

# 学习神经隐式曲面 通过卷渲染进行多视图重建

王鹏, 刘灵杰<sup>\*,†</sup>, 刘元, 基督教神! 王文平<sup>°, \*</sup>

<sup>†</sup>香港大学+马克斯·普朗克信息学研究所

○德州农工大学

t{pwang3, yliu, taku}@cs.hku.hk {lliu, theobalt}@mpi-  
inf.mpg.de

Owenping@tamu.edu

## 摘要

我们提出了一种新的神经表面重建方法。称为NeuS。用于从二维图像输入中高保真地重建物体和场景。现有的神经表面重建方法, 如DVR [Niemeyer等人。2020]和IDR [Yariv等人。2020], 需要前景面具作为监督。很容易陷入局部极小值。因此, 很难重建具有严重自遮挡或薄结构的物体。同时, 最近的新观点合成的神经方法, 如NeRF[米尔登霍尔等人。2020]及其变体。使用体积渲染来产生具有鲁棒性的神经场景表示, 即使是高度复杂的对象。然而, 从学习到的隐式表示中提取高质量的曲面是困难的, 因为在表示中没有足够的曲面约束。在NeuS中, 我们提出将曲面表示为有符号距离函数(SDF)的零水平集, 并开发了一种新的体积绘制方法来训练神经SDF表示。我们观察到, 传统的体积绘制方法会导致表面重建产生固有的几何误差(i.e.bias), 因此提出了一种在一阶近似中无偏差的新公式, 从而在没有掩模监督的情况下实现更精确的表面重建。在DTU数据集和BlendedMVS数据集上的实验表明, NeuS在高质量的表面重建方面优于最先进的水平, 特别是对于具有复杂结构和自遮挡的对象和场景。

## 1介绍

从多视点图像中重建表面是计算机视觉和计算机图形学中的一个基本问题。利用神经隐式表示的三维重建最近成为经典重建方法[37.8.2]的替代方法, 因为其高重建质量和重建复杂物体的潜力是经典方法, 如非兰伯表面和薄结构。最近的研究成果将曲面表示为有符号的距离函数(SDF)[49.52.17.23或入住率。来训练他们的神经模型。这些方法使用可微表面渲染方法将三维对象渲染成图像, 并将其与输入图像进行比较以进行监督。例如, IDR[49]产生了令人印象深刻的重建结果, 但它不能重建具有复杂结构而导致深度突变的物体。造成这种限制的原因是, IDR中使用的曲面渲染方法只考虑每条光线的单个曲面交点。因此, 梯度只存在于这个单点上, 这对于有效的反向传播过于局部性, 当图像的深度发生突变时, 优化会陷入较差的局部最小值。此外, 对象

\*相应的作者。

第35届神经信息处理系统会议(NeurIPS 2021)。

掩模作为收敛到一个有效表面的监督。如图1 (a) top. with所示，由于孔洞引起的根本深度变化，神经网络会错误地预测前表面附近的点为蓝色，无法找到远后的蓝色表面。图1 (b) 中的实际测试示例表明，IDR不能正确地重建深度突变的边缘附近的表面。

最近，NeRF[29]及其变体已经探索了使用体积渲染方法来学习体积辐射场的新视图合成。这种体积渲染方法沿着每条光线对多个点进行采样，并对采样点的颜色进行组合，以产生用于训练目的的输出像素颜色。体积渲染方法的优点是它可以处理突然的深度变化，因为它考虑了沿光线的多个点，因此所有的样本点。无论是靠近表面还是在远表面上，都会产生用于反向传播的梯度信号。例如，参考图1(a)底部，当发现近表面（黄色）与输入图像的颜色不一致时，体积渲染方法能够训练网络找到远背表面，产生正确的场景表示。然而，由于它是用于新的视图合成，而不是表面重建。NeRF只学习一个体积密度场，从中很难提取出一个高质量的表面。图1 (b) 显示了NeRF学习的密度场水平集表面的表面。虽然表面正确地解释了突变的突变，但在一些平面区域包含明显的噪声。



图1: (a) 曲面渲染和掩模的小示意图。(b) 的 1 视图的 4 个 竹制种植机，在种植机的顶部有遮挡。与现有的方法相比，我们的方法可以处理遮挡，并实现更好的重建质量。

在这项工作中，我们提出了一种新的神经渲染方案，称为NeuS，用于多视图表面重建。NeuS使用有符号距离计算（SDF）进行曲面表示，并使用一种新的体积渲染方案来学习神经SDF表示。具体来说，通过引入密度分布引起的自卫队，我们可以应用体积渲染方法学习一个隐式自卫队表示，因此最好的worlds. i. e. an准确的表面表示使用神经自卫队模型和健壮的网络训练的突变深度变化使体积呈现。注意，简单地将标准体积绘制方法应用于与SDF相关的密度，就会在重建表面中产生可识别的偏差（i. e. inherent几何误差）。这是一个新的和重要的观察结果，我们稍后将详细阐述。因此，我们提出了一种新的体积渲染方案，以确保在SDF的一阶近似下的无偏表面重建。在DTU数据集和BlendedMVS数据集上的实验表明，即使没有前景掩模作为监督，DTU也能够重建具有严重遮挡和精细结构的复杂三维物体和场景。在重建质量方面，它优于最先进的神经场景表示方法，即IDR[49]和NeRF[29]。

## 2相关工程

传统的多视点表面和体积重建传统的多视点三维重建方法可以大致分为两类：基于点和表面的重建[2, 8, 9, 37]和体积重建[6, 3, 38]。基于点和表面的重建方法利用图像间光度一致性[8]估计每个像素的深度图，然后将深度图融合成一个全局致密点[26, 51]。表面重建通常采用筛选泊松表面重建[16]等方法进行后处理。这个

重构质量在很大程度上依赖于对应匹配的质量，而没有丰富纹理的对象的对应匹配困难往往会导致重构结果中严重的伪影和部分缺失。另外，体积重建方法通过从多视图图像中估计体素网格中的占用率和颜色，并评估每个体素的颜色一致性，来规避显式对应匹配的困难。由于可实现的体素分辨率有限，这些方法不能达到较高的精度。

**神经隐式表示。**一些方法通过引入归纳偏差，在深度学习框架中加强了三维理解。这些归纳偏差可以是显式表示，如体素网格 [13, 5, 47]. point 云 [7, 25, 19]. meshes [44, 46, 14]. and 隐式表示。神经网络编码的隐式表示由于是连续的，可以实现较高的空间分辨率，近年来得到了广泛的关注。该表示方法已成功应用于形状表示方法 [27, 28, 32, 4, 1, 10, 50, 33]. novel 视图合成 [40, 24, 15, 29, 22, 34, 35, 43, 39] 和多视图三维重建 [49, 30, 17, 12, 23]。

我们的工作主要是通过经典的渲染技术从二维图像中学习编码在三维空间中的几何形状和外观的隐式神经表示。在这个范围内有限。可以根据绘制技术、used, i. e. surface 绘制方法和卷绘制方法对相关工作进行大致分类。基于表面渲染的方法 [30]。17, 49, 23] 假设光线的颜色只依赖于光线与场景几何图形的交点的颜色，这使得梯度只反向传播到交点附近的局部区域。因此，这种方法难以重建具有严重自遮挡和突然深度变化的复杂物体。此外，它们通常需要物体面具作为监督。相反，我们的方法在没有挑战性的情况下有很好的效果。

基于体渲染的方法，例如通过沿每条光线合成采样点的颜色来合成图像。如在引言中所解释的，它可以处理突然的深度变化，并合成高质量的图像。然而，由于基于密度的场景表示对其水平集缺乏足够的约束，从学习到的隐式场中提取高保真表面是困难的。相比之下，我们的方法结合了基于表面渲染和基于体积绘制的方法的优点，将场景空间约束为一个有符号的距离函数，但应用体积渲染来训练具有鲁棒性的表示。用户 [31]，一个并发工作，也通过体积渲染学习隐式曲面。在优化过程中，通过缩小体积渲染的样本区域来提高重构质量。我们的方法不同于精灵，精灵代表表面占用值，而我们的方法代表场景的自卫队，因此可以自然地提取表面的零水平集，产生更好的重建精度比精灵，稍后将在实验部分看到。

### 3方法

给定一组三维物体的姿态图像，我们的目标是重建它的表面 $S$ 。该曲面由一个神经隐式SDF的零水平集表示。为了学习神经网络的权值，我们开发了一种新的体积渲染方法，从隐式SDF渲染图像，并最小化渲染图像和输入图像之间的差异。这种体积渲染方法确保了NeuS中对复杂结构对象的鲁棒优化。

#### 3.1渲染程序

**场景表示。**NeuS对象重建的场景由两个功能表示： $f: \mathbb{R}^3 \rightarrow \mathbb{R}$  映射空间位置 $x \in \mathbb{R}^3$ 到对象的符号距离，和 $c: \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow \mathbb{R}$  编码的颜色与点 $x \in \mathbb{R}^3$ 和观看方向 $v \in \mathbb{S}^2$ 。这两个函数都是由多层感知器（MLP）编码的。该对象的曲面 $S$ 由其SDF的零级集表示，即，

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

为了应用体积绘制方法对SDF网络进行训练。我们首先引入一个概率密度函数 $o$ ， $(f(x))$  . called S-densiry, where  $f(x)$ .  $x \in \mathbb{R}^3$ . 是有符号距离

函数和  $\sigma(x) = se^{-sx} / (1 + e^{-sx})^2$ ，commonly称为逻辑密度分布。是  $\phi(x) = (1 + e^{-sx})^{-1}$ ，i.e.  $\sigma(x) = \phi'(x)$ 。原则上， $\phi(x)$  可以是任何以0为中心的单峰（i.e. bell-shaped）密度分布；这里我们选择逻辑密度分布是因为它的计算方便。请注意， $\phi(x)$  的标准差由  $1/s$  给出，这也是一个可训练的参数，即。当网络训练收敛时， $1/s$  趋近于零。

直观地说，NeuS的主要思想是，在  $s$ -密度场  $\phi_s(f(x))$  的帮助下，使用体积渲染来训练SDF网络，只有二维输入图像作为监督。在基于这种监督成功最小化损失函数后，网络编码的SDF的零级集有望代表一个精确重建的表面  $s$ ，其诱导的  $s$  密度  $\sigma_s(f(x))$  假设在表面附近有明显的高值。

**渲染。**为了学习神经SDF和色场的参数。我们建议用一个体积渲染方案来渲染所提出的SDF表示中的图像。给定一个像素，我们将从这个像素发出的射线表示为  $\{p(t) = o + tv \mid t \geq 0\}$ ，where  $o$  是相机的中心， $v$  是射线的单位方向向量。我们沿着光线积累颜色

$$C(o, v) = \int_0^{+\infty} u(t) c(p(t), v) dt. \quad (2)$$

其中  $C(o, v)$  是该像素的输出颜色， $u(t)$  为点  $p(t)$  的权重。和  $C(p(t), v)$  为沿观察方向  $v$  的点  $p$  处的颜色。

**重量功能要求。**从二维图像中学习精确的SDF表示的关键是在输出颜色之间建立适当的连接，而SDF，i.e.  $t_o$  根据场景的SDF  $f$  在光线上推导出适当的权重函数  $x(t)$ 。在下面，我们列出了对权重函数  $w(t)$  的要求。

1. **无偏** 给定一个相机射线  $p(t)$ ， $w(t)$  在曲面交点  $p$  处获得局部最大值 ( $t^*$ )，i.e. with  $f(p(t^*)) = 0$ ，that为。点  $p(t^*)$  在SDF ( $x$ ) 的零级集上。
2. **感知空间。** 给定任意两个深度值满足  $f$  和  $t$  ( $t_o = f(t_1)$ )。  $u(t_o) > 0$ ， $\&(t_1) > 0$ ，and  $t_o < t_1$ ，there is  $u(t_o) > u(t_1)$ 。也就是说，当两个点具有相同的SDF值（因此相同的SDF诱导的  $s$  密度值）。靠近视点的点对最终输出颜色的贡献应该比其他点有更大的贡献。

无偏权值函数  $w(t)$  保证了相机射线与零级SDF集的交集对像素颜色的贡献最大。遮挡感知特性确保当光线连续通过多个表面时，渲染过程将正确地使用最接近照相机的表面的颜色来计算输出的颜色。

接下来，我们将首先介绍一种简单的方法来定义权重函数  $u(t)$ ，它直接使用体呈现的标准管道。并在引入我们的  $\&(t)$  结构之前，解释为什么它不适合重建。

**朴素的解决方案。**使重量

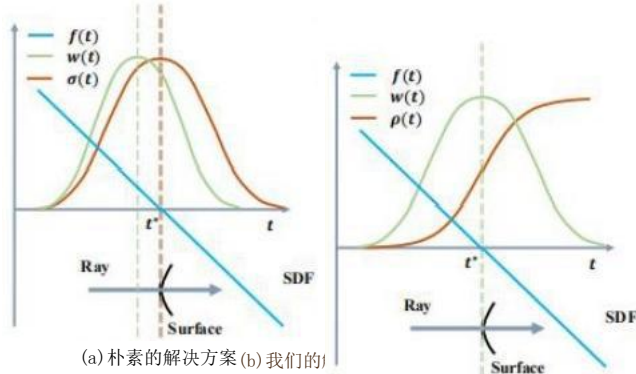


图2: 朴素解的(a)权重偏差示意图，(b)是我们的解中定义的权值函数，它在SDF的一阶近似中是无偏的。

功能遮挡感知，一个自然的解决方案是基于标准的体积渲染公式—它定义了值[29]的权重函数

$$u(t) = T(t) o(t). \quad (3)$$

其中， $o(t)$  是经典体积渲染中所谓的体积密度和  $T(t) = \exp(-\int_0^t o(u) du)$  这里表示沿射线的累积透射率。采用斯坦

标准体积密度公式[29]。这里的 $\sigma(t)$ 被设置为等于s-密度值，即。 $\sigma(t) = \sigma_s(f(p(t)))$ 和权重函数 $w(t)$ 由Eqn. 3. Although计算，得到的权重函数是遮挡感知的，由于它引入了重构曲面的固有误差而有偏倚。如图2(a)所示，权重函数 $w(t)$ 在光线到达表面点 $p(t)$ 之前的某一点达到局部最大值，满足 $f(p(t^*)) = 0$ 。这一事实将在补充材料中得到证明。

我们的解决方案。为了引入我们的解决方案，我们首先引入了一种直接的方法来构造一个无偏权重函数，它直接使用归一化的s密度作为权重

$$w(t) = \frac{\sigma_s(f(p(t)))}{\int_0^{+\infty} \sigma_s(f(p(u))) du} \quad (4)$$

这种权重函数的构造是无偏的，但没有遮挡感知。例如，如果光线穿透两个表面。SDF函数 $f$ 将在光线上有两个零点，这导致权重函数 $w(t)$ 上出现两个峰值，由此得到的权重函数将相同地混合两个表面的颜色，而不考虑遮挡。

为此，现在我们将设计一个权重函数 $w(t)$ ，它在SDF的一阶近似中既是遮挡感知的，又是无偏的。为了确保权重函数 $w(t)$ 的遮挡感知特性。我们仍然将遵循Eqn. 3. However的基本框架，不同于传统的处理，我们以一种新的方式从s密度定义我们的函数 $w(t)$ 。我们首先定义一个不透明的密度函数 $p(t)$ ，它是标准体积渲染中的体积密度 $\sigma$ 的对应物。然后我们计算新的权重函数 $w(t)$  by

(5)

我们如何推导出不透明的密度 $p$ 。我们首先考虑一个简单的理想情况，其中只有一个曲面的交点，并且曲面只是一个平面，接近无限远处的相机。自Eqn. 4以来。在这个假设下，4确实满足上述要求。我们推导出与Eqn. 4的权重定义对应的底层不透明密度。4. 使用卷渲染的框架。然后将这种不透明密度推广到多表面交点的一般情况。

具体来说，在单平面相交的简单情况下，很容易看到有符号的距离函数 $f(p(t))$ 是 $-|\cos(\theta)| \cdot (t - t^*)$ ，其中 $f(p(t^*)) = 0$ 和 $\theta$ 是视图方向 $v$ 和向外表面法向量 $n$ 之间的夹角。因为假设表面是一个平面。 $|\cos(\theta)|$ 是一个常数。它来自于Eqn. 4个

$$w(t) = \lim_{t^* \rightarrow +\infty} \frac{\sigma_s(f(p(t)))}{\int_0^{+\infty} \sigma_s(f(p(u))) du} \quad (6)$$

$$\begin{aligned} &= \lim_{t^* \rightarrow +\infty} \frac{\sigma_s(f(p(t)))}{\int_0^{+\infty} \sigma_s(-|\cos(\theta)|(u - t^*)) du} \\ &\equiv \lim_{t^* \rightarrow +\infty} \frac{|\cos(\theta)| \sigma_s(f(p(t)))}{\int_0^{+\infty} |\cos(\theta)| \sigma_s(-|u - t^*|) du} \end{aligned}$$

回想一下，体积渲染的w框架是由 $w(t)$ 给出的

$T(t) = \exp(-\int_0^t p(u) du) T(t) p(t)$  式中表示累计透光率。因此，推导出 $p(t)$ 。我们有

$$T'(t) p(t) = -\frac{dT(t)}{dt} \quad \text{注}$$

因为 $T(t) = \exp(-\int_0^t p(u) du)$  it is 这很容易验证

意同时整合到两

$|\cos(\theta)| \sigma_s(f(p(t))) = p(t)$  由此可以得出：

者中

$$\begin{aligned} T(t) p(t) &= -\frac{dT(t)}{dt} \\ \frac{dT(t)}{dt} &= -\frac{dT(t)}{dt} \sigma_s(f(p(t))) \end{aligned}$$

这个方程的侧边产生了结果

$$T(t) = \Phi_s(f(\mathbf{p}(t))). \quad (8)$$

取对数，然后区分两边。我们有

$$\begin{aligned} \int_0^t \rho(u) du &= -\ln(\Phi_s(f(\mathbf{p}(t)))) \\ \Rightarrow \rho(t) &= \frac{-\frac{d\Phi_s}{dt}(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))}. \end{aligned} \quad (9)$$

这是在单平面相交的理想情况下，不透明密度 $\rho(t)$ 的公式。由 $\rho(t)$ 诱导的权重函数 $w(t)$ 如图2(b)所示。现在我们将不透明密度推广到沿射线 $\mathbf{p}(t)$ 有 $r$ 截面的一般情况。在这个case中，随着SDF值的增加， $-\frac{d\Phi_s}{dt}(f(\mathbf{p}(t)))$ 在射线段上变为负值。因此，我们就把它夹了起来，以确保 $\rho$ 的值始终是非负的。这在一般情况下给出了以下不透明的密度函数 $\rho(t)$ 。

$$\rho(t) = \max \left( \frac{-\frac{d\Phi_s}{dt}(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))}, 0 \right) \quad (10)$$

基于这个方程，权重函数 $w(t)$ 可以用标准体积绘制来计算，如Eqn. 5。多曲面相交的情况如图3所示。

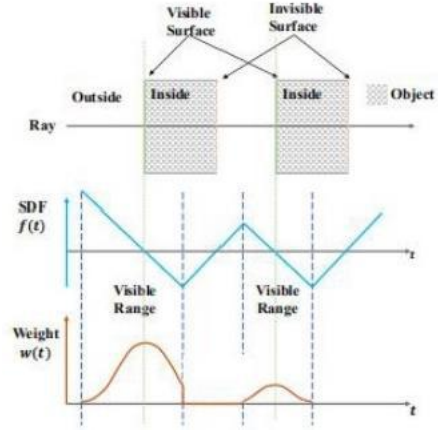


图3: 重量说明  
在发生多重曲面相交时的分布。

下面的定理说明，在一般情况下(即。

包括单个曲面交点和多个曲面交点)由Eqn定义的权函数。10和等式。5在SDF的一阶近似中是无偏的。证明在补充材料中提供。

**定理1** 假设一个光滑曲面 $S$ 由有符号距离的零能级集定义的

函数 $f(x)=0$ 。一个光线 $\mathbf{p}(t)$ 从外到内进入表面在 $\mathbf{p}(t^*)$ 处的交点。也就是， $f(\mathbf{p}(t^*))=0$ ，存在一个区间 $t^* \in [t_r, t_l]$ 和 $f(\mathbf{p}(t))$ 在中呈单调递减趋势。假设在这个局部区间内 $[t_r, t_l]$ ，the表面可以用一个惊人的小平面的小块来切向近似，即。 $\forall t$ 是， $f(\mathbf{p}(t))$ 是负的。然后，得到了由Eqn和Eqn10计算出的权值函数 $w(t)$ 。10. 5它的最大值为 $w(t^*)$ 。

离散化。为了获得不透明度和权值函数的离散对应物，我们采用了与NeRF [29]相同的近似方案。该方案沿射线采样 $n$ 个点 $\{t_i; i=1, \dots, n, t_{i+1}\}$ ，计算射线的近似像素颜色为

$$C = \sum_{i=1}^n T_i \alpha_i c_i. \quad (11)$$

其中 $T$ 为由所定义的离散累积透射率 $T_i = \prod_{j=1}^{i-1} (1 - \alpha_j)$ 和不透明度值由

是离散的

$$\alpha_i = 1 - \exp \left( - \int_{t_i}^{t_{i+1}} \rho(t) dt \right) \quad (12)$$

这可以进一步证明是

$$\alpha_i = \max \left( \frac{\Phi_s(f(\mathbf{p}(t_i))) - \Phi_s(f(\mathbf{p}(t_{i+1})))}{\Phi_s(f(\mathbf{p}(t_i)))}, 0 \right) \quad (13)$$

在补充材料中给出了关于 $\alpha_i$ 的这个公式的详细推导。



### 3.2 培训

为了训练NeuS，我们最小化渲染颜色和地面真实颜色之间的差异，没有任何3D监督。除了颜色之外，我们还可以利用面具进行监督。

具体来说，我们通过在三维世界空间 $P = \{C_x\}$ 中随机采样一批像素及其相应的射线来优化我们的神经网络和逆标准差 $s$ 。其中 $C_g$ 是其像素颜色， $M_g \in \{0, 1\}$ 是其可选掩模值。从每次迭代中的图像。我们假设点采样大小为 $n$ ，批处理大小为 $m$ 。损失函数的定义为

$$= C_{lor} + \lambda C_{reg} + 3C_{mask}. \quad (14)$$

颜色损失 $c$ 颜色被定义为

$$\mathcal{L}_{color} = \frac{1}{m} \sum_k \mathcal{R}(\hat{C}_k, C_k). \quad (15)$$

与IDR[49]，我们经验选择 $R$ 作为L1损失相同，在我们的观察中对异常值具有鲁棒性，在训练中稳定。

我们在采样点上添加一个元素项[10]，以正则化 $f_o$ 的SDF

$$\mathcal{L}_{reg} = \frac{1}{nm} \sum_{k,i} (\|\nabla f(\hat{\mathbf{p}}_{k,i})\|_2 - 1)^2. \quad (16)$$

可选的掩模丢失 $C_{mask}$ 被定义为

$$C_{mask} = \text{BCE}(\mathbf{M}_x, 0), \quad (17)$$

$\hat{O}_k = \sum_{i=1}^n T_{k,i} \alpha_k$  其中 $t$ 是沿相机射线的权值之和，BCE是二值交叉熵损失。

分层抽样。在这项工作中，我们遵循了与NeRF [29]中类似的分层抽样策略。我们首先对射线上的点进行均匀采样，然后在粗概率估计的基础上迭代地进行重要采样。不同之处在于。不像NeRF，同时优化粗网络和精细网络。我们只保持一个网络，其中粗采样的概率是基于固定标准差的 $s$ 密度 $os(f(x))$ 计算的，而精细采样的概率是基于学习到的 $o$ ， $(f(x))$ 计算的。分层抽样策略的细节在补充材料中提供。

## 4 实验

### 4.1 实验性设置

数据集。为了评估我们的方法和基线方法，我们使用了来自DTU数据集[11]的15个场景，与IDR[49]使用的场景相同，具有多种材料、外观和几何形状，包括具有挑战性的重建算法案例，如非兰伯表面和薄结构。每个场景包含49或64张图像，图像分辨率为 $1600 \times 1200$ 。每个场景都用IDR [49]提供的前景面具进行测试。我们进一步测试了来自BlendedMVS数据集[48]（CC-4许可）的低分辨率集的7个具有挑战性的场景。每个场景有31-143张图像，为 $768 \times 576$ 像素，掩模由BlendedMVS数据集提供。我们进一步用32个输入图像捕获了两个薄物体，以测试我们的薄结构重建方法。

基线(1)最先进的表面渲染方法-IDR [49]：IDR可以重建高质量的表面，但需要前景掩模作为监督：由于IDR比另一种基于表面渲染的方法-DVR[30]具有更好的质量，因此我们没有与DVR进行比较。(2)最先进的体积渲染方法-NeRF[29]：我们使用25的阈值从学习到的密度场中提取网格。我们在补充材料中验证了这一选择。(3)一种广泛使用的经典MVS方法COLMAP[37]：利用筛选泊松曲面重建[16]从COLMAP的输出点云重建网格。(4)将曲面渲染和体积渲染与占用字段作为场景表示的并发工作-统一[31]。关于基线方法的更多细节包括在补充材料中。

**实施细节。**我们假设感兴趣的重复是在一个单位范围内。我们每批采样512条光线，并在一个NVIDIA RTX2080Ti GPU上对模型进行300小时（w/掩码设置）和16小时（w/o掩码设置）的迭代训练。对于w/o掩码的设置，我们通过NeRF++[53]对背景进行建模。我们的网络架构和初始化方案类似于IDR [49]。更多关于网络体系结构和训练参数的细节可以在补充材料中找到。

#### 4.2比较

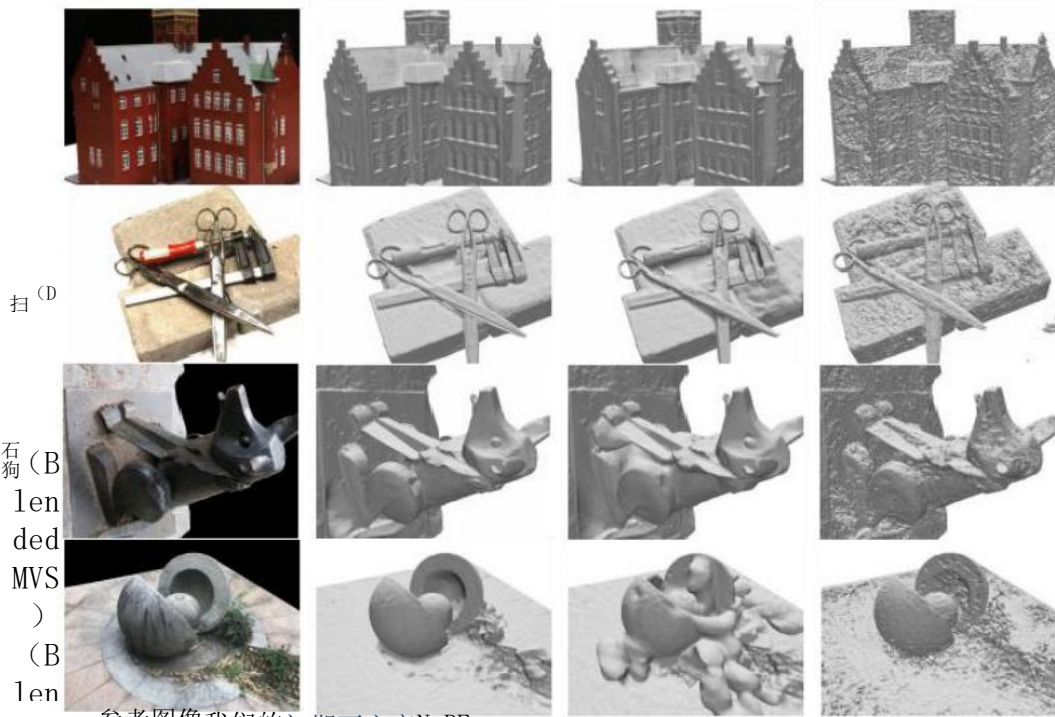
斯坎德	w/mask			无掩模			
	DR	NeRF	ours	COLMAP	eR	吉尼苏尔	在
an24	1.63	8	1.84	C8	1.90	1.2	
can3?	1.87	2.39		2.05	1.60	1.36	1.37
scan4	0.63	1.79	0.56	0.75	1.85	1.72	0.95
st an55	1.48	0.6n	0.37	1.22	0.58	0.44	0.43
	1.04	1.79	1.13	1.79	2.28	1.35	1.10
snh	0.79	1.44	0.59	1.58	1.27	0.79	0.65
	0.77	1.50	0.60	1.02	1.47	0.80	0.57
下沉	1.33	1.20	1.4	3.05	1.67	1.49	1.48
扫描	1.16	1.96	0.95	1.40	2.0	1.37	1.09
97 5an	0.76	1.27	0.78	2.05	5	0.89	0.83
! 0s	0.67	1.44	0.52	1.00	1.07	0.59	0.52
an!	0.90	2.61	1.43	1.32	0.88	1.47	1.20
三明治	0.42	1.0H	0.36	0.49	2.53	0.46	0.35
kcan!1	0.51	1.13	0.45	0.78	1.06	0.59	0.49
4	0.53	0.99	0.45	1.17	1.15	0.62	0.5
san11					0.96		
a							
scan12							
2							
意思	0.0	1.54		1.56	1.49	1.02	0.84

表1: 对DTU数据集的定量评价。COLMAP的结果是通过修剪=0来实现的。

我们在两种情况下进行了比较。

有口罩监督（无口罩），无口罩监督（无口罩）。我们用与单神经元[31]和IDR [49]相同的方法，用倒角距离来测量重建质量，并在表1中报告了分数。结果表明，我们的方法在DTU数据集上的w/和w/o掩模设置上都优于基线方法。请注意，在w/掩码设置、NeRF和无掩码设置中报告的IDR分数来自IDR[49]和单神经元[31]。

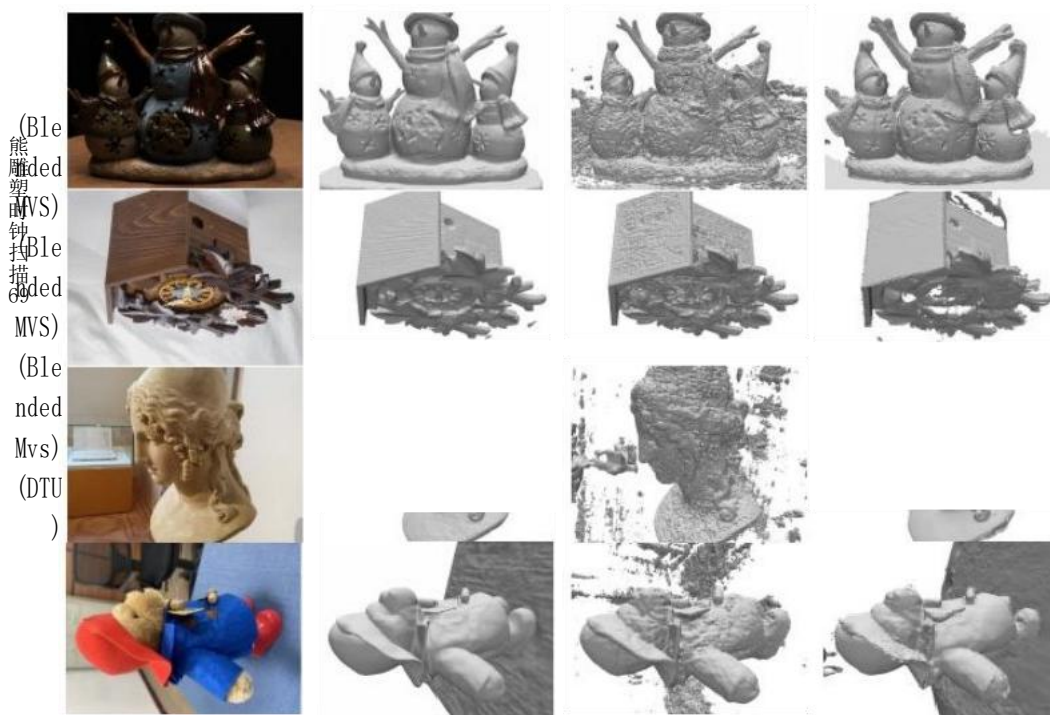
我们对DTU数据集和BlendedMVS数据集进行了定性比较，分别有/屏蔽和无屏蔽，分别在图4和图5中。如图4所示的w/掩码的设置。IDR在Scan37（DTu）中重建薄金属零件的性能有限，由于表面渲染的局部优化过程，无法处理石材（BlendedMVS）的突然深度变化。由于体积密度场对其三维几何没有足够的约束，因此提取的NeRF网格有噪声。关于w/o掩模设置，我们在图5中我们将我们的方法与w/o掩模设置的NeRF和COLMAP进行了直观比较，这表明我们重建的表面比基线更保真度。我们进一步展示了一个比较



参考图像我们的初期死亡率NeRF

图4: 表面重建与口罩监督的比较。





参考图像我们的网络COLMAP

图5：无口罩监督的表面重建的比较。

在两个例子中。请注意，我们使用联合用户用户报告他们的论文的定性结果进行比较。我们的方法更适合于深度突变对象。更多的定性图像包括在补充材料中。

### 4.3分析

#### 消融研究。

为了评估权重计算的效果，我们测试了Sec. 3.1：(a)朴素解中描述的三种不同的权重结构。(b)直接结构，如公式中所示。4. (e)全模型。如图6所示，在倒角距离方面，朴素解的定量结果比我们的权重选择(e)要差。这是因为它对表面重建引入了一种偏差。如果使用直接施工，就会有严重的文物。



(a) 朴素解决方案 (b) 直接解决方案 (c) 和首席 (d) 和地理信息。 (e) 完整的模型参考图像

1.49 4.45 0.64 0.62 0.59 倒角的距离

1.75 44.34 88.94 6.19 0.93 MAE

图6：消融研究。我们展示了定性的结果，并报告了在倒角距离和MAE（平均绝对误差）之间的定量指标。

我们还研究了二元正则化[10]和几何初始化[1]的影响。不进行正则化或几何初始化，倒角距离的结果与全模型相当。然而，它们都不能正确地输出一个有符号的距离函数。这可以由SDF预测和相应的地面真实SDF之间的MAE（平均绝对误差）来表示。如图6的底线所示。MAE被计算在物体的边界球体中的均匀采样点上。定性结果

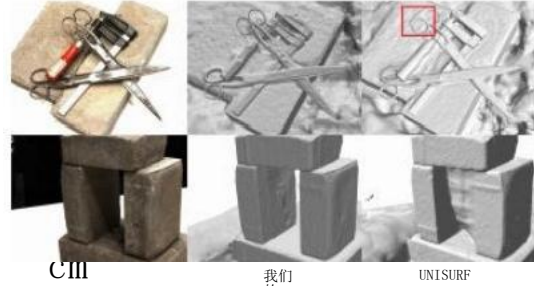
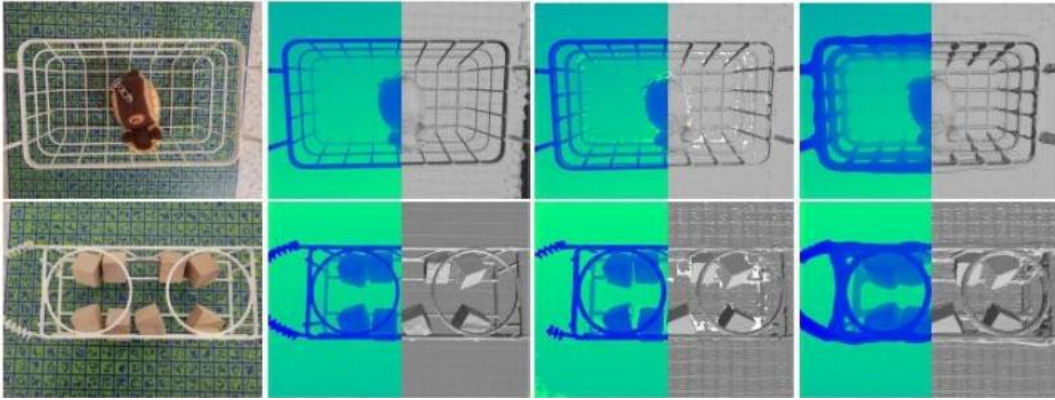


图7：与UNISURF的视觉比较。

SDF预测在补充材料中提供。

**薄结构。**此外，我们还展示了用32张输入图像对两个具有挑战性的薄物体的结果。对象下纹理丰富的平面用于相机校准。如图8所示，我们的方法能够准确地重建这些薄结构，特别是在深度突变的边缘。此外，与仅针对高质量薄结构重建的[41, 20, 45, 21]方法不同，我们的方法可以处理具有薄结构和一般对象混合的场景。



参考1mag (a) Ours (b) COLMAPtrim=10 (c) COLMAPPrim=7

图8：场景与薄结构对象的比较。左半部分是深度图，右半部分是重建的曲面。

## 5结论

我们提出了NeuS，一种新的多视图表面重建方法，将三维曲面表示为神经SDF，并开发了一种新的体积渲染方法，用于训练隐式SDF表示。NeuS可以产生高质量的重建，并成功地重建具有严重遮挡和复杂结构的物体。它在质量上和数量上都优于最先进的技术。我们的方法的一个局限性是，虽然我们的方法不严重依赖于纹理特征的对应匹配，但对于无纹理对象的性能仍然会下降（我们在补充材料中展示了失败的情况）。此外，NeuS只有一个单一的尺度参数，用于模拟所有空间位置的概率分布的标准偏差。因此，未来一个有趣的研究课题是对不同空间位置的不同方差的概率进行建模，并优化场景表示。这取决于不同的局部几何特征。负面的社会影响：像许多其他基于学习的工作一样，我们的方法需要大量的计算资源进行网络训练，这可能是全球气候变化的问题。

## 认股权证

我们感谢迈克尔·奥克斯尔提供了联合国基金会的结果。克里斯蒂安·西奥堡特得到了ERC合并者拨款770784的支持。刘灵杰获得梅特纳博士后资助。计算资源主要由HKU GPU Farm提供。

## 参考文献

- [1]Matan Atzmon和雅龙·利普曼。Sal: 从原始数据中学习形状的符号无关性。在 *IEEE/CVF计算机视觉和模式识别会议论文集* 2565–2574. 2020.
- [2]康纳利巴恩斯。伊莱谢克特曼。亚当·芬克尔斯坦。和丹·戈德曼。补丁匹配: 一种用于结构图像编辑的随机对应算法。ACM跨。图.. 28(3): 24. 2009.
- [3]Adrian Broadhurst. 汤姆·德拉蒙德和罗伯托·奇波拉。一个概率框架空间雕刻。在 *第八届IEEE国际计算机视觉会议的会议记录中*。ICCV2001, 第1卷。第388–393页。电器和电子工程师学会2001.
- [4]Z. Chen和H. 张。学习生成式形状建模的隐式领域。2019年 *IEEE/CVF计算机视觉和模式识别 (CVPR) 会议*。第5932–5941页。2019.
- [5]克里斯托弗B乔伊。许丹飞。Jun Young格瓦克。陈凯文。和西尔维奥·萨瓦雷斯。3d-r2n2: 一种用于单视图和多视图三维物体重建的统一方法。在欧洲会议上 *关于计算机视觉*。第628–644页。蹦跳的人2016.
- [6], 杰里米·德·博内和保罗·维奥拉。像素: 概率体素化的体积重建。在 *国际计算机视觉会议的会议记录 (ICCV)*, 第418–425页, 1999年。
- [7]郝强风扇。郝苏和列奥尼达斯·J吉巴斯。一个三维对象的点集生成网络从单个图像中进行重建。在IEEE关于计算机视觉的会议的论文集中 *和模式识别*。第605–613页, 2017年。
- 古川安田[8]和潘塞。准确。密集, 稳健的多视图立体视觉。电器和电子工程师学会 *模式分析和机器intelligence*, 32(8): 1362–1376. 2009.
- [9]西尔瓦诺加利亚尼。卡特琳·拉辛格和康拉德·辛德勒。大规模并行的多种方式查看立体声重建。德国出版的电影。 *费和地理信息e. V. 25(361–369):2. 2016.*
- [10]阿莫斯Gropp。Lior亚里夫。尼夫海姆, 马坦阿兹蒙, 和雅隆利普曼。学习用于 shapes. arXiv预印本arXiv: 2002. 10099, 2020.
- [11]拉斯谟詹森。安德斯达尔, 乔治沃吉亚齐斯, 恩吉尔托拉, 和亨里克阿内斯。大尺度多视角立体视觉评价。2014年IEEE计算机视觉与模式会议 *识别能力*。页面406–413. 2014.
- [12]JueJiang。南通吉。韩志忠和兹威克。Sdfdifff: 可区分渲染对有符号距离字段的三维形状优化。在IEEE/CVF计算机会议上 *视觉和模式识别 (CVPR)*。2020年6月。
- [13]Abhishek Kar. 克里斯蒂安·汉内, 和吉丁德拉·马利克。学习多视角立体声机。 *arXiv预印本, arXiv: 1708. 05375, 2017年*。
- [14]广野加藤。上谷吉坂和原田达也。神经三维网格渲染器。在 *诉讼中 IEEE关于计算机视觉和模式识别会议*, 第3907–3916. 2018页。[15]Srinivas Kaza等人。使用有符号距离函数的可分割卷渲染。博士论文麻省理工学院, 2019年。
- [16], 迈克尔·卡兹丹和胡格斯·霍普。经过筛选的泊松表面重建。ACM跨。图.. 32(3), 2013年7月。
- [17]Petr Kellnhofer. 拉尔斯杰比, 安德鲁琼斯, 瑞恩斯派塞。卡里·普利, 和戈登·韦茨斯坦。神经发光图rendering. arXiv预印本arXiv: 2103. 11571. 2021.

- [18] 饮食协会, 金马和吉米·巴。一种制作随机optimization.arXiv预印本的方法  
*arXiv:1412.6980.2014.*

- [19]Chen-Hsuan林。陈孔和西蒙·露西。学习高效的点云生成，以进行密集的三维物体重建。在人工智能AAAI人工智能会议上。卷32. 2018。
- [20]Lingjie刘。杜伊古锡兰。程林。王文平，和米特拉。基于图像的电线艺术的重建。36(4):63:1-63:11. 2017。
- [21]Lingjie刘。陈能伦。杜伊古锡兰。西奥博，王文平和尼洛。米特拉。曲线融合：从rgbd序列中重建薄结构。37(6), 2018。
- [22]Lingjie刘。顾雅高。林考。蔡美儿和神督。神经稀疏体素字段。神经信息处理研究进展。
- [23]邵辉刘。张殷达。彭松友，石博欣。波尔菲斯和崔兆鹏。区域：渲染具有可微球面跟踪的深隐式有符号距离函数。在计算机视觉和模式识别会议论文集 2019-2028. 2020。
- [24]，斯蒂芬·伦巴迪，托马斯·西蒙，杰森·萨拉吉，加布里埃尔·施瓦茨。安德烈亚斯·莱尔曼和耶瑟尔·谢赫。神经体积：从图像中学习动态可渲染的体积。自动控制模块在图形上的交易 (TOG). 38(4): 65. 2019。
- [25]Priyanka Mandikal. K L Navaneet. 玛雅克·阿加瓦尔。和文卡特什先生。3D-LMNet：潜在嵌入匹配精确和多样的三维点云重建从单一图像在英国机器视觉会议会议记录中 (BMVC) . 2018。
- [26]保罗·梅雷尔。阿米尔阿克巴扎德。梁王。菲利普斯莫多海。迈克尔、杨瑞刚、尼斯特和波勒菲。基于实时可见性的深度地图融合。页面1-8. 012007。
- [27]Lars Mescheder. 迈克尔·奥克尔，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。在诉讼程序中 IEEE/CVF计算机视觉和模式识别会议。页面4460-4470. 2019。
- [28]Mateusz米哈尔凯维奇。庞特斯，多米尼克杰克，巴克塔什莫特拉格。安德斯和埃里克森。神经网络中隐含的表面表示。在IEEE国际计算机视觉会议 (ICCV) 。 2019年10月。
- [29]本·米尔登霍尔。普拉图尔P斯里尼瓦桑。马修Tancik。乔纳森T巴伦。拉维拉马莫奥蒂。和任宁集团。Nerf：将场景表示为视图合成的神经辐射场。在欧洲计算机视觉会议。第405-421页。施普林格， 2020年。
- [30]迈克尔·尼梅耶。拉尔斯·梅切德，迈克尔·奥克斯尔和安德烈亚斯·盖格。可差的体积metrie渲染：在没有三维监督的情况下学习隐式三维表示。在诉讼中 IEEE/CVF计算机视觉和模式识别会议。第3504-3515页。 2020。
- [31]迈克尔·奥克斯尔。彭鹏和盖革。统一的神经元：统一的神经隐式表面和辐射场的多视图reconstruction. arXiv预印本arXiv: 2104. 10078. 2021
- [32]Jeong Joon公园。彼得弗洛伦斯。朱利安·斯特劳布，理查德·纽科姆和史蒂文·洛夫格罗夫。DeepSDF：学习连续有符号的距离函数的形状表示。在继续 IEEE/CVF计算机视觉和模式识别会议的简介。第165-174页。 2019。
- [33]松友彭。迈克尔尼梅耶。梅斯切德，马克波勒菲斯和安德烈亚斯盖格。卷积占用网络。ArXiv, abs/2003. 04618. 2020。
- [34]SunsukeSaito。曾黄。Ryota自然。三岛志贺。金泽安珠和李浩。Pifu：像素对齐的隐式功能，用于高分辨率覆盖的人类数字化。ICCV, 2019。
- 齐藤顺助，托马斯西蒙，杰森萨拉吉。和俊。多级像素对齐隐式功能用于高分辨率三维人体数字化。在诉讼程序中 IEEE/CVF计算机视觉和模式识别会议。页面84-93. 2020。
- [36]，蒂姆·萨利曼斯和迪德里克·P·金马。权重归一化：一个简单的重新参数化，以加速训练的深度神经网络。arXiv预印本arXiv: 1602. 07868, 2016。



- [37], 约翰内斯·申伯格, 郑英良。迈克尔弗拉姆和马克波尔菲斯。非结构化多视图立体的像素视图选择。在欧洲计算机视觉会议上。第501-518页。蹦跳的人2016。
- [38] 史蒂文·塞茨和查尔斯·戴尔。通过体素着色重建逼真的场景。*《国际计算机视觉杂志》*。35(2):151-173. 1999。
- [39] 文森特西茨曼。Justus奖杯。费利克斯海德, 马提亚斯Nießner。戈登·韦茨斯坦和迈克尔Zollhofer。深度体素: 学习持久的3d特征嵌入。在诉讼程序中*IEEE/CVF计算机视觉和模式识别会议*。页面2437-2446. 2019。
- [40] 文森特西茨曼。迈克尔·佐尔霍弗和戈登·韦茨斯坦。场景表示网络: 连续的三维结构感知的神经场景表示。在神经信息处理系统的研究进展中。页1119-1130. 2019。
- [41] A. Tabb。来自轮廓概率图的形状: 在存在轮廓提取和校准误差的情况下, 薄物体的重建。第161-168页。2013年6月。
- [42] Matthew Tancik。普拉图尔P斯里尼瓦桑。本米尔登霍尔。莎拉·弗里多维奇-凯尔。Nithin拉哈万。乌特卡什辛格尔。拉维拉马莫奥蒂。巴伦和任吴。傅里叶特征使网络在低维domains。arXiv预印本arXiv: 2006.10739. 2020中学习高频函数。
- [43] 亚历克斯·特雷维蒂克和博杨。Grf: 学习3d场景表示和rendering。arXiv预印本arXiv: 2010.04595. 2020的一般辐射场。
- [44] 南阳王。张殷达。李朱文, 傅彦伟。刘伟二、姜玉刚。像素2网格: 从单个rgb图像生成3d网格模型。发表在欧洲计算机视觉会议(ECCV)的会议记录中, 2018年, 第52-67页。
- [45] Peng王。刘玲杰, 陈能伦。朱红国。西奥博和王文平。Vid2曲线: 同时估计摄像机运动和从rgb视频重建。ACM跨。图.. 39(4)。2020年7月。
- [46] Chao文。张殷达。李朱文和傅彦伟。像素2网格++: 通过变形生成多视图三维网格。在IEEE/CVF国际会议会议记录中*计算机视觉*。页面1042-1051. 2019。
- [47] 谢浩哲。姚宏勋, 孙小帅。周尚辰, 张升平。Pix2vox: 从单个和多视图图像进行上下文感知的三维重建。在诉讼程序中*IEEE/CVF计算机视觉国际会议*。页2690-2698. 2019。
- [48] 姚姚, 罗子欣, 李世伟, 张景阳。任宇凡。周磊、田芳、龙全。一个用于广义多视图立体声网络的大规模数据集。在*IEEE/CVF计算机视觉和模式识别会议论文集*1790-1799. 2020。
- [49] Lior亚里夫。Yoni卡斯汀。恐怖莫兰。梅拉夫加伦。Matan阿兹蒙。巴斯里·罗南和亚伦·利普曼。多视图神经表面重建, 通过解开几何形状和外观。*神经信息处理系统的研究进展*。33. 2020。
- [50] 王一凡。吴世浩。奥兹蒂雷利和奥尔加霍金。等点: 使用混合representations。arXiv预印本arXiv: 2012.06434. 2020优化神经隐式曲面。
- [51] 克里斯托弗扎克。托马斯·波克和霍斯特·比肖夫。一种鲁棒电视范围图像集成的全局最优算法。2007年IEEE第一届计算机视觉国际会议。2007年第1-8页
- [52] Kai张。卢军。王倩倩。卡维塔巴拉。和诺亚蛇毒的。Physg: 逆渲染与球面高斯数为基于物理的材料编辑和relighting。arXiv预印arXiv: 2104.00674. 2021。
- [53] Kai张。Gernot Riegler。和科尔顿。Nerf++: 分析和改善神经辐射, fields。arXiv预印本arXiv: 2010.07492. 2020。

## 补充-

一个计算不透明度 $\alpha$ 的推导：

在本节中，我们将推导出方程式中的公式。13篇计算离散不透明度 $\alpha$ ；回想一下，不透明的密度函数 $\rho(t)$ 被定义为

$$\begin{aligned}\rho(t) &= \max \left( \frac{-\frac{d\Phi_s(f(\mathbf{p}(t)))}{dt}}{\Phi_s(f(\mathbf{p}(t)))}, 0 \right) \\ &= \max \left( \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))}, 0 \right)\end{aligned}\quad (18)$$

其中， $\phi_s(x)$  和  $\Phi_s(x)$  分别为逻辑分布的概率密度函数（PDF）和累积分布函数（CDF）。首先考虑样本点间隔 $[t_i, t_{i+1}]$ 位于一个范围 $[t_e, t]$ ，在这个范围内，相机光线从外部进入表面到inside，i.e. the符号距离函数在相机光线 $\mathbf{p}(t)$ 上减少。那么很容易看到 $-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) > 0$ 。它来自于Eqn. 论文中的12篇。

$$\begin{aligned}\alpha_i &= 1 - \exp \left( - \int_{t_i}^{t_{i+1}} \rho(t) dt \right) \\ &= 1 - \exp \left( - \int_{t_i}^{t_{i+1}} \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))} dt \right)\end{aligned}\quad (19)$$

请注意，积分项是由

$$\int \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))} dt = -\ln(\Phi_s(f(\mathbf{p}(t)))) + C, \quad (20)$$

其中 $C$ 是一个常数。因此，离散的不透明度可以通过计算

$$\begin{aligned}\alpha_i &= 1 - \exp[-(-\ln(\Phi_s(f(\mathbf{p}(t_{i+1})))) + \ln(\Phi_s(f(\mathbf{p}(t_i))))] \\ &= 1 - \frac{\Phi_s(f(\mathbf{p}(t_{i+1})))}{\Phi_s(f(\mathbf{p}(t_i)))} \\ &= \frac{\Phi_s(f(\mathbf{p}(t_i))) - \Phi_s(f(\mathbf{p}(t_{i+1})))}{\Phi_s(f(\mathbf{p}(t_i)))}.\end{aligned}\quad (21)$$

接下来考虑的情况， $[t_i, t_{i+1}]$ 在一个范围内 $[t_r, t]$ 。在 $[t_i, t]$ 上，相机光线退出surface，i.e. the符号距离函数的增加 $[t_i, t_{i+1}]$ 。然后我们有 $-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) < 0$ 在 $[t_i, t_{i+1}]$ 。然后，根据Eqn. 18。我们有 $\rho(t) = 0$ 。因此，由Eqn. 在12篇论文中，我们有

$$\alpha_i = 1 - \exp \left( - \int_{t_i}^{t_{i+1}} \rho(t) dt \right) = 1 - \exp \left( - \int_{t_i}^{t_{i+1}} 0 dt \right) = 0$$

因此， $\alpha$ 值在这种情况下，是由

$$\alpha_i = \max \left( \frac{\Phi_s(f(\mathbf{p}(t_i))) - \Phi_s(f(\mathbf{p}(t_{i+1})))}{\Phi_s(f(\mathbf{p}(t_i)))}, 0 \right) \quad (22)$$

这就完成了方程式的推导。论文中的13篇。

### b一阶偏差分析

#### B. 1对我们的解的无偏性质的证明

定理1的证明：假设射线是从表面的外部移动到内部的。因此，我们有 $-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) > 0$ . because 约定符号距离函数 $f(x)$ 在表面外部是正的，内部是负的。

回想一下，我们的s密度场 $\phi_s(f(x))$ 是用逻辑密度函数 $\phi_s(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ 定义的，它是s型函数 $\Phi$ 的导数。 $\phi_s(x) = (1+e^{-x})^{-1}$ 。i. e.  $\phi_s(x) = \Phi'(x)$ 。

根据Eqn. 本文的5中，权值函数 $u(t)$ 由

$$u(t) = T(t)p(t),$$

在哪里

$$\rho(t) = \max \left( \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))}, 0 \right)$$

根据假设， $-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) > 0$  for  $t \in [t_i, t_f]$ 。因为 $\phi_s$ 是 $\phi_s(f(\mathbf{p}(t))) > 0$ 。Clearly,  $\phi_s$ 那个密度函数，我们那个 $(f(\mathbf{p}(t))) > 0$ 。但它是这样的

$$\rho(t) = \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))}.$$

这是积极的。因此，

$$\begin{aligned} u(t) &= T(t)p(t) \\ &= \exp \left( - \int_0^t \rho(t') dt' \right) \rho(t) \\ &= \exp \left( - \int_0^{t_i} \rho(t') dt' \right) \exp \left( - \int_{t_i}^t \rho(t') dt' \right) \rho(t) \\ &= T(t_i) \exp \left( - \int_{t_i}^t \rho(t') dt' \right) \rho(t) \\ &= T(t_i) \exp [ - (-\ln(\Phi_s(f(\mathbf{p}(t_i)))) + \ln(\Phi_s(f(\mathbf{p}(t)))) ] p(t) \\ &= T(t_i) \frac{\Phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t_i)))} \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi_s(f(\mathbf{p}(t)))}{\Phi_s(f(\mathbf{p}(t)))} \\ &= \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) T(t_i)}{\Phi_s(f(\mathbf{p}(t_i)))} \phi_s(f(\mathbf{p}(t))). \end{aligned} \tag{23}$$

作为有符号距离函数 $f$ 的一阶近似，假设曲面局部近似一个足够小的平面块，其向外单位法向量表示为 $\mathbf{n}$ 。因为 $f(x)$ 是组合距离函数。局部它有一个单位梯度向量 $\nabla f = \mathbf{n}$ 。

然后我们有

$$\begin{aligned} u(t) &= \frac{-(\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) T(t_i)}{\Phi_s(f(\mathbf{p}(t_i)))} \phi_s(f(\mathbf{p}(t))) \\ &= \frac{|\cos(\theta)| T(t_i)}{\Phi_s(f(\mathbf{p}(t_i)))} \phi_s(f(\mathbf{p}(t))). \end{aligned} \tag{24}$$

式中， $\theta$ 为视图方向 $\mathbf{v}$ 与单位法向量 $\mathbf{n}$ 之间的夹角，即 $\cos(\theta) = \mathbf{v} \cdot \mathbf{n}$ 。这里是 $|\cos(\theta)| T(t_i) \cdot \Phi_s(f(\mathbf{p}(t_i))) > 1$ 可以看作是一个常数。因此，当 $f(\mathbf{p}(t)) = 0$ 时， $w(t)$ 达到一个局部最大值，因为 $\Phi_s(x)$ 是一个单峰密度函数，在 $x = 0$ 处达到最大值。

我们注意到，在这个证明中，我们没有对相机和样本点 $\mathbf{p}(t_r)$ 之间的表面的存在作出任何假设。因此，该结论适用于在相机射线上有多个表面交点的情况。这就完成了证明。□

## B. 2幼稚解中的偏差

在本节中，我们证明了在朴素解中导出的权值函数是有偏的。根据Eqn. 3,  $e(t) = \langle \sigma \rangle - \sigma(t)$ ，不透明度  $\sigma(t)$ 。然后我们就有了

$$\begin{aligned}\frac{dw}{dt} &= \frac{d(T(t)\sigma(t))}{dt} \\ &= \frac{dT(t)}{dt}\sigma(t) + T(t)\frac{d\sigma(t)}{dt} \\ &= \left[ \exp\left(-\int_0^t \sigma(t)dt\right)(-\sigma(t)) \right] \sigma(t) + T(t)\frac{d\sigma(t)}{dt} \\ &= T(t)(-\sigma(t))\sigma(t) + T(t)\frac{d\sigma(t)}{dt} \\ &= T(t)\left(\frac{d\sigma(t)}{dt} - \sigma(t)^2\right)\end{aligned}\quad (25)$$

现在，我们对曲面交点附近的有符号距离函数  $f$  进行相同的一阶近似。在此条件下，上述方程可以改写为

$$\begin{aligned}\frac{dw}{dt} &= T(t) \left( (\nabla f(\mathbf{p}(t)) \cdot \mathbf{v}) \phi'_s(f(\mathbf{p}(t))) - \phi_s(f(\mathbf{p}(t)))^2 \right) \\ &= T(t) (\cos(\theta) \sigma - \sigma^2) \phi'_s(f(\mathbf{p}(t))) - \phi_s(f(\mathbf{p}(t)))^2.\end{aligned}\quad (26)$$

这里的  $\cos(\theta)$  可以看作是一个常数。现在假设  $\mathbf{p}(t^*)$  是表面  $S$  上的一个点，即  $f(\mathbf{p}(t^*)) = 0$ 。接下来，我们将检查的值。首先，显然， $T(t^*) > 0$  和  $\frac{dw}{dt}(t)$  at  $t = t^*$   $\phi_s(f(\mathbf{p}(t^*)))^2 > 0$ 。那么，既然我们有  $\phi'_s(0) = 0$ ，

$$\frac{dw}{dt}(t^*) = T(t^*) (\cos(\theta) \phi'_s(0) - \sigma(t^*)^2) = -T(t^*) \phi_s(0)^2 < 0.$$

因此，在朴素解中， $u(t)$  没有达到局部最大值，它对应于曲面  $S$  上的一个点。这就完成了证明。□

## C 二阶偏差分析

在本节中，我们简要介绍在表面交点附近的区间的局部分析。在二阶近似中。在这种情况下，我们遵循与B节类似的假设，即有符号的距离函数  $f(\mathbf{p}(t))$  在区间  $[t_r, t_r]$  内沿射线单调地减小。

根据Eqn. 25,  $u(t)$  的导数为：

$$\frac{dw}{dt} = T(t) \left( \frac{d\sigma(t)}{dt} - \sigma(t)^2 \right)$$

显然，我们有  $T(t) > 0$ 。因此，当  $w(f)$  达到局部最大值 at  $t$  时，就存在

0.

这就是我们的解决方案的情况。在我们的解中，体积密度由Eqn之后的  $\sigma(t) = p(t)$  给出。18. 在组织之后，我们有

$$\frac{d^2 f}{dt^2}(\mathbf{p}(\tilde{t})) \cdot \phi_s(f(\mathbf{p}(\tilde{t}))) + \left( \frac{df}{dt}(\mathbf{p}(\tilde{t})) \right)^2 \phi'_s(f(\mathbf{p}(\tilde{t}))) = 0.$$

这里我们在表面交点  $t$  的  $f$  附近进行局部分析，其中  $f(\mathbf{p}(t^*)) = 0$ ,  $F = t^* + \Delta t$ 。

和我们让。作为一个三阶分析，我们假设在  $\tau$

这个本地  $inc$  经过替换和组织后， $p$  就会被诱导。

局部最大点  $F$  的方程  $-\left(s(\mu\Delta_t + \frac{1}{2}\tau\Delta_t^2)\right) = (\mu + \tau\Delta_t)^2 \cdot \left(s\left(1 - e^{-s(\mu\Delta_t + \frac{1}{2}\tau\Delta_t^2)}\right)\right)$ 。

(27)

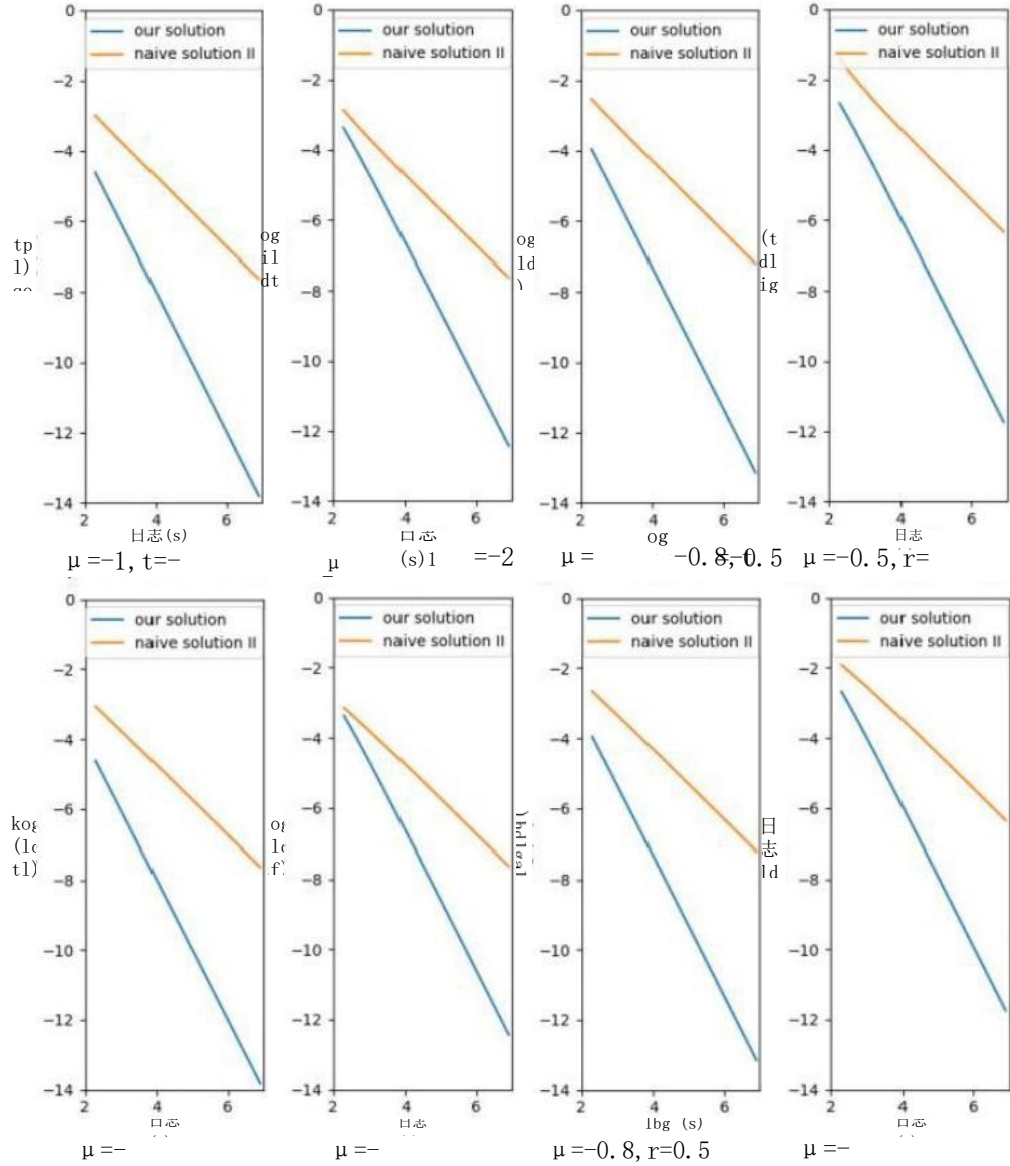


图9:  $\Delta$ 与 $s$ 的曲线, 给定固定的 $\mu$ ,  $T$ 。请注意, 这些轴在 $\ln(|\Delta|)$ 和 $\ln(s)$ 中所示。

我们稍后会分析。

这个天真的解的例子。在这里, 我们进行了与我们的解决方案的情况下类似的局部分析。对于朴素解, 当 $u(t)$ 在 $f$ 处达到局部最大值时, 有:

$$(\mu + \tau \Delta_t) \cdot \left( - \left( 1 - e^{-2s(\mu \Delta_t + \frac{1}{2} \tau \Delta_t^2)} \right) \right) = e^{-s(\mu \Delta_t + \frac{1}{2} \tau \Delta_t^2)} \quad (28)$$

比较。基于Eqn. 27和方程式. 28。我们可以数值求解任何给定的 $\mu$ ,  $r$ 和 $s$ 值的 $\Delta t$ 上的方程。下面我们绘制了 $\mu$ ,  $r$ 中不同(固定)值的 $\Delta$ 相对于增加 $s$ 的曲线. 9。

如图9所示。我们的解的局部最大位置的误差 $\Delta_t = 0(s-2)$ 和朴素解的误差 $\Delta_t = 0(s-1)$ 。也就是说, 当 $s$ 密度的标准差 $1/s$ 接近于0时, 我们的误差比朴素解的误差更快于零。这是二次收敛和线性收敛。



## d附加的实验细节

### D. 1附加的实施细节

**网络架构。**我们使用与IDR[49]类似的网络架构，它由两个mlp组成，分别编码SDF和颜色。有符号距离函数f由一个MLP建模，该MLP由8个隐藏层组成，隐藏大小为256。我们用3=100的Softplus替换了原来的ReLU，作为所有隐藏层的激活功能。跳过连接[32]用于连接输入和第四层的输出。颜色预测的函数c由MLP建模4隐藏层大小256，不仅需要空间位置p作为输入也视图方向v法向量 $SDF_n = \nabla f(p)$ 和SDF MLP 256维特征向量。位置编码应用于6个频率的空间位置p和4个频率的方向v。与IDR一样，我们使用权重归一化[36]来稳定训练过程。

**培训细节。**我们使用ADAM优化器[18]来训练我们的神经网络。在前5k次迭代中，学习速率首先从0到 $5 \times 10^{-4}$ 线性升温，然后由余弦衰减计划控制到最小学习速率为 $2.5 \times 10^{-5}$ 。我们在一个Nvidia 2080Ti GPU上对每个模型进行300k次迭代训练14小时（“w/掩码”设置）和16小时（“w/o掩码”设置）。

**阿尔法和颜色计算。**在实现中，我们实际上有两个

点-采样剖面点=、+和采样中点

切片长度为0；=ti+1-ti。10. 为了计算alpha值aj，我们

使用最高的截面点  
甲点p的颜色；。

0). 为了计算颜色cyp我们使用  
$$p_i = o + \frac{t_i + t_{i-1}}{2} v$$

分层抽样。具体地说，我们首先沿射线均匀采样64个点，然后对k=进行4次重要采样。第i次迭代中的粗概率估计由一个固定的s值计算，该值设为 $32 \times 2^i$ 。在每次迭代中，我们额外抽取16个点。因此，NeuS的采样点总数为128个。对于“w/o掩模”设置，我们在球体外采样额外的32个点。外部场景用NeRF++[53]表示。

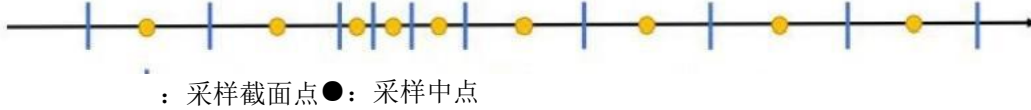


图10：在射线上定义的截面点和中点。

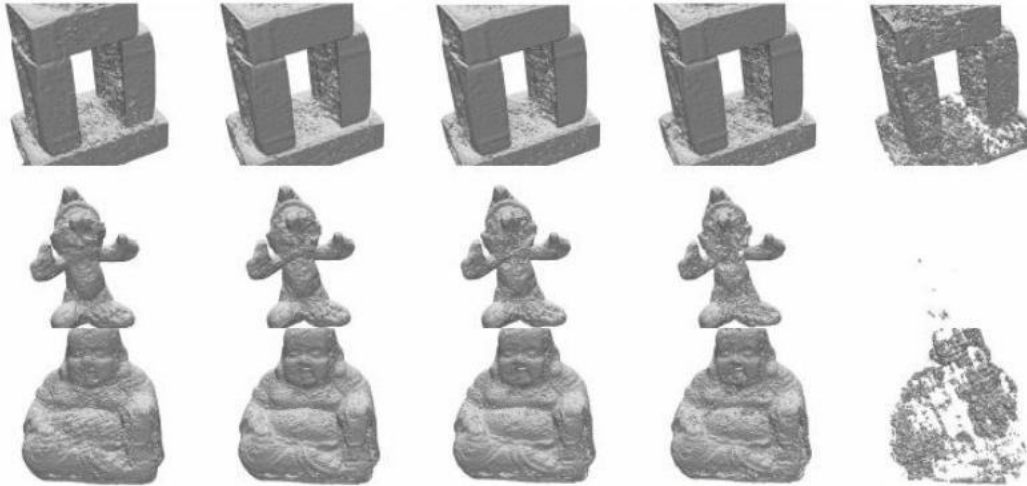
### D. 2基线

我或[49]。为了实现IDR，我们在DTU数据集上使用了他们正式发布的codes<sup>2</sup> and 预训练模型。

<sup>2</sup><https://github.com/lioryariv/idr>

扫描 ID	阈值0	阈值25	阈值50	阈值100	阈值500
扫描 40	2.36	.79	1.86	2.07	4.26
扫描 83	1.65	1.20	1.37	2.24	29.10
扫描114	1.62	1.04	1.10	1.4	8.66

表2：在从DTU数据集的三个场景上使用不同的阈值从NeRF结果中提取的地面真实值和水平集表面之间的倒角距离。



阈值0 阈值25 阈值50 阈值100 阈值500

图11: 使用不同的阈值从NeRF结果中提取的水平集表面的可视化。

NeRF[29] . 为了实现NeRF, 我们使用了来自nerf-pytorch<sup>2</sup> 中的代码。为了从NeRF中提取曲面, 我们使用了密度水平集25, 通过实验验证, 这是重建误差最小的最佳水平集, 如表2和图11所示。

**COLMAP[37]**. 我们使用经常提供的CLI (命令行界面) 版本的COLMAP。密集的点云是通过按顺序运行以下命令而产生的: (1) `feature_extractor`, (2) `dense_matching`, (3) `patch_match_stereo`, 和 (4) `stereo_fusion`。给定致密点云。网格由 (5) 泊松公司生产。

统一用户[31]。本文中的定量和定性的结果是由统一神经元的作者提供的。

## E附加的实验结果

### E. 1渲染的质量和速度

除了重建的表面外, 我们的方法还呈现了高质量的图像, 如图12所示。在单个Nvidia 2080Ti GPU上的默认体积渲染设置中, 以1600x1200的分辨率渲染图像大约需要320秒。此外, 我们还测试了另一种采样策略, 首先应用球体追踪来寻找表面附近的区域, 并且只寻找这些区域内的采样点。使用此策略, 渲染相同分辨率的图像只需要大约60秒。表3报告了默认体渲染设置和球体跟踪设置中的PSNR和SSIM的定量结果。

S “<https://github.com/yenchenlin/nerf-py>” 火炬

扫描ID	2437	40	55	6365	69	8397	105	106	110	114	118	122	麦肯
PSNR (我们的)	28.2027	1028.1328	8032.0533	7530.9634	4729.5732	9835.0732	7431.6936	9737.07					31.97
PSNR (Oursr)	27.0726	5827.7028	3731.3231	3930.2031	7928.5830	8733.6132	4031.3335	5535.96					30.85
SSIMiOurs	0.7640	0.8130	0.7370	0.7680	0.9170	0.8350	0.8450	0.8500	0.8370	0.8370	0.8750	0.8760	0.840
SSIMiOursr)	0.7570	0.8110	0.7360	0.7590	0.9150	0.7880	0.8130	0.8120	0.7940	0.8110	0.8520	0.8620	0.820

表3: 不同抽样strategies. -ST的定量比较表明了采用球面追踪的抽样策略。



图12：使用我们的方法在DTU数据集上使用不同的采样策略渲染图像。-ST表示使用球体跟踪的采样策略。

扫描ID	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	平均
PSNR (NeRF)	24.83	25.35	26.87	27.27	28.64	29.30	29.24	29.65	28.28	29.94	26.76	29.26	28.61	29.32	28.85	29.00
PSNR (我们的)	23.98	22.79	25.21	26.03	28.32	29.32	29.80	27.45	28.89	26.03	28.93	27.47	29.78	29.37	23.95	28.55
SSIM (NeRF)	0.75	0.73	0.79	0.40	0.78	0.00	0.50	0.80	0.30	0.81	0.50	0.87	0.00	0.85	0.78	0.82
SSIM (我们的)	0.73	0.20	0.77	0.80	0.22	0.73	0.90	0.81	0.10	0.81	0.15	0.86	0.60	0.86	0.47	0.82

表4：与NeRF在没有掩模监督的新视图合成任务上的定量比较

## E. 2新视图合成

在这个实验中，我们将DTU数据集中10%的图像作为测试集，其他的图像作为训练集。我们比较了PSNR和SSIM与NeRF的测试集的定量结果。如表4所示。我们的方法达到了与NeRF相当的性能。

## E. 3 SDF定性评价

虽然我们的方法没有元素正则化[10]或几何初始化[1]，产生可信的表面重建结果，我们的完整模型可以预测更准确的符号距离函数，如图13所示。此外，使用随机初始化由于位置编码[42]的光谱偏差而产生轴对齐的伪影，而几何初始化[1]没有这种伪影。

#### E. 4 培训进展

我们在BlendedMVS数据集中展示了榴莲在不同训练阶段的重建表面。如图14所示。曲面在训练过程中变得更加清晰。同时，我们还在图中提供了一条曲线来显示o的可训练标准差在训练过程中的变化。正如我们所看到的，优化过程将自动减少标准偏差，从而使表面随着更多的训练步骤而变得更清晰、更清晰。

#### E. 5 限制

图16显示了一个失败的情况，我们的方法不能正确地重建金属兔模型表面的无纹理区域。原因是这种无纹理区域在神经渲染中的重建是模糊的。



#### E. 6 其他结果

在这节中，我们展示广告 图16：一个包含无文本区域的故障重建案例。

不同的定性结果

在DTU数据集和BlendedMVS数据集上。图17显示了在w/掩模设置和w/o掩模设置中与基线方法的比较。图15显示了无掩码设置中的附加结果。

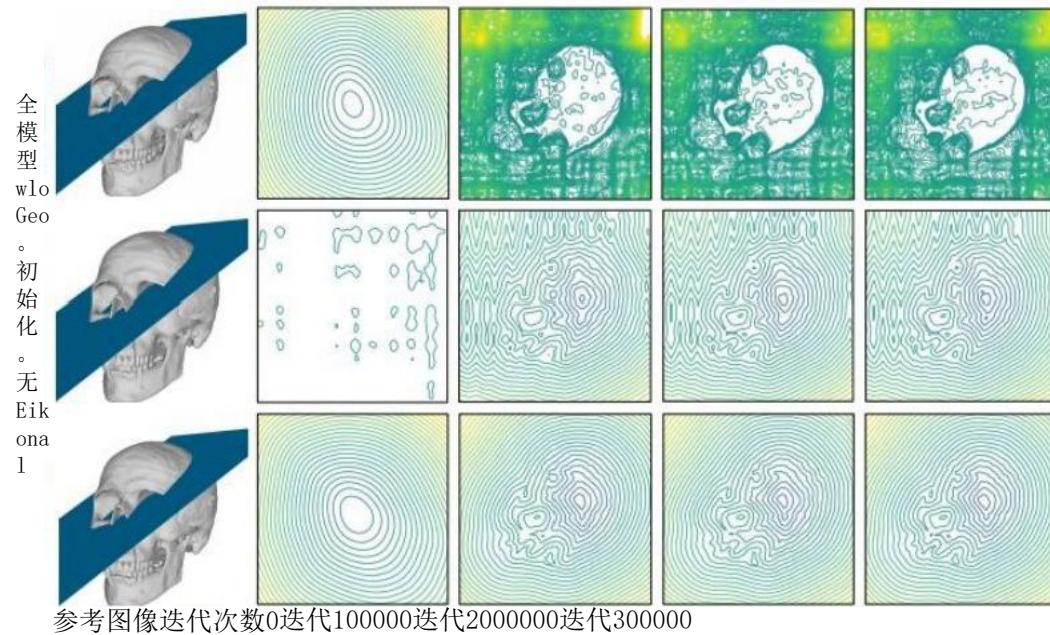


图13：不同训练迭代中切割平面（左图蓝色平面）上有符号距离场的可视化。



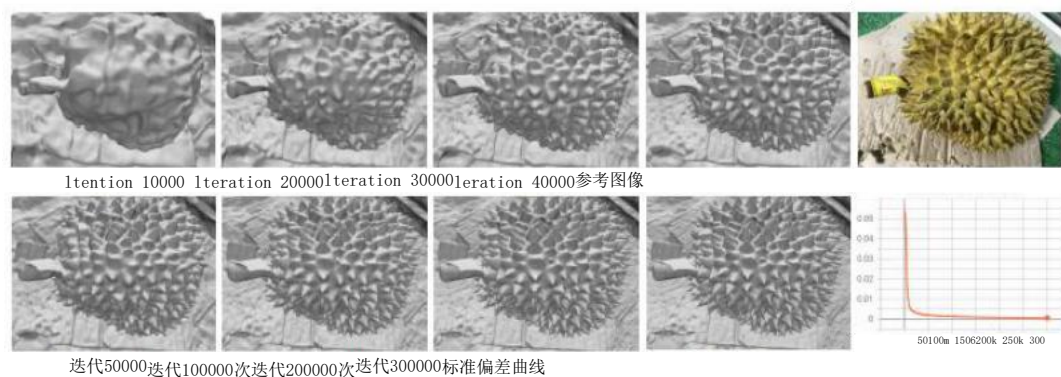


图14：在BlendedMVS数据集中，榴莲的训练进展。右下角显示了训练过程中可训练标准差的曲线。



图15：在没有掩码监督的BlendedMVS数据集上的附加重建结果。



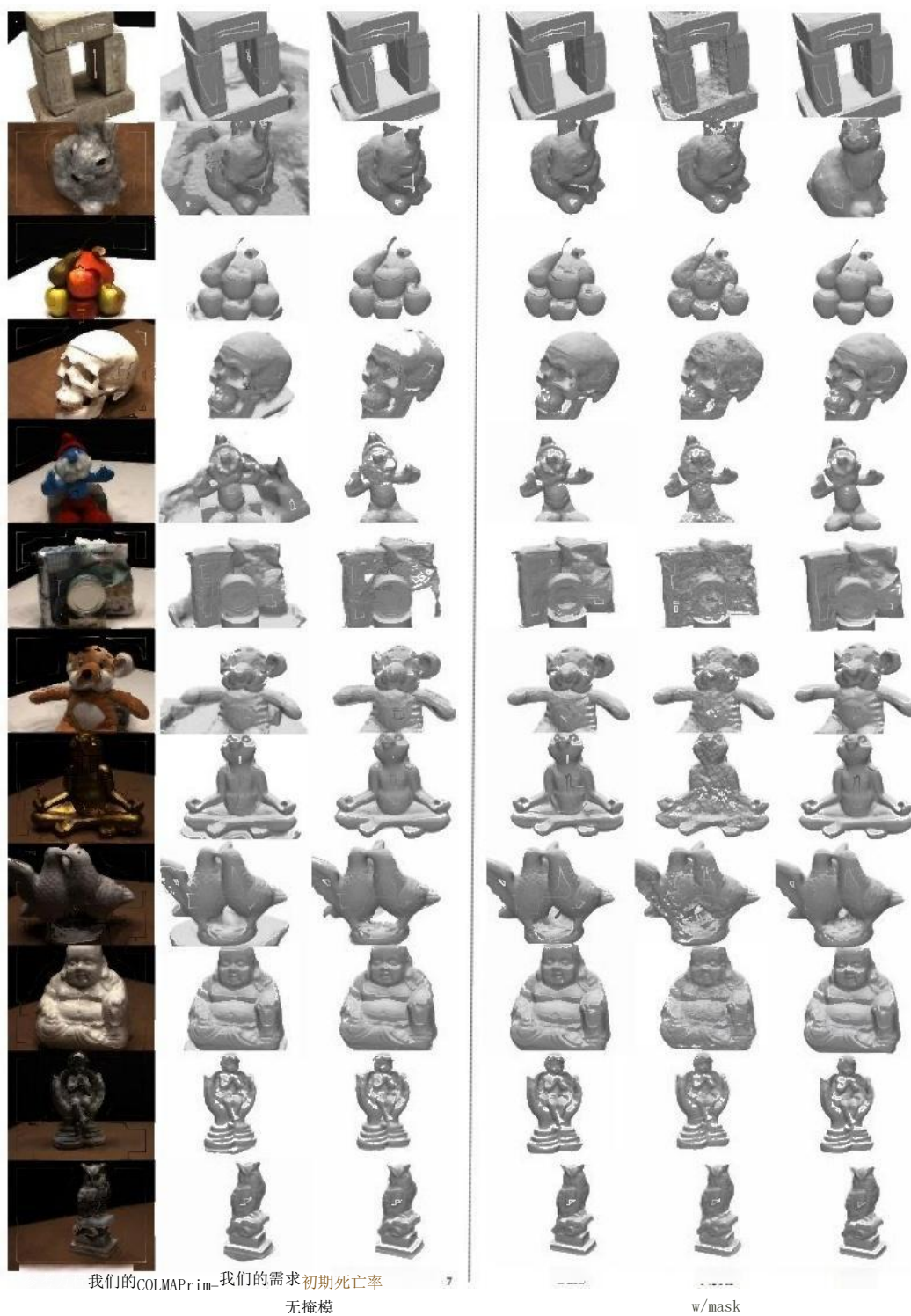


图17：在DTU数据集上的其他重建结果。