Reflection and Noise Separation from Polarized Images via Joint Nonnegative Matrix Factorization and Plug-and-Play Denoising

Maharu Oda and Ryo Matsuoka*
The University of Kitakyushu, Japan
*r-matsuoka@kitakyu-u.ac.jp

***Abstract*—We propose a robust method for separating reflections and noise in polarized images captured through glass surfaces and affected by mixed noise, including impulse and Gaussian components. The method formulates an optimization problem based on an observation matrix constructed from multiple polarization-angle images and employs non-negative matrix factorization (NMF) with regularization terms that enforce smoothness in transmission and reflection components and sparsity in noise. To solve this problem efficiently, we adopt an alternating optimization strategy combined with the plug-and-play alternating direction method of multipliers (PnP-ADMM) framework, enabling the integration of advanced denoising operators. Experiments demonstrate that the proposed method consistently outperforms conventional approaches in both quantitative and qualitative evaluations.**

I. Introduction

With the widespread use of in-vehicle cameras and surveillance systems, images captured through glass surfaces have become increasingly common. Such images often suffer from unwanted reflections and sensor noise, particularly under low-light conditions, severely degrading visual quality and hindering downstream tasks like object detection.

Polarization filters reduce reflections by attenuating light in specific directions. As the optimal angle varies by scene, manual adjustment is needed. Several methods address this by using multiple polarized images taken at different angles.

Existing approaches fall into two main categories: component separation and machine learning-based methods. Component separation techniques, such as ICA [1] and NMF [2], attempt to decompose an image into transmission and reflection components. However, they often struggle with noise separation or yield physically implausible results. Deep learning methods [3]–[6] learn image-to-image mappings from large datasets and have shown impressive performance under clean conditions, but their robustness degrades in the presence of impulse or mixed noise.

To address these limitations, we propose a robust framework for separating transmission, reflection, and noise components in polarized images affected by both impulse and Gaussian noise. Our method formulates a regularized NMF-based optimization problem and solves it using the Plug-and-Play ADMM (PnP-ADMM) [7]–[11], which integrates an edge-preserving denoiser as a smoothness prior. This approach enables accurate decomposition while maintaining structural details in the transmission component.

Compared to prior work [2] that directly embeds total variation (TV) [12] regularization into the NMF framework, our decoupled PnP-ADMM formulation improves stability and flexibility, especially under challenging noise conditions. Experiments on polarized images with synthetic noise confirm that our method outperforms conventional techniques in both quantitative and qualitative evaluations.

The remainder of this paper is organized as follows. Section II reviews the mathematical background, including NMF and the PnP-ADMM framework. Section III describes the proposed reflection and noise separation method, which integrates NMF with sparsity and smoothness regularization, and solves the resulting optimization problem using PnP-ADMM. Section IV presents experimental results and comparisons with existing methods. Finally, Section V summarizes the findings and concludes the paper.

II. Preliminaries

Throughout this paper, bold lowercase and uppercase letters denote vectors and matrices, respectively. The N-dimensional real-valued vector space is denoted by $\mathbb{R}^N$, and the space of real-valued $N \times M$ matrices is denoted by $\mathbb{R}^{N\times M}$.

A. Polarization Imaging Camera

In conventional polarization photography, a polarizing filter is mounted in front of the camera lens, and multiple images are captured by manually rotating the filter to different angles. In contrast, the polarization camera used in this study (FLIR BFS-U3-51S5P-C) is equipped with a Sony IMX250 polarization image sensor, enabling the simultaneous acquisition of four images at different polarization angles in a single shot.

The sensor achieves this by spatially arranging linear polarizing filters at $0°, 45°, 90°$, and $135°$ directly on the sensor surface in a repetitive lattice pattern. As a result, the captured images do not maintain a pixel-wise correspondence across polarization angles, and interpolation is required after image acquisition.

This structure eliminates the need for mechanical switching of polarization angles and allows stable acquisition of multiple polarization images while minimizing the effects of positional

979-8-3315-7206-8/25/$31.00

misalignment and temporal variation during shooting. Moreover, the four polarization-direction images allow for straightforward calculation of Stokes parameters, which quantitatively describe the physical polarization characteristics.

B. Non-negative Matrix Factorization (NMF)

NMF [2], [13] is a multivariate analysis technique that factorizes an observation matrix with all non-negative elements into two low-rank matrices, both constrained to be non-negative. Given an observation matrix $\mathbf{X} \in \mathbb{R}^{3N\times M}$, NMF approximates it as the product of two non-negative low-rank matrices:

$$\mathbf{X} \approx \mathbf{WH}, \tag{1}$$

where $\mathbf{W} \in \mathbb{R}^{3N\times K}$ is the basis matrix, $\mathbf{H} \in \mathbb{R}^{K\times M}$ is the coefficient matrix, and K denotes the number of bases.

This decomposition is achieved by minimizing the following objective function:

$$\min_{\mathbf{W},\mathbf{H}} \frac{1}{2}\|\mathbf{X}-\mathbf{WH}\|_F^2 \quad \text{s.t.} \quad W_{i,j}\geq 0, \quad H_{i,j}\geq 0, \quad \forall i,j \tag{2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. This optimization problem is typically solved using multiplicative update rules, gradient descent, or the Alternating Least Squares (ALS) method.

C. Plug-and-Play Alternating Direction Method of Multipliers (PnP-ADMM)

PnP-ADMM is a flexible framework that replaces the proximal operator in conventional ADMM [14], [15] with an external denoising algorithm.

ADMM is a proximal splitting algorithm that can treat convex optimization problems of the form

$$\min_{\mathbf{x}\in\mathbb{R}^{N_1},\,\mathbf{z}\in\mathbb{R}^{N_2}} F(\mathbf{x})+G(\mathbf{z}) \;\text{ s.t. }\; \mathbf{z}=\mathbf{Lx}, \tag{3}$$

where $F(\cdot)$ and $G(\cdot)$ are assumed to be quadratic and proximable functions, respectively, and $\mathbf{L}\in\mathbb{R}^{N_2\times N_1}$ is a matrix with full-column rank. For any $\mathbf{x}^{(0)}\in\mathbb{R}^{N_1}, \mathbf{z}^{(0)}, \mathbf{b}^{(0)}(\in\mathbb{R}^{N_2})$, and $\rho>0$, the ADMM algorithm is given by

$$\begin{cases} \mathbf{x}^{(t+1)} = \arg\min_{\mathbf{x}} \left\{ F(\mathbf{x}) + \dfrac{\rho}{2}\|\mathbf{z}^{(t)}-\mathbf{Lx}-\mathbf{b}^{(t)}\|_2^2 \right\}, \\ \mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \left\{ G(\mathbf{z}) + \dfrac{\rho}{2}\|\mathbf{z}-\mathbf{Lx}^{(t+1)}-\mathbf{b}^{(t)}\|_2^2 \right\}, \\ \qquad = \mathrm{prox}_{1/\rho G}(\mathbf{Lx}^{(t+1)}+\mathbf{b}^{(t)}) \\ \mathbf{b}^{(t+1)} = \mathbf{b}^{(t)} + \mathbf{Lx}^{(t+1)} - \mathbf{z}^{(t+1)}, \end{cases} \tag{4}$$

where the superscript (t) denotes the iteration number. The update of $\mathbf{z}$ reduces to the proximity operator of the function $G(\cdot)$ [1]. The sequence generated by (4) quickly converges to an optimal solution of (3).

[1]The proximity operator [16] is a fundamental tool in proximal splitting methods. Let $\mathbf{x}\in\mathbb{R}^N$ be a given input vector. For any $\gamma>0$, the proximity operator of f over $\mathbb{R}^N$ is defined as

$$\mathrm{prox}_{\gamma f}(\mathbf{x}) := \arg\min_{\mathbf{y}\in\mathbb{R}^N} f(\mathbf{y}) + \frac{1}{2\gamma}\|\mathbf{x}-\mathbf{y}\|^2.$$

In PnP-ADMM [7]–[11], the solution to the sub-problem with respect to $\mathbf{z}$ (assuming $\mathbf{L}$ is the identity matrix) is replaced by an off-the-shelf denoising algorithm. This yields the update:

$$\mathbf{z}^{(t+1)} = \mathcal{D}_\sigma\left(\mathbf{x}^{(t+1)}+\mathbf{b}^{(t)}\right), \tag{5}$$

where $\mathcal{D}_\sigma$ denotes the Gaussian denoiser, and σ is the standard deviation of the assumed Additive white Gaussian noise (AWGN). Auto-encoder-based networks are often used as a denoiser [9].

III. Proposed Method

A. Observation Model

When capturing images through glass using a polarizing filter, the observed image $\mathbf{y}_\theta\in\mathbb{R}^{3N}$ at polarization angle θ can be modeled as:

$$\mathbf{y}_\theta = \alpha_\theta\mathbf{t} + \beta_\theta\mathbf{r} + \mathbf{s}_\theta + \mathbf{n}_\theta, \tag{6}$$

where $\mathbf{t}$ and $\mathbf{r}$ represent the transmission and reflection components, respectively, $\mathbf{s}_\theta$ represents impulse noise, and $\mathbf{n}_\theta$ denotes Gaussian noise. Here, N denotes the number of pixels per channel, and pixel values are normalized to $[0,1]$.

Stacking M images at angles $\theta_1,\ldots,\theta_M$ yields the matrix formulation:

$$\mathbf{Y} = \mathbf{WH} + \mathbf{S} + \mathbf{N}, \tag{7}$$

where $\mathbf{Y} =: [\mathbf{y}_{\theta_1}\cdots\mathbf{y}_{\theta_M}] \in \mathbb{R}^{3N\times M}$ is the observation matrix, $\mathbf{W}=[\mathbf{t}\ \mathbf{r}]\in\mathbb{R}^{3N\times 2}$ is the basis matrix, $\mathbf{H}\in\mathbb{R}^{2\times M}$ contains the combination coefficients, $\mathbf{S}=[\mathbf{s}_{\theta_1}\cdots\mathbf{s}_{\theta_M}]$ represents the sparse impulse noise, $\mathbf{N}=[\mathbf{n}_{\theta_1}\cdots\mathbf{n}_{\theta_M}]$ represents the Gaussian noise.

In this model, impulse noise $\mathbf{S}$ is treated as sparse and localized, whereas Gaussian noise $\mathbf{N}$ is assumed to be small and distributed throughout the image.

B. Optimization Problem

We estimate $\mathbf{W}$, $\mathbf{H}$, and $\mathbf{S}$ by solving the following constrained minimization problem

$$\begin{aligned} &\min_{\mathbf{W},\mathbf{H},\mathbf{S}} \ \|\mathbf{Y}-\mathbf{WH}-\mathbf{S}\|_F^2 + \lambda_1 R(\mathbf{W}) + \lambda_2\|\mathbf{S}\|_1 \\ &\text{s.t.} \quad W_{i,j}\geq 0, \ \ H_{i,j}\geq 0, \ \ -1\leq S_{i,j}\leq 1, \ \ \forall i,j, \end{aligned} \tag{8}$$

where $R(\mathbf{W})$ is a regularization term that promotes spatial smoothness in the basis components (transmission and reflection), respectively, implemented using an edge-preserving denoiser within the PnP-ADMM framework. The ℓ_1 norm term enforces sparsity in the impulse noise component $\mathbf{S}$. The parameters λ_1 and λ_2 (>0) control the trade-offs among data fidelity, smoothness, and sparsity. In this formulation, the residual $\mathbf{Y}-\mathbf{WH}-\mathbf{S}$ is expected to absorb the remaining small Gaussian noise $\mathbf{N}$, assuming it follows a low-energy distribution.

To solve (8), we adopt an alternating optimization strategy.

- Update $\mathbf{H}$ using the ADMM under non-negative constraints.
- Jointly update $\mathbf{W}$ and $\mathbf{S}$ by solving a subproblem via PnP-ADMM, where an edge-preserving smoothing denoiser is

TABLE I
COMPARISON OF QUANTITATIVE EVALUATION

	PSNR				SSIM				SI			
Method \ Scene	1	2	3	4	1	2	3	4	1	2	3	4
Input image (90°)	23.3822	21.1409	23.1202	21.7895	0.7320	0.8062	0.6926	0.7318	0.7436	0.8713	0.8235	0.7543
Denoise+NMF	27.0938	27.9459	34.8642	32.6095	0.8673	0.9409	0.8741	0.9291	0.8790	0.9661	0.9588	0.9400
Ours-BLF	28.4158	30.1384	35.2188	33.5487	0.9299	0.9654	0.9062	0.9627	0.9405	0.9829	0.9696	0.9662
Ours-TV	**28.8483**	**30.2404**	**35.9491**	**34.4383**	**0.9408**	**0.9671**	**0.9074**	**0.9733**	**0.9505**	**0.9842**	**0.9765**	**0.9757**

applied to $\mathbf{W}$ under non-negative constraints, and soft-thresholding is applied to $\mathbf{S}$ under bounded constraints to enforce sparsity.

Through these steps, robust separation of transmission, reflection, and impulse noise components is achieved.

NMF algorithms typically initialize the variables $\mathbf{W}$ and $\mathbf{H}$ with random non-negative values. However, such initialization can lead to instability and slow convergence.

To address this, we explicitly initialize the basis matrix $\mathbf{W}$ as follows. Each polarization image is first denoised using bilateral and median filters to suppress noise and enhance spatial smoothness. The denoised images are then vectorized and stacked to form a matrix, from which two dominant non-negative components are extracted via NMF. Since the transmission component generally has a higher average intensity than the reflection, we use the mean intensity of each extracted basis as a heuristic to identify the transmission component.

The variable $\mathbf{H}$ is initialized with random non-negative values, and the sparse noise matrix $\mathbf{S}$ is initialized to zero.

IV. Experimental Results

To evaluate the effectiveness of the proposed method, we performed controlled experiments using polarized images captured through a glass plate. A FLIR BFS-U3-51S5P-C polarization camera was used to acquire four polarization-angle images in a single shot. To simulate real-world degradations, salt-and-pepper noise (density = 0.002) and additive Gaussian noise (mean = 0, variance = 10) were synthetically added to each image. We compared the proposed method with a conventional NMF-based approach that includes denoising as a preprocessing step. For quantitative evaluation, ground-truth transmission images were obtained by capturing the same scenes without the glass plate, thus eliminating reflection components.

The estimated transmission components were quantitatively evaluated using PSNR, SSIM [17], and SI [18]. PSNR measures pixel-wise differences between the restored and reference images, with higher values indicating better reconstruction accuracy. SSIM assesses structural similarity in terms of local structure, luminance, and contrast, reflecting perceptual image quality. SI evaluates image sharpness based on edge strength and is useful for assessing the preservation of fine details. A Gaussian filter of size 11 was used for both SSIM and SI.

In all experiments, the parameter ρ was fixed at $\rho = 1$. The regularization parameters λ_1 and λ_2 were manually adjusted for each scene to obtain the most visually pleasing results in terms of reflection suppression. In the proposed method, we evaluate two variants based on the choice of denoiser used in the PnP-ADMM framework: (1) Ours-BLF, which employs a bilateral filter for edge-preserving smoothing, and (2) Ours-TV, which incorporates TV regularization as a prior [12], [19].

Table I summarizes the quantitative evaluation results for the input images, the conventional method [2], and the proposed method. Across all scenes and evaluation metrics, the proposed method consistently outperforms the conventional approach. Although Ours-BLF shows slightly lower quantitative scores than Ours-TV, it preserves finer text edge structures more effectively in qualitative evaluations.

Figures 1 and 2 present the reference images without reflections, the input images at a polarization angle of $90°$, the results of the conventional method with preprocessing, and those of the proposed variants. While the conventional method reduces some reflections, it tends to amplify noise and fails to preserve fine details such as text edges (e.g., Fig. 1(d)). In contrast, the proposed method achieves better reflection suppression while retaining structural details, owing to the denoising-based regularization term $R(\mathbf{W})$ incorporated via PnP-ADMM (see Figs. 1(e), (f) and 2(e), (f)).

V. Conclusion

This paper proposed a method for separating transmission, reflection, and noise components from polarized images affected by impulse noise. The approach extends conventional NMF-based reflection removal by introducing ℓ_1 regularization for sparse anomalies such as impulse noise and specular reflections, and incorporating the PnP-ADMM framework with an edge-preserving denoiser to ensure structural smoothness. Experimental results on synthetically degraded images demonstrated that the method consistently outperforms conventional approaches in both visual quality and quantitative accuracy.

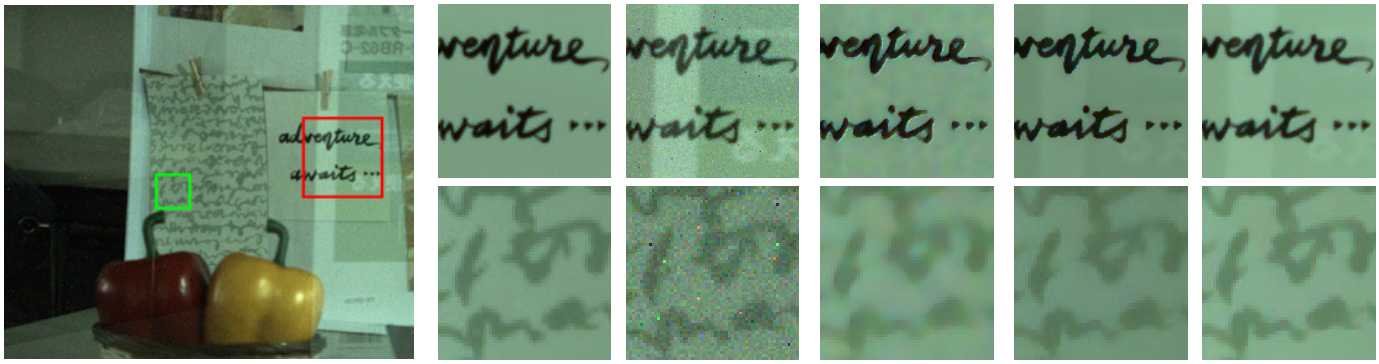
Future work will focus on evaluating the method under outdoor and real-world conditions with significant noise, improving robustness through polarization images captured under varied exposure settings, and addressing dynamic range limitations such as underexposure and glare.

Acknowledgment

This work was partially supported by JSPS KAKENHI Grant Number 24K15075.

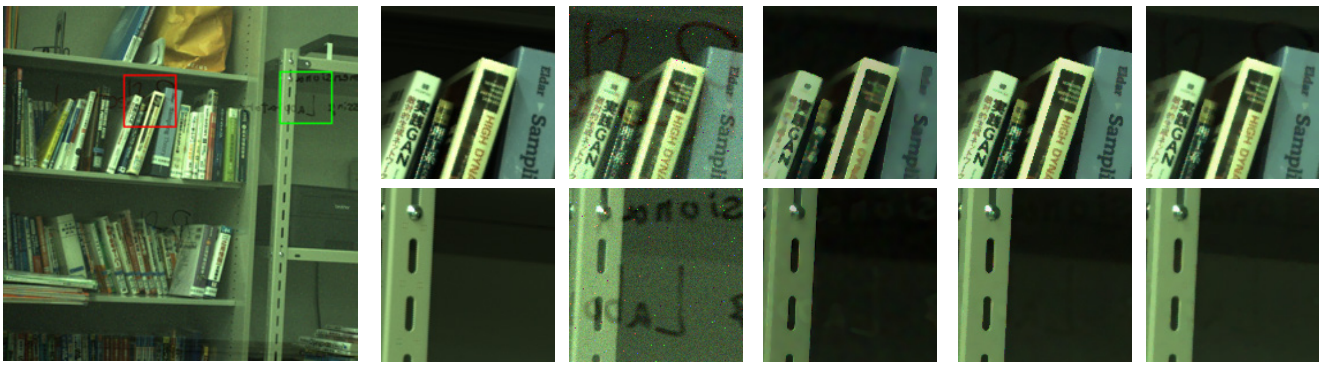
References

[1] H. Farid and E. H. Adelson, "Separating reflections from images by use of independent component analysis," *Journal of the Optical Society of America A*, vol. 16, no. 9, pp. 2136–2145, 1999.



(a) Input image ($\theta = 90^\circ$) (b) Reference (c) Input (d) Denoise+NMF (e) Ours-BLF (f) Ours-TV

Fig. 1. Comparison of Reflection Removal Results in Scene 1



(a) Input image ($\theta = 90^\circ$) (b) Reference (c) Input (d) Denoise+NMF (e) Ours-BLF (f) Ours-TV

Fig. 2. Comparison of Reflection Removal Results in Scene 2

[2] T. Aizu and R. Matsuoka, "Reflection removal using multiple polarized images with different exposure times," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, IEEE, 2022, pp. 498–502.

[3] R. Wan, B. Shi, L.-Y. Duan, A.-H. Tan, and A. C. Kot, "Crrn: Multi-scale guided concurrent reflection removal network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 4777–4785.

[4] K. Wei, J. Yang, Y. Fu, D. Wipf, and H. Huang, "Single image reflection removal exploiting misaligned training data and network enhancements," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 8178–8187.

[5] C. Lei, X. Huang, M. Zhang, Q. Yan, W. Sun, and Q. Chen, "Polarized reflection removal with perfect alignment in the wild," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 1750–1758.

[6] Y. Zhu, X. Fu, P.-T. Jiang, *et al.*, "Revisiting single image reflection removal in the wild," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024, pp. 25 468–25 478.

[7] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. IEEE Glob. Conf. Signal Inf. Process. (GlobalSIP)*, IEEE, 2013, pp. 945–948.

[8] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play admm for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 84–98, 2016.

[9] Z. Meng, X. Yuan, and S. Jalali, "Deep unfolding for snapshot compressive imaging," *Int. J. Comput. Vis.*, vol. 131, no. 11, pp. 2933–2958, 2023.

[10] H. Yamamoto, S. Anami, and R. Matsuoka, "Optimizing dynamic mode decomposition for video denoising via plug-and-play alternating direction method of multipliers," *Signals*, vol. 5, no. 2, pp. 202–215, 2024, ISSN: 2624-6120. DOI: 10.3390/signals5020011.

[11] A. Emoto and R. Matsuoka, "Unsupervised anomaly detection in hyperspectral imaging: Integrating tensor robust principal component analysis with autoencoding adversarial networks," *IEEE Access*, vol. 13, pp. 21 422–21 433, 2025. DOI: 10.1109/ACCESS.2025.3534981.

[12] P. L. Combettes and J.-C. Pesquet, "A proximal decomposition method for solving convex variational inverse problems," *Inverse Probl.*, vol. 24, no. 6, p. 065 014, 2008.

[13] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Comput. Stat. Data Anal.*, vol. 52, no. 1, pp. 155–173, 2007.

[14] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Comput. Math. Appl.*, vol. 2, no. 1, pp. 17–40, 1976.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011, ISSN: 1935-8237. DOI: 10.1561/2200000016.

[16] J. J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," *C. R. Acad. Sci. Paris*, vol. 255, pp. 2897–2899, 1962. [Online]. Available: https://hal.science/hal-01867195.

[17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[18] R. Wan, B. Shi, L.-Y. Duan, A.-H. Tan, and A. C. Kot, "Benchmarking single-image reflection removal algorithms," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 3922–3930.

[19] R. Matsuoka, S. Ono, and M. Okuda, "Transformed-domain robust multiple-exposure blending with huber loss," *IEEE Access*, vol. 7, pp. 162 282–162 296, 2019. DOI: 10.1109/ACCESS.2019.2951817.