

# Reflection Removal via Realistic Training Data Generation

Youxin Pang  
NLPR-CASIA  
School of AI, UCAS  
pangyouxin2020@ia.ac.cn

Mengke Yuan  
NLPR-CASIA  
School of AI, UCAS  
mengke.yuan@nlpr.ia.ac.cn

Qiang Fu  
KAUST  
qiang.fu@kaust.edu.sa

Dong-Ming Yan  
NLPR-CASIA  
School of AI, UCAS  
yandongming@gmail.com

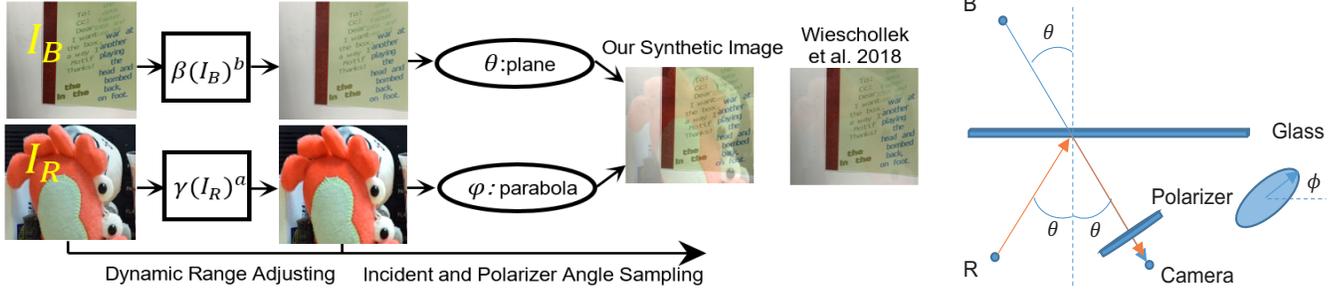


Figure 1: Left: training data synthesis pipeline.(Image courtesy Dennis Sng(Rose Lab)) Right: polarization model.

## ABSTRACT

We present a valid polarization-based reflection contaminated image synthesis method, which can provide adequate, diverse and authentic training dataset. Meanwhile, we enhance the neural network by introducing the reflection information as guidance and utilizing adaptive convolution kernel size to fuse multi-scale information. We demonstrate that the proposed approach achieves convincing improvements over state of the arts.

## KEYWORDS

Reflection removal, Polarization, Network enhancement

### ACM Reference Format:

Youxin Pang, Mengke Yuan, Qiang Fu, and Dong-Ming Yan. 2020. Reflection Removal via Realistic Training Data Generation. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '20 Posters)*, August 17, 2020. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3388770.3407419>

## 1 INTRODUCTION

Recent data-driven reflection removal neural networks [Fan et al. 2017; Wei et al. 2019; Wieschollek et al. 2018] achieve promising results for removing distractive reflection from single image captured behind glass, but their performances show great inconsistency when dealing with real-world scenarios. The challenges stem from : (i) Ground-truths are scant for training, since collecting precise and dense labels in real scenarios is labor-intensive, existing synthetic datasets possess different characteristics comparing with natural

images. (ii) Neither reflection information nor multi-scale context is utilized to regularize the reflection/transmission decomposition.

In this poster, we take the physical polarization property of reflection light into account, thus mimic the attenuation of reflection under different incident and polarizer angles to synthesize more realistic and diverse training data. Furthermore, we integrate reflection information as guidance in training procedure and multi-scale convolution kernel selection units to enhance the neural network. The experiments show the newly constructed dataset and enhanced neural network improve the reflection removal results significantly.

## 2 OUR APPROACH

**Realistic dataset generation.** We first propose a more accurate physically-based dataset construction pipeline to enhance the verisimilitude and diversity of sythetic dataset. Two randomly selected images,  $I_R$  and  $I_B$ , are treated as reflection layer and background layer to compose the target image  $I$ . As shown in Fig. 1(right), placing a polarizer ahead of the camera, with incidence angle  $\theta$  and polarization angle  $\phi$ , the intensity of each pixel  $I(x)$  can be computed according to [Kong et al. 2013] as follows:

$$I(x) = \alpha(\theta, \phi) \frac{I_R(x)}{2} + (1 - \alpha(\theta, \phi)) \frac{I_B(x)}{2}$$

$$\alpha(\theta, \phi) = R_{\perp}(\theta) \sin^2(\phi_{\parallel} - \phi) + R_{\parallel}(\theta) \cos^2(\phi_{\parallel} - \phi) \quad (1)$$

$$R_{\perp}(\theta, \kappa) = \frac{\sin^2(\theta - \theta_t(\theta, \kappa))}{\sin^2(\theta + \theta_t(\theta, \kappa))}, R_{\parallel}(\theta, \kappa) = \frac{\tan^2(\theta - \theta_t(\theta, \kappa))}{\tan^2(\theta + \theta_t(\theta, \kappa))},$$

where  $x$  is the index of input images.  $\alpha(\theta, \phi)$  represents the mixing coefficient.  $\phi_{\parallel}$  is the angle for the orientation of the intersection line between the polarizer and the plane of incidence.  $R_{\perp}$  and  $R_{\parallel}$  are orthogonal decomposing coefficients of reflection light. Moreover, according to Snell's law, we use  $\theta_t(\kappa = \frac{1}{1.474}, \theta) = \arcsin(\frac{1}{1.474} \sin \theta)$ .

To promote the reflection/transmission decomposing capacity of the neural network, we discard the assumption that the intensity of transmitted light should surpass the reflected portion [Wieschollek et al. 2018], and utilize a more flexible formulation to adjust the

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGGRAPH '20 Posters, August 17, 2020, Virtual Event, USA

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-7973-1/20/08.

<https://doi.org/10.1145/3388770.3407419>

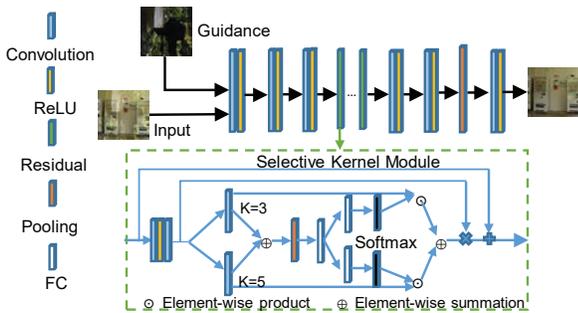


Figure 2: Reflection removal network."©DennisSng(RoseLab)"

dynamic range:

$$L_R = \gamma(I_R)^a, L_B = \beta(I_B)^b, \quad (2)$$

where  $\gamma$ ,  $\beta$ ,  $a$  and  $b$  are random numbers uniform sampled within a certain range [0.7,1.6]. Additionally, we observe that sampling the incidence and polarization angle from the same distribution such as parabola [Wieschollek et al. 2018], the composition may fail (reflection layer is unclear or fades away). Instead, we adopt plane sampling for the incident angle and parabola for the polarization angle, so as to ensure the quality of the image under the premise of diversified collection. See Fig. 1(left) for the generation process.

*Neural network enhancement.* Then, we integrate the reflection information and dynamic convolution kernel selection mechanism into the reflection removal neural network to achieve more accurate reflection/transmission decomposition results.

During the training stage, we feed the neural network with synthesized image, and reflection image as guidance, to facilitate the removal of local strong reflection (see Fig. 2 for details). The architecture proposed in [Fan et al. 2017] is employed as our backbone, and reflection image guidance loss (3) is assembled:

$$\mathcal{L}_R = \lambda \|R - R_0\|_2^2 + \mu (\|\nabla_x R - \nabla_x R_0\|_1 + \|\nabla_y R - \nabla_y R_0\|_1), \quad (3)$$

where  $R_0$  and  $R$  are the predicted and ground truth of reflection image, respectively.  $\nabla_x$  and  $\nabla_y$  are gradients in horizontal and vertical directions.  $\lambda, \mu$  are hyper-parameters. Moreover, we find that dynamic kernel selection unit [Li et al. 2019], which exploits multi-scale global context and adaptively adjusts receptive field size, can be used to enhance the reflection removal results.

### 3 EXPERIMENTAL RESULTS

We verify the advantages of our dataset (POL) in neural network training and network we proposed. ERRNet [Wei et al. 2019] is re-trained on POL and evaluated on real-world dataset. Table. 1 shows the quantitative measurement improvements by using POL and our enhanced neural network. Moreover, in Fig. 3 we qualitatively compares the results after training with POL and CEIL.

### 4 CONCLUSION AND FUTURE WORK

We propose a polarization-based reflection image synthesis method, which provides adequate and authentic training dataset. Furthermore, we enhance the neural network by introducing the reflection



Figure 3: Qualitative results obtained with CEIL (mid) and POL (right). Left is the input. "©DennisSng(RoseLab)"

Table 1: Quantitative evaluation of proposed dataset and network. "Real", "Object", "SIR<sup>2</sup>" [Wan et al. 2017] are real-world testing datasets, "CEIL" is a synthetic dataset proposed in [Fan et al. 2017]. Better results are in bold.

Index	Dataset	Dataset		Dataset	Network	
		CEIL	POL		ERRNet	Ours
PSNR	Real	21.66	<b>22.23</b>	Real	21.66	<b>21.79</b>
SSIM		0.79	<b>0.80</b>		0.79	<b>0.79</b>
PSNR	Object	23.84	<b>25.13</b>	SIR <sup>2</sup>	23.38	<b>23.54</b>
SSIM		0.88	<b>0.89</b>		0.88	<b>0.88</b>
PSNR	Average	23.64	<b>24.87</b>	Average	23.31	<b>23.47</b>
SSIM		0.87	<b>0.88</b>		0.88	<b>0.88</b>

information as guidance and utilizing adaptive convolution kernel to fuse multi-scale information. The experiments demonstrate the proposed approach achieves convincing improvements over state of the arts. In the future, we plan to explore new data acquisition methods as well as the neural network architecture for further improving the performance of single image reflection removal.

### ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (2019YFB2204104 and 2018YFB2100602). (Portions of) the research in this paper used the 'SIR<sup>2</sup>' Dataset made available by the ROSE Lab at the Nanyang Technological University, Singapore.

### REFERENCES

- Qingnan Fan, Jiaolong Yang, Gang Hua, Baoquan Chen, and David Wipf. 2017. A generic deep architecture for single image reflection removal and image smoothing. In *Proc. Int. Conf. on Comp. Visi.* 3238–3247.
- Naejin Kong, Yu-Wing Tai, and Joseph S Shin. 2013. A physically-based approach to reflection separation: from physical modeling to constrained optimization. *IEEE Trans. Patt. Anal. Mach. Inte.* 36, 2 (2013), 209–221.
- Xiang Li, Wenhai Wang, Xiaolin Hu, and Jian Yang. 2019. Selective kernel networks. In *Proc. Conf. on Comp. Visi. and Patt. Reco.* 510–519.
- Renjie Wan, Boxin Shi, Ling-Yu Duan, Ah-Hwee Tan, and Alex C Kot. 2017. Benchmarking single-image reflection removal algorithms. In *Proc. Int. Conf. on Comp. Visi.* 3922–3930.
- Kaixuan Wei, Jiaolong Yang, Ying Fu, David Wipf, and Hua Huang. 2019. Single image reflection removal exploiting misaligned training data and network enhancements. In *Proc. Conf. on Comp. Visi. and Patt. Reco.* 8178–8187.
- Patrick Wieschollek, Orazio Gallo, Jinwei Gu, and Jan Kautz. 2018. Separating reflection and transmission images in the wild. In *Proc. Euro. Conf. on Comp. Visi.* 89–104.