与传输提供了可行的技术路径，展现出在机器人、自动驾驶等实时感知系统中广阔的应用前景[4-7]。在这些学习型方案中，多尺度稀疏表示（multi scale representation，MSR）[8-11]通过有效利用静态 PCG 同一帧内的跨尺度和同尺度相关性，实现了前所未有的性能提升。

本研究将最初用于静态 PCG 的 MSR 框架发展用于压缩动态 PCG[8-9]。在此方面，本研究引入多尺度时间先验用于帧间条件编码。对于先前重建的 PCG 帧（例如参考帧），本研究逐步对其进行下采样并提取分层特征，随后将其作为额外的时间先验传输给当前帧的同尺度 PCG 张量以辅助压缩。为此，本研究将来自帧间参考的同尺度时间先验与来自同一帧内较低尺度的空间先验拼接，形成上下文信息，以更好逼近压缩过程中的条件占用概率。这种针对动态 PCGC 的帧间条件编码方案是在最初为静态 PCG 设计的 SparsePCGC[9]基础上实现的，用于定量评估其效率。实验结果表明，在遵循 MPEG 标准化委员会使用的通用测试条件（CTC）[12]的情况下，所提方法在有损模式优于现有基于规则的和基于学习的方法。

除现有的 G-PCC 和 V-PCC 标准及其他基于规则的 PCC 方法[13-18]外，近年来涌现出大量基于学习的 PCC 解决方案。因此，ISO/IEC MPEG 三维图形编码组启动了基于人工智能的点云压缩（artificial intelligence，AI-PCC）项目，旨在研究更好的点云压缩潜在技术。

最近的研究主要致力于研究静态 PCG 的压缩[4]，也称为静态 PCGC，产生了基于体素的[19-20]、基于点的[21-24]、基于八叉树的[25-28]和基于稀疏张量的方法[8-10]。

基于体素的压缩方法曾主导早期点云压缩研究，典型代表如文献[19-20]。这类开创性方法通常通过体素化将点云转换为三维体积表征，随后将其分割为 64×64×64 体素的立方体区块。压缩过程采用自编码器架构，通过三维卷积运算将这些区块编码为紧凑的潜在表征。在模型优化方面，这些方法主要采用专门的损失函数，特别是焦点损失或加权二元交叉熵损失，以解决数据固有的不平衡问题。然而，这类方法存在显著的计算效率低下问题——由于大部分体素为空，导致计算资源和内存资源的浪费。

基于点的方法在点云处理领域开创了独特范式，无需借助体素化或其他结构化转换即可保持原始点云表征。这类方法主要采用 PointNet[21]及其改进版 PointNet＋＋[22]架构，通过逐点全连接层直接处理原始点云。典型实现采用基于补丁的处理方式，通过远端点采样实现高效子采样，并结合 K 近邻（K nearest neighbors，KNN）搜索算法构建逐点特征嵌入，最终形成基于多层感知机（multi-layer perception，MLP）的自动编码器框架。然而多项研究[23-24]表明，这类逐点模型存在显著局限性，尤其在编码效率方面仍逊色于其他方法。此外，这些方法在大规模密集点云场景中存在扩展性不足的问题，泛化能力也较为有限。另一个缺陷在于计算开销较大，由于需要大量预处理和后处理操作，整体编码效率显著降低。

基于八叉树的方法[25-28]已成为点云表示的高效解决方案，兼具优异的存储效率和计算性能。这些技术通过复杂的熵上下文建模，结合节点的层级依赖关系（父节点）和空间邻近关系（子节点），精准预测各节点的占用概率。该领域的发展历程展现了上下文建模的持续进步：OctSqueeze[25]和 MuSCLE[28]率先采用 MLP 捕捉父节点与子节点间的层级依赖关系；在此基础上，VoxelContextNet[27]通过整合父节点、邻近节点及体素化邻域点信息，拓展了上下文建模范式，显著提升了概率估计精度。最新突破 OctAttention[28]通过引入大规模 Transformer 架构的上下文注意力模块，大幅扩展了占用概率估计的感知域，实现了性能的质的飞跃。这些无损编码方法在稀疏激光雷达点云处理中展现出卓越性能，为该领域树立了新的基准。然而，基于八叉树的方法在稠密点云压缩领域仍存在应用难题。

近期基于稀疏卷积的方法[8-9]首先将原始点云序列转换为 Minkowski 稀疏张量[29]，通过利用稀疏性特征实现高效的大规模数据深度网络处理，从而提升局部与全局三维特征提取能力。其中，基于稀疏张量的方法不仅具有领先的性能，而且复杂度低。第 1 个有代表性的工作是 PCGCv2[8]，其中使用基于稀疏卷积神经网络（sparse convolutional neural network，SparseCNN）的自动编码器对静态 PCG 张量进行分层下采样和有损压缩。后来，SparsePCGC[6]在统一的 MSR 框架下对 PCGCv2 进行了极大改进，通过广泛利用跨尺度和同尺度相关性，使用基于 SparseCNN 的占用概率近似（SparseCNN-based occupancy probability approximation，SOPA）模型进行更好的上下文建模，以支持各种点云的无损和有损压缩，并在静态点云压缩领域处于领先地位。

传统基于规则的点云压缩技术主要以二维投影方法为主导。二维投影技术[30-31]充分利用现有视频编解码技术，多数研究聚焦于投影算法开发与补丁配置优化。例如，He 等[30]实现了立方体投影技术，而 Zhu 等[31]提出了一种结合视图全局投影与补丁局部投影的混合方案。此外，MPEG 提出的专为动态点云（dynamic point cloud，DPC）压缩设计的 V-PCC 标准[2]。该标准通过整合立方体投影、补丁生成与打包流程，生成几何图和纹理图。V-PCC 在性能表现上堪称业界标杆，其压缩效率在 2D 投影压缩方案中处于领先地位。

而深度学习方法由于其独特的非线性学习优势，在动态

点云几何压缩中有超越传统方法的趋势。在 PCGCv2 的基础上,Akhtar 等[32-33]提出通过编码时间连续的 PCG 帧之间的残差来实现动态 PCGC。他们的主要区别在于 inter 预测信号的生成方式不同。Akhtar 等[32]提出利用前一帧预测当前帧的方法,通过减少码流冗余实现了动态点云压缩。Akhta 等[32]设计了一种预测器,通过在目标坐标上执行卷积操作,融合前一帧的多尺度特征以实现运动补偿。然而,其网络仅基于相邻帧的几何相似性,从前一帧特征中进行空间域预测,既未隐式也未显式地学习和利用运动等时序信息,缺乏用于指导帧间预测的显式运动估计/运动补偿网络。缺乏时序信息会制约当前帧预测的准确性。

为此,Fan 等[33]提出了带有端到端特征域运动估计/运动补偿模块的 D-DPCC 框架,该模块通过计算连续两帧的潜在特征进行后续运动估计,并提出可推导运动补偿的三维自适应加权插值。尽管 D-DPCC 在 DPC 压缩框架中表现优异,但其仅通过单次运动估计/运动补偿实现帧间预测,且运动估计网络同样基于目标坐标卷积,仍存在大量时间依赖关系有待探索。为提升所提帧间预测模块的运动估计效率,Xia 等[34]进一步设计了 KNN-注意力块匹配(KNN attention block matching,KABM)网络,该网络通过几何特征和特征相关性来评估潜在对应点的影响。然而,与目标位置卷积类似,KABM 使用的 KNN 算法仍是在空间域上索引近邻,当不相关的点距离较近,引入 KNN 索引点的特征会降低上下文的准确性。

进一步,Zhou 等[35]提出了一种基于时空变换器建模的动态点云压缩框架 DPCC-STTM,用于在潜在空间中压缩点云序列。为有效提取并充分利用时间上下文,DPCC-STTM 引入了变换器 Transformer 模块,该模块通过时域内容的相关性对时空信息进行有效建模。此外,DPCC-STTM 还开发了多尺度处理模块,能够捕捉多帧点云中短程和长程的时间相关性。DPCC-STTM 采用空间域上算术操作,直接将当前帧的点和参考帧的同坐标点的特征进行加减操作,仍未摆脱空间域上索引的局限性。

为克服上述问题,本文提出特征域索引近邻的方法,利用部分已传输特征和参考帧特征的相关性来索引近邻,使网络更易找到参考帧中与待编码点相关的点,提升上下文的相关性和准确性。

1　所提方法

1.1　整体框架

所提出的基于 MSR 的动态 PCGC 如图 1 所示。以两帧为例,其中参考帧已编码并重建作为时间参考,当前帧即将被编码。显然,这两帧的编解码示例可轻松扩展到一串帧序列。

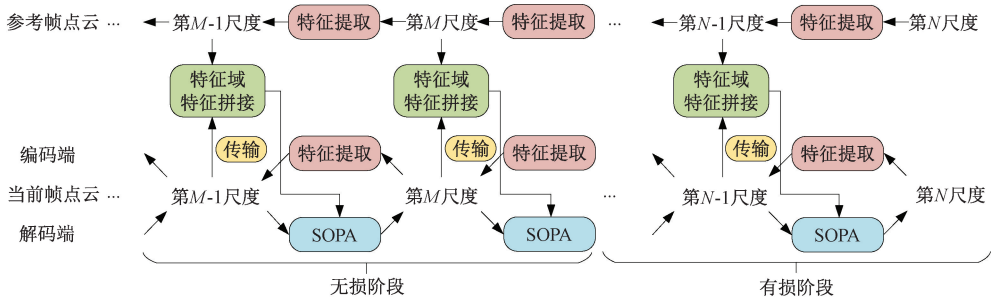

图 1　整体框架

Fig. 1　Overall framework

为了压缩当前帧,一种直接的方法是独立编码每帧 PCG,即帧内编码(intra coding),使用默认的 SparsePCGC 仅能利用同一帧内的跨尺度和同尺度相关性。由于相邻帧间存在强时间相关性,通常采用帧间预测(inter prediction)以提高压缩效率。为此,本研究遵循 MSR 原理,首先,在编码端使用特征提取器(Extractor)从当前帧提取特征,然后将当前帧的部分特征进行编码传输。其次,在解码端将参考帧重建结果经过同样的特征提取器提取特征并与编码端传输过来的特征在特征域执行最近邻索引算法,将被索引的特征与传输过来的特征进行特征拼接,生成时间先验,用于当前帧的条件编码。

类似于 SparsePCGC,二进制重采样被用于多尺度计算。假设输入点云有在第 N 帧的最高尺度,该 PCG 的有损压缩由 M 个尺度的无损压缩和(N-M)个尺度的有损压缩组成。调整 M 可以平衡有损率失真权衡。如图 2 所示,在无损阶段,时间先验与同帧低尺度空间先验拼接后输入 SOPA 模型,以更好地逼近占用概率用于无损编码;而在有损阶段,此类拼接的时空先验可选择与解码后的局部邻域信息增强,或直接用于 SOPA 模型以改善几何重建效果。感知残差网络(inception-ResNet,IRN)包含多个卷积核大小为 1 和 3 的卷积层及多个分支的残差连接[5]。

接下来将详细说明为在条件编码中使用时间先验而开发的各个模块。

1.2　特征提取模块

特征提取模块通常设计为分辨率下采样结构,通过聚合局部邻域信息形成空间内先验,以增强 SOPA 模型性能。相应地,SOPA 模型从低到高尺度逐步估算占用概率,用于几何重建(即体素占用状态),其利用了同一帧中

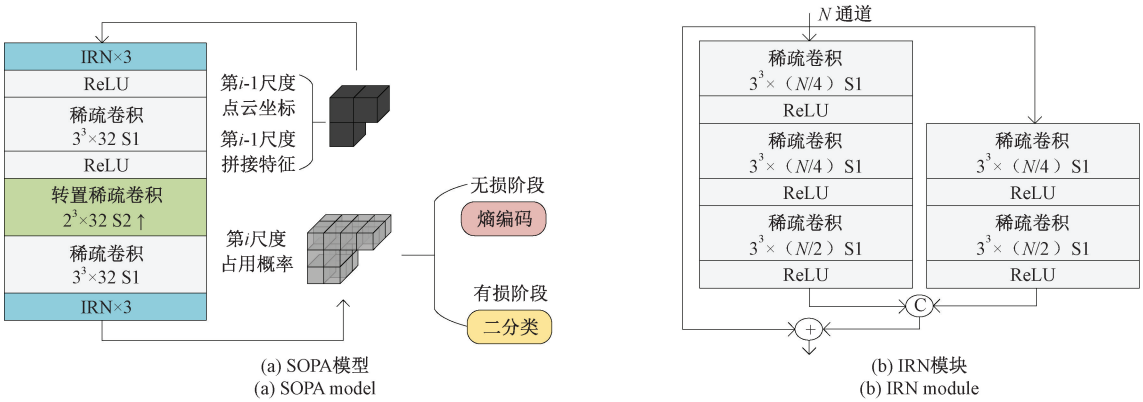

图 2　概率估计模型

Fig. 2　The probability estimation model

的空间先验(例如解码的潜在特征、低尺度输入)和来自参考帧的时间先验。

特征提取模块采用稀疏卷积和非线性激活函数进行计算,如图 3 所示,包括一个卷积体素下采样层(每个维度的卷积核大小和步长均为 2),以及堆叠的 IRN 块用于深层特征聚合。

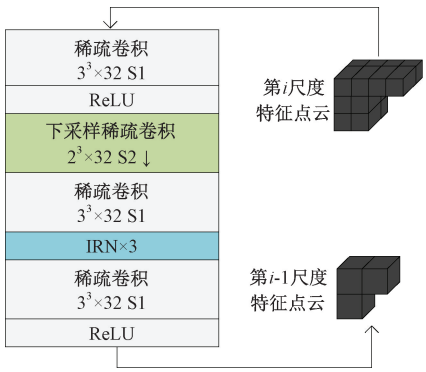

图 3　特征提取模块

Fig. 3　The feature extraction module

1.3　特征拼接

本研究使用在特征域的特征拼接将参考帧的信息传递给当前帧以实现压缩。如图 4 所示,特征拼接由两部分实现。

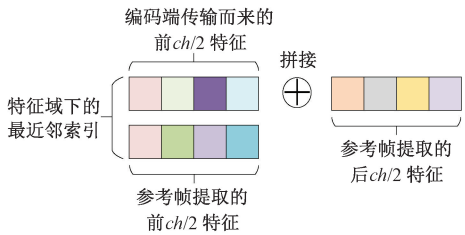

图 4　特征索引与拼接

Fig. 4　The feature searching and splicing

首先,对于编码端传输而来的特征,将其中任意一点的特征与同尺度参考帧提取的所有点的前 $ch/2$ 个特征(ch 是特征的通道数)执行 L2 距离下的特征域的最近邻索引算法:

$$k^* = \mathrm{argmin}_{i\in\{1,\cdots,n\}} \| \boldsymbol{x}_i - \boldsymbol{q} \|_2 \tag{1}$$

其中,$\boldsymbol{q} = (q_1, q_2, \cdots, q_{ch/2}) \in \mathbb{R}^{ch/2}$ 是当前帧的任意一个查询点的特征。$\boldsymbol{x}_i = (x_{i,1}, x_{i,2}, \cdots, x_{i,ch/2}) \in \mathbb{R}^{ch/2}$ 是参考帧的第 i 个特征。k^* 是索引到的特征序号。

其次,将 $\boldsymbol{q}$ 与 $\boldsymbol{x}_{k^*}$ 执行拼接操作,形成一个完整的特征向量作为该点的特征。在解码端对所有传输来的特征执行上述算法,得到当前帧的时间先验。

该方法灵活且适用于 MSR 框架下的所有尺度。参考帧提取用于生成逐尺度的时间先验,这些先验随后与同一帧中的空间先验拼接,以帮助无损和有损压缩。通过这种方式,本研究突破了目标位置卷积的只能在局部邻域获取参考信息的限制,并将其用于当前帧的压缩,使编解码器能够自适应地选取任意空间处的参考信息用于占用概率估计。在无损模式下,它能生成比特率消耗更低的比特流;而在有损模式下,它有助于用更少失真更好地重建几何结构。

1.4　损失函数

为了量化体素占用概率,本研究使用二元交叉熵(binary cross entropy,BCE)损失来衡量编码占用状态所需的比特率。同时,BCE 损失也代表了有损压缩中的几何失真。对于编码器中潜在特征的压缩,本研究使用一个简单的全分解熵模型[33]来估计其概率,并使用交叉熵损失计算比特率 R。总损失函数是 BCE 损失 D 与比特率消耗 R 的组合,即 $J = R + \lambda D$,其中 λ 是用于调整率失真权衡的权重。

2　实验与分析

2.1　实验设置

1)训练与测试数据集

本研究使用 8i Voxelized Full Bodies(8iVFB)数据集[36]进行训练,使用 Owlii 动态人体序列数据集[37]进行测试。训

练数据集包含 5 个序列:longdress、loot、redandblack、soldier、queen,每个序列有 300 帧,几何精度为 10 bit。测试数据集包含 4 个序列:basketball_player、dancer、model、exercise,它们均量化为 10 bit 几何精度。训练与测试样本的划分遵循 MPEG AI-PCC 组[12]使用的探索实验建议。

2)训练策略

在训练过程中,本研究将每帧划分为 4 个块,使用 kdtree 进行分割,并逐步降采样到 4 个不同尺度以实现数据增强。本研究训练一个用于无损编码的模型和 5 个用于有损编码的模型。通过调整有损阶段的 M 以及损失函数中的率失真权重 λ,本研究获得 5 个不同的有损编码模型,覆盖从 0.01～0.18 bpp(每点比特数的比特率范围)。

3)测试条件

测试遵循 AI-PCC 组针对动态 PCGC 定义的通用测试条件(common test condition,CTC)。第 1 帧以帧内模式编码,后续所有 99 帧使用时间上最接近的重建作为参考。结果对使用 100 帧的情况取平均值。比特率通过每输入点的平均比特数(bit per point,bpp)评估每个序列。几何失真通过每帧的 D1-PSNR 和 D2-PSNR 评估,以得到序列级平均值(包括第 1 帧)。

2.2　编码性能

对于有损编码,选择 V-PCC[38]进行比较,因为它在动态有损 PCGC(点云几何压缩)中表现最佳;在此本研究应用 V-PCC 中的默认低延迟 HEVC 视频编码[39]。此外,本研究还与其他基于学习的 PCGC 方法进行对比,包括最初为静态 PCGC 设计的 SparsePCGC[9],以及由 Fan 等[33]、Xia 等[34]提出的两种动态 PCGC 方法(D-DPCC、LDPCC)和 Zhou 等[35]提出的基于 Transformer 的方法 TransDPCC。对于 SparsePCGC,每一帧 PCG 都独立编码为帧内模式(intra mode),不使用帧间预测。关于基于学习的方法[32-35],由于它们均在 MPEG AI-PCC 工作组中研究,并遵循 CTC 训练和测试标准[12],本研究直接引用最新标准中报告的结果[40-41]以确保公平比较。

如表 1 和 2 所示,所提方法相对于 V-PCC 基准平均获得 75.05%和 71.60%增益。在与学习的 PGCC 方法的比较中,本研究相对与目前静态 PCGC 的最先进学习方法(SparsePCGC,SPCGC)的 BD-Rate 增益为 10.31%和 15.13%。由于 Akhtar 等[32]所提方法和 TransDPCC 并未提供源码,因此本文使用其论文中仅提供的 D1-PSNR 数据进行对比。

表 1　相对于基线方法在 D1-PSNR 下的 BD-Rate 增益

Table 1　The BD-Rate gains under D1-PSNR compared with baselines　%

序列	player	dancer	model	exercise	average
SPCGC	8.52	9.68	13.29	9.75	10.31
文献[32]	40.72	44.82	47.31	31.95	41.20
D-DPCC	23.70	30.57	27.53	20.78	25.65
LDPCC	14.71	22.02	16.69	9.83	15.81
TransDPCC	34.68	—	46.77	41.46	40.97
V-PCC	74.66	73.82	76.40	71.49	75.05

表 2　相对于基线方法在 D2-PSNR 下的 BD-Rate 增益

Table 2　The BD-Rate gains under D2-PSNR compared with baselines　%

序列	player	dancer	model	exercise	average
SPCGC	14.80	14.04	18.18	13.49	15.13
文献[32]	—	—	—	—	—
D-DPCC	40.93	45.92	42.04	35.88	41.19
LDPCC	33.74	39.89	33.27	26.91	33.45
TransDPCC	—	—	—	—	—
V-PCC	69.74	76.40	71.23	77.64	71.60

在与学习的动态 PCGC 方法的比较中,本研究相对于目前最先进学习方法,包括文献[32]、D-DPCC[33]、LDPCC[34]和 TransDPCC[35]的 D1-PSNR 下的 BD-Rate 增益分别为 41.20%、25.65%、15.81%、40.97%。由于文献[32]和 TransDPCC 未提供 D2-PSNR 下的数据,因此与 D-DPCC 和 LDPCC 相比,D2-PSNR 下的 BD-Rate 增益为

41.19%和 33.45%。

本研究还在图 5 中可视化了相应的率失真曲线。结果表明，本研究的方法在更广泛的比特率范围内始终优于其他方法。还可观察到，LDPCC 和 D-DPCC 专注于高比特率，无法达到低于 0.1 bpp 的比特率；而文献[32]主要适用于低比特率但在高比特率下表现较差。这主要是由于各自有损阶段的固定尺度设置所致，即 LDPCC 和 D-DPCC 在有损压缩时下采样 2 次，文献[32]下采样 3 次。相比之下，本研究的方法提供了灵活的尺度调整（即高/中/低比特率下自适应 M，即 $M\in(1,2,3)$），并通过简单但有效的特征拼接实现多尺度帧间条件编码。这些改进实现了最先进的性能。

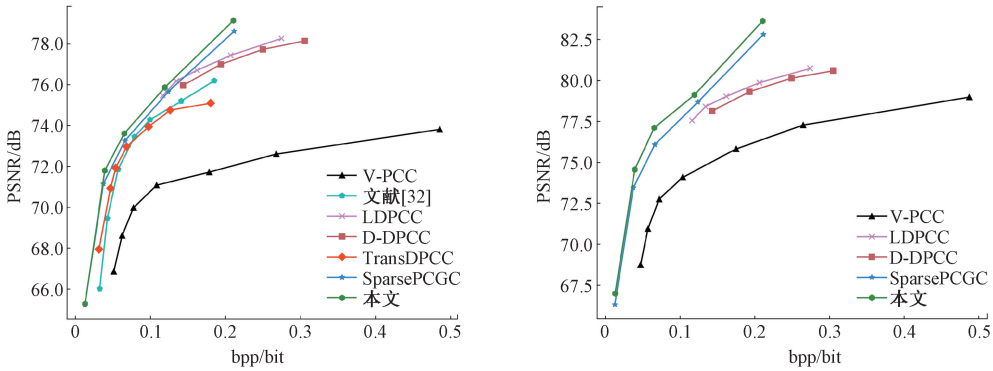

图 5　与基线方法比较的 D1-PSNR 和 D2-PSNR 下的率失真曲线

Fig. 5　The R-D curves under D1-PSNR and D2-PSNR compared with baselines

在主观视觉质量方面，如图 6 所示，本研究比较各个方法对点云序列 basketball_player 的重建效果。为确保评估的公平性，本研究为每种方法选取了相似的码率点。实验结果表明：DDPCC 的码率为 0.196 6 bpp，D1-PSNR 和 D2-PSNR 分别为 76.43、79.13 dB；LDPCC 码率略升至 0.200 3 bpp，对应的 D1-PSNR 和 D2-PSNR 提升至 77.39、79.77 dB；SparsePCGC 码率进一步增加至 0.210 3 bpp，D1-PSNR 和 D2-PSNR 分别达到 78.56、82.10 dB；所提方法的码率为 0.208 2 bpp，略低于 SparsePCGC，其 D1-PSNR 为 78.23 dB，稍逊于 SparsePCGC，但 D2-PSNR 高达 83.70 dB，显著优于所有对比方法。这一结果说明所提方法在保证较低码率的前提下，能够实现更优的点云重构精度，尤其是在 D2-PSNR 指标上具备明显优势。

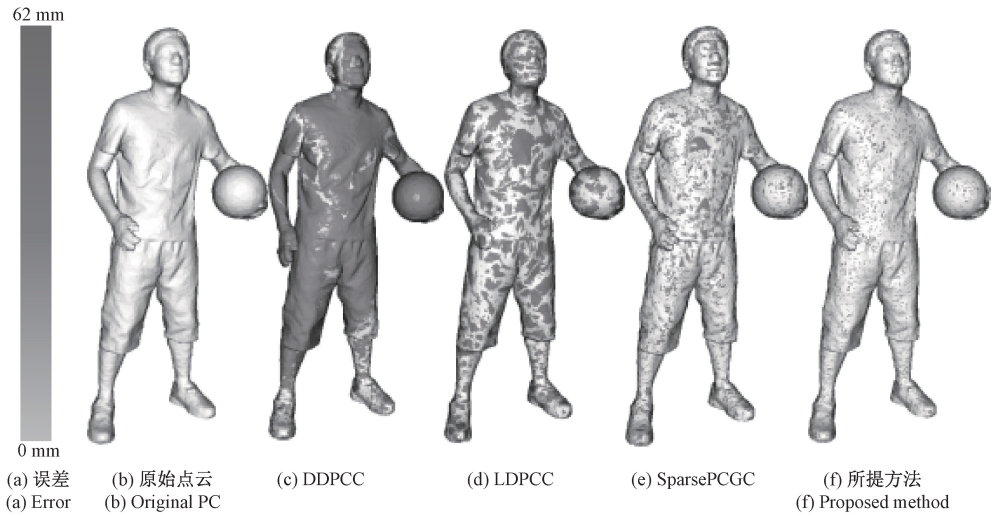

图 6　与基线方法比较的主观质量

Fig. 6　The subjective quality compared with baselines

为了表明使用特征域近邻索引参考帧上下文的有效性，如图 7 所示，本研究展示了使用空间域近邻与特征域近邻索引到的参考帧上的点。其中，灰色点表示当前帧或者参考帧的原始 10 bit 点云；蓝色点表示当前帧原始点云下采样至 9 bit 的点，由于近邻索引是在 N-1 或更低尺度上进行的，因此算法需要为蓝色点寻找近邻；红色点为在

9 bit 参考帧点云上使用空间域的最近邻查找得到的点,由胳膊和篮球处的的放大图可以看到,红色点大多分布在边界上,与均匀分布的蓝色明显无法对应的很好;绿色点为在 9 bit 参考帧点云上使用特征域最近邻索引到的点,由局部放大图可以看出,绿色分布较为均匀与蓝色点的分布可以对应,但是分布也较为稀疏,这是由于多个蓝色点可能索引到同一个绿色点。

如表 3 和 4 所示,本研究还评估了不同方法的编解码时间复杂度和显存占用。所有点云序列均在固定配置的 GPU 和 CPU 上进行顺序处理,操作系统为 Ubuntu 20.04。编解码时间和内存占用均为测试中所有码率点的平均耗时。从表中可以看出,传统方法 V-PCC 的编码时间最长,基于学习的方法编码时间相当。在解码时间方面,所有方法的耗时均相当。在显存方面,所有方法的编码/解码占用均相当。

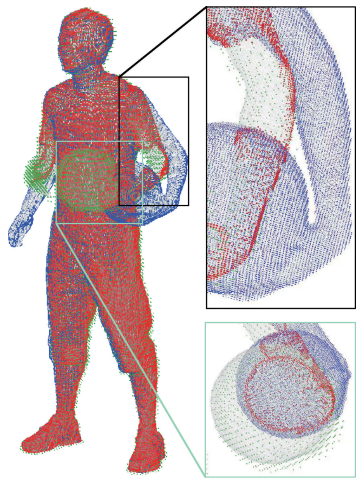

图 7　空间域和特征域近邻索引的示意

Fig. 7　The of searching in spatial domain and feature domain

表 3　不同方法的编码/解码时间

Table 3　The encode/decode time of different methods　s

序列	player	dancer	model	exercise	average
SPCGC	1.29/1.54	1.46/1.53	1.33/1.49	1.47/1.54	1.39/1.53
文献[32]	—	—	—	—	—
D-DPCC	1.43/1.45	1.39/1.45	1.11/1.14	1.37/1.39	1.33/1.36
LDPCC	1.44/1.46	1.36/1.40	1.12/1.14	1.35/1.38	1.32/1.35
TransDPCC	—	—	—	—	—
V-PCC	80.53/1.31	81.49/1.35	80.78/1.06	81.24/1.26	81.01/1.25
所提方法	1.36/1.93	1.34/1.67	1.39/1.67	1.34/1.60	1.36/1.71

表 4　不同方法的编码/解码的显存占用

Table 4　The encode/decode memory occupancy of different methods　GB

序列	player	dancer	model	exercise	average
SPCGC	9.60/9.78	7.59/7.61	8.13/8.47	8.32/8.65	8.41/8.63
文献[32]	—	—	—	—	—
D-DPCC	8.84/10.65	8.19/9.74	7.64/9.16	7.85/9.74	8.13/9.82
LDPCC	8.78/11.06	7.88/9.93	7.42/9.43	7.73/9.65	7.95/10.02
TransDPCC	—	—	—	—	—
所提方法	8.71/10.99	7.82/9.87	7.36/9.37	7.68/9.60	7.89/9.96

3　结　　论

本研究提出了一种基于特征域最近邻搜索与拼接的动态点云几何压缩方法,将多尺度稀疏表示框架从静态点云拓展至动态序列,通过引入特征域时间先验建模机制,有效突破了传统空间域索引在局部几何约束下的局限性,提升了上下文建模能力,实验结果表明该方法在 MPEG 通用测试条件下相较于 V-PCC、SparsePCGC、D-DPCC、LDPCC 及 TransDPCC 等主流方法在 D1/D2-PSNR 指标下均有较大增益,验证了其在动态点云几何压缩中的先进性与实用价值。

然而本研究仍存在特征域匹配鲁棒性受限、多帧参考机制尚未挖掘、动态属性压缩未涉及等问题,未来需进一步探索学习型度量空间,引入时序注意力机制与多帧融合策略以挖掘长时序依赖,拓展联合几何-属性压缩框架以实现更完整的动态点云编码系统。

参考文献

[1]　GRAZIOSI D, NAKAGAMI O, KUMA S, et al. An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and

geometry-based(G-PCC)[J]. APSIPA Transactions on Signal and Information Processing, 2020, 9: e13.

[2] SCHWARZ S, PREDA M, BARONCINI V, et al. Emerging MPEG standards for point cloud compression[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2018, 9(1): 133-148.

[3] CAO CH, PREDA M, ZAKHARCHENKO V, et al. Compression of sparse and dense dynamic point clouds—Methods and standards[J]. IEEE, 2021, 109(9): 1537-1558.

[4] QUACH M, PANG J H, TIAN D, et al. Survey on deep learning-based point cloud com-pression[J]. Frontiers in Signal Processing, 2022, 2: 846972.

[5] 吴永豪，李胜，邹文成. 动态场景中的多传感器融合 SLAM 算法研究[J]. 电子测量与仪器学报，2025, 39(11): 1-10.
WU Y H, LI SH, ZOU W CH. Research on multi-sensor fusion SLAM algorithm in dynamic scenes[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(11): 1-10.

[6] 侯余鑫，介婧，侯北平，等. 动态场景下基于稀疏光流的语义视觉 SLAM 算法[J]. 电子测量技术，2025, 48(16): 60-69.
HOU Y X, JIE J, HOU B P, et al. Semantic visual SLAM algorithm based on sparse optical flow in dynamic scenes[J]. Electronic Measurement Technology, 2025, 48(16): 60-69.

[7] 范明泽，徐晓苏. 基于多阶段动态过滤的静态点云地图生成算法[J]. 仪器仪表学报，2025, 46(4): 193-205.
FAN M Z, XU X S. Static point cloud map generation algorithm based on multi-stage dynamic filtering[J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 193-205.

[8] WANG J Q, DING D D, LI ZH, et al. Multiscale point cloud geometry compression[C]. 2021 Data Compression Conference(DCC), 2021: 73-82.

[9] WANG J Q, DING D D, LI ZH, et al. Sparse tensor-based multiscale representation for point cloud geometry compression[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(7): 9055-9071.

[10] LIU G X, WANG J Q, DING D D, et al. PCGFormer: Lossy point cloud geometry compression via local self-attention[C]. 2022 IEEE International Conference on Visual Communications and Image Processing(VCIP), 2022: 1-5.

[11] XUE R X, WANG J Q, MA ZH. Efficient LiDAR point cloud geometry compression through neighborhood point attention[J]. ArXiv preprint arXiv: 2208.12573, 2022.

[12] WG7, MPEG 3D graphics coding, description of exploration experiment 5.3 on AI-based dynamic pc coding[R]. ISO/IEC JTC 1/SC 29/WG7 N00386, 2022.

[13] PEIXOTO E. Intra-frame compression of point cloud geometry using dyadic decomposition[J]. IEEE Signal Processing Letters, 2020, 27: 246-250.

[14] GU SH, HOU J H, ZENG H Q, et al. 3D point cloud attribute compression via graph prediction[J]. IEEE Signal Processing Letters, 2020, 27: 176-180.

[15] RAMALHO E, PEIXOTO E, MEDEIROS E. Silhouette 4D with context selection: Lossless geometry compression of dynamic point clouds[J]. IEEE Signal Processing Letters, 2021, 28: 1660-1664.

[16] THANOU D, CHOU P A, FROSSARD P. Graph-based compression of dynamic 3D point cloud sequences[J]. IEEE Transactions on Image Processing, 2016, 25(4): 1765-1778.

[17] GARCIA D C, FONSECA T A, FERREIRA R U, et al. Geometry coding for dynamic voxelized point clouds using octrees and multiple contexts[J]. IEEE Transactions on Image Processing, 2020, 29: 313-322.

[18] GU SH, HOU J H, ZENG H Q, et al. 3D point cloud attribute compression using geometry-guided sparse representation[J]. IEEE Transactions on Image Processing, 2019, 29: 796-808.

[19] WANG J Q, ZHU H, LIU H J, et al. Lossy point cloud geometry compression via end-to-end learning[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(12): 4909-4923.

[20] GUARDA A F R, RODRIGUES N M M, PEREIRA F. Adaptive deep learning-based point cloud geometry coding[J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15(2): 415-430.

[21] QI CH R, SU H, MO K CH, et al. Pointnet: Deep learning on point sets for 3D classification and segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652-660.

[22] QI C R, YI L, SU H, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space[J]. Advances in Neural Information Processing Systems, 2017, 30: 5099-5108.

[23] YOU K, GAO P. Patch-based deep autoencoder for

point cloud geometry compression[C]. 3rd ACM International Conference on Multimedia in Asia, 2021: 1-7.

[24] XIE L, GAO W, FAN S L, et al. Pdnet: Parallel dual-branch network for point cloud geometry compression and analysis[C]. 2024 Data Compression Conference(DCC), 2024: 596-596.

[25] HUANG L, WANG SH L, WONG K, et al. Octsqueeze: Octree-structured entropy model for lidar compression[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1313-1323.

[26] BISWAS S, LIU J, WONG K, et al. Muscle: Multi sweep compression of lidar using deep entropy models[J]. Advances in Neural Information Processing Systems, 2020, 33: 22170-22181.

[27] QUE Z ZH, LU G, XU D. Voxelcontext-net: An octree based framework for point cloud compression[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 6042-6051.

[28] FU CH Y, LI G, SONG R, et al. Octattention: Octree-based large-scale contexts model for point cloud compression[C]. AAAI Conference on Artificial Intelligence, 2022, 36(1): 625-633.

[29] CHOY C, GWAK J Y, SAVARESE S. 4D spatio-temporal convnets: Minkowski convolutional neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3075-3084.

[30] HE L Y, ZHU W J, XU Y L. Best-effort projection based attribute compression for 3D point cloud[C]. 2017 23rd Asia-Pacific Conference on Communications (APCC), 2017: 1-6.

[31] ZHU W J, MA ZH, XU Y L, et al. View-dependent dynamic point cloud compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(2): 765-781.

[32] AKHTAR A, LI ZH, VAN DER AUWERA G, Inter-frame compression for dynamic point cloud geometry coding[J]. IEEE Transactions on Image Processing, 2024, 33: 584-594.

[33] FAN T Y, GAO L Y, XU Y L, et al. D-DPCC: Deep dynamic point cloud compression via 3D motion prediction[C]. 31st International Joint Conference on Artificial Intelligence, 2022: 898-904.

[34] XIA SH T, FAN T Y, XU Y L, et al. Learning dynamic point cloud compression via hierarchical inter-frame block matching[C]. 31st ACM International Conference on Multimedia, 2023: 7993-8003.

[35] ZHOU Y CH, ZHANG X F, MA X Q, et al. Dynamic point cloud compression with spatio-temporal transformer-style modeling[C]. 2024 Data Compression Conference(DCC), 2024: 53-62.

[36] D'EON E, HARRISON B, MYERS T, et al. 8i voxelized full bodies-a voxelized point cloud dataset[R]. Geneva: ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) m38673/M72012, 2017.

[37] XU Y, LU Y, WEN Z Y. Owlii dynamic human mesh sequence dataset[R]. Macau: ISO/IEC JTC1/SC29/WG11 m41658, 120th MPEG Meeting, 2017.

[38] ISO/IEC 23090-5: 2023. Information technology—Coded representation of immersive media—Part 5: Visual volumetric video-based coding(V3C) and video-based point cloud compression(V-PCC)[S]. 2023. ISO/IEC 23090-5:2023 | IEC.

[39] SULLIVAN G J, WIEGAND T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Magazine, 1998, 15(6): 74-90.

[40] AKHTAR A, LI Z, VAN DER AUWERA G, et al. Results dynamic point cloud compression[R]. Online: 8th WG7 Meeting, 139th MPEG Meeting, 2022.

[41] XU Y L, FAN T Y, GAO L Y, et al. D-DPCC test results on 10 bit owlii[R]. Online: 8th WG7 Meeting, 139th MPEG Meeting, 2022.

作者简介

山寿(通信作者),硕士,高级工程师,主要研究方向为计算机视觉、新质遥测技术、试飞遥测监控技术等。

E-mail:shaaanshou@163.com