

## 2、如何对每个code再生成一个shape?

1、2互为编码解码过程? 3、瘦surface怎么办吗? 4、大的scene怎么处理吗? 可以不用sample? 5、如何动画形变? CVF这篇cvpr2021论文是开放版本, 由计算机视觉基金会提供。

除此水印外, 其特征应与已接受的版本相同;  
会议的最终出版版本可在本网站上找到。

详细的神经几何水平: 具有隐式三维形状的实时渲染高川, 2,4,2,3,3,3,2,2,4

1 NVIDIA2 多伦多大学3 麦吉尔大学4 向量研究所

共振函数 (nerualSDFs) 被合并为

---

\*作者贡献相同。

神经符号距离函数 (SDFs) 是3dshape的有效表示。最先进的方法通常对SDF与固定大小的神经网络进行编码, 以近似复杂的sha类型。然而, 使用这些大型网络进行渲染, 在计算上很昂贵, 因为它需要每个像素的网络, 使得这些表示对于实时图形不切实际。我们介绍了效率神经表示, 首次实现了高保真度神经SDFs的实时渲染, 同时实现了最先进的几何重建质量。我们使用一个八叉树基特征体来表示隐式表面, 该特征体自适应具有多个双屏幕细节层次 (LOD) 的形状, 并支持具有SDF插值的连续LOD。我们开发了一种有效的算法, 通过只查询稀疏遍历的必要lod, 直接实时呈现我们新的神经SDF表示。我们表明, 与之前的作品相比, 我们的表示的渲染速度要高2-3个数量级。此外, 它在3d几何和二维图像空间度量下产生了art重建质量f或复杂形状。

### 1. 介绍

在计算机图形中使用的高级几何建模和渲染技术, 使用具有复杂细节、任意拓扑和质量的三维形状, 通常利用多角网格。然而, 适应基于学习的方法是很重要的, 因为它们缺乏不同的能力, 因此不能轻易地用于子宫补偿应用, 如基于学习的3直肠教学。最近, 信号器的神经近似

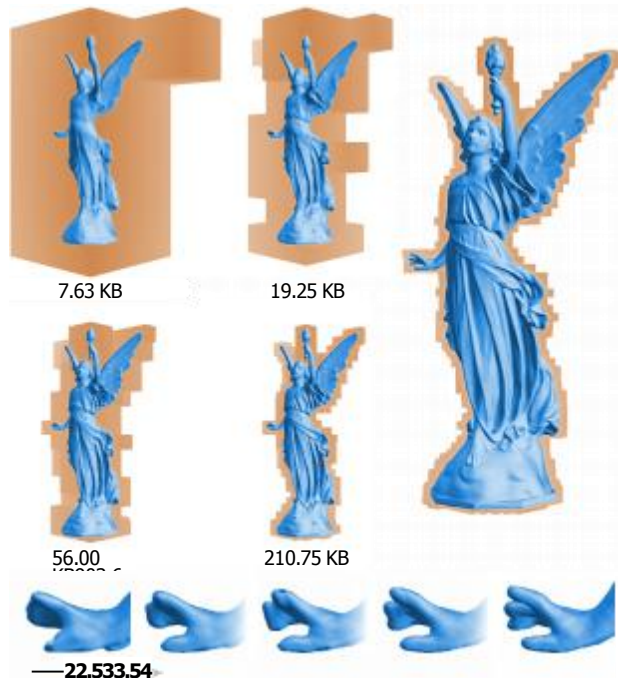


图1：详细程度。我们的表示池特征从多个尺度到自适应地重建具有连续细节水平（LOD）的大腿保真度几何。这些图显示了不同lod下的面（蓝色），叠加在相应的粗糙的稀疏八叉树（橙色）上，其中包含了所获得的有符号距离函数的特征。这些都是直接呈现在实时使用我们的高效稀疏球跟踪算法。

扩大计算机视觉和图形应用程序的选择性选择。[36,31,6,8] 之前的工作表明，神经网络可以通过学习sdf来编码精确的三维几何，而不限拓扑或分辨率，sdf通过其零层集定义曲面。这些工作方法只使用一个大的，固定大小的多层p感知器（MLP）作为学习的距离函数。

直接渲染和探测神经SDFs通常在球面跟踪[17]，算法-识别算法可以



图2: 我们能够适应不同的复杂性、风格、比例的形状, 具有一致的粘质量, 同时能够利用几何阴影, 环境遮挡[11], 甚至与二次光线的阴影。最好的维维佐米丁。

需要数百个SDF评估才能收敛。由于单个向前通过基于大型mlp的SDFcan需要数百万次操作, 神经SDFs对于实时图形应用程序来说是不切实际的, 因为计算单个像素的方法涉及数以百万计的操作。诸如Davies等人的工作。[8]通过使用一个小的神经网络来支持单链来解决这个问题, 但这是以通用性和重建质量为代价的。以前的方法也使用了固定大小的神经网络, 这使得它们无法表达复杂度超过网络容量的地理测量。

本文提出了一种新的表示方法可以自适应规模到不同级别的de-的ralsdf尾部(lod)和重建高度细节的几何结构。我们的该方法可以在不同的尺度之间平滑地进行插值的几何图形(见图1), 并可以实时渲染具有合理的内存占用空间。类似于戴维斯等al.[8], 我们也使用一个小的MLP来进行球体跟踪实用, 但不牺牲质量或通用性。

我们的灵感来自于经典的表面提取使用正交和空间数据的机制[26,12]存储距离值的结构可以离散欧盟客户空间, 使简单的线性基函数可以重新生成构造几何图形。在这样的作品中, 决议或树深度决定了细节的几何级别(LOD)和不同的lod可以用插值结束。如何曾经, 它们通常需要高树的深度来重建aso-润滑剂具有满意的质量。

相反, 我们使用稀疏空间来离散空间体素体树(SVO), 我们存储学习到的特征向量代替有符号的距离值。这些向量可以编码成标量距离使用一个较浅的MLP, 允许我们要截断树的深度, 同时继承其优势有一些经典的方法(例如, LOD)。我们还另外开发了一些产品

一种针对我们的架构量身定制的射线遍历算法使我们的扭转器几何形状接近100×比DeepSDF[36]。虽然直接与神经系统进行比较据我们报道, 体积渲染方法是不可能的比NeRF [32]快500×, 快50×比NSVF[24]在类似的实验设置中更好。

综上所述, 我们有贡献的领域如下:

我们介绍了第一个具有神经sdf的复杂几何的实时渲染方法。

我们提出了一种神经SDF表征离子, 它能够有效地捕获多个lod, 并以最先进的质量重建三维几何图形(见图2)。

我们展示了我们的架构可以以比传统方法更高的视觉保真度表示3D形状压缩格式, 并概括不同的学习例子, 甚至从单一的几何图形。

由于我们的方法的实时性, 我们将其设想为大量应用的模块化构建程序, 如基于时代的场景重建、机器人导航和形状分析。

## 2.相关工作

我们的工作与细节层次的网格褶皱化、三维神经形状表征和隐式神经渲染的初步研究最为相关。

详细程度。细节水平(LOD)[27]计算机图形是指限制特征变化的三维形状, 通常在图像空间中像素大小的左右。这种减轻锯齿叠造成的攻击, 并通过降低模型复杂性来加速渲染。虽然信号处理技术可以过滤纹理[46], 几何过滤是特定的表示和变化。一种方法是网格抽取, 其中网格被简化为预算数的离移点、顶点或边。经典的方法[14,18]贪婪地去除对几何精度影响最小的中间网格。更多的重构方法优化了感知度量[23,22,7]或专注于拓扑[29]。网格在低内存约束下存在离散化错误, 在lod之间存在混合困难。相比之下, sdf可以用更少的内存再现平滑的表面, 并平滑地连接lod, 以减少混叠。在过去的工作中, [40,44]在压缩gridsdf方面也表现出了令人印象深刻的性能。我们的神经sdf继承了这些属性。

神经隐式表面。隐式曲面基基值在潜在向量或神经网络权值中编码几何, 通过水平集参数化曲面。

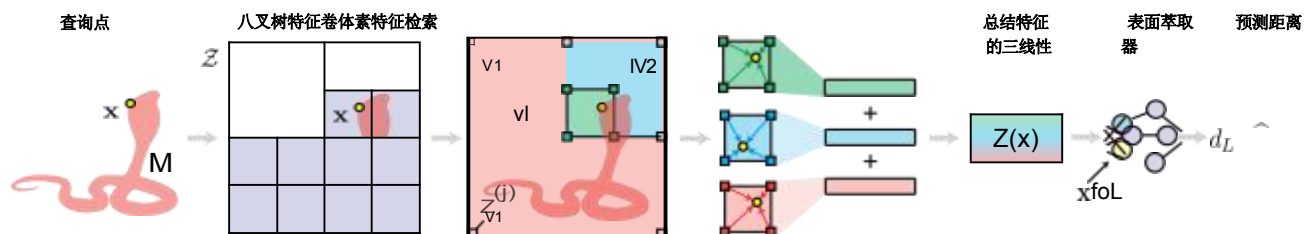


图3：建筑。我们使用稀疏体素八树（SVO）对神经SDF进行编码，它包含一个集合距离 $z$ 。SVO的级别定义 $\text{lod}$ ，体素角包含定义局部表面分段 $ts$ 的特征向量。给定查询点和LOD  $L$ ，我们找到相应的体素 $V1$ ： $L$ ，三内插值它们的角 $V(j)$ 到 $\text{toL}$ 和 $\text{sum}$ ，得到一个特征向量 $z(x)$ 。和 $x$ 放在一起，将该特征输入一个小MLP  $f_\theta$ ，获得组合距离 $\hat{d}$ 。我们共同优化MLP parameters  $\theta$ 和端到端特征。

开创性的工作[36,31,6]通过使用自动解码器MLP将这些向量转换为潜在向量，它输出一个基于向量和位置的标量。另一个并行工作的工作线[43,41]使用周期函数，从而导致重建质量的巨大改进。戴维斯等人。[8]支持将神经网络过度拟合到单一形状，允许一个通信协议MLP来表示几何形状。类似课程DeepSDF[10]在渐进生长网络中编码几何，但丢弃中间表示。BSP-Net和CvxNet[5,9]使用空间分区树学习隐含几何。PIFu[38,39]以深度作为附加的输入参数在密度2Dgrid上学习特征，而其他工作在稀疏规则[15,3,20]或变形的[13]3D网格上学习这些特征。补丁网[45]学习表面补丁，由特征的点云定义。这些工作主要依赖于等表面提取算法，如标记立方体[26]来创建一个密集的表面网格来渲染图像对象。相反，在本文中，我们提出了一种直接以交互速率呈现形状的方法。

隐式曲面的神经渲染。许多工作都是渲染神经隐式表示。SRN[42]通过直接预测深度来渲染隐式曲面。尼迈耶等人。[34]提出了一种使用射线行进的隐式曲面的微分渲染器。DIST [25]和SDFD- iff [21]使用分支跟踪为sdf提供了可微的渲染器。这些可微渲染器是热跟踪算法；它们只需要关于热表面相交的可发酵性。因此，我们利用在这些工作中提出的名称，使我们的渲染器也可区分。NeRF

[32]学习几何形状作为密度精灵，并使用光线行进来可视化它们。IDR [47]将一个以MLP为基础的阴影功能附加到一个神经SDF上，解开几何形状和阴影。NSVF[24]与我们的工作类似，它也用稀疏八叉树编码特征表示。与NSVF相比，我们的工作使细节级别成为可能，并使用球体追踪，这允许我们从阴影中分离出几何图形，从而优化光线追踪，这在体积渲染框架中是不可能实现的。正如之前所述，我们的渲染器是两个量级的顺序

比在NSVF[24]中报告的数字更快。

### 3.方法

我们的目标是设计一种表示，可以重构详细的几何结构并实现连续的细节水平，同时可以在交互中呈现。图3显示了我们的方法的可视化概述。Section3.1pro-vides是一个关于神经sdf及其局限性的背景。然后，我们提出了在Section3.2and中对神经稀疏体素八叉树进行编码的方法，并在第3.3节中提供了训练去尾。我们的渲染算法为我们的表示在第3.4节中描述。

#### 3.1.神经有符号距离函数（SDFs）

SDFs是函数 $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ ，其中 $d = f(x)$ 是从点 $x$ 到体积 $M \subset \mathbb{R}^3$ 的表面的最短符号距离，其中符号表示 $x$ 是否在 $M$ 的外侧。因此， $\partial$ 隐含地表示为 $f$ 的零水平集：

$$S = \{x \in \mathbb{R}^3 \mid f(x) = 0\}. \quad (1)$$

$f_\theta(\text{los}(x, s)) (f_\theta(x) (a_\theta))$ ， $s_\theta$ ，这是他所掌握的 $\theta = d$ ， $\theta =$ 有符号距离和 $L$ 是某个距离度量，如12-距离。一个可选的输入“形状”特征向量 $z \in \mathbb{R}$ 可以用 $a_{\text{fixed}}$ 使网络适应不同的形状。

为了直接渲染神经元sdf，光线跟踪可以使用根固定算法，如球体跟踪[17]。该算法可以实现多达100个距离查询，如果网络较大且距离查询太慢，则可以实现标准的神经sdf的抑制。使用小型网络可以加快这一迭代过程，但重建的形状可能会速度。此外，固定大小的网络不能适应高度复杂的形状，也不能适应不需要视觉细节的简单或远物体。



在下一节中，我们描述了一个框架，通过使用稀疏素八叉树编码SDF来解决这些问题，允许表示适应不同的细节水平，并使用浅神经网络的脚趾几何，同时保持几何精度。

### 3.2. 神经几何层次的细节

框架。与标准的神经sdf相似，我们使用参数神经网络和数据学习输入特征来编码形状，从而预发送sdf。在DeepSDF[36]中，我们使用一个特征向量来编码形状，我们使用一个包含特征向量集合的特征体积，我们用Z表示。

八叉树（SVO），跨越边界体积 $B=[-1,1]^3$ 。svo中的每个体素V都拥有其八角（由j索引）的可指定特征向量 $z_V(j) \in \mathbb{R}^3$ ，如果邻域体素存在，它们是共享的。只有当体素V包含表面时，体素才被分配，使SVO稀疏。

定义的每一级LOD强度测量。随着SVO中的树深度增加，树面用离散化表示，允许侦察构造质量随记忆使用情况进行缩放。我们将最大树的深度表示为 $L_{max}$ 。我们另外使用小MLP神经网络 $f_{\theta_1}$ ： $L_{max}$ ，表示解码器，与parameters $\theta_1$ ： $L_{max}=\{\theta_1, \dots, \theta_{L_{max}}\}$ 为每个LOD。

为了计算满足所需LOD L的查询点 $x \in \mathbb{R}^3$ 的SDF，我们遍历树到level L，以找到所有体素 $V_1: L = \{V_1, \dots, V_L\}$ 包含x。预测level $\ell \in \{1, \dots, L\}$ ，我们通过三期插值x处体素的角特征来计算每体素形状的vector $\psi(x; \ell, Z)$ 。我们总结了各个层次的特征

获取 $z(x; L, Z) = \Sigma \ell = 1 \psi(x; \ell, Z)$ ，并将其传递到具体计算的SDF为

$$\hat{d}_L = f_{\theta_L}([x, z(x; L, Z)]). \quad (2)$$

其中， $\hat{d}_L$ 表示连接。这个跨越lod的总结允许有意义的梯度来支持玛瑙交叉lod，帮助特别是更粗的lod。

由于我们的形状向量 $z_V(j)$ 现在只代表小的表面片段，而不是整个的形状，因此我们将计算复杂度从神经网络 $f_{\theta}$ 移到特征向量 $query \psi: \mathbb{R}^3 \rightarrow \mathbb{R}^m$ 中，即aSVO遍历和体素特征的三线性插值。这个关键的设计决策允许我们使用非常小的mlp，提高显著的重建质量。

水平混合。虽然八叉树的水平是稳定的

得到一个期望的连续LOD $L \geq 1$ ，我们通过线性插入 $\hat{d}_L$ 来混合不同离散的八叉树LODs $L$ ，我们能够平稳地插入它们之间。向

对应的预测距离：

$$\hat{d}_L = (1-\alpha) \hat{d}^{L*} + \alpha \hat{d}^{L*+1} + \frac{1}{3} \hat{d}^{L*+2}$$

其中 $L^*=L$  and  $\alpha=L-L$ 是分数部分，允许我们在lod之间平稳过渡（见图1）。这种简单的混合方案只适用于sdf，而不适用于密度或占用和分布

定义为网格和云。我们讨论如何

在Section 3.4.3.3中设置渲染时的连续LOD L。训练

我们通过联合训练each LOD，确保每个旋转预测的离散水平都具有有效的几何形状。我们通过计算每个级别的个人损失，并在各个层次进行它们：

$$J(\theta, Z) = \mathbb{E}_{x,d} \sum_{L=1}^{L_{max}} f_{\theta_L}([x, z(x; L, Z)]) - d \quad (4)$$

然后，我们随机地将损失函数优化为both $\theta_1: L_{max}$ 和 $z$ 。通过对点 $\in B$ 的重要抽样来估计期望。我们使用了三种分布的混合样本： $B$ 中的均变分布，表面样本和扰动表面样本。我们在补充材料中详细介绍了这些采样算法和具体的训练超参数。

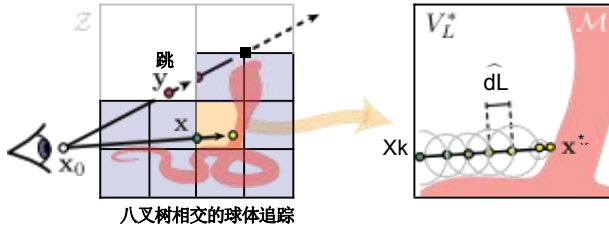
### 3.4. 交互式渲染

Sphere跟踪。我们使用领域跟踪[17]来直接呈现我们的表示。然而，呈现一个基于svo的sdf内层跟踪，提出了一些需要解决的技术建议。典型的sdf被定义在所有的 $\mathbb{R}^3$ 上。相比之下，我们的SVO sdf只适用于与表面几何形状相交的体素sv。因此，需要正确地处理在空白空间中进行的距离查询。一种选择是使用恒定的步长，即光线进行，但不能保证轨迹会转换，因为步长会超调。

相反，在帧的开始，我们首先执行ray-SVO交集（详情见下文）来检索每个与射线相交的分辨率。形式上，如果 $r(t) = x_0 + t d$ ， $t > 0$ 是一条原点为 $x_0 \in \mathbb{R}^3$ ，方向为 $d \in \mathbb{R}^3$ 的射线，我们让 $V_{\ell}(r)$ 表示r在 $\ell$ 级相交体素的深度序集。

$V_{\ell}(r)$ 中的每个体素都包含相交的光线指数、体素位置、父体素以及指向八个角特征向量 $z_V(j)$ 的指针。我们检索特征向量的指针。特征向量存储在固定数组中，指针在初始化步骤中通过迭代g个整体体素，并对每个角落的特征找到相应的索引。

自适应射线步进。对于球面轨迹迭代k中的给定射线，我们执行一个ray-AABB交点[28]



八叉树相交的球体追踪

图4：自适应区域步骤。当查询点位于一个体素内（例如， $x$ ）时，对所有响应的体素进行三线性插值，直到基本的八叉树分辨率，以计算非球面跟踪步骤（右）。当查询点在体素外部（例如， $y$ ）时，将使用ray-AABB交集跳道到下一个体素。

针对目标LODIL中的体素检索

命中第一个体素 $V^*L \in V_L(r)$ 。如果 $x_k \in V^*L$ ，我们先推进 $x_k$

到ray-AABB的交点。如果 $x_k \in V^*L$ ，我们就知道

我们的功能体积。我们递归地检索所有对应于粗levels $\ell \in \{1, \dots, L-1\}$ 的解析体 $V_\ell^*$ ，得到体素集合 $V_1^* : L$ 。然后我们在每个节点上的三线性插值特征相加。

注意，副节点总是通过构造而存在的。MLP f0的支持

du w1 c r1 ces a conse rv ative distane  $\hat{d}_L$  t m ve in dretion d and

o so iic ee t ake a stand ard sphere ta ng tep:  $x_{k+1} \leftarrow x_k + \hat{d}_L d$

1 If  $x_{k+1}$  is no w in empty space we sk p to th next v oxe

在 $V_L(r)$ 中，沿着光线移动，如果不存在，则丢弃therayr。如果 $x_{k+1}$ 在体素内，我们执行s前跟踪。这将重复，直到所有光线都错过，或者达到停止条件以恢复命中点 $x^* \in s$ 。该过程如图4所示。这种自适应跳跃使体素从不查询稀疏，允许我们的表示提供最小的存储。我们在补充材料中详细说明了停止标准。

SparseRay-OctreeIntersection.我们现在描述了我们新的射线-八叉树交叉算法，它利用宽fi第一次遍历策略和并行扫描内核[30]来在现代图形硬件上实现高性能。算法1提供了我们的算法的伪代码。我们将在补充材料中提供子程序的细节。

算法1：迭代、平行、宽-冷杉树遍历

```

1: 程序RAYTO CTREE (L, R)
2:   N i(0)  $\leftarrow \{i, 0\}, i = 0, \dots, |R|-1$ 
3: for  $\ell=0$  to L do
4: D  $\leftarrow$  DECIDE ( R, N( $\ell$ ),  $\ell$ )
5: 它包括 $\leftarrow$ 和(D)
6: if  $\ell=L$ , 然后
7: N ( $\ell$ )  $\leftarrow$  C压缩 (N ( $\ell$ )、D、S) 8: else
9: N( $\ell+1$ ) $\leftarrow$  S UBDIVIDE (N( $\ell$ ), D,S)

```

该算法首先生成了由  
i)，并将它们存储在数组N (0)磨损体素对中

是射线体素交叉的建议。我们用八叉树初始化每个 $N_i(0) \in N(0)$ 的根节点的停止级体素（第2行）。接下来，我们遍历八叉树levels $\ell$ （第3行）。在每次迭代中，我们确定在DECIDE中导致相交的射线-体素对，如果ray与体素相交，则为 $D_j=0$ ，则返回一个决策列表和 $D_j=1$ ，否则（第4行）。然后，我们使用E扩展的SUM来计算列表D的唯一和，我们将其输入接下来的两个子例程（第5行）。如果我们还没有达到我们想要的LOD级别L，我们使用Subspien来显示下一个列表 $N(\ell+1)$ 和光线相交的子变量 $N_j(\ell)$ 并继续迭代（第9行）。否则，我们使用COMPACTIFYto来计算所有不导致交集的 $N_j(\ell)$ （第7行）。结果是一个紧凑的，深度有序的射线-体素交叉点列表。注意，通过分析热起源位于体素内部的空间，我们可以对子体素进行排序，这样射线-体素对 $N(L)$ 将按到射线起源的距离排列。

LOD选择。用深度 $\tilde{\cdot}$ 表示渲染，其中 $\tilde{\cdot}$ 与基于距离输出的

## 4.实验

我们做了几个实验来展示我们的架构的有效性。我们首先将我们的模型匹配到数据集的三维网格模型，包括形状网[4]、Thing10K[48]模型，并从TurboSquid1中选择模型，并基于基于3D几何的度量和渲染的图像空间度量对它们进行评估。我们还证明了我们的模型能够拟合具有独特的性质的复杂解析符号距离函数。我们还展示了实时渲染、推广到多种形状和几何简化的结果。

我们实验中使用的MLP只有维为 $h=128$ 的隐藏层，中间层有恢复激活，比我们比较的基线中使用的网络运行得更快，如实验所示。我们使用了一个名为 $m=32$ 的SVO特征维度。我们使用高斯先验with $\sigma=0.01$ 初始化其特征 $z \in z$ 。

### 4.1.重建三维数据集

在几个不同的3个数据集上建立我们的架构，以评估重建数据表面的质量。我们比较了基线，包括DeepSDF[36]，傅里叶特征网络[43]，SIREN [41]，和神经隐式-

<sup>1</sup><https://www.turbosquid.com>

<sup>2</sup><https://www.shadertoy.com>

	储存 (KB)	#推理参数。	形状网150[4]		三叶草32[48]		涡轮机16			
			gIoU ↑	倒角-L1 ↓	gIoU ↑	倒角-L1 ↓	iIoU ↓	正常-L2 ↓	gIoU ↑	倒角-L1 ↓
DeepSDF[36]	7186	1839614	86.9	0.316	96.8	0.0533	97.6	0.180	93.7	0.211
纤维素[43]	2059	526977	88.5	0.077	97.7	0.0329	95.5	0.181	92.2	0.362
警报器[41]	1033	264449	78.4	0.381	95.1	0.0773	92.9	0.208	82.1	0.488
神经隐式[8]	30	7553	82.2	0.500	96.0	0.0919	93.5	0.211	82.7	0.354
我们/ LOD1	96	4737	84.6	0.343	96.8	0.0786	91.9	0.224	79.7	0.471
我们/ LOD 2	111	4737	88.3	0.198	98.2	0.0408	94.2	0.201	87.3	0.283
我们/ LOD 3	163	4737	90.4	0.112	99.0	0.0299	96.1	0.184	91.3	0.162
我们的4号线	391	4737	91.6	0.069	99.3	0.0273	97.1	0.170	94.3	0.111
我们/ LOD 5	1356	4737	91.7	0.062	99.4	0.0271	98.3	0.166	95.8	0.085
我们的6号线	9826	4737	—	—	—	—	98.5	0.167	96.7	0.076

表1：网格重建。这个表展示了与三个不同的数据集进行的架构和每个形状的重建比较。在所有的评估方案中，我们从LOD3开始的架构尽管存储和推理参数要低得多，但性能要好得多。我们的表示的存储是基于所有数据集中所有形状的平均解析体素计数加上解码器的大小来计算的，并且#推理参数测量用于单个距离的网络参数。

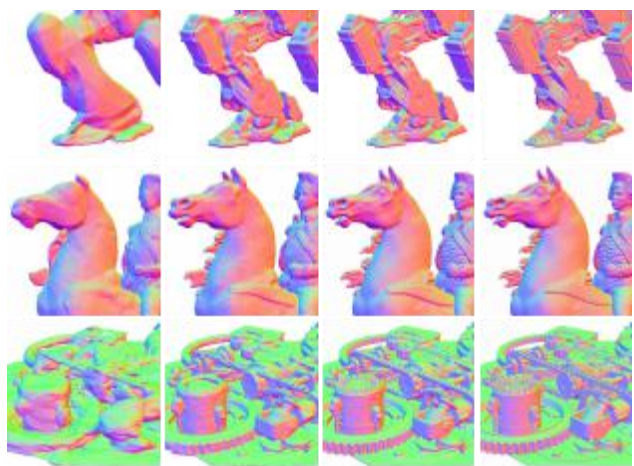
其（NI）[8]。这些架构显示了对3D形状的过拟合的最先进的性能，并且也有可用的原始代码。我们使用它们的源代码作为参考，并在补充材料中提供细节，以重新实现这些基线。

网格数据集。表1显示了沙网、Thingi10K和涡轮增压乌贼的总体结果。我们分别从每个数据集中抽取150、32和16个形状，并使用100、100和600个周期对每个形状进行采样。对于ShapeNet150，我们分别使用了来自汽车、飞机和椅子类别的50个形状。ForThingi32，我们使用

32个形状被标记为扫描。ShapeNet150和Thingi32使用倒角-l1距离（乘以103）和均匀采样点（gIoU）上的交叉过并集进行评估。TurboSquid具有更有趣的表面特征，因此我们使用了3个几何基础的歌词以及基于32个多视图渲染反射的图像空间指标。具体地说，我们计算了分割掩模（iIoU）和分割空间归一化对齐的并集相交。形状的复杂性在数据集上大致增加。我们为沙佩网150和Thingi32训练5个LOD，为TurboSquid训练6个LOD。对于数据准备，我们遵循DualSDF [16]并对其进行归一化，去除神经间三角形，并用射线刺[35]表示距离。

存储（KB）对应于解码器大小和表示的和，假设精度为32位。对于我们的架构，解码器对米由90kb的存储影响组成，因此较低的lod的有效存储大小更小，因为解码器能够推广多个形状。#推理参数。是距离序列所需的参数的数量，它与推理所需的fiops的数量大致相关。

在所有的数据集和指标上，我们实现了最先进的结果。值得注意的是，我们的r演示显示了从第三个lod开始的更好的结果，在那里我们的存储影响最小。我们还注意到，我们的推断成本估计为4737



NI[8] FFN[43]我们的/ LOD 6参考资料

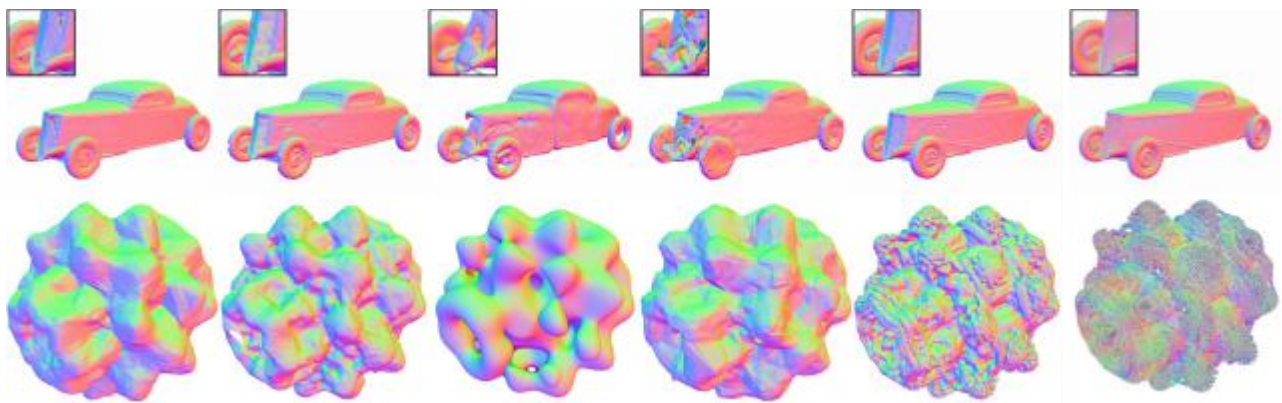
图5：对TurboSquid的比较。我们定性地比较了网格重建。只有我们的系统能够恢复细节，速度比FFN快50×，与toNI相当。我们通过渲染表面规范化来突出显示几何细节。

与FFN[43]相比，FFN[43]需要的推理参数少99%，比神经隐式[8]少37%，同时显示出更好的重建质量（定性评估见图5）。

特例分析SDFs。我们还评估了在两个特别困难的分析sdf同事上的重建。老年人模型是一种高同步度量的SDF，它不满足电子方程 $|\nabla f|=1$ ，并包含不连续点。这是一个关键的情况，因为基于度量的sdf经常被用于社会效应和更容易建模的sdf。具有无限分辨率的曼德尔布尔比萨递归分形。这两个sdf都是由数学表达式定义的，我们从数学表达式中提取和采样距离值。我们对这些分析形状进行了100个时期的训练，每个时期有5个×106个样本。

只有我们的架构能够捕获这些复杂例子的高频尾，以达到合理的精度。无





DeepSDF[36] FFN[43] SIREN[41]神经隐式[8]我们的/ LOD5参考

Figure6: AnalyticSDFs.我们测试了shadertoy的催化SDF例子；它包含一个高度量符号距离场，以及曼德尔灯泡，它是一个递归分形结构，只能用隐曲面表示。只有我们的架构才能合理地重建这些困难的情况。我们将表面归一化渲染为高光几何细节。

Method/ LOD	1	2	3	4	5
DeepSDF (100 epochs)[36]	0.0533	0.0533	0.0533	0.0533	0.0533
FFN (100 epochs)[43]	0.0329	0.0329	0.0329	0.0329	0.0329
我们的(30个时代)	0.1197	0.0572	0.0345	0.0285	0.0278
我们的(30个时代, 经过预先训练)	0.1018	0.0499	0.0332	0.0287	0.0279
我们的(100个时期)	0.0786	0.0408	0.0299	0.0273	0.0271

Table2: Chamfer- L1Convergence.我们在不同的训练设置下评估我们在Thing32d数据集上的架构，并报告更高lod的收敛速度。

值得注意的是，FFN[43]和SIREN[41]似乎都完全失败了；这可能是因为他们都只能适应平滑距离场，而不能处理不连续和递推结构。质量量分析见图6。

收敛我们进行了实验来评估我们的架构的训练收敛速度。表2显示了我们的模型上Thing32的构建结果，我们完全训练了100个时代，训练了30个时代，训练了30个时代（图1）。我们发现，即使DeepSDF[36]和FFN[43]的训练时间约为45%，我们的架构也能转换迅速，并实现更好的重建，它们已经训练了整整100个时代的时间。从训练前的权重进行微调有助于较低的lod，但差异很小。我们的表示法迅速收敛到好的解。

## 4.2. 渲染性能

我们也评估了我们的原始结构的推理性能，无论有和没有我们的渲染算法。我们首先使用PyTorch [37]中基于简单的python的球体跟踪算法来评估性能，并在所有基线上进行相同的公平比较。对于我们表示的Python版本，我们将它们存储在一个密集的体素网格上，因为一个天真球跟踪器——而不是手手性。对于优化的实现，我们使用排序器的表现

使用 $\eta$ [37]、CUB[30]和CUDA实现。

表3显示了具有不同分辨率的涡轮m场景的帧时间。在这里，我们将确定为球面跟踪和标准化计算的cuda时间。#可见像素列显示模型在图像中占据的像素数量。我们可以看到，我们的朴素的PyTorch渲染器和稀疏优化的CUDA渲染器都比基线性能更好。特别是，稀疏帧时间比DeepSDF快100×以上，同时以更少的参数实现了更好的可视化质量。我们还注意到，我们的帧时间显著减少，而我们的天真渲染器的LOD显著减少，但我们的优化渲染器的帧时间较小。这是由于渲染的瓶颈不是在射线-八树相交中，这依赖于体素的数量，而是在MLP推断和杂项记忆I/O中。我们相信，通过缓存小的MLP解码器来最小化数据传输，仍然有很大的改进空间。尽管如此，较低的LODs仍然受益于较低的内存消耗和存储。

## 4.3. 泛化

我们现在证明，我们的表面提取机制可以推广到多重形，即使是在单重形上进行训练。这一点很重要，因为在[8,41]中为每个对象加载不同的权重会导致大量的内存移动，这是非常昂贵的。通过一般的表面提取机制，权重可以预加载，多分辨率体素可以按需流入。表4显示了关于Thing32的结果。DeepSDF[36], FFN [43]和我们的形状（过度适合）在每个形状上都是过度适合的。我们的（一般的）首先在斯坦福Lucy模型上覆盖了架构，修复了表面提取网络权重，并只训练了解析特性。我们认为我们的代表表现更好，即使是对大型网络的每个特定形状的例子。在最低的LOD，



分辨率	#可见像素	帧数时间 (ms) /改进系数							
		DeepSDF[36]	纤维素[43]	警报器[41]	尼[8]	我们的(N) / LOD4	我们的(N) / LOD6	我们的(S) / LOD4	我们的(S) / LOD6
640×480	94624	1693/57×	1058/35×	595/20×	342/ 11×	164/5×	315/ 11×	28	30/1×
1280×720	213937	4901/96×	2760/54×	1335/26×	407/8×	263/5×	459/9×	50	51/1×
1920×1080	481828	10843/ 119×	5702/62×	2946/32×	701/8×	473/5×	784/9×	93	91/1×

表3：渲染帧时间。在不同表示之间的运行时比较，其中(N)和(S)分别对应于我们的天真者和稀少者。我们在LOD6时将基线与我们(Sparse)进行了比较。#可见像素显示了基准场景(TurboSquid VMech)所占据的像素数，帧时测量光线追踪和表面法线计算。

Method/ L	倒角-L1↓				
	1	2	3	4	5
DeepSDF (overfi tper shape)[36]	0.0533	0.0533	0.0533	0.0533	0.0533
FFN (overfi tper shape)[43]	0.0322	0.0322	0.0322	0.0322	0.0322
我们的(过件形状)	0.0786	0.0408	0.0299	0.0273	0.0271
我们的(一般)	0.0613	0.0378	0.0297	0.0274	0.0272

表4：广义。我们评估了对Thing32的泛化。我们的(一般的)冻结了在单轴上预先训练的表面提取器的权重，并且只训练特征卷。即使是大型的过拟合网络，我们在高lod下也表现得更好。

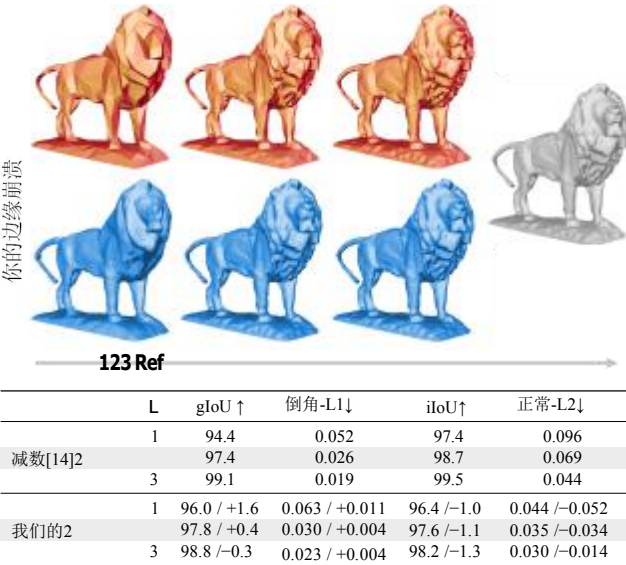


图7：与网格分解过程的比较。在较低的备忘录预算下，我们的模型能够比网格抽取更好地保持视觉细节，从较低的正常-12误差中可以看出。

表面提取器努力储存建造良好的表面，预期：越来越高水平和复杂的特点。

#### 4.4.几何图形简化

在这个激光环境中，我们评估我们的低边缘形式与经典网格抽取算法，在里比格[19]中的特定边缘崩溃[14]。我们比较了网格压缩而不是网格压缩

[1]，因为我们的模型也可以受益于压缩和网格，解码产生额外的r非时间间隔。我们首先评估我们的内存影响，其中M=(m+1) |V1： L最大|字节，其中m+1是特征特征

z阶曲线[33]，|V1： |最大|是八叉树大小。然后我们计算F=M/3，因为连通性可以是任意的。因此，我们选择辅助服务预算来增加他们的利益。

图7显示了来自西藏32的狮子雕像的结果。我们看到，随着记忆预算的减少，我们的方法的感知质量的变化就会增加，这可以从基于图像的归一化误差中得到证明。sdf可以容易表示平滑特征，而离散误差导致面部预算减少。我们的表示也可以平滑地混合LODsby构造，这是网格很难做的事情。

#### 5.限制和未来的工作

总之，我们引入了Neur al几何LOD，这是一种隐式3d形状的代表，它实现了最先进的几何重建质量，同时允许使用可接受的内存足迹进行实时渲染。我们的模型结合了一个小的表面提取神经网络和参数-八叉树数据结构，该结构编码地理测量和自然启用lod。与一种有效的分支跟踪算法相结合，该方法具有一定的计算性能和较高的表达能力。

我们的方法在很大程度上依赖于在训练过程中使用的点样本。因此，将我们的表示缩放到非常大的场景或真实的无体积几何是困难的。此外，我们不能轻易地动画或变形我们的几何使用传统的方法。让这些挑战忽略了未来搜索的方向。尽管如此，我们相信我们的模型在基于神经隐式函数的几何结构中向前迈进了一步，据我们所知，这是这种模型中第一个可以实时呈现和查询的表征。我们希望它将成为许多下游应用程序的重要组成部分，如场景构造、超精确的机器人整体规划、交互激活帐篷的创建等等。

致谢我们感谢让-弗朗索瓦·拉菲切、雪莉、谢凯文、乔纳森·格兰斯科格、亚历克斯·埃文斯和亚历克斯·毕在整个项目中进行的有趣的讨论。我们也感谢雅各布·蒙克伯格、彼得·雪莉、亚历山大、大卫·卢布克、约纳菲利翁和高俊提供的论文评论。

## 参考文献

- [1] Pierre Alliez和克雷格和戈斯曼。3个网格网格的最新进展。在几何模型的进步中，第3-26页。施普林格，2005年。8
- [2] 约翰·阿曼那提德。带锥体的光线追踪。ACMSig图形计算机图形学，18(3): 129-135,1984.5
- [3] Rohan查布拉、简·埃里克·伦森、埃迪·伊尔格、坦纳·施密特、朱利安·斯特劳布、史蒂文·洛夫格罗夫和理查德·纽科姆。深度局部形状：L获得局部SDF先验，进行详细的三维重建。欧洲人会议压缩。Vis., 2020.3
- [4] 天使X·张、托马斯·芬克豪瑟、列奥尼达·吉巴斯、汉拉汉、黄七星、李齐莫、瑞士、马诺利斯·萨瓦、宋秀兰、郝苏等。ShapeNet: An information-rich 3D model repository. arXiv preprint arXiv:1512.03012, 2015.5, 6
- 陈[5]志勤、安德烈、张浩。通过二进制空间划分生成紧凑的网格。在IEEE Conf. 压缩。Vis. 模式记录。，2020.3
- 陈[6]志勤、张浩勤。学习隐含领域为生成的形状建模。在IEEE Conf. 压缩。Vis. 模式。2019年6月。1, 3
- [7] 科西尼，穆罕默德·主席拉拉比，老√，Libor√。静态和动态三角形网格的感知度量值。发表在计算机图形论坛上，第32卷，第101-125页，2013.2页sa, and Kai Wang
- [8] 托马斯·戴维斯，德里克·诺鲁泽扎哈雷，和亚历克·雅各布森。论权重编码的神经隐式三维形状的有效性。arXiv预印本arXiv: 2009.09808, 2020.1、2、3、6、7、8
- [9] 博阳、凯尔·热诺娃、索罗什·雅兹达尼、布阿扬、杰弗里·辛顿和安德烈·塔利亚萨奇。可学习的曲线外部分解。在IEEE Conf. 压缩。Vis. 模式记录。，2020.3年6月
- [10] 月奇段、朱海东、何王、李毅、内瓦提、列奥尼达。吉巴斯群岛。Curriculum DeepSDF. Eur. 会议压缩。Vis., 2020.3
- [11] 亚历克斯·埃文斯。在动态场景上的全局照明的快速近似。在ACM签名图2006年课程中，签名图‘06页，第153-171页，2006.2页
- 莎拉F. 弗里斯肯，罗纳德·佩里，艾林P. 洛克伍德和图伊斯R. 琼斯。自适应采样距离场：计算机图形学形状的一般表示。本文介绍了第27届计算机图形学家和交互式技术年会，第249-254页，2000.2页
- [13] Jun Gao、温正陈、汤米翔、阿莱克·雅各布森、摩根·麦圭尔和桑贾·菲德勒。学习可变形的四面体网格进行3次重建。在Adv. 神经信息。过程。Syst., 2020.3
- [14] 迈克尔·加兰和保罗·s·赫克伯特。利用二次像差度量进行表面简化。在第24届计算机G图形和交互技术年会上的介绍中，符号图‘97，第209-216页，1997.2, 8
- [15] 凯尔·热诺娃，弗雷斯特科尔，阿苏德，亚伦萨纳，和托马斯·芬克豪瑟。三维形状的局部深度隐式函数。在IEEE Conf. 压缩。Vis. 模式记录。，2020.3年6月
- [16] Zekun Hao，哈达，诺亚，和贝隆吉。DualSDF：使用两级表示的语义形状操作。IEEE框架。压缩。Vis. 模式记录。，2020.6
- [17] John C. 哈特。球面追踪：隐式曲面抗锯齿光线追踪的几何方法。第12（10）：527-545, 1996.1、3、4
- [18] Hugues Hoppe. Progre ssivemeshes. 在第23届计算机图形和交互活动技术年会的会议记录中，符号图‘96，第99-108页，1996.2
- [19]：一个简单的C++几何处理库，2018年。[https:// libigl. 吉图布。io.8](https://libigl.github.io)
- [20] Chiyu 江，大卫，马卡迪亚，黄精卫，马蒂亚斯·尼斯纳，和芬克豪瑟。3个数据场景的局部隐式网格表示。在IEEE Conf. 压缩。Vis. 模式记录。，2020.3年6月
- [21] 越江，丹济，韩志忠，马提亚·茨威克。SDFDiff：用于三维形状优化的有符号距离场的可微渲染。在IEEE Conf. 压缩。Vis. Pat- tern 记录。，2020.3
- [22] 密歇根州的eal拉金和卡罗罗·沙利文。对动画人物的简单虚构作品的感知。在ACM符号图研讨会在图形和可视化中的应用中，第93-100页，2011.2
- [23] 彼得·林德斯滕伦和格雷格·特克。图像驱动简化阳离子。ACM跨。图表。，19(3): 204-241, 2000.2.2
- 刘[24]玲杰，顾嘉雄、林觉觉、蔡达成、神堡基督教。神经稀疏体素场。Adv. 神经信息。过程。Syst., 2020.2, 3
- [25] 刘少辉、张银达、彭宋友、施信、波勒菲、赵朋。DIST：渲染具有可微球面跟踪的深度多符号距离函数。在IEEE Conf. 压缩。Vis. 模式记录。，2020年。3
- [26] 威廉e洛伦森和哈维e克萊恩。标记立方体：一种高分辨率的三维曲面构造算法。在第14届计算机图形和交互技术年会上，符号图‘87，第163-169页，1987.2, 3页
- [27]，大卫·卢布克，马丁·雷迪，乔纳桑德。科恩，阿米塔布瓦什尼，本杰明沃森和罗伯特韦布纳。针对3D图形的详细信息级别。摩根·考夫曼出版公司，2002.2
- [28] 亚历山大先生，西里尔克拉辛，彼得雪莉，和莫尔·根·麦圭尔。一种有效的动态体素渲染交叉算法。《计算机图形学技术杂志》（JCGT），7(3): 66-81, 2018.4年9月
- [29] Ravish Mehra，周清南，杰里米·朗，真主安拉·谢弗，阿米古奇，和尼洛伊J. 米特拉。人的形状的抽象。ACM跨。图表。，28(5): 2009.2年12月1-10日
- [30] Duane 美林。CUB：一个扭曲宽、块宽和设备宽的GPU并行原语库，2017.5, 7
- [31] 拉尔斯·梅切德，迈克尔·奥克斯尔，迈克尔·尼迈耶，塞巴斯蒂安·诺沃津和安德烈亚斯·盖格。占用网络：在功能空间中学习三维重建。在IEEE Conf. 压缩。Vis. 模式重组。，2019年6月。1, 3

- [32]本·米尔登霍尔, 普拉图尔·斯里尼瓦桑, 马修·坦西克, 乔纳森·巴伦, 拉维·拉莫莫蒂, 和任吴。将场景表示为视点论文的神经辐射场。在欧元。会议压缩。Vis., 2020.2, 3
- [33] Guy M. 莫顿。一种面向计算机的大地测量数据库及其定向测序新技术。技术报告, 1966年。
- 8
- [34]迈克尔·尼迈耶, 拉斯梅切德, 迈克勒和安德烈亚斯·盖格。可微体积渲染: 学习隐式三维表示。在IEEE Conf. 压缩。Vis. 模式记录。 , 2020.3
- [35] Fakir S. 诺鲁丁和格雷格·特克。利用体积技术简化离子和修复多边形模型。  
《IEEE的可视化与计算学报》, 9(2): 191-205年, 4月。2003.6
- [36] Jeong 琼公园, 彼得弗洛伦斯, 朱利安斯特劳布, 理查德纽康姆, 和史蒂文洛夫格罗夫。DeepSDF: 学习连续有符号距离函数。在IEEE Conf. 压缩。Vis. 模式重组。 , 2019.1、2、3、4、5、6、7、8
- [37] 亚当·帕斯克, 山姆·格罗斯, 弗朗西斯科·马萨, 亚当·勒, 詹姆斯布拉德伯里、格雷戈里、特雷沃基林、泽明林、纳塔利吉梅尔辛、卢卡安提加等。PyTorch: 一个不稳定的、高性能的深度学习库。在Adv. 神经信息。过程。Syst., pages8026–8037, 2019.
- 7
- [38]俊齐藤、曾黄、夏敏、三岛茂、金泽安珠、李浩。像素像素功能, 用于高分辨率的人体数字化。在Int. 会议压缩。Vis., 第2304-2314页, 2019.3
- [39]分流齐藤, 汤斯西门, 茉莉萨拉吉, 和汉比尔乔。PIFuHD: 多级像素对齐的隐式功能, 用于高分辨率的三维人类数字化。在IEEE Conf. Com- put. Vis. 模式重组。 , 第84-93页, 2020.3页
- [40] 索拉布·辛格、唐、唐、周、唐、明实、杜、杜、杜、安、索菲恩布阿齐兹和森克斯金。深度隐式体积压缩。在CVPR, 2020.2
- [41]文森特西茨曼, 朱利安N.P.马特尔, 亚历山大W.伯格曼, 大卫B.林德尔, 和戈登韦茨斯坦。具有周期性激活函数的内隐神经表征。在Adv. 神经信息。过程。Syst., 2020.3, 5, 6, 7,8
- [42]文森特·西茨曼, 迈克尔·佐尔·弗和戈登·韦特·兹斯坦。场景表示网络: 连续的三维结构感知的神经场景表示。InAdv. 神经信息。过程。Syst., 2019.3
- [43] 马修坦西克, 普拉斯里尼瓦桑, 本米尔登霍尔, 萨拉弗里多维·基尔, 尼廷拉加万, 乌特卡什辛哈尔, 拉维拉·马莫蒂, 乔纳桑·巴伦, 和伦格。让网络在低维域中学习高频函数。Adv. 神经信息。过程。Syst., 2020.3, 5,6,7,8
- 唐[44]、窦明松、彼得·林肯、菲利普·大卫森、凯文郭、乔纳坦泰勒、肖恩·法内罗、阿达什科德尔、索菲恩、布阿齐兹、沙赫拉米扎迪和安德里亚·塔利亚萨奇。实时压缩和流。在ACM翻译。《图》, 2018.2
- [45]埃德加·特雷茨克, 阿尤什·特瓦里, 弗拉迪斯拉夫·戈利亚尼克, 迈克尔·佐尔, 卡斯滕·斯托尔和克里斯蒂安·西奥奥巴特。补丁: 基于补丁的通用深度嵌入它3形状抑制。欧洲人会议压缩。Vis., 2020.3
- [46]兰斯威廉姆斯。金字塔参数。发表在第十届计算机图形和实际技术年会上, 第1-11页, 1983.2页
- [47] Lior Yariv、约尼·卡斯滕、德莫兰博士、梅拉夫·加伦、马坦·阿茨蒙、罗恩巴斯里和亚隆利普曼。具有隐式照明和材质的多视点表面重建。Adv. 神经信息。过程。Syst., 2020.3
- [48]和亚历克·雅各布森。Thing10K: 一个包含10,000个3d打印模型的数据集。arXiv预印本  
arXiv:1605.04797, 2016.5, 6