动态点云压缩技术的研究发展与综述

杨柏林[1)], 应瑞琦[1)], 丁丹丹[2)], 南方哲[3)], 赵俊杰[4)], 郑东[5)]

[1)] (浙江工商大学计算机科学与技术学院 杭州 310018)
[2)] (杭州师范大学信息科学与技术学院 杭州 311121)
[3)] (浙江工商大学统计与数学学院 杭州 310018)
[4)] (中国移动(杭州)信息技术有限公司 杭州 310023)
[5)] (杭州宇泛智能科技股份有限公司 杭州 311121)
(ybl@zjgsu.edu.cn)

摘 要: 动态点云数据因具备时间连续性和高维数据特性, 能准确地描述物体动态变化, 广泛应用于自动驾驶、增强现实、虚拟现实、机器人导航和三维视频等. 动态点云压缩在有效地管理不断增长的点云数据量的存储、传输和感知方面发挥着重要作用, 当前缺乏专门针对动态点云压缩方法的全面调查. 针对该问题, 对动态点云压缩方法进行了系统的综述. 首先介绍动态点云压缩的研究意义和理论支撑; 然后从几何压缩、属性压缩和联合压缩 3 个方面阐述了主流模型的基本原理及优势; 之后介绍了主流的动态点云压缩数据集及评价指标, 总结各种方法在不同数据集上的性能; 最后分析了当前方法在算法复杂度与场景适应性方面的局限性, 并指出轻量化网络设计、联合时空建模及跨模态信息融合是未来的主要研究方向. 文中总结了动态点云压缩的最新进展, 有助于更深入理解动态点云压缩, 促进基于动态点云压缩系统的进一步发展.

关键词: 动态点云压缩; 深度学习; 几何压缩; 属性压缩; 联合压缩
中图法分类号: TP391.9 **DOI:** 10.3724/SP.J.1089.2024-00699

Research Progress and Review of Dynamic Point Cloud Compression Technology

Yang Bailin[1)], Ying Ruiqi[1)], Ding Dandan[2)], Nan Fangzhe[3)], Zhao Junjie[4)], and Zheng Dong[5)]

[1)] (*School of Computer Science and Technology, Zhejiang Gongshang University, Hangzhou* 310018)
[2)] (*School of Information Science and Technology, Hangzhou Normal University, Hangzhou* 311121)
[3)] (*School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou* 310018)
[4)] (*China Mobile (Hangzhou) Information Technology Company Limited, Hangzhou* 310023)
[5)] (*Universal Ubiquitous Co., Ltd., Hangzhou* 311121)

Abstract: Dynamic point cloud data, characterized by temporal continuity and high-dimensional features, can accurately capture dynamic changes of objects and are thus widely applied in fields such as autonomous driving, augmented reality, virtual reality, robotic navigation, and 3D video. Dynamic point cloud compression plays a crucial role in efficiently managing the storage, transmission, and perception of the ever-increasing volume of point cloud data. Despite this demand, there remains a clear lack of comprehensive surveys specifically focused on dynamic point cloud compression methods. To address this gap, this paper presents a systematic review of existing approaches, aiming to summarize recent advancements in the field and provide insights for future research.

收稿日期: 2024-11-16; 修回日期: 2025-06-24. **基金项目**: 国家自然科学基金(62172366); 浙江省“尖兵”“领雁”研发攻关计划(2023C01150). **杨柏林**(1975—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为数字媒体、大数据可视分析、虚拟/增强现实、人工智能; **应瑞琦**(2001—), 男, 硕士研究生, 主要研究方向为动态点云压缩; **丁丹丹**(1983—), 女, 博士, 副教授, 硕士生导师, CCF 会员, 主要研究方向为计算机视觉、智能视频图像处理、智能视频编码、三维点云压缩编码与重建; **南方哲**(1994—), 女, 博士研究生, CCF 会员, 主要研究方向为动态点云压缩; **赵俊杰**(1985—), 男, 硕士, 高级工程师, 主要研究方向为视联网; **郑东**(1990—), 男, 博士研究生, 高级工程师, 主要研究方向为智能物联.

The paper first introduces the significance and theoretical foundations of dynamic point cloud compression. It then elaborates on the fundamental principles and advantages of mainstream models from three perspectives: geometry compression, attribute compression, and joint compression. Subsequently, commonly used datasets and evaluation metrics are reviewed, and the performance of various methods across different datasets is summarized. Finally, the paper analyzes the limitations of current methods in terms of algorithmic complexity and scene adaptability, and points out that lightweight network design, joint spatiotemporal modeling, and cross-modal information fusion are major directions for future research. This survey is expected to facilitate a deeper understanding of dynamic point cloud compression and promote further development of related systems.

Key words: dynamic point cloud compression; deep learning; geometric compression; attribute compression; joint compression

随着点云硬件采集设备的快速发展, 点云数据作为一种重要的三维数据表示形式广泛应用于自动驾驶、虚拟现实、建筑等领域, 如实时感知环境、用户互动、建筑物监测等. 点云由具有 x, y, z 坐标及颜色、法线、反射率等属性的点组成, 与网格模型[1]和体素数据[2]相比, 其结构简单、表达灵活. 点云分为静态和动态 2 类, 静态点云描述固定时间点的三维结构, 动态点云则通过连续帧展示对象的时间变化. 然而, 点云数据量通常非常庞大, 特别是在处理动态点云时数据量更是呈指数级增长, 对存储和传输带来了巨大的挑战, 因此高效的动态点云压缩技术成为了研究热点. 近年来, 研究人员提出多种高效的动态点云压缩方法, 通过结合实时数据处理技术、高效的压缩方法和先进的重建方法, 可以有效地应对动态点云压缩面临的挑战, 在各种应用中充分地发挥动态点云的优势.

本文对当前动态点云压缩方向的主要方法和技术进行总结和归纳. 首先讨论其基本原理、优缺点及适用场景; 然后详细阐述动态点云在几何和属性压缩方面的研究进展; 再介绍相关的公开数据集及其评价标准, 对当前代表性方法的性能进行总结; 最后对动态点云压缩的发展方向进行展望.

1 相关工作

动态点云压缩技术旨在对三维点云数据进行编码与压缩, 以减少多帧点云数据存储空间和传输带宽, 同时尽可能地保留原始数据的几何特征和重要信息. 动态点云压缩技术性能的优化对高效地传输和存储大规模三维数据至关重要. 借助先进的技术, 可以显著地提升动态点云压缩的效率和质量.

根据压缩方法的理论基础和实现方式, 动态点云压缩可细分为基于规则的方法和基于深度学习的方法. 基于规则方法包括八叉树、投影(图像与视频编码)和熵编码等, 基于深度学习的方法包括稀疏卷积、自动编码器和生成对抗网络等. 在实际研究中, 不同方法之间存在一定程度的交叉与融合, 这些方法为静态点云压缩[3-9]和动态点云压缩技术提供了理论基础和技术支撑.

1.1 八叉树方法

八叉树是一种广泛应用于三维空间中的空间划分数据结构, 具有空间分层管理、自适应分辨率、数据简化与降采样、快速邻域搜索和层次化数据访问等特性, 使动态点云压缩方法能够更高效、灵活地处理大规模三维点云数据. 八叉树具有递归分割空间的特点, 能够根据不同的分辨率需求调整结构, 在处理大规模数据时具有明显的优势.

八叉树编码[10-11]将空间递归地划分为 8 个子立方体, 并进一步细分这些子立方体, 直至达到预设的分辨率或某个终止条件, 这种递归划分方式非常适用于处理三维点云数据, 能有效地表示和管理稀疏和密集区域. 利用八叉树的层次结构, 可以借助父节点的信息对当前的体素的占用率进行编码, 这种分层上下文显著地提高了压缩性能, 使得压缩不仅考虑点云数据的几何结构, 还能根据点的分布自适应地调整细节处理, 进一步优化压缩效果. 根据点云的结构, 可以自适应地将其划分为多分辨率体素块, 并使用八叉树进行标记, 有效地解决稀疏性带来的空间冗余问题.

动态点云中, 八叉树同样具有显著优势, 核心是将时间序列中每帧的点云构建为对应的八叉树结构, 通过在连续帧之间建立节点级别的对应关系实现帧间的冗余建模. 在八叉树编码过程中引入时空上下文模型, 利用前一帧的八叉树节点状

态预测当前帧的编码概率分布, 降低熵编码成本, 使得压缩框架不仅具备空间适应性, 还具备时间连续性, 有效地提升动态点云的压缩效率.

1.2　图像与视频压缩方法

动态点云由一系列时序相关的三维点组成, 类似于三维空间中的视频序列, 因此传统的图像和视频压缩技术在动态点云压缩中具有天然的应用潜力. 通过将动态点云投影到二维平面, 其空间结构可以转化为适合图像和视频压缩技术[12-14]处理的形式, 使得三维点云数据能够利用现有的图像和视频编码技术进行压缩. 在处理包含时序变化的动态点云数据时, 投影技术不仅能够保持较高的还原质量, 还提供了灵活且高效的压缩框架, 使得基于图像和视频压缩技术的动态点云压缩方法能够充分地发挥其优势. 因此, 投影技术为图像和视频编码方法在动态点云压缩中的应用起到了重要的桥梁作用.

最初, 图像压缩的实现依赖自动编码器. Ballé 等[15]联合非线性分析变换、均匀量化器和非线性综合变换, 并利用随机梯度下降的变体进行优化, 实现了在各个比特率下显著的视觉质量提升; 进一步, Ballé 等[16]提出基于变分自编码器的图像压缩模型, 通过结合超先验和侧信息捕捉潜在表示中的空间依赖性, 并与底层自编码器联合训练复杂先验.

当前, 在视频压缩方向已有多个标准和方法应运而生[17-19], 大多数方法采用自动编码器, 同时引入帧插值、运动估计与补偿等技术减少时间和空间的冗余. 帧插值方法首先利用神经网络对视频中的帧进行时间插值, 然后对残差[20]进行编码. 基于光流的运动补偿方法通过估计光流并对其进行压缩, 再通过双线性插值预测当前帧, 该方法能够更准确地捕捉视频序列中的运动信息, 进一步提高了视频压缩的效果[21].

1.3　熵编码方法

熵编码是动态点云压缩方向的重要技术之一, 通过优化数据表示减少数据量, 在保持原始信息完整性的同时, 显著地提高了压缩效率和质量. 熵编码的基本原理是利用输入数据的统计特性, 将频率较高的符号用较短的编码表示, 而频率较低的符号用较长的编码表示, 从而在整体上减少数据的平均编码长度. 熵编码方法的优势是能够自适应地根据输入数据的分布特性进行编码, 大大提高了编码效率. 在动态点云数据中, 由于点的空间分布和属性变化具有一定的统计规律, 熵编码能够有效地捕捉并利用这些规律实现高效的数据压缩[22].

OctAttention[23]中的熵编码方案如图 1 所示, 通过对兄弟节点和祖先节点的上下文进行建模, 以利用相邻节点之间的强依赖关系. 熵编码的无损性, 使得动态点云压缩模型能够在保持高压缩比的同时减少数据恢复过程中的损失, 为后续研究中采用熵编码方法打下了坚实的基础.

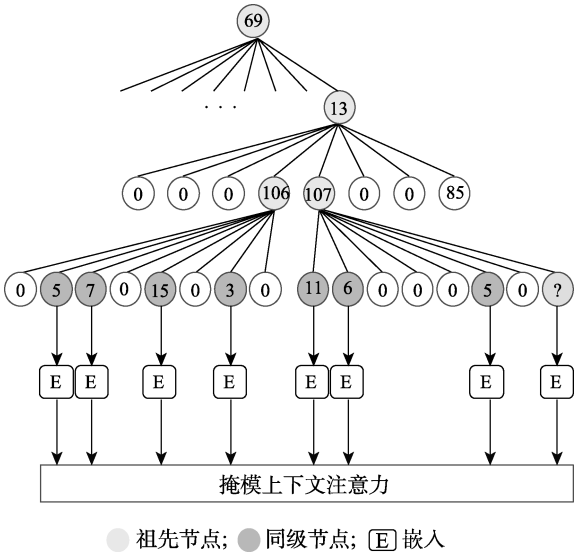


图 1　OctAttention[23]熵编码方案

1.4　深度学习方法

深度学习在动态点云压缩中的作用和重要性日益凸显, 通过自动学习数据的复杂特征, 提供了高度灵活和高效的压缩方法. 深度学习方法能够从大量的点云数据中提取出有效的特征表示, 减少冗余信息, 并在重建过程中保持较高的精度. 与传统卷积不同, 稀疏卷积利用数据的稀疏性只在特定的非零元素上进行运算, 显著地减少计算量和内存占用. 这种方法通过专门设计的稀疏卷积核, 使得网络能够在处理稀疏数据时保持较高的效率和精度.

自适应编码器指能够根据输入数据的特性动态地调整其结构或参数的编码器. 通过深度学习技术, 自适应编码器可以根据数据分布的变化自主地选择合适的编码策略, 在压缩性能和重建质量之间取得最佳平衡.

基于卷积神经网络(convolutional neural network, CNN)的自编码器将输入转换为较低维度的潜在空间, 并重建与输入相同的输出. Quach 等[24]是最早将 CNN 引入点云压缩的工作之一, 为后续研究奠定了基础; Choy 等[25]使用稀疏卷积网络, 解决了在体素网格上使用 CNN 导致的时间和空间

复杂性问题. 通过这种方法, 点云压缩的速度得到显著提升, 且在较低的比特率下依然保持较高的重建精度.

对于动态点云, 深度学习不仅能提取单帧的空间特征, 还能建模帧间的时间相关性, 其核心是构建时空联合的特征表示, 通过学习点在连续帧中的运动趋势与分布模式, 有效地去除时间维度上的冗余. 例如, 通过预测当前帧与参考帧之间的残差或运动场, 结合深度网络进行帧间补偿编码, 在保持压缩效率的同时提升重建质量[26].

1.5 小 结

动态点云压缩技术包括基于规则的方法和基于深度学习的方法, 它们通过不同的技术手段实现对点云数据的高效压缩和重建. 基于规则的方法, 如八叉树、投影、熵编码等具有理论清晰、稳定性强的优点; 基于深度学习的方法, 如稀疏卷积、自动编码器和生成对抗网络等, 通过特征提取和表示提高压缩效率和重建质量, 特别是在处理复杂场景和动态变化的点云数据时, 表现出显著的优势.

2 动态点云压缩方法

本文中, 将点云压缩方法分为几何、属性和联合 3 类, 如图 2 所示, 对各类动态点云压缩方法进行系统性的分析. 每类方法又细分为多种方法, 如基于八叉树的压缩、基于距离图像的压缩等. 通过该分类结构, 深入探讨每类方法的理论基础、技术实现及其在动态点云压缩中的应用与发展, 为读者提供全面且严谨的研究进展分析, 为后续的技术创新和应用奠定坚实的理论基础.

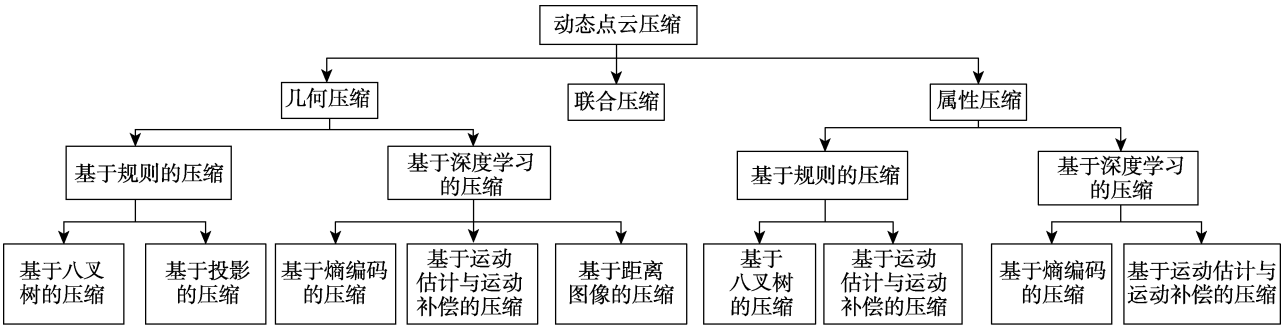


图 2 点云压缩方法分类

2.1 动态点云几何压缩

几何压缩是动态点云压缩研究中的一个关键方向. 几何信息包括点的位置、密度和形状等, 通过高效的几何压缩, 能够实现更快的数据传输速度和更低的存储成本; 同时, 确保在解压缩后恢复出高质量的点云数据, 满足各种应用的需求. 由于动态点云数据量大且冗余度高, 因此有效的几何压缩技术对于减小存储空间和传输带宽至关重要.

早期, 动态点云几何压缩的研究主要集中于静态点云压缩, 即在无时间维度变化的情况下对单帧点云进行压缩. 传统的静态几何压缩方法, 如八叉树划分[27]和三维体素化[28], 通过空间分层和区域划分显著地减少了数据量. 2006 年, Schnabel 等[29]首次将八叉树用于点云压缩, 通过局部表面近似实现了高压缩率; Octsqueeze[30]中利用点之间的稀疏性和结构冗余, 将 LiDAR 点云编码为八叉树, 并设计条件熵模型降低比特率; Zhang 等[31]提出的无损帧内编码器结合投影和上下文自适应二进制算术编码器, 简化当前帧内压缩方法, 并通过半自适应编码器进一步提升了性能.

然而, 因为传统的静态点云压缩方法仅考虑单帧数据的空间结构, 没有考虑随时间变化的几何信息, 所以在处理动态点云时存在局限性. 在静态点云压缩的基础上, 动态点云几何压缩方法考虑了时间相关性. 目前, 基于几何的动态点云压缩技术分为 2 大类: 基于规则的方法和基于深度学习的方法.

2.1.1 基于规则的几何压缩

基于规则的几何压缩方法中, 利用点云数据的空间结构和统计特性对动态点云数据进行压缩. 动态点云由连续时刻的三维点集构成, 随着时间的推移, 这些点云数据不仅包含空间分布信息, 还反映出时序相关性, 基于规则的几何压缩通过识别和利用动态点云中反复出现的模式和规律, 能够在不显著损失精度的前提下减少存储和传输的开销.

基于规则的点云压缩方法的流程如图 3 所示, 具体实现过程中, 通常包含 3 个关键步骤: 首先对动态点云数据的预处理, 包括去除噪声和冗余数

据, 并根据时间序列特性进行时域优化; 然后通过分析点云的几何特征, 如点与点之间的距离、法向量和曲率等, 结合时序变化构建适应动态变化的压缩规则, 这些规则可以是显式的几何关系, 也可以是通过机器学习方法从大量动态点云数据中提取的隐式时序模式; 最后利用这些规则对动态点云数据进行编码和压缩, 以大幅减少数据量.

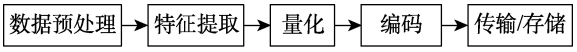


图 3　基于规则的点云压缩流程示意图

(1) 基于八叉树的动态点云几何压缩

基于八叉树的动态点云压缩方法的流程如图 4 所示, 利用八叉树数据结构对点云进行分割和组织, 该方法适应动态场景的特性, 通过递归地细分空间, 八叉树根据点云的密度和分布情况灵活地调整划分精度, 在高密度区域提供精细划分, 而在低密度区域则采用较低分辨率的划分, 这种自适应的空间划分方式使得八叉树能够有效地应对动态场景中的变化, 如物体的运动或遮挡. 当物体移动时, 八叉树能够对空间结构重新划分或调整分辨率, 仅对变化较大的区域进行细化, 而对静态区域使用较低分辨率, 优化数据存储和传输; 通过有效地组织和表示稀疏的点云数据, 实现了点云数据的高效表示和传输. 该技术不仅可提高数据的处理效率, 还能保留点云的几何结构、优化数据的存储和传输, 对于动态点云数据的处理具有重要意义[32].



图 4　基于八叉树的动态点云压缩流程示意图

2012 年, Kammerl 等[33]提出一种基于八叉树的数据结构进行空间分解点云流有损压缩的方法, 有效地检测并消除点云数据流中的时间冗余; 2014 年, Saranya 等[34]提出一种实时压缩策略, 采用基于八叉树的空间分解方法有效地消除空间冗余, 并通过编码相邻点云之间的结构差异, 成功地去除了时间冗余.

2020 年, Garcia 等[35]提出一种逐步增加八叉树分辨率的动态点云压缩方法, 在八叉树框架内引入局部体素上下文, 利用基于体素上下文的深度熵模型对非叶子八叉树节点进行无损压缩, 显著地提高了压缩效率; 同时, 重建后的点云在对象检测和语义分割等下游任务中表现出更高的精度.

(2) 基于投影的动态点云几何压缩

通常, 处理三维点云数据需要消耗大量的存储空间和计算资源, 而将其投影为二维数据后, 图像序列在时间与空间上表现出高度一致性[36], 降低数据复杂性并提高了处理效率; 当前, 图像和视频处理方向已有许多成熟且经过验证的技术, 可以高效地处理转换为二维格式的三维数据. 在动态点云压缩中, 通过如图 5 所示的流程, 将三维点云投影为二维图像[37]或视频[38], 然后利用现有图像和视频编码技术进行压缩, 这类方法具有显著优势.

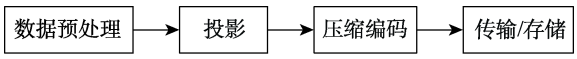


图 5　基于投影的点云压缩流程示意图

在 MPEG 发布的点云压缩标准 V-PCC[39]中, 通过将三维点云数据转化为二维视频, 利用现有的视频编码技术实现了高压缩率和高质量重建, 其架构如图 6 所示. V-PCC 采用几何投影和包装技术

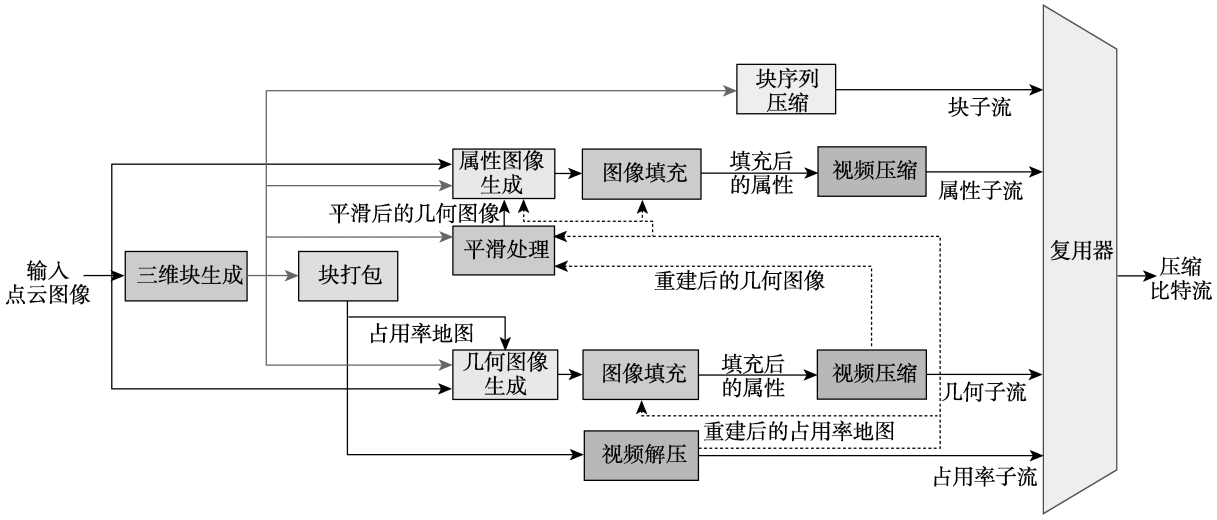


图 6　V-PCC[39]架构

将三维点云投影到二维平面上, 生成包含几何位置和颜色信息的深度图和属性图, 这些二维图像与视频编码器兼容, 简化了数据处理; 其中的帧间预测技术借鉴视频编码中的运动补偿方法, 利用相邻帧之间的几何时间相关性减少冗余. 该方法显著地降低点云数据的存储和传输成本, 同时保持较高的重建质量, 适用于虚拟现实、增强现实和自主驾驶等应用场景.

在 V-PCC 环境下, Li 等[40]提出一种估计二维属性视频中的二维运动矢量的方法, 其基于准确的二维几何视频重建几何体, 不仅丰富了点云数据压缩的理论, 还为实际应用提供了有价值的参考, 但其未利用局部投影; 相比之下, 2021 年, Zhu 等[41]提出 View-PCC, 在多视角和多层的全局投影基础上采用基于块的局部投影, 同时进行帧内位移排列和帧间时间对齐, 旨在最大化有效投影点的数量并保留时空连贯性, 实现更高效的压缩.

基于非深度学习的方法在计算复杂度和实现难度上相对较低, 适用于资源受限的应用场景; 而基于深度学习的方法则在压缩效率和处理复杂动态场景的能力上更具优势, 但需要较高的计算和数据资源支持. 在实际应用中, 应根据具体需求和条件选择合适的几何压缩方法, 实现最优的压缩效果和性能.

2.1.2 基于深度学习方法的几何压缩

深度学习模型能够有效地学习点云数据的特征和结构, 自适应地处理各种类型和复杂度的点云数据. 通过实现端到端的训练[42], 深度学习模型简化了处理流程, 同时捕捉数据的非线性关系和上下文信息, 这些模型具有强大的泛化能力, 能够更有效地压缩动态点云数据, 满足实际应用需求.

近年来, 深度学习方法在几何信息提取方面展现了显著优势, 并广泛应用于静态点云和动态点云压缩. 图 7 所示为基于深度学习的点云压缩流程, 与传统的依赖手工设计特征和规则的几何方法相比, 深度学习模型通过自动学习特征, 能够更好地捕捉点云数据的细节与整体结构, 实现点云的动态恢复[43]. CNN[44-46]、变分自编码[47]等深度模型通过处理点云数据的空间关系和拓扑结构实现高效的压缩, 并提供比传统方法更灵活和高效的解决方案. 进一步, 这些方法细分为基于熵编码、基于运动估计与运动补偿, 以及基于距离图像的动态点云几何压缩方法.

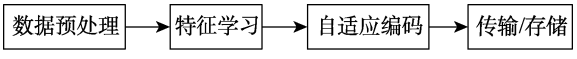


图 7 基于深度学习的点云压缩流程示意图

(1) 基于熵编码的动态点云几何压缩

在动态点云压缩中, 熵编码与深度学习的结合发挥了重要作用, 通过捕获高阶特性实现了数据的高效压缩和重要特性的保留. 如图 8 所示, 深度学习通过神经网络将点云数据编码为隐含表示, 捕捉其高阶特性, 为压缩提供丰富的信息. 这些隐含表示经过量化后送入熵编码器进行无损压缩, 熵编码器利用隐含特征的统计特性为频率高的特征分配较短编码, 实现高效的点云压缩.

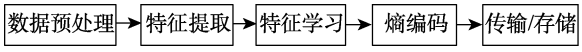


图 8 基于熵编码的点云压缩流程示意图

深度熵模型 VoxelContext-Net[48]中, 利用节点的局部体素上下文对八叉树中非叶节点的符号进行无损压缩, 使用深度神经网络估计每个节点符号的概率分布, 并通过最小化交叉熵损失优化模型, 有效地捕捉三维空间中的上下文信息, 显著地提高点云数据的压缩性能.

为了减少压缩过程中的损失, Riddle[49]通过预测栅格扫描序列中的下一个像素值, 并对预测值与实际值之间的残差进行熵编码, 实现无损压缩; 该方法依赖预测模型的准确性, 预测越精确残差的熵越小, 提高了压缩效率. 然而, 这种单向预测方法仅考虑前一个时间步的帧, 存在一定的局限性. 如图 9 所示, BIRD-PCC 模型[50]中引入基于深度学习的残差编码模块, 利用残差帧中的空间冗余信息, 通过双向策略处理无效像素, 进一步提高了压缩效率, 该方法通过双向处理策略克服单向预测方法的局限性, 实现了更高效的动态点云压缩.

(2) 基于运动估计与运动补偿的动态点云几何压缩

近年来, 针对动态点云数据的非均匀分布与时变性的特点, 图 10 所示的深度学习结合运动估计与运动补偿的方法逐渐受到研究者的关注. 2021 年, Gao 等[51]采用基于卷积的框架实现对点云几何的高效压缩, 通过逐点卷积实现所有计算, 更好地利用了输入的无约束几何相关性.

Akhtar 等[52]使用稀疏卷积结合分层多尺度的三维特征学习, 将先前帧的信息编码至当前帧中, 通过对目标坐标进行卷积操作, 将前一帧的潜在表示映射到当前帧的下采样坐标, 实现对当前帧特征嵌入的预测. 该方法是首个在所有比特率下都优于 V-PCC 帧间模式的基于深度学习的方法, 为后续工作提供了新的思路.

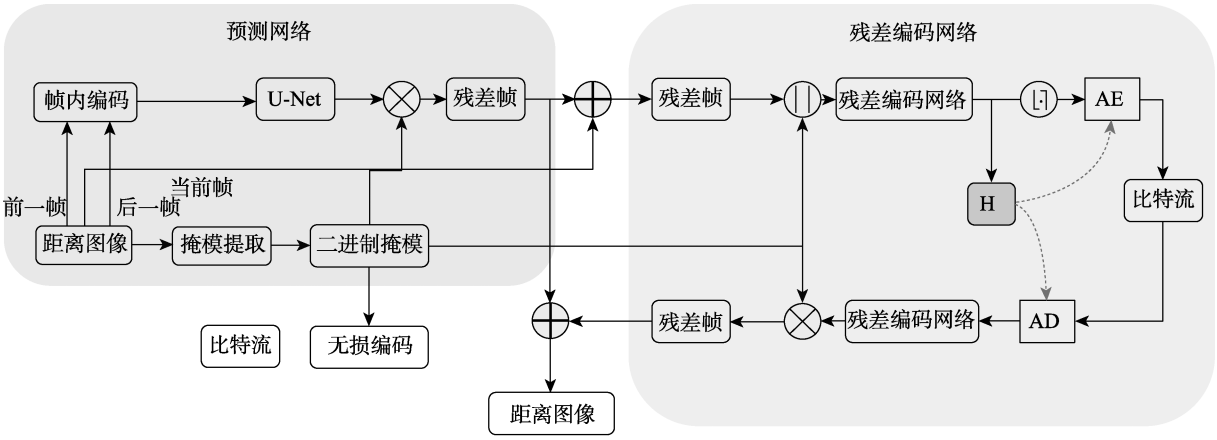


图 9　BIRD-PCC 模型[50]架构

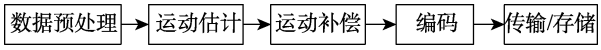


图 10　基于运动估计与运动补偿的点云压缩流程示意图

Fan 等[53]提出基于三维稀疏卷积的深度动态点云压缩网络 D-DPCC，其架构如图 11 所示，通过三维运动估计和运动补偿压缩点云几何信息；利用多尺度运动融合模块和三维自适应加权插值算法进行准确的三维光流估计和运动补偿，并采用有损自编码器网络压缩运动嵌入和残差，是首个端到端深度动态点云压缩框架，具有重要意义．尽管 D-DPCC 在许多动态点云压缩框架中表现出色，但其仅通过一次性的运动估计与运动补偿执行帧间预测，仍有许多时间依赖性问题需要进一步探索．基于此，2023 年，Xia 等[54]选择光流的粒度准确地封装运动信息，并提出了一个 *k*NN-注意力块匹配网络进行更精确的运动估计，解决了 D-DPCC 中网络过于简单的问题，显著地提高了压缩性能与精度；Jiang 等[55]则提出一种结合多尺度运动估计网络和残差压缩网络的端到端的网络．这些网络都利用深度学习的优势，实现更准确的运动估计和更有效的运动补偿．

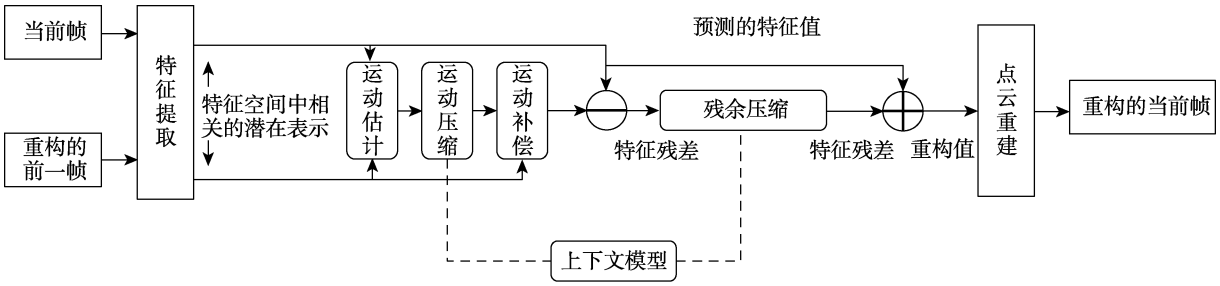


图 11　D-DPCC 模型[53]的架构

(3) 基于距离图像的动态点云几何压缩

在动态点云压缩中，将距离图像与深度学习结合展现出显著的优势，如图 12 所示．利用有损的距离图像可以大幅提高字节流压缩的效率，同时保持点云的高质量；此外，距离图像能够用于自适应特征调整，提取局部点特征，具有适应复杂局部结构的独特优势，因此在广泛的应用场景中展现出巨大的潜力．

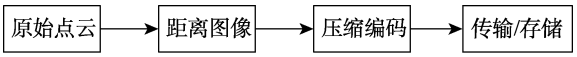


图 12　基于距离图像的点云压缩流程示意图

Sun 等[56]提出一种基于距离图像和预测神经网络的方法，将 LiDAR 点云数据压缩为距离图像，并通过高效视频编码算法实现点云序列的高效压缩，利用点云数据的结构特性通过簇内预测技术去除空间冗余，并使用卷积长短期记忆(long short-term memory, LSTM)网络预测未来帧，消除时间冗余．Beemelmanns 等[57]提出一种类似的方法，将三维点云数据转换为二维距离图像进行压缩，利用簇内预测技术和卷积 LSTM 网络分别去除空间和时间冗余，在毫米级精度下实现显著的压缩比，并充分地利用点云数据的空间信息，其架构

如图 13 所示. 这些方法通过结合距离图像和深度学习技术, 显著地提升了动态点云数据的压缩效率和质量.

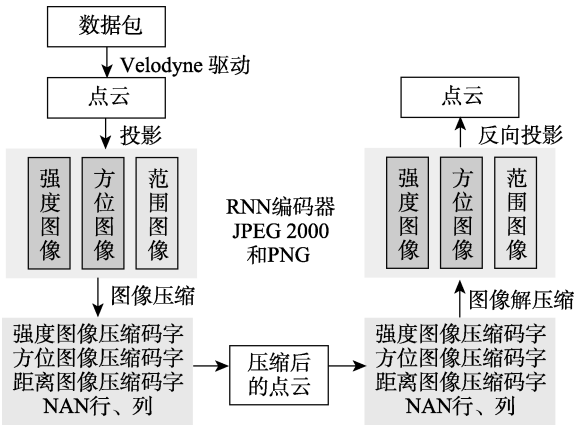


图 13 文献[57]模型的架构

深度学习的几何压缩方法中, 基于熵编码的技术利用神经网络精确地建模数据分布实现高效的压缩, 但训练和推理过程需要大量的计算资源. 基于运动估计的压缩方法通过深度学习模型自动学习点云中的运动模式, 能够适应复杂的动态变化, 但模型的训练数据需求较高且泛化能力受限. 基于距离图像的压缩方法将点云转换为距离图像进行处理, 利用 CNN 的优势进行高效压缩, 但在点云密度变化较大的场景中可能面临挑战.

2.2 动态点云属性压缩

点云属性压缩通过压缩点携带的属性, 如颜色、法线、纹理和反射率等, 解决稀疏性、不规则性和动态性等问题. 通常, 静态属性压缩方法[58-59]包括 2 个关键步骤: 一是属性的变换编码, 该步骤捕获数据空间中的冗余信息, 并将其转化为变换系数; 二是变换系数的熵编码, 通过符号无损压缩将这些变换系数转化为紧凑的比特流.

动态属性压缩方法在处理点云数据的时间和空间变化方面具有显著优势, 它们通常依赖时间相关性和空间相似性, 通过捕捉和利用这些特性提高压缩效率和重建质量. 本文将动态点云属性压缩分为基于规则和基于深度学习的方法.

2.2.1 基于规则的属性压缩

在动态点云属性压缩中, 基于规则的方法通常使用明确的数学模型和算法, 如迭代最近邻算法[60]和基于图的表示[61]等, 这些方法的内部逻辑和操作过程清晰, 易于理解和实现, 并在实际应用中表现出稳定的效果.

2015 年, Thanou 等[62]研究动态点云序列中的运动估计问题, 提出将点云序列视为图形集合, 并对这些图形进行特征提取和匹配, 最终在三维点云序列的压缩中利用运动估计进行颜色补偿. 2016 年, Thanou 等[61]利用时间连续的相似性和运动估计, 解决了点云帧中具有不同数量的点且缺乏明确对应信息的挑战性问题, 该方法使用图形表示时变几何, 并利用谱图小波描述子进行运动估计, 实现对点云序列颜色的有效压缩. 2022 年, Hong 等[63]提出基于块的分数体素运动估计方案, 考虑了点云的帧内和帧间体素分布的不规则性, 并证明运动补偿可以从分数精度提供的更高分辨率参考和更准确的位移中受益.

将八叉树和运动估计的方法相结合, 能够有效地捕捉动态点云序列中的运动信息, 提升数据压缩的效率与精度. 2017 年, Mekuria 等[60]提出一种通用的实时动态点云编解码器, 将八叉树顶点属性的颜色映射到图像网格, 并利用传统的基于 JPEG 的图像编码技术进行压缩; 实验结果表明, 该编码器在混合现实应用中表现出色, 与原始重建点云相比, 预测失真可忽略不计.

区域自适应分层变换(region-adaptive hierarchical transform, RAHT)是 MPEG 的几何点云压缩测试模型 G-PCC[39]的重要组成部分, 用于压缩颜色信号以及如反射率等其他属性, 通过自下而上的方式反向跟踪八叉树, 并在每步中进行变换和重组, 实现点云数据的高效压缩. 该方法充分利用点云数据的空间结构, 提高了压缩效率; 将 RAHT 与运动估计和运动补偿的结合, 在动态点云属性压缩中展现出显著的优势. 2021 年, Xu 等[64]将时间信息纳入用于颜色编码的简化 RAHT 的上下文中, 利用点云簇之间的运动构建基于图的属性压缩中的时空邻接关系, 其架构如图 14 所示. 2020 年, Souto 等[65]提出一种应用于 RAHT 的预测编码方法, 通过编码残差而非直接编码体素属性提高动态点云的压缩效率; 实验结果表明, 与纯 RAHT 或仅帧内预测相比, 结合帧间预测和帧内预测能显著地提升压缩性能, 使其适用于实时应用. 2023 年, Souto 等[66]进一步提出自适应零运动矢量和自适应运动补偿预测 RAHT 方案, 结合帧间信息和运动补偿预测扩展 RAHT 的应用范围, 并显著地提高了编码效益.

通常, 基于非深度学习的动态点云压缩方法在计算效率和实时性能上表现优异, 更适用于自动驾驶等需要处理实时数据的场景; 非深度学习方法不依赖大量标注数据进行训练, 在获取大量

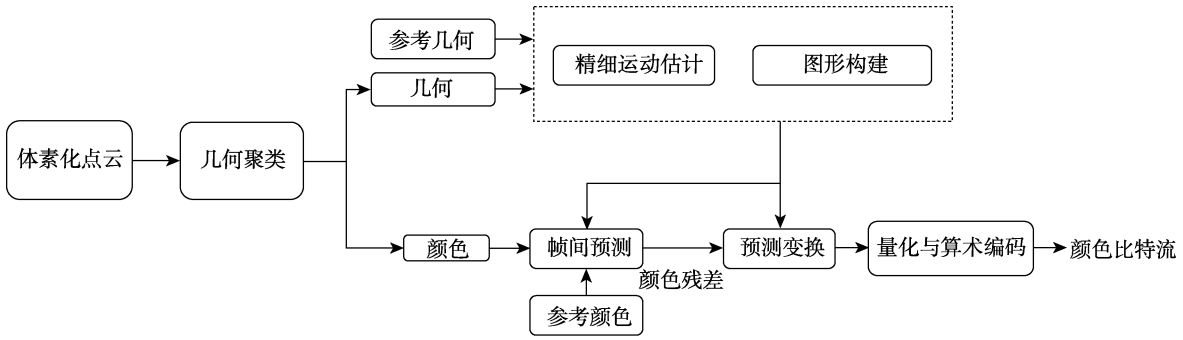


图 14　文献[65]方案的架构

标注数据具有挑战性的情况下尤为重要. 然而, 非深度学习的方法也存在明显不足, 如在优化点云的空间关联性方面具有一定的局限性, 无法自适应地学习和利用数据中的冗余信息; 此外, 这些方法通常仅使用树结构或块结构的编码方式进行压缩, 未能充分地利用点云之间的空间相关性, 大量的冗余信息被编码, 限制了其压缩性能.

2.2.2　基于深度学习的属性压缩

深度学习已广泛地应用于动态点云压缩中. 如图 15 所示, He 等[67]提出一种直接用于动态点云属性压缩的深度学习方法, 保留了大多数传统点云压缩方法忽视的局部密度信息, 采用自编码器策略; 其中, 编码器对点进行下采样并学习点特征, 解码器利用这些特征对点进行上采样, 实现了当时最先进的码率失真权衡. 该方法通过自动学习点云数据中的复杂特征和空间结构信息, 能够显著地提高压缩效率和重建质量.

Pan 等[68]提出一个以块生成模块为核心的点云压缩模型框架 patchDPCC, 通过生成逐点特征进行熵编码, 保证重建速度; 此外, 引入特征转移模块, 利用帧间相似性来提高特征质量. patchDPCC 在基于深度学习的压缩模型中表现出显著的解码速度优势, 展示了其在动态点云压缩中的潜力和实用性. PCAC[69]则采用多尺度的结构, 通过平均池化生成多个尺度的属性张量, 并对每个尺度内的属性进行概率估计; 还考虑了同一尺度内的跨组相关性和不同颜色间的相关性; 最终可以有效地压缩点云属性, 并且具有较低的计算复杂度和可接受的编解码运行时间. DAE-MP[70]通过预测点云帧之间的运动变化, 考虑几何信息和运动一致性将相邻帧的信息进行融合, 进一步增强点云的属性质量; 此外, 通过在不同频率域提取特征区分高频和低频特征, 并在融合过程中调整其权重; 在优化过程中, 采用多个损失函数最小化原始帧与预测帧之间的差异.

2.3　动态点云联合压缩

动态点云压缩已在几何与属性联合压缩方面取得了显著的进展. 如图 16 所示, 将几何和属性

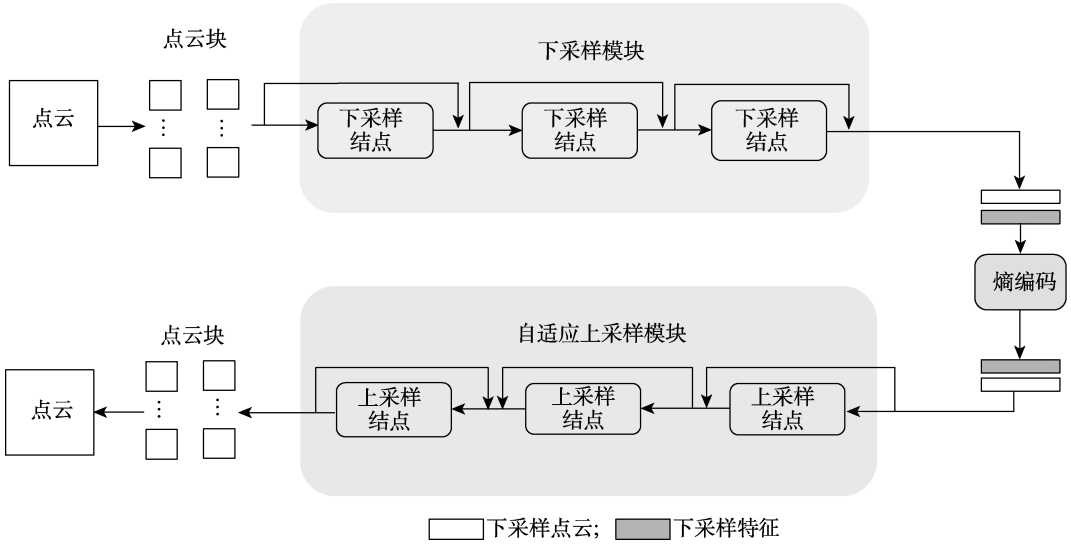


图 15　文献[67]方法的架构

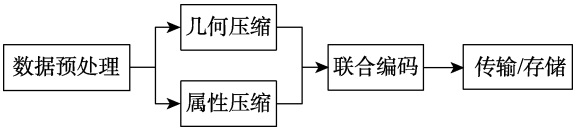


图 16 联合点云压缩流程示意图

信息联合压缩, 可以更有效地减少数据冗余, 提高压缩效率和重建质量. 近年来, 研究人员提出了不同的方法, 利用几何和属性信息之间的相关性优化压缩算法.

G-PCC 方法[39]利用八叉树结构对点云数据进行递归分割, 其中, 八叉树将三维空间划分为更小的立方体节点, 利用父节点和邻居节点的上下文信息提高压缩效率; 父节点的存在与否及其空间位置、状态, 以及邻居节点的位置和占用情况, 都为当前节点的压缩提供了重要的参考, 减少编码所需的比特数. 此外, G-PCC 还利用体素网格与八叉树级别之间的对偶性, 构建基于体素的上下文信息, 将当前级别和父级别的信息结合提升了压缩性能. G-PCC 中, 几何信息的编码采用上下文自适应二进制算术编码等熵编码技术, 通过预测和编码相结合实现几何信息的高效压缩, 不仅显著地减少点云数据的存储和传输需求, 而且保持较高的解码质量, 其架构如图 17 所示.

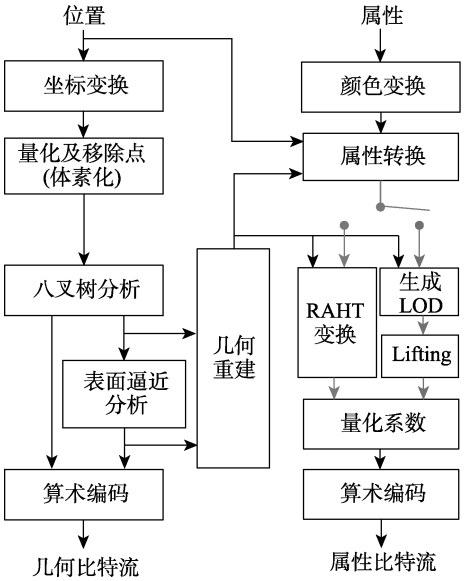


图 17 G-PCC 方法[39]架构

2016 年, Anis 等[71]提出一种通过创建一致的细分三角形网格实现几何与属性联合压缩的方法, 首先对给定的基本网格进行多次细分, 生成不同层次的子网格; 然后在这些子网格中计算顶点位置和颜色的矩阵; 再根据每个点到子网格的距离和表面法向量, 确定最佳匹配的子网格中的顶点; 最后比较点云和其对应的子网格之间的几何和颜色偏差, 找到最优的子网格模型, 实现几何与属性的联合压缩. 2017 年, de Queiroz 等[72]针对时间冗余提出一种基于低比特率编码的动态点云运动补偿方法, 能够准确地表示和传输几何形状的变化, 同时对颜色属性进行残差编码, 进一步降低数据量.

2020 年, MuSCLE[73]利用深度熵模型捕捉多个扫描之间的时间相关性, 并利用先前扫描的几何和强度信息预测当前扫描的符号概率; 然后通过这些学习到的概率编码整个数据流实现压缩. 该方法可以同时对 LiDAR 数据中的几何形状和强度值进行压缩.

2023 年, G-PCC++[74]通过分离几何与属性 2 个部分提高压缩质量. (1) 在几何压缩阶段, 采用八叉树编码对点云的几何信息进行量化; 为了改善几何细节的恢复, 引入基于 *k*NN 的线性插值方法对局部几何进行稠密化. 该方法通过迭代方式逐步增强点云的局部邻域信息, 有效地提升了几何重建的分辨率. (2) 在属性恢复阶段, 经过几何压缩后, 使用基于 *k*NN 的高斯距离加权映射恢复点云的属性数据. 该联合压缩方案不仅有效地降低数据的存储和传输开销, 而且保证了在压缩与恢复过程中几何与属性的高质量重建. 同年, Zhang 等[75]提出基于端到端学习的压缩框架 YOGA, 采用 G-PCC (八叉树)对几何信息进行编码, 并结合 G-PCC (RAHT)与自适应量化对属性信息进行压缩; 在压缩后的属性值的优化方面, 引入自适应滤波器, 通过最小均方误差(mean squared error, MSE)优化对属性值进行细化, 使压缩后的属性值更接近原始值, 显著地减少有损压缩带来的伪影, 提高压缩质量. 2025 年, Zhang 等[76]提出同样基于端到端学习的 DeepPCC, 以提升大规模物体点云的压缩效率. (1) 首先对点云的几何信息进行编码, 利用多尺度下采样和邻域信息聚合提取几何特征, 然后通过压缩编码器对几何信息进行有效的压缩. (2) 解码过程中, 在逐层恢复几何信息的同时, 利用低尺度信息进行精细化修复; 恢复的几何坐标被用作后续属性压缩的隐式几何先验, 在此基础上, 采用邻域信息聚合与自回归模型对属性进行编码和压缩. (3) 在解码阶段, 通过 *k*NN 算法找到恢复的几何坐标与原始点云属性之间的匹配关系, 并利用最近邻点的属性均值确定目标点的属性值.

2025 年, Wang 等[77-78]提出 Unicorn, 支持任意类型、静态或动态点云数据的几何和属性信息的高

效压缩. 该方法基于多尺度稀疏张量表示，并使用跨尺度条件编码利用空间和时间相关性；对于动态点云，可以在多个尺度上进行有损压缩；解决了现有学习型压缩器无法有效地处理多样化点云数据的问题. 通常，传统学习型压缩器只能处理特定类型的点云数据，而 Unicorn 则能够适应各种类型的数据，包括密集物体点云、稀疏场景点云，以及静态和动态点云. 此外，Unicorn 还能够同时压缩点云的几何和属性信息，提高了压缩效率. 因此，Unicorn 是一种具有广泛应用前景的学习型压缩器.

在动态点云压缩研究中，几何与属性联合压缩是一个关键的方向，具有重要的现实意义和广阔的应用前景. 随着传感器技术、编解码技术和机器学习技术的快速发展，对高效点云压缩方法的需求也日益增长. 几何与属性联合压缩通过充分地利用几何信息和属性信息之间的相关性，不仅在理论研究上提供新的思路和方法，也在实际应用中展示了巨大的潜力.

2.4　小　结

动态点云压缩方法的研究进展如图 18 所示，涵盖了 2012—2024 年的重要研究成果和关键文献，从早期的基础研究到近年来的先进方法，系统地展示了各个研究方向的演变过程. 其中，通过颜色编码区分不同类型的压缩方法，“联合”压缩方法通常结合了多种技术.

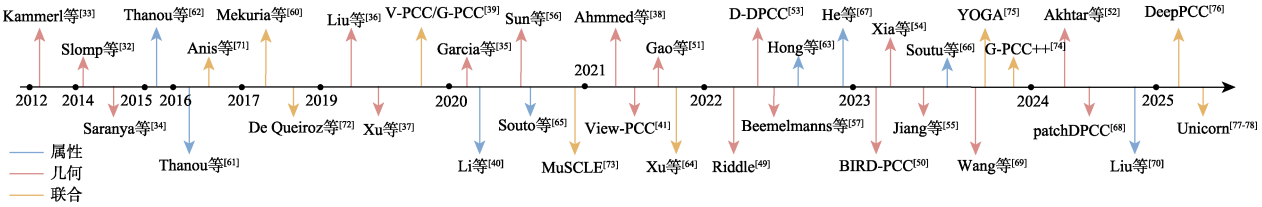


图 18　动态点云压缩方法研究时间线

3　动态点云压缩的分析与对比

通过对现有的动态点云压缩方法进行全面、系统的对比分析，揭示各种方法的优缺点及其适用场景. 首先对当前可靠的公开点云数据集进行分类与整理，探讨动态点云压缩效果的评价指标；在此基础上，对现有的动态点云压缩方法进行综合的对比.

3.1　公共数据集

随着点云研究的不断深入与发展，越来越多有价值的数据集被公开并广泛使用，这些公共数据集不仅丰富了研究资源，也为动态点云压缩方法的开发和评估奠定了坚实的基础. 由于高质量的数据集能够确保实验结果的可靠性和可重复性，因此合理选择和使用这些公共数据集对于推动动态点云压缩技术的发展至关重要.

8i VSLF 是由 8i 公司提供的高质量三维点云数据集[79]，支持和推动计算机视觉和计算机图形学的研究. 其中，包含一系列高分辨率的 300 帧视频和 6 个高分辨率的单帧点云，每个场景都经过精细捕捉和处理；包含 Boxer, LongDress, Loot, Red and Black, Soldier 和 Thaidancer 这 6 个序列；每个点云帧包含数百万个点的高分辨率点云数据、多视角捕捉系统提供的丰富视角信息和高质量的纹理映射，确保数据的丰富性和准确性，为视觉分析提供了可靠的基础；采用标准化格式存储，包括.ply 文件(点云数据)，.png 文件(纹理图像)和.json 文件(元数据). 该数据集广泛应用于三维重建、虚拟现实和增强现实、计算机视觉及深度学习等.

Owlii 数据集[80]由 Owlii 公司提供，是一个广泛应用于计算机视觉和深度学习中的高质量、高精度、多样化三维人体姿态数据集，包含 Basketball player, Dancer, Exercise 和 Model 这 4 个序列；包括多个场景和动作，涵盖了日常活动、体育运动和复杂的人体互动；数据通过多摄像头系统捕捉，采用先进的图像处理和三维重建技术确保数据的准确性和多样性；每个序列中，帧速为 30 帧/s，持续 20 s. 该数据集在动作识别、虚拟现实、影视制作等方面有重要应用，并为学术研究提供标准的基准数据，推动了三维姿态估计和动作识别技术的发展.

MVUB 数据集[81]是由 Microsoft 公司提供的高精度点云数据集，包含 Andrew, David, Phil, Ricardo 和 Sara 这 5 个序列，这些被摄人物的上半身由 4 个正面 RGBD 摄像头以 30 帧/s 的速度拍摄；包含从多个视角采集的高精度点云，确保了数据的全面性和丰富性. 这些视角涵盖了不同的角度和距离，提供了详细的三维信息. 该数据集具有高精度、多视角和多样性的特点，为三维重建和计算机

视觉方向提供了宝贵的资源, 是研究和应用三维点云技术的优秀数据集.

SemanticKITTI 数据集[82]由波恩大学的摄影测量与机器人学组、计算机视觉组和自主智能系统组联合提供, 是一个大规模、真实场景的三维点云数据集, 主要用于自动驾驶中的语义分割和场景理解任务. 该数据集包含对点云数据的语义标签, 每个点云数据中的点被标注为 28 个不同的语义类别, 如道路、车辆、行人和建筑物等, 成为训练和评估机器学习模型, 尤其在自动驾驶方向中, 是三维场景理解、物体检测与跟踪等任务的关键资源; 广泛应用于目标检测[83]、道路测绘与建模[84]、语义分割[85]和三维目标检测等研究中.

Ford Campus Vision 1 数据集(简称 Ford)数据集[86]由福特公司开源, 提供了在真实城市环境中采集的带有时间戳的视觉和导航数据, 对于研究动态点云压缩具有重要意义. 该数据集由一辆配备多种传感器的车辆采集, 包括 Velodyne 激光扫描仪、全方位摄像头和导航传感器, 这些传感器提供了多视角、多时间和不同天气条件下的多种交通场景信息, 涵盖城市街道、高速公路等多样化的驾驶环境. 该数据集不仅提供详尽的传感器数据, 还对点云中的各个物体进行了精确的标注, 便于自动驾驶算法的训练与评估, 尤其是在实时感知与决策任务中起到关键应用.

ShapeNet[87]是一个大型的三维 CAD 模型库, 由斯坦福大学、普林斯顿大学和美国芝加哥的丰田技术研究所的研究团队共同开发, 涵盖从日常物品到复杂物体的多样化形状, 每个物体模型不仅提供精确的几何表示, 还附带了语义标签和结构化元数据. 这种多层次、多角度的标注方式, 使该数据集成为研究物体形状理解和生成的理想平台, 其开放性和广泛适用性也大大促进了三维数据处理技术的发展.

本文对当前常见的公共数据集进行统计与整理, 如表 1 所示.

表 1 常见的公共数据集信息

名称	帧速/(帧·s^{-1})	点数	网格数	几何精度/bit
Loot_vox10_1200[79]	1	≈0.805M		10
Soldier_vox10_0690[79]	1	≈1.089M		10
Longdress_vox10_1300[79]	1	≈0.857M		32
Red and Black_vox10_1050[79]	1	≈0.787M		10
Owlii[80]	30		≈40k	10/11
MVUB[81]	30	≈0.300M		9/10
SemanticKITTI[82]	10	≈0.100M		16
Ford Campus Vision 1[86]	10	≈1.000M		32
ShapeNet[87]	1		≈2k	

3.2 点云压缩指标

3.2.1 每点平均位数

在点云压缩研究中, 压缩比使用每点平均位数(bits per point, BPP)作为评估指标, 用于衡量存储点云数据时每个点所需的平均编码位数. 点云压缩的目标是在保持整体形状和结构的同时减少每个点的数据量, 即 BPP 值越低, 表示压缩效果越好. 在相似重建质量条件下, 通过比较不同方法的 BPP 值, 可以有效地评估各类压缩方法的编码效率和实用性.

3.2.2 MSE

MSE 是一种衡量点云数据的重建质量的常见指标, 采用点对点的度量方式, 计算公式为

$$\mathrm{MSE}=\sum_{i=1}^{n}(Y-X)^2/n.$$

其中, n 表示点云中点的数量; Y 表示重建点云中的点; X 表示对应的参考点云中的最近点. MSE 的值越低, 表示重建结果与原始数据之间的差异较小. 该指标易于计算, 能够直接反映重建点云与参考点云之间差异的特点, 广泛应用于点云压缩的质量评价.

3.2.3 峰值信噪比

峰值信噪比(peak signal to noise ratio, PSNR)是评价点云重建质量的常用指标, 尤其适用于衡量点云在压缩与重建过程中引入的失真. 通常, 几何误差依据 2 种距离度量进行评估: 点到点距离 D1-PSNR 和点到平面距离 D2-PSNR, 其值越高, 表示重建质量越高.

D1-PSNR 通过计算点云中每个点与其在参考点云中最近邻点之间的欧几里得距离, 反映重建点云中的几何误差, 计算公式为

$$\mathrm{D1\text{-}PSNR}=10\lg\left(\frac{d_{\max}}{\mathrm{RMSE}_{\mathrm{p2p}}^2}\right).$$

其中, $\mathrm{RMSE}_{\mathrm{p2p}}$ 表示基于点到点距离的均方根误差(root mean square error, RMSE), 公式为

$$\mathrm{RMSE}_{\mathrm{p2p}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\boldsymbol{E}_{i,j}\right|^2}.$$

$\boldsymbol{E}_{i,j}=P_i-Q_j$ 表示点云中第 i 个点与参考点云中最近邻点之间的误差向量, P_i 表示重建点云中的第 i 个点, Q_j 表示参考点云中距离 P_i 最近的点; $d_{\max}$ 表示参考点云与重建点云中所有最近点对之间的最大欧几里得距离, 用于对信噪比进行标准化, 确保不同点云的测量结果具有可比性.

D2-PSNR 考虑点到平面的距离, 反映点云的局部几何形状误差, 重建点云中每个点投影到参考点云法向量上的距离误差的计算, 公式为

$$\text{D2-PSNR}=10\lg\left(\frac{d_{\max}}{\mathrm{RMSE}_{\mathrm{p2pl}}^2}\right).$$

其中, $\mathrm{RMSE}_{\mathrm{p2pl}}=\sqrt{\frac{1}{n}\sum_{i-1}^{n}\left\langle\boldsymbol{E}_{i,j},\boldsymbol{N}_j\right\rangle^2}$ 表示点到平面距离的 RMSE, $\boldsymbol{N}_j$ 为参考点云中第 j 个点的法向量, 表示局部表面的朝向; $\boldsymbol{E}_{i,j}$ 表示点云中第 i 个点的误差向量, 其投影到法向量上 $\boldsymbol{N}_j$ 的投影长度 $\left\langle\boldsymbol{E}_{i,j},\boldsymbol{N}_j\right\rangle$ 表示点到平面的误差.

3.2.4　Y-PSNR

在动态点云压缩中, Y-PSNR 用于衡量压缩后点云数据颜色信息的保留情况, 常用于评估属性压缩质量, 其值越高, 表示重建质量越高. Y-PSNR 的计算公式为

$$\text{Y-PSNR}=10\lg\left(\frac{C_{\max}^2}{\mathrm{MSE}}\right).$$

其中, $\mathrm{MSE}=\frac{1}{N}\sum_{i=1}^{N}\left(Y_i-\hat{Y}_i\right)^2$, N 表示颜色值的总数, Y_i 和 $\hat{Y}_i$ 分别表示原始颜色值和压缩后颜色值; $C_{\max}$ 表示颜色值的最大值, 通常为 255. 通过计算 MSE 衡量压缩过程中颜色信息的损失; Y-PSNR 对比原始点云和解压缩后点云的颜色信息评估压缩质量, 能够准确、有效地评估压缩方法在保留动态点云颜色信息方面的表现, 对动态点云压缩方法的研究和评价具有重要意义.

3.2.5　BD-RATE

BD-RATE(Bjøntegaard delta rate)指标被广泛用于评估点云压缩方法的性能差异, 由 Bjøntegaard[88] 提出, 用于比较视频编码的性能, 通过计算相同视觉质量下不同编码方法的平均码率差异量化其压缩效率. 首先在不同码率下测量重建点云的质量, 并绘制码率-失真(rate-distortion, RD)曲线; 然后将每组数据点进行四阶多项式拟合后, 对拟合曲线在相同质量范围内积分, 计算每种压缩方案的总码率; 最后比较不同方案的积分结果, 计算平均码率差异, 即 BD-RATE. BD-RATE 以百分比表示, 反映在相同质量下 2 种压缩方法的码率差异. 当 BD-RATE 为负值时, 表示在相同质量下所需码率更少, 压缩效率更高; 其为正值时则相反. BD-RATE 的计算公式为

$$\text{BD-RATE}=10\lg\left(\frac{\int_{Q_1}^{Q_2}R_2(Q)-\int_{Q_1}^{Q_2}R_1(Q)}{Q_2-Q_1}\right)\times100\%.$$

其中, R_1 和 R_2 分别表示 2 种压缩方法在质量 Q 下的码率函数; Q_1 和 Q_2 分别表示质量范围的上限和下限. BD-RATE 能够定量地评估点云压缩方法的性能, 为方法选择和优化提供依据.

3.3　实验结果分析与对比

本文从多个角度对不同方法在各个指标上进行统计与分析. 其中, 部分实验结果来源于原文献; 使用 MPEG 提出的动态点云压缩标准统一为 V-PCCv18.0 和 G-PCCv21.0, G-PCC 采用 9 级八叉树编码和 RAHT. 通过对比各种方法在不同测试条件下的性能, 可以更全面地了解它们在实际应用中的优劣, 为后续研究提供参考.

3.3.1　复杂度分析

在动态点云压缩技术的发展过程中, 复杂度分析是评估各类压缩方法性能的关键环节之一. 本节对实验结果进行分析和对比, 重点关注各方法在编解码时间和存储空间方面的表现. 通过对这些关键指标的综合考察, 能够更全面地了解各种压缩方法的优缺点及其在实际应用中的可行性. 实验平台的参数配置如表 2 所示. 9 种代表性方法编解码时间和存储空间的对比结果如表 3 所示, 每种方法均基于其原论文设定的数据及实现, 为动态点云压缩方法的选择提供了重要的参考依据.

表 2　实验环境的参数配置

参数	数值
操作系统	Ubuntu 20.04
CPU	Intel(R) i9-13900K
GPU	GTX3090/24 GB
CUDA 版本	11.7
Python 版本	3.8
Pytorch 版本	1.10.1

表 3 动态点云压缩方法复杂度对比

方法	压缩对象	数据集	时间/s		模型大小/MB
			编码	解码	
V-PCC[39]	几何	8i VSLF	19.67	12.63	
G-PCC[39]	联合	8i VSLF	10.40	10.20	
D-DPCC[53]	几何	8i VSLF	0.23	0.13	33.00
Jiang 等[55]	几何	8i VSLF	0.77	0.24	435.00
He 等[67]	属性	SemanticKITTI	0.08	0.31	0.44
patchDPCC[68]	几何	8i VSLF	5.06	1.34	
G-PCC++[74]	联合	8i VSLF	1.79	3.00	8.90
YOGA[75]	几何	8i VSLF	0.51	0.56	4.40
	属性		7.64	2.67	165.10
Unicorn[77-78]	几何	Owlii	0.91	1.16	51.00
	属性		2.90	1.60	65.00

编解码时间是评估压缩方法效率的核心参数. 随着点云数据规模的不断增大, 快速、高效的编解码过程变得至关重要. 本文系统地对多种主流动态点云压缩方法的编码和解码时间进行实验, 评估其在实时处理中的表现. 实验结果表明, 在处理大规模点云数据时, 不同方法的时间开销存在显著差异, 这是因为算法的复杂度和实现优化程度不同.

存储空间占用是衡量压缩效果的重要指标. 尽管高压缩率意味着较小的存储空间占用, 但通常也伴随着较长的编解码时间. 因此在实际应用中, 压缩方法的选择需要在压缩率和复杂度之间进行权衡. 在不同数据集上, 通过对各类压缩方法的压缩率和存储空间进行比较分析, 可以深入地了解其性能表现, 并探讨其在不同应用场景中的适用性.

3.3.2 性能对比

对本文所述动态点云压缩方法的压缩结果进行比较和总结, 选取 G-PCC(TMC13v21)作为基准, 以相同数据集下的 D1-PSNR, D2-PSNR 和 Y-PSNR 的 BD-RATE 值作为评价指标, 以更清晰地了解各种压缩方法的性能.

动态点云几何压缩效果对比如表 4 所示. 可以看出, 深度学习方法通常表现出色; 在 D1-PSNR 的 BD-RATE 上, D-DPCC 方法[53]达到−98.17%, 在应对复杂动态场景时展现出显著的优势, 尤其适用于自动驾驶场景中对车辆与行人进行实时检测与跟踪等延迟敏感任务, 可在保证压缩效率的同时保持高精度重建; 在 D1-PSNR 的 BD-RATE 上, Jiang 等[55]的方法达到−96.84%, 在工业环境中的远程三维建模与操作任务中显示出良好的鲁棒性, 能够有效地支持对设备状态的精确还原与监控; 在 D1-PSNR 的 BD-RATE 上, AuxGR[43]达到−97.37%, 在面向边缘端部署的轻量化感知系统中表现良好, 适用于资源受限环境下的数据采集与压缩传输. 相比之下, 传统方法 Riddle[49]采用八叉树策略, 在

表 4 动态点云几何压缩效果对比

处理方法	文献	数据结构	数据集		BD-RATE/%		平台
			训练	测试	D1-PSNR	D2-PSNR	
规则	V-PCC[39]	投影		Owlii			
	Riddle[49]	八叉树		SemanticKITTI	−47.69		
学习	AuxGR[43]	体素	ShapeNet	Owlii	−97.37	−95.89	RTX4090
	PCGFormer[46]	体素	ShapeNet	8i VSLF	−85.81	−85.20	RTX3090
	Gao 等[51]	图	ShapeNet	ShapeNet	18.81		GTX 1080 Ti
	Akhtar 等[52]	八叉树	8i VSLF	8i VSLF	−86.00		
	D-DPCC[53]	体素	Owlii	8i VSLF	−98.17		
	Xia 等[54]	体素	8i VSLF	Owlii	−64.32	−82.23	
	Jiang 等[55]	体素	8i VSLF	8i VSLF	−96.84	−74.60	
	patchDPCC[68]	体素	Owlii	8i VSLF	−66.38	−79.70	RTX3090

D1-PSNR 的 BD-RATE 上也达到−47.69%, 该方法结构简单且计算量低, 仍适用于在移动设备或车载终端等资源受限的边缘计算场景中作为快速压缩的前处理模块使用, 但面对数据密度高、动态变化剧烈的复杂场景时, 其压缩性能和适应能力明显弱于深度学习方法.

动态点云属性压缩效果对比如表 5 所示. 可以看出, 传统方法依然展现出显著的性能优势; Hong 等[63]的块处理方法达到−85.04%, 适用于大规模静态监控场景中对纹理与颜色属性的压缩, 如室内安防系统或固定视角的多媒体传输等. 由于许多深度学习方法的评价指标并不统一, 因此在表中未以 BD-RATE 的形式列出. 尽管如此, 深度学习方法的潜力依然值得关注, 其在处理复杂属性分布方面具备自动学习能力, 尤其在虚拟现实或增强现实等需要高度细节还原的沉浸式应用中表现出色, 能够自适应地建模颜色和纹理变化.

动态点云联合压缩效果对比如表 6 所示. 可以看出, 深度学习方法同样展现出较强的性能; YOGA[75]在几何和属性的联合压缩中均表现突出, D1-PSNR 和 Y-PSNR 的 BD-RATE 达到−85.11%和−72.24%, 能够在保留几何结构信息的同时有效地维持颜色与纹理的还原质量, 特别适用于虚拟现实、远程协作等多模态交互场景, 满足对空间感与细节还原的双重需求; G-PCC++[74]在几何压缩上 BD-RATE 也有−73.25%的表现, 显示出其联合处理的潜力.

表 5　动态点云属性压缩效果对比

处理方法	文献	数据结构	数据集		Y-PSNR 的 BD-RATE/%
			训练	测试	
规则	Thanou 等[61]	图			
	Thanou 等[62]	图			
	Hong 等[63]	块		8i VSLF	−85.04
	Xu 等[64]	体素		8i VSLF	−37.23
	Soutu 等[65]	图		8i VSLF	−77.27
	Soutu 等[66]	图		8i VSLF	−83.54
学习	He 等[67]	块	SemanticKITTI/ShapeNet	SemanticKITTI/ShapeNet	
	Liu 等[70]	块	8i VSLF	8i VSLF	

表 6　动态点云联合压缩效果对比

处理方法	文献	数据结构	数据集		BD-RATE			平台
			训练	测试	D1-PSNR/%	D2-PSNR/%	Y-PSNR/%	
规则	G-PCC[39]	八叉树		Owill				GTX 3090
	Anis 等[71]	网格		8i VSLF	−75.63%		−8.27%	
	MuSCLE[73]	八叉树		SemanticKITTI		−66.88%		
学习	G-PCC++[74]	体素	ShapeNet	8i VSLF	−73.25%			GTX 3090Ti
	YOGA[75]	体素	ShapeNet	8i VSLF	−85.11%		−72.24%	
	DeepPCC[76]	体素	ShapeNet	8i VSLF	−90.00%	−85.89%	−51.15%	RTX 3090Ti
	Unicorn[77-78]	体素	8iVFB	Owlii	−96.22%	−74.15%	−31.86%	

整体来看, 深度学习方法正逐步在动态点云的几何与属性综合压缩任务中超越传统方法, 标志着该方向正在迈向更高效、更智能的压缩技术发展路径.

3.4　小　结

本节详细介绍并分析了动态点云压缩中常用的数据集和评价指标, 对现有的方法进行系统性的评估和比较, 不仅指出当前动态点云压缩技术的优势与不足, 还为未来的研究指明了方向. 结合不同压缩方法的优点发展混合型压缩算法, 以及进一步优化现有深度学习模型的效率和精度, 是未来值得深入探索的重要方向. 此外, 随着应用场景的不断拓展, 研究更具通用性和适应性的压缩算法也显得尤为关键. 这些算法不仅能够提升动态点云压缩效果, 还能更好地适应多样化的应用需求.

4　结　语

动态点云压缩技术在现代多媒体和通信领域中发挥着至关重要的作用. 未来, 随着硬件技术的

进步和计算能力的提升, 动态点云压缩技术将继续向高效、高质量和实时方向演进; 随着网络传输技术的进步, 实时传输和处理海量点云数据将成为可能, 进一步推动动态点云压缩系统的实用化落地.

当前, 尽管已有多种先进方法取得了良好的效果, 但动态点云压缩仍面临诸多挑战: (1) 压缩算法复杂度高, 计算资源需求大, 限制了其在资源受限设备上的部署; (2) 不同应用场景对压缩精度、延迟与鲁棒性的需求差异较大[89], 当前方法在通用性和适应性方面仍有不足; (3) 随着点云应用的多样化与智能化发展, 如何在压缩过程中保障数据的安全性与鲁棒性, 也成为亟待解决的重要问题[90].

为了应对上述挑战, 未来可以在以下几个方向展开研究: (1) 设计轻量级神经网络结构, 结合模型剪枝、知识蒸馏等技术降低推理开销; (2) 发展端到端的自监督学习框架, 减少对人工标签或显式运动建模的依赖, 提升泛化能力; (3) 引入联合时空建模机制, 如时序图网络或 Transformer 架构, 更充分地挖掘帧间冗余信息; (4) 融合图像、深度图等跨模态信息, 引导点云压缩策略向更高层语义对齐; (5) 采用动态比特分配策略, 根据点云内容的重要性灵活地调整压缩资源分配.

语义感知压缩、流式增量压缩、安全保护机制与多任务压缩策略, 也将成为未来的研究热点. 通过这些路径, 动态点云压缩技术可望实现多项关键目标, 包括支持在边缘设备上的实时压缩与传输、在极低比特率下保持高重建质量、适应复杂多样的应用场景, 并在保障隐私与安全的前提下广泛应用于自动驾驶、增强现实/虚拟现实、智能监控等.

动态点云压缩技术是一个充满活力和潜力的研究方向, 随着研究的不断深入和技术的不断进步, 未来该技术将会更加高效、智能和实用, 为各种应用场景提供强有力的支持. 动态点云压缩技术不仅在当前的多媒体和通信领域中具有重要意义, 其未来发展也将极大地依赖持续的学术研究与技术创新, 因此, 推动这一研究的深入探索将实现更广泛应用的关键. 通过不断地探索和解决现有的技术瓶颈, 动态点云压缩技术必将在今后的应用中发挥更加举足轻重的作用.

参考文献(References):

[1] Alliez P. Recent advances in compression of 3D meshes[C] //Proceedings of the 13th European Signal Processing Conference. Los Alamitos: IEEE Computer Society Press, 2005: 1-4

[2] Aleksandrov M, Zlatanova S, Heslop D J. Voxelisation algorithms and data structures: a review[J]. Sensors, 2021, 21(24): Article No.8241

[3] Garcia D C, de Queiroz R L. Intra-frame context-based octree coding for point-cloud geometry[C] //Proceedings of the 25th IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2018: 1807-1811

[4] Jiang Zhaoyi, Zou Wenqin, Song Chao, *et al*. Variable rate compression of point cloud based on channel gain[J]. Journal of Computer-Aided Design & Computer Graphics, 2024, 36(11): 1816-1824(in Chinese)
(江照意, 邹文钦, 宋超, 等. 基于通道增益的可变比特率点云压缩[J]. 计算机辅助设计与图形学学报, 2024, 36(11): 1816-1824)

[5] Akhtar A, Gao W, Zhang X, *et al*. Point cloud geometry prediction across spatial scale using deep learning[C] //Proceedings of the IEEE International Conference on Visual Communications and Image Processing. Los Alamitos: IEEE Computer Society Press, 2020: 70-73

[6] Malvar H S. Adaptive run-length/Golomb-Rice encoding of quantized generalized Gaussian sources with unknown statistics[C] //Proceedings of the Data Compression Conference. Los Alamitos: IEEE Computer Society Press, 2006: 23-32

[7] Jiang Z Y, Ding L Y, Tam G K L, *et al*. C2SPoint: a classification-to-saliency network for point cloud saliency detection[J]. Computers & Graphics, 2023, 115: 274-284

[8] Yan W, Shao Y T, Liu S, *et al*. Deep AutoEncoder-based lossy geometry compression for point clouds[OL]. [2024-11-16]. https://arxiv.org/abs/1905.03691

[9] Wiesmann L, Milioto A, Chen X Y L, *et al*. Deep compression for dense point cloud maps[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 2060-2067

[10] Yamaguchi K, Kunii T L, Fujimura K, *et al*. Octree-related data structures and algorithms[J]. IEEE Computer Graphics and Applications, 1984, 4(1): 53-59

[11] Meagher D. Geometric modeling using octree encoding[J]. Computer Graphics and Image Processing, 1982, 19(2): 129-147

[12] Wu C Y, Singhal N, Krähenbühl P. Video compression through image interpolation[C] //Proceedings of the 15th European Conference on Computer Vision. Heidelberg: Springer, 2018: 425-440

[13] Yang R H, Yang Y B, Marino J, *et al*. Hierarchical autoregressive modeling for neural video compression[OL]. [2024-11-16]. https://arxiv.org/abs/2010.10258

[14] Hu Z H, Lu G, Xu D. FVC: a new framework towards deep video compression in feature space[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 1502-1511

[15] Ballé J, Laparra V, Simoncelli E P. End-to-end optimized image compression[OL]. [2024-11-16]. https://arxiv.org/abs/1611.01704v3

[16] Ballé J, Minnen D, Singh S, *et al*. Variational image compression with a scale hyperprior[OL]. [2024-11-16]. https://arxiv.org/abs/1802.01436v2

[17] Bross B, Chen J L, Ohm J R, *et al*. Developments in international video coding standardization after AVC, with an overview of versatile video coding (VVC)[J]. Proceedings of the IEEE, 2021, 109(9): 1463-1493

[18] Schwarz H, Marpe D, Wiegand T. Overview of the scalable video coding extension of the H.264/AVC standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(9): 1103-1120

[19] Sullivan G J, Ohm J R, Han W J, *et al*. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668

[20] Djelouah A, Campos J, Schaub-Meyer S, *et al*. Neural inter-frame compression for video coding[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2019: 6420-6428

[21] Lu G, Ouyang W L, Xu D, *et al*. DVC: an end-to-end deep video compression framework[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 10998-11007

[22] Nan F Z, Li F, Wang Z Y, *et al*. Multi-modal dynamic point cloud geometric compression based on bidirectional recurrent scene flow[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2025: 1-5

[23] Fu C Y, Li G, Song R, *et al*. OctAttention: octree-based large-scale contexts model for point cloud compression[C] //Proceedings of the 36th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2022: 625-633

[24] Quach M, Valenzise G, Dufaux F. Learning convolutional transforms for lossy point cloud geometry compression[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2019: 4320-4324

[25] Choy C, Gwak J Y, Savarese S. 4D spatio-temporal ConvNets: minkowski convolutional neural networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 3070-3079

[26] Jiang Z Y, Han D, Song C, *et al*. MP-DPCC: a motion proxy-based dynamic point cloud compression framework[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2025: 1-5

[27] de Oliveira Rente P, Brites C, Ascenso J, *et al*. Graph-based static 3D point clouds geometry coding[J]. IEEE Transactions on Multimedia, 2019, 21(2): 284-299

[28] Nguyen D T, Quach M, Valenzise G, *et al*. Learning-based lossless compression of 3D point cloud geometry[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2021: 4220-4224

[29] Schnabel R, Klein R. Octree-based point-cloud compression[C] //Proceedings of the 3rd Eurographics/IEEE VGTC conference on Point-Based Graphics. Aire-la-Ville: Eurographics Association Press, 2006: 111-121

[30] Huang L L, Wang S L, Wong K, *et al*. Octsqueeze: octree-structured entropy model for liDAR compression[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 1310-1320

[31] Zhang C, Florêncio D, Loop C. Point cloud attribute compression with graph transform[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2014: 2066-2070

[32] Slomp M, Kawasaki H, Furukawa R, *et al*. Temporal octrees for compressing dynamic point cloud streams[C] //Proceedings of the 2nd International Conference on 3D Vision. Los Alamitos: IEEE Computer Society Press, 2014: 49-56

[33] Kammerl J, Blodow N, Rusu R B, *et al*. Real-time compression of point cloud streams[C] //Proceedings of the IEEE International Conference on Robotics and Automation. Los Alamitos: IEEE Computer Society Press, 2012: 778-785

[34] Saranya M R, Senthil S T. Real-time compression strategy on various point cloud streams[J]. International Journal of Computer Science and Mobile Computing, 2014, 3(3): 351-358

[35] 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

[36] Liu J Q, Yao J, Tu J M, *et al*. Data-adaptive packing method for compression of dynamic point cloud sequences[C] //Proceedings of the IEEE International Conference on Multimedia and Expo. Los Alamitos: IEEE Computer Society Press, 2019: 904-909

[37] Xu Y Z, Zhu W J, Xu Y L, *et al*. Dynamic point cloud geometry compression via patch-wise polynomial fitting[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2019: 2287-2291

[38] Ahmmed A, Paul M, Murshed M, *et al*. Dynamic point cloud compression using a cuboid oriented discrete cosine based motion model[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2021: 1935-1939

[39] 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, 2019, 9(1): 133-148

[40] Li L, Li Z, Zakharchenko V, *et al*. Advanced 3D motion prediction for video-based dynamic point cloud compression[J]. IEEE Transactions on Image Processing, 2020, 29: 289-302

[41] Zhu W J, Ma Z, Xu Y L, *et al*. View-dependent dynamic point cloud compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(2): 765-781

[42] Wang J Q, Ding D D, Li Z, *et al*. Multiscale point cloud geometry compression[C] //Proceedings of the Data Compression Conference. Los Alamitos: IEEE Computer Society Press, 2021: 73-82

[43] Liu G X, Zhu J H, Ding D D, *et al*. Encoding auxiliary information to restore compressed point cloud geometry[C] //Proceedings of the 33rd International Joint Conference on Artificial Intelligence. Freiburg: IJCAI, 2024: 2189-2197

[44] Wang J Q, Ding D D, Li Z, *et al*. Sparse tensor-based multiscale representation for point cloud geometry compression[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(7): 9055-9071

[45] Quach M, Valenzise G, Dufaux F. Improved deep point cloud geometry compression[C] //Proceedings of the 22nd IEEE International Workshop on Multimedia Signal Processing. Los Alamitos: IEEE Computer Society Press, 2020: 1-6

[46] Liu G X, Wang J Q, Ding D D, *et al*. PCGFormer: lossy point cloud geometry compression via local self-attention[C] //Proceedings of the IEEE International Conference on Visual Communications and Image Processing. Los Alamitos: IEEE Computer Society Press, 2022: 1-5

[47] 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

[48] Que Z Z, Lu G, Xu D. VoxelContext-Net: an octree based framework for point cloud compression[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 6038-6047

[49] Zhou X Y, Qi C R, Zhou Y, *et al*. Riddle: lidar data compression with range image deep delta encoding[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2022: 17191-17200

[50] Liu C S, Yeh J F, Hsu H, *et al*. BIRD-PCC: bi-directional range image-based deep lidar point cloud compression[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2023: 1-5

[51] Gao L Y, Fan T Y, Wan J Q, *et al*. Point cloud geometry compression via neural graph sampling[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2021: 3373-3377

[52] Akhtar A, Li Z, van der Auwera G. Inter-frame compression for dynamic point cloud geometry coding[J]. IEEE Transactions on Image Processing, 2024, 33: 584-594

[53] Fan T Y, Gao L Y, Xu Y L, *et al*. D-DPCC: deep dynamic point cloud compression via 3D motion prediction[C] //Proceedings of the 31st International Joint Conference on Artificial Intelligence. Freiburg: IJCAI, 2022: 898-904

[54] Xia S T, Fan T Y, Xu Y L, *et al*. Learning dynamic point cloud compression via hierarchical inter-frame block matching[C] //Proceedings of the 31st ACM International Conference on Multimedia. New York: ACM Press, 2023: 7993-8003

[55] Jiang Z Y, Wang G L, Tam G K L, *et al*. An end-to-end dynamic point cloud geometry compression in latent space[J]. Displays, 2023, 80: Article No.102528

[56] Sun X B, Wang S K, Wang M H, *et al*. A novel coding architecture for LiDAR point cloud sequence[J]. IEEE Robotics and Automation Letters, 2020, 5(4): 5637-5644

[57] Beemelmanns T, Tao Y C, Lampe B, *et al*. 3D point cloud compression with recurrent neural network and image compression methods[C] //Proceedings of the IEEE Intelligent Vehicles Symposium. Los Alamitos: IEEE Computer Society Press, 2022: 345-351

[58] de Queiroz R L, Chou P A. Compression of 3D point clouds using a region-adaptive hierarchical transform[J]. IEEE Transactions on Image Processing, 2016, 25(8): 3947-3956

[59] Cohen R A, Tian D, Vetro A. Attribute compression for sparse point clouds using graph transforms[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2016: 1374-1378

[60] Mekuria R, Blom K, Cesar P. Design, implementation, and evaluation of a point cloud codec for tele-immersive video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(4): 828-842

[61] 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

[62] Thanou D, Chou P A, Frossard P. Graph-based motion estimation and compensation for dynamic 3D point cloud compression[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2015: 3235-3239

[63] Hong H R, Pavez E, Ortega A, *et al*. Fractional motion estimation for point cloud compression[C] //Proceedings of the Data Compression Conference. Los Alamitos: IEEE Computer Society Press, 2022: 369-378

[64] Xu Y Q, Hu W, Wang S S, *et al*. Predictive generalized graph Fourier transform for attribute compression of dynamic point clouds[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(5): 1968-1982

[65] Souto A L, de Queiroz R L. On predictive RAHT for dynamic point cloud coding[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2020: 2701-2705

[66] Souto A L, De Queiroz R L, Dorea C. Motion-compensated predictive RAHT for dynamic point clouds[J]. IEEE Transactions on Image Processing, 2023, 32: 2428-2437

[67] He Y, Ren X L, Tang D H, *et al*. Density-preserving deep point cloud compression[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2022: 2323-2332

[68] Pan Z R, Xiao M B, Han X, *et al*. patchDPCC: a patchwise deep compression framework for dynamic point clouds[C] //Proceedings of the 38th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2024: 4406-4414

[69] Wang J Q, Ding D D, Ma Z. Lossless point cloud attribute compression using cross-scale, cross-group, and cross-color prediction[C] //Proceedings of the Data Compression Conference. Los Alamitos: IEEE Computer Society Press, 2023: 228- 237

[70] Liu W, Gao W, Mu X M. Fast inter-frame motion prediction for compressed dynamic point cloud attribute enhancement[C] //Proceedings of the 38th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2024: 3720-3728

[71] Anis A, Chou P A, Ortega A. Compression of dynamic 3D point clouds using subdivisional meshes and graph wavelet transforms[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2016: 6360-6364

[72] de Queiroz R L, Chou P A. Motion-compensated compression

of dynamic voxelized point clouds[J]. IEEE Transactions on Image Processing, 2017, 26(8): 3886-3895

[73] Biswas S, Liu J, Wong K, *et al*. MuSCLE: multi sweep compression of LiDAR using deep entropy models[C] //Proceedings of the 34th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc, 2020: Article No.1859

[74] Zhang J Z, Chen T, Ding D D, *et al*. G-PCC++: enhanced geometry-based point cloud compression[C] //Proceedings of the 31st ACM International Conference on Multimedia. New York: ACM Press, 2023: 1352-1363

[75] Zhang J T, Chen T, Ding D D, *et al*. YOGA: yet another geometry-based point cloud compressor[C] //Proceedings of the 31st ACM International Conference on Multimedia. New York: ACM Press, 2023: 9070-9081

[76] Zhang J Z, Liu G X, Zhang J T, *et al*. DeepPCC: learned lossy point cloud compression[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2025, 9(2): 1897-1909

[77] Wang J Q, Xue R X, Li J X, *et al*. A versatile point cloud compressor using universal multiscale conditional coding–Part Ⅰ: geometry[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(1): 269-287

[78] Wang J Q, Xue R X, Li J X, *et al*. A versatile point cloud compressor using universal multiscale conditional coding–Part Ⅱ: attribute[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(1): 252-268

[79] d'Eon E, Harrison B, Myers T, *et al*. 8i voxelized full bodies-a voxelized point cloud dataset[R]. Geneva: MPEG, 2017

[80] Xu Y, Lu Y, Wen Z Y. Owlii dynamic human mesh sequence dataset[OL]. [2024-11-16]. https://mpeg-pcc.org/index.php/pcc-content-database/owlii-dynamic-human-textured-mesh-sequence-datase

[81] Loop C, Cai Q, Escolano S O, *et al*. Microsoft voxelized upper bodies-a voxelized point cloud dataset[R]. Geneva: MPEG, 2016

[82] Behley J, Garbade M, Milioto A, *et al*. SemanticKITTI: a dataset for semantic scene understanding of LiDAR sequences[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2019: 9296-9306

[83] Liu T, Zhao Y, Wei Y C, *et al*. Concealed object detection for activate millimeter wave image[J]. IEEE Transactions on Industrial Electronics, 2019, 66(12): 9909-9917

[84] Yang B S, Huang R G, Dong Z, *et al*. Two-step adaptive extraction method for ground points and breaklines from lidar point clouds[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 119: 373-389

[85] Liu X Y, Qi C R, Guibas L J. FlowNet3D: learning scene flow in 3D point clouds[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 529-537

[86] Pandey G, McBride J R, Eustice R M. Ford campus vision and lidar data set[J]. The International Journal of Robotics Research, 2011, 30(13): 1543-1552

[87] Chang A X, Funkhouser T, Guibas L, *et al*. ShapeNet: an information-rich 3D model repository[OL]. [2024-11-16]. https://arxiv.org/abs/1512.03012

[88] Bjøntegaard G. Calculation of average PSNR differences between RD-curves[C] // Proceedings of the ITU-T Video Coding Experts Group Thirteenth Meeting. Austin: ITU-T SG16, 2001: 1-15

[89] Wang Changshuo, Wang Han, Ning Xin, *et al*. 3D point cloud classification method based on dynamic coverage of local area[J]. Journal of Software, 2022, 34(4): 1962-1976(in Chinese)
(王昌硕, 王含, 宁欣, 等. 基于局部区域动态覆盖的 3D 点云分类方法[J]. 软件学报, 2022, 34(4): 1962-1976)

[90] Wang C S, Cao R S, Wang R P. Learning discriminative topological structure information representation for 2D shape and social network classification via persistent homology[J]. Knowledge-Based Systems, 2025, 311: Article No.113125