

# Project Progress Report:

## Aug-NeLF: Augmented Neural Light Fields with Stronger Multi-View Consistency

Zhiyuan Ren, Han Meng, Maolin Gan, Lanpeng Li  
Michigan State University  
428 S Shaw Ln 3115, East Lansing, MI 48824

renzhiy1@msu.edu, menghan1@msu.edu, ganmaoli@msu.edu, lilanpen@msu.edu

### Abstract

*Compared with Neural Radiance Field(NeRF), Neural Light Field(NeLF) can render 2D images with novel views from neural representation of 3D scenes without multiple sampling along the ray, which makes real-time rendering possible. One shortcoming of the NeLF is lack of the multi-view consistency or geometric structure due to the sparse feature representation space. Recently, robust data augmentation have reached great success for improving the generalization and robustness of the model. Inspired by this, we hope robust data augmentation can bring its power to NeLF by learning from perturbed data, thus improving the multi-view consistency. For the next 45 days, we plan to establish the baseline setting within two weeks and then tune the basic methods as we get more inspiring experimental results within three weeks. Last we use around a week to finish the writing job.*

### 1. Introduction

Inverse rendering from 3D scene is a fundamental problem both in computer vision and computer graphics. Recent research in modeling implicit neural representations [3, 4, 6, 8] and rendering in the differential way [5] has advanced the progress of this problem. Especially, Neural Radiance Field(NeRF) and its variants have reached great success due to the straightforward optimization by the MLP and differential volumetric rendering. However, the inference of NeRF needs to sample many points along each ray and feed these points into the neural network one by one, which make it difficult to do real-time rendering.

To tackle this problem, people proposed Neural Light Field(NeLF) [7] which directly map a ray passing through one pixel to the radiance value. Different from NeRF(shown in Fig.1), NeLF does not need to query opacity at 3D locations along a ray and regress the color value

with the four dimensional space of light rays, which speed up rendering process a lot. As every coin has two sides, NeLF suffers from bad multi-view consistency because it does not model the density explicitly as NeRF, leading to sparse feature representation space and difficult optimization. There has been several works studying to solve this problem. [7] uses meta learning to learn a strong prior to improve the geometry consistency. [1] projects the low dimensional input into higher dimensional embedding to feed richer representation into neural network. [9] adopts distillation to help NeLF learn the multi-view consistency and geometry information from well pre-trained NeRF.

Inspired by recent work Aug-NeRF [2], which applies robust data augmentation strategy to NeRF and helps improve inconsistent and visually non-smooth geometric results, we think that NeLF lacks more multi-view consistency and geometry information and naturally we propose Aug-NeLF. Training procedure will injects different level perturbations including input coordinates, feature level representation and output color. Hopefully, our Aug-NeLF can efficiently solve the multi-view inconsistency problem in Neural Light Field area. Our codes are available in <https://github.com/Zhiyuan-VisionKing/Aug-NeLF>

### 2. Related Work

**Aug-NeRF.** Aug-NeRF addresses the inherent non-smooth geometries of NeRF. Specifically, based on solid physical grounds, Aug-NeRF seamlessly injects worst-case perturbations into three levels of the NeRF pipeline, leading to substantially improved geometry continuity and generalization ability. Moreover, the implicit smooth prior induced by triple-level augmentation enables NeRF to recover scenes from noisy supervision images. Considering that Aug-NeRF has reached great success and NeLF need more geometry consistency than NeRF, we judge that it is feasible to improve the performance of the NeLF model.

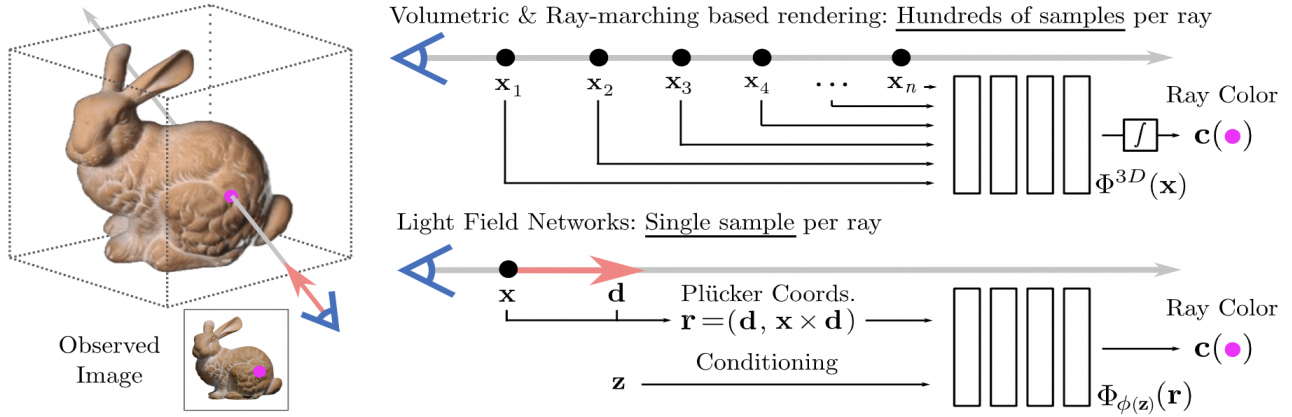


Figure 1. Comparison between Neural Radiance Field(NeRF) and Neural Light Field(NeLF).

### 3. Method

Designing dedicated perturbations for NeLF is far from trivial due to its inherent physics. We propose to inject worst-case perturbations into all three levels: coordinates, intermediate features of MLP, and pre-rendering MLP output: all with clear physical meanings. Formally, our approach

$$\min_{\Theta} \mathbb{E}_{(\mathbf{r}, \hat{\mathbf{C}}) \sim \mathbb{P}(\mathcal{R})}^{\max_{\delta}} \left\| \mathbf{C}^{\dagger}(\mathbf{r} \mid \Theta, \delta) - \hat{\mathbf{C}} \right\|_2^2,$$

where  $\delta = (\delta_p, \delta_f, \delta_r) \in \mathcal{S}_p \times \mathcal{S}_f \times \mathcal{S}_r$  where  $\delta_p, \delta_f$ , and  $\delta_r$  are the perturbations to be learned and injected to the input coordinate, intermediate MLP feature, and pre-rendering RGB- $\sigma$  output, respectively, where  $\mathcal{S}_p \subseteq \mathbb{R}^6$ ,  $\mathcal{S}_f \subseteq \mathbb{R}^D$ , and  $\mathcal{S}_r \subseteq \mathbb{R}^4$  are the corresponding perturbation search range,  $D$  is the hidden dimension of the MLP. We elaborate on each perturbation as below.

The perturbations can be accurately estimated by multi-step Projected Gradient Descent (PGD). Taking  $\ell_{\infty}$  norm ball for example:

$$\delta^{(t+1)} = \prod_{\|\delta\|_{\infty} \leq \epsilon} \left[ \delta^{(t)} + \alpha \cdot \text{sgn}(\nabla_{\delta} \mathcal{L}(\Theta \mid \mathcal{R}, \delta)) \right]$$

where  $\alpha$  is the step size of the inner maximization,  $\Pi[\cdot]$  denotes a projection operator,  $\text{sgn}(\cdot)$  takes the sign of the input, and  $\mathcal{L}(\Theta \mid \mathcal{R}, \delta)$  represents the MSE loss between perturbed color  $\mathbf{C}^{\dagger}$  and ground-truth color  $\hat{\mathbf{C}}$ .

### 4. Experiments

NeLF.

#### References

[1] Benjamin Attal, Jia-Bin Huang, Michael Zollhöfer, Johannes Kopf, and Changil Kim. Learning neural light fields with

ray-space embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19819–19829, 2022. 1

[2] Tianlong Chen, Peihao Wang, Zhiwen Fan, and Zhangyang Wang. Aug-nerf: Training stronger neural radiance fields with triple-level physically-grounded augmentations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15191–15202, 2022. 1

[3] Frank Dellaert and Lin Yen-Chen. Neural volume rendering: Nerf and beyond. *arXiv preprint arXiv:2101.05204*, 2020. 1

[4] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4460–4470, 2019. 1

[5] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021. 1

[6] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 165–174, 2019. 1

[7] Vincent Sitzmann, Semon Rezchikov, Bill Freeman, Josh Tenenbaum, and Fredo Durand. Light field networks: Neural scene representations with single-evaluation rendering. *Advances in Neural Information Processing Systems*, 34:19313–19325, 2021. 1

[8] Towaki Takikawa, Joey Litalien, Kangxue Yin, Karsten Kreis, Charles Loop, Derek Nowrouzezahrai, Alec Jacobson, Morgan McGuire, and Sanja Fidler. Neural geometric level of detail: Real-time rendering with implicit 3d shapes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11358–11367, 2021. 1

[9] Huan Wang, Jian Ren, Zeng Huang, Kyle Olszewski, Menglei Chai, Yun Fu, and Sergey Tulyakov. R2l: Distilling neural radiance field to neural light field for efficient novel view synthesis. *arXiv preprint arXiv:2203.17261*, 2022. 1