SOL-NeRF: Sunlight Modeling for Outdoor Scene Decomposition and Relighting

Supplementary Material

JIA-MU SUN, Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences, China
TONG WU, Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences, China
YONG-LIANG YANG, Department of Computer Science, Bath University, United Kingdom
YU-KUN LAI, School of Computer Science and Informatics, Cardi University, United Kingdom
LIN GAO*, Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences , China

The supplementary material contains details of our network architecture, more implementation details, and additional results.

1 NETWORK ARCHITECTURE

We show the architecture of our scene decomposition network in Fig. 1.

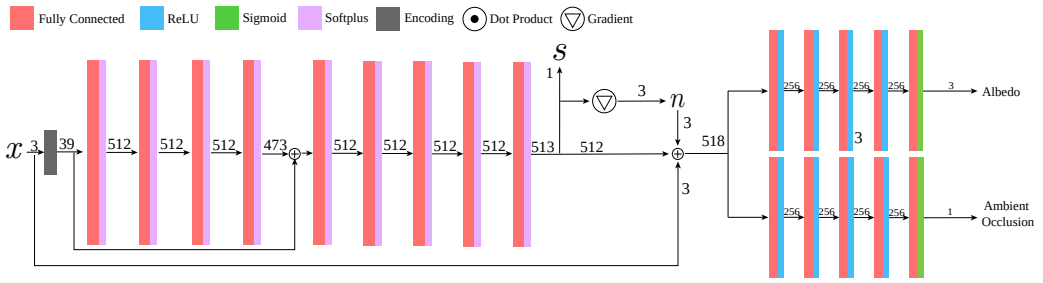


Fig. 1. Architecture of our scene decomposition network.

*Corresponding author is Lin Gao (gaolin@ict.ac.cn).

Authors addresses: Jia-Mu Sun, Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences, China, sunjiamu21s@ict.ac.cn; Tong Wu, Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences, China, wutong19s@ict.ac.cn; Yong-Liang Yang, Department of Computer Science, Bath University, United Kingdom, y.yang@cs.bath.ac.uk; Yu-Kun Lai, School of Computer Science and Informatics, Cardi University, United Kingdom, LaiY4@cardi.ac.uk; Lin Gao, Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences , China, gaolin@ict.ac.cn.

2 MORE IMPLEMENTATION DETAILS

2.1 Sunlight Color Prior

As mentioned in Sec. 3.4, we approximate the relationship between the sun's elevation angle and the sun's color with a function f_{sun}. Here we give more details on this function.

$$r = \begin{cases} 3.0405 \times 10^8\theta^{10} - 7.2707 \times 10^8\theta^9 + 7.5520 \times 10^8\theta^8 - 4.4699 \times 10^8\theta^7 & \\ +1.6620 \times 10^8\theta^6 - 4.0319 \times 10^7\theta^5 + 6.3914 \times 10^6\theta^4 - 6.3317 \times 10^5\theta^3 & \\ +3.2104 \times 10^4\theta^2 + 4.3170 \times 10^2\theta + 2.3430, & 0 \le \theta < 0.43633 \\ & \\ 20.5526\theta^5 - 123.5332\theta^4 + 300.1427\theta^3 & \\ -379.3047\theta^2 + 260.7181\theta + 108.1877, & 0.43633 \le \theta \le \pi/2 \end{cases} \tag{1}$$

$$g = \begin{cases} -8.7786 \times 10^7\theta^{10} + 1.9038 \times 10^8\theta^9 - 1.7260 \times 10^8\theta^8 + 8.3335 \times 10^7\theta^7 & \\ -2.1914 \times 10^7\theta^6 + 2.3827 \times 10^6\theta^5 + 2.6284 \times 10^5\theta^4 + 1.1738 \times 10^5\theta^3 & \\ +1.3415 \times 10^4\theta^2 + 1.5841 \times 10^1\theta + 0.5643, & 0 \le \theta < 0.43633 \\ & \\ 24.3718\theta^5 - 147.5083\theta^4 + 361.6367\theta^3 & \\ -463.3131\theta^2 + 352.0770\theta + 54.2336, & 0.43633 \le \theta \le \pi/2 \end{cases} \tag{2}$$

$$b = \begin{cases} -5.2340 \times 10^7\theta^{10} + 1.3193 \times 10^8\theta^9 - 1.4495 \times 10^8\theta^8 + 9.0902 \times 10^7\theta^7 & \\ -3.5713 \times 10^7\theta^6 + 9.0300 \times 10^6\theta^5 - 1.4312 \times 10^6\theta^4 + 1.2493 \times 10^5\theta^3 & \\ -3.1073 \times 10^3\theta^2 + 7.7017\theta - 0.1208, & 0 \le \theta < 0.43633 \\ & \\ 25.3379\theta^5 - 157.3345\theta^4 + 398.5353\theta^3 & \\ -536.5092\theta^2 + 404.7429\theta - 12.2936, & 0.43633 \le \theta \le \pi/2 \end{cases} \tag{3}$$

This function is the direct result of applying the "polyfit" function of numpy [Harris et al. 2020] on the values calculated by the nishita model[Nishita et al. 1993].

2.2 Training Details

When training the decomposition network, we set the loss multipliers in Eqn. 6 of the main paper as follows: $\lambda_1 = 0.005, \lambda_2 = 0.1, \lambda_3 = 0.1, \lambda_4 = 0.05, \lambda_5 = 0.005$. When calculating the Sampled Ambient Occlusion, we sample inside a hemisphere of radius 0.5-1.5 meters (depending on the size and scale of the details of the scene). We optimize the network for 1×10^8 steps using the ADAM optimizer with a learning rate of 5×10^{-4}. The whole training process takes 24-48 hours (depending on the size of the scene) on a single 3090 GPU.

2.3 Synthetic Dataset Details

We collected three scenes [1,2,3] from Blenderswap.com and rendered each scene under 20 different illuminations, each of which is a "Sky Texture" of "Nishita" type provided by Blender. The parameters of each "Sky Texture" are randomly sampled. For each illumination, we render each scene from 20 different viewpoints, the viewpoints are randomly sampled as well.

3 COMPARISON WITH FEGR [Wang et al. 2023]

Very recently, Wang et al. [2023] proposed FEGR, which is dedicated to decomposing and relighting outdoor scenes. However, the source code of FEGR has not been made available. In order to compare with FEGR, we reproduced a baseline similar to FEGR, which uses: 1) A deferred rendering technique, which renders albedo and normal from the radiance field, and performs rendering on the rendered maps. 2) Sampling-based shading and shadow calculation. Rays are sampled randomly and those samples are used for calculating the shadow and rendering equation. 3) An MLP for predicting the environment map. We show the reconstruction result of the baseline, and compare it to our pipeline qualitatively as shown in Fig. 2. We observed some noisy shading results and slightly bumpy surfaces in this baseline. The light appears overly bright, causing the shadow to be inaccurately decomposed and the albedo to look gray and lack saturation.

4 ADDITIONAL RESULTS

4.1 Geometry Error of the w/o Manhattan Prior (M.P.) Ablation

In the main paper, we show the reconstruction result and the normal Mean Average Error of the w/o Manhattan prior ablation of our pipeline. Here, we additionally show the geometry error map of the full pipeline and the w/o M.P. ablation with the ground truth mesh in Fig. 3.

The Chamfer distances from the ground truth (GT) geometry to the extracted geometry of the full model and the w/o M.P. ablation are **0.0462** and 0.0507 respectively.

4.2 Scene Decomposition

We show more scene decomposition results in Fig. 4 and compare them with NeRF-OSR [Rudnev et al. 2022].

4.3 Scene Relighting

We show more scene relighting results in Fig. 5 and compare them with NeRF-OSR [Rudnev et al. 2022]. We also show the applied lighting for OSR scenes and the ground truth rendered result for synthetic scenes.

Additionally, we modify Fig. 9 of the main paper by including the applied lighting (environment maps) for OSR scenes and the ground truth rendered results for synthetic scenes in Fig. 6.

4.4 Scene Editing

Our method also supports editing the appearance of the input scene by adjusting the network's output. We show such results in Fig. 7. We edit the scene by directly editing the albedo. We also show two relighting results of the edited scene in Fig. 7 to show the consistent relighting effects after editing.

[1]https://blendswap.com/blend/29892
[2]https://blendswap.com/blend/4616
[3]https://blendswap.com/blend/21136

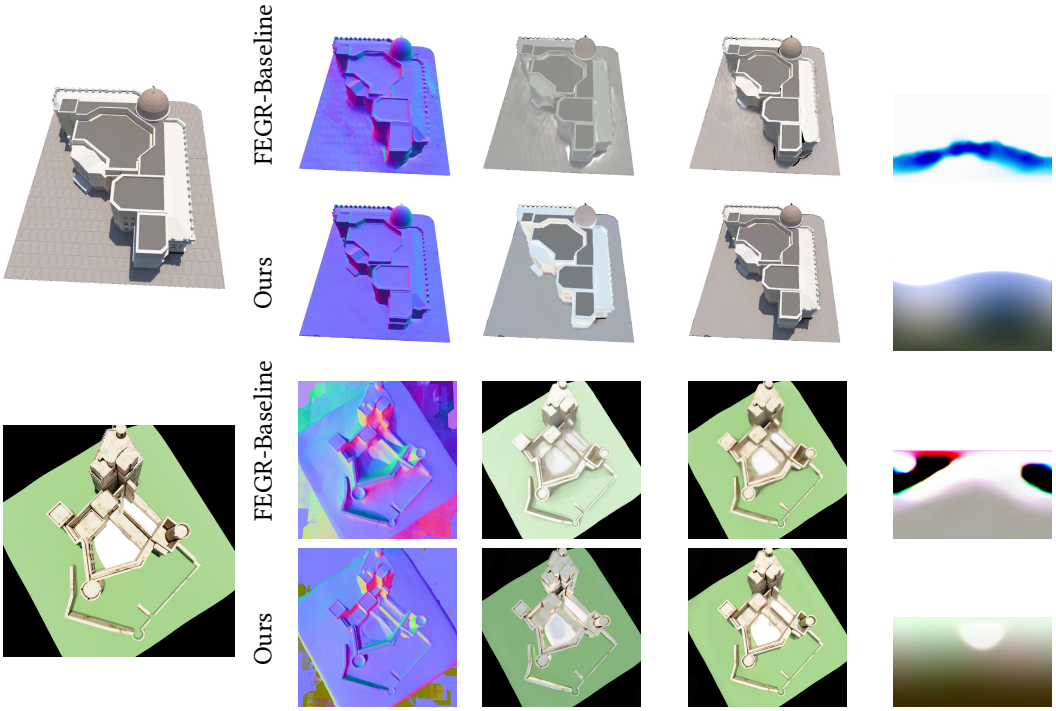


Fig. 2. **Qualitative comparison to an FEGR [Wang et al. 2023]-like baseline.** We show the reconstruction result, albedo and normal of the baseline and our method.

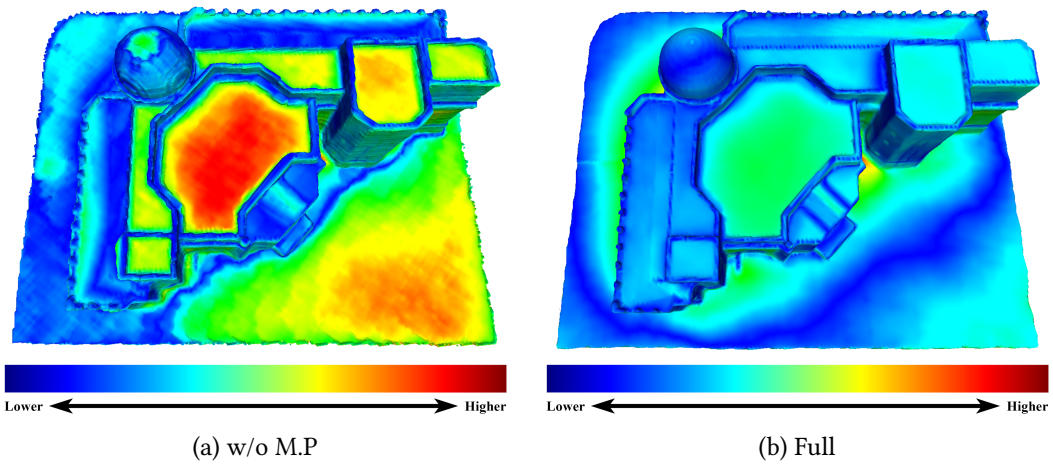


Fig. 3. **Error maps of w/o M.P. ablation and full pipeline.** We calculate the geometry error maps of w/o M.P. ablation and full pipeline and visualize them on the geometry.

REFERENCES

Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk,

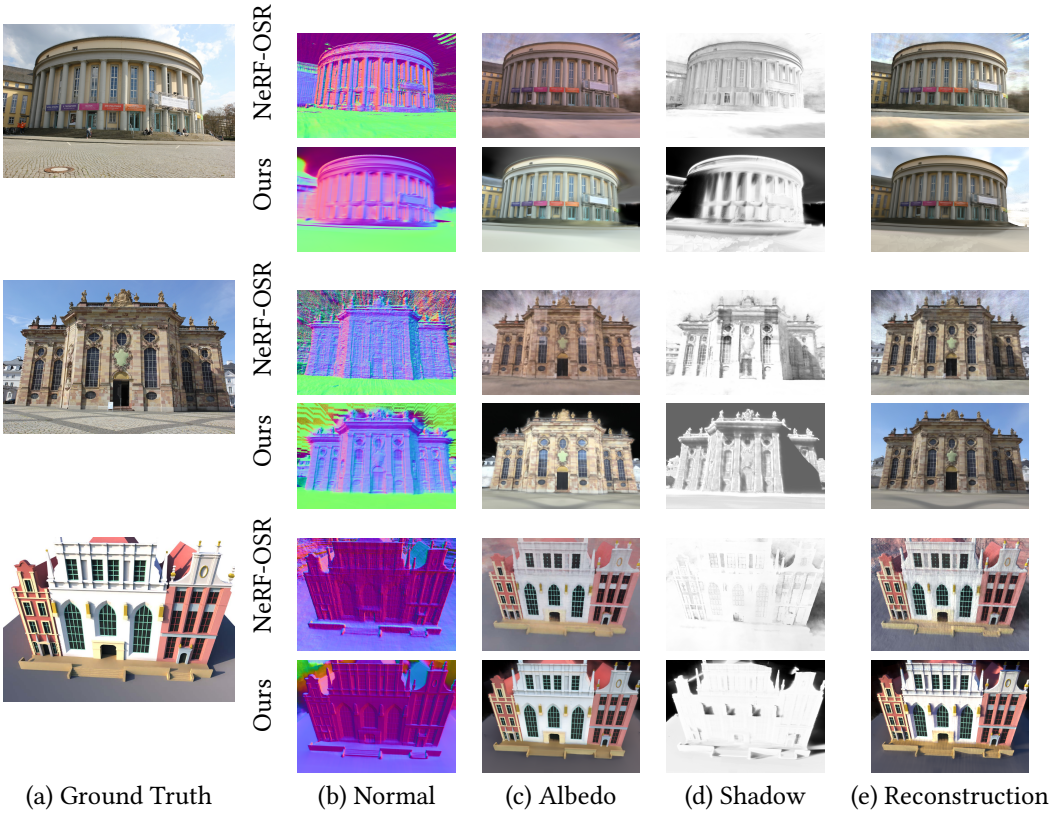


Fig. 4. **Decomposition results of NeRF-OSR and our method.** For each scene, we show different decomposed components (normal, albedo, and shadow) and the reconstructed image.

Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. Array programming with NumPy. *Nature* 585, 7825 (Sept. 2020), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Tomoyuki Nishita, Takao Sirai, Katsumi Tadamura, and Eihachiro Nakamae. 1993. Display of the Earth Taking into Account Atmospheric Scattering *(SIGGRAPH '93)*. 175–182.

Viktor Rudnev, Mohamed Elgharib, William Smith, Lingjie Liu, Vladislav Golyanik, and Christian Theobalt. 2022. NeRF for Outdoor Scene Relighting. In *ECCV*.

Zian Wang, Tianchang Shen, Jun Gao, Shengyu Huang, Jacob Munkberg, Jon Hasselgren, Zan Gojcic, Wenzheng Chen, and Sanja Fidler. 2023. Neural Fields meet Explicit Geometric Representations for Inverse Rendering of Urban Scenes. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

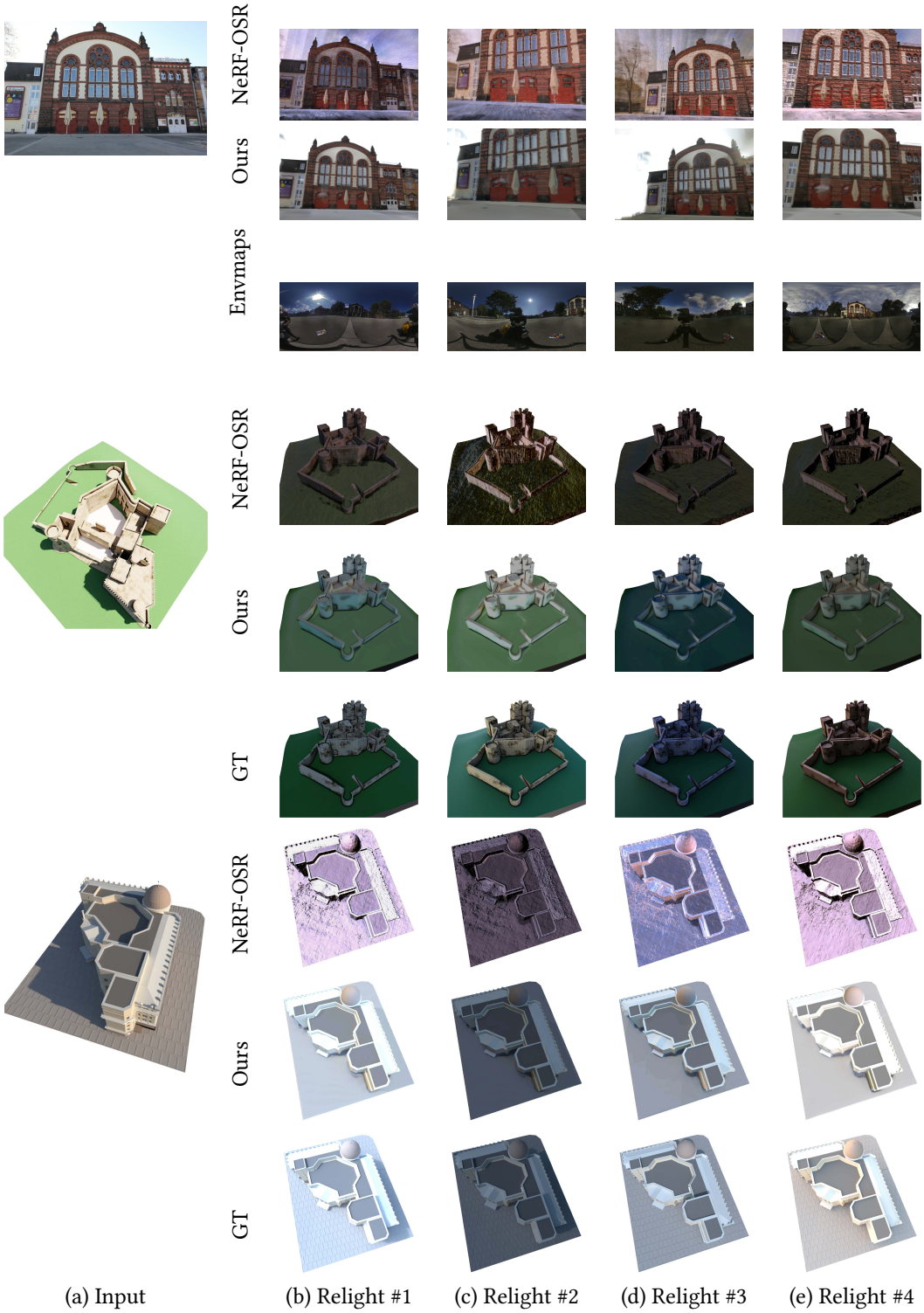


Fig. 5. **Relighting results of NeRF-OSR and our method.** For each input scene, we relight it with four different lighting conditions and show rendered images after relighting.

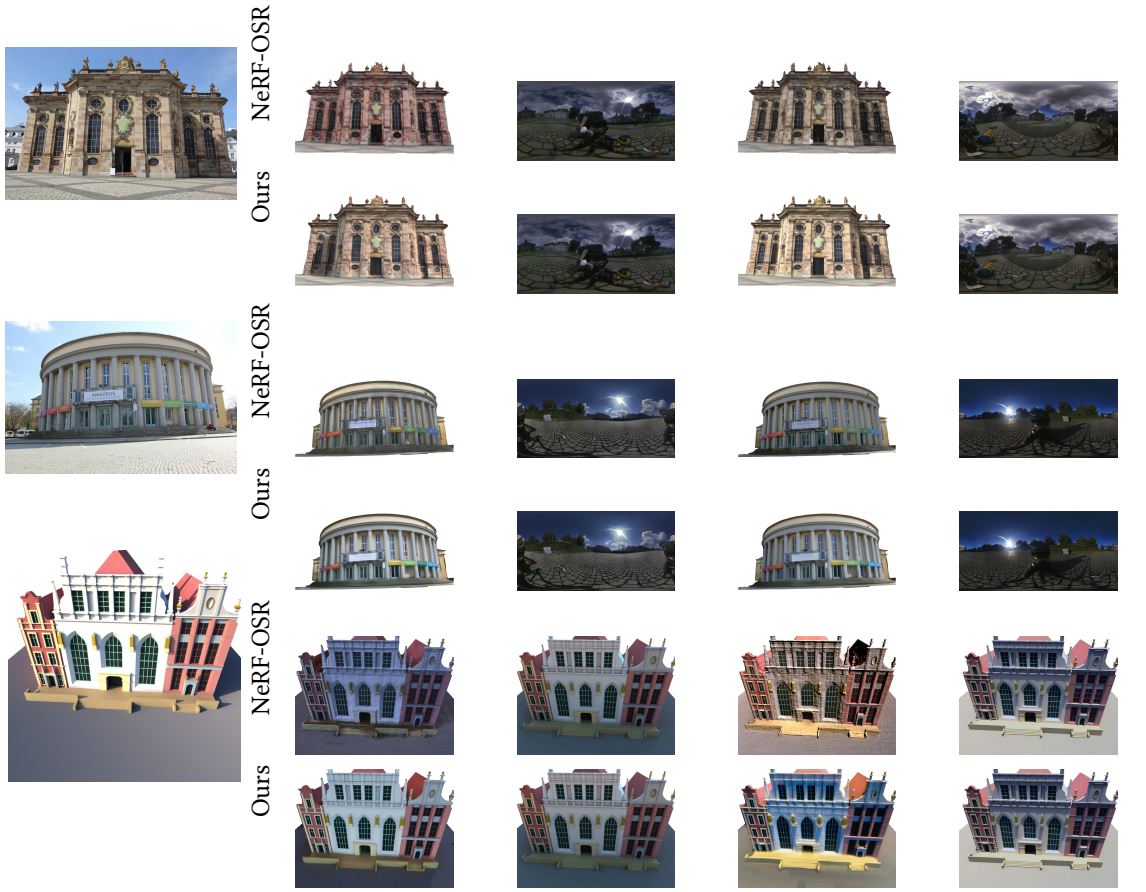
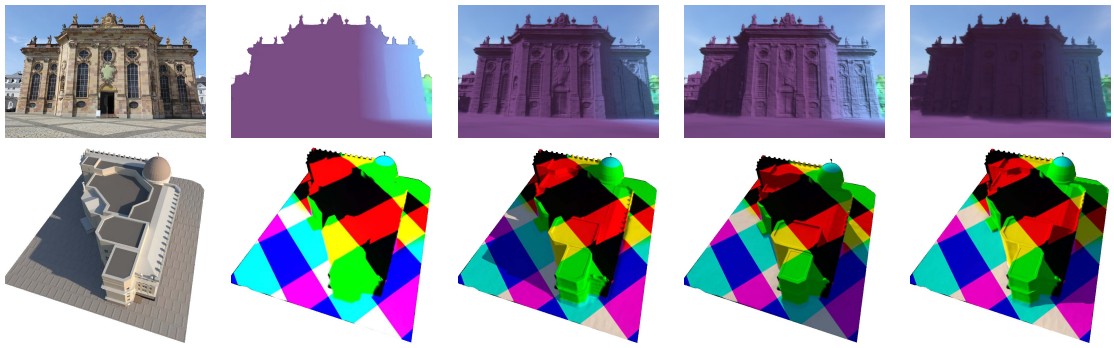


Fig. 6. **Relighting results of NeRF-OSR and our method.** For each input scene, we relight it with two different lighting conditions (as visualized using either environment maps for real scenes, or ground truth rendering for synthetic scenes) and show rendered results of both methods.



(a) Input Image (b) Edited Albedo (c) Rendered (d) Relighting #1 (e) Relighting #2

Fig. 7. **Editing results by our method.** For each scene, we edit its albedo and render the edited scene under the original lighting and two novel lightings.