Reconstructing and Rendering Shiny Surface with Hybrid Neural Fields

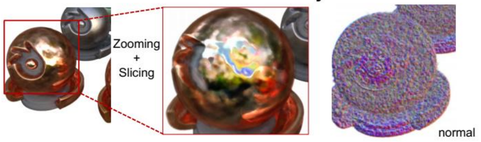
Ruofan Liang

Department of Computer Science, University of Toronto

Motivation

Problem: Reconstructing high-quality shinny 3D objects is still a problem for NeRF. *Surface* quality and *rendering* quality cannot be obtained at the same time (i.e., *appearance & geometry ambiguity* [5]).

- Some work achieves higher rendering quality but has erroneous surface.
- Some work achieves smooth surface reconstruction but lacks object details.



Existing work that tackle this problem:
Ref-NeRF [1], VolSDF [2], NeuS [3], …
These methods rely on ***fully implicit MLP*** and take ***very long time to train*** (days of). However, simply replacing pure MLP model with ***hybrid model*** (e.g., Instant-NGP [4]) for speedup will deteriorate the surface quality a lot because of discrete neural feature in hybrid model. Hence,

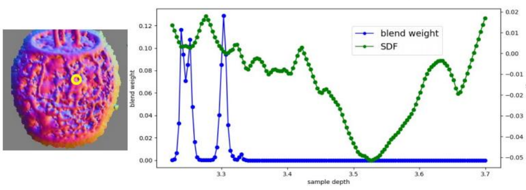
- Extra constraints need to be considered.

Goal: We want to achieve better *rendering* and *surface* reconstruction results for shinny objects, with much *faster speed* by using hybrid NeRF models.

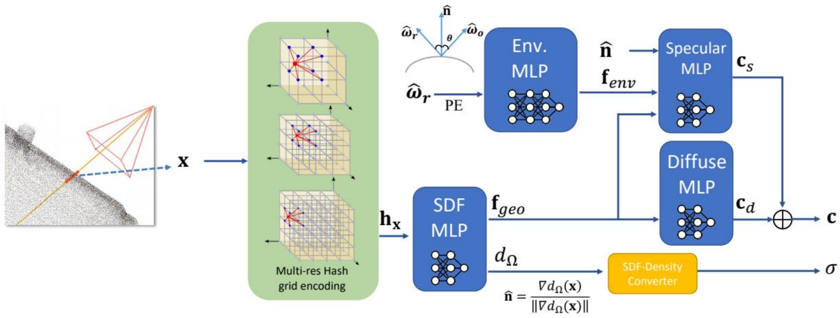
Observations

What if we simply replace the fully implicit MLP with Instant-NGP and keep the rest parts unchanged (including Eikonal regularization)?

GT VolSDF VolSDF+NGP



Overview



- Instant-NGP like hybrid model with SDF-based volume density (e.g., VolSDF).
- Decomposing color with diffuse and specular color with two shallow MLP branches.
- Using a relatively large env. MLP with reflected direction as the only input to implicitly learn global illumination from environment.

SDF Constraints

Additional constraints for hybrid NeRF models:

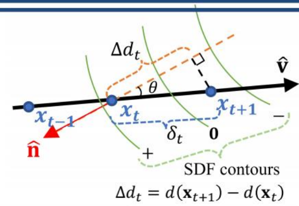
- SDF continuity regularization

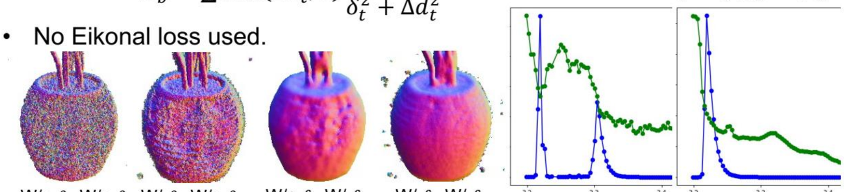
$$\mathcal{L}_d = \sum \|\Delta d_\Omega(\mathbf{x}_t) - (\hat{\mathbf{n}} \cdot \hat{\mathbf{v}})\delta_t\|^2$$

- Back face suppression

$$\mathcal{L}_b = \sum \max(\Delta d_t, 0)\frac{w_i \Delta d_t}{\delta_t^2 + \Delta d_t^2}$$

- No Eikonal loss used.

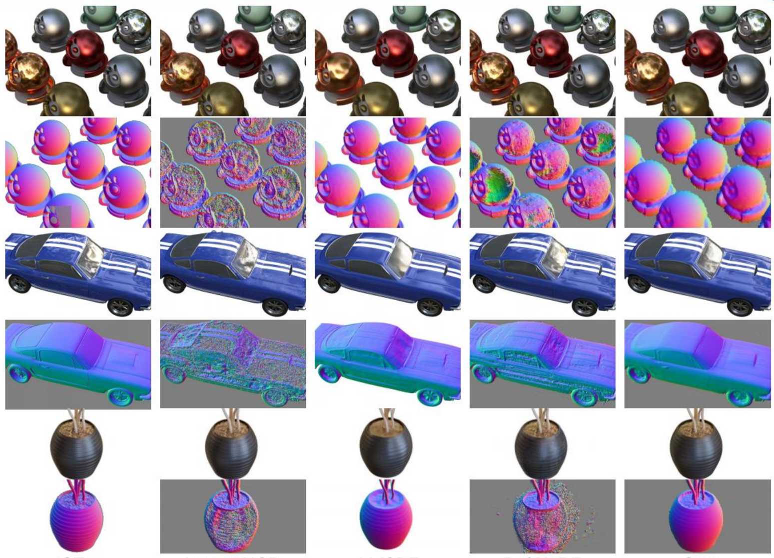




W/o $\mathcal{L}_d$ W/o $\mathcal{L}_b$ | W/ $\mathcal{L}_d$ W/o $\mathcal{L}_b$ | W/o $\mathcal{L}_d$ W/ $\mathcal{L}_b$ | W/ $\mathcal{L}_d$ W/ $\mathcal{L}_b$ | W/ $\mathcal{L}_d$ W/o $\mathcal{L}_b$ | W/ $\mathcal{L}_d$ W/ $\mathcal{L}_b$

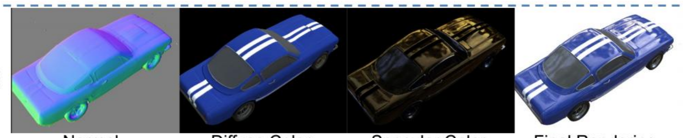
* Evaluated only after ~100 epoch, no env MLP is used.

Experimental Results



GT Instant-NGP VolSDF Ref-NeRF Ours

Our Decomposition:



Normal Diffuse Color Specular Color Final Rendering

	PSNR↑			SSIM↑			LPIPS↓			MAE°↓		
	ficus	materials	car	ficus	materials	car	ficus	materials	car	ficus	materials	car
iNGP	**31.68**	28.07	26.14	**0.988**	0.962	0.939	**0.015**	0.034	0.061	61.35	70.81	73.25
VolSDF	24.26	28.56	27.02	0.957	0.966	0.951	0.059	0.041	0.055	<u>46.58</u>	<u>32.49</u>	**46.92**
RefNeRF	<u>29.88</u>	**33.06**	**30.01**	<u>0.982</u>	**0.987**	<u>0.967</u>	0.021	**0.013**	**0.026**	68.00	42.48	57.19
Ours	29.23	<u>29.44</u>	<u>29.50</u>	<u>0.982</u>	<u>0.971</u>	**0.969**	<u>0.020</u>	<u>0.026</u>	<u>0.036</u>	**41.46**	**21.09**	<u>50.64</u>
RefNeRF*	33.91	35.41	30.82	0.983	0.983	0.955	0.019	0.022	0.041	41.05	9.53	14.93

RefNeRF* are the report scores in the original paper.

References

[1] Verbin, Dor, et al. "Ref-nerf: Structured view-dependent appearance for neural radiance fields." CVPR, 2022.
[2] Yariv, Lior, et al. "Volume rendering of neural implicit surfaces." NeurIPS, 2021.
[3] Wang, Peng, et al. "Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction." NeurIPS, 2021.
[4] Müller, Thomas, et al. "Instant neural graphics primitives with a multiresolution hash encoding." SIGGRAPH, 2022
[5] Zhang, Kai, et al. "Nerf++: Analyzing and improving neural radiance fields." preprint arXiv:2010.07492 (2020).