Ghost-free High Dynamic Range Imaging with Context-aware Transformer

Zhen Liu[1,*], Yinglong Wang[2,*], Bing Zeng[3], and Shuaicheng Liu[3,1,†]

[1] Megvii Technology, Beijing, China
liuzhen03@megvii.com
[2] Noah's Ark Lab, Huawei Technologies, Shenzhen, China
ylwanguestc@gmail.com
[3] University of Electronic Science and Technology of China, Chengdu, China
{zengbing,liushuaicheng}@uestc.edu.cn
https://github.com/megvii-research/HDR-Transformer
*Joint First Author, †Corresponding Author

1 Empirical Study on CA-ViT

To better explore the effectiveness of the proposed Context-aware Vision Transformer (CA-ViT), we conduct an empirical study on the number of CA-ViT in each Context-aware Transformer Block (CTB). The results are depicted in Fig. 1. To balance the training computational budgets and the performance, we set the number of CA-ViT in each RTB to 6 in our proposed HDR-Transformer.

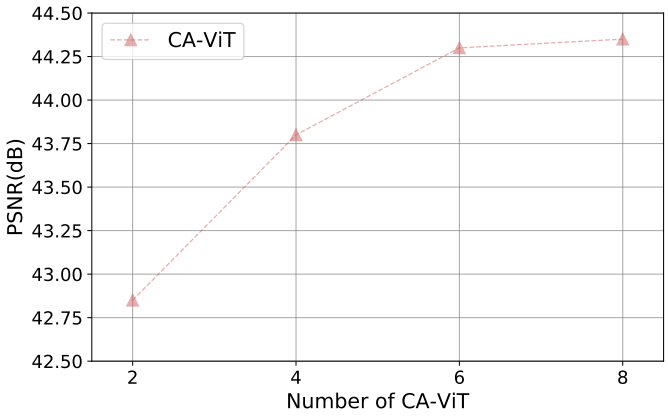

Fig. 1. Empirical study on the CA-ViT.

2 Full Results on the Kalantari *et al.*'s Dataset

For better comparison with other methods, we also show the results with full images on Kalantari *et al.*'s [2] dataset. As shown in Fig. 3 and Fig. 4, the first row lists the input LDR images and the HDR result directly fused by Photomatix Pro. The last two rows list the comparison results and the ground truth HDR image. For each method, we crop several patches (with a range of about 200 pixels) for better comparison. As seen, the result directly fused by Photomatix Pro produces undesired ghosting artifacts caused by long-range hands movements. The patch match based methods and CNN-based methods also fail to effectively remove the ghosting artifacts and cause distortion artifacts as described in our paper. On the contrary, our results are free of ghosting artifacts and more visually pleasing.

3 Results on the Prabhakar *et al.*'s Dataset

For the Prabhakar *et al.*'s [5] dataset, we train our method on it from scratch and report the quantitative results in Table 1. It is observed that our proposed HDR-Transformer is up to 6.2dB and 4.7dB higher than Prabhakar20's [4] method in terms of PSNR-l and PSNR-μ, respectively, which further demonstrates the superiority of our method over existing methods.

Table 1. Quantitative results on the Prabhakar *et al.*'s dataset. All the results are averaged over 116 testing samples.

Metrics	Sen12	Hu13	Kalantari17	DeepHDR	AHDRNet	Prabhakar20	Ours
PSNR-l	29.57	28.87	32.08	30.72	31.83	<u>32.52</u>	**38.77**
PSNR-μ	32.09	30.82	35.34	32.31	33.72	<u>35.84</u>	**40.53**
HDR-VDP-2	62.43	60.47	64.47	64.03	64.32	<u>64.76</u>	**66.81**

4 Visual Results on the Perceptual Loss

To verify the effectiveness of the perceptual loss, we conduct experiments by training the HDR-Transformer both with and without the loss term. The qualitative results are shown in Fig. 2. As can be seen, when suffering occlusion and intensity variation, such regions (highlighted by red boxes) produce residual distortions and inconsistent details. By using the perceptual loss, the results (highlighted by blue boxes) are more reasonable and visually consistent.

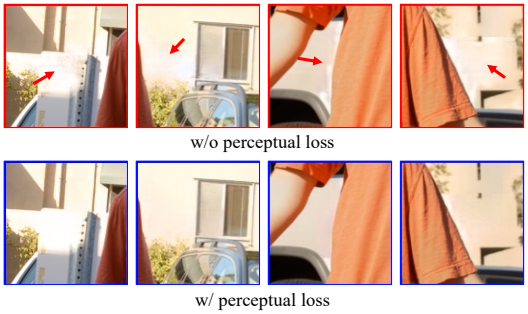

Fig. 2. Visual results on the perceptual loss.

Fig. 3. Visual Comparison of our method with the state-of-the-art methods [6,1,2,7,8,3] on Kalantari *et al* [2]'s dataset.

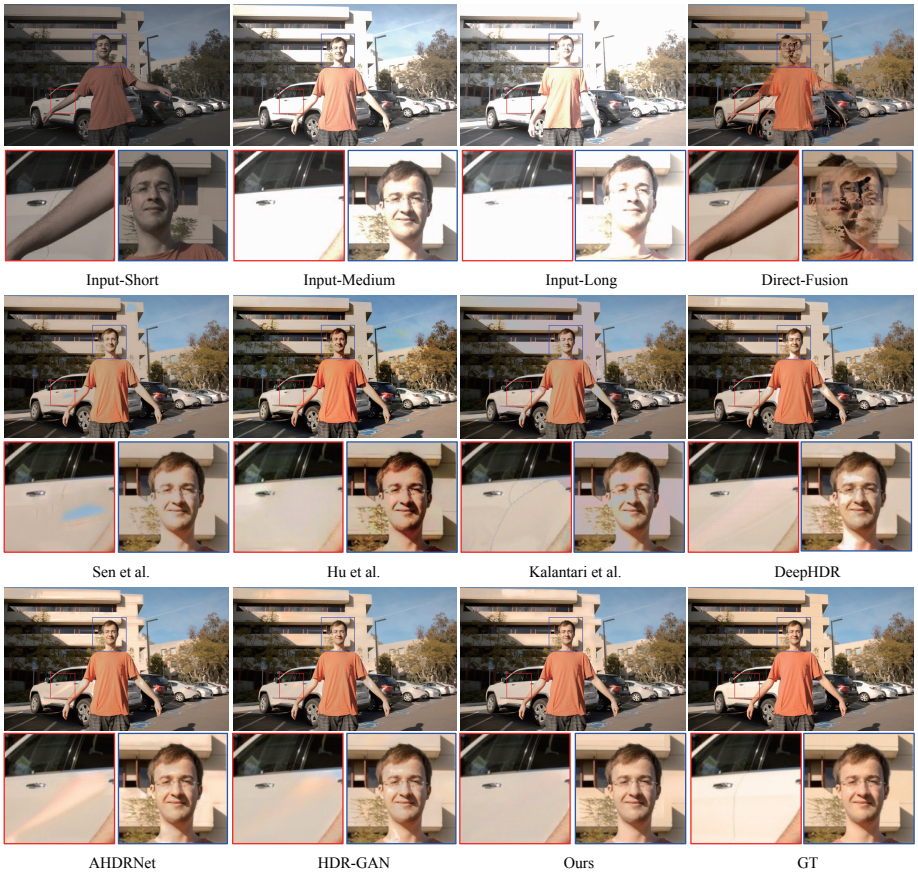

Fig. 4. Visual Comparison of our method with the state-of-the-art methods [6,1,2,7,8,3] on Kalantari *et al* [2]'s dataset.

References

1. Hu, J., Gallo, O., Pulli, K., Sun, X.: Hdr deghosting: How to deal with saturation? In: Proc. CVPR. pp. 1163–1170 (2013) 3, 4
2. Kalantari, N.K., Ramamoorthi, R.: Deep high dynamic range imaging of dynamic scenes. ACM Trans. Graphics **36**(4), 144 (2017) 2, 3, 4
3. Niu, Y., Wu, J., Liu, W., Guo, W., Lau, R.W.: Hdr-gan: Hdr image reconstruction from multi-exposed ldr images with large motions. IEEE Trans. on Image Processing **30**, 3885–3896 (2021) 3, 4
4. Prabhakar, K.R., Agrawal, S., Singh, D.K., Ashwath, B., Babu, R.V.: Towards practical and efficient high-resolution hdr deghosting with cnn. In: European Conference on Computer Vision. pp. 497–513. Springer (2020) 2
5. Prabhakar, K.R., Arora, R., Swaminathan, A., Singh, K.P., Babu, R.V.: A fast, scalable, and reliable deghosting method for extreme exposure fusion. In: 2019 IEEE International Conference on Computational Photography (ICCP). pp. 1–8. IEEE (2019) 2
6. Sen, P., Kalantari, N.K., Yaesoubi, M., Darabi, S., Goldman, D.B., Shechtman, E.: Robust patch-based hdr reconstruction of dynamic scenes. ACM Trans. Graphics **31**(6), 203 (2012) 3, 4
7. Wu, S., Xu, J., Tai, Y.W., Tang, C.K.: Deep high dynamic range imaging with large foreground motions. In: Proc. ECCV. pp. 117–132 (2018) 3, 4
8. Yan, Q., Gong, D., Shi, Q., Hengel, A.v.d., Shen, C., Reid, I., Zhang, Y.: Attention-guided network for ghost-free high dynamic range imaging. In: Proc. CVPR. pp. 1751–1760 (2019) 3, 4