HybridNeRF: Efficient Neural Rendering via Adaptive Volumetric Surfaces

Supplementary Material

A. Videos

We include several videos as supplementary materials under the '**vids**' folder. We recommend watching these videos via '**index.html**', which also includes descriptions and further details on the results and comparisons.

B. Color Distillation

We distill the MLP used to represent distance from our 256-wide MLP to a 16-wide network during the finetuning stage (Section 3.2). It is possible to further accelerate rendering by similarly distilling the color MLP. We found this to provide a significant boost in rendering speed (from 46 to 60 FPS) at the cost of a minor but statistically significant decrease in rendering quality (see Table 5). We observed qualitatively similar results when decreasing width from 64 to 32 channels with more notable changes in color when decreasing the width to 16 channels (see Figure 9). As our initial results suggest that MLP evaluation remains a significant rendering bottleneck, replacing our scene-wide color MLP with a collection of smaller, location-specific MLPs as suggested by KiloNeRF [26] is potential future work that could boost rendering speed at a smaller cost in quality.

C. Anti-Aliasing

We model rays as cones [1] and use a similar anti-aliasing strategy to VR-NeRF [35] by dampening high-resolution grid features based on pixel footprint. For a given sample $\boldsymbol{x}$, we derive a pixel radius $p(\boldsymbol{x})$ in the contracted space, and calculate the optimal feature level $L(\boldsymbol{x})$ based on the Nyquist–Shannon sampling theorem:

$$L(\boldsymbol{x}) := -\log_2(2s \cdot p(\boldsymbol{x})), \tag{7}$$

where s is our base grid resolution (128). We then multiply grid features at resolution level L with per-level weights w_L:

$$w_L = \begin{cases} 1 & \text{if } L < \lfloor L(\boldsymbol{x}) \rfloor \\ L(\boldsymbol{x}) - \lfloor L(\boldsymbol{x}) \rfloor & \text{if } \lfloor L(\boldsymbol{x}) \rfloor < L \leq \lceil L(\boldsymbol{x}) \rceil \\ 0 & \text{if } \lceil L(\boldsymbol{x}) \rceil < L. \end{cases} \tag{8}$$

D. ScanNet++

We evaluate 9 scenes from ScanNet++ [38] in Section 4.3 (5fb5d2dbf2, 8b5caf3398, 39f36da05b, 41b00feddb, 56a0ec536c, 98b4ec142f, b20a261fdf, f8f12e4e6b, fe1733741f). We undistort the fisheye DSLR captures to pinhole images using the official dataset toolkit [39] to facilitate comparisons against 3D Gaussian splatting [12] (whose implementation does not support fisheye projection). We use the official validation splits, which consist of entirely novel trajectories that present a more challenging novel-view synthesis problem than the commonly used pattern of holding out every eighth frame [19]. The dataset authors note that their release is still in the beta testing phase, and that the final layout is subject to change. Our testing reflects the dataset as of November 2023.

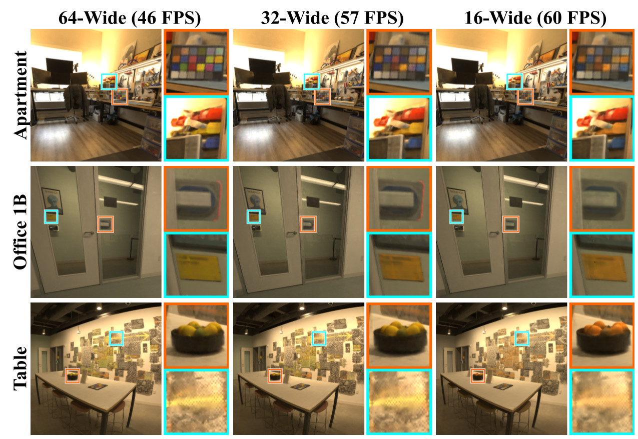

Figure 9. **Color Distillation.** Distilling the color MLP to a smaller width during the finetuning stage (Section 3.2) accelerates rendering at the cost of a minor decrease in quality. We observe largely similar results when decreasing the width to 32 channels, and more noticeable changes in color when further decreasing to 16.

Table 5. **Color distillation.** We evaluate the effect of color MLP distillation on the Eyeful Tower dataset [35], and find a significant increase in rendering speed at the cost of quality.

Color Width	↑PSNR	↑SSIM	↓LPIPS	↑FPS
16-wide (distilled)	30.88	0.888	0.236	**60.13**
32-wide (distilled)	31.17	0.900	0.220	57.05
64-wide (original)	**31.57**	**0.913**	**0.198**	45.78

E. Societal Impact

Our technique facilitates the rapid generation of high-quality neural representations. Consequently, the risks associated with our work parallel those found in other neural rendering studies, primarily centered around privacy and security issues linked to the deliberate or unintentional capture of sensitive information, such as human facial features and vehicle license plate numbers. Although we did not specifically apply our approach to data involving privacy or security concerns, there exists a risk, akin to other neural rendering methodologies, that such sensitive data could become incorporated into the trained model if the datasets utilized are not adequately filtered beforehand. It is imperative to engage in pre-processing of the input data employed for model training, especially when extending its application beyond research, to ensure the model's resilience against privacy issues and potential misuse. However, a more in-depth exploration of this matter is beyond the scope of this paper.

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 *ICCV*, 2021. 1

[2] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded anti-aliased neural radiance fields. In *CVPR*, 2022. 1, 3, 4, 5, 7, 8

[3] Ang Cao and Justin Johnson. HexPlane: A fast representation for dynamic scenes. In *CVPR*, 2023. 4

[4] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. TensoRF: Tensorial radiance fields. In *ECCV*, 2022. 1, 2

[5] Zhiqin Chen, Thomas Funkhouser, Peter Hedman, and Andrea Tagliasacchi. MobileNeRF: Exploiting the polygon rasterization pipeline for efficient neural field rendering on mobile architectures. In *CVPR*, 2023. 2, 7

[6] Sara Fridovich-Keil, Giacomo Meanti, Frederik Warburg, Benjamin Recht, and Angjoo Kanazawa. K-planes: Explicit radiance fields in space, time, and appearance. In *CVPR*, 2023. 2, 4

[7] Stephan J. Garbin, Marek Kowalski, Matthew Johnson, Jamie Shotton, and Julien Valentin. FastNeRF: High-fidelity neural rendering at 200FPS. In *ICCV*, 2021. 2

[8] Yuan-Chen Guo, Yan-Pei Cao, Chen Wang, Yu He, Ying Shan, Xiaohu Qie, and Song-Hai Zhang. VMesh: Hybrid volume-mesh representation for efficient view synthesis. arXiv:2303.16184, 2023. 3

[9] Kunal Gupta, Miloš Hašan, Zexiang Xu, Fujun Luan, Kalyan Sunkavalli, Xin Sun, Manmohan Chandraker, and Sai Bi. MCNeRF: Monte Carlo rendering and denoising for real-time NeRFs. In *SIGGRAPH Asia*, 2023. 3

[10] Peter Hedman, Julien Philip, True Price, Jan-Michael Frahm, George Drettakis, and Gabriel Brostow. Deep blending for free-viewpoint image-based rendering. *ACM Trans. Graph.*, 37(6), 2018. 7

[11] Peter Hedman, Pratul P. Srinivasan, Ben Mildenhall, Jonathan T. Barron, and Paul Debevec. Baking neural radiance fields for real-time view synthesis. *ICCV*, 2021. 2

[12] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139:1–14, 2023. 1, 3, 6, 7, 8

[13] Leonid Keselman and Martial Hebert. Approximate differentiable rendering with algebraic surfaces. In *ECCV*, 2022.

[14] Leonid Keselman and Martial Hebert. Flexible techniques for differentiable rendering with 3D Gaussians. arXiv:2308.14737, 2023. 3

[15] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015. 6

[16] Andreas Kurz, Thomas Neff, Zhaoyang Lv, Michael Zollhöfer, and Markus Steinberger. AdaNeRF: Adaptive sampling for real-time rendering of neural radiance fields. In *ECCV*, 2022. 3

[17] Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H. Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *CVPR*, pages 8456–8465, 2023. 5

[18] Shaohui Liu, Yinda Zhang, Songyou Peng, Boxin Shi, Marc Pollefeys, and Zhaopeng Cui. Dist: Rendering deep implicit signed distance function with differentiable sphere tracing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2019–2028, 2020. 1

[19] Ben Mildenhall, Pratul P. Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *ACM Transactions on Graphics (TOG)*, 2019. 1

[20] 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. In *ECCV*, 2020. 1, 3, 7

[21] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–15, 2022. 2, 4, 5, 6, 7, 8

[22] Thomas Neff, Pascal Stadlbauer, Mathias Parger, Andreas Kurz, Joerg H. Mueller, Chakravarty R. Alla Chaitanya, Anton S. Kaplanyan, and Markus Steinberger. DONeRF: Towards Real-Time Rendering of Compact Neural Radiance Fields using Depth Oracle Networks. *Computer Graphics Forum*, 2021. 3

[23] Michael Oechsle, Songyou Peng, and Andreas Geiger. UNISURF: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction. In *ICCV*, 2021. 2, 3

[24] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. In *NeurIPS*, pages 8024–8035, 2019. 6

[25] Martin Piala and Ronald Clark. TermiNeRF: Ray termination prediction for efficient neural rendering. In *3DV*, 2021. 3

[26] Christian Reiser, Songyou Peng, Yiyi Liao, and Andreas Geiger. KiloNeRF: Speeding up neural radiance fields with thousands of tiny MLPs. In *ICCV*, 2021. 1

[27] Christian Reiser, Richard Szeliski, Dor Verbin, Pratul P. Srinivasan, Ben Mildenhall, Andreas Geiger, Jonathan T. Barron, and Peter Hedman. MERF: Memory-efficient radiance fields for real-time view synthesis in unbounded scenes. *ACM Trans. Graph.*, 42(4):89:1–12, 2023. 2, 4, 5, 6, 7, 8

[28] Gernot Riegler and Vladlen Koltun. Stable view synthesis. In *CVPR*, 2021. 7

[29] Radu Alexandru Rosu and Sven Behnke. PermutoSDF: Fast multi-view reconstruction with implicit surfaces using permutohedral lattices. In *CVPR*, 2023. 5

[30] SMPTE. High dynamic range electro-optical transfer function of mastering reference displays. SMPTE Standard ST 2084:2014, Society of Motion Picture and Television Engineers, 2014. 6

[31] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin,

Kamyar Salahi, Abhik Ahuja, David Mcallister, Justin Kerr, and Angjoo Kanazawa. Nerfstudio: A modular framework for neural radiance field development. In *SIGGRAPH*, pages 72:1–12, 2023. 4

[32] Angtian Wang, Peng Wang, Jian Sun, Adam Kortylewski, and Alan Yuille. VoGE: A differentiable volume renderer using Gaussian ellipsoids for analysis-by-synthesis. In *ICLR*, 2023. 3

[33] 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. In *NeurIPS*, 2021. 1, 2, 3, 4, 5

[34] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004. 7

[35] Linning Xu, Vasu Agrawal, William Laney, Tony Garcia, Aayush Bansal, Changil Kim, Samuel Rota Bulò, Lorenzo Porzi, Peter Kontschieder, Aljaž Božič, Dahua Lin, Michael Zollhöfer, and Christian Richardt. VR-NeRF: High-fidelity virtualized walkable spaces. In *SIGGRAPH Asia*, 2023. 1, 2, 4, 6, 7, 8

[36] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. In *NeurIPS*, 2021. 1, 2, 3, 4, 5, 6, 7, 8

[37] Lior Yariv, Peter Hedman, Christian Reiser, Dor Verbin, Pratul P. Srinivasan, Richard Szeliski, Jonathan T. Barron, and Ben Mildenhall. BakedSDF: Meshing neural SDFs for real-time view synthesis. In *SIGGRAPH*, 2023. 1, 3, 4, 5, 7

[38] Chandan Yeshwanth, Yueh-Cheng Liu, Matthias Nießner, and Angela Dai. ScanNet++: A high-fidelity dataset of 3D indoor scenes. In *ICCV*, 2023. 7, 8, 1

[39] Chandan Yeshwanth, Yueh-Cheng Liu, Matthias Nießner, and Angela Dai. ScanNet++ Toolkit. https://github.com/scannetpp/scannetpp, 2023. Accessed: 2023-11-01. 1

[40] Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. PlenOctrees for real-time rendering of neural radiance fields. In *ICCV*, 2021. 2

[41] Zehao Yu, Anpei Chen, Bozidar Antic, Songyou Peng, Apratim Bhattacharyya, Michael Niemeyer, Siyu Tang, Torsten Sattler, and Andreas Geiger. Sdfstudio: A unified framework for surface reconstruction, 2022. 8

[42] Kai Zhang, Gernot Riegler, Noah Snavely, and Vladlen Koltun. NeRF++: Analyzing and improving neural radiance fields. arXiv:2010.07492, 2020. 5, 7

[43] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018. 7