Depth Light Field Training (DeLFT)

Author: Mihnea Toader
m.toader@student.tudelft.nl
Supervisors: Petr Kellnhofer, Michael Weinmann
Responsible professor: Elmar Eisemann

1 Background

- Neural Radiance Fields (NeRFs) create a 3D model from a set of 2D images of that object
- Rendering techniques are generally expensive for NeRF models
- Traditional rasterization does not require the full detailed model for some steps of the process, such as collision detection or shadow casting

2 Research topic

Given a NeRF or a 3D mesh, can we create a more compact neural representation that can directly produce the required silhouettes for any desired view angle without the need for expensive 3D integration of the classical volumetric NeRF? Can we use it to render shadows for a NeRF object in a simple CG scenario?

3 Methods

Neural Light Fields (NeLF) estimate the color of a ray with a single network forward. Implicitly, NeLFs are more complex than NeRFs - not enough training data. Model: 128 wide; 44 deep;

Knowledge Distillation using a pre-trained NeRF model, generate training data for the NeLF model. Generate the rays, save the output and the rays. In this case, the output is just the estimated depth.

Ray reprojection: in order to achieve full viewpoint freedom, rays that have their origin outside of the bounding sphere of the scene need to be reprojected. Figure 2 shows this process.

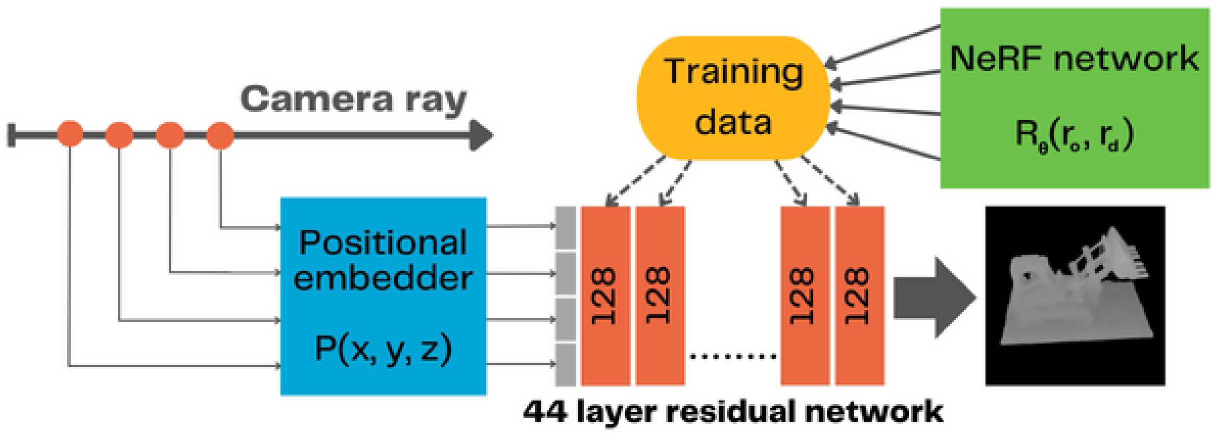

Figure 1: Network architecture

Project page

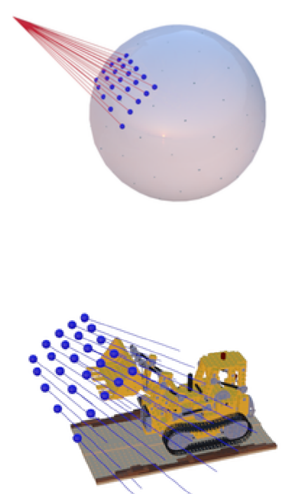
Figure 2: ray reprojection

4 Analysis


Full view | Ground truth | NeRF (~20 s) | DeLFT (~0.2 s)

	Scene	PSNR	MSE	SSIM
Synthetic	Lego	28.3468	0.0015	0.9821
Synthetic	Microphone	26.5106	0.0022	0.8832
Synthetic	Chair	24.9331	0.0032	0.8855
LLFF*	Fern	29.8775	0.0010	0.9925
LLFF*	Horns	27.0494	0.0020	0.9635

Table 1: Evaluation metrics for multiple converged scenes, trained for 100k iterations. (*) LLFF scenes do not have ground truth depth data available, so train PSNR is displayed.

5 Conclusions

- NeLF combined with knowledge distillation prove to be a viable solution for the proposed problem, generating depth maps with sufficient granularity.
- Not only are the outputs close to the NeRF teacher, but the noise is reduced and silhouette edges are better approximated.
- Further speedups could be achieved by merging training and data generation, or incorporating speedup structures.