

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2024)12-3612-16

论文引用格式: Chen J H, Fang G C, Li H R, Zhang Y, Huang X H and Guo Y L. 2024. An implicit coding network for point cloud geometry compression. Journal of Image and Graphics, 29(12):3612-3627(陈佳慧, 方广驰, 李浩然, 张晔, 黄小红, 郭裕兰. 2024. 面向点云几何压缩的隐式编码网络. 中国图象图形学报, 29(12):3612-3627)[DOI:10.11834/jig.230906]

面向点云几何压缩的隐式编码网络

陈佳慧, 方广驰, 李浩然, 张晔, 黄小红*, 郭裕兰

中山大学电子与通信工程学院, 深圳 518107

摘要: 目的 现有点云几何压缩算法通常将点云转换为八叉树或带潜在特征的稀疏点, 从而提高数据结构的存储效率。这些方法将点云量化至三维网格点, 导致点云所在表面的精度受限于量化分辨率。针对这一问题, 本文将点云转化为连续的隐式表征, 提出一种基于隐式表征的点云几何压缩算法框架, 以克服量化分辨率对压缩质量的不利影响。**方法** 该框架由基于符号对数距离场的隐式表征与带乘性分支结构的神经网络组成。具体来说, 本文在编码阶段利用神经网络拟合隐式表征, 并对该网络进行模型压缩, 然后在解码阶段结合改进的 Marching Cube(MC) 算法重建点云所在表面, 采样恢复点云数据。**结果** 本文在 ABC(a big CAD model dataset)、Famous 与 MPEG PCC(MPEG point cloud compression dataset) 数据集上进行了点云表面压缩实验。与基准算法 INR(implicit neural representations for image compression) 相比, 本文算法的 L1 倒角损失平均下降了 12.4%, Normal Consistency 与 F-score 指标平均提升了 1.5% 与 13.6%, 压缩效率随模型参数量增大而提升, 平均增幅为 12.9%。与几何压缩标准算法 G-PCC(geometry-based point cloud compression) 相比, 本文算法在存储大小为 10 KB 下依然保持 55 dB 以上的 D1-PSNR 重建性能, 有效压缩上限高于 G-PCC。此外, 消融实验分别验证了本文提出的隐式表征和神经网络结构的有效性。**结论** 实验结果表明, 本文提出的点云压缩算法克服了现有算法的分辨率限制, 不仅提升了表面重建精度, 而且提升了点云表面的压缩效率与有效压缩上限。

关键词: 点云几何压缩; 隐式表征; 三维重建; 模型压缩; 表面提取算法

An implicit coding network for point cloud geometry compression

Chen Jiahui, Fang Guangchi, Li Haoran, Zhang Ye, Huang Xiaohong*, Guo Yulan

School of Electronic and Communication Engineering, Sun Yet-sen University, Shenzhen 518107, China

Abstract: Objective Point clouds captured by depth sensors or generated by reconstruction algorithms are essential for various 3D vision tasks, including 3D scene understanding, scan registration, and 3D reconstruction. However, a simple scene or object contains massive amounts of unstructured points, leading to challenges in the storage and transmission of these point cloud data. Therefore, developing point cloud geometry compression algorithms is important to effectively handle and process point cloud data. Existing point cloud compression algorithms typically involve converting point clouds into a storage-efficient data structure, such as an octree representation or sparse points with latent features. These intermediate representations are then encoded as a compact bitstream by using either handcrafted or learning-based entropy coders. Although the correlation of spatial points effectively improves compression performance, existing algorithms may not

收稿日期: 2024-01-15; 修回日期: 2024-03-29; 预印本日期: 2024-04-05

* 通信作者: 黄小红 huangxh75@sysu.edu.cn

基金项目: 国家自然科学基金项目(U20A20185); 广东省基础与应用基础研究基金项目(2022B1515020103)

Supported by: National Natural Science Foundation of China(U20A20185); Basic and Applied Basic Research Foundation of Guangdong Province, China(2022B1515020103)

fully exploit these points as representations of the object surface and topology. Recent studies have addressed this problem by exploring implicit representations and neural networks for surface reconstruction. However, these methods are primarily tailored for 3D objects represented as occupancy fields and signed distance fields, thus limiting their applicability to point clouds and non-watertight meshes in terms of surface representation and reconstruction. Furthermore, the neural networks used in these approaches often rely on simple multi-layer perceptron structures, which may lack capacity and compression efficiency for point cloud geometry compression tasks. **Method** To deal with these limitations, we proposed a novel point cloud geometry compression framework, including a signed logarithmic distance field, an implicit network structure with the multiplicative branches, and an adaptive marching cube algorithm for surface extraction. First, the point cloud surface (serving as the zero level-set) maps the arbitrary points in space to the distance values of their nearest points on the point cloud surface. We design an implicit representation called signed logarithmic distance field (SLDF), which utilizes the thickness assumption and logarithmic parameterization to fit arbitrary point cloud surfaces. Afterward, we apply a multiplicative implicit neural encoding network (MINE) to encode the surface as a compact neural representation. MINE combines sinusoidal activation functions and multiplicative operators to enhance the capability and distribution characteristics of the network. The overfitting process transforms the mapping function from point cloud coordinates to implicit distance fields into a neural network, which is subsequently utilized for model compression. Through the decompressed network, the continuous surface is reconstructed using the adaptive marching cubes algorithm (AMC), which incorporates a dual-layer surface fusion process to further enhance the accuracy of surface extraction for SLDF. **Result** We compared our algorithm with six state-of-the-art algorithms, including the surface compression approaches based on implicit representation and point cloud compression methods, on three public datasets, namely, ABC, Famous, and MPEG PCC. The quantitative evaluation metrics included the rate-distortion curves of chamfer-L1 distance (L1-CD), normal consistency (NC), F-score for continuous point cloud surface, and the rate-distortion curve of D1-PSNR for quantized point cloud surface. Compared with the suboptimal method (i. e., INR), our proposed method reduces L1-CD loss by 12.4% and improves the NC and F-score performance by 1.5% and 13.6% on the ABC and Famous datasets, respectively. Moreover, the compression efficiency increases by an average of 12.9% along with the growth of model parameters. On multiple MPEG PCC datasets with samples taken from the 512-resolution MVUB dataset, 1024-resolution 8iVFB dataset, and 2048-resolution OwlII dataset, our method achieves a D1-PSNR performance of over 55 dB within the 10 KB range, which highlights its higher effective compression limit compared with G-PCC. Ablation experiments show that in the absence of SLDF, the L1-CD loss increases by 18.53%, while the D1-PSNR performance increases by 15 dB. Similarly, without the MINE network, the L1-CD loss increases by 3.72%, and the D1-PSNR performance increases by 2.67 dB. **Conclusion** This work explores the implicit representation for point cloud surfaces and proposes an enhanced point cloud compression framework. We initially design SLDF to extend the implicit representations of arbitrary topologies in point clouds, and then we use the multiplicative branches network to enhance the capability and distribution characteristics of the network. We then apply a surface extraction algorithm to enhance the quality of the reconstructed point cloud. In this way, we obtain a unified framework for the geometric compression of point cloud surfaces at arbitrary resolutions. Experimental results demonstrate that our proposed method achieves a promising performance in point cloud geometry compression.

Key words: point cloud geometry compression; implicit representation; surface reconstruction; model compression; surface extraction algorithm

0 引言

随着自动驾驶、虚拟现实与增强现实等技术的发展,点云作为一种三维数据表征方式,得到了广泛应用(龙霄潇等,2021)。由于高精度与高密度扫描数据的增加,点云的体量呈现出快速增长的趋势,给

其存储与传输带来挑战(Fang等,2022)。目前,以G-PCC(geometry-based point cloud compression)(Schwarz等,2019)与PCGC(multiscale point cloud geometry compression)(Wang等,2021)为代表的点云几何压缩算法日趋成熟。然而,现有算法通常将点云量化至一定分辨率的坐标点,再对量化点云进行压缩。这一过程不仅增大了量化点云与原始点云

的误差,而且导致点云的精度受限于量化分辨率。事实上,点云通常为物体表面的连续采样点,而现有算法忽视了点云所在表面的连续一致性,因此难以有效应对高精度的点云表面压缩任务。由于原始采样点代表着3D空间连续平面上某坐标位置的占用情况,因此如果在压缩过程中能够直接对连续表面进行压缩,则能够使压缩后的重建点云不受限于分辨率,实现更高的保真度。

现有点云几何压缩算法通常将点云转换为存储高效的数据结构,如八叉树(Schwarz等,2019)或带潜在特征的稀疏点(Wang等,2021;Yu等,2023)。通过熵编码器,数据被进一步编码为二进制码流。尽管现有算法可以充分利用空间点的相关性来降低编码冗余并提高压缩效率,但仍需将点云量化为256~1 024像素范围的网格点,导致点云原始表示的细节丢失(Yang等,2018)。此外,现有算法未充分利用连续点云表面的拓扑特性,存在内存随分辨率剧增的问题(Cao等,2019)。一些算法(Mamou等,2009;Maglo等,2015)将离散点云(图1(a))视为具有连续信息的离散点云表面(图1(b)),从而可以利用更低分辨率的网格点及其连接面信息来表示点云。然而,连接面的引入使得压缩效率下降,且依然难以缓解前文提及的量化损失与分辨率受限等困难。

针对上述问题,本文将离散点云转换为连续的隐式表征(图1(c))。隐式表征通过定义一个连续的场函数来描述点云表面性质。场函数的零值集合隐含了表面的几何形状、颜色或反射率等属性信息。隐式表征无需对点云进行量化,可以克服分辨率对于压缩性能的局限性。现有的隐式表征包括符号距离场(signed distance field, SDF)(Park等,2019)、无符号距离场(unsigned distance field, UDF)(Chibane等,2020)等。SDF将空间中的点坐标映射到该点的最邻近表面点的距离值,符号代表点位于表面内或表面外。UDF仅有距离值,没有符号信息,可以作用于任意物体表面。然而,将SDF与UDF直接作用于点云表面仍然存在一些困难:1) SDF需要界限清晰的内表面和外表面,因此只能作用于水密物体而无法应用于离散点云;2)UDF虽然可以作用于离散点云,但在表面处不可导,难以用网络对其进行优化学习,无法直接作用于点云几何压缩。

此外,点云几何压缩在保证表面重建性能(图1

(d))的同时,还要尽可能减少内存开销,以获得较高的压缩效率。本文利用神经网络对隐式表征进行拟合,将场函数进一步转换为网络参数,这不仅需要网络具备足够的函数空间,而且需要网络参数量较小。然而,现有的隐式神经网络通常采用全连接结构,通过增加网络参数来扩大函数空间的方式,难以找到重建性能与压缩效率之间的平衡点。

为了解决现有隐式表征及其神经网络在点云几何压缩任务中存在的问题,本文在UDF基础上进行位置偏移与对数变换,得到符号对数距离场(signed logarithmic distance field, SLDF)。SLDF表征不仅可以有效表示任意分辨率的点云表面,而且在重新定义的表面上具有可导性,从而克服了UDF难以用网络优化学习的问题。然后,本文将乘性分支引入SIREN(sinusoidal representation network)网络(Sitzmann等,2020),增大了SIREN网络的函数空间,得到更加高效的乘性隐式神经编码网络(multiplicative implicit neural encoding network, MINE)。最后,本文利用表面融合算法提高重建表面精度,再对表面采样即可得到编解码后的点云。

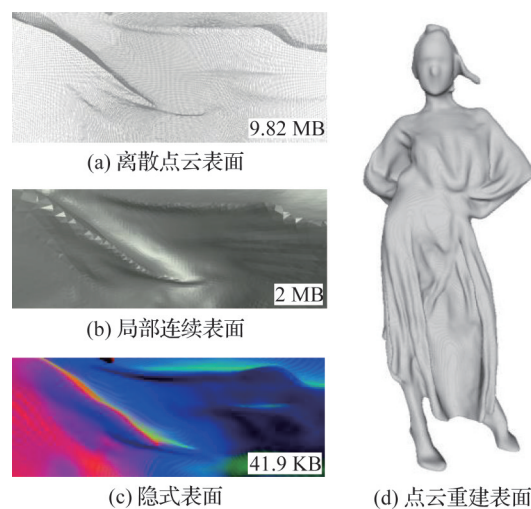


图1 点云表征与点云重建表面

Fig. 1 Point cloud representations and reconstructed surface from point cloud ((a) discrete point cloud surface; (b) continuous point cloud surface; (c) implicit surface; (d) reconstructed surface from point cloud)

本文的贡献如下:1)提出一种基于隐式表征的点云几何压缩框架,不仅适用于现有离散点云几何压缩,而且支持连续的点云表面压缩任务。2)设计了SLDF表征用于描述点云表面信息。SLDF表征可

以描述任意分辨率的点云表面,解决了现有算法的压缩效率受限于分辨率的问题。3)构造了MINE网络用于点云表面编码。MINE网络通过引入乘性分支结构,扩大了网络的函数空间并改善网络参数分布特性,有效提升了点云几何压缩效率。4)在ABC(a big CAD model dataset)、Famous、MPEG PCC(MPEG point cloud compression dataset)和ShapeNet这4个数据集上的实验结果表明,本文算法在点云表面重建精度和压缩效率等方面均取得了性能提升。

1 相关工作

不同于现有传统或深度学习压缩方法,本文将三维重建算法中的隐式表征引入至点云几何压缩任务。因此,本文按照点云的编码表征方式对现有算法进行分类,即基于显式表征的点云几何压缩算法与基于隐式表征的三维重建算法。

1.1 基于显式表征的点云几何压缩算法

1.1.1 传统点云几何压缩方法

传统点云几何压缩方法利用经典的数据结构表示点云几何信息,再结合预测编码、熵编码等压缩算法进一步降低冗余。Schwarz等人(2019)提出的G-PCC存在Octree与Trisoup两种编码模式。Octree模式采用八叉树结构将三维空间递归地划分为8个均匀的子空间,再对树节点由浅至深进行预测编码与熵编码;Trisoup模式利用三角形网格的拓扑结构,对三角面片的顶点坐标以及构成三角面片的邻接索引进行算术编码与熵编码。Jang等人(2019)提出的V-PCC(video-based point cloud compression)则是将点云投影为不同角度的深度图像,再用高效率视频编码算法(high efficiency video coding, HEVC)(Sullivan等,2012)压缩。传统算法关注点云的空间坐标及其拓扑关系,因此能够实现高保真度有损压缩或无损压缩。然而,此类方法需将点云量化至预设分辨率的网格点,此过程会导致无法逆转的信息丢失。此外,随着网格分辨率提高,点云数据量呈现指数级增长,给压缩效率与计算效率带来不利影响。

1.1.2 基于深度学习的点云几何压缩方法

由于传统方法均为基于手工设计的方法,提取点云局部相关性的性能有限。Quach等人(2020)利用卷积网络将点云映射到低维特征向量,根据特征向量分布将其熵编码为比特流,再用反卷积解码器

将低维特征向量重构为原始点云。Wang等人(2021)采用稀疏卷积将原始点云下采样为带特征向量的稀疏点,进一步降低时空复杂度。朱威等人(2023)引入密集残差结构与通道注意力机制,提升了网络的特征提取能力。Huang等人(2020)与Que等人(2021)则在八叉树编码的基础上,分别从父节点及兄弟节点获取上下文信息来提高熵模型的预测概率。与传统方法相比,基于深度学习的点云几何压缩算法捕捉点云中的局部结构和特征能力更强。然而,这类方法沿用了传统方法的显式表征,依然难以避免量化误差与内存随分辨率增加而剧增的问题。

1.2 基于隐式表征的三维重建算法

现有点云几何压缩算法因采用了离散的显式表征而导致其性能受限于表面分辨率。而隐式表征作为连续的表面函数,跳过了点云量化过程,因此可以实现任意分辨率的点云表面重建,克服现有点云几何压缩算法受限于点云分辨率的问题。

1.2.1 隐式表征

隐式表征作为一种高效的连续表征,已广泛应用于三维重建(Erler等,2020)、图像压缩(Dupont等,2021)和视频压缩(Chen等,2021)等领域。Erler等人(2020)与Chibane等人(2020)利用神经网络将点云映射到空间距离场,从点云中重建连续表面。Dupont等人(2021)与Strümpfer等人(2022)则是将图像像素映射到对应RGB值,结合模型压缩,实现超高压缩比的图像压缩。Chen等人(2021)将视频编码为神经网络,提高了编码速度与解码质量。

常见的三维隐式表征包括二值化占有场(occupancy field, OF)(Mescheder等,2019)、符号距离场(signed distance field, SDF)(Park等,2019)、无符号距离场(unsigned distance field, UDF)(Chibane等,2020)、神经辐射场(neural radiance field, NeRF)(Mildenhall等,2020)和混合场(neural implicit surface, NeuS)(Wang等,2021)等。这些表征将点云映射为值域连续的空间场,再用网络学习映射函数。其中,NeRF、NeuS用于多视图的三维重建,本文更加关注用于点云三维重建的SDF、OF与UDF表征。OF与SDF表征由于仅适用于存在严格内外之分的闭合点云表面,需要对非闭合点云表面进行水密化的预处理。这一局限性使得这类表征难以作用于非水密表面。

为了解决非水密表面的表征问题,Chibane 等人(2020)提出了无符号距离场UDF,UDF支持任意3D形状的表征。然而,UDF在表面上存在梯度消失现象,导致表面上的法向量模糊与精度丢失问题。为此,Wang等人(2022)与Ye等人(2022)在UDF基础上,分别引入语义特征与表面相交函数,增加额外信息以提升表面精度。Ren等人(2023)利用网络学习点云特征预测UDF及其梯度,并结合重建表面的倒角损失优化拟合性能。然而,这些改进方法通常采用基于学习的思路,作用于点云几何压缩任务时,需要同时编码原始点云与其重建网络,才能在解码端中有效恢复点云表面。这一问题导致此类方法难以直接用于点云几何压缩任务。近期,Zhou等人(2023)通过改进网络的损失函数实现了拟合表面的精度提升。Long等人(2023)将UDF表征与表面概率函数相结合,通过优化渲染图像的颜色损失,实现了渲染性能提升。这类基于优化的方法将UDF与其他领域的损失函数相结合,避免了直接优化UDF而表面梯度不存在的问题。因此,这类方法虽然可以解决UDF在表面上不可微的难题,但在作用于点云几何压缩任务时,依然存在边界定义模糊导致表面精度受限以及难以高效提取表面的问题。

与上述UDF及其改进方法相比,本文的SLDF表征不仅提升了隐式表征的泛化性,而且存在表面梯度与可学习的表面临界值。因此,SLDF可利用网络学习到精确表面信息,并将原始点云作为编码对象,在解码端无需先验点云进行表面重建。此外,SLDF在定义的表面上具有内外差异与数值连续的特性,可以通过改进Marching Cube(MC)算法高效提取表面,避免了边界模糊问题。综合分析,SLDF比现有方法更加适用于点云几何压缩任务。

1.2.2 隐式神经网络

上述的隐式表征为连续的场函数,需要使用网络空间为连续函数的神经网络进行拟合。现有算法通常采用多层感知器(multilayer perceptron, MLP)学习隐式表征。针对神经网络具有低频特性的问题,Sitzmann等人(2020)提出的SIREN网络在线性层后添加正弦激活函数,提高网络的高频拟合能力。Fathony等人(2021)提出的MFN(multiplicative filter network)网络利用多级位置编码输入,增加了网络的高频特性。Lindell等人(2022)在MFN基础上对多级位置编码层的权重进行约束,使网络具有

可控的频率特性。此类网络可以用既定的网络参数量表征任意分辨率的点云表面。然而,随着点云表面复杂度的提升,网络参数量需随之增加以提高网络的非线性拟合能力。这一困难导致现有网络结构难以平衡重建精度与模型压缩比的要求。

1.2.3 表面提取算法

神经网络拟合隐式表征后,需要利用表面提取算法从神经网络中恢复原始表面。Lorensen和Cline(1987)提出的Marching Cubes算法根据量化网格点所在立方体的顶点函数值,生成与等值面相合的三角形面片,并对相邻体素的面片进行拼接,得到等值面对应的连续点云表面。Hart(1996)提出的Sphere Tracing算法通过光线追踪来确定光线与物体表面的相交点,进而生成点云表面。由上述原理可知,此类表面提取算法仅支持具有单调一致性的等值面,无法直接作用于本文提出的SLDF表征。

在此研究背景下,本文围绕上述隐式表征及其神经网络在点云几何压缩的不足展开研究。

2 本文算法

2.1 整体框架

本文提出了一种基于隐式表征的点云几何压缩框架,包含SLDF表征变换、MINE网络、AMC表面提取3个模块,如图2所示。具体来说,该框架首先将点云转换为SLDF表征,并使用MINE网络将空间点坐标映射至对应的SLDF值。然后,本文算法在编码阶段采用模型压缩算法将MINE的网络参数编码为二进制码流,并在解码阶段利用AMC算法从解压缩后的编码网络中提取连续表面,进而采样得到点云表面。

2.2 符号对数距离场

2.2.1 问题阐述

隐式表征将对象描述为一个连续的空间函数。其中,二值化占有场OF、符号距离场SDF和无符号距离函数UDF是3种常用的隐式表征。给定一个空间点 $\mathbf{x} \in \mathbf{R}^3$,二值化占有场定义了空间的占有情况,表面点被表示为1,非表面点被表示为0,表达式为

$$\text{OF}(\mathbf{x}) = \{0, 1\} \quad \mathbf{x} \in \mathbf{R}^3 \quad (1)$$

符号距离场为空间点 $\mathbf{x}$ 提供了一个实数值 s ,表示该点到最近物体表面的距离。 s 的符号表示点 $\mathbf{x}$ 相对于表面的位置:若 s 为负,则点 $\mathbf{x}$ 位于表面内部;

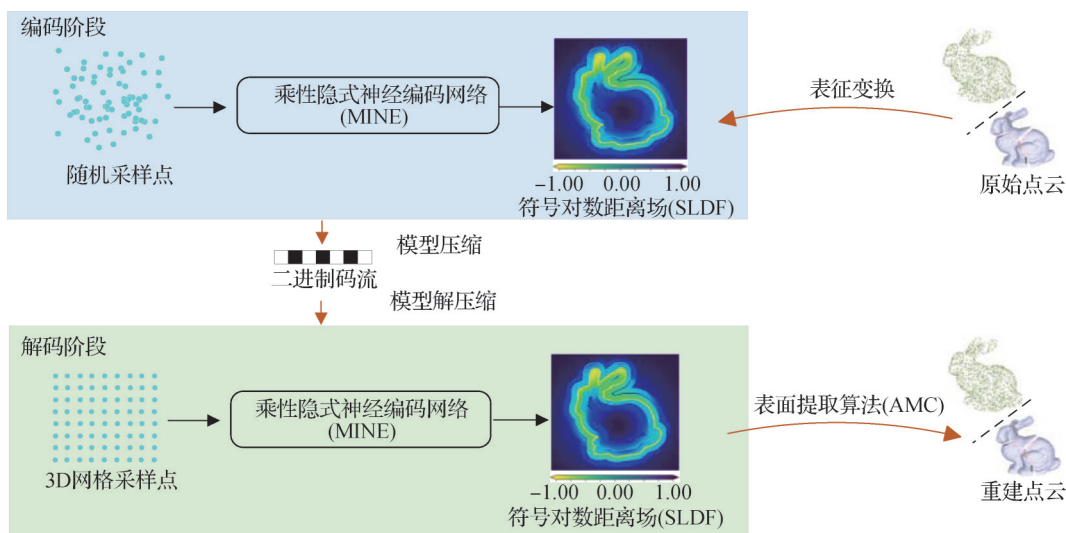


图2 本文算法整体流程

Fig. 2 The whole pipeline of our proposed method

若 s 为正,则点 $\mathbf{x}$ 位于表面外部。数学表达式为

$$\text{SDF}(\mathbf{x}) = s \quad \mathbf{x} \in \mathbf{R}^3 \quad (2)$$

与符号距离场相比,无符号距离场提供了到表面最近点的距离绝对值,没有表面内外之分,表达式为

$$\text{UDF}(\mathbf{x}) = |s| \quad \mathbf{x} \in \mathbf{R}^3 \quad (3)$$

为了实现有效的压缩,隐式表征应具备以下特性:1)泛化性:能够处理任意拓扑结构的点云;2)准确性:能从隐式表征中重建高精度的原始数据。由于SF和SDF需要界限分明的内外表面,无法用于离散点云;UDF虽然能够处理任意拓扑结构的点云,却在表面上不可微分,无法采用基于优化方法的网络结构。

2.2.2 符号对数距离场的定义

为了解决上述问题,本文设计了一种基于表面厚度假设的符号对数距离场SLDF。具体而言,将离散点云所在的表面看做具有一定厚度的表面,在UDF基础上引入表面距离偏移量。这一操作使得该隐式表征具备如下特性:1)在表面外具有UDF特性,能够表征任意拓扑结构的点云表面;2)在重新定义的表面具有SDF特性,在零等值面上可微分,能被隐式神经网络高效学习。

由于此处引入了厚度假设,该隐式表征的零等值面与实际表面真值存在差距。为了减少厚度假设带来的影响,本文算法对该隐式表征进行对数变换,扩大表面邻近值的差异且扁平化表面外的函数值,使重新定义的表面与实际表面真值更加贴合。此外,由于对数函数在靠近表面的几何位置上具有更大的优化梯度,因此有利于提高表面的拟合性能。

最终,符号对数距离场的定义为

$$\text{SLDF}(\mathbf{x}) = -[\log(|s| + \varepsilon) + \delta] \quad \mathbf{x} \in \mathbf{R}^3 \quad (4)$$

式中, s 为在归一化后的空间点到最邻近表面点的距离。 δ 为点云表面厚度,本文将其设置为 10^{-3} ,取对数后计入原式。 ε 为极小偏移值,本文将其设置为 10^{-5} ,以确保对数函数在零等值面处存在定义。通过控制 δ 的值,符号对数距离场有效表征了具有一定厚度的表面区域,并将空间分为表面内部与表面外部两部分,如图3所示。最后,本文对SLDF进行了一次符号反转,使本文在表面外的符号定义与UDF相同,均为正值。

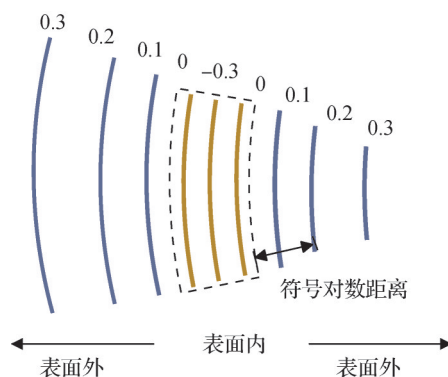


图3 符号对数距离场

Fig. 3 Signed logarithmic distance field

2.3 乘性隐式神经编码网络

2.3.1 网络结构

本文提出了一个神经编码网络,使用网络的分布式参数拟合SLDF表征,其隐式神经网络的函数空

间为

$$f(\mathbf{x}_i; \Theta) \approx \text{SLDF}(\mathbf{x}_i) \quad (5)$$

式中, $\mathbf{x}_i \in \mathbf{R}^3$ 表示输入的三维点坐标, Θ 表示神经网络的可学习参数。

由于神经网络倾向于学习网络输出函数的低频信息, SIREN网络(Sitzmann等, 2020)在MLP的线性层后添加正弦激活层以提高网络的高频特征提取能力。由于SIREN网络可以有效实现任意连续函数的拟合, 在三维重建(Strümler等, 2022)、图像压缩(Dupont等, 2021)等领域均取得了优异的性能。然而, 由于点云表面的维度高、空间复杂度大, 将SIREN网络用于拟合点云表面需要较大的函数空间, 需要通过增加网络层数或增大隐藏层大小来提高拟合精度。这使得SIREN网络难以平衡点云几何压缩的压缩效率与重建精度的要求。

针对SIREN网络的不足, 本文将乘性分支结构(Fathony等, 2021)引入到SIREN网络(Sitzmann等, 2020)中, 进而提出了MINE网络。如图4所示, SIREN网络将输入坐标 $\mathbf{x}$ 位置编码为隐藏层向量 $\mathbf{z}_0 = \sin(\mathbf{w}\mathbf{x} + \varphi_0)$, 再通过一系列带正弦激活函数的线性层, 数学表示为 $\mathbf{z}_i = \sin(\mathbf{W}_i \mathbf{z}_{i-1} + \mathbf{b}_i)$ 。在此基础上, MINE网络在各个隐藏层的输出添加原始输入的单层分支, 这些分支网络通过Hadamard乘积(记做 $\otimes$)并入到主干网络, 数学表示为 $\mathbf{z}_i = \sin(\mathbf{W}_i \mathbf{z}_{i-1} + \mathbf{b}_i) \otimes \sin(\mathbf{w}_i \mathbf{x} + \varphi_i)$ 。最终, 通过线性全连接层输出对应函数值 $\mathbf{y} = \mathbf{W}_i^{\text{out}} \mathbf{z}_i + \mathbf{b}_i^{\text{out}}$ 。

完成编码网络的训练后, 本文采用模型压缩算法提高网络模型的内存效率。由于剪枝(Han等, 2016)与蒸馏(Hinton等, 2015)对模型精度影响大且时间成本高, 本文沿用COIN(compression with implicit neural representations)(Dupont等, 2021)的模型压缩策略, 即量化与算术编码。

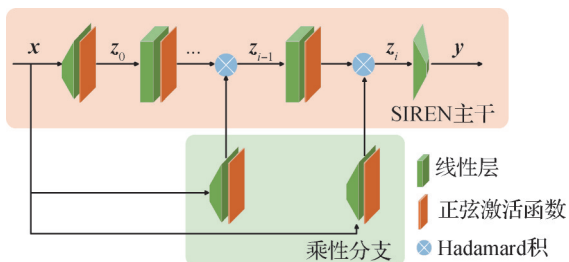


图4 乘性隐式神经编码网络

Fig. 4 Multiplicative implicit neural encoding network

2.3.2 理论分析

MINE框架与SIREN网络差异在于隐藏层特征 $\mathbf{z}_i$, 本文使用泰勒展开对网络的 $\mathbf{z}_i$ 进行分析。SIREN(Sitzmann等, 2020)与MINE的隐藏层特征 $\mathbf{z}_i$ 可以整理为式(6)与式(7)。由式(6)可知, SIREN网络的线性层系数构成了隐藏层特征的高阶项系数。

$$\mathbf{z}_i = \sin(\mathbf{W}_i \mathbf{z}_{i-1} + \mathbf{b}_i) = \sum_{k=0}^{\infty} \sum_{u=0}^{2k+1} \eta(k, u) \times \mathbf{z}_{i-1}^u \quad (6)$$

式中, $\eta(k, u)$ 为 $\mathbf{W}_i$ 与 $\mathbf{b}_i$ 的高幂次乘积项函数,

$$\eta(k, u) = \frac{(-1)^k}{(2k+1)!} C_{2k+1}^u \mathbf{W}_{2k+1}^u \mathbf{b}_i^{2k+1-u}。$$

乘性分支可以增加隐藏层特征的高阶项及其系数自由度, 具体为

$$\mathbf{z}_i = \sin(\mathbf{W}_i \mathbf{z}_{i-1} + \mathbf{b}_i) \otimes \sin(\mathbf{w}_i \mathbf{x} + \varphi_i) = \sum_{k=0}^{\infty} \sum_{u=0}^{2k+1} \sum_{v=0}^{2k+1-u} (\eta_1 + \eta_2)(k, u, v) \times \mathbf{x}^u \times \mathbf{z}_{i-1}^v \quad (7)$$

式中, $\eta_1(k, u, v)$ 与 $\eta_2(k, u, v)$ 为 $\mathbf{W}_i, \mathbf{w}_i, \mathbf{b}_i$ 与 φ_i 的高

$$\text{幂次乘积项函数, } \eta_1(k, u, v) = \frac{(-1)^k}{(2k+1)!} \mathbf{w}_i^u \mathbf{W}_i^v (\varphi_i + \mathbf{b}_i - \frac{\pi}{2})^{2k+1-u-v}, \quad \eta_2(k, u, v) = \frac{(-1)^{k+v}}{(2k+1)!} \mathbf{w}_i^u \mathbf{W}_i^v (\varphi_i + \mathbf{b}_i + \frac{\pi}{2})^{2k+1-u-v}。$$

乘性分支与线性层系数 ($\mathbf{W}_i, \mathbf{w}_i, \mathbf{b}_i, \varphi_i$) 均作为输入特征的系数组成部分, 提升了隐藏层特征的高阶项系数自由度。 $\mathbf{x}^u$ 作为乘积项, 使隐藏层特征的高阶项更加丰富。因此, 与SIREN网络相比, MINE网络具有更大的模型空间, 可以拟合更复杂的隐式表征, 提升点云表面的拟合精度。

2.4 表面提取算法

在解码阶段, 本文利用表面提取算法从隐式表征中重建点云表面。由于现有的表面提取算法(Lorensen和Cline, 1987; Hart, 1996)利用等值面确定表面位置, 需要隐式表征在表面上具有唯一的零等值面。然而, 由于本文2.2节的SLDF表征进行了厚度假设, 导致SLDF表征的零等值面为与实际表面等距的对称偏移表面, Marching Cube算法提取的表面与实际表面存在误差, 如图5所示。

为解决此问题, 在Marching Cube算法后添加表面融合步骤, 具体过程如算法1所示。首先, 选择两个不同的等值面阈值 δ_c 和 δ_f 结合Marching Cube算法以生成两个原始表面。其中, δ_c 是待优化表面 S_c 所在等值面阈值, δ_f 是更加贴合真值的表面 S_f 的等值面

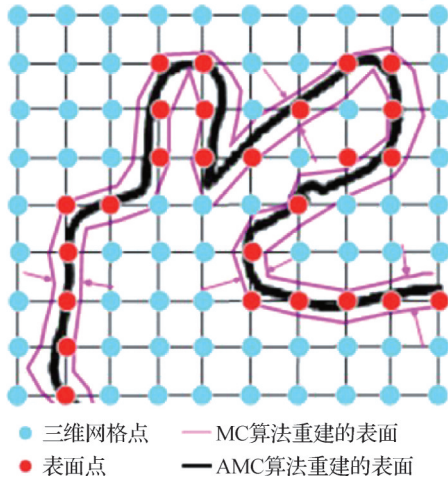


图5 MC算法与AMC算法重建的表面
Fig. 5 The surface reconstructed by MC algorithm and AMC algorithm

阈值。本文将 δ_e 和 δ_f 分别设置为0与 10^{-3} 。然后,将表面 S_f 的点云构建为KD(k-dimensional)树。对于表面 S_e 上的每个点,利用KD树找到表面 S_f 中的最近邻点集,由此计算表面 S_e 到 S_f 的搜索方向 n_i 和距离 d_i 。最终,本文设定的点数阈值截取局部表面,以确保局部表面涵盖足够丰富的双层表面信息,并对每个局部表面点计算最近邻点群的平均方差。将局部表面点数阈值设为 2×10^3 ,表面点的最近邻点群设为 10^2 ,再以最小化表面局部的平均方差为目标更新 S_e ,重复迭代步骤直至达到收敛条件,得到最终的重建表面。

由于SLDF表征具有表面两侧符号相异的零界面,本文设计的表面提取算法可以结合MC算法提取到精确的等值表面。此外,表面融合步骤利用了偏移表面的对称特性来消除厚度表面假设带来的误差影响并重建最终的表面,从而有效避免了UDF临界值模糊和难以提取连续表面的问题。值得注意的是,表面融合算法虽然降低了从SLDF中重建的连续表面与原始点云表面的位置差距,但存在表面重叠问题。由于本文关注的是点云表面压缩,因此AMC表面采样后依然为点云表面的有效点,可以有效缓解SLDF表征的厚度假设带来的精度下降问题。

算法1如下所示。

输入:三维网格点集合 X ,最大迭代次数 T ,外层表面的阈值 δ_e ,内层表面的阈值 δ_f ;

输出:优化后的重建表面 S_e ;

1) 初始化表面 $S_e = MC(X, \delta_e)$, $S_f = MC(X, \delta_f)$;

- 2) 初始化 S_f 的KD树;
- 3) for each 顶点 v_i in S_e do;
- 4) 计算最近邻点 v_i :对 S_f 的KD树最近邻搜索 v_i ;
- 5) 计算递进距离 $d_i = \|v_i' - v_i\|$;
- 6) 计算递进方向 $n_i = (v_i' - v_i) / \|v_i' - v_i\|$;
- 7) end;
- 8) 初始化搜索距离 $d = \sum_{i=0}^N d_i / N$;
- 9) $t = 0$;
- 10) while $t < T$ do;
- 11) 计算更新后表面 $S_e' = \{v_i + n_i d \mid v_i \in S_e\}$;
- 12) 截取局部表面 $P \subseteq S_e, P' \subseteq S_e'$;
- 13) 计算局部表面点的平均方差;
- $\sigma_e^2 = E\{\text{var}(x_i \text{ 在 } P_e \text{ 的 KD 树中的最近邻点群}) \mid x_i \in P_e\}$;
- $\sigma_e'^2 = E\{\text{var}(x_i \text{ 在 } P_e' \text{ 的 KD 树中的最近邻点群}) \mid x_i \in P_e'\}$;
- 14) if $\sigma_e'^2 < \sigma_e^2$ do;
- 15) 更新表面 $S_e = S_e'$;
- 16) else;
- 17) 更新搜索距离 $d = d/2$;
- 18) end;
- 19) $t = t + 1$;
- 20) end。

3 实验与分析

在本节中,将本文算法与现有算法在点云的几何压缩性能和三维重建精度等方面进行了详细对比。此外,本节也提供了SLDF表征与MINE网络的消融实验结果。

3.1 实验设置

3.1.1 数据集

为保证实验的公平性,本文分别在连续表面与量化点云数据集上进行实验,并对数据集进行归一化处理。ABC(Koch等,2019)、Famous(Turk和Levoy,1994)(<http://www-graphics.stanford.edu/data/3Dscanrep/>)和MEPG PCC数据集(<https://mpeg-pcc.org/index.php/pcc-content-database/>)用于评估现有算法的点云几何压缩性能,ShapeNet数据集(Chang等,2015)用于评估现有算法的三维重建精度。其中,ABC是包含孔洞的工件CAD(computer-aided design)模型数据

集, Famous 是从多角度图像中重建的连续表面数据集, 这两个数据集允许在任意分辨率下对表面点进行采样。MPEG PCC 中的 MVUB (Loop 等, 2016)、8iVFB (d'Eon 等, 2017)、Owlii (Xu 等, 2017) 分别对应分辨率为 512、1 024、2 048 的量化点云数据集。ShapeNet 数据集常用于衡量算法的表面重建性能, 是具有平滑表面的三维 CAD 模型集, 涵盖丰富的物体类别。

3.1.2 基准算法

在点云的连续表面压缩方面, 本文算法与 NP (learning signed distance functions from point clouds by learning to pull space onto surfaces) (Ma 等, 2021)、BACON (band-limited coordinate networks for multi-scale scene representation) (Lindell 等, 2022)、INR (implicit neural representations for image compression) (Strümppler 等, 2022)、LevelSetUDF (Zhou 等, 2023) 算法进行了比较。此四者均为基于优化的方法, NP 与 LevelSetUDF 采用带 ReLU 激活函数的 MLP 网络, 结合网络梯度补偿误差, 从而降低点到邻近表面的位置误差; BACON 与 INR 分别采用了带多级位置编码的 MFN 网络与带正弦激活函数的 SIREN 网络, 优化 SDF 对点云的表征能力。

在点云的量化压缩方面, 本文算法和 G-PCC (Schwarz 等, 2019) 与 INR (UDF) 算法比较。G-PCC 是传统点云几何压缩标准算法, 采用八叉树结构以降低空间信息冗余。由于 INR 算法的 SDF 表征无法直接作用于点云表面, 本文在量化点云压缩任务中比较的 INR 算法为 UDF 表征与 SIREN 网络的组合。

在点云表面重建的补充实验中, 本文与 P2S (learning implicit surfaces from point cloud) (Erler 等, 2020)、NDF (neural unsigned distance fields for implicit function learning) (Chibane 等, 2020)、GeoUDF (Ren 等, 2023)、NP (Ma 等, 2021)、BACON (Lindell 等, 2022)、INR (Strümppler 等, 2022)、LevelSetUDF (Zhou 等, 2023) 算法进行比较。其中, P2S、NDF、GeoUDF 为基于学习的方法, P2S 采用 SDF 表征, NDF 与 GeoUDF 采用 UDF 表征。

3.1.3 评价指标

点云几何压缩的性能评估包含两个方面: 重建精度与压缩效率。常用的评价指标为率-失真优化曲线, 该曲线表示在不同压缩强度下对应的重建精度。

对于重建精度, 连续点云表面的评价指标为 Chamfer-L1 distance (L1-CD) (Fan 等, 2017)、NC (normal consistency) (Mescheder 等, 2019) 与 F-score (Tatarchenko 等, 2019)。其中, L1-CD 度量了两组点云之间最近点对的距离平均值; NC 度量了两组点云之间最近点对的法向量角度差异; F-score 评估了两组点云之间最近点对的距离落在给定阈值内的分类准确性, 本文将 F-score 的阈值 μ 设为 0.005。量化点云表面的评价指标为 D1-PSNR (Wang 等, 2021), 该指标度量了两组离散点云的误差距离信噪比 (peak signal to noise ratio, PSNR)。具体而言, L1-CD 越低, NC、F-score、D1-PSNR 越高, 重建精度越优。

对于压缩效率, 现有几何压缩算法沿用 G-PCC (Schwarz 等, 2019) 中的 BPP (bit-per-point), 使得量化分辨率对内存的影响被忽略, 本文沿用 INR (Strümppler 等, 2022) 的实际内存占用指标 (KB), 使分辨率对内存的影响也被充分考虑。内存占用越小, 压缩强度越大。

3.1.4 实验参数

本文在一台配备 Intel® Core (TM) i7 CPU, NVIDIA GeForce RTX 3090 GPU, Ubuntu20.04 操作系统的计算机上开展算法的实验验证。对比实验均使用相关论文的开源代码进行实现。

本文分别在 ABC、Famou、MPEG PCC 数据集上进行连续点云表面与离散点云压缩实验。其中, 本文从 MPEG PCC 数据集中选取分辨率分别为 512、1 024、2 048 的 MVUB、8iVFB、Owlii 数据集进行离散点云的几何压缩实验。此外, 本文从 ShapeNet 数据集中选取 8 个物体类别验证本文在表面重建任务上的有效性。

实验的具体数据及类别如表 1 所示, 本文对上述样本的结果取统计平均作为最终的实验结果。在数据集预处理过程中, 本文对点云表面下采样后添加噪声作为伪随机点输入网络。SLDF 表征通过在连续点云表面上计算点到最邻近表面的垂直距离以及在离散点云表面上计算点到最邻近点的距离, 具体变换过程如式(4)所示。随后, SLDF 表征作为网络输出的真值。

本文采用均方误差 (mean square error, MSE) 损失函数优化的网络输出与隐式表征 SLDF 的差异。在点云表面压缩任务中, 本文与对比方法分别采用 3 层 64 维、3 层 128 维、3 层 256 维等隐藏层参数, 代

表不同的模型压缩效率。在三维重建任务中,本文与对比方法均采用3层256维的隐藏层。由于LevelSetUDF受模型参数影响较大,此处沿用8层256维隐藏层进行对比。在训练过程中,batch大小为 10^4 ,学习率初始化为 10^{-3} ,使用ReduceLR-OnPlateau调度器调整学习率,如损失函数连续500 epochs未下降,则学习率减半。此外,本文采用AI Model Effi-

ciency Toolkit库将模型量化至8位精度,并用二进制算术编码算法对量化后的权重进行无损压缩。

此外,本文在ABC数据集上进行了连续点云表面的消融实验,并分别在MPEG PCC的8iVFB数据集与ShapeNet数据集上进行了离散点云压缩与三维重建的消融实验,验证SLDF表征、MINE网络与AMC表面提取算法的有效性。

表1 实验数据集

Table 1 Datasets used for experiments

数据集	数据或数据类别	数量
ABC(Koch等,2019)	-	96
Famous(Turk和Levoy,1994)	armadillo,dragon,hand,happy,horse,bunny,liberty,tortuga,netsuke,angel	10
MVUB(Loop等,2016)	andrew,david,phil,sarah	4
8iVFB(d'Eon等,2017)	longdress,loot,redandblack,soldier	4
Owlii(Xu等,2017)	basketball player,dancer,exercise model	4
ShapeNet(Chang等,2015)	airplane,cabinet,chair,display,lamp,sofa,car,vessel	80

注:“-”表示该数据集没有明确的分类信息。

3.2 实验结果

3.2.1 点云的连续表面压缩实验

本文算法在ABC与Famous数据集上的几何压缩性能如图6和图7所示。表面重建性能均随着模型参数的内存增加(隐藏层大小从64增加到256)而增大。与现有算法相比,本文算法在更高的压缩效率下达到最优的重建性能。与基准方法INR相比,在ABC数据集上,L1-CD损失平均下降了18.5%,NC与F-score性能平均提高了1.48%与20.3%。在Famous

数据集上,L1-CD损失平均下降了6.31%,NC与F-score性能平均提高了1.52%与6.88%。这表明本文算法在同等压缩强度下,提高了连续点云表面的重建精度。此外,本文算法虽然比INR的SIREN网络增加了乘性分支,在相同隐藏层大小与数量下,压缩后的内存开销低于INR。随隐藏层大小的增加,内存开销的下降幅度从11%提升至14.7%。实验结果表明,本文算法有效改善了模型参数分布,并提高了现有算法在连续点云表面的压缩效率。

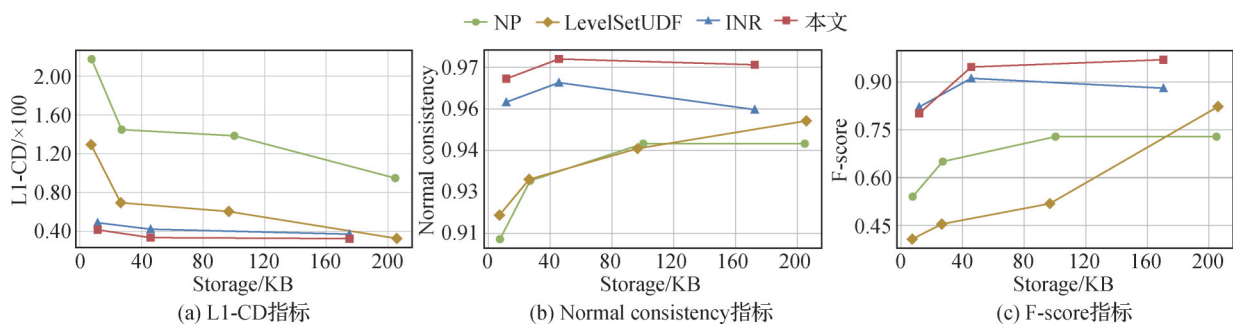


图6 在ABC数据集上的连续点云几何压缩性能曲线

Fig. 6 Result of continuous point cloud geometry compression on the ABC dataset((a) L1-CD; (b) Normal consistency; (c) F-score)

连续点云的压缩表面可视化结果见图8所示,本文算法在重建ABC数据集的孔洞结构与Famous数据集的复杂表面上更加精确。此外,本文在隐藏层大小为64(存储大小为11.6 KB)的重建效果优于隐藏层大

小为128的INR(存储大小为46.3 KB),表明本文可以更高效地实现点云表面的压缩。在“Dragon”模型中,本文算法在保证重建精度的同时实现了1395:1的压缩比,证明本文具有较高的点云表面压缩上限。

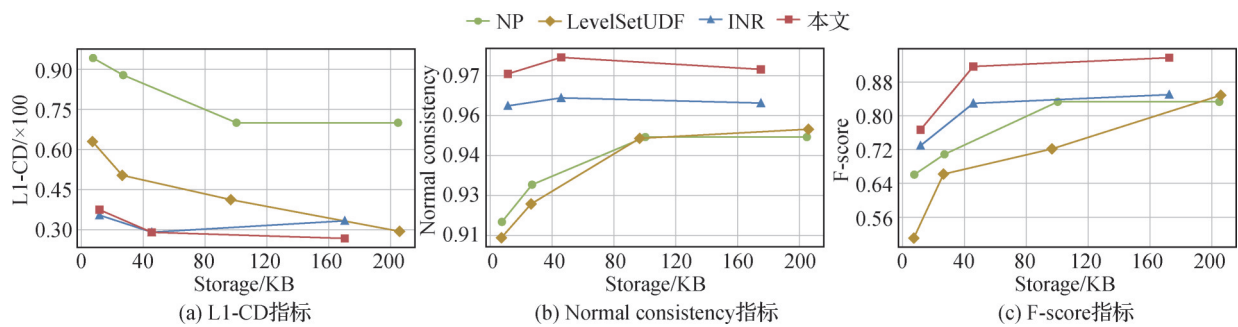


图7 在Famous数据集上的连续点云几何压缩性能曲线
Fig. 7 Result of continuous point cloud geometry compression on the Famous dataset
((a) L1-CD; (b) Normal consistency; (c) F-score)

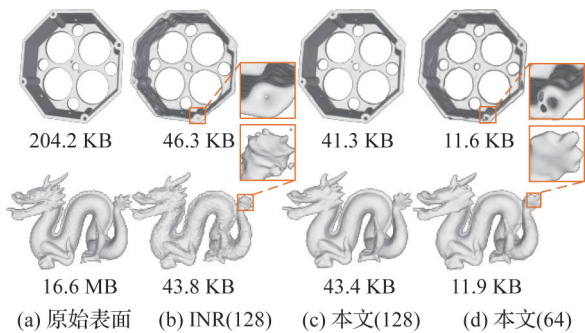


图8 在ABC、Famous数据集上的连续点云几何压缩重建可视化结果
Fig. 8 Visualization of continuous point cloud geometry compression on ABC and Famous dataset ((a) original surfaces; (b) INR (128); (c) ours (128); (d) ours (64))

3.2.2 点云的量化表面压缩实验

在MPEG PCC数据集上,本文算法的几何压缩性

能如图9所示。在相同的压缩框架下,与结合UDF表征的INR网络相比,本文算法在不同压缩强度下D1-PSNR平均提高15 dB,这表明本文提出的SLDF表征能够有效改善UDF在零等值面处不可导性对网络拟合精度的影响。此外,在10 KB的内存占用下,本文算法在512、1 024、2 048分辨率的数据集上均保持了55 dB以上的重建精度,远优于G-PCC。这些实验结果表明本文算法克服了分辨率对编码精度的限制,比G-PCC支持更高的压缩分辨率。然而,从图9可以发现,随着内存增加,本文算法的重建精度逐渐劣于G-PCC。这是由于G-PCC针对点云坐标进行压缩,随内存增大趋于无损压缩,而本文算法拟合的是离散网格点所在的连续表面,重建表面的采样点难以与原始网格点完全重合,因此本文算法比G-PCC更加适合高压压缩率情况下的量化点云表面压缩。

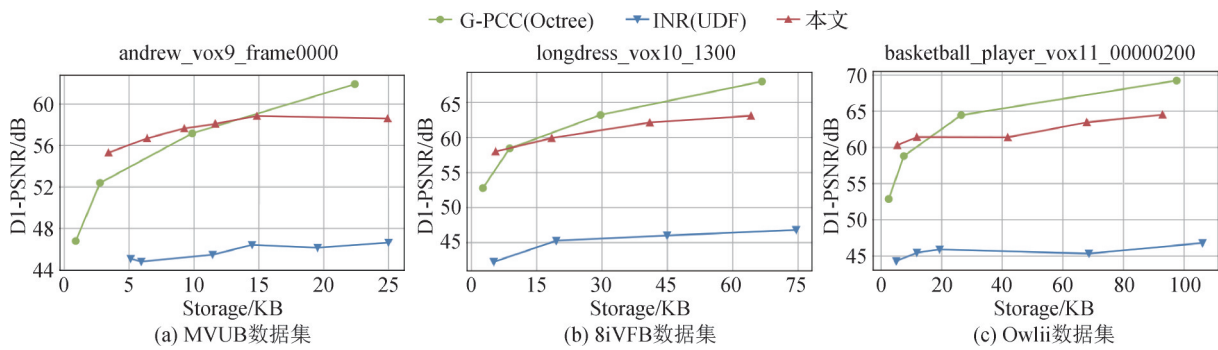


图9 在MPEG PCC数据集上的离散点云几何压缩性能曲线
Fig. 9 Result of discrete point cloud geometry compression on the MPEG PCC dataset ((a) MVUB; (b) 8iVFB; (c) OwlII)

量化点云表面的重建误差可视化结果如图10所示。与G-PCC相比,INR与本文的距离误差更加一致连续。这是由于隐式表征无需对原始点云做量化处理,没有量化误差。与基准方法INR相比,本文算法将误差距离降低了一个数量级(从0.01降至0.001),这表明本文算法有效提升了现有隐式表征

在离散点云压缩中的性能。

3.2.3 点云的连续表面重建实验

在ShapeNet数据集上,本文算法的表面重建性能如表2所示。与现有优化算法(NP、LevelSetUDF)相比,基于学习的算法GeoUDF取得了更高的NC性能,但L1-CD损失与F-score性能要次于现有优化方

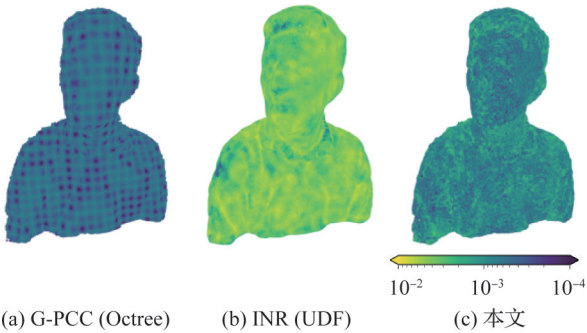


图 10 在“Andrew”(MVUB数据集)上的点云几何压缩误差可视化
Fig. 10 Visualization of point cloud geometry compression error on “Andrew” (MVUB)
(a) G-PCC (Octree); (b) INR (UDF); (c) ours

法。实验结果表明基于学习的算法虽然可以利用网络学习到的先验信息来提高表面与GT法向量的一致连续性,但实际的表面拟合精度要劣于优化方法。

表 2 在 ShapeNet 数据集上的表面重建性能
Table 2 Surface reconstruction results achieved on the ShapeNet dataset

方法	L1-CD ($\times 100$)	NC	F-score
P2S (Erler 等, 2020)	1.03	0.79	0.41
NDF (Chibane 等, 2020)	0.42	0.90	0.72
GeoUDF(Ren 等, 2023)	0.36	0.93	0.80
NP (Ma 等, 2021)	0.44	0.93	0.80
INR (Strümpler 等, 2022)	0.30	0.92	0.88
LevelSetUDF (Zhou 等, 2023)	0.34	0.92	0.84
本文	0.27	0.95	0.92

注:加粗字体表示各列最优结果。

与现有三维重建算法相比,本文采用的基准算法 INR 在 L1-CD、NC 与 F-score 指标上均更优。这表明在优化方法中,通过向网络添加正弦激活函数可以提高网络的非线性拟合能力,可以有效提升表面重建性能。与基准算法 INR 相比,本文算法的 L1-CD 损失下降 10. 0%, NC 和 F-score 指标分别提高了 3. 2% 与 4. 5%。这些实验结果表明本文算法进一步提高了 INR 在多个物体种类上的平均表面重建精度。

图 11 为 LevelSetUDF、INR 和本文算法的可视化

对比。与 LevelSetUDF 和本文相比,基于 SDF 的优化方法 INR 难以在“船”与“飞机”类别的扁平表面上取得较为精确的点云重建表面。这是由于 SDF 在扁平表面处具有剧烈高频分量,难以被网络拟合。与本文算法相比,基于 UDF 的优化方法 LevelSetUDF 在“椅子”、“汽车”的复杂表面上存在空洞。这是由于 LevelSetUDF 在全连接网络仅用了 ReLU (rectified linear unit) 与 sigmoid 激活层来提高网络的非线性,难以拟合高频表面信息。本文算法不仅可以生成更加平滑的表面,而且能够准确捕获高频表面信息。这是由于本文的 SLDF 表征没有内外表面之分,避免了在复杂表面上因符号频繁突变带来的剧烈高频分量。此外,本文在全连接网络及其分支结构中引入了正弦激活层,提高了网络的高频拟合特性,因此能够重建复杂的局部表面。

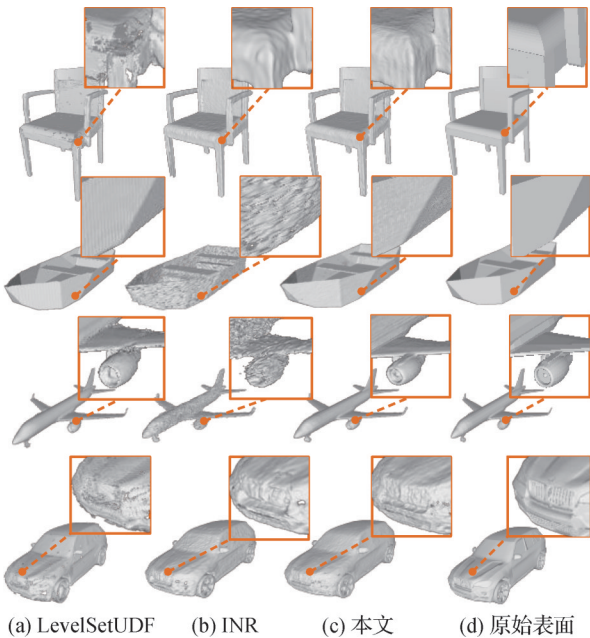


图 11 在 ShapeNet 数据集上的表面重建可视化结果
Fig. 11 Visualization of surface reconstruction on ShapeNet dataset((a) LevelSetUDF; (b) INR; (c) ours; (d) origin surfaces)

3. 2. 4 消融实验

为分别验证本文提出的 SLDF 表征、MINE 网络与 AMC 表面提取算法的有效性,设计了相应的消融实验,并将实验结果展示在表 3、图 12 和图 13 中。下面对消融实验结果进行详细的阐述与分析。

1)SLDF 表征的有效性。引入 SLDF 表征将连续点云表面压缩的 L1-CD 平均误差下降 18. 53%,将量

表3 表面重建消融实验

Table 3 Ablation study for surface reconstruction

方法	L1-CD($\times 100$)	NC	F-score
INR	0.310	0.924	0.881
w/o AMC	0.304	0.943	0.884
w/o SLDF	0.281	0.949	0.901
w/o MINE	0.273	0.945	0.914
本文	0.272	0.954	0.916

注:加粗字体表示各列最优结果,w/o表示未使用该模块。

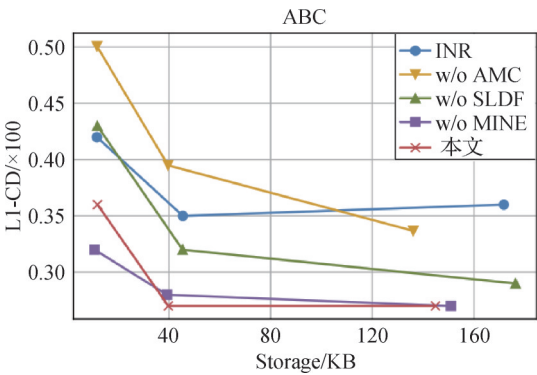


图12 连续点云压缩消融实验

Fig. 12 Ablation study for continuous point cloud compression

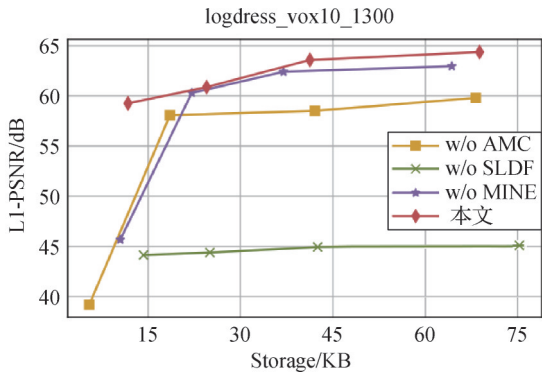


图13 量化点云压缩消融实验

Fig. 13 Ablation study for quantized point cloud compression

化点云表面压缩的D1-PSNR指标平均提升15 dB,使三维重建的L1-CD误差下降2.91%。实验结果表明SLDF表征有效提升了表面拟合精度。

2)MINE网络的有效性。从图12可以看出,随网络参数量的增大,MINE网络的内存效率增幅从2.5%增至3.3%,表明MINE网络可以有效改善现有模型的参数分布,提高了压缩效率。在重建精度上,引入MINE将三维重建的L1-CD误差下降0.36%,连续点云表面压缩的L1-CD误差平均下降3.72%,量化点云表面压缩的D1-PSNR指标平均提

升2.67 dB。这表明MINE同样有效提升了重建表面精度。为了进一步说明MINE网络有效提升网络非线性拟合能力,本文可视化了MINE网络在连续点云表面压缩的实验结果。如图14所示,MINE网络有利于网络学习到更加丰富的表面细节。

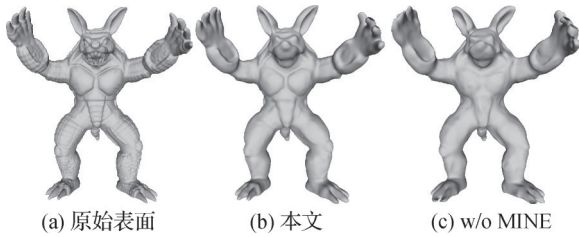


图14 在连续点云表面压缩的任务中运用
的MINE网络可视化效果

Fig. 14 Visualization of the MINE network on continuous point cloud surface compression ((a) original surface; (b) ours; (c) w/o MINE)

3)AMC表面提取算法的有效性。从图12与图13可以看出,AMC算法使连续点云表面压缩的L1-CD平均误差下降了26.28%,量化点云表面压缩的D1-PSNR指标平均提升4.8 dB,并使三维重建的L1-CD误差下降了10.53%。这表明AMC算法有效改善了MC算法直接作用于SLDF表征时,因厚度假设导致的精度损失。

4 结 论

针对现有点云几何压缩算法难以适用于连续点云表面、内存随点云的量化分辨率提高而剧烈增加的问题,本文提出了一种基于隐式表征的点云几何压缩算法。该算法由SLDF表征与MINE网络组成。SLDF表征首先将点云表面所在空间映射为连续的场函数,实现了任意分辨率与任意拓扑结构的点云表面的函数表述。然后,MINE网络利用分布式网络参数拟合SLDF表征,结合模型压缩实现了点云表面的高效编码,并提升了压缩效率。最后,AMC算法对Marching Cube算法生成的双层表面进行融合,从而有效地从解码后的网络中重建点云表面。

本文算法可用于任意点云表面的几何压缩,且不受分辨率影响。与基准算法INR相比,本文算法将连续点云表面压缩中L1-CD损失下降了10.0%,NC与F-score指标分别提升了3.2%与4.5%。此

外,本文算法将量化点云压缩的有效压缩内存下限降至5~10 KB,并在该内存区间范围的D1-PSNR指标维持在55 dB以上。实验结果表明,本文算法不仅拓展了点云几何压缩的应用范围,增加了点云表面的有效压缩上限,而且提升了同等压缩强度下的表面重建性能。

由于本文算法在点云压缩过程中的编码时间较长,拟在未来探索结合先验信息的权重初始化方法,进一步提升压缩效率,降低编码时间。此外,将进一步研究基于隐式表征的点云几何与属性信息的混合压缩算法。

参考文献(References)

- Cao C, Preda M and Zaharia T. 2019. 3D point cloud compression: a survey//Proceedings of the 24th International Conference on 3D Web Technology. Los Angeles, USA: ACM: 1-9 [DOI: 10.1145/3329714.3338130]
- Chang A X, Funkhouser T, Guibas L, Hanrahan P, Huang Q X, Li Z M, Savarese S, Savva M, Song S R, Su H, Xiao J X, Yi L and Yu F. 2015. ShapeNet: an information-rich 3D model repository [EB/OL]. [2024-01-15]. <http://arxiv.org/pdf/1512.03012.pdf>
- Chen H, He B, Wang H Y, Ren Y X, Lim S N and Shrivastava A. 2021. NeRV: neural representations for videos//Proceedings of the 35th International Conference on Neural Information Processing Systems. Online: Curran Associates Inc.: 21557-21568
- Chibane J, Mir A and Pons-Moll G. 2020. Neural unsigned distance fields for implicit function learning//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: 21638-21652
- 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 WG11M 40059/WG1M74006. Geneva, Switzerland: MPEG
- Dupont E, Golinski A, Alizadeh M, Teh Y W and Doucet A. 2021. COIN: compression with implicit neural representations [EB/OL]. [2024-01-15]. <http://arxiv.org/pdf/2103.03123.pdf>
- Erler P, Guerrero P, Ohrhallinger S, Mitra N J and Wimmer M. 2020. Points2Surf learning implicit surfaces from point clouds//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer: 108-124 [DOI: 10.1007/978-3-030-58558-7_7]
- Fan H Q, Su H and Guibas L. 2017. A point set generation network for 3D object reconstruction from a single image//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 2463-2471 [DOI: 10.1109/CVPR.2017.264]
- Fang G C, Hu Q Y, Wang H Y, Xu Y L and Guo Y L. 2022. 3DAC: learning attribute compression for point clouds//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 14799-14808 [DOI: 10.1109/CVPR52688.2022.01440]
- Fathony R, Sahu A K, Willmott D and Kolter J Z. 2021. Multiplicative filter networks//Proceedings of the 9th International Conference on Learning Representations. Virtual Event, Austria: OpenReview.net
- Han S, Mao H Z and Dally W J. 2016. Deep compression: compressing deep neural network with pruning, trained quantization and Huffman coding//Proceedings of the 4th International Conference on Learning Representations. San Juan, USA: OpenReview.net
- Hart J C. 1996. Sphere tracing: a geometric method for the antialiased ray tracing of implicit surfaces. The Visual Computer, 12 (10): 527-545 [DOI: 10.1007/s003710050084]
- Hinton G, Vinyals O and Dean J. 2015. Distilling the knowledge in a neural network [EB/OL]. [2024-01-15]. <http://arxiv.org/pdf/1503.02531.pdf>
- Huang L L, Wang S L, Wong K, Liu J and Urtasun R. 2020. OctSqueeze: octree-structured entropy model for LiDAR compression//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 1310-1320 [DOI: 10.1109/CVPR42600.2020.00139]
- Jang E S, Preda M, Mammou K, Tourapis A M, Kim J, Graziosi D B, Rhyu S and Budagavi M. 2019. Video-based point-cloud-compression standard in MPEG: from evidence collection to committee draft [Standards in a Nutshell]. IEEE Signal Processing Magazine, 36(3): 118-123 [DOI: 10.1109/MSP.2019.2900721]
- Koch S, Matveev A, Jiang Z S, Williams F, Artemov A, Burnaev E, Alexa M, Zorin D and Panozzo D. 2019. ABC: a big CAD model dataset for geometric deep learning//Proceedings of 2019 IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 9593-9603 [DOI: 10.1109/CVPR.2019.00983]
- Lindell D B, Van Veen D, Park J J and Wetzstein G. 2022. Bacon: band-limited coordinate networks for multiscale scene representation//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 16231-16241 [DOI: 10.1109/CVPR52688.2022.01577]
- Long X X, Cheng X J, Zhu H, Zhang P J, Liu H M, Li J, Zheng L T, Hu Q Y, Liu H, Cao X, Yang R G, Wu Y H, Zhang G F, Liu Y B, Xu K, Guo Y L and Chen B Q. 2021. Recent progress in 3D vision. Journal of Image and Graphics, 26(6): 1389-1428 (龙霄潇, 程新景, 朱昊, 张朋举, 刘浩敏, 李俊, 郑林涛, 胡庆拥, 刘浩, 曹汛, 杨睿刚, 吴毅红, 章国锋, 刘焯斌, 徐凯, 郭裕兰, 陈宝权. 2021. 三维视觉前沿进展. 中国图象图形学报, 26(6): 1389-1428) [DOI: 10.11834/jig.210043]
- Long X X, Lin C, Liu L J, Liu Y, Wang P, Theobalt C, Komura T and Wang W P. 2023. NeuralUDF: learning unsigned distance fields for multi-view reconstruction of surfaces with arbitrary topologies//

- Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 20834-20843 [DOI: 10.1109/CVPR52729.2023.01996]
- 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/M72012.Geneva, Switzerland: MEPEG
- Lorensen W E and Cline H E. 1987. Marching cubes: a high resolution 3D surface construction algorithm//Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques [s. l.]: ACM: 163-169 [DOI: 10.1145/37401.37422]
- Ma B, Han Z, Liu Y S and Zwicker M. 2021. Neural-pull: learning signed distance functions from point clouds by learning to pull space onto surfaces//Proceedings of the 38th International Conference on Machine Learning. Online: PMLR: 139: 7246-7257 [DOI: 10.48550/arXiv.2011.13495]
- Maglo A, Lavoué G, Dupont F and Hudelot C. 2015. 3D mesh compression: survey, comparisons, and emerging trends. ACM Computing Surveys, 47(3): #44 [DOI: 10.1145/2693443]
- Mamou K, Zaharia T and Prêteux F. 2009. A triangle-fan-based approach for low complexity 3D mesh compression//Proceedings of the 16th IEEE International Conference on Image Processing. Cairo, Egypt: IEEE: 3513-3516 [DOI: 10.1109/ICIP. 2009. 5414075]
- Mescheder L, Oechsle M, Niemeyer M, Nowozin S and Geiger A. 2019. Occupancy networks: learning 3D reconstruction in function space//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 4455-4465 [DOI: 10.1109/CVPR.2019.00459]
- Mildenhall B, Srinivasan P P, Tancik M, Barron J T, Ramamoorthi R and Ng R. 2020. NeRF: representing scenes as neural radiance fields for view synthesis//Proceedings of the 16th European Conference on Computer Vision. Online: Springer: 405-421 [DOI: 10.1007/978-3-030-58452-8_24]
- Park J J, Florence P, Straub J, Newcombe R and Lovegrove S. 2019. DeepSDF: learning continuous signed distance functions for shape representation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 165-174 [DOI: 10.1109/CVPR.2019.00025]
- Quach M, Valenzise G and Dufaux F. 2020. Improved deep point cloud geometry compression//Proceedings of 2020 IEEE 22nd International Workshop on Multimedia Signal Processing. Tampere, Finland: IEEE: 1-6 [DOI: 10.1109/MMSP48831.2020.9287077]
- Que Z Z, Lu G and Xu D. 2021. VoxelContext-Net: an octree based framework for point cloud compression//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 6038-6047 [DOI: 10.1109/CVPR46437. 2021.00598]
- Ren S Y, Hou J H, Chen X D, He Y and Wang W P. 2023. GeoUDF: surface reconstruction from 3D point clouds via geometry-guided distance representation//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 14168-14178 [DOI: 10.1109/ICCV51070.2023.01307]
- Schwarz S, Preda M, Baroncini V, Budagavi M, Cesar P, Chou P A, Cohen R A, Krivokuća M, Lasserre S, Li Z, Llach J, Mammou K, Mekuria R, Nakagami O, Siahaan E, Tabatabai A, Tourapis A M and Zakharchenko V. 2019. 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]
- Sitzmann V, Martel J N P, Bergman A W, Lindell D B and Wetzstein G. 2020. Implicit neural representations with periodic activation functions//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: 7462-7473 [DOI: 10.48550/arXiv.2006.09661]
- Strömpler Y, Postels J, Yang R, van Gool L and Tombari F. 2022. Implicit neural representations for image compression//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 74-91 [DOI: 10.1007/978-3-031-19809-0_5]
- Sullivan G J, Ohm J R, Han W J and Wiegand T. 2012. Overview of the high efficiency video coding (HEVC) standard. IEEE Transactions on Circuits and Systems for Video Technology, 22(12): 1649-1668 [DOI: 10.1109/TCSVT.2012.2221191]
- Tatarchenko M, Richter S R, Ranftl R, Li Z W, Koltun V and Brox T. 2019. What do single-view 3D reconstruction networks learn?//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3400-3409 [DOI: 10.1109/CVPR.2019.00352]
- Turk G and Levoy M. 1994. Zippered polygon meshes from range images//Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM: 311-318 [DOI: 10.1145/192161.192241]
- Wang B, Yu Z D, Yang B, Qin J, Breckon T, Shao L, Trigoni N and Markham A. 2022. RangeUDF: semantic surface reconstruction from 3D point clouds [EB/OL]. [2024-01-15]. <http://arxiv.org/pdf/2204.09138.pdf>
- Wang J Q, Ding D D, Li Z and Ma Z. 2021. Multiscale point cloud geometry compression//Proceedings of 2021 Data Compression Conference. Snowbird, USA: IEEE: 73-82 [DOI: 10.1109/DCC50243.2021.00015]
- Xu Y, Lu Y and Wen Z. 2017. OwlII dynamic human mesh sequence dataset. ISO/IEC JTC1/SC29/WG11 m41658. Macau, China: MPEG
- Yang Y Q, Feng C, Shen Y R and Tian D. 2018. FoldingNet: point cloud auto-encoder via deep grid deformation//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 206-215 [DOI: 10.1109/CVPR. 2018.00029]

- Ye J L, Chen Y T, Wang N Y and Wang X L. 2022. GIFS: neural implicit function for general shape representation//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 12819-12829 [DOI: 10.1109/CVPR52688.2022.01249]
- Yu P P, Zuo D, Huang Y E, Huang R S, Wang H Y, Guo Y L and Liang F. 2023. Sparse representation based deep residual geometry compression network for large-scale point clouds//Proceedings of 2023 IEEE International Conference on Multimedia and Expo. Brisbane, Australia: IEEE: 2555-2560 [DOI: 10.1109/ICME55011.2023.00435]
- Zhou J S, Ma B R, Li S J, Liu Y S and Han Z Z. 2023. Learning a more continuous zero level set in unsigned distance fields through level set projection//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 3158-3169 [DOI: 10.1109/ICCV51070.2023.00295]
- Zhu W, Zhang Y H, Ying Y, Zheng Y Y and He D F. 2023. A dense residual structure and multi-scale pruning-relevant point cloud compression network. Journal of Image and Graphics, 28(7): 2105-

2119 (朱威, 张雨航, 应悦, 郑雅羽, 何德峰. 2023. 结合密集残差结构和多尺度剪枝的点云压缩网络. 中国图象图形学报, 28(7): 2105-2119) [DOI: 10.11834/jig.220047]

作者简介

陈佳慧,女,硕士研究生,主要研究方向为点云几何压缩。

E-mail:chenjh283@mail2.sysu.edu.cn

黄小红,通信作者,男,副教授,主要研究方向为图像智能解译。E-mail:huangxh75@sysu.edu.cn

方广驰,男,硕士研究生,主要研究方向为点云属性压缩。

E-mail:fanggch@mail2.sysu.edu.cn

李浩然,男,博士研究生,主要研究方向为点云几何压缩。

E-mail:lihr67@mail2.sysu.edu.cn

张晔,男,博士后,主要研究方向为人机交互。

E-mail:zhangy2658@sysu.edu.cn

郭裕兰,男,副教授,主要研究方向为三维视觉与模式识别。

E-mail:guoyulan@sysu.edu.cn