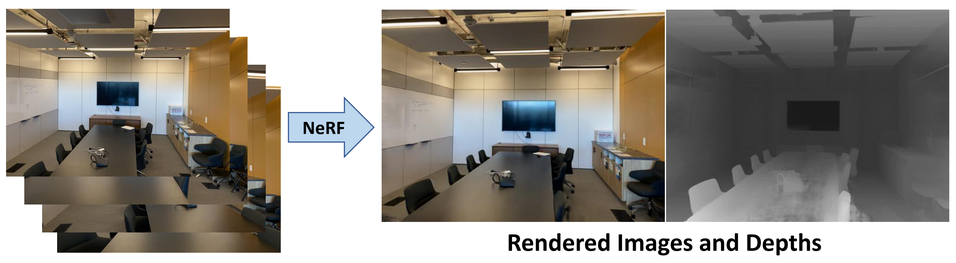
RoomNeRF: Representing Empty Room as Neural Radiance Fields for View Synthesis

ManKyu Kong, Seongwon Lee, Euntai Kim
School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea

Motivation

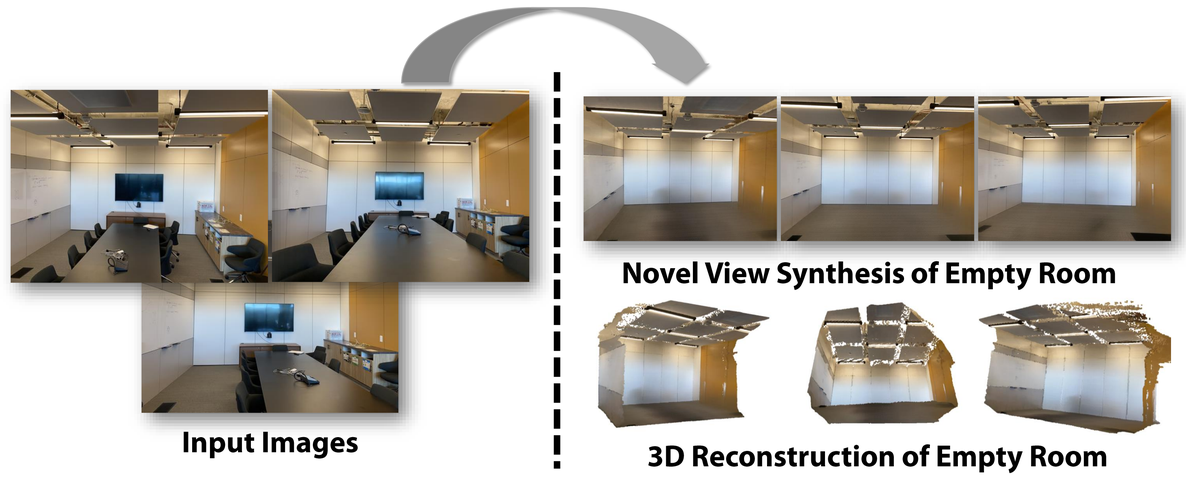
- One of the typical scenes to utilize NeRF is a **room.**
- The experience of being in a virtual room makes users curious about an empty room without furniture.



- **2D Inpainting techniques** are utilized to remove objects from room images.
- However, 2D inpainted images do not have **3D information**.
- The 3D information of an empty room can help NeRF to reconstruct a **visually and geometrically consistent** scene.

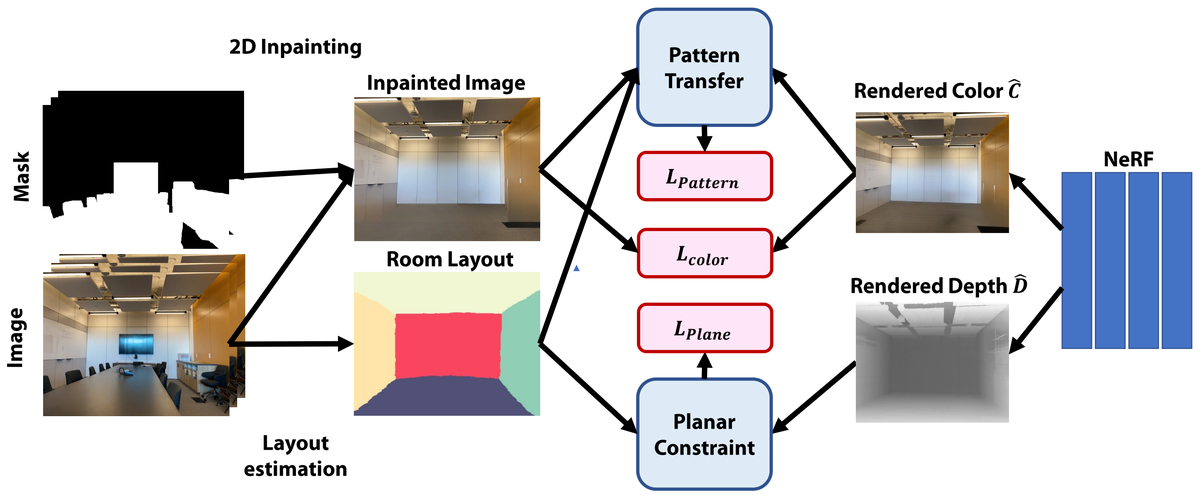
Problem

- Reconstructing **an empty room** with **the object-existing images**



Method

Overview



- Object masks, inpainted images, and room layouts are obtained in advance using Video object segmentation, 2D inpainting module, and layout estimation.

- Our Network exploits **the intrinsic properties of a room** shared by each plane
- Losses from Pattern transfer module and Planar Constraint Module help NeRF to reconstruct an empty room.

$$L_{PT} = \|M \odot \{\hat{R}_i - K_{j'}\}\|^2 \qquad j' = \underset{j \in (M^c \cap L)}{\operatorname{argmax}}\ d(Q_i, K_j),$$

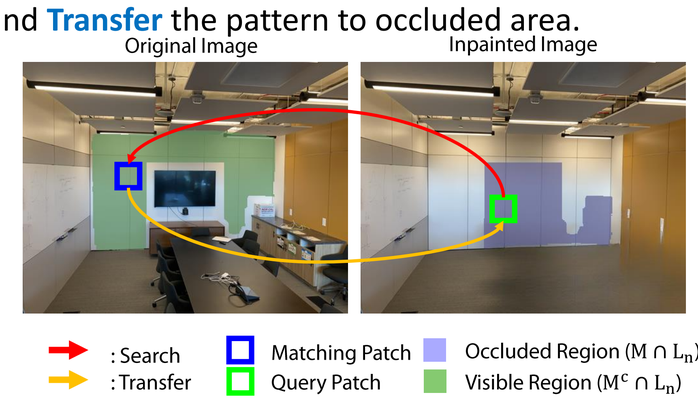
$$L_{PC} = \frac{1}{N_{Mask}} \sum_{n=1}^{N_{layout}} \sum_{D_n \in M \cap L_n} \left| \overrightarrow{A_n D_n} \times \overrightarrow{B_n D_n} \cdot \overrightarrow{C_n D_n} \right|$$

- Our total loss is as follows.

$$L_{total} = L_{color} + L_{depth} + \lambda_{PT} L_{PT} + \lambda_{PC} L_{PC}.$$

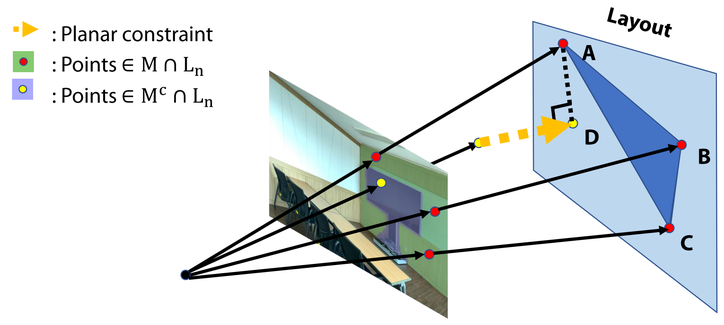
Pattern Transfer

- All the walls in the room all have a certain pattern.
- To perform a **consistent visual reconstruction**, PT transfers the visual pattern of room wall to occluded regions.
- Two Step : **Search** the closest pattern from visible region and **Transfer** the pattern to occluded area.



Planar Constraint

- The Manhattan world assumes that all walls are flat.
- The occluded part of the wall should be **geometrically consistent** as a plane.

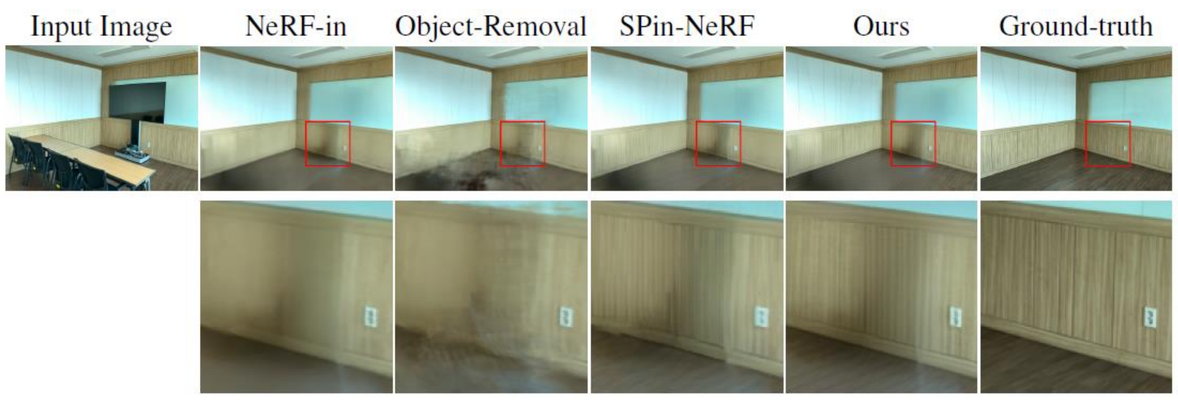


Experimental Results

Quantitative Results

	Seminar room			Office room		
	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)
Masked NeRF	21.69	0.8383	0.3781	20.61	0.9387	0.3409
Inpainted NeRF	23.44	0.8672	0.3722	21.07	0.9509	0.3199
NeRF-in [15]	23.35	0.8754	0.3420	21.42	0.8778	0.3310
Object-removal [31]	22.60	0.8591	0.3644	21.26	0.9425	0.3237
SPin-NeRF [20]	22.96	0.8643	0.2454	21.26	0.9467	0.2841
RoomNeRF (Ours)	**23.82**	**0.9148**	**0.1546**	**21.58**	**0.9580**	**0.2434**

Qualitative Results



Qualitative Results on LLFF and Replica dataset