Learning Environment and Reflectance from Neural Radiance Fields

Ruofan Liang
Department of Computer Science
University of Toronto
ruofan@cs.toronto.edu

1 Introduction

Neural Radiance Fields (NeRF) [7] nowadays become a hot topic in the area of computer vision and graphics. By combining deep learning and traditional volume rendering methods, NeRF achieves great improvements in modeling and reconstructing 3D scenes with photo-realistic rendering quality. NeRF's strong ability to represent 3D objects is favored by a wide range of related 3D vision or graphics tasks, including novel view synthesis [1, 2], 3D content creation [6, 11, 9], visual SLAM [5, 16], relighting [3, 15, 4], etc.

Although NeRF method is able to synthesize novel view images with promising visual qualities, NeRF models often fail to accurately represent shiny surfaces with high specular reflectance. Instead of learning a solid, smooth surface for those shiny regions, NeRF tends to misinterpret the view-dependent specular reflectance as emitting light sources under the real surfaces. This erroneous behavior of NeRF also results in the poor quality of the extracted surface on the shiny regions, because the fake internal light sources have to be able to transmit through the internal volume for synthesizing view-dependent reflectance effects. Verbin et al. also discussed this issue in their work [10], please refer to their paper for a more detailed analysis.

There are two lines of work that attempt to address the problem of learning specularity in NeRF. The first direction is to make NeRF's rendering more like a surface rendering [13, 14, 12]. Instead of using volume density to represent 3D geometry, the methods in this line use surface-based implicit geometry representation to ensure higher surface reconstruction quality for their surface-based rendering The work in the second direction sticks with NeRF's volume representation and puts more effort into the modeling of view-dependent radiance prediction MLP (e.g., Ref-NeRF [10]). They thus improve the model's capability for representing complex reflectance on the pure object surface.

These abovementioned works are able to improve the quality of representation for the challenging specular reflectance, but they do have limitations. First, these methods are costly to train. Because their optimization process involves the normal (the computed gradient), the actual training time for these methods could be even longer than vanilla NeRF [7]. Second, these methods cannot be easily used in existing rendering frameworks. Since the environment lights are implicitly encoded into the neural parameters, we cannot change the surface reflectance to accommodate the new environment in the game or movie, though these models learn good surface reflectance for a specific static scene. To this end, we want to propose an efficient NeRF model that is able to capture more accurate specular reflectance and is also able to decompose an accurate environment representation from NeRF's neural models.

The ideas this work might refer to are from instant-NGP [8] and Ref-NeRF [10]. Instant-NGP combines multi-level spatial hash tables and neural implicit representation to achieve fast training and inference speed for NeRF. We will consider using instant-NGP as our backbone model to achieve training speedup [1]. Since such a hybrid NeRF model with discrete features may hurt the continuity of reconstructed surfaces, we might need to propose extra regularizations to ensure the high quality

[1]Fast training is super important for the time-limited course project.

Figure 1: NeRF's fake emitters under the object surface.

of reconstructed surfaces. Ref-NeRF is the current state-of-the-art method for learning specular reflectance. This method uses reflected/outgoing radiance directions to condition the directional MLP that outputs the specular radiance, however, this work does not explicitly learn or represent the environment light, though its large directional MLP model does implicitly learn the environment lighting information, otherwise it cannot synthesize the shiny surfaces with the best quality. It would be beneficial if we can extract the environment light from such models. We hope the explicit parameterization of the environment light could help NeRF model get even higher rendering quality and also enable image-based relighting.

References

[1] Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5855–5864, 2021.

[2] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5470–5479, 2022.

[3] Mark Boss, Raphael Braun, Varun Jampani, Jonathan T Barron, Ce Liu, and Hendrik Lensch. Nerd: Neural reflectance decomposition from image collections. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12684–12694, 2021.

[4] Mark Boss, Varun Jampani, Raphael Braun, Ce Liu, Jonathan Barron, and Hendrik Lensch. Neural-pil: Neural pre-integrated lighting for reflectance decomposition. *Advances in Neural Information Processing Systems*, 34:10691–10704, 2021.

[5] Robin Dhamankar, Yoonkyong Lee, AnHai Doan, Alon Halevy, and Pedro Domingos. imap: Discovering complex semantic matches between database schemas. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, pages 383–394, 2004.

[6] Steven Liu, Xiuming Zhang, Zhoutong Zhang, Richard Zhang, Jun-Yan Zhu, and Bryan Russell. Editing conditional radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5773–5783, 2021.

[7] 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.

[8] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *arXiv preprint arXiv:2201.05989*, 2022.

[9] Ben Poole, Ajay Jain, Jonathan T Barron, and Ben Mildenhall. Dreamfusion: Text-to-3d using 2d diffusion. *arXiv preprint arXiv:2209.14988*, 2022.

[10] Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T Barron, and Pratul P Srinivasan. Ref-nerf: structured view-dependent appearance for neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5481–5490. IEEE, 2022.

[11] Can Wang, Menglei Chai, Mingming He, Dongdong Chen, and Jing Liao. Clip-nerf: Text-and-image driven manipulation of neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3835–3844, 2022.

[12] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. *arXiv preprint arXiv:2106.10689*, 2021.

[13] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. *Advances in Neural Information Processing Systems*, 33:2492–2502, 2020.

[14] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. *Advances in Neural Information Processing Systems*, 34:4805–4815, 2021.

[15] Xiuming Zhang, Pratul P Srinivasan, Boyang Deng, Paul Debevec, William T Freeman, and Jonathan T Barron. Nerfactor: Neural factorization of shape and reflectance under an unknown illumination. *ACM Transactions on Graphics (TOG)*, 40(6):1–18, 2021.

[16] Zihan Zhu, Songyou Peng, Viktor Larsson, Weiwei Xu, Hujun Bao, Zhaopeng Cui, Martin R Oswald, and Marc Pollefeys. Nice-slam: Neural implicit scalable encoding for slam. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12786–12796, 2022.