

Nerf-04：动态场景NeRF（dynamic）

1.D-NeRF: Pumarola A, Corona E, Pons-Moll G, et al. D-nerf: Neural radiance fields for dynamic scenes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10318-10327.

2. Hdr-nerf: Huang X, Zhang Q, Feng Y, et al. Hdr-nerf: High dynamic range neural radiance fields[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 18398-18408.

3. NR-NeRF: Tretschk E, Tewari A, Golyanik V, et al. Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 12959-12970.

第一篇：D-Nerf：Neural radiance fields for dynamic scenes（引入宏观时间信息）

回答如下问题【带着问题去思考】：

- 理解 deformation 网络如何建模物体的移动。你认为此文章的存在什么问题？

阅读过程中笔记:

1.核心：

1) 解决了什么问题：

NeRF只能重建静态场景，论文提出的方法可以把神经辐射场扩展到动态领域，可以在单相机围绕场景旋转一周的情况下重建物体的刚性和非刚性运动。

2) 主要贡献：

把时间作为附加维度加到输入中，同时把学习过程分为两个阶段：第一个把场景编码到规范空间，另一个在特定时间把这种规范表达形式map到变形场景中。两个map都用mlp学习，训练完后可以通过控制相机视角和时间变量达到物体运动的效果。

3) 实验效果如何：

Method	Hell Warrior				Mutant				Hook				Bouncing Balls			
	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓
NeRF	44e-3	13.52	0.81	0.25	9e-4	20.31	0.91	0.09	21e-3	16.65	0.84	0.19	94e-4	20.26	0.91	0.2
T-NeRF	47e-4	23.19	0.93	0.08	8e-4	30.56	0.96	0.04	18e-4	27.21	0.94	0.06	16e-5	37.81	0.98	0.12
D-NeRF	31e-4	25.02	0.95	0.06	7e-4	31.29	0.97	0.02	11e-4	29.25	0.96	0.11	12e-5	38.93	0.98	0.1

Method	Lego				T-Rex				Stand Up				Jumping Jacks			
	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓
NeRF	9e-3	20.30	0.79	0.23	3e-3	24.49	0.93	0.13	1e-2	18.19	0.89	0.14	1e-2	18.28	0.88	0.23
T-NeRF	3e-4	23.82	0.90	0.15	9e-3	30.19	0.96	0.13	7e-4	31.24	0.97	0.02	6e-4	32.01	0.97	0.03
D-NeRF	6e-4	21.64	0.83	0.16	6e-3	31.75	0.97	0.03	5e-4	32.79	0.98	0.02	5e-4	32.80	0.98	0.03

Table 1: Quantitative Comparison. We report MSE/LPIPS (lower is better) and PSNR/SSIM (higher is better).

4) 存在的问题：

引入时间t作为维度，而t在其官方提供的数据集中取值范围是[0,1]的浮点数，对于每一个相机视角而言需要新增一个时间t参数，那么该参数应该如何标注？官方并没有给出回答。

经查阅，有人提到是手工标注，啊这

2. 具体method细节

1.规范网络canonical Network和变形网络deformation Network：拆分 $(x, y, z, \theta, \phi, t) \rightarrow (rgb, \sigma)$ 为：

1) deformation Network Ψ_t ：

$$\Psi_t(x, t) = \begin{cases} \Delta x, & \text{if } t \neq 0 \\ 0, & \text{if } t = 0 \end{cases} \quad (1)$$

可以认为是在不同时刻将 $\Psi_t : (x, t) \rightarrow \Delta x$

2) canonical Network Ψ_x ：可以理解为原始nerf

将由上变形网络得到的 $(\Delta x, \Delta y, \Delta z)$ 送入网络，即：

$\Psi_x : (x + \Delta x, y + \Delta y, z + \Delta z, \theta, \phi) \rightarrow (rgb, \sigma)$.

2.加入时间t参数的体渲染：

$$\begin{aligned} C(p, t) &= \int_{h_n}^{h_f} \mathcal{T}(h, t) \sigma(\mathbf{p}(h, t)) \mathbf{c}(\mathbf{p}(h, t), \mathbf{d}) dh, \\ \text{where } \mathbf{p}(h, t) &= \mathbf{x}(h) + \Psi_t(\mathbf{x}(h), t), \\ [\mathbf{c}(\mathbf{p}(h, t), \mathbf{d}), \sigma(\mathbf{p}(h, t))] &= \Psi_x(\mathbf{p}(h, t), \mathbf{d}), \\ \text{and } \mathcal{T}(h, t) &= \exp \left(- \int_{h_n}^h \sigma(\mathbf{p}(s, t)) ds \right). \end{aligned}$$

3.损失函数

$$\mathcal{L} = \frac{1}{N_s} \sum_{i=1}^{N_s} \|\hat{C}(p, t) - C(p, t)\|_2^2 \quad (2)$$

第二篇：Hdr-nerf : High dynamic range neural radiance fields

[阅读过程中笔记:](#)

1.核心：

1) 解决了什么问题：

1.任务：从一组具有不同曝光（曝光定义为曝光时间和辐射度的乘积）的LDR（低动态范围Low Dynamic Range）图像中恢复高动态范围神经辐射场。

2.目前nerf中所获得的辐射动态范围仍在一个低动态范围里(0,255)，而现实的物理世界范围更高 $(0, +\infty)$.本论文解决了该问题，从辐射场中渲染恢复了高动态范围。

3.解决了传统重建HDR由固定姿势产生伪影的问题。

2) 主要贡献：

- 1.提出了对物理成像过程的建模：相机响应函数建模。
- 2.提出了端到端方法Hdr-nerf：场景点的辐射度通过两个隐式函数（辐射场、色调映射器）转换为LDR图像中的像素值。
- 3) 实验效果如何:

Table 1. Quantitative comparisons with baseline methods on synthetic and real scenes. Metrics are averaged over the scenes from our dataset (per-scene metrics are shown in supplementary material). LDR-OE denotes the LDR results with exposure t_1 , t_3 , and t_5 . LDR-NE denotes the LDR results with exposure t_2 , and t_4 . HDR denotes the HDR results. We color code each column as best and second best.

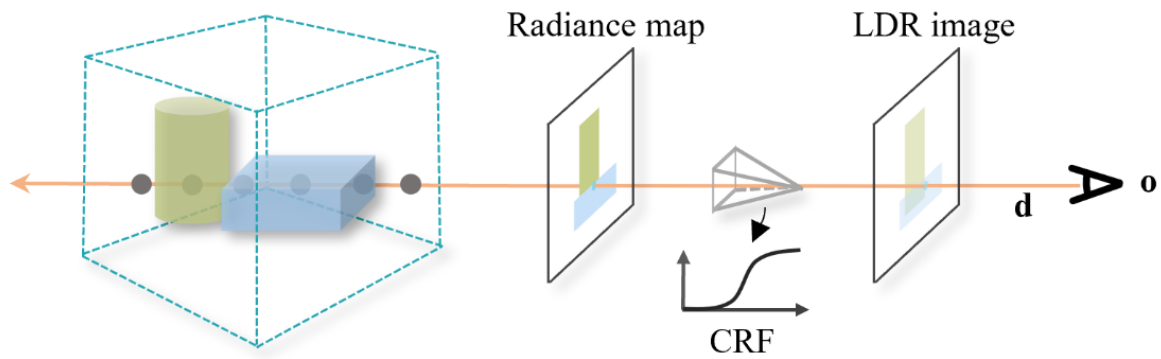
		LDR-OE (t_1, t_3, t_5)			LDR-NE (t_2, t_4)			HDR		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
NeRF [42]	Syn.	13.97	0.555	0.376	—	—	—	—	—	—
	Real	14.95	0.661	0.308	—	—	—	—	—	—
NeRF-W ^l [37]	Syn.	29.83	0.936	0.047	29.22	0.927	0.050	—	—	—
	Real	28.55	0.927	0.094	28.64	0.923	0.089	—	—	—
NeRF-GT ² [42]	Syn.	37.66	0.965	0.028	35.87	0.955	0.032	37.80	0.964	0.029
	Real	34.55	0.958	0.057	34.59	0.956	0.051	—	—	—
Ours†	Syn.	—	—	—	—	—	—	—	—	—
	Real	30.37	0.944	0.075	29.37	0.938	0.078	—	—	—
Ours	Syn.	39.07	0.973	0.026	37.53	0.966	0.024	36.40	0.936	0.018
	Real	31.63	0.948	0.069	31.43	0.943	0.069	—	—	—

2. 具体method细节

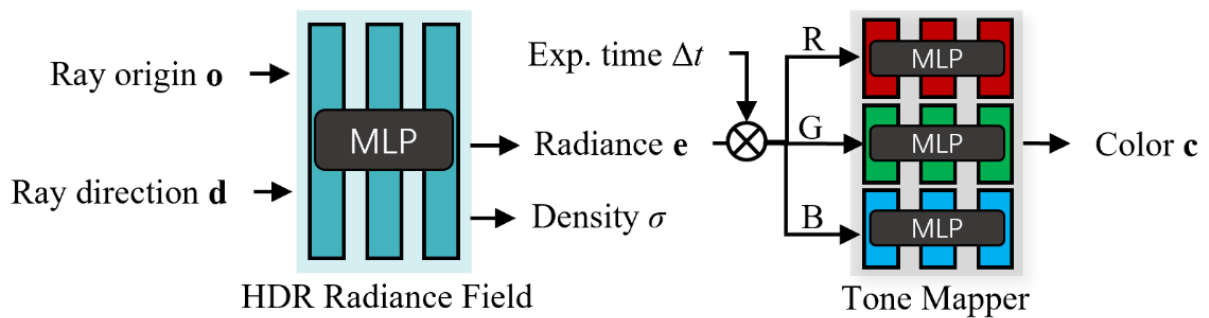
核心：是对捕获场景亮度并将它们映射到像素值的过程进行建模。

- 1.将场景点的辐射度通过两个隐式函数（辐射场、色调映射器）转换为LDR图像中的像素值：
 - 1) **辐射场**：对场景辐射度进行编码(值从0到 $+\infty$ 变化)，通过给出 相应的射线源和射线方向 输出射线的密度和辐射度。
 - 2) **色调映射器**：对映射过程进行建模，射线照射到相机传感器上成为像素值。通过将**辐射度**和**相应的曝光时间**输入到色调映射器来预测光线的颜色。

Physical process



HDR-NeRF



2.体渲染：使用传统体渲染技术。

3.损失函数

仅使用输入的LDR的图像作为监督。

总的loss 为： $\mathcal{L} = \mathcal{L}_{color} + \lambda_D \mathcal{L}_u$

第三篇：NR-NeRF: Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video

阅读过程中笔记:

1.核心:

1) 解决了什么问题:

- 1.仅需要动态场景的单目视频作为输入（不需要多视图输入）
- 2.可以渲染动态体积的变形或非变形部分。

2) 主要贡献:

1. 使用两个组件表示非刚性场景：（1）用于捕获几何和外观的规范神经辐射场；（2）场景变形场。
2. 场景变形是通过光线弯曲实现的，光线弯曲由MLP进行建模。
3. 为每个点分配刚度得分，允许变形不影响场景的静态区域。
4. 引入多个正则化项作为额外的软约束：（1）限制体积内变形的大小；（2）保持局部形状。

3) 实验效果如何:

		Ours	Naïve	Rigid (cond.)	Rigid (no cond.)	NV	NV (mod.)
PSNR	↑	<i>24.70</i>	25.83	22.24	21.88	14.13	14.10
SSIM	↑	0.758	<i>0.738</i>	0.662	0.659	0.259	0.263
LPIPS	↓	0.197	<i>0.226</i>	0.309	0.313	0.580	0.583

2. 具体method细节

不同于传统nerf经典流程的特点是：

- 1) Deformation Model
- 2) loss函数：引入了多个正则化项来额外约束体积内变形的大小，以此来保持局部形状。