Robustifying the Multi-Scale Representation for Neural Radiance Fields (RM-NeRF)

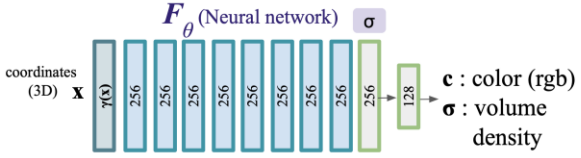
Nishant Jain Suryansh Kumar Luc Van Gool

BMVC 2022

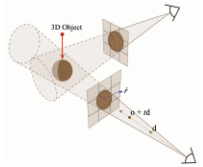
ETH zürich | CVL

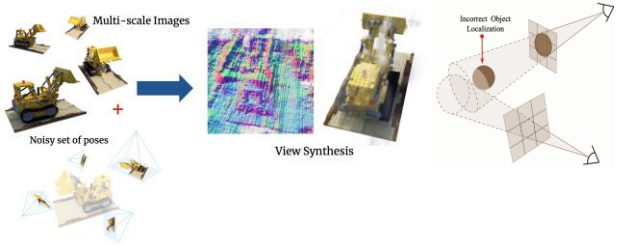
Overview

Mip-NeRF[1]

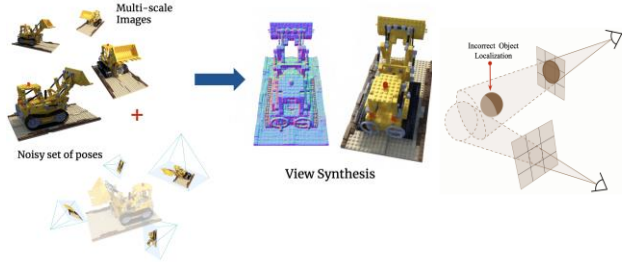


$$\gamma^*(\mathbf{o},\mathbf{d},\dot{r},t_0,t_1)=\frac{\int\gamma(\mathbf{x})\mathbf{F}(\mathbf{x},\mathbf{o},\mathbf{d},\dot{r},t_0,t_1)d\mathbf{x}}{\int\mathbf{F}(\mathbf{x},\mathbf{o},\mathbf{d},\dot{r},t_0,t_1)d\mathbf{x}}$$





RM-NeRF



- Handles camera pose estimation and multi-scale representation difficulties in day-to-day multi-view images.
- Effective optimization of a joint pose and rendering loss function, assuming the scene is rigid.

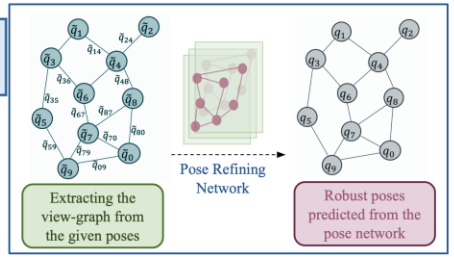
Approach

Pose-refining : Rotation Averaging[2]

Image-Set as a view graph, vertex denotes rotation of each image, edge is the relative rotation b/w 2.

$$\underset{\Theta}{\operatorname{argmin}} \sum_{\mathcal{G}\in\mathcal{D}} \sum_{\mathcal{E}_{ij}\in\mathcal{E}} d(R_{ij}^f, R_{ij}) + \beta \sum_{\mathcal{V}_j\in\mathcal{V}} d(R_j^f, R_j)$$

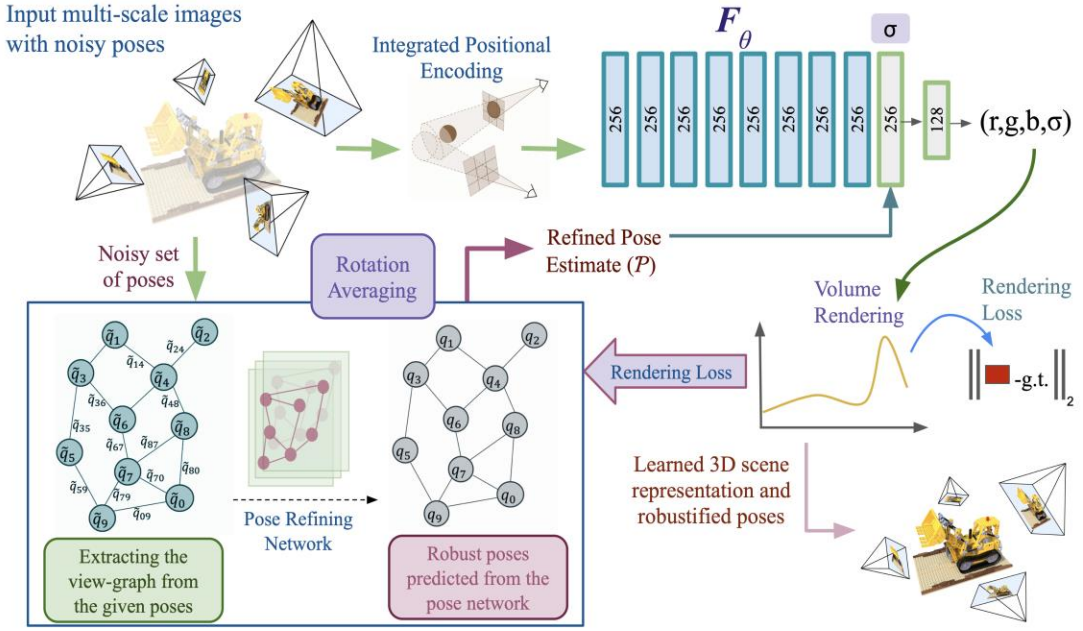
params | view graphs set | edge error | relative rotations | vertex set | absolute rotations



Proposed Cost Function and Pipeline

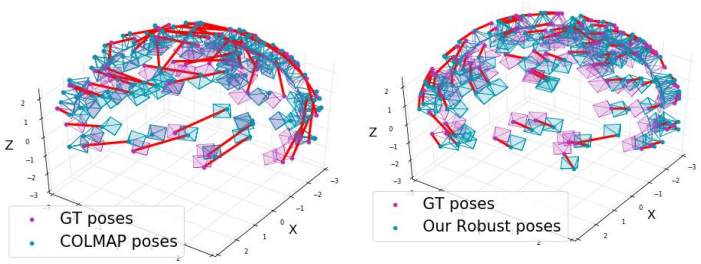
$$\mathcal{L}=\overbrace{\sum_{\mathbf{r}\in\mathcal{R}}\|C(\mathbf{r})-\hat{C}(\mathbf{r})\|_2^2}^{\text{rendering loss}}+\overbrace{\sum_{\mathcal{E}_{ij}\in\mathcal{E}} d_Q(q_{ij}^f,q_{ij})+\beta\sum_{\mathcal{V}_j\in\mathcal{V}} d_Q(q_j^f,q_j)}^{\text{pose-estimation loss1}}$$

quaternions

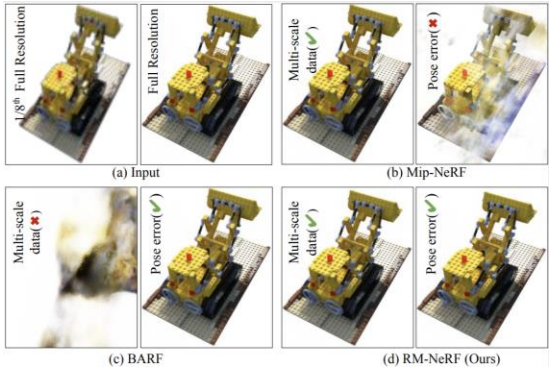


Results

Pose estimation with noisy keypoints



Multi-Scale Blender Data[1]



Perturbed Poses

	Lego		Ship		Drums		Mic		Chair	
	PSNR↑	LPIPS↓	PSNR↑	LPIPS↓	PSNR↑	LPIPS↓	PSNR↑	LPIPS↓	PSNR↑	LPIPS↓
Mip-NeRF	21.52	0.06	24.54	0.07	13.34	0.075	24.71	0.05	29.1	0.049
BARF	10.88	0.55	8.81	0.74	11.56	0.76	12.35	0.57	14.35	0.47
Base A	11.67	0.49	14.28	0.28	13.25	0.67	12.28	0.41	15.12	0.20
Base B	12.46	0.37	13.43	0.31	11.32	0.58	14.26	0.29	13.71	0.42
NeRF–	16.89	0.094	19.89	0.118	15.67	0.074	18.35	0.08	20.22	0.098
Base C	18.28	0.089	16.32	0.22	17.25	0.070	19.42	0.073	18.67	0.114
Ours	**27.01**	**0.044**	**26.59**	**0.067**	**26.07**	**0.043**	**32.8**	**0.008**	**35.23**	**0.031**

A,B: Mip-NeRF + BARF[3] C: Mip-NeRF+NeRF--[4]

[1]Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields. ICCV'21
[2]NeuRoRA: Neural Robust Rotation Averaging. ECCV'20
[3]BARF: Bundle Adjusting Radiance Fields. ICCV'21
[4]NeRF--: Neural Radiance Fields Without Known Camera Parameters. Arxiv'21