arXiv:2304.00335v1 [eess.SP] 1 Apr 2023

采用具有几何注意力的前馈网络进行三维点云体属性压缩

Tam Thuc Do[⋆,†]、*Philip A. Chou*[⋆]、*Gene Cheung*[†]

[⋆]Google Research，西雅图，华盛顿州，美国
[†]约克大学，多伦多，加拿大

摘要

我们研究采用体积式方法进行三维点云属性压缩：给定目标体属性函数 f：$\mathbb{R}^3 \to \mathbb{R}$，我们在编码器处对表征 f 的参数向量 θ 进行量化和编码，以便在解码器处根据已知三维点 $\mathbf{x}$' 重建 $f_\theta(\mathbf{x})$。在先前工作区域自适应分层变换（RAHT）的基础上，该方法采用分段常数函数来构成函数空间嵌套序列；我们提出一种前馈线性网络，无需进行特征分解即可实现张成函数空间的高阶B样条基。前馈网络架构使系统适用于端到端神经网络学习。我们网络的关键是空间变化卷积，类似于图算子，其权重根据已知三维几何计算，以实现归一化。我们证明，编码器归一化中的层数等于矩阵逆泰勒级数中的项数。在真实世界三维点云上的实验结果表明，与RAHT相比，能量集中效果提高了2–3 dB，码率降低了20–30%。

索引词— 三维点云压缩、深度学习

1. 引言

我们对三维点云执行属性压缩。三维点云是一组点 $\{(\mathbf{x}_i, \mathbf{y}_i)\}$，其位置为 $\mathbf{x}_i \in \mathbb{R}^3$，相应的属性为 $\mathbf{y}_i \in \mathbb{R}^r$。点云几何 和属性 压缩分别涉及点位置和属性的有损或无损压缩。尽管几何和属性压缩可以联合进行（例如，[1]），但近期研究中的主流方法 [2]–[8] 以及 MPEG 基于几何的点云压缩（G-PCC）标准[9]–[12] 均是先压缩几何，再根据几何对属性进行压缩。[1] 因此，即使几何是有损的，属性压缩问题也可以归结为属性编码器和属性解码器均已知解码后的几何的情况。本文采用这一假设。

属性压缩有许多种方法。我们采用一种体积式方法。如果 f：$\mathbb{R}^d \to \mathbb{R}^r$ 满足 $d = 3$（或者满足 $d > 3$），则称实值向量函数为体积式（或者称为超体积式）。体积函数的例子包括吸收密度、到最近表面的有符号距离、辐射场和光场。[2] 在

本研究部分由 Google 学生研究项目资助。Gene Cheung 感谢 NSERC 拨款 RGPIN-2019-06271、RGPAS-2019-00110 提供的额外支持。[1]如果几何压缩是有损的，那么在压缩之前，必须先将原始属性转移到重建后的几何上。[2]辐射度或光场不仅取决于位置 $\mathbf{x} \in \mathbb{R}^3$，还取决于单位球面上的方向 (θ, ϕ)，可用关于 (x, y, z, θ, ϕ) 的超体积函数表示，该函数可生成一个属性向量



图 1：RAHT($p = 2$) 在 0.22 bpp 下（右）以及 RAHT($p = 1$) 在 0.20 bpp 下（左）的重建点云

点云属性压缩的体积式方法中，将函数 f_θ 拟合到点云位置 $\mathbf{x}_i$（假定已知）处的属性 $\mathbf{y}_i$。随后，通过压缩并传输量化后的 $\hat{\theta}$，对 f_θ 进行压缩和传输。最后，在解码器处重建体积函数 $f_{\hat{\theta}}$，并在点 $\mathbf{x}_i$ 处查询，以获得重建属性 $\hat{\mathbf{y}}_i = f_{\hat{\theta}}(\mathbf{x}_i)$。例如，也可以在任意点 $\mathbf{x} \in \mathbb{R}^3$ 处查询，从而实现无限空间缩放。

文献中较少采用体积式方法进行点云属性压缩。然而，[7] 指出，体积式属性压缩是 MPEG G-PCC 的核心，该标准采用区域自适应层次变换（RAHT）[4, 13]。RAHT 隐式地将任何在位置 $\mathbf{x}_i$ 处插值属性 $\mathbf{y}_i$ 的体积函数 f 投影到一个函数空间嵌套序列 $\mathcal{F}_0^{(1)} \subseteq \cdots \subseteq \mathcal{F}_L^{(1)}$ 上，该序列对应不同的细节层级 $\ell = 0, \ldots, L$。具体而言，$\mathcal{F}_\ell^{(1)}$ 是由宽度为 $2^{L-\ell}$ 个体素的立方体上的分段常函数构成的空间，其中 2^L 是包含该点云的立方体的体素边长。也就是说，$\mathcal{F}_\ell^{(1)}$ 是尺度为 $2^{L-\ell}$的一阶（3D）B样条函数空间。此外，[7] 展示了如何改进 RAHT：改用 $\mathcal{F}_\ell^{(p)}$ 这一$p > 1$阶 B 样条空间。本文将其称为 p阶 RAHT，或 RAHT(p)。然而，对于 $p > 1$，将 f 投影到 $\mathcal{F}_\ell^{(p)}$ 需要在 [7] 中计算大型特征分解。

与此同时，继学习型神经网络在图像压缩方面取得成功之后 [14]–[23], ，许多研究利用学习型神经网络，显著改进了 MPEG G-PCC 中的点云几何压缩 [24]–[30]。也有少数研究尝试将学习型神经网络用于点云属性压缩 [1][31]–[33]。还有一些研究更广泛地将神经网络用于体积属性压缩 [34]–[36]。然而，在点云属性压缩任务上，学习型神经网络方法尚未胜过 MPEG G-PCC。原因

$\mathbf{a}$ 或一个体积函数，其输入为 (x, y, z)，输出为球面上属性值函数的参数向量（例如球谐函数）$\mathbf{a}(\theta, \phi)$。

可能在于点云具有稀疏、不规则且不断变化的特性，这带来了根本性的困难。神经网络用于图像压缩以及点云几何压缩时，通常采用稠密、平移不变的卷积；当属性仅在特定位置已知时，这些卷积就会变成稀疏卷积。稀疏卷积（将缺失点视为属性值为零的卷积）应当进行归一化，但迄今为止，如何进行归一化尚不明确。

本文通过展示如何将 RAHT(p) 实现为一种无需特征分解的前馈线性网络，向学习用于点云属性压缩的网络迈出了一步。我们前馈网络的骨干卷积（类似于 UNet [37] 中的下采样卷积和上采样转置卷积）采用具有空间不变卷积核的稀疏卷积。前馈网络的“肋条”（类似于 UNet 中的跳跃连接）通过空间变化卷积进行归一化。空间变化卷积可以被视为具有广义图拉普拉斯算子 [38] 的图算子。我们表明，该算子的权重本身可以通过使用具有空间不变卷积核的稀疏卷积，对点上的权重进行计算而获得。这使得属性变换能够以几何为条件，并利用几何信息对稀疏卷积的结果进行归一化。该归一化过程可以解释为一种注意力机制，对应于 RAHT 中的“区域自适应”。我们表明，编码器中归一化操作的层数等价于矩阵逆的泰勒级数项数，或者等价于展开梯度下降优化以计算解码所需系数时的步数。在实验结果部分，我们研究了前馈网络的能量集中和比特率，并将其与 RAHT 进行了比较，结果显示能量集中最高提升 2–3 dB，比特率降低 20–30%。尽管目前我们的前馈网络中既没有非线性，也没有学习过程，但我们相信，这是理解和训练用于点云属性压缩的前馈神经网络的重要第一步。

2. 编码框架

假设属性为标量。扩展到向量属性也很直接。用 $f\colon \mathbb{R}^3 \to \mathbb{R}$ 表示一个在 $\mathbb{R}^3$上定义属性场的体积函数。用 $\{(\mathbf{x}_i, y_i)\}$ 表示一个带有属性 $y_i = f(\mathbf{x}_i)$的点云，其中 $\mathbf{x}_i \in \mathbb{R}^3$ 是点 i的三维坐标。给定位置对属性进行编码时，RAHT(p)将 f投影到函数空间嵌套序列 $\mathcal{F}_0^{(p)} \subseteq \cdots \subseteq \mathcal{F}_L^{(p)}$ 上。每个函数空间 $\mathcal{F}_\ell^{(p)}$ 都是由一组基函数 $\{\phi_{\ell,\mathbf{n}}^{(p)} : \mathbf{n} \in \mathcal{N}_\ell\}$张成的参数函数族，即体积 B 样条基函数 $\phi_\ell^{(p)}$,n(x) $= \phi_{\mathbf{0}}^{(p)}$, (2$^\ell$x $-$n)，其阶数为 p、尺度为 ℓ，偏移量为 $\mathbf{n} \in \mathcal{N}_\ell \subset \mathbb{Z}^3$。具体而言，对于 $p = 1$, $\phi_{\mathbf{0}}^{(p)}$, (x) 是各分量 x 上“箱形”函数 $1_{[0,1]}(x)$的乘积；而对于 $p = 2$, 它是各分量 x 上“帽形”函数 $1_{[-1,1]}(x)(1 - |x|)$的乘积。因此，RAHT(1) 将 f 投影到函数空间 $\mathcal{F}_\ell^{(1)}$，该空间中的函数在宽度为 $2^{-\ell}$的块上分段常数；而 RAHT(2) 将 f 投影到函数空间 $\mathcal{F}_\ell^{(2)}$，该空间中的连续函数在宽度为 $2^{-\ell}$的块上分段三线性。无论哪种情况，显然 $f_\ell \in \mathcal{F}_\ell$ 蕴含 $f_\ell \in \mathcal{F}_{\ell+1}$。因此，这些空间是嵌套的。下文中，为简化记法，我们有时省略阶数（p）。

任何 $f_\ell \in \mathcal{F}_\ell$ 都可以表示为基函数的线性组合$\phi_{\ell,\ \mathbf{n}}$:

$$f_\ell(\mathbf{x}) = \sum_{\mathbf{n}\in\mathcal{N}_\ell} F_{\ell,\mathbf{n}}\phi_{\ell,\mathbf{n}}(\mathbf{x}), \tag{1}$$

也可以写作 $f_\ell = \mathbf{\Phi}_\ell F_\ell$，其中 $\mathbf{\Phi}_\ell = [\phi_{\ell,\mathbf{n}}]$ 是基函数的行向量$\phi_{\ell,\ \mathbf{n}}$，而 $F_\ell = [F_{\ell,\mathbf{n}}]$ 是系数的列向量$F_{\ell,\ \mathbf{n}}$，每个向量的长度均为 $N_\ell = |\mathcal{N}_\ell|$。

RAHT 将 f投影到希尔伯特空间中（等价类的）函数空间里的 $\mathcal{F}_\ell$ 上，该空间的内积为 $\langle f, g\rangle = \sum_i f(\mathbf{x}_i)g(\mathbf{x}_i)$。因此，将 f 投影到 $\mathcal{F}_\ell$上的投影 f_ℓ^*（记作 $f_\ell^* = f\circ\mathcal{F}_\ell$）使平方误差 $||f - f_\ell||^2 = \langle f - f_\ell, f - f_\ell\rangle = \sum_i (y_i - f_\ell(\mathbf{x}_i))^2$,最小，这正是属性编码所关注的平方误差。此外，表示基 $\mathbf{\Phi}_\ell$ 下函数 $f_\ell^* = f\circ\mathcal{F}_\ell$ 的系数 F_ℓ^*（称为低通系数）可通过正规方程计算:

$$F_\ell^* = (\mathbf{\Phi}_\ell^\top\mathbf{\Phi}_\ell)^{-1}\mathbf{\Phi}_\ell^\top f. \tag{2}$$

这里，$\mathbf{\Phi}_\ell^\top f$表示由内积 $[\langle\phi_{\ell,\mathbf{n}}, f\rangle$]组成的长度为$N_\ell$的向量，而 $\mathbf{\Phi}_\ell^\top\mathbf{\Phi}_\ell$ 表示内积 $[\langle\phi_{\ell,\mathbf{n}}, \phi_{\ell,\mathbf{m}}\rangle]$的 $N_\ell \times N_\ell$ 格拉姆矩阵 。此外，由于所有投影 f_ℓ^*都满足勾股定理 [39, 第 3 节3],，可以证明 [7] 对于所有 $\ell' \geq \ell$，都有$f_\ell^* = f_{\ell'}^* \circ \mathcal{F}_\ell$，且所有残差 函数

$$g_\ell^* = f_{\ell+1}^* - f_\ell^* \tag{3}$$

都属于子空间 $\mathcal{G}_\ell$，该子空间是$\mathcal{F}_{\ell+1}$中与 $\mathcal{F}_\ell$ 正交的补空间。因此，

$$\begin{aligned} \mathcal{F}_L &= \mathcal{F}_0 \oplus \mathcal{G}_0 \oplus \cdots \oplus \mathcal{G}_\ell \oplus \cdots \mathcal{G}_{L-1} \quad \text{and} & (4)\\ f_L^* &= f_0^* + g_0^* + \cdots + g_\ell^* + \cdots + g_{L-1}^*. & (5) \end{aligned}$$

在我们的编码框架中，我们假设细节层次 L 足够高，使得 $f \in \mathcal{F}_L$。因此，我们取 $f_L^* = f$，并通过编码 f_0^*, g_0^*、.....、 g_{L-1}^*来编码 f。为了编码这些函数，我们将它们表示为某个基下的系数，然后进行量化和熵编码。接下来，我们将说明如何高效地计算这些系数。

2.1. 低通系数

假设存在足够精细的分辨率 L，使得第 i 个基函数 $\phi_{L,\ \mathbf{n}_i}$ 在 $\mathbf{x}_i$ 上为 1，并在 $\mathbf{x}j$ 上为 0，其中 $j \neq i$，即 $\phi_{L,\ \mathbf{n}_i}(\mathbf{x}j) = \delta_{ij}$。于是，我们有 $\mathbf{\Phi}_L^\top\mathbf{\Phi}_L = \mathbf{I}$。接下来，记 $\mathbf{A}_\ell = [a_{ij}^l]$ 为一个 $N_\ell \times N_{\ell+1}$ 矩阵，其第 i 行表示 i 个基函数 $\phi_{\ell,\ \mathbf{n}_i}$ 如何用 $\mathcal{F}_\ell$ 的基函数 $\phi_{\ell+1,\ \mathbf{m}_j}$ 表示，即

$$\phi_{\ell,\mathbf{n}_i} = \sum_j a_{ij}^\ell\phi_{\ell+1,\mathbf{m}_j}. \tag{6}$$

注意，对于$\mathbf{d}_{ij} = \mathbf{m}_j - 2\mathbf{n}_i$,

$$a_{ij}^\ell = \begin{cases} 1 & \text{if } \mathbf{d}_{ij} \in \{0,1\}^3 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

对于 $p = 1$ 和

$$a_{ij}^\ell = \begin{cases} 2^{-\|\mathbf{d}_{ij}\|_1} & \text{if } \mathbf{d}_{ij} \in \{-1,0,1\}^3 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

对于 $p = 2$。用向量形式表示为

$$\mathbf{\Phi}_\ell = \mathbf{\Phi}_{\ell+1}\mathbf{A}_\ell^\top, \tag{9}$$

其中，$\mathbf{A}_\ell$ 是一个普通（即空间不变）但稀疏的卷积矩阵，具有 2× 空间下采样；而$\mathbf{A}_\ell^\top$是对应的转置卷积矩阵，具有 2×上采样。利用这些矩阵，我们可以计算从第 $\ell+1$ 层到第 ℓ层的基函数内积矩阵:

$$\begin{aligned} \mathbf{\Phi}_\ell^\top\mathbf{\Phi}_\ell &= \mathbf{A}_\ell\mathbf{\Phi}_{\ell+1}^\top\mathbf{\Phi}_{\ell+1}\mathbf{A}_\ell^\top & (10)\\ \mathbf{\Phi}_{\ell+1}^\top\mathbf{\Phi}_\ell &= \mathbf{\Phi}_{\ell+1}^\top\mathbf{\Phi}_{\ell+1}\mathbf{A}_\ell^\top & (11)\\ \mathbf{\Phi}_\ell^\top\mathbf{\Phi}_{\ell+1} &= \mathbf{A}_\ell\mathbf{\Phi}_{\ell+1}^\top\mathbf{\Phi}_{\ell+1}. & (12) \end{aligned}$$

从这些表达式可以看出，矩阵 $\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell$也是一个稀疏（三对角）卷积矩阵，不过它是空间变化的，对应于一个具有 27 邻域连接（包括自环）的图的边权重。

因此，从最高细节层次 L 开始，其中 $\phi_{L,\mathbf{n}_i,j}(\mathbf{x}) = \delta_{ij}$ 和 $\mathbf{I}$，我们有，因而 $\mathbf{\Phi}_L^\top \mathbf{\Phi}_L = \mathbf{I}$（根据 $\mathbf{\Phi}_L^\top f = [y_i]$）。转到 $\ell < L$，我们有

$$F_\ell^* = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-1} \mathbf{\Phi}_\ell^\top f_{\ell+1}^* \tag{13}$$

$$= (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-1} \mathbf{\Phi}_\ell^\top \mathbf{\Phi}_{l+1} F_{\ell+1}^* \tag{14}$$

$$= (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-1} \mathbf{A}_\ell (\mathbf{\Phi}_{\ell+1}^\top \mathbf{\Phi}_{\ell+1}) F_{\ell+1}^*. \tag{15}$$

令 $\tilde{F}_\ell^* = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell) F_\ell^*$，我们称之为未归一化低通系数，则可以轻松计算所有层级 ℓ 的 F_ℓ^*：

$$\tilde{F}_\ell^* = \mathbf{A}_\ell \tilde{F}_{\ell+1}^*. \tag{16}$$

这是一个普通的稀疏卷积。最后，可以计算 F_ℓ^*：

$$F_\ell^* = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-1} \tilde{F}_\ell^*. \tag{17}$$

稍后我们将介绍如何使用前馈网络计算这一结果。

2.2. 高通系数

为了对残差函数 g_0^*、……、g_{L-1}^*进行编码，必须在一个基下表示它们。由于 $g_\ell^* \in \mathcal{G}_\ell \subset \mathcal{F}_{\ell+1}$，表示 g_ℓ^*的一种可能基是针对 $\mathcal{F}_{\ell+1}$的基 $\mathbf{\Phi}_{\ell+1}$。在此基下，$g_\ell^* = f_{\ell+1}^* - f_\ell^* = \mathbf{\Phi}_{\ell+1} F_{\ell+1}^* - \mathbf{\Phi}_\ell F_\ell^*$ 的系数由正规方程（2）和表达式（11）给出：

$$(\mathbf{\Phi}_{\ell+1}^\top \mathbf{\Phi}_{\ell+1})^{-1} \mathbf{\Phi}_{\ell+1}^\top g_\ell^* = F_{\ell+1}^* - \mathbf{A}_\ell^\top F_\ell^* \triangleq \delta F_{\ell+1}^*. \tag{18}$$

然而，这种表示中的系数数量为 $N_{\ell+1}$，而 $\mathcal{G}_\ell$ 的维数仅为 $N_{\ell+1} - N_\ell$。因此，使用这种表示对 g_ℓ^* 进行编码称为**过完备（或非临界）残差编码**。

或者，g_ℓ^* 可以表示为 $\mathcal{G}_\ell$的基 $\mathbf{\Psi}_\ell$ 中的线性组合。用 G_ℓ^* 表示与长度均为 $N_{\ell+1} - N_\ell$的基函数行向量 $\mathbf{\Psi}_\ell$相对应的系数列向量，使得$g_\ell^* = \mathbf{\Psi}_\ell G_\ell^*$。系数 G_ℓ^*——此处称为高通系数——可通过正规方程计算得到

$$G_\ell^* = (\mathbf{\Psi}_\ell^\top \mathbf{\Psi}_\ell)^{-1} \mathbf{\Psi}_\ell^\top g_\ell^*. \tag{19}$$

尽管 $\mathbf{\Psi}_\ell$ 的基 $\mathcal{G}_\ell$有许多种选择，但我们知道$\mathcal{G}_\ell$ 与 $\mathcal{F}_\ell$ 正交，因此 $\mathbf{\Phi}_\ell^\top \mathbf{\Psi}_\ell = \mathbf{0}$。此外，由于 $\mathcal{G}_\ell \subset \mathcal{F}_{\ell+1}$，$\mathbf{\Psi}_\ell$ 可以写成基函数 $\mathbf{\Phi}_{\ell+1}$的线性组合：

$$\mathbf{\Psi}_\ell = \mathbf{\Phi}_{\ell+1} \mathbf{Z}_\ell^\top. \tag{20}$$

因此，$\mathbf{0} = \mathbf{\Phi}_\ell^\top \mathbf{\Psi}_\ell = \mathbf{A}_\ell \mathbf{\Phi}_{\ell+1}^\top \mathbf{\Phi}_{\ell+1} \mathbf{Z}_\ell^\top = \mathbf{A}_\ell \tilde{\mathbf{Z}}_\ell^\top$，其中 $\mathbf{Z}_\ell^\top = (\mathbf{\Phi}_{\ell+1}^\top \mathbf{\Phi}_{\ell+1})^{-1} \tilde{\mathbf{Z}}_\ell^\top$。因此，$\tilde{\mathbf{Z}}_\ell^\top$ 的列张成了$\mathbf{A}_\ell$ 的零空间。一种可能的选择是

$$\tilde{\mathbf{Z}}_\ell^\top = \begin{bmatrix} -(\mathbf{A}_\ell^a)^{-1} \mathbf{A}_\ell^b \\ \mathbf{I}^b \end{bmatrix} \tag{21}$$

其中，我们将$\mathbf{A}_\ell$ 拆分为两个子矩阵，大小分别为 $N_\ell \times N_\ell$ 和 $N_\ell \times (N_{\ell+1} - N_\ell)$：

$$\mathbf{A}_\ell = \begin{bmatrix} \mathbf{A}_\ell^a & \mathbf{A}_\ell^b \end{bmatrix}. \tag{22}$$

使用（19）中的表示 G_ℓ^* 对 g_ℓ^* 进行编码，称为**正交归一残差编码（或临界残差编码）**。第4节比较了正交归一残差编码和过完备残差编码的压缩性能。

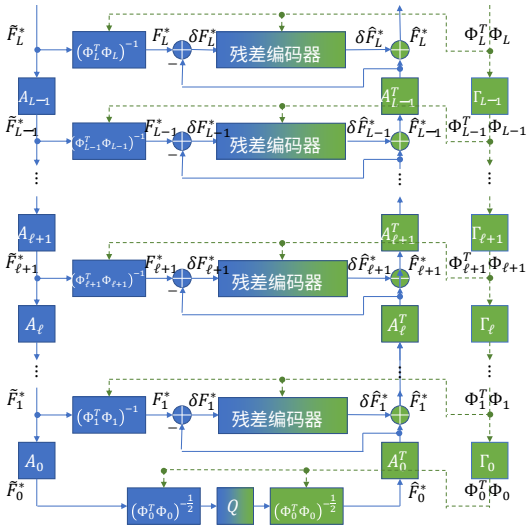


图 2：实现点云属性编码器（蓝色和绿色部分）及解码器（绿色部分）的多层前馈网络。

2.3. 正交归一化

直接对系数 F_ℓ^*和 G_ℓ^*进行标量量化无法获得良好的率失真性能，因为它们的基函数 $\mathbf{\Phi}_\ell$ 和 $\mathbf{\Psi}_\ell$虽然彼此正交（即 $\mathbf{\Phi}_\ell^\top \mathbf{\Psi}_\ell = \mathbf{0}$），但各自并非正交归一（即 $\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell \neq \mathbf{I}$，且 $\mathbf{\Psi}_\ell^\top \mathbf{\Psi}_\ell \neq \mathbf{I}$）。因此，我们将它们正交归一化如下：

$$\bar{\mathbf{\Phi}}_\ell = \mathbf{\Phi}_\ell \mathbf{R}_\ell = \mathbf{\Phi}_\ell (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-\frac{1}{2}} \tag{23}$$

$$\bar{\mathbf{\Psi}}_\ell = \mathbf{\Psi}_\ell \mathbf{S}_\ell = \mathbf{\Psi}_\ell (\mathbf{\Psi}_\ell^\top \mathbf{\Psi}_\ell)^{-\frac{1}{2}}. \tag{24}$$

由此，$\bar{\mathbf{\Phi}}_\ell^\top \bar{\mathbf{\Phi}}_\ell = \mathbf{I}$和 $\bar{\mathbf{\Psi}}_\ell^\top \bar{\mathbf{\Psi}}_\ell = \mathbf{I}$。现在可以计算与正交归一基函数 $\bar{\mathbf{\Phi}}_\ell$ 和 $\bar{\mathbf{\Psi}}_\ell$ 对应的系数：

$$\bar{F}_\ell^* = \bar{\mathbf{\Phi}}_\ell^\top f = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-\frac{1}{2}} \mathbf{\Phi}_\ell^\top f = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{\frac{1}{2}} F_\ell^* \tag{25}$$

$$\bar{G}_\ell^* = \bar{\mathbf{\Psi}}_\ell^\top g = (\mathbf{\Psi}_\ell^\top \mathbf{\Psi}_\ell)^{-\frac{1}{2}} \mathbf{\Psi}_\ell^\top g = (\mathbf{\Psi}_\ell^\top \mathbf{\Psi}_\ell)^{\frac{1}{2}} G_\ell^*. \tag{26}$$

3. 前馈网络实现

我们在图 2 所示的前馈网络中实现该编码框架。编码网络由蓝色和绿色两部分组成，而解码网络仅由绿色部分组成。编码网络以左上方的点云属性 $F_L^* = [\mathbf{y}_i]$ 作为输入，该属性表示为一个 $N \times r$张量，其中 N 为点云中的点数，r 为属性通道数。（在第 2 节中，为简单起见，我们假设 $r = 1$，但很容易将其推广。）属性 $[\mathbf{y}_i]$ 可视为稀疏体素网格中位置 $[\mathbf{x}_i]$ 处的特征，也可视为图中顶点 $[\mathbf{x}_i]$ 处的特征。网络还以右上方的内积矩阵 $\mathbf{\Phi}_L^\top \mathbf{\Phi}_L$ 作为输入，该矩阵表示为一个 $N \times 27$ 张量，其第 ij 个元素为位于体素或顶点 $\mathbf{n}_i = \mathbf{x}_i$ 处的基函数 $\phi_{L,\mathbf{n}_i}$ 与位于其 27 邻域中第 j个邻居 $\mathbf{n}_j$处的基函数 $\phi_{L,\mathbf{n}_j}$ 之间的内积。回想一下，按构造，该元素在 L级别上为 δ_{ij}。

然后，从第 $\ell = L-1$ 层到第 $\ell = 0$ 层，网络通过对$N_{\ell+1} \times r$ 张量 $\tilde{F}_{\ell+1}^*$ 先使用空间不变卷积核进行稀疏卷积，再进行 2^3 下采样，计算 $N_\ell \times r$张量 $\tilde{F}_\ell^* = \mathbf{A}_\ell F_{\ell+1}^*$（$\mathbf{A}_\ell$），如式（16）所示。同样，网络计算

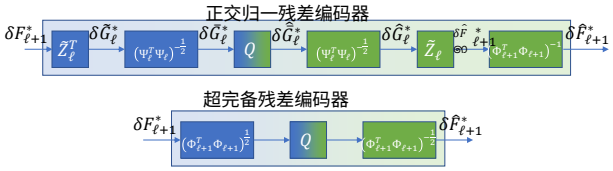


图 3：图 2 中实现残差编码器子网络的多层前馈网络。

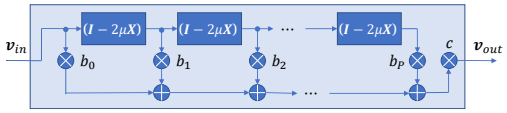


图 4：图 2 和图 3 中实现子网络 $h(\mathbf{X})$ 的多层前馈网络，其中 $h(x) = x^{-1}$、$x^{-1/2}$、或 $x^{1/2}$ 和 $\mathbf{X} = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)$、$(\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_{\ell+1}^a)$ 或 $(\mathbf{\Psi}_\ell^\top \mathbf{\Psi}_\ell)$。这里，$b_0, \ldots, b_P, c$ 是 $h(x)$ 在 $1/(2\mu)$ 附近的 P 阶泰勒展开的系数，其中 μ 是 $\mathbf{X}$ 的特征值上界。

$N_\ell \times 27$ 张量形式的 $\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell$，具体方法是使用空间不变 $3^3 27^2$ 卷积核进行稀疏卷积（$\mathbf{\Gamma}_\ell$），然后进行下采样，这可由式（10）推导得到。在此过程中，网络通过图 4 所示的子网络，按照式（17）计算 $N_\ell \times r$ 张量 $F_\ell^* = (\mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell)^{-1} \tilde{F}_\ell^*$。该子网络使用 $h(x)$ 在 $x_0 = 1/(2\mu)$ 附近的截断泰勒级数近似，实现矩阵 $\mathbf{X} = \mathbf{\Phi}_\ell^\top \mathbf{\Phi}_\ell$ 的算子 $\mathbf{X}^{-1}\mathbf{v}$，其中 μ 是$\mathbf{X}$的特征值上界。

在层级 $\ell = 0$，特征 F_0^* 在编码器处按照 $(\mathbf{\Phi}_0^\top \mathbf{\Phi}_0)^{-1/2}$ 进行归一化（如式（25）所示），然后进行均匀标量量化和熵编码。在解码器（以及编码器）处，这些特征经过解码后再次按照 $(\mathbf{\Phi}_0^\top \mathbf{\Phi}_0)^{-1/2}$ 重新归一化为 $\hat{} F_0^*$。这里，算子 $(\mathbf{\Phi}_0^\top \mathbf{\Phi}_0)^{-1/2}$ 由一个子网络实现，如图 4 所示。

然后，从级别 $\ell = 0$ 到级别 $\ell = L-1$，网络使用稀疏转置卷积，随后进行 2^3 上采样和减法，计算以基 $\mathbf{\Phi}_{\ell+2}$ 表示的残差 $g_{\ell+1}^*$，即 $N_{\ell+1} \times r$ 张量 $\delta F_{\ell+1}^* = F_{\ell+1}^* - \mathbf{A}_{\ell+1}^\top \hat{} F_\ell^*$，如式（18）所示。表示 $\delta F_{\ell+1}^*$ 由图 3 所示的残差编码器子网络之一进行编码，并解码为 $\delta \hat{} F_{\ell+1}^*$。残差编码器子网络同样是前馈网络，其算子通过适当的泰勒级数近似实现，如图 4 所示。

请注意，图 4 中使用截断泰勒展开近似算子 $h(\mathbf{X})$ 的前馈网络，可以被视为具有稀疏边权重矩阵 $\mathbf{X}$ 的图的广义图拉普拉斯算子的多项式。因此，前馈网络可以通过图卷积网络实现。或者，也可以将图卷积视为空间变化核的稀疏 3D 卷积，其权重由超网络 $\mathbf{\Gamma}_\ell$ 确定。从这个意义上说，超网络实现了一种注意力机制，它利用作为 $\mathbf{\Phi}_L^\top \mathbf{\Phi}_L$ 中信号提供的几何来调制属性的处理过程。

4. 实验

在本文中，为了证明我们的归一化近似的有效性，我们将权重设置为 $\mathbf{A}_\ell = [a_{ij}^\ell]$ 无需训练即可直接使用，因此它们对应于 $p = 2$ 阶体积 B样条（如 (8) 中所述）。随后，我们比较预定义的神经网络 $RAHT(2)$ 与 $RAHT(1)$ [4, 13] 在属性表示和压缩方面的表现。评估在具有 10位分辨率 [40] 的 *Longdress* 和 *Redandblack* 数据集上进行。

在图 1 中，我们以给定的比特率重建点云，以

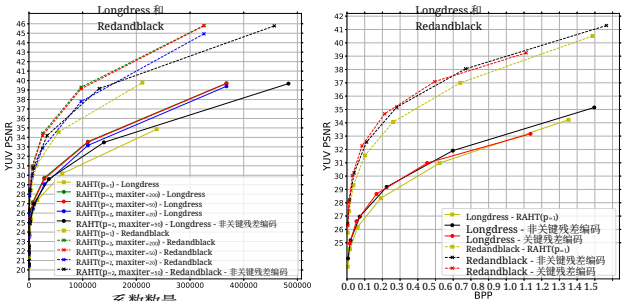


图 5：RAHT($p = 2$) 和 RAHT($p = 1$) 在 *Longdress* 和 *Redandblack* 数据集上的（左）能量集中和（右）编码增益。

展示高阶 B样条的有效性。显然，在给定的比特率下，分段三线性 基函数在视觉上能更好地拟合点云。使用 RAHT(2) 时，块状伪影不太明显，因为高阶基函数可确保块与块之间的颜色平滑过渡。

在图5（左）中，我们展示了失真随对应于0至9不同细节层级的系数数量变化的情况。曲线之间的间隔表示能量集中的改善程度。我们还比较了近似中泰勒展开的不同阶数。显然，RAHT(2)具有更好的能量集中性能，相较于RAHT(1)最高可提升2-3 dB。我们还表明，尽管超完备残差编码需要更多系数来表示每个层级的残差，尤其是在较高细节层级上，但其能量集中性能仍可与RAHT(1)相媲美。

在图5（右）中，我们展示了YUV峰值信噪比随每个输入体素比特数变化的情况。这里，我们采用均匀标量量化，随后使用自适应游程长度Golomb-Rice（RLGR）熵编码 [41]，分别对每个颜色分量（Y、U、V）的系数进行编码。图中显示，在一定比特率范围内，RAHT(2)在失真方面的改善幅度可达1 dB；或者，对于超完备残差编码和正交归一残差编码，比特率均可降低20-30%。在更高比特率下，正交归一（临界）残差编码的性能有所下降。我们认为，这可能与正交归一化矩阵有关：当点云几何结构具有某些局部配置时，这些矩阵可能存在病态条件。进一步研究留待未来开展。令人意外的是，超完备（非临界）残差编码在这些比特率下表现更好。然而，在更高的比特率下（图中未显示），其性能也会下降，表现与其能量集中性能一致。我们认为，原因在于高细节层级的大多数超完备高通系数在低比特率下往往为零，而现在变为非零；这会对编码性能造成严重惩罚。

5. 结论

我们研究了采用体积式方法进行三维点云属性编码，其中对表征目标属性函数 f 的参数 θ 进行量化和编码，使解码器能够恢复 $\hat{}\theta$，以便在已知三维坐标 $\mathbf{x}$ 处计算 $f_{\hat{}\theta}(\mathbf{x})$。对于信号表示，我们定义了一个由高阶 B 样条基张成的函数空间嵌套序列，编码参数即通过稀疏卷积计算得到的 B 样条系数。实验表明，我们的前馈网络相比基于一阶 B 样条基的先前方案，峰值信噪比最高可提升 2–3 dB。

6. 致谢

感谢 S. J. Hwang 和 N. Johnston 提供的宝贵讨论。

7. 参考文献

[1] E. Alexiou、K. Tung 和 T. Ebrahimi，《面向点云压缩的神经网络方法》，载于《数字图像处理应用 XLIII》，A. G. Tescher 和 T. Ebrahimi 编。国际光学与光子学学会，2020 年，第 11510 卷，第 1151008 页，SPIE。[2] C. Zhang、D. Florˆêncio 和 C. Loop，《基于图变换的点云属性压缩》，载于 2014 年 IEEE 国际图像处理会议（ICIP），2014 年 10 月。[3] D. Thanou、P. A. Chou 和 P. Frossard，《动态三维点云序列的基于图的压缩》，《IEEE 图像处理汇刊》，第 25 卷，第 4 期，2016 年 4 月。[4] R. L. de Queiroz 和 P. A. Chou，《使用区域自适应分层变换压缩三维点云》，《IEEE 图像处理汇刊》，第 25 卷，第 8 期，第 3947–3956 页，2016 年。[5] R. L. de Queiroz 和 P. A. Chou，《动态体素化点云的运动补偿压缩》，《IEEE 图像处理汇刊》，第 26 卷，第 8 期，第 3886–3895 页，2017 年 8 月。[6] R. A. Cohen、D. Tian 和 A. Vetro，《使用图变换压缩稀疏点云属性》，载于 IEEE 国际图像处理会议（ICIP），2016 年 9 月。[7] P. A. Chou、M. Koroteev 和 M. Krivoku´ca，《点云压缩的体积式方法——第一部分：属性压缩》，《IEEE 图像处理汇刊》，第 29 卷，第 2203–2216 页，2020 年。[8] M. Krivoku´ca、P. A. Chou 和 M. Koroteev，《点云压缩的体积式方法——第二部分：几何压缩》，《IEEE 图像处理汇刊》，第 29 卷，第 2217–2229 页，2020 年。[9] S. Schwarz、M. Preda、V. Baroncini、M. Budagavi、P. Cesar、P. A. Chou、R. A. Co- hen、M. Krivoku´ca、S. Lasserre、Z. Li、J. Llach、K. Mammou、R. Mekuria、O. Nak- agami、E. Siahaan、A. Tabatabai、A. Tourapis 和 V. Zakharchenko，《新兴的 MPEG 点云压缩标准》，《IEEE 电路与系统新兴主题期刊》，第 9 卷，第 1 期，第 133–148 页，2019 年 3 月。[10] D. Graziosi、O. Nakagami、S. Kuma、A. Zaghetto、T. Suzuki 和 A. Tabatabai，《点云压缩标准化工作的概述：基于视频的（V-PCC）和基于几何的（G-PCC）》，《APSIPA 信号与信息处理汇刊》，第 9 卷，第 e13 页，2020 年。[11] E. S. J ang、M. Preda、K. Mammou、A. M. Tourapis、J. Kim、D. B. Graziosi、S. Rhyu 和 M. Budagavi，《MPEG 中的基于视频的点云压缩标准：从证据收集到委员会草案 [标准概要]》，《IEEE 信号处理杂志》，第 36 卷，第 3 期，第 118–123 页，2019 年。[12] 3DG，《G-PCC 编解码器说明 v12》，已批准的 WG 11 文件 N18891，ISO/IEC MPEG JTC1/SC29/WG11，瑞士日内瓦，2020 年 10 月。[13] G. P. Sandri、P. A. Chou、M. Krivoku´ca 和 R. L. d. Queiroz，《区域自适应分层变换的整数替代方案》，《IEEE 信号处理快报》，第 26 卷，第 9 期，第 1369–1372 页，2019 年。[14] G. Toderici、S. M. O'Malley、S. J. Hwang、D. Vincent、D. Minnen、S. Baluja、M. Covell 和 R. Sukthankar，《使用循环神经网络进行可变码率图像压缩》，载于第 4 届国际学习表征会议（ICLR），2016 年。[15] J. Ball´e、V. Laparra 和 E. P. Simoncelli，《面向感知质量的非线性变换编码端到端优化》，载于 2016 年图像编码研讨会（PCS），2016 年。[16] G. Toderici、D. Vincent、N. Johnston、S. J. Hwang、D. Minnen、J. Shor 和 M. Covell，《使用循环神经网络进行全分辨率图像压缩》，载于 2017 年 IEEE 计算机视觉与模式识别会议（CVPR），2017 年。[17] J. Ball´e、V. Laparra 和 E. P. Simoncelli，《端到端优化的图像压缩》，载于第 5 届国际学习表征会议（ICLR），2017 年。[18] J. Ball´e，《用于有损图像压缩的高效非线性变换》，载于 2018 年图像编码研讨会（PCS），2018 年。[19] J. Ball´e、D. Minnen、S. Singh、S. J. Hwang 和 N. Johnston，《采用尺度超先验的变分图像压缩》，载于第 6 届国际学习表征会议（ICLR），2018 年。[20] D. Minnen、J. Ball´e 和 G. Toderici，《用于学习式图像压缩的联合自回归与分层先验》，载于《神经信息处理系统进展 31》，2018 年。[21] J. Balle、P. A. Chou、D. Minnen、S. Singh、N. Johnston、E. Agustsson、S. J. Hwang 和 G. Toderici，《非线性变换编码》，《IEEE 信号处理精选主题期刊》，第 1–1 页，2020 年。[22] F. Mentzer、G. D. Toderici、M. Tschannen 和 E. Agustsson，《高保真生成式图像压缩》，《神经信息处理系统进展》，第 33 卷，2020 年。[23] Y. Hu、W. Yang、Z. Ma 和 J. Liu，《端到端学习有损图像压缩：基准测试》，《IEEE 模式分析与机器智能汇刊》，2021 年。[24] W. Yan、Y. Shao、S. Liu、T. H. Li、Z. Li 和 G. Li，《基于深度自编码器的点云有损几何压缩》，CoRR，第 abs/1905.03691 卷，2019 年。

[25] M. Quach、G. Valenzise 和 F. Dufaux，“用于有损点云几何压缩的学习式卷积变换”，载于 *2019 IEEE*国际图像处理会议（*ICIP*），2019年。[26] A. F. R. Guarda、N. M. M. Rodrigues 和 F. Pereira，“基于深度学习的点云编码：行为与性能研究”，载于 *2019*年第*8*届欧洲视觉信息处理研讨会（*EUVIP*），2019年，第34–39页。[27] A. F. R. Guarda、N. M. M. Rodrigues 和 F. Pereira，“点云编码：采用基于深度学习的方法”，载于 *2019*图像编码研讨会（*PCS*），2019年，第1–5页。[28] A. F. R. Guarda、N. M. M. Rodrigues 和 F. Pereira，“基于深度学习的点云几何编码：通过隐式量化和显式量化进行率失真控制”，载于 *2020 IEEE*多媒体与博览会国际会议工作坊（*ICMEW*），2020年。[29] D. Tang、S. Singh、P. A. Chou、C. H¨ane、M. Dou、S. Fanello、J. Taylor、P. Davidson、O. G. Guleryuz、Y. Zhang、S. Izadi、A. Tagliasacchi、S. Bouaziz 和 C. Keskin，“深度隐式体积压缩”，载于 *2020 IEEE/CVF*计算机视觉与模式识别会议（*CVPR*），2020年。[30] M. Quach、G. Valenzise 和 F. Dufaux，“改进的深度点云几何压缩”，*2020 IEEE*第*22*届多媒体信号处理国际研讨会（*MMSP*），第1–6页，2020年。[31] X. Sheng、L. Li、D. Liu、Z. Xiong、Z. Li 和 F. Wu，“Deep-PCAC：面向点云属性的端到端深度有损压缩框架”，*IEEE*多媒体汇刊，第1–1页，2021年。[32] B. Isik、P. A. Chou、S. J. Hwang、N. Johnston 和 G. Toderici，“LVAC：使用基于坐标的网络进行点云学习式体积属性压缩”，*arXiv*预印本 *arXiv:2111.08988*，2021年。[33] J. Wang 和 Z. Ma，“基于稀疏张量的点云属性压缩”，载于 *2022 IEEE*多媒体信息处理与检索国际会议（*MIPR*），2022年。[34] T. Bird、J. Ball´e、S. Singh 和 P. A. Chou，“通过熵惩罚神经表示函数实现三维场景压缩”，载于 图像编码研讨会（*PCS*），2021年6月。[35] B. Isik，“通过模型压缩实现神经三维场景压缩”，*2021 IEEE*计算机视觉与模式识别会议（*CVPR*）*WiCV*研讨会。*arXiv:2105.03120*，2021年。[36] T. Takikawa、A. Evans、J. Tremblay、T. M¨uller、M. McGuire、A. Jacobson 和 S. Fidler，“可变比特率神经场”，载于计算机图形学与交互技术特别兴趣小组会议论文集，美国纽约州纽约市，2022年，SIGGRAPH22会议论文集，计算机协会。[37] O. Ronneberger、P. Fischer 和 T. Brox，“U-Net：用于生物医学图像分割的卷积网络”，载于 医学图像计算与计算机辅助干预——*MICCAI 2015*，N. Navab、J. Hornegger、W. M. Wells 和 A. F. Frangi 编，2015年，第234–241页，施普林格国际出版公司。[38] G. Cheung、E. Magli、Y. Tanaka 和 M. Ng，“图谱图像处理”，载于*IEEE*会刊，2018年5月，第106卷，第5期，第907–930页。[39] D. G. Luenberger，《向量空间方法优化》，约翰·威立父子出版公司，1969年。[40] E. d'Eon、B. Harrison、T. Meyers 和 P. A. Chou，“8i体素化全身数据——一个体素化点云数据集”，输入文档 M74006 和 m42914，ISO/IEC JTC1/SC29 WG1 和 WG11（JPEG 和 MPEG），斯洛文尼亚卢布尔雅那，2017年1月。[41] H. Malvar，“未知统计特性的量化广义高斯源的自适应游程/Golomb-Rice编码”，载于 数据压缩会议（*DCC'06*），2006年，第23–32页。