

NERF++: 分析和改进 神经辐射场

张凯
康奈尔理工学院

Gernot Riegler
英特尔实验室

诺亚斯奈弗利
康奈尔理工学院

弗拉德伦科尔顿
英特尔实验室

摘要

神经辐射场 (NeRF) 实现了令人印象深刻的视图合成结果的各种捕获设置, 包括 360^0 有界场景的捕获和有界和无界的场景的向前捕获。NeRF将表示视图不变的不透明度和与视图相关的颜色体积的多层感知器 (MLPs) 匹配到一组训练图像中, 并基于体渲染技术对新的视图进行采样。在本技术报告中, 我们首先评论了辐射场及其潜在的歧义, 即形状-辐射歧义, 并分析了NeRF在避免这种歧义方面的成功程度。其次, 我们解决了在将NeRF应用于 360^0 时所涉及的一个参数化问题⁰在大规模的、无限的3D场景中捕捉物体。我们的方法在这个具有挑战性的场景中提高了视图合成的保真度。代码可以在<https://github.com/Kai-46/nerfplusplus>上获得。

1 介绍

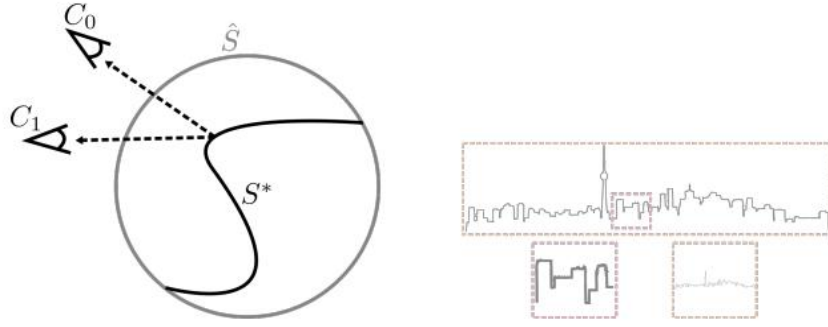
回想一下你上次的假期, 你拍下了几张你最喜欢的地方的照片。现在在家里, 你希望再次在这个特殊的地方四处走动, 如果只是虚拟的。这要求您在可能无限的场景中从不同的、自由放置的视点渲染相同的场景。这种新颖的观点合成任务是计算机视觉和图形学中长期存在的问题 (陈和威廉姆斯, 1993; Debevec等人, 1996; 莱沃伊和汉拉汉, 1996年; 戈特勒等人, 1996年; 舒姆和康, 2000年)。

近年来, 基于学习的方法在逼真的、新颖的观点合成方面取得了重大进展。特别是神经辐射场 (NeRF) 的方法, 已经引起了广泛的关注 (Mildenhall等., 2020)。NeRF是一个隐式的基于MLP的模型, 它将5D向量——3D坐标加上2D查看方向——映射到不透明度和颜色值, 通过将模型拟合到一组训练视图来计算。所得到的5D函数可以用传统的体渲染技术生成新的视图。

在本技术报告中, 我们首先分析了NeRF中存在潜在的失效模式, 并分析了为什么NeRF在实践中避免了这些失效模式。其次, 我们提出了一种新的空间参数化方案, 我们称之为倒球参数化, 它允许NeRF处理一类新的无界场景捕获。

特别地, 我们发现在理论上, 在一组训练图像没有任何正则化的情况下, 从一组训练图像优化5D函数可能会遇到关键的退化解, 而不能推广到新的测试视图。这种现象被封装在形状-辐射模糊中 (图1, 左), 其中通过在每个表面点上适当选择输出的二维辐射, 可以完美地为任意图像匹配一组错误的几何图像。我们的经验表明, NeRF中使用的特定MLP结构在避免这种歧义方面发挥了重要作用, 产生了一种令人印象深刻的合成新观点的能力。我们的分析为NeRF令人印象深刻的成功提供了一个新的视角。

我们还解决了一个在涉及 360^0 的挑战性的场景中出现的空间参数化问题⁰在无限的环境中捕获周围的对象 (图1, 右)。对于 360^0 捕获, NeRF假设整个场景可以打包成一个有限的体积, 这对于大尺度场景来说是有问题的: 要么我们将一小部分场景匹配到体积中并进行详细采样, 要么完全不能捕获背景元素; 或者, 我们将整个场景匹配到体积中, 由于采样分辨率有限, 到处都缺乏细节。我们提出了一个简单而有效的解决方案, 分别建模前景和背景, 解决建模无界三维的挑战



无界场景的形状-辐射模糊性参数化

图1: 形状辐射模糊度 (左) 和无边界场景的参数化 (右)。形状偏差模糊性: 我们的理论分析表明, 在没有显式或隐式正则化的情况下, 一组训练图像可以独立于恢复的几何图形进行拟合 (e.g., 通过利用与视图相关的亮度来模拟正确的几何图形的效果)。无界场景的参数化: 使用标准的参数化方案, (1) 只对部分场景进行了建模 (红色轮廓), 导致背景元素中的重要伪影, 或者 (2) 对整个场景进行了建模 (橙色轮廓), 这导致由于有限的采样分辨率而导致细节的整体损失。

具有倒置球体场景参数化的背景内容。我们从坦克和模板数据集 (Knapitsch 等人, 2017) 和 Ycer 等人的光场数据集中获得了改进的定量和定性结果。ü (2016)。

总之, 我们分析了 NeRF 如何解决形状-辐射模糊, 以及 360 情况下无界场景参数化的补救方法⁰ 俘获

2 项准备工作

给定一个静态场景的姿态多视图图像, NeRF 重建了一个表示软形状的不透明度场 a , 以及一个表示与视图相关的表面纹理的辐射场 c 。 a 和 c 都被隐式地表示为多层感知器 (MLPs); 不透明度场被计算为一个三维位置, $x \in \mathbb{R}^3$ 的函数³, 辐射场由三维位置和观察方向进行参数化² (即, 单元 3 个向量的集合)。因此, 我们使用 $a(x)$ 将不透明度作为位置的函数, 而使用 $c(x, d)$ 将亮度作为位置和观察方向的函数。

理想情况下, 对于不透明材料, a 应该在地面真实表面位置达到峰值, 在这种情况下, c 减少为表面光场 (Wood et al., 2000)。给定 n 个训练图像, NeRF 使用随机梯度下降, 通过最小化地面真实观测图像 I 之间的差异来优化 a 和 c , 以及预测的图像 $\hat{I}_i(a, c)$ a 和 c 的观点:

$$\min_{\sigma, c} \frac{1}{n} \sum_{i=1}^n \|I_i - \hat{I}_i(\sigma, c)\|_2^2. \quad (1)$$

隐式体积 a 和 c 被光线跟踪来渲染每个像素 $\hat{I}(a, c)$ (Kajiya & Von Herzen, 1984)。对于给定的射线, $x \in \mathbb{R}^3$, $d \in \mathbb{S}^2$, $t \in \mathbb{R}^+$, 它的颜色是由积分决定的

$$C(r) = \int_{t=0}^{\infty} a(o+t) \cdot c(o+t, d) \cdot e^{-\int_{s=0}^t a(o+sd) ds} dt. \quad (2)$$

为了补偿网络的光谱偏差和合成更清晰的图像, NeRF 使用了一个位置编码 V , 将 x 和 d 映射到它们的傅里叶特征 (Tancik et al., 2020):

$$V^k: p \mapsto (\sin(2^0 p), \cos(2^0 p), \sin(2^1 p), \cos(2^1 p), \dots, \sin(2^k p), \cos(2^k p)), \quad (3)$$

其中 k 是一个指定傅里叶特征向量维数的超参数。



GT训练视图预测 (PSNR: 25.4) GT测试视图预测 (PSNR: 10.7)

图2: 演示形状辐射模糊, 我们预训练网络合成数据集的不透明度字段优化模型一个不正确的三维形状 (一个单位球, 而不是一个推土机形状), 而辐射场 c 优化映射训练射线的交集与球体和视图方向他们的像素颜色。在本例中, 我们使用3个MLP层来建模视图依赖的影响 (见图3中的MLP结构), 并拟合到50个视点随机分布在一个半球上的合成训练图像。由此产生的不正确的解决方案很好地解释了训练图像 (左两张图像), 但不能推广到新的测试视图 (右两张图像)。

3形状-辐射模糊

NeRF建模视图相关的外观的能力导致了三维形状和一个固有的亮度之间的模糊性, 可以允许退化解, 在没有正则化。对于一个任意的、不正确的形状, 我们可以证明存在一系列的辐射场, 它可以完美地解释训练图像, 但却不能很好地推广到新的测试视图。

为了说明这种模糊性, 假设对于一个给定的场景, 我们将几何图形表示为一个单位球体。换句话说, 让我们将NeRF的不透明度场固定在单位球面的表面为1, 在其他地方为0。然后, 对于每个训练图像中的每个像素, 我们将穿过该像素的光线与球体相交, 并将交点 (以及沿着光线方向) 处的辐射值定义为该像素的颜色。这个人工构造的解决方案是一个有效的NeRF重建, 完美地适合输入的图像。然而, 这种解决方案合成新视图的能力非常有限: 准确地生成这样的视图将涉及到在每个表面点上重建一个任意复杂的与视图相关的函数。该模型不太可能准确地插值如此复杂的函数, 除非训练视图非常密集, 如传统的光场渲染工作 (Buehler等人, 2001; 列沃&汉拉汉, 1996; Gortler等人, 1996年)。这种形状-辐射的模糊性如图2所示。

为什么NeRF要避免这种退化解? 我们假设有两个相关的因素影响了NeRF: 1) 不正确的几何形状迫使辐射场具有更高的内在复杂性 (i. e., 相比之下, NeRF的特定的MLP结构隐含地编码了表面反射率的平滑BRDF先验。

因子1: 当 a 偏离正确的形状时, c 一般必须相对于 d 成为一个高频函数, 才能重建输入的图像。对于正确的形状, 表面光场通常会更平滑 (实际上, 对于兰伯特材料是恒定的)。在有限的容量MLP下, 不正确形状所需的更高复杂性更难表示。

因子2: 特别是, NeRF的特定MLP结构编码了一个隐式的先验, 有利于光滑的表面反射函数, 其中在任何给定的表面点 x 上, c 相对于 d 是光滑的。这个MLP结构, 如图3所示, 处理了场景的位置

x 和观看方向 d 不对称地: d 被注入到靠近MLP末端的网络中, 这意味着有更少的MLP参数, 以及更少的非线性激活, 参与了观看依赖效应的创建。此外, 用于编码观看方向的傅里叶特征只包含低频分量, 即 $v^4(\cdot)$

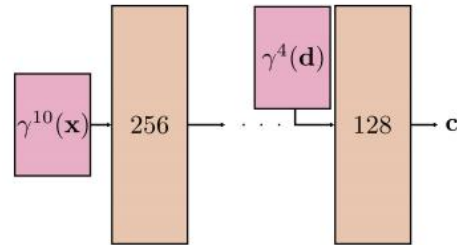


图3: 辐射 c 的NeRF MLP结构。造型

	扫描65			扫描106			扫描118		
	tLPIPS	"ssim	"PSNR	tLPIPS	"ssim	"PSNR	tLPIPS	"ssim	信号-噪音功
NeRF MLP香	0.0258/0.08	0.988/0.9	33.47/28.	0.0641/0.09	0.978/0.9	35.12/30.	0.0380/0.06	0.987/0.9	率比
草MLP	25	67	41	62	59	65	38	69	37.53/31.82
	0.0496/0.139	0.975/0.93	28.55/21.65	0.116/0.173	0.948/0.91	30.29/26.38	0.0609/0.094	0.979/0.95	35.31/29.58
		7			5			2	

表1: 在DTU场景上 (Jensen等人, 2014), 用普通的MLP取代NeRF的MLP显著减少了对新视图的泛化。我们使用与Riegler & Koltun (2020) 相同的数据分割。左边的数字是插值, 右边的数字是外推。他们在背景被掩盖的完整图像上进行评估。

对抗 $V^{10}(\cdot)$ 用于编码 d vs. x (见等式3). 换句话说, 对于一个固定的 x , 辐射 $c(x, d)$ 对 d 有有限的表达性。

为了验证这一假设, 我们进行了一个实验, 用普通的MLP对称处理 x 和 d 来表示 c ——也就是说, 接受两者作为第一层的输入, 用 V 编码 $V^{10}(\cdot)$ —消除从网络结构中产生的任何涉及观察方向的隐式先验。如果我们用这个替代模型从头训练NeRF, 我们观察到与NeRF的特殊MLP相比, 测试图像质量降低, 如图4和表1所示。这一结果与我们的假设相一致, 即NeRF的MLP辐射度 c 模型中反射率的隐式正则化有助于恢复正确的解。

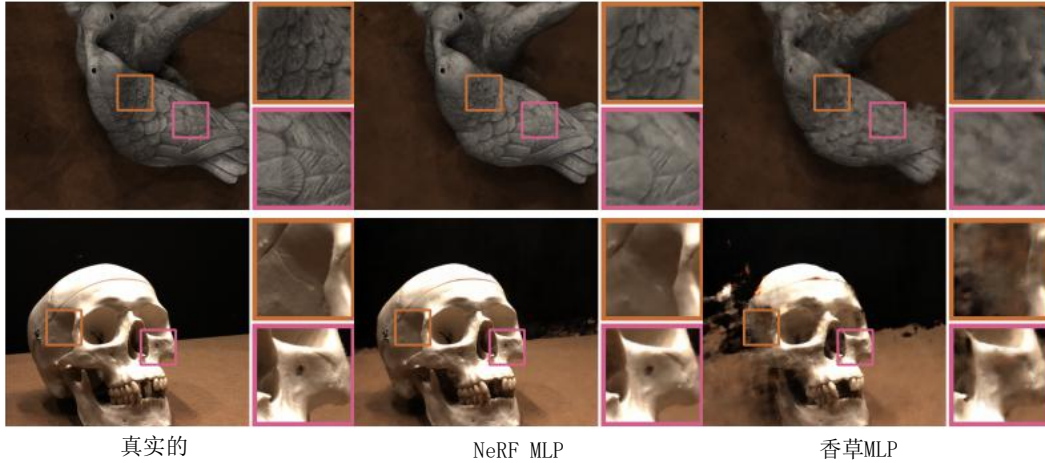


图4: 在DTU场景中 (Jensen等人, 2014; 里格勒和科尔顿, 2020年), 该图显示了用香草MLP替换辐射场 c 的NeRF模型的效果 (同时保持结构相同, 并从头开始训练两个场)。普通的MLP损害了NeRF概括到新观点的能力。

4倒球参数化

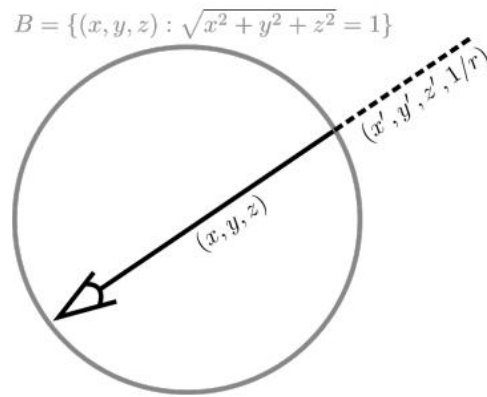
等式中的体积渲染公式2整合了欧几里得的深度。当真实场景深度的动态范围较小时, 该积分可以用有限数量的样本进行很好的数值逼近。然而, 对于户外, 360^0 捕捉时, 围绕着附近的物体, 也可以观察到周围的环境, 动态的深度范围可以非常大, 如背景 (建筑、山脉、云等)。可以任意遥远。如此高的动态深度范围导致了NeRF的体积场景表示中严重的分辨率问题, 因为为了合成逼真的图像, 在等式中的积分2在前景和背景区域都需要足够的分辨率, 这对于根据三维空间的欧氏参数化来简单地采样点来实现是具有挑战性的。图5演示了在场景覆盖和捕捉细节之间的权衡。在一个更有限的场景中, 所有的摄像机都朝向一个平面, 将摄像机与场景内容分开, NeRF通过投影映射欧几里得空间的一个子集, 即参考摄像机的视图, 到标准化设备坐标 (NDC) (McReynolds & Blythe, 2005), 并在这个NDC空间中集成。然而, 这种NDC参数化也从根本上限制了可能的观点, 因为它没有覆盖参考视图障碍之外的空间。



(a) 边界体积为卡车只有 (b) 边界体积为整个场景

图5: 360°对无界场景的捕获, NeRF的空间参数化要么只建模场景的一部分, 导致背景元素(a)中的显著伪影, 要么对整个场景进行建模, 但由于有限的采样分辨率(b). 而遭受整体的细节损失

我们用倒置来解决这个限制
便于自由查看的球体参数化
综合在我们的代表中, 我们首先进行部分-
将场景空间分成两卷, 一个是内部的
单位球体和外部体积表示
一个覆盖了...的补体的反向球体
内部体积(有关插图, 请参见图6
和图7, 为一个真实世界的场景示例
以这种方式建模)。内部卷图标
包含了前景和所有的照相机, 而
外部卷包含其余的
环境



这两个卷是用两个9月-
仲裁nerf. 为了渲染一个射线的颜色, 他们图6: NeRF++应用不同的参数化-
是单独的石膏, 然后是最后的网站-单元内外场景内容的文件
position. 不需要重新参数化球
内部的NeRF, 因为场景的那部分很好
有界限的对于外部的NeRF, 我们应用了一个在-
转换球体参数化。

具体来说, 一个3D点 (x, y, X) , $T = \sqrt{x^2 + y^2 + X^2}$ 外部卷中的 >1 可以重新
参数化由四倍 $(xyx \setminus \setminus, 1/T)$, $\sqrt{x^2 + y^2 + X^2} = 1$, 其中 $(x \setminus, y \setminus, X \setminus)$ 是一个与 (x, y, X)
相同方向的单位矢量, 表示球体上的一个方向, $0 < 1/T < 1$ 是沿这个方向的逆半径, 指定点
 $T \cdot (x \setminus, y \setminus, X \setminus)$ 在球体之外。与欧几里得空间不同, 物体可以与原点保持无限的距离, 重新参
数化的四重数中的所有数字都是有界的, i. e., $x \setminus, y \setminus, X \setminus \in [-1, 1]$, $1/T \in [0, 1]$. 这
不仅提高了数值的稳定性, 而且还尊重了更远的物体应该得到更低的分辨率的事实。我们可以
直接用光线投射这个4D有界的体积(只有3个自由度)来渲染相机射线的颜色。请注意, 前景和
背景的复合物相当于打破了等式中的积分2分为两部分, 整合在内部和外部卷内。特别地, 考虑
光线=被单位球划分为两部分: 在第一个, $t \in (0, t \setminus)$ 在里面



图7: NeRF++将前景和背景的建模分开。NeRF++ (a) 合成的图像是预测的前景(b)和背景(c). 的合成

在第二个球体中， $t \in (t', 1)$ 在球体之外。我们可以在等式中重写体积渲染积分2作为

$$C(r) = \underbrace{\int_{t=0}^{t'} \sigma(\mathbf{o} + t\mathbf{d}) \cdot c(\mathbf{o} + t\mathbf{d}, \mathbf{d}) \cdot e^{-\int_{s=0}^t \sigma(\mathbf{o} + s\mathbf{d}) ds} dt}_{(i)} + \underbrace{e^{-\int_{s=0}^{t'} \sigma(\mathbf{o} + s\mathbf{d}) ds}}_{(ii)} \cdot \underbrace{\int_{t=t'}^{\infty} \sigma(\mathbf{o} + t\mathbf{d}) \cdot c(\mathbf{o} + t\mathbf{d}, \mathbf{d}) \cdot e^{-\int_{s=t'}^t \sigma(\mathbf{o} + s\mathbf{d}) ds} dt}_{(iii)}. \quad (4)$$

项(i)和(ii)在欧几里得空间中计算，而项(iii)在倒置的球面中计算

$\frac{1}{r}$ 以空间作为积分变量。换句话说，我们使用 \mathbf{a} 在 $(\mathbf{o} + t\mathbf{d})$ ， c 在 $(\mathbf{o} + t\mathbf{d}, \mathbf{d})$ 的条款

(i)和(ii)，以及 \mathbf{a} 出局 $(x', y', z', 1/r)$ ， c 出局 $(x', y', z', 1/r, \mathbf{d})$ 期限(iii)。

为了计算光线 $r = \mathbf{o} + t\mathbf{d}$ 的项(iii)，我们首先需要能够计算 \mathbf{a} 出局， c 出局换句话说，我们需要一种计算的方法 (x', y', z') 对应于一个给定的 $1/r$ ，因此，该 \mathbf{a} 出局， c 出局可以采取 $(x', y', z', 1/r)$ 作为输入。这可以实现。

如图所示。8、让光线与

单位球面在点 \mathbf{a} 处，和中点的中点

与光线对齐的和弦是重点

b. 点一个 \mathbf{b} 或 t_{ad} 由 solv -计算

$\text{ing } |\mathbf{o} + t\mathbf{d}| = 1$ ，而 $\mathbf{b} = \mathbf{o} + t\mathbf{d}$ 是

通过求解 d 获得 $T(\mathbf{o} + t\mathbf{d}) = 0$ 。然后

获取 (x', y', z') 给定 $1/r$ ，我们可以旋转

向量 \mathbf{a} 沿着向量 \mathbf{bd} 的角度 α

$\alpha = \arcsin |\mathbf{b}| - \arcsin(|\mathbf{b}| \cdot \frac{1}{r})$ 。一旦我们可以

评估一个出局， c 出局在任何一个 $1/r$ ，我们只是 sam -

从这个区间中得到有限数量的点

$[0, 1]$ 来计算项(iii)。

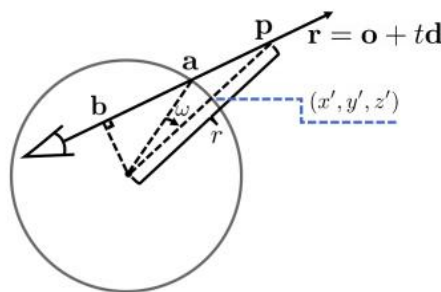


图8：说明 (x', y', z') ，对于倒球参数化中已知 $1/r$ 的点 p 。

的反向球面参数化

外部体积有一个直观的物理解释-

国家它可以用虚拟的角度来看待

图像平面为单位球的相机

场景的起源。因此，3D点 (x, y, z) 被投影到像素 (x', y', z') ，而术语 $1/r^2 (0, 1)$ 作为该点的(逆)深度或视差。从这个角度来看，仅适用于正向捕获的NDC参数化与我们的表示有关，因为它使用了一个虚拟的针孔相机，而不是一个球形的投影面。从这个意义上说，我们的倒置球参数化与最近在视图合成工作中提出的多球图像(由嵌套的同心球采样的场景表示)的概念有关(Attal等人, 2020; Broxton等人, 2020)。

5个实验

我们验证了NeRF++，并将其与使用手持相机捕获的两个真实数据集上的NeRF进行了比较：坦克和寺庙(T&T)(Knapitsch等人, 2017)和Ycer等人的光场(LF)数据集。ü(2016). 我们报告了PSNR、SSIM和LPIPS(Zhang et al., 2018)作为我们测量合成测试图像质量的定量指标。

T&T数据集。我们使用由Riegler&Koltun(2020年)提供的训练/测试图像和SfM姿态。这个数据集由手持的360个数据集组成^o捕捉四个大规模场景；相机的姿势由COLMAP SfM估计(施恩伯格&Frahm, 2016)。o对于NeRF，我们将场景归一化，使所有的摄像机都在半径的球体内。1/8这种规范化功能确保了单元球体覆盖了大部分的背景场景内容(尽管一些背景几何图形仍然位于边界单元球体之外)。为了数值计算每个相机射线的体积渲染积分，我们均匀地从射线原点到它与单位球的交点的采样点。由于NeRF++射线同时投射内部和外部体积，因此每个相机射线使用的样本数量是单个体积的两倍，为了公平起见，我们也将NeRF使用的样本数量增加了一倍。具体来说，NeRF的粗级MLP使用了128个均匀的样本，而

	卡车			列车			M60			操场		
	tLPIPS	“ssim”	psnr	tLPIPS	“ssim”	psnr	tLPIPS	“ssim”	psnr	tLPIPS	“ssim”	psnr
NeRF	0.513	0.747		0.651	0.635		0.602	0.702		0.529	0.765	
		20.85			16.64			16.86			21.55	
NeRF++	0.298	0.823		0.523	0.672		0.435	0.738		0.391	0.799	
		22.77			17.17			17.88			22.37	
	非洲			篮子			火炬			船		
	tLPIPS	“ssim”	psnr	tLPIPS	“ssim”	psnr	tLPIPS	“ssim”	psnr	tLPIPS	“ssim”	psnr
NeRF	0.217	0.894	26	0.377	0.805	20	0.347	0.811	22	0.372	0.801	23
		16			83			81			24	
NeRF++	0.163	0.923		0.254	0.884		0.226	0.867		0.241	0.867	
		27.41			21.84			24.68			25.35	

表2：我们比较了NeRF++和NeRF在四个坦克和寺庙上的场景：卡车、火车、M60和操场，以及四个LF场景：非洲、火炬、船和篮子。NeRF++在所有指标上都始终优于NeRF。

精细级的MLP使用了256个额外的重要样本。在这种超参数设置下，NeRF在训练和测试过程中的计算成本和GPU内存占用与+++大致相同。我们在每次训练迭代中随机抽取2048条摄像机射线，并训练NeRF和NeRF++进行250k次迭代，学习速率为 $5e-4$ 。

LF数据集。我们使用了来自LF数据集的四个场景：非洲、篮子、船和火炬。每个场景都被2k-4k手持式捕获的图像密集覆盖，并使用SfM恢复相机参数。我们通过对图像进行40倍的时间子采样来构造一个稀疏的周围捕获。特别地，训练图像是第0、40、80、...帧，而测试图像是第20、60、100帧。.....场景归一化方法、每条摄像机射线的样本数量、批处理大小、训练迭代次数和学习率与T&T数据集相同。

结果如表2所示，NeRF++在涉及360的具有挑战性的场景中显著优于NeRF⁰捕获大规模无界场景中的对象。在图9中，我们可以看到，由NeRF++合成的图像具有明显更高的保真度。

6个开放的挑战

NeRF++改进了无界场景的参数化，其中前景和背景都需要真实地表示。然而，仍存在一些公开的挑战。首先，在NeRF和NeRF++的一个单一的大规模场景上的训练和测试是相当耗时和记忆密集型的。在一个使用4个RTX 2080Ti GPU的节点上训练NeRF++需要24小时。~在一个这样的GPU上渲染一个1280x720的图像在测试时需要30秒。~刘等人。（2020）加快了推理，但渲染仍远非实时。其次，较小的相机校准误差可能会阻碍逼真的合成。可以应用稳健的损失函数，如上下文损失（Mechrez et al., 2018）。第三，光度效果，如自动曝光和渐晕也可以考虑，以提高图像保真度。这一调查线与马丁-布鲁拉等人在正交工作中提到的照明变化有关。（2020）。

参考文献

本杰明阿塔耳，赛琳娜林，亚伦哥卡斯兰，克里斯蒂安理查特，和詹姆斯汤普金。MatryODShka：实时6自由度使用视频视图合成多球图像。在ECCV，2020年。

迈克尔·布罗克斯顿、约翰·弗林、瑞恩·奥弗贝格、丹尼尔·埃里克森、彼得·海德曼、马修·杜瓦尔、杰森·杜尔加里安、杰伊·布施、马特·惠伦和保罗·德贝维克。沉浸式光场视频与分层网格表示。ACM图形事务处理，第39(4)页，2020年。

克里斯·比勒，迈克尔·波斯，伦纳德·麦克米兰，史蒂文·戈特勒和迈克尔·科恩。非结构化的发光图形渲染。在签名，2001年。

陈申昌和威廉姆斯。查看图像合成的内插值。在签名，1993年。

保罗E. Debevec，卡米洛J. 泰勒和吉登德拉·马利克。从照片中建模和渲染体系结构：一种基于几何图形和图像的混合方法。在签名，1996年。



图9：我们定性地比较了两个T&T场景（卡车、游乐场）和两个LF场景（非洲、火炬）上的NeRF++和NeRF。NeRF++比NeRF产生更清晰的图像，能够更好地表示前景和背景。

史蒂文戈特勒，格泽斯祖克，理查德泽利斯基和迈克尔科恩。卢米格拉普。在签名，1996年。

拉斯姆斯·詹森，安德斯·达尔，乔治·沃吉亚齐斯，恩吉尔·托拉和亨里克·Aanæs。大尺度多视角立体视觉评价。在CVPR，2014年。

詹姆斯T Kajiya和布莱恩P冯赫尔森。光线追踪体积密度。在签名，1984年。

国王、朴杰熙、周千毅、科尔敦。坦克和寺庙：基准测试大规模的場景重建。ACM图形事务处理，第36(4)页，2017年。

马克·列沃伊和帕特·汉拉汉。灯光字段渲染。在签名，1996年。

刘灵杰、顾嘉道、林觉觉、蔡达成、神堡。神经稀疏体素字段。arXiv:2007.11571, 2020.

里卡多马丁-布鲁拉，诺哈拉德万，迈赫迪S萨贾迪，乔纳森T巴伦，阿列克谢多索维茨基，和丹尼尔达克沃斯。野外的NeRF：无约束照片集合的神经辐射场。arXiv:2008.02268, 2020.

汤姆·麦克雷诺兹和大卫·布莱斯。使用OpenGL的高级图形编程。爱思唯尔，2005年。

罗伊·梅克雷斯基, 伊塔玛尔·塔尔米, 和利希·泽尔尼克庄园。使用非对齐数据的图像转换的上下文丢失。在ECCV, 2018年。

本·米尔登霍尔, 普拉图尔。斯里尼瓦桑, 马修·坦西克, 乔纳森·T.巴伦, 拉维拉马莫蒂和任。NeRF: 将场景表示为视图合成的神经辐射场。在ECCV, 2020年。

格诺特·里格勒和弗拉德伦·科尔顿。自由视图合成。在ECCV, 2020年。

约翰内斯卢茨施恩伯格和简-迈克尔弗拉姆。从运动结构重新审视。在CVPR, 2016。

沈先生和唱冰康。回顾一下基于图像的渲染技术。在视觉通信和图像处理方面, 2000年。

马修·坦西克、斯里尼瓦桑、本米尔登霍尔、萨拉弗里多维奇-基尔、尼廷拉加万、乌特卡什辛哈尔、拉维拉莫蒂、乔纳森·巴伦和任吴。傅里叶特征让网络在低维域中学习高频函数。
arXiv:2006.10739, 2020.

丹尼尔·伍德, 丹尼尔·阿祖马, 肯阿尔丁格, 布莱恩柯利斯, 汤姆杜尚, 大卫H萨莱辛, 和沃纳斯图茨尔。用于三维摄影的表面光场。在签名, 2000年。

凯恩耶、亚尔山大霍农、王和奥尔加霍农。从密集采样光场的高效三维目标分割, 并应用到三维重建。*ACM图形事务处理, 第35(3)页, 2016年。*

张理查德, 菲利普伊索拉, 阿列克谢埃弗罗斯, 伊莱谢克特曼, 和王奥利弗。深度特征作为一种感知度量的不合理有效性。在CVPR, 2018年。