Supplementary Materials for Deblurring Neural Radiance Fields with Event-driven Bundle Adjustment

Yunshan Qi
State Key Laboratory of Virtual Reality Technology and Systems, SCSE, Beihang University
Beijing, China
qi_yunshan@buaa.edu.cn

Lin Zhu*
School of Computer Science and Technology, Beijing Institute of Technology
Beijing, China
linzhu@bit.edu.cn

Yifan Zhao
State Key Laboratory of Virtual Reality Technology and Systems, SCSE, Beihang University
Beijing, China
zhaoyf@buaa.edu.cn

Nan Bao
State Key Laboratory of Virtual Reality Technology and Systems, SCSE, Beihang University
Beijing, China
nbao@buaa.edu.cn

Jia Li*
State Key Laboratory of Virtual Reality Technology and Systems, SCSE, Beihang University
Beijing, China
jiali@buaa.edu.cn

6 Supplementary Qualitative Comparison on Synthetic Scenes

As shown in Figure 9, in the Cozyroom scene there is no apparent qualitative difference in the results of the BAD-NeRF [4], E^2NeRF [1] and EBAD-NeRF because the camera motion is even and slight. However, in the other four scenes, the results of our EBAD-NeRF have been significantly improved compared to BAD-NeRF and E^2NeRF, proving that when facing complicated camera motion, our proposed event-driven bundle adjustment can effectively provide accurate estimated poses for the NeRF network and with event-enhanced imaging simulation our method can recover a sharp NeRF from the severely blurry images and corresponding events.

In the Factory scene, EBAD-NeRF's reconstruction of scars on the wall is closest to ground truth. In comparison, the results of E^2NeRF are too thick, and the results of BAD-NeRF are still blurry. In the Pool scene, E^2NeRF misses the lotus leaf in the pool and generates blurrier results in the Tanabata scene than EBAD-NeRF. Artifacts appear on the white flag in the results of E^2NeRF in the Wine scene. Since there is no supervision of event data for pose estimation and image motion blur modeling, the results of BAD-NeRF are not sharp enough compared to ERGB-based methods E^2NeRF and EBAD-NeRF. The pre-deblurring-based method MPR-NeRF [5], SRN-NeRF [3], and D2Net-NeRF [2] is not good enough due to the limitations of the deblurring method itself. Overall, our EBAD-NeRF achieved the best results in all synthetic scenes.

7 Supplementary Qualitative Comparison on Real-World Scenes

Figure 10 shows the two scenes of the real data. In the Bar scene, the synthesis novel view results of EBAD-NeRF are closest to ground truth, especially for the light spot in the upper left corner and the characters on the billboard. Other methods have obvious reconstruction errors of the light points and blur results of the characters. In the Classroom scene, our EBAD-NeRF obtains sharper letters and fewer artifacts in the synthesis results than other methods. The experiments on real-world scenes demonstrate that our EBAD-NeRF is more effective than the ERGB-based deblurring NeRF method E^2NeRF, which without pose optimizing during the training. Also, our method has significantly improved over the image-based deblurring NeRF methods and NeRF with pre-deblurring images.

8 Supplementary Video

We show the video rendering results for the ablation study, synthetic scenes, and real-world scenes on our project website: https://icvteam.github.io/EBAD-NeRF.html.

The video rendering results for the ablation study prove the effectiveness of our proposed event-driven bundle adjustment. For the synthetic scene, our method achieves the sharpest video rendering results. The results of E^2NeRF are a little misaligned in space because the poses estimated by its pre-deblurring framework are slightly different from the initial poses given for the other methods. For the real-world scene, the results of NeRF and NeRF with pre-deblurring images are severely blurry and contain many cloudy materials. In comparison, our EBAD-NeRF has sharper rendering results with less cloudy materials than Deblur-NeRF, BAD-NeRF, and E^2NeRF. Overall, our EBAD-NeRF achieves the best video rendering results in both real-world and synthetic scenes.

References

[1] Yunshan Qi, Lin Zhu, Yu Zhang, and Jia Li. 2023. E2NeRF: Event Enhanced Neural Radiance Fields from Blurry Images. In *Proceedings of the IEEE/CVF ICCV*. 13254–13264.

[2] Wei Shang, Dongwei Ren, Dongqing Zou, Jimmy S Ren, Ping Luo, and Wangmeng Zuo. 2021. Bringing events into video deblurring with non-consecutively blurry frames. In *Proceedings of the IEEE/CVF ICCV*. 4531–4540.

[3] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. 2018. Scale-Recurrent Network for Deep Image Deblurring. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8174–8182.

[4] Peng Wang, Lingzhe Zhao, Ruijie Ma, and Peidong Liu. 2023. BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields. In *Proceedings of the IEEE/CVF CVPR*. 4170–4179.

[5] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. 2021. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF CVPR*. 14821–14831.

*Correspondence should be addressed to Jia Li and Lin Zhu.
Website: https://cvteam.buaa.edu.cn

Figure 9: Supplementary Reconstruction Qualitative Comparison on Synthetic Scenes.

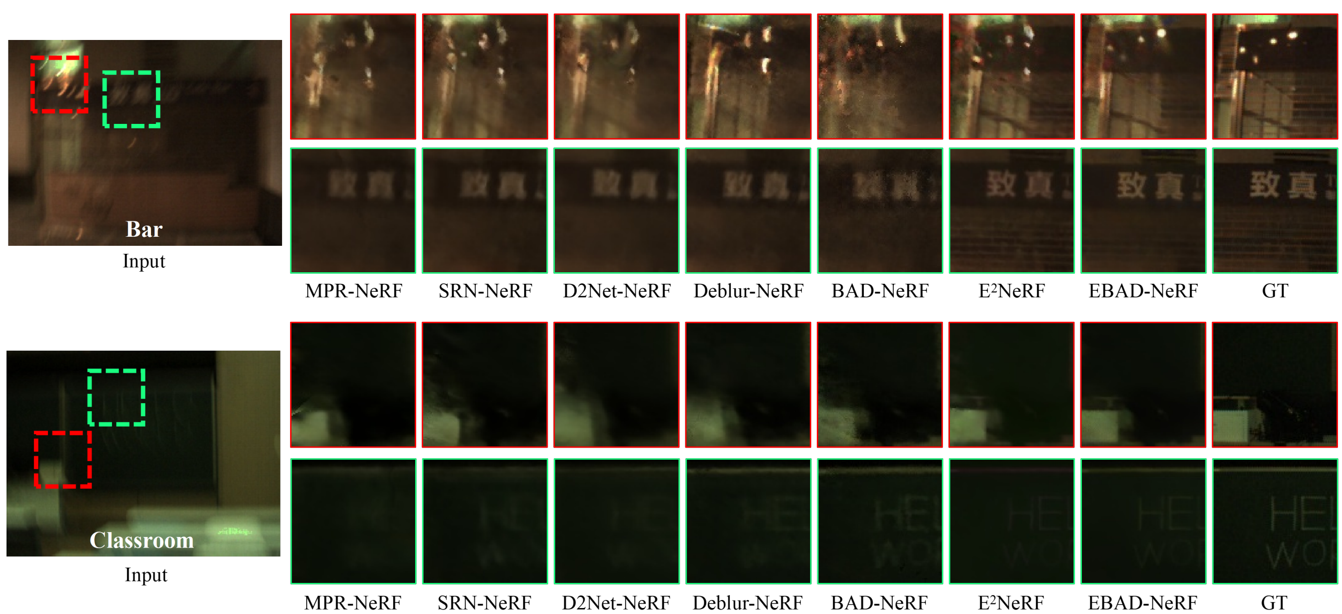

Figure 10: Supplementary Reconstruction Qualitative Comparison on Real Scenes.