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

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

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).
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 synthetic dataset. Two randomly selected images, \(I_R\) and \(I_P\), 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 polarizer angle \(\phi\), the intensity of each pixel \(I(x)\) can be computed according to [Kong et al. 2013] as follows:

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

\[
\alpha(\theta, \phi) = R_{\perp}(\theta) \sin^2(\phi_B - \phi) + R_{\parallel}(\theta) \cos^2(\phi_B - \phi) \tag{1}
\]

\[
R_{\perp}(\theta, \kappa) = \frac{\sin^2(\theta - \theta_1(\theta, \kappa))}{\sin^2(\theta + \theta_1(\theta, \kappa))} \quad R_{\parallel}(\theta, \kappa) = \frac{\tan^2(\theta - \theta_1(\theta, \kappa))}{\tan^2(\theta + \theta_1(\theta, \kappa))}
\]

where \(x\) is the index of input images. \(\alpha(\theta, \phi)\) represents the mixing coefficient. \(\phi_B\) 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 \(\theta_1(\kappa = \frac{\sin \theta}{n_2}, \theta) = \arcsin \frac{n_1 \sin \theta}{n_2}\).

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
We verify the advantages of our dataset (POL) in neural network which provides adequate and authentic training dataset. Further, we adopt plane reflection layer is unclear or fades away. Instead, we enhance the neural network by introducing the reflection removal results. The archi-

cited image, and reflection image as guidance, to facilitate the reflection/transmission decomposition results.

During the training stage, we feed the neural network with syn-
thesized 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:

$$L_R = \lambda ||R - R_0||^2 + \mu (||\nabla_x R - \nabla_x R_0||_1 + ||\nabla_y R - \nabla_y R_0||_1),$$

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-

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\textsuperscript{2}’ Dataset made available by the ROSE Lab at the Nanyang Technological University, Singapore.

REFERENCES


