Supplementary Material for
BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields

Peng Wang[1,2] Lingzhe Zhao[2] Ruijie Ma[2] Peidong Liu[2†]
[1]Zhejiang University [2]Westlake University
{wangpeng, zhaolingzhe, maruijie, liupeidong}@westlake.edu.cn

1. Introduction

In this supplementary material, we present the formulation of cubic-spline, additional quantitative evaluation of BAD-NeRF under different configurations, visualization of the optimized camera motion trajectories and additional qualitative experimental results.

2. Cubic B-Spline Formulation

Compared to a linear-spline, a more complex camera trajectory within exposure time can be controlled by four control knots, denoted as $\mathbf{T}_0$, $\mathbf{T}_1$, $\mathbf{T}_2$ and $\mathbf{T}_3 \in \mathbf{SE}(3)$. We assume that the four control knots are uniformly sampled at a fixed time interval and the trajectory begins with $\mathbf{T}_0$ at time $t = 0$ and terminates with $\mathbf{T}_3$ at time $t = \tau$, where τ represents the exposure time. We introduce the parameter $u = \frac{t}{\tau}$, where u lies within the interval $[0, 1)$. Based on the De Boor-Cox formula [3], the matrix representation of the cumulative basis $\tilde{\mathbf{B}}(u)$ can be written as

$$\tilde{\mathbf{B}}(u) = \mathbf{C} \begin{bmatrix} 1 \\ u \\ u^2 \\ u^3 \end{bmatrix}, \quad \mathbf{C} = \frac{1}{6} \begin{bmatrix} 6 & 0 & 0 & 0 \\ 5 & 3 & -3 & 1 \\ 1 & 3 & 3 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

The virtual camera pose at time t can then be represented as

$$\mathbf{T}(u) = \mathbf{T}_0 \cdot \prod_{j=0}^{2} \exp(\tilde{\mathbf{B}}(u)_{j+1} \cdot \mathbf{\Omega}_j), \tag{2}$$

where $\tilde{\mathbf{B}}(u)_{j+1}$ denotes the $(j+1)^{th}$ element of the vector $\tilde{\mathbf{B}}(u)$, $\mathbf{\Omega}_j = \log(\mathbf{T}_j^{-1} \cdot \mathbf{T}_{j+1})$. Assuming there are n virtual sharp images within exposure time, it can be derived that $\mathbf{T}_i$, which corresponds to the i^{th} virtual pose, is differentiable with respect to $\mathbf{T}_0$, $\mathbf{T}_1$, $\mathbf{T}_2$ and $\mathbf{T}_3$. The goal of BAD-NeRF is now to estimate $\mathbf{T}_0$, $\mathbf{T}_1$, $\mathbf{T}_2$ and $\mathbf{T}_3$ for each frame, as well as the learnable parameters of NeRF.

†Corresponding author.

Gamma-distorted Non-distorted

Figure 1. **Results of BAD-NeRF on the *Factory* scene when training with gamma-distorted and non-distorted blurry images. First row:** Input blurry (training) images. **Second row:** Rendering results of BAD-NeRF. **Third row:** Groundtruth.

3. Additional Quantitative Evaluations

In this section, we study the robustness of BAD-NeRF against the non-linearity (i.e. camera response function) of real image formation process. In particular, we apply gamma operation (i.e. with $\gamma = 2.2$ on the *Factory* sequence) on both the blurry and sharp images, to introduce non-linearity in the color space. We then train our model (i.e. without explicitly modeling the gamma correction) with the gamma-distorted blurry images, and then compare the rendered sharp images against the gamma-distorted sharp images. The resulting metric of gamma operation (i.e. with a PSNR metric as 32.07) is almost the same as that without gamma distortion (i.e. with a PSNR metric as 32.08), qualitative results are shown in Fig. 1. It demon-

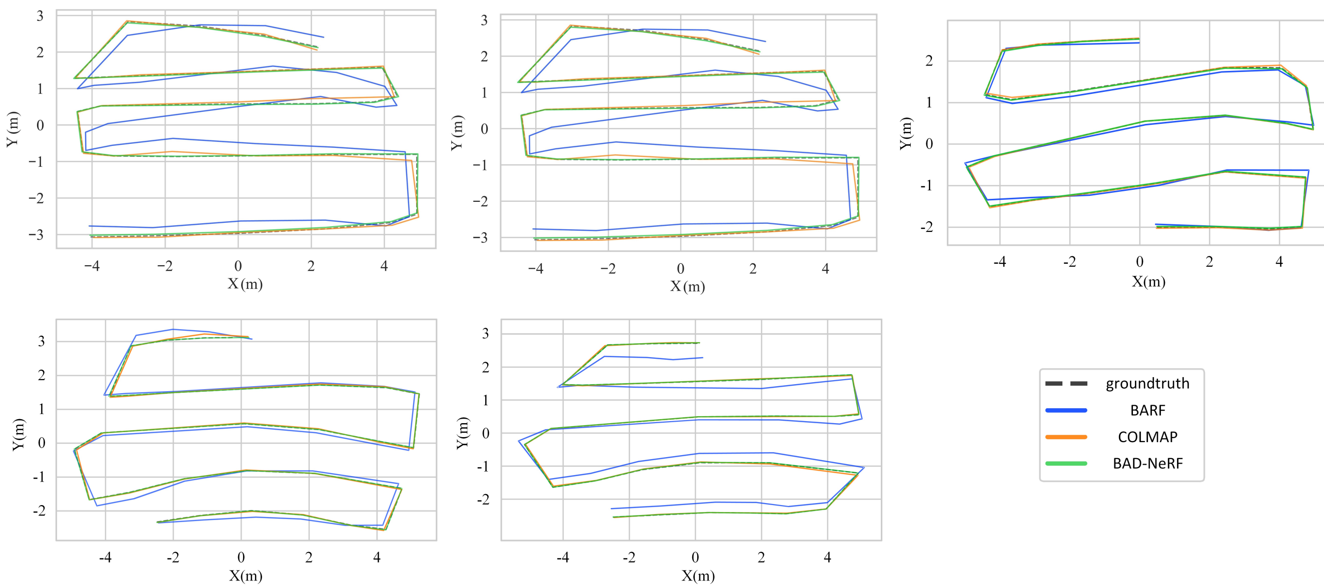

Figure 2. **Qualitative Comparisons of estimated camera poses on Deblur-NeRF dataset.** These are results on *Cozy2room*, *Factory*, *Pool*, *Tanabata* and *Trolley* sequences respectively. The results demonstrate that our method delivers reasonable camera pose estimations and performs better than both COLMAP [4] and BARF [1].

strates the robustness of our method against the nonlinear camera response function.

4. Trajectory Visualization

In this section, we present the visualization results in terms of camera pose estimation. The experiments are conducted on the synthetic dataset of Deblur-NeRF [2]. We present the comparison result of BAD-NeRF against BARF [1] and COLMAP [4] in Fig. 2. It demonstrates that our method recovers reasonable motion trajectories and performs better than both COLMAP [4] and BARF [1].

5. Additional Qualitative Evaluations

We present additional qualitative experimental results on both the synthetic and real datasets. The results are presented in Fig. 3 and Fig. 4 respectively. It further demonstrates the superior performance of our method against prior state-of-the-art approaches.

References

[1] Lin, Chen-Hsuan and Ma, Wei-Chiu and Torralba, Antonio and Lucey, Simon. BARF: Bundle-Adjusting Neural Radiance Fields. In *International Conference on Computer Vision (ICCV)*, 2021. 2

[2] Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and Pedro V Sander. Deblur-NeRF: Neural Radiance Fields from Blurry Images. In *Computer Vision and Pattern Recognition (CVPR)*, pages 12861–12870, 2022. 2, 3, 4

[3] Kaihuai Qin. General Matrix Representations for B-Splines. In *Sixth Pacific Conference on Computer Graphics and Applications*, 1998. 1

[4] Johannes L Schonberger and Jan-Michael Frahm. Structure-from-motion Revisited. In *Computer Vision and Pattern Recognition (CVPR)*, pages 4104–4113, 2016. 2, 3

[5] Hyeongseok Son, Junyong Lee, Jonghyeop Lee, Sunghyun Cho, and Seungyong Lee. Recurrent Video Deblurring with Blur-Invariant Motion Estimation and Pixel Volumes. *ACM Transactions on Graphics (TOG)*, 40(5):1–18, 2021. 4

[6] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In *Computer Vision and Pattern Recognition (CVPR)*, 2018. 3, 4

[7] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage Progressive Image Restoration. In *Computer Vision and Pattern Recognition (CVPR)*, pages 14821–14831, 2021. 4

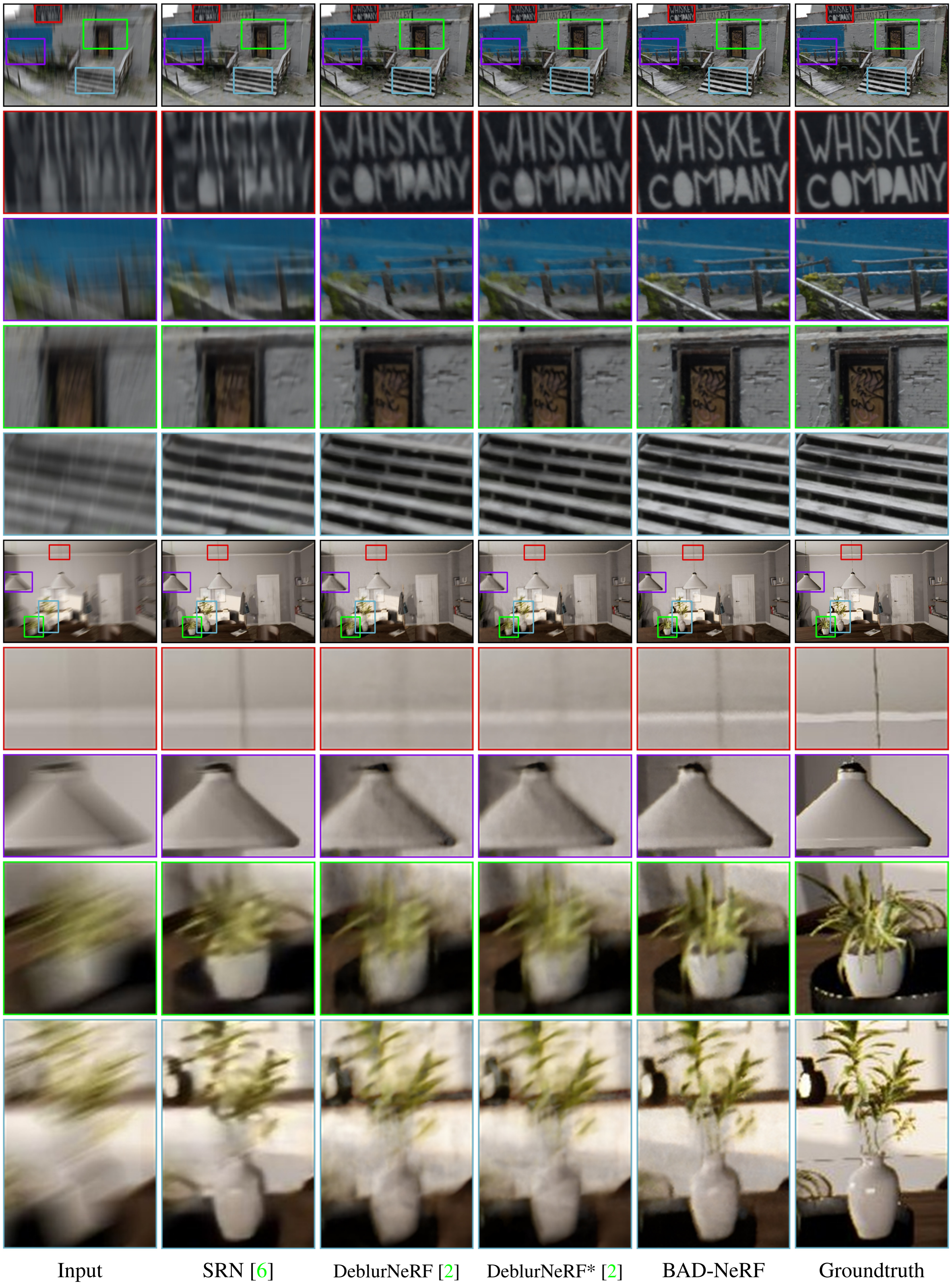

Figure 3. **Qualitative results of different methods with synthetic datasets.** Note that DeblurNeRF* is trained with the ground-truth poses, while the other one is trained with the estimated poses by COLMAP [4]. It demonstrates that our method can deliver sharper results compared to prior works.

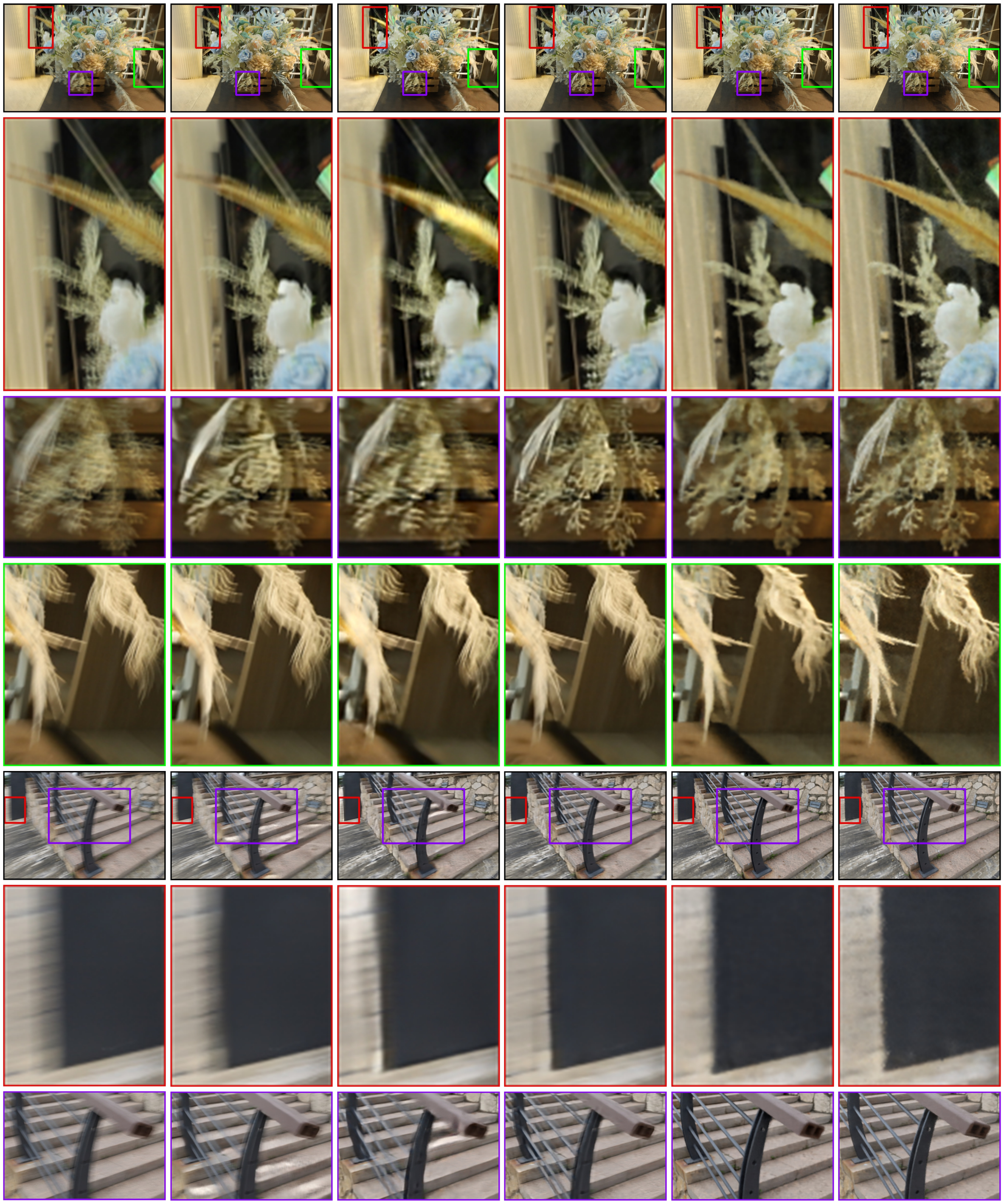

Figure 4. **Qualitative results of different methods with real datasets.** It demonstrates that our method recovers sharper results compared to prior works, especially for boundaries with large depth variations compared to Deblur-NeRF [2].