摘要

随着虚拟现实、自动驾驶、数字孪生与沉浸式通信等新兴应用的快速发展，三维点云作为能够直接记录空间几何与属性信息的数据形式，正逐渐成为未来三维视觉处理与交互系统的核心数据基础。然而，点云具有非结构化、稀疏且数据量庞大的特征，其存储与传输负担远超传统二维图像或视频。如何在保证重建质量的前提下实现高效压缩，已成为点云编码领域长期亟需解决的问题。针对这一问题，国内外标准化组织（如 MPEG、JPEG、AVS）陆续提出了多种点云压缩框架并取得初步成果，但在预测机制的精度、预测结构的灵活性、几何与属性协同编码及视觉感知质量优化等方面仍存在明显不足，难以满足新一代沉浸式应用对低码率高质量重建的需求。

针对点云编码中存在的上述问题，从算法改进、结构优化到框架创新三个层次展开了系统研究，构建了由基于空间分布特性的预测优化、联合几何和纹理特性的预测结构改进，到基于高斯表征的几何属性联合编码的多层技术体系。主要研究内容与创新贡献包括以下三个方面：

首先，针对现有点云属性编码中预测精度不足、参考节点分布不均衡的问题，提出了基于空间分布特性的点云空域与变换域预测方法。在 MPEG G-PCC 框架下，设计了两种预测优化方案：其一为基于子节点的变换域预测算法，通过引入同层及相邻层的已编码节点，构建更加精细的预测关系，从而有效降低预测残差能量；其二为基于平滑滤波的预测增强算法，通过统计分析预测残差分布特性，在不增加搜索复杂度的前提下隐式扩展参考范围，提高预测精度并增强局部一致性。实验结果表明，所提方法在多类点云数据集上均取得稳定性能增益，在保证复杂度可控的前提下显著提升了属性编码性能与预测效率。

其次，针对传统点云编码中过度依赖几何驱动的扫描顺序、预测结构单一且缺乏灵活性等固化问题，提出了联合几何与纹理特性的点云预测结构优化及编码方法。本文系统分析了几何—属性一致性分布特征及其对预测性能的影响规律，提出基于仿射空间变换的多候选扫描顺序生成方法，并通过预测残差反馈机制实现自适应选择与优化，从而实现几何与属性的协同预测。此外，为兼顾性能与复杂度，本文提出了基于局部优化的快速剪枝算法，在保持优越性能的同时显著降低了编码时间。大量实验验证表明，该方法在复杂几何与多样纹理场景下均取得明显编码性能提升，表现出较强的通用性和实用性。

最后，针对传统点云编码在低码率条件下重构质量不足、几何与属性联合编码能力有限的问题，提出了基于高斯表征的点云几何与属性联合编码框架（GS-PCC）。该

框架首次将三维高斯泼溅（3D Gaussian Splatting, 3DGS）技术引入点云压缩，实现了从“离散点表示”到“高效高斯参数化表征”的范式转变。在编码端，GS-PCC 通过分层结构将输入点云分解为二维纹理和三维结构两部分，分别采用视频编码器和点云编码器实现联合压缩；在解码端，引入基于超分辨率的高斯训练优化与基于掩膜融合的渲染策略，有效提升了低码率条件下的视觉重建质量。同时，本文采用了基于二维渲染的感知质量评价体系，通过多视角渲染和图像质量指标对重建效果进行评估，更符合人眼主观体验。实验结果显示，所提出的 GS-PCC 框架在低码率场景下显著优于现有主流点云压缩方法 G-PCC 和 Draco，在结构边界与纹理细节的视觉保真度方面表现尤为突出。

综上所述，本研究围绕点云编码中的关键科学问题，从预测机制、预测结构到编码架构进行了系统研究，提出了多种具有创新性和实用价值的点云高效编解码方法。研究工作在预测精度提升、预测结构优化以及几何属性联合编码等方面取得了积极进展，并促进了点云编码技术从几何驱动向联合驱动、从独立处理向协同压缩的转变。总体而言，所形成的方法和技术为点云压缩标准的完善、三维内容的高效传输以及沉浸式视觉媒体的应用提供了可行的技术方案和理论支持，具有一定的学术意义和工程应用价值。

关键词：3D 点云压缩，点云编码标准，MPEG G-PCC，属性预测，几何属性联合编码

Research on Efficient Point Cloud Coding Methods Using Joint Geometry and Attribute Information

Qian Yin (Computer Application Technology)
Directed by Prof. Siwei Ma

ABSTRACT

With the rapid development of emerging applications such as virtual reality, autonomous driving, digital twins, and immersive communications, three-dimensional point clouds, which can directly record spatial geometry and attribute information, have gradually become a fundamental data representation for high-fidelity spatial perception and interaction. However, point clouds are characterized by their unstructured, sparse, and massive nature, making their storage and transmission burdens far greater than those of traditional two-dimensional images or videos. How to achieve efficient compression while maintaining reconstruction quality has long been an urgent problem in the field of point cloud coding. To address this issue, domestic and international standardization organizations such as MPEG, JPEG, and AVS have successively proposed a variety of point cloud compression frameworks and achieved preliminary progress. However, limitations remain in terms of prediction accuracy, prediction structure flexibility, joint geometry–attribute coding, and visual perceptual quality optimization, which makes it difficult to meet the demands of next-generation immersive applications for high-quality reconstruction at low bitrates.

To address these challenges, this dissertation conducts systematic research on point cloud coding from three perspectives: algorithmic improvement, structural optimization, and architectural innovation. A multi-layered technical framework is constructed, encompassing spatial prediction optimization based on spatial distribution, prediction structure improvement based on joint geometric and texture features, and joint geometry–attribute coding based on Gaussian representation. The main research contents and innovative contributions of this work are summarized as follows.

First, to tackle insufficient prediction accuracy and uneven reference node distribution in existing point cloud attribute coding, this work proposes spatial- and transform-domain prediction methods based on spatial distribution characteristics. Within the MPEG G-PCC

framework, two prediction optimization schemes are designed. The first is a transform-domain prediction algorithm based on sibling nodes, which introduces already-coded nodes from the same or adjacent levels to build more precise prediction relationships, thereby effectively reducing prediction residual energy. The second is a prediction enhancement algorithm based on smoothing filters, which performs statistical analysis of residual distributions and implicitly expands the reference range without increasing search complexity, thus improving prediction precision and enhancing local consistency. Experimental results show that the proposed methods achieve consistent performance gains across various point cloud datasets, significantly improving attribute coding performance and prediction efficiency under controllable complexity.

Second, to overcome the rigid limitations of traditional geometry-driven scanning orders and single-structure prediction schemes, this work proposes a point cloud prediction structure optimization and coding method that jointly considers geometric and texture characteristics. The consistency between geometry and attribute distribution and its influence on prediction performance are systematically analyzed. A multi-candidate scanning order generation method based on affine spatial transformation is proposed, and an adaptive selection and optimization mechanism is realized through residual feedback, achieving cooperative prediction between geometry and attributes. Moreover, a locally optimized pruning algorithm is proposed to balance performance and complexity, significantly reducing encoding time while maintaining competitive coding efficiency. Extensive experiments demonstrate that the proposed method achieves notable coding performance improvement across complex geometric and textured scenes, showing strong generalization ability and practical value.

Finally, to overcome the limitations of traditional point cloud coding in low-bitrate scenarios, particularly the insufficient reconstruction quality and lack of joint geometry–attribute compression, the dissertation introduces GS-PCC, a novel point cloud compression framework based on Gaussian representation. This framework pioneers the integration of 3D Gaussian Splatting (3DGS) into point cloud compression, transforming the representation paradigm from discrete points to compact Gaussian parameterization. On the encoder side, GS-PCC decomposes the input point cloud into two parts: two-dimensional texture and three-dimensional structure, which are jointly compressed by a video encoder and a point cloud encoder, respectively. On the decoder side, super-resolution-based Gaussian optimization and mask-fusion-based rendering strategies are employed to effectively improve visual reconstruction quality under low bitrate conditions. Meanwhile, a perceptual quality evaluation system based on

two-dimensional rendering is adopted, which evaluates reconstruction performance through multi-view rendering and image quality metrics, better aligning with human visual perception. Experimental results demonstrate that the proposed GS-PCC framework significantly outperforms existing mainstream point cloud compression methods, such as G-PCC and Draco, under low bitrate conditions, especially in preserving structural boundaries and texture details.

In summary, this research focuses on the key scientific problems in point cloud coding, and conducts a systematic investigation from prediction mechanisms to prediction structures and coding architectures. This research also proposes multiple innovative and practical high-efficiency point cloud coding and decoding methods. The work achieves notable progress in enhancing prediction accuracy, optimizing prediction structures, and enabling joint geometry–attribute coding. It also drives the evolution of point cloud compression technology from geometry-driven to jointly-driven approaches, and from independent processing to collaborative compression. Overall, the developed methods and technologies provide feasible technical solutions and theoretical support for the enhancement of point cloud compression standards, the efficient transmission of 3D content, and the applications of immersive visual media, demonstrating both academic significance and engineering value.

KEYWORDS: 3D point cloud compression, point cloud coding standard, MPEG G-PCC, attribute prediction, joint geometry-attribute coding