A. Additional Implementation Details

When training, we used $\lambda_d = 0.1$, $\gamma = 0.7$, and $\lambda_{\text{gr}} = 10^{-3}/2$.

A.1. Fitting Hyper-Parameters

Phase A. The SDVAE/R32 NeRFs were optimized for 500/3000 iterations, using learning rates of 10^{-4}/4×10^{-4}. The learning rates were halved at 150, 300, and 450 iterations (SDVAE) and every 500 iterations for R32. Patches of size 512^2 were used, with batch sizes of 3/5.

Phase B. The joint optimization was run for 20K iterations. We used 4096 rays for the colour and DS-NeRF losses, each. The latent loss, $\mathcal{L}_p$, is computed via 32^2 latent-space patches. The learning rate (excluding the VAE) starts from 10^{-2} and is decayed according to $10^{-2} \times (10^{-1})^{t/\tau}$, where t is the step iteration and $\tau = 10^4$. The VAE is optimized with a fixed learning rate of 10^{-4}.

Phase C. Decoder fine-tuning proceeds for 3000/10000 iterations for the SDVAE/R32 architecture. A batch size of three was used (one from S_I and two from $\widetilde{S}_I$). Note that we render 512 images from the RGB-NeRF to act as supervision (i.e., $|\widetilde{S}_I| = 512$). The process starts from a learning rate of 10^{-4}, and is decayed by 0.5 every 1000/2500 iterations.

A.2. R32 Architecture

The encoder, E, has the following structure: `C5`, `RBIN`, `HD`, `RBIN`, `HD`, `RBIN`, `HD`, `RBIN`, `C1`. The components are as follows: `C5` is a conv-5×5-`norm`-elu block; `RBIN` is two residual blocks [23], each using conv-3×3 and `norm`; `HD` is a bilinear halving downscaler; and `C1` is just a conv-1×1. The encoder has layer sizes of `(32,128,128,256,256)`.

The decoder, D, has the following structure: `C1`, `RBIN`, `HU`, `RBIN`, `HU`, `RBIN`, `HU`, `RBIN`, `C1`, `sigmoid`. Components are the same, except that `HU` is a bilinear doubling upscaler. The decoder has layer sizes of `(256,256,128,128,32)`.

Both networks use the ELU non-linearity [12] and instance normalization [73] as `norm`.

References

[1] Michal Adamkiewicz, Timothy Chen, Adam Caccavale, Rachel Gardner, Preston Culbertson, Jeannette Bohg, and Mac Schwager. Vision-Only Robot Navigation in a Neural Radiance World. *IEEE Robotics and Automation Letters*, 2022. 1

[2] Tristan Aumentado-Armstrong, Stavros Tsogkas, Sven Dickinson, and Allan D Jepson. Representing 3D shapes with probabilistic directed distance fields. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[3] 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. *International Conference on Computer Vision (ICCV)*, 2021. 1

[4] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded anti-aliased neural radiance fields. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1, 2, 3

[5] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Zip-NeRF: Anti-aliased grid-based neural radiance fields. *arXiv preprint arXiv:2304.06706*, 2023. 1, 2

[6] Kenneth Blomqvist, Francesco Milano, Jen Jen Chung, Lionel Ott, and Roland Siegwart. Neural implicit vision-language feature fields. *arXiv preprint arXiv:2303.10962*, 2023. 2

[7] Caroline Chan, Frédo Durand, and Phillip Isola. Learning to generate line drawings that convey geometry and semantics. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 3

[8] Eric R Chan, Koki Nagano, Matthew A Chan, Alexander W Bergman, Jeong Joon Park, Axel Levy, Miika Aittala, Shalini De Mello, Tero Karras, and Gordon Wetzstein. Generative novel view synthesis with 3D-aware diffusion models. *arXiv preprint arXiv:2304.02602*, 2023. 2

[9] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. TensoRF: Tensorial radiance fields. In *European Conference on Computer Vision (ECCV)*, 2022. 2

[10] Anpei Chen, Zexiang Xu, Xinyue Wei, Siyu Tang, Hao Su, and Andreas Geiger. Factor fields: A unified framework for neural fields and beyond. *arXiv preprint arXiv:2302.01226*, 2023. 2

[11] Zhiqin Chen, Thomas Funkhouser, Peter Hedman, and Andrea Tagliasacchi. MobileNeRF: Exploiting the polygon rasterization pipeline for efficient neural field rendering on mobile architectures. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 1, 2

[12] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (ELUs). *arXiv preprint arXiv:1511.07289*, 2015. 9

[13] Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised NeRF: Fewer views and faster training for free. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022. 2

[14] Danny Driess, Ingmar Schubert, Pete Florence, Yunzhu Li, and Marc Toussaint. Reinforcement learning with neural radiance fields. *Neural Information Processing Systems (NeurIPS)*, 2022. 1

[15] S. M. Ali Eslami, Danilo Jimenez Rezende, Frederic Besse, Fabio Viola, Ari S. Morcos, Marta Garnelo, Avraham Ruderman, Andrei A. Rusu, Ivo Danihelka, Karol Gregor, David P. Reichert, Lars Buesing, Theophane Weber, Oriol Vinyals, Dan Rosenbaum, Neil Rabinowitz, Helen King, Chloe Hillier, Matt Botvinick, Daan Wierstra, Koray Kavukcuoglu, and Demis Hassabis. Neural scene representation and rendering. *Science*, 360(6394):1204–1210, 2018. 2

[16] Jiemin Fang, Lingxi Xie, Xinggang Wang, Xiaopeng Zhang, Wenyu Liu, and Qi Tian. Neusample: Neural sample field for efficient view synthesis. *arXiv preprint arXiv:2111.15552*, 2021. 2

[17] Brandon Y Feng, Yinda Zhang, Danhang Tang, Ruofei Du, and Amitabh Varshney. PRIF: Primary ray-based implicit function. In *European Conference on Computer Vision (ECCV)*, 2022. 2

[18] Stephanie Fu*, Netanel Tamir*, Shobhita Sundaram*, Lucy Chai, Richard Zhang, Tali Dekel, and Phillip Isola. DreamSim: Learning new dimensions of human visual similarity using synthetic data. *arXiv preprint arXiv:2306.09344*, 2023. 3

[19] Stephan J Garbin, Marek Kowalski, Matthew Johnson, Jamie Shotton, and Julien Valentin. FastNeRF: High-fidelity neural rendering at 200fps. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 2

[20] Jiatao Gu, Lingjie Liu, Peng Wang, and Christian Theobalt. StyleNeRF: A style-based 3D-aware generator for high-resolution image synthesis. *arXiv preprint arXiv:2110.08985*, 2021. 2, 3

[21] 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 preprint arXiv:2303.16184*, 2023. 2

[22] Ayaan Haque, Matthew Tancik, Alexei Efros, Aleksander Holynski, and Angjoo Kanazawa. Instruct-NeRF2NeRF: Editing 3D scenes with instructions. *arXiv preprint arXiv:2303.12789*, 2023. 1

[23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. 2, 9

[24] Yingqing He, Tianyu Yang, Yong Zhang, Ying Shan, and Qifeng Chen. Latent video diffusion models for high-fidelity video generation with arbitrary lengths. *arXiv preprint arXiv:2211.13221*, 2022. 3

[25] Peter Hedman, Pratul P Srinivasan, Ben Mildenhall, Jonathan T Barron, and Paul Debevec. Baking neural radiance fields for real-time view synthesis. In *International Conference on Computer Vision (ICCV)*, 2021. 1, 2

[26] Jonathan Ho, William Chan, Chitwan Saharia, Jay Whang, Ruiqi Gao, Alexey Gritsenko, Diederik P Kingma, Ben

Poole, Mohammad Norouzi, David J Fleet, and Tim Salimans. Imagen video: High definition video generation with diffusion models. *arXiv preprint arXiv:2210.02303*, 2022. 3

[27] Trevor Houchens, Cheng-You Lu, Shivam Duggal, Rao Fu, and Srinath Sridhar. NeuralODF: Learning omnidirectional distance fields for 3D shape representation. *arXiv preprint arXiv:2206.05837*, 2022. 2

[28] Huaibo Huang, Zhihang Li, Ran He, Zhenan Sun, and Tieniu Tan. IntroVAE: Introspective variational autoencoders for photographic image synthesis. *Neural Information Processing Systems (NeurIPS)*, 2018. 2

[29] Muhammad Zubair Irshad, Sergey Zakharov, Rares Ambrus, Thomas Kollar, Zsolt Kira, and Adrien Gaidon. ShAPO: Implicit representations for multi object shape appearance and pose optimization. *European Conference on Computer Vision (ECCV)*, 2022. 1

[30] Mohammad Mahdi Johari, Yann Lepoittevin, and François Fleuret. GeoNeRF: Generalizing nerf with geometry priors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[31] Justin Kerr, Chung Min Kim, Ken Goldberg, Angjoo Kanazawa, and Matthew Tancik. LeRF: Language embedded radiance fields. *arXiv preprint arXiv:2303.09553*, 2023. 2

[32] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, 2015. 3

[33] Sosuke Kobayashi, Eiichi Matsumoto, and Vincent Sitzmann. Decomposing NeRF for editing via feature field distillation. *Neural Information Processing Systems (NeurIPS)*, 2022. 2

[34] Naruya Kondo, Yuya Ikeda, Andrea Tagliasacchi, Yutaka Matsuo, Yoichi Ochiai, and Shixiang Shane Gu. VaxNeRF: Revisiting the classic for voxel-accelerated neural radiance field. *arXiv preprint arXiv:2111.13112*, 2021. 2

[35] Mikhail Kurenkov, Andrei Potapov, Alena Savinykh, Evgeny Yudin, Evgeny Kruzhkov, Pavel Karpyshev, and Dzmitry Tsetserukou. NFOMP: Neural field for optimal motion planner of differential drive robots with nonholonomic constraints. *IEEE Robotics and Automation Letters*, 2022. 1

[36] Andreas Kurz, Thomas Neff, Zhaoyang Lv, Michael Zollhöfer, and Markus Steinberger. AdaNeRF: Adaptive sampling for real-time rendering of neural radiance fields. In *European Conference on Computer Vision (ECCV)*, 2022. 2

[37] Soomin Lee, Chen Le, Wang Jiahao, Alexander Liniger, Suryansh Kumar, and Fisher Yu. Uncertainty guided policy for active robotic 3D reconstruction using neural radiance fields. *IEEE Robotics and Automation Letters*, 2022. 1

[38] Haotong Lin, Sida Peng, Zhen Xu, Yunzhi Yan, Qing Shuai, Hujun Bao, and Xiaowei Zhou. Efficient neural radiance fields for interactive free-viewpoint video. In *SIGGRAPH Asia 2022 Conference Papers*, 2022. 2

[39] David B Lindell, Julien NP Martel, and Gordon Wetzstein. AutoInt: Automatic integration for fast neural volume rendering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 2

[40] Gal Metzer, Elad Richardson, Or Patashnik, Raja Giryes, and Daniel Cohen-Or. Latent-NeRF for shape-guided generation of 3D shapes and textures. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 2

[41] 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. 3, 4

[42] 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 *European Conference on Computer Vision (ECCV)*, 2020. 1

[43] Ashkan Mirzaei, Tristan Aumentado-Armstrong, Marcus A. Brubaker, Jonathan Kelly, Alex Levinshtein, Konstantinos G. Derpanis, and Igor Gilitschenski. Reference-guided controllable inpainting of neural radiance fields. In *International Conference on Computer Vision (ICCV)*, 2023. 1

[44] Ashkan Mirzaei, Tristan Aumentado-Armstrong, Konstantinos G. Derpanis, Jonathan Kelly, Marcus A. Brubaker, Igor Gilitschenski, and Alex Levinshtein. SPIn-NeRF: Multiview segmentation and perceptual inpainting with neural radiance fields. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 1, 2

[45] Ashkan Mirzaei, Yash Kant, Jonathan Kelly, and Igor Gilitschenski. LaTeRF: Label and text driven object radiance fields. In *European Conference on Computer Vision (ECCV)*, 2022. 2

[46] Thomas Müller. Tiny CUDA neural network framework, 2021. https://github.com/nvlabs/tiny-cuda-nn. 3

[47] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (TOG)*, 2022. 1, 2, 3

[48] Thomas Neff, Pascal Stadlbauer, Mathias Parger, Andreas Kurz, Joerg H Mueller, Chakravarty R Alla Chaitanya, Anton Kaplanyan, and Markus Steinberger. DONeRF: Towards real-time rendering of compact neural radiance fields using depth oracle networks. In *Computer Graphics Forum*, 2021. 2

[49] Ruiqi Ni and Ahmed H. Qureshi. NTFields: Neural time fields for physics-informed robot motion planning. *International Conference on Learning Representations (ICLR)*, 2023. 1

[50] Michael Niemeyer and Andreas Geiger. GIRAFFE: Representing scenes as compositional generative neural feature fields. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 2, 3

[51] Martin Piala and Ronald Clark. TermiNeRF: Ray termination prediction for efficient neural rendering. In *International Conference on 3D Vision (3DV)*, 2021. 2

[52] Ri-Zhao Qiu, Yixiao Sun, Joao Marcos Correia Marques, and Kris Hauser. Real-time semantic 3D reconstruction for high-touch surface recognition for robotic disinfection. In *Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, 2022. 1

[53] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning (ICML)*, 2021. 3

[54] Christian Reiser, Songyou Peng, Yiyi Liao, and Andreas Geiger. KiloNeRF: Speeding up neural radiance fields with thousands of tiny MLPs. In *International Conference on Computer Vision (ICCV)*, 2021. 2

[55] 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. In *arXiv preprint arXiv:2302.12249*, 2023. 1, 2

[56] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[57] Sara Fridovich-Keil and Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1, 2

[58] Hoigi Seo, Hayeon Kim, Gwanghyun Kim, and Se Young Chun. DITTO-NeRF: Diffusion-based iterative text to omni-directional 3D model. *arXiv preprint arXiv:2304.02827*, 2023. 2

[59] Nur Muhammad Mahi Shafiullah, Chris Paxton, Lerrel Pinto, Soumith Chintala, and Arthur Szlam. CLIP-fields: Weakly supervised semantic fields for robotic memory. *arXiv preprint arXiv:2210.05663*, 2022. 2

[60] Dongseok Shim, Seungjae Lee, and H Jin Kim. SNeRL: Semantic-aware neural radiance fields for reinforcement learning. *arXiv preprint arXiv:2301.11520*, 2023. 1

[61] Vincent Sitzmann, Semon Rezchikov, Bill Freeman, Josh Tenenbaum, and Fredo Durand. Light field networks: Neural scene representations with single-evaluation rendering. *Neural Information Processing Systems (NeurIPS)*, 2021. 2

[62] Cameron Omid Smith, Hong-Xing Yu, Sergey Zakharov, Fredo Durand, Joshua B Tenenbaum, Jiajun Wu, and Vincent Sitzmann. Unsupervised discovery and composition of object light fields. *Transactions on Machine Learning Research*, 2023. 2

[63] Cheng Sun, Min Sun, and Hwann-Tzong Chen. Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[64] Cheng Sun, Min Sun, and Hwann-Tzong Chen. Improved direct voxel grid optimization for radiance fields reconstruction. *arXiv preprint arXiv:2206.05085*, 2022. 2

[65] Andrea Tagliasacchi and Ben Mildenhall. Volume rendering digest (for NeRF). *arXiv preprint arXiv:2209.02417*, 2022. 2

[66] Matthew Tancik, Vincent Casser, Xinchen Yan, Sabeek Pradhan, Ben Mildenhall, Pratul P Srinivasan, Jonathan T Barron, and Henrik Kretzschmar. Block-NeRF: Scalable large scene neural view synthesis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1, 2

[67] Matthew Tancik, Ben Mildenhall, Terrance Wang, Divi Schmidt, Pratul P Srinivasan, Jonathan T Barron, and Ren Ng. Learned initializations for optimizing coordinate-based neural representations. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 2

[68] 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 *Proceedings of SIGGRAPH*, 2023. 1

[69] Ayush Tewari, Ohad Fried, Justus Thies, Vincent Sitzmann, Stephen Lombardi, Kalyan Sunkavalli, Ricardo Martin-Brualla, Tomas Simon, Jason M. Saragih, Matthias Nießner, Rohit Pandey, Sean Ryan Fanello, Gordon Wetzstein, Jun-Yan Zhu, Christian Theobalt, Maneesh Agrawala, Eli Shechtman, Dan B. Goldman, and Michael Zollhofer. State of the art on neural rendering. *Computer Graphics Forum*, 39(2):701–727, 2020. 2

[70] Ayush Tewari, Justus Thies, Ben Mildenhall, Pratul Srinivasan, Edgar Tretschk, Yifan Wang, Christoph Lassner, Vincent Sitzmann, Ricardo Martin-Brualla, Stephen Lombardi, Tomas Simon, Christian Theobalt, Matthias Niessner, Jonathan T. Barron, Gordon Wetzstein, Michael Zollhoefer, and Vladislav Golyanik. Advances in neural rendering. In *Proceedings of SIGGRAPH*, 2021. 1

[71] Vadim Tschernezki, Iro Laina, Diane Larlus, and Andrea Vedaldi. Neural feature fusion fields: 3D distillation of self-supervised 2D image representations. *arXiv preprint arXiv:2209.03494*, 2022. 2

[72] Haithem Turki, Deva Ramanan, and Mahadev Satyanarayanan. Mega-NeRF: Scalable construction of large-scale nerfs for virtual fly-throughs. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[73] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016. 9

[74] 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. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1

[75] Yael Vinker, Yuval Alaluf, Daniel Cohen-Or, and Ariel Shamir. CLIPascene: Scene sketching with different types and levels of abstraction. *arXiv preprint arXiv:2211.17256*, 2022. 3

[76] Krishna Wadhwani and Tamaki Kojima. SqueezeNeRF: Further factorized FastNeRF for memory-efficient inference. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[77] Ziyu Wan, Christian Richardt, Aljaž Božič, Chao Li, Vijay Rengarajan, Seonghyeon Nam, Xiaoyu Xiang, Tuotuo Li, Bo Zhu, Rakesh Ranjan, and Jing Liao. Learning neural duplex radiance fields for real-time view synthesis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 2

[78] Peng Wang, Yuan Liu, Zhaoxi Chen, Lingjie Liu, Ziwei Liu, Taku Komura, Christian Theobalt, and Wenping Wang. F^2-NeRF: Fast neural radiance field training with free camera trajectories. *arXiv preprint arXiv:2303.15951*, 2023. 2

[79] Qianqian Wang, Zhicheng Wang, Kyle Genova, Pratul P Srinivasan, Howard Zhou, Jonathan T Barron, Ricardo Martin-Brualla, Noah Snavely, and Thomas Funkhouser. IBRNet: Learning multi-view image-based rendering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 2

[80] Frederik Warburg, Ethan Weber, Matthew Tancik, Aleksander Holynski, and Angjoo Kanazawa. Nerfbusters: Removing ghostly artifacts from casually captured NeRFs. *arXiv preprint arXiv:2304.10532*, 2023. 1

[81] Bowen Wen, Jonathan Tremblay, Valts Blukis, Stephen Tyree, Thomas Muller, Alex Evans, Dieter Fox, Jan Kautz, and Stan Birchfield. BundleSDF: Neural 6-DoF tracking and 3D reconstruction of unknown objects. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 1

[82] Haoning Wu, Erli Zhang, Liang Liao, Chaofeng Chen, Jingwen Hou Hou, Annan Wang, Wenxiu Sun Sun, Qiong Yan, and Weisi Lin. Exploring video quality assessment on user generated contents from aesthetic and technical perspectives. In *International Conference on Computer Vision (ICCV)*, 2023. 3

[83] Liwen Wu, Jae Yong Lee, Anand Bhattad, Yu-Xiong Wang, and David Forsyth. Diver: Real-time and accurate neural radiance fields with deterministic integration for volume rendering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[84] Yang Xue, Yuheng Li, Krishna Kumar Singh, and Yong Jae Lee. GIRAFFE-HD: A high-resolution 3D-aware generative model. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[85] Tarun Yenamandra, Ayush Tewari, Nan Yang, Florian Bernard, Christian Theobalt, and Daniel Cremers. FIRe: Fast inverse rendering using directional and signed distance functions. *arXiv preprint arXiv:2203.16284*, 2022. 2

[86] Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. Plenoctrees for real-time rendering of neural radiance fields. In *International Conference on Computer Vision (ICCV)*, 2021. 2

[87] Alex Yu, Vickie Ye, Matthew Tancik, and Angjoo Kanazawa. PixelNeRF: Neural radiance fields from one or few images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 2

[88] Kai Zhang, Nick Kolkin, Sai Bi, Fujun Luan, Zexiang Xu, Eli Shechtman, and Noah Snavely. ARF: Artistic radiance fields. In *European Conference on Computer Vision (ECCV)*, 2022. 1

[89] Kai Zhang, Gernot Riegler, Noah Snavely, and Vladlen Koltun. NeRF++: Analyzing and improving neural radiance fields. *arXiv preprint arXiv:2010.07492*, 2020. 1

[90] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 3

[91] Shuaifeng Zhi, Tristan Laidlow, Stefan Leutenegger, and Andrew J Davison. In-place scene labelling and understanding with implicit scene representation. In *International Conference on Computer Vision (ICCV)*, 2021. 2

[92] Shuaifeng Zhi, Edgar Sucar, Andre Mouton, Iain Haughton, Tristan Laidlow, and Andrew J Davison. iLabel: Interactive neural scene labelling. *arXiv preprint arXiv:2111.14637*, 2021. 2

[93] Shaohong Zhong, Alessandro Albini, Oiwi Parker Jones, Perla Maiolino, and Ingmar Posner. Touching a NeRF: Leveraging neural radiance fields for tactile sensory data generation. In *Conference on Robot Learning*, 2023. 1