

Supplementary Material

Polarization-Aware Low-Light Image Enhancement

Chu Zhou¹, Minggui Teng², Youwei Lyu³, Si Li³, Chao Xu¹, Boxin Shi^{2*}

¹Key Laboratory of Machine Perception (MOE), School of Intelligence Science and Technology, Peking University

²National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

³School of Artificial Intelligence, Beijing University of Posts and Telecommunications

Details of the Simulation Experiment about the Average Error Rates

In this section, we provide details of the simulation experiment about the average error rates, corresponding to Footnote 4 of the paper.

First, we take the polarized images captured with a long exposure time in the PLIE dataset as the ground truth normal-light images $\mathbf{I}_{\alpha_{1,2,3,4}}$. Then, according to Eq. (6) of the paper, we simulate the process of generating the corresponding low-light images $\hat{\mathbf{I}}_{\alpha_{1,2,3,4}}(\gamma)$ using

$$\hat{\mathbf{I}}_{\alpha_i}(\gamma) = \begin{cases} \frac{1}{\gamma}\mathbf{I}_{\alpha_i} + \mathbf{N}_i & \text{if } \gamma > 1 \\ \mathbf{I}_{\alpha_i} & \text{if } \gamma = 1 \end{cases} \quad (i = 1, 2, 3, 4),$$

where γ is a linear scaling factor denoting the image irradiance reduction caused by decreasing the exposure time or scene radiance, and $\mathbf{N}_i = \mathcal{N}(\frac{1}{\gamma}\mathbf{I}_{\alpha_i})$ is a noise term (we adopt the same settings as Lv, Li, and Lu (2021)). We generate $\hat{\mathbf{I}}_{\alpha_i}(\gamma)$ with 10 different γ ($\gamma \in [1, 10]$ and $\gamma \in \mathbb{Z}^+$), and compute the average error rates of the polarized images, AoP, DoP, and Stokes parameters for each γ . Finally, we get the relationships between the average error rates and γ , as shown in Fig. 1 (c) of the paper.

More Information about the PLIE Dataset

In this section, we provide more information about the PLIE dataset, corresponding to Footnote 5 of the paper.

First, we use a Lucid Vision Phoenix polarization camera to capture 130 different indoor scenes with variant objects, materials, and scene depths as the data source to make the PLIE dataset (some of the source images are shown in Fig. 6). Each scene is captured twice, with a short exposure time t_{short} (as the low-light one) and a long exposure time $t_{\text{long}} = 10t_{\text{short}}$ (as the normal-light one). Note that we should not choose a very large linear scaling factor (e.g., $\gamma = 30$) like other single-image low-light enhancement datasets, since the DoP and AoP could be already significantly degenerated when $\gamma = 10$; and if γ is too large, the polarization information could be too lacking. In each capture, the camera outputs four spatially-aligned

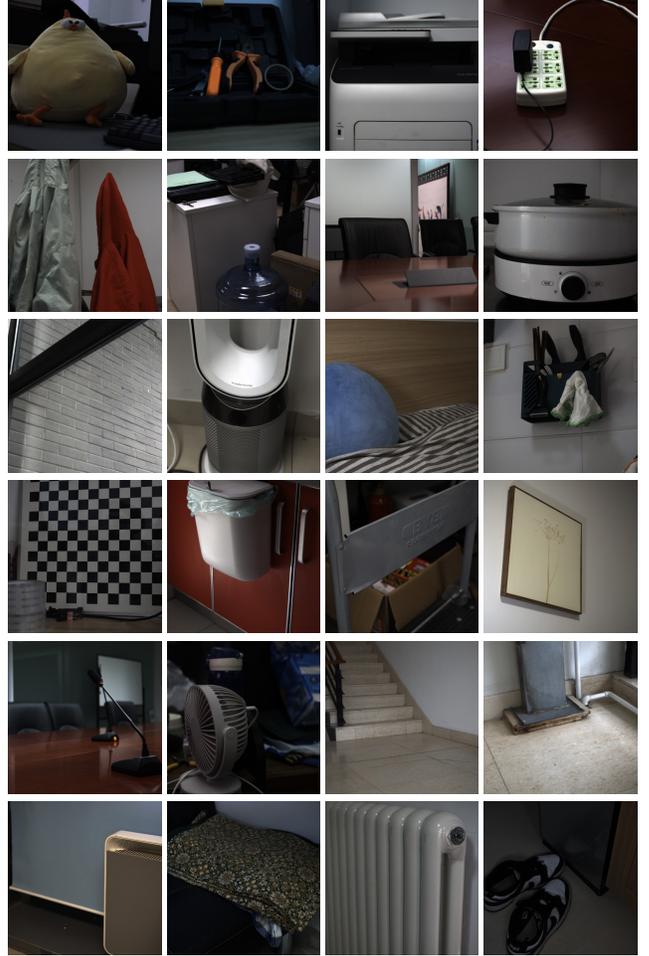


Figure 6: A gallery showing some example scenes of the PLIE dataset.

and temporally-synchronized polarized images with different polarizer angles $\alpha_{1,2,3,4} = 0^\circ, 45^\circ, 90^\circ, 135^\circ$. For keeping the camera untouched to avoid misalignment caused by camera motion during the capturing procedure, we place the camera on a sturdy tripod and use software to change the exposure time. The original spatial resolution of the captured

*Corresponding author: shiboxin@pku.edu.cn

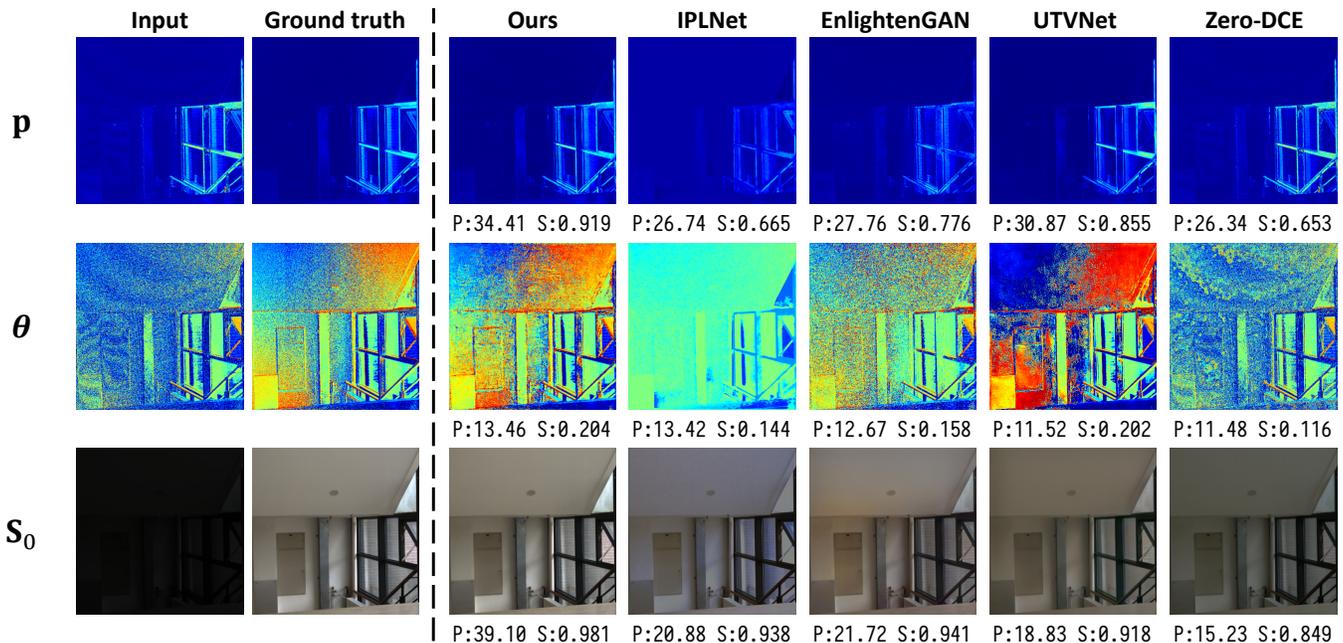


Figure 7: Additional visual quality comparisons (part 1) on the PLIE dataset among our method, IPLNet (Hu et al. 2020), EnlightenGAN (Jiang et al. 2021), UTVNet (Zheng, Shi, and Shi 2021), and Zero-DCE (Guo et al. 2020). Quantitative results evaluated using PSNR (P) and SSIM (S) are displayed below each image.

	PSNR-p	SSIM-p	PSNR- θ	SSIM- θ	PSNR- S_0	SSIM- S_0
Ours	5.41	0.11	2.44	0.08	2.81	0.01
IPLNet (Hu et al. 2020)	4.31	0.08	2.63	0.10	6.35	0.04
EnlightenGAN (Jiang et al. 2021)	4.73	0.11	1.65	0.04	4.85	0.04
UTVNet (Zheng, Shi, and Shi 2021)	4.97	0.11	2.78	0.09	4.02	0.09
Zero-DCE (Guo et al. 2020)	5.47	0.15	1.31	0.03	4.39	0.04

Table 3: The standard deviations of the quantitative evaluation results on the PLIE dataset among our method, IPLNet (Hu et al. 2020), EnlightenGAN (Jiang et al. 2021), UTVNet (Zheng, Shi, and Shi 2021), and Zero-DCE (Guo et al. 2020).

polarized images is 1224×1024 , and we crop the images in the middle to discard dark corners caused by vignetting effect so that the spatial resolution becomes 1024×1024 .

Then, we randomly split the data source into two parts that contain 100 and 30 scenes for making the training and test sets respectively. When making the training set, we crop each 1024×1024 source image in the middle to obtain four 512×512 images, and perform data augmentation (*e.g.*, random cropping, flipping, and rotating) on them, so that we have 6000 different 256×256 images for training. The process of making the test set is similar to the one of making the training set, while in which we do not perform data augmentation so that the spatial resolution of test images is 512×512 .

Additional Visual Quality Comparisons on the PLIE Dataset

In this section, we provide additional visual quality comparisons on the PLIE dataset among our method, IPLNet

(Hu et al. 2020), (the only existing method designed for enhancing polarized low-light images as far as we know), and three state-of-the-art single-image low-light enhancement methods including EnlightenGAN (Jiang et al. 2021), UTVNet (Zheng, Shi, and Shi 2021), and Zero-DCE (Guo et al. 2020), as shown in Fig. 7 and Fig. 8, corresponding to Footnote 7 of the paper.

The Standard Deviations

In this section, we provide the standard deviations of the quantitative evaluation results on the PLIE dataset among our method, IPLNet (Hu et al. 2020), EnlightenGAN (Jiang et al. 2021), UTVNet (Zheng, Shi, and Shi 2021), and Zero-DCE (Guo et al. 2020), as shown in Tab. 3.

Additional Results of the Applications of Polarization-Based Vision

In this section, we provide additional results of the applications of polarization-based vision, including reflection re-

	Input	Ground truth	Ours	IPLNet	EnlightenGAN	UTVNet	Zero-DCE
p							
			P:25.11 S:0.718	P:17.65 S:0.598	P:21.11 S:0.569	P:19.95 S:0.581	P:15.53 S:0.414
θ							
			P:15.73 S:0.509	P:15.68 S:0.355	P:11.51 S:0.268	P:15.43 S:0.496	P:12.71 S:0.221
S_0							
			P:37.94 S:0.982	P:15.97 S:0.947	P:17.58 S:0.861	P:19.63 S:0.798	P:14.41 S:0.854
p							
			P:25.32 S:0.739	P:24.72 S:0.687	P:23.73 S:0.636	P:23.26 S:0.680	P:16.69 S:0.454
θ							
			P:19.16 S:0.402	P:18.17 S:0.308	P:15.17 S:0.225	P:19.12 S:0.397	P:11.90 S:0.123
S_0							
			P:37.72 S:0.979	P:32.52 S:0.969	P:24.37 S:0.871	P:16.91 S:0.759	P:23.26 S:0.882

Figure 8: Additional visual quality comparisons (part 2) on the PLIE dataset among our method, IPLNet (Hu et al. 2020), EnlightenGAN (Jiang et al. 2021), UTVNet (Zheng, Shi, and Shi 2021), and Zero-DCE (Guo et al. 2020). Quantitative results evaluated using PSNR (P) and SSIM (S) are displayed below each image.

removal (using PRRPAW (Lei et al. 2020)) and shape from polarization (using DP3I (Deschaintre, Lin, and Ghosh 2021)), as shown in Fig. 9 and Fig. 10 respectively, corresponding to Footnote 9 and Footnote 10 of the paper respectively. Note that for shape from polarization, we do not have the ground truth normal maps for these two objects, instead we provide the normal maps estimated from normal-light polarized images as references without computing quantitative metrics. From Fig. 9 we can see that our method can im-

prove the performance of reflection removal both quantitatively and qualitatively, and outperforms IPLNet (Hu et al. 2020) consistently. Although from Fig. 10 we cannot quantitatively tell how many degrees of mean angular error (MAE) the enhanced normal maps decrease, it is obvious that the estimated normal maps after enhancement are cleaner and smoother by a large margin (note that these two objects both have smooth surfaces), while IPLNet (Hu et al. 2020) brings negative effects since it generates over-smooth DoP

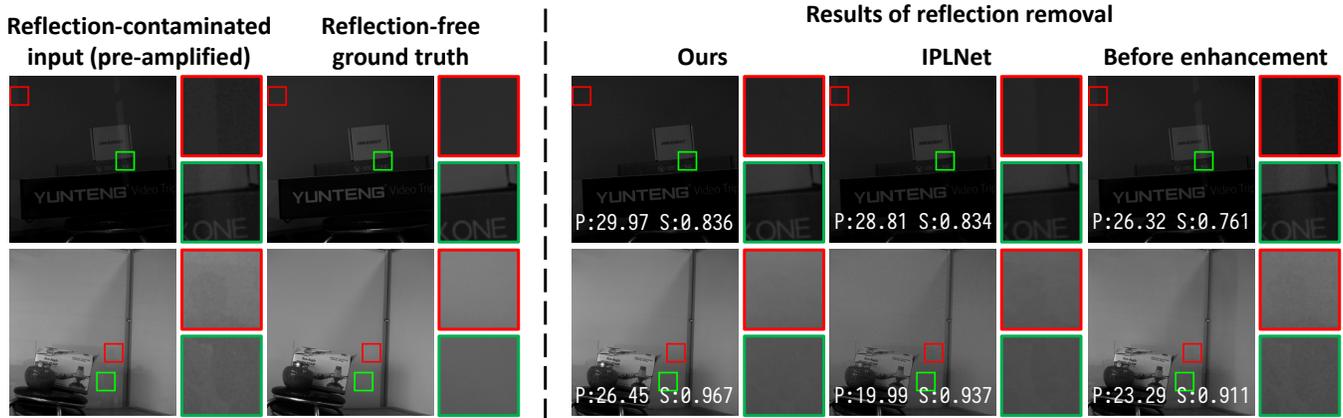


Figure 9: Additional results of reflection removal (using PRRPAW (Lei et al. 2020)) before and after enhancement by our method and IPLNet (Hu et al. 2020). Quantitative results evaluated using PSNR (P) and SSIM (S) are displayed below each image. Please zoom-in for better details.

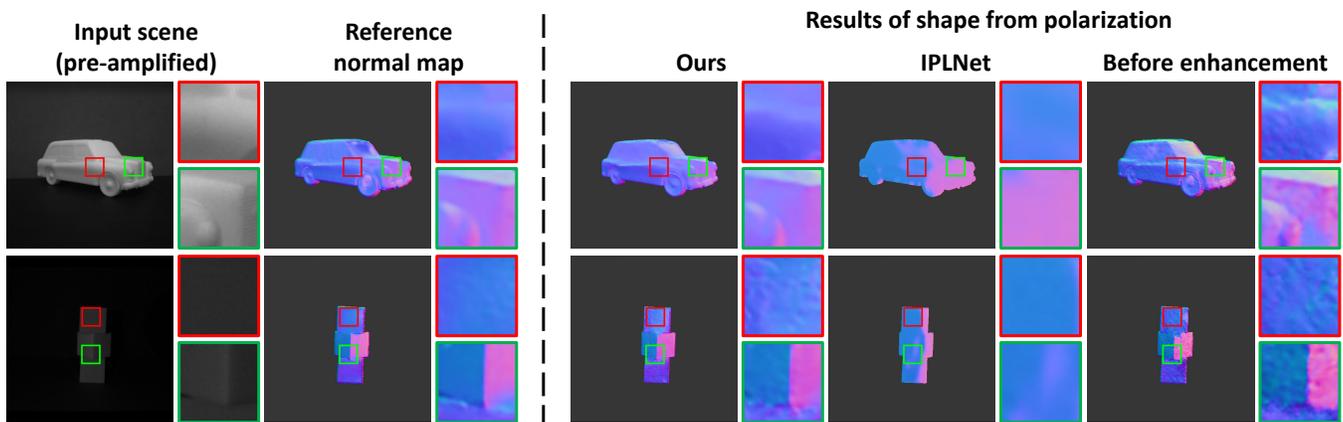


Figure 10: Additional results of shape from polarization (using DP3I (Deschaintre, Lin, and Ghosh 2021)) before and after enhancement by our method and IPLNet (Hu et al. 2020). Please zoom-in for better details.

and AoP, leading to degenerated normal maps.

References

- Deschaintre, V.; Lin, Y.; and Ghosh, A. 2021. Deep polarization imaging for 3D shape and SVBRDF acquisition. In *Proc. of Computer Vision and Pattern Recognition*.
- Guo, C.; Li, C.; Guo, J.; Loy, C. C.; Hou, J.; Kwong, S.; and Cong, R. 2020. Zero-reference deep curve estimation for low-light image enhancement. In *Proc. of Computer Vision and Pattern Recognition*.
- Hu, H.; Lin, Y.; Li, X.; Qi, P.; and Liu, T. 2020. IPLNet: A neural network for intensity-polarization imaging in low light. *Optics Letters*, 45(22): 6162–6165.
- Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; and Wang, Z. 2021. EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30: 2340–2349.
- Lei, C.; Huang, X.; Zhang, M.; Yan, Q.; Sun, W.; and Chen, Q. 2020. Polarized reflection removal with perfect alignment in the wild. In *Proc. of Computer Vision and Pattern Recognition*.
- Lv, F.; Li, Y.; and Lu, F. 2021. Attention guided low-light image enhancement with a large scale low-light simulation dataset. *International Journal of Computer Vision*, 129(7): 2175–2193.
- Zheng, C.; Shi, D.; and Shi, W. 2021. Adaptive unfolding total variation network for low-light image enhancement. In *Proc. of International Conference on Computer Vision*.