



基於二元輻射場之混合編碼對動態神經輻射場的優化

Dynamic Hybrid Encoding for Binary Radiance Fields

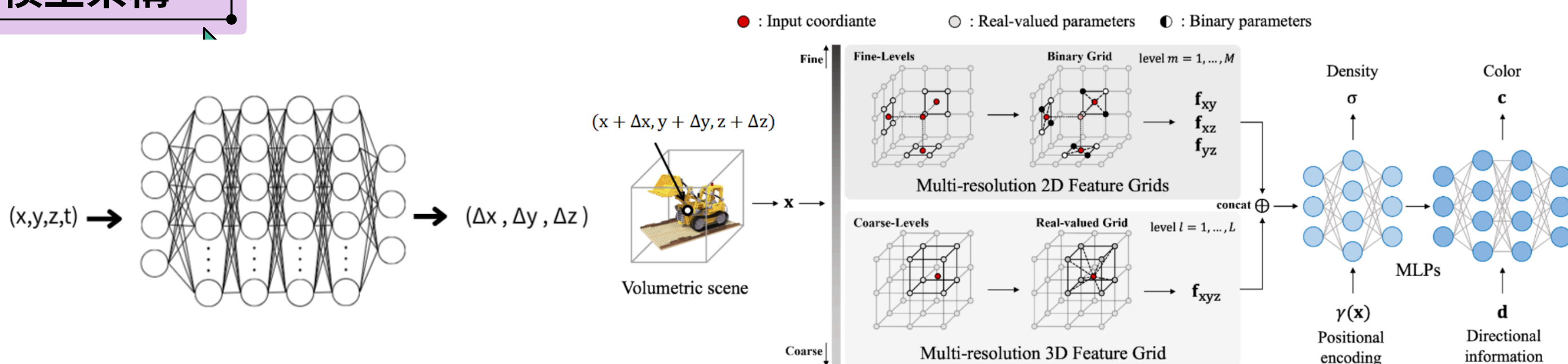
指導教授: 江瑞秋教授

專題生: 賴奕臻、林寬皇

摘要

本研究提出一種低訓練時長、低參數儲存量的高渲染品質動態神經輻射場 (Neural Radiance Fields, 簡稱NeRF) 之新視角合成(Novel View Synthesis)技術, 稱為Dynamic Hybrid BiRF, 我們通過動靜分離的變形網路與典型網路乘載包含時空資訊的輸入影像, 並將場景使用雙重混合架構表示, 有效利用網格儲存特徵限縮了後續所需學習的神經網路規模, 同時搭配二元化高解析度的特徵參數與哈希表避免模型參數大幅增加。研究結果中, 與現有同樣應用變形/典型網路的動態神經輻射場技術D-NeRF架構相比, 我們獲得了更精細的渲染品質, 提升約13%, 與僅僅1/3的訓練時長, 同時模型大小的增長也限制在2倍之內。

模型架構



實作方式

在實現中, 網路會學習時間資訊為0的訊息作為典型網路, 其可視為一個標準靜態場景, 而變形網路中學習的位移量則表示當前時空位置與典型網路中的差向量。

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

因此對任意時空位置(x,t), 需通過將位置與位移量相加, 將其映射回典型網路後再進行取樣。兩個神經網路會同步進行學習並持續優化。

接著在典型網路中, 我們使用了雙重混合架構:

1. 粗細層次分級

我們使用多解析度網格描述三維場景, 其中低解析度族群直接使用三維網格儲存特徵, 而高解析度族群會通過將三維資訊投影到三個方向上的二維平面, 大幅減少參數儲存量

$$N^3 \rightarrow N^2 \times 3, N > \text{threshold}$$

2. 特徵編碼型態

此外, 對於二維平面儲存的高頻細節, 對於場景精緻度仍有貢獻, 但影響相對較小, 因此我們使用二元化替代昂貴的浮點數參數

$$\theta' = \text{sign}(\theta) = \begin{cases} +1 & \text{if } \theta \geq 0 \\ -1 & \text{if } \theta < 0 \end{cases}$$

最終特徵會以Hash的方式進行儲存, 並接續通過兩個小型多層感知器學習出密度與顏色訊息, 呈現出影像。

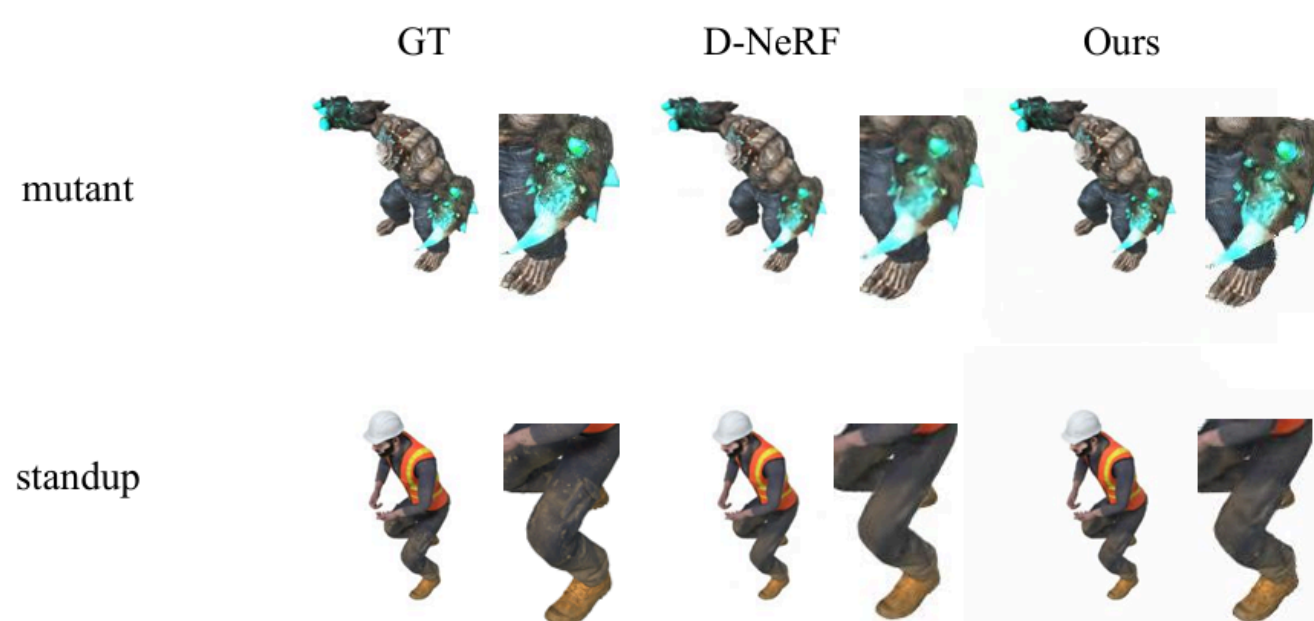
$$f = \{f_{xyz}, f_{xy}, f_{xz}, f_{yz}\}$$

$$(\sigma, e) = \text{MLP}_{\text{density}}(\gamma(x), f) \quad c = \text{MLP}_{\text{color}}(e, d)$$

專題成果

我們使用了D-NeRF提供的合成場景資料集進行評估, 並整理了多種方法在該資料集中的平均表現, 並列出部分Ground Truth與渲染影像以供對比。

Performance Methods	D-NeRF Synthetic Dataset		
	PSNR ↑	SSIM ↑	LPIPS ↓
NeRF	19.00	0.87	0.18
T-NeRF	29.50	0.95	0.08
D-NeRF	29.67	0.95	0.07
Ours	30.21	0.96	0.10



結論

在本研究中我們針對純隱式動態場景進行優化, 引入雙重混合架構, 以改善深層多層感知器帶來的高額訓練時間, 成功為渲染品質、訓練速度帶來了提升。與此同時我們的模型增長了較小的幅度, 如何通過模型壓縮進一步得到全方面的提升會是我們持續關注的課題。