Supplementary Materials to "Joint HDR Denoising and Fusion: A Real-World Mobile HDR Image Dataset"

Shuaizheng Liu[1,2], Xindong Zhang[1,2], Lingchen Sun[1,2], Zhetong Liang[2], Hui Zeng[2], Lei Zhang[1,2]
[1]The HongKong Polytechnic University, [2]OPPO Research
{*shuaizhengliu21, cshzeng*}*@gmail.com*; *ling-chen.sun@connect.polyu.hk*;
liangzhetong@oppo.com; {*csxdzhang, cslzhang*}*@comp.polyu.edu.hk*

In this supplementary material, we show more visual comparisons on the test scenes with and without ground-truth (GT), respectively. We compare our Joint-HDRDN method with state-of-the-art deep learning based HDR image reconstruction methods, including DeepHDR [3], NHDRRNet [5], AHDRNet [4] and HDRTransformer [1]. All the test data are captured by mobile phones in real-world scenes, covering various lighting conditions and noise levels. The test data with GT are acquired under controlled motion, including global motion and subject motion, while the test data without GT are captured by hand-held mobile phones in practical scenarios, including walking pedestrian and moving vehicles. All the following results are visualized by a simple ISP followed by the Reinhard [2] tone-mapping operator.

1. Visual comparisons on test scenes with ground-truth

Figs. 1 5 show the visual comparisons on test scenes with GT. As shown in Fig. 1 and Fig. 4, for the scenes where the reference images have over-exposed regions, the competing other methods are susceptible to ghosting artifacts. This is because these methods lack modules for feature alignment, and hence they are prone to introducing unaligned contents from the under-exposed frames. For the dark scenes corrupted by severe noise, as shown in Fig. 2, Fig. 3 (patches highlighted in red box) and Fig. 5, our Joint-HDRDN method restores more details than other methods and generates less noisy artifacts. This is because our pyramid cross-attention modules can better aggregate the correlated features from other frames.

2. Visual comparisons on test scenes without ground-truth

Figs. 6 9 show the visual comparisons on test scenes without GT. Similar conclusions to the above section can be made. In addition, the test data without GT can more faithfully reflect the real motion between different frames. The models trained on our Mobile-HDR dataset captured under controlled motion perform well on these real-world scenes, demonstrating the effectiveness of our collected LDR-HDR paired dataset.

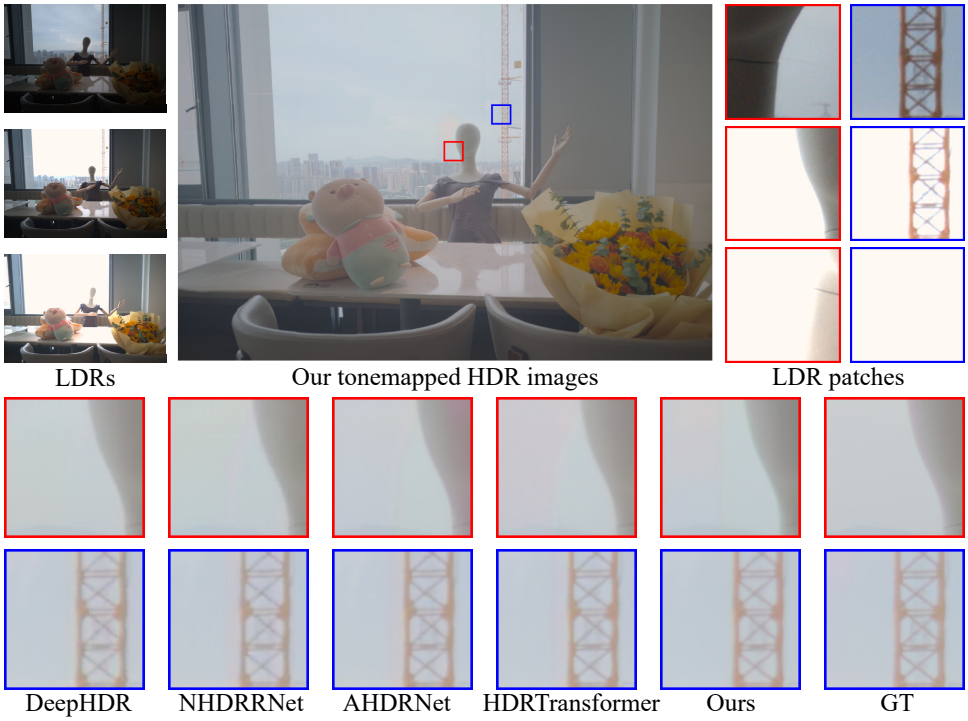

Figure 1. Visual comparison on test Scene A with GT.

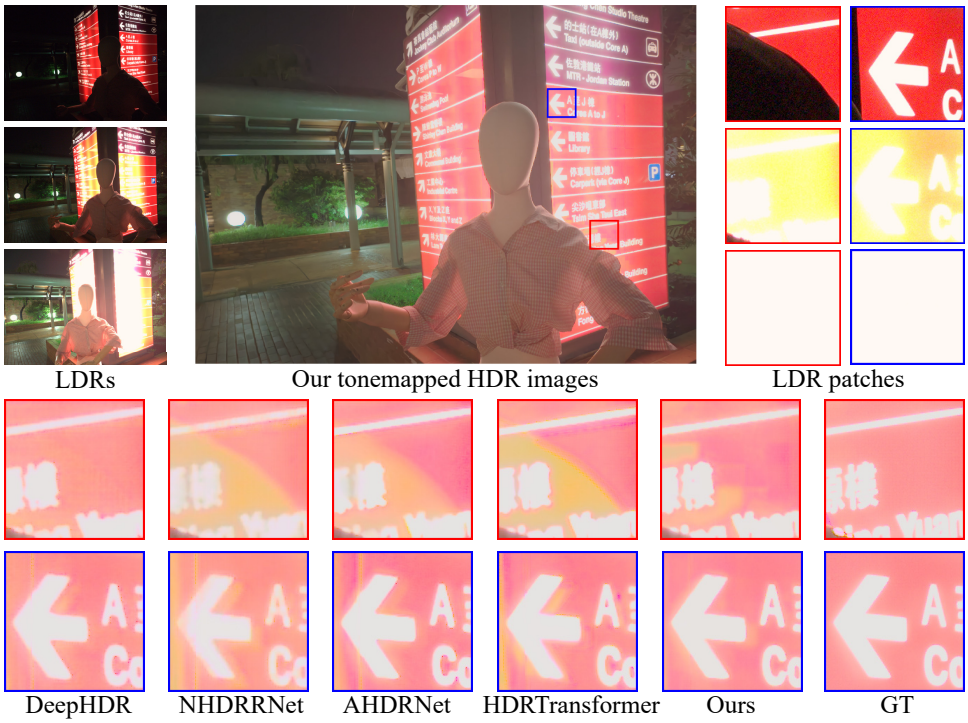

Figure 2. Visual comparison on test Scene B with GT.

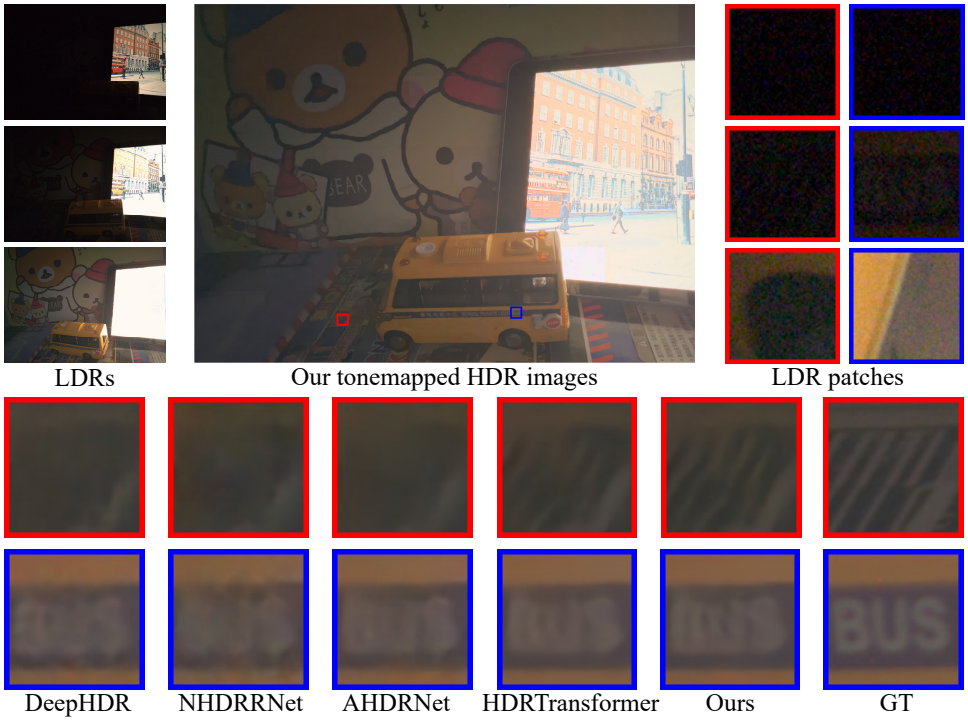

Figure 3. Visual comparison on test Scene C with GT.

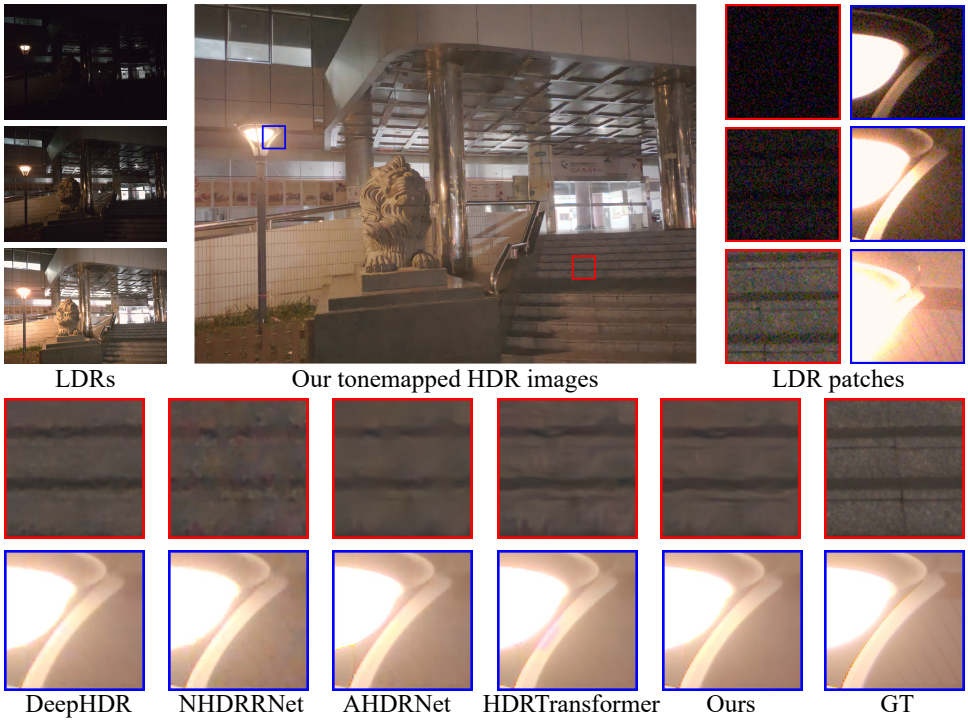

Figure 4. Visual comparison on test Scene D with GT.

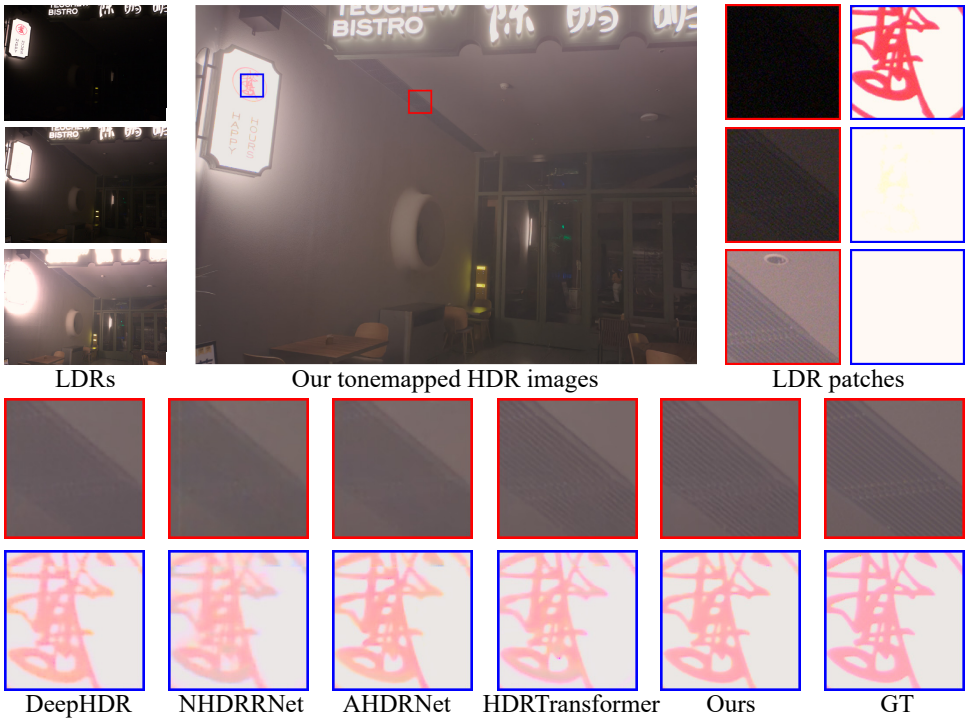

Figure 5. Visual comparison on test Scene E with GT.

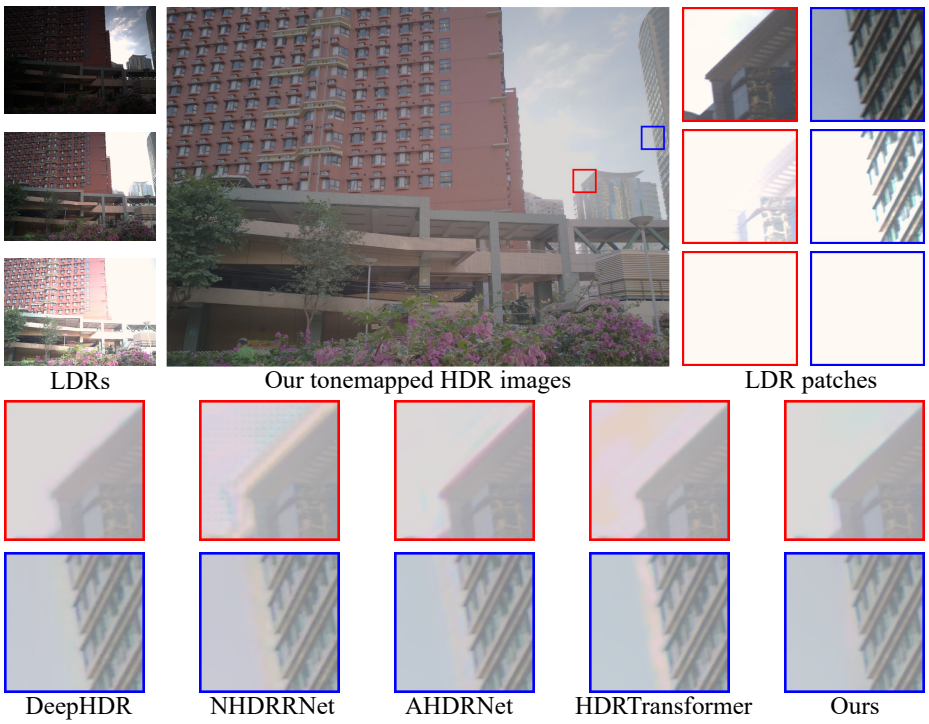

Figure 6. Visual comparison on test Scene A without GT.

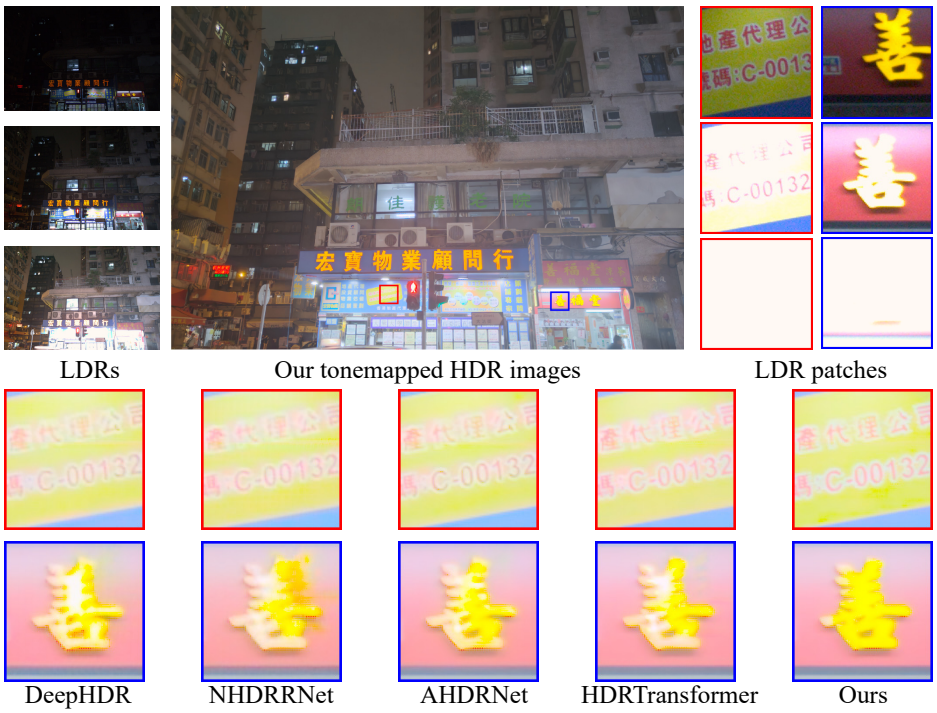

Figure 7. Visual comparison on test Scene B without GT.

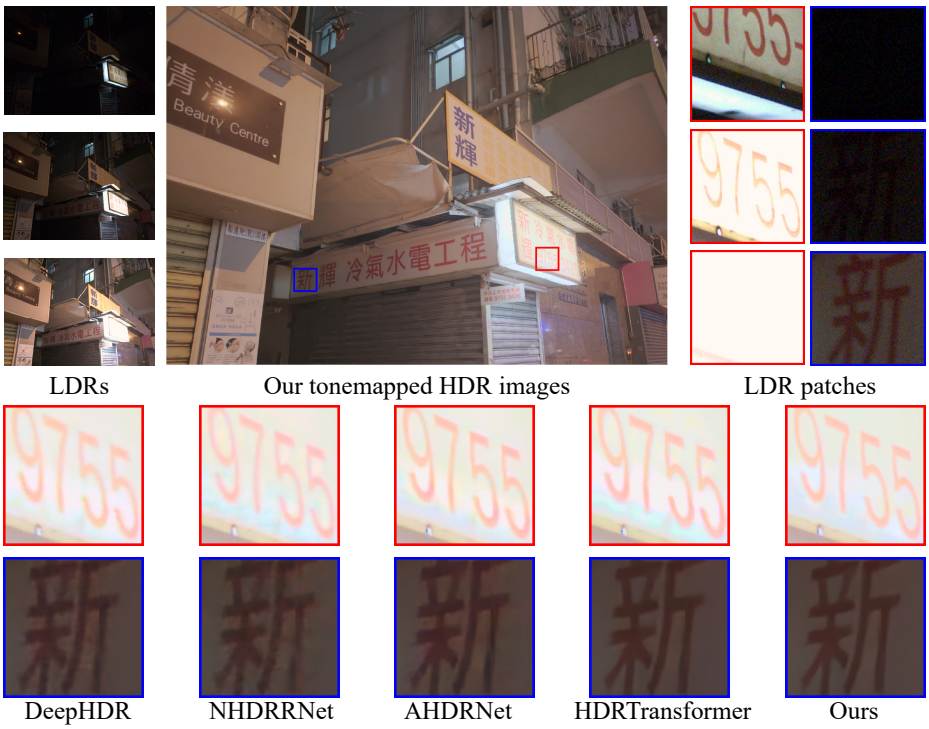

Figure 8. Visual comparison on test Scene C without GT.

Figure 9. Visual comparison on test Scene D without GT.

References

[1] Zhen Liu, Yinglong Wang, Bing Zeng, and Shuaicheng Liu. Ghost-free high dynamic range imaging with context-aware transformer. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIX*, pages 344–360. Springer, 2022. 1

[2] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 267–276, 2002. 1

[3] Shangzhe Wu, Jiarui Xu, Yu-Wing Tai, and Chi-Keung Tang. Deep high dynamic range imaging with large foreground motions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018. 1

[4] Qingsen Yan, Dong Gong, Qinfeng Shi, Anton van den Hengel, Chunhua Shen, Ian Reid, and Yanning Zhang. Attention-guided network for ghost-free high dynamic range imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1751–1760, 2019. 1

[5] Qingsen Yan, Lei Zhang, Yu Liu, Yu Zhu, Jinqiu Sun, Qinfeng Shi, and Yanning Zhang. Deep hdr imaging via a non-local network. *IEEE Transactions on Image Processing*, 29:4308–4322, 2020. 1