

# Supplementary Material for Ray Deformation Networks for Novel View Synthesis of Refractive Objects

**I. Visual Examples of Our Collected Datasets.** We collect the datasets using iPhone-11 and calculate the camera poses with COLMAP. We randomly show some images of each datasets in Figures A-1 and A-2.



Figure A-1. Visual example of datasets: Cup - A/B/C/D.

Table A-1. Ablation study. We report the average metrics across the six refractive object datasets. The results demonstrate that the size of the cuboid has a limited effect on our method.

Method	PSNR ( $\uparrow$ )	SSIM ( $\uparrow$ )	LIPIS ( $\downarrow$ )
Nerfacto	22.80	0.826	0.150
# Ours	26.73	0.861	0.109
A Expand	26.70	0.858	0.110
B Narrow	26.50	0.851	0.112

**II. Cuboid for Deformable Ray.** We employ a cuboid to estimate and pinpoint the refractive object’s area. If a ray interacts with the cuboid, it is reviewed as a deformable ray. To obtain the cuboid, we project rough bounding box annotations from 2D training images into 3D space using known camera poses. We randomly selected 10 training images for this purpose. Alternatively, we can use ns-viewer of nerfstudio [6] to draw a 3D cuboid easily.

Additionally, we assess the impact of the cuboid on our method. We validate this by both expanding and narrowing



Figure A-2. Visual example of datasets: Cup - E/F/G.

its scale by 5%. Results in Table A-1 demonstrate that the size of the cuboid has a limited effect on our method.

**III. Evaluation on Foreground Region.** Following [1,5], we report the metrics of the foreground region with the provided object segmentation masks on Ball and Glass. We use the evaluation code provided by [5]. As shown in Table A-2, the three models that are designed for refraction (*i.e.*, MS-NeRF, Eikonal Fields [1], and SampleNeRFRO [5]) achieve better performance than standard NeRF models. Compared with them, our method performs competitively. Moreover, we show a qualitative evaluation in Figure A-3. The refractive object regions in our renderings appear smoother and cleaner on both datasets.

**IV. Implementation Details.** As depicted in Fig. A-4, our deformation network comprises a three-hidden layer MLP with ReLU activation. The inputs (position, normal, and direction) are encoded with position encoding using three frequencies. Our method’s implementation is based on nerfacto, provided by nerfstudio [6] (Version 0.3.3), which can be found at <https://github.com/nerfstudio-project/nerfstudio>. For reference, the official code for SampleNeRFRO can be accessed at <https://github.com/nerfstudio-project/samplerfro>

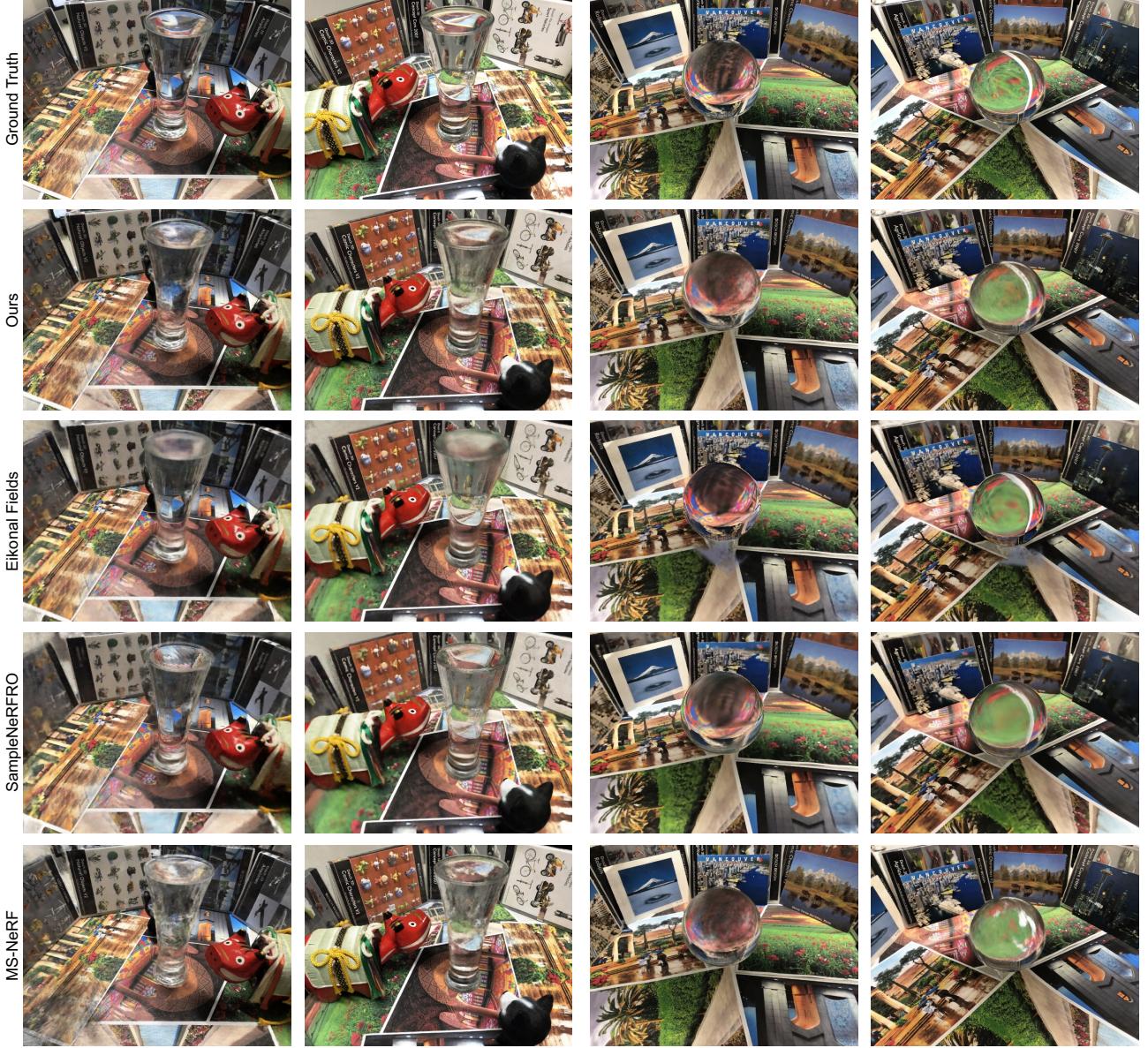
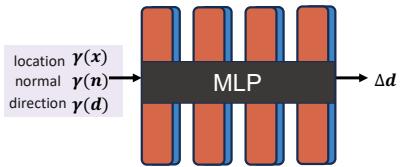


Figure A-3. Qualitative comparison of novel view synthesis with four NeRF models that are designed for refraction. Our method achieves promising novel view synthesis, resulting in smoother and cleaner results for refractive objects.



[github.com/alexkeroro86/SampleNeRFRO](https://github.com/alexkeroro86/SampleNeRFRO), and the official code for Eikonal Fields is available at <https://github.com/m-bemana/eikonalfield>.

Figure A-4. The direction deformation network architecture. It comprises a three-hidden layer MLP with 128 units per layer, utilizing ReLU activation. Inputs include position, normal, and direction, with position encoding using three frequencies. The Position Deformation Network follows an identical structure.

Table A-2. Quantitative evaluation on the test set of Ball and Glass. We evaluate the **foreground region** with the segmentation mask. We report PSNR ( $\uparrow$ ), SSIM ( $\uparrow$ ), and LPIPS ( $\downarrow$ ), across various NeRF models: TensoRF [2], Instant-NGP [4], MS-NeRF [3], Nerfacto [6], Eikonal Fields [1], SampleNeRFRO [5], and ours.

<sup>†</sup> Assumes known geometry/masks and refractive indices.

Model	Ball [1]			Glass [1]		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
TensoRF	25.25	0.906	0.071	25.48	0.926	0.0519
Instant-NGP	24.35	0.897	0.071	25.26	0.926	0.054
Nerfacto	24.90	0.901	0.065	25.48	0.926	0.052
MS-NeRF	25.16	0.910	0.057	26.19	0.936	0.043
SampleNeRFRO <sup>†</sup>	25.73	0.896	0.063	26.97	0.932	0.039
Eikonal Fields	26.13	0.906	0.064	26.23	0.935	<b>0.039</b>
Ours	<b>27.18</b>	<b>0.918</b>	<b>0.051</b>	<b>27.50</b>	<b>0.939</b>	0.040

## References

- [1] Mojtaba Bemana, Karol Myszkowski, Jeppe Revall Frisvad, Hans-Peter Seidel, and Tobias Ritschel. Eikonal fields for refractive novel-view synthesis. In *ACM SIGGRAPH 2022 Conference Proceedings*, pages 1–9, 2022. [A-1](#), [A-3](#)
- [2] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. Tensorf: Tensorial radiance fields. In *European Conference on Computer Vision (ECCV)*, 2022. [A-3](#)
- [3] Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, Peter Hedman, Ricardo Martin-Brualla, and Jonathan T. Barron. Multi-NeRF: A Code Release for Mip-NeRF 360, Ref-NeRF, and RawNeRF, 2022. [A-3](#)
- [4] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)*, 41(4):1–15, 2022. [A-3](#)
- [5] Jen-I Pan, Jheng-Wei Su, Kai-Wen Hsiao, Ting-Yu Yen, and Hung-Kuo Chu. Sampling neural radiance fields for refractive objects. In *SIGGRAPH Asia 2022 Technical Communications*, pages 1–4. 2022. [A-1](#), [A-3](#)
- [6] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Justin Kerr, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, David McAllister, and Angjoo Kanazawa. Nerfstudio: A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings, SIGGRAPH '23*, 2023. [A-1](#), [A-3](#)