Proceedings of 13th IEEE International Conference
on CYBER Technology in Automation, Control, and Intelligent Systems
July 11 - 14, 2023, Qinhuangdao, China

2023 IEEE 13th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER) | 979-8-3503-1519-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/CYBER59472.2023.10256552

A Continuous Illumination Estimator for Shadow Removal

Jiawei Liu, Huijie Fan*, Yiming Su, Zhen Yan, Qiang Wang, Wentao Li, and Yandong Tang

***Abstract*— Shadow removal methods often encounter difficulties in completely removing shadows or avoiding excessively bright removal results in shadow areas. To address this challenge, we propose a continuous-intensity shadow synthesis model and an illumination estimator for shadow removal. Our shadow synthesis model generates visually realistic images with continuously varying shadow intensity to augment the current shadow removal dataset. By utilizing the synthesized shadow images, we train a network to accurately estimate the illumination intensity in shadow areas. The estimated illumination intensity is positive for a shadow input image and negative for an over-lit input image, thus our method can illuminate shadow areas or suppress over-lit areas. Experimental results demonstrate that our method outperforms current state-of-the-art shadow removal methods, particularly when handling shadows with a wide range of intensity variations.**

I. Introduction and Background

Shadows are a common illumination phenomenon in natural scenes, which can cause computer vision methods to fail or perform poorly, e.g., object detection, semantic segmentation, and tracking [1]–[9]. Shadow removal, as a pre-processing step, can improve the robustness of these tasks to illumination variations.

This work is supported by the National Natural Science Foundation of China (62273339, 61991413, 62073205, U20A20200), and the Youth Innovation Promotion Association of Chinese Academy of Sciences(2019203).

Huijie Fan is the corresponding author.

J. Liu, Y. Su, and Z. Yan are with State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, 110169, China; University of Chinese Academy of Sciences, Beijing, 100049, China (email: liujiawei@sia.cn).

Q. Wang is with the Key Laboratory of Manufacturing Industrial Integrated, Shenyang University, Shenyang, 110044, China; the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China (email: wangqiang@sia.cn).

H. Fan, W. Li and Y. Tang are with State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, 110169, China (email: {fanhuijie,liwentao,ytang}@sia.cn).

Current shadow removal methods mainly consist of traditional methods [10]–[14] based on physical modeling and deep learning-based shadow removal networks [15]–[20]. The goal of these methods is to recover the illumination of shadow regions, making them visually consistent with non-shadow areas. However, accurately estimating the illumination intensity of shadow regions is challenging, leading to low or high estimated illumination intensities. This often results in visually residual shadows or over-illumination of shadow areas.

In this paper, we propose a continuous-intensity shadow synthesis model and an illumination estimator for shadow removal. We use shadow and shadow-free image pairs to synthesize realistic shadow images with continuous illumination intensities, even over-lit images. Our shadow synthesis method significantly extends the current pairwise shadow removal dataset by incorporating a wider range of illumination variations. We employ a convolutional network to learn the illumination intensity of shadow regions from the synthesized shadow images for shadow removal. The estimated illumination varies with the input and is used to compensate or suppress inconsistent illumination regions. In fact, acquiring pairwise shadow images is time-consuming, and obtaining shadow images with continuous illumination variations is more challenging. Thus, the proposed method offers a solution to address complex shadow intensity variations.

II. Method

A. Continuous-intensity Shadow Synthesis Model

The current popular shadow synthesis method [17], [21], [22] involves two steps that require a shadow-free image (I_f), a soft shadow mask (β) and parameters (k,b). In the first step, the shadow intensity is modified by adjusting the illumination parameters, i.e., $I_{dark} = (I_f - b) \cdot k^{-1}$. In the second step, a new shadow image (I_{sys}) is synthesized using the formula $I_{sys} = I_f \cdot \beta + I_{dark} \cdot (1-\beta)$. The proposed method eliminates the need for creating soft masks (β) and calculating parameter (k,b) ranges from the shadow removal dataset.

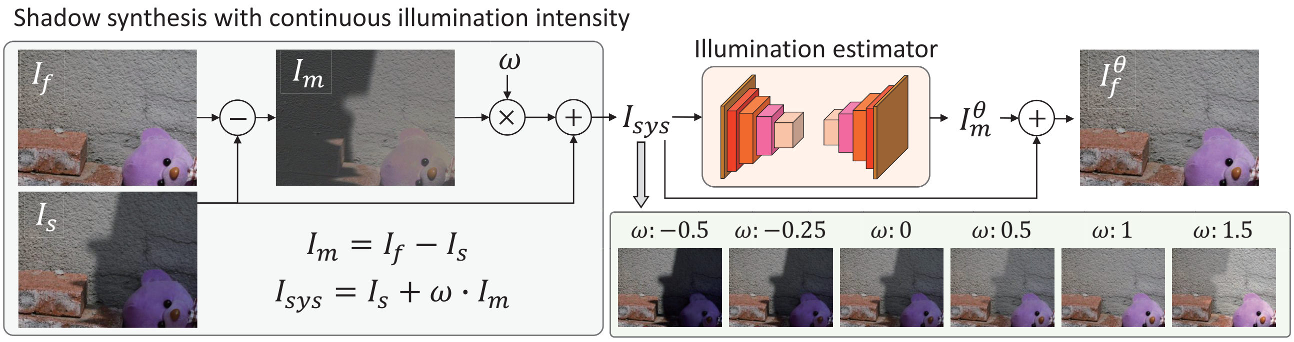

Fig. 1. Overview of our continuous illumination estimator (CIE) for shadow removal. Initially, we obtain a shadow matte image (I_m) that represents the intensity and location of the shadows, utilizing a pair of shadow (I_s) and shadow-free (I_f) images. Subsequently, we use a factor (ω) to control the relative shadow effect, which allows the synthetic image (I_{sys}) to have shallower or darker shadow regions compared to the original shadow image (I_s). Finally, we train an illumination estimator using the synthetic shadow image to learn how to balance inconsistent illumination between the shadow and non-shadow regions. The relationship between the ground truth of the estimated illumination (I_m^θ) and the original shadow matte (I_m) is proportional, i.e., $I_m^\theta \to (1-\omega)\cdot I_m$.

Specifically, we utilize a pair of shadow (I_s) and shadow-free images (I_f) to obtain a shadow matte image (I_m) that represents the location and intensity of the shadows, as shown in Fig. 1. A synthetic shadow image (I_{sys}) can be generated by the following formula,

$$I_{sys} = I_s + \omega \cdot I_m, \tag{1}$$

where I_m is the residual image between I_f and I_s (i.e., $I_m = I_f - I_s$), and the parameter ω can be used to control the relative shadow effect, allowing the synthetic image to have darker ($\omega < 0$), shallower ($0 < \omega < 1$) or "over-lit" ($1 < \omega$) shadows compared to those in the original shadow image (I_s). In a special case, when $\omega = 0$, the synthetic image (I_{sys}) is identical to the shadow image (I_s), and when $\omega = 1$, I_{sys} becomes the shadow-free image (I_f). Our proposed single-step approach not only simplifies the shadow synthesis process but also produces a wider range of illumination variations (see Fig. 1), potentially aiding in the development of more effective shadow removal techniques to address challenging shadow intensity variations.

B. Illumination Estimator

We can generate synthetic images with varying shadow intensities using Eq. 1. Our goal is to train a continuous illumination estimator (CIE) that can learn to handle local illumination variations from these synthetic images. Specifically, given a synthetic image (I_{sys}) as input, the network (G) should output a shadow matte image (I_m^θ) that represents illumination variations.

$$I_m^\theta = G(I_{sys}), \tag{2}$$

We can obtain a shadow-free image (I_f^θ) by using the following equation.

$$I_f^\theta = I_m^\theta + I_{sys}. \tag{3}$$

In fact, the estimated shadow matte (I_m^θ) is close to $(1-\omega)\cdot I_m$, representing the location and intensity of the shadows in the synthetic image.

$$I_m^\theta \to (1-\omega)\cdot I_m. \tag{4}$$

This means that our illumination estimator learns a changing shadow matte ($(1-\omega)\cdot I_m$), rather than the current networks [15], [23]–[25] learning a fixed shadow matte (I_m).

C. Loss Functions

To recover a high quality shadow-free image without artifacts, we employ multiple loss functions to train our illumination estimator, including ℓ_1 loss, adversarial loss [18], [26], and perceptual loss L_{perc} [27].

$$L_f = \lambda_1 L_{\ell_1} + \lambda_2 L_{adv} + \lambda_3 L_{perc}. \tag{5}$$

The ℓ_1 loss is defined as

$$L_{\ell_1} = \mathbb{E}_{I_f^\theta, I_f}\left[\left\| I_f^\theta - I_f \right\|_1\right]. \tag{6}$$

The adversarial loss is

$$L_{adv} = \mathbb{E}_{\tilde{I}_f, I_s}\left[\log\left[1 - D(\tilde{I}_f, I_s)\right]\right], \tag{7}$$

where D is a discriminator. The perceptual loss is computed by

$$L_{perc} = \sum_{k=1}^{5} \mathbb{E}_{\tilde{I}_f, I_f}\left[\left\| \Phi_k(I_f^\theta) - \Phi_k(I_f) \right\|_1\right], \tag{8}$$

where Φ_k is the activation map extracted from the k-th layer of the pretrained VGG19 [28].

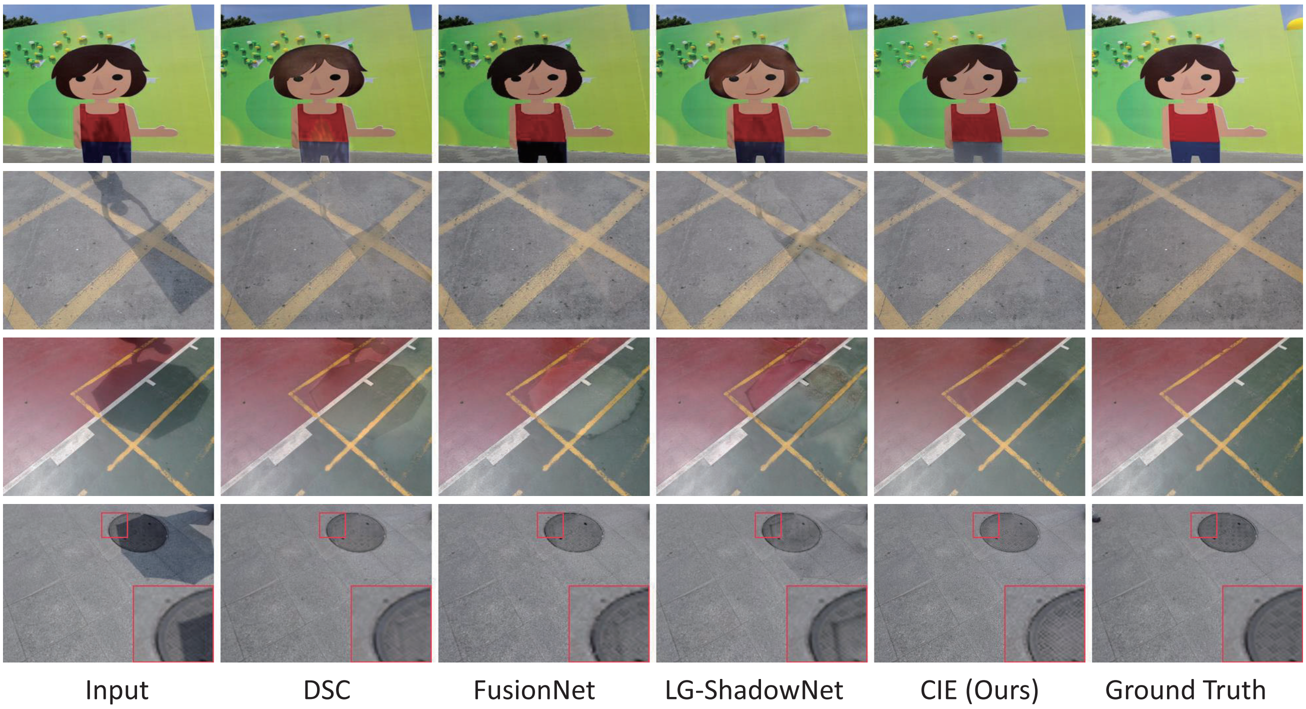

Fig. 2. Visual comparison results on the ISTD dataset [16].

III. EXPERIMENTS

A. Datasets and Evaluation Metrics

We assessed the effectiveness of our proposed approach on the ISTD dataset [16], a widely used benchmark for shadow removal. ISTD [16] is a large-scale dataset that includes two tasks, namely shadow detection and shadow removal. The dataset is composed of 1,330 training and 540 testing triplets, consisting of a shadow image, a shadow mask, and a corresponding shadow-free image. The images in the ISTD dataset have a resolution of 640×480 pixels. To evaluate the quantitative results of shadow removal, we utilized the Mean Absolute Error (MAE) as the evaluation metric.

B. Network Architecture and Implementation Details

We adapted the structure of the existing shadow removal network, DMTN [25], as the illumination estimator in our approach. As we do not require multi-task learning, we excluded the multi-task decoupling layer from the DMTN. To estimate the illumination in shadow regions, we generate shadow-free images using Eq. 3, while the network focuses on learning the shadow matte I_m^{θ}. To handle under or over lighting during shadow synthesis using Eq. 1, we set the value of ω within the range of $[-1,2)$. Empirically, we set the hyperparameters (in Eq. 1) as $\lambda_1 = 2550$, $\lambda_2 = 1$, and $\lambda_3 = 1000$.

C. Comparison with the State-of-the-art Methods

We present a comparative analysis of our proposed CIE with several state-of-the-art methods, including Yang *et al.* [29], Guo *et al.* [30], Gong *et al.* [31], ST-CGAN [16], DSC [32], RIS-GAN [23], DHAN [18], LG-ShadowNet [33], FusionNet [34], and CANet [35]. Table I demonstrates that our proposed method outperforms the compared methods on the ISTD dataset, achieving a 5.6% reduction in MAE for the whole image and an 8.0% reduction in MAE for the non-shadow areas. Additionally, as shown in Fig. 2, our proposed method successfully recovers the illumination of shadow regions with high accuracy.

We show the results of removing shadows with continuous intensity variations in Fig. 3. Compared to DHAN [18], our method is able to cope with both dark and light shadows, and even correct over-lit areas.

IV. CONCLUSION

In this paper, we propose a continuous-intensity shadow synthesis model and a continuous illumination estimator (CIE). The synthesis model can generate images with varying shadow intensity, which augments the current shadow removal dataset and allows for the synthesis of over-lit images. Our CIE accurately estimates the illumination intensity of shadow regions for effective shadow removal and can handle

TABLE I

QUANTITATIVE COMPARISON RESULTS ON THE **ISTD** DATASET [16]. WE REPORT THE MEAN ABSOLUTE ERROR (MAE) IN THE SHADOW AREA, NON-SHADOW AREA, AND THE WHOLE IMAGE (ALL). THE RESULTS MARKED WITH $*$ AND ¶ ARE REPORTED BY [16] AND [18], RESPECTIVELY. THE BEST AND SECOND-BEST RESULTS ARE **HIGHLIGHTED** AND UNDERLINED, RESPECTIVELY. THE SYMBOL $\downarrow$ INDICATES THAT SMALLER VALUES ARE BETTER.

Method	Shadow	Non-Shadow	ALL
Original	32.67	6.83	10.97
Yang [29] $*$	19.82	14.83	15.63
Guo [30] $*$	18.95	7.46	9.30
Gong [31] $*$	14.98	7.29	8.53
ST-CGAN [16]	10.33	6.93	7.47
DSC [32] ¶	9.48	6.14	6.67
RIS-GAN [23]	8.99	6.33	6.95
DHAN [18]	<u>8.14</u>	6.04	6.37
LG-ShadowNet [33]	10.23	<u>5.38</u>	6.18
FusionNet [34]	**7.77**	5.56	<u>5.92</u>
CANet [35]	8.86	6.07	6.15
Ours (CIE)	8.99	**4.95**	**5.59**

images with both shadows and over-lit regions. Our experimental results demonstrate that the proposed method outperforms current state-of-the-art methods in removing shadows and correcting over-lit regions, particularly when the shadows exhibit wide intensity variations. This approach provides a promising solution to the challenging task of continuous-intensity shadow removal and has the potential to be applied to other related computer vision tasks. Overall, our proposed method represents a significant step forward in the field of shadow removal and has practical implications in various real-world applications.

REFERENCES

[1] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE PAMI*, vol. 25, no. 10, pp. 1337–1342, 2003.

[2] S. Nadimi and B. Bhanu, "Physical models for moving shadow and object detection in video," *IEEE PAMI*, vol. 26, no. 8, pp. 1079–1087, 2004.

[3] C. R. Jung, "Efficient background subtraction and shadow removal for monochromatic video sequences," *IEEE Transactions on Multimedia*, vol. 11, no. 3, pp. 571–577, 2009.

[4] S.-C. Huang and B.-H. Chen, "Highly accurate moving object detection in variable bit rate video-based traffic monitoring systems," *IEEE TNNLS*, vol. 24, no. 12, pp. 1920–1931, 2013.

[5] S. H. Khan, M. Bennamoun, F. Sohel, and R. Togneri, "Automatic shadow detection and removal from a single image," *IEEE PAMI*, vol. 38, no. 3, pp. 431–446, 2015.

[6] W. Zhang, X. Zhao, J.-M. Morvan, and L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE PAMI*, vol. 41, no. 3, pp. 611–624, 2018.

[7] C. Li, W. Xia, Y. Yan, B. Luo, and J. Tang, "Segmenting objects in day and night: Edge-conditioned cnn for thermal image semantic segmentation," *IEEE TNNLS*, vol. 32, no. 7, pp. 3069–3082, 2020.

[8] Y. Zhong, X. Liu, D. Zhai, J. Jiang, and X. Ji, "Shadows can be dangerous: Stealthy and effective physical-world adversarial attack by natural phenomenon," in *Proc. CVPR*, 2022.

[9] B. Fan, H. Zhang, Y. Cong, Y. Tang, H. Fan, and J. Tian, "Dual aligned siamese dense regression tracker," *IEEE TIP*, vol. 31, pp. 3630–3643, 2022.

[10] E. Arbel and H. Hel-Or, "Texture-preserving shadow removal in color images containing curved surfaces," in *Proc. CVPR*, 2007, pp. 1–8.

[11] Y. Shor and D. Lischinski, "The shadow meets the mask: Pyramid-based shadow removal," in *Computer Graphics Forum*, vol. 27, no. 2, 2008, pp. 577–586.

[12] J. Tian, J. Sun, and Y. Tang, "Tricolor attenuation model for shadow detection," *IEEE TIP*, vol. 18, no. 10, pp. 2355–2363, 2009.

[13] E. Arbel and H. Hel-Or, "Shadow removal using intensity surfaces and texture anchor points," *IEEE PAMI*, vol. 33, no. 6, pp. 1202–1216, 2011.

[14] J. Tian and Y. Tang, "Linearity of each channel pixel values from a surface in and out of shadows and its applications," in *Proc. CVPR*, 2011, pp. 985–992.

[15] L. Qu, J. Tian, S. He, Y. Tang, and R. W. Lau, "Deshadownet: A multi-context embedding deep network for shadow removal," in *Proc. CVPR*, 2017, pp. 4067–4075.

[16] J. Wang, X. Li, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," in *Proc. CVPR*, 2018, pp. 1788–1797.

[17] H. Le and D. Samaras, "Shadow removal via shadow image decomposition," in *Proc. ICCV*, 2019, pp. 8578–8587.

[18] X. Cun, C.-M. Pun, and C. Shi, "Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting gan." in *AAAI*, 2020, pp. 10 680–10 687.

[19] Y. Zhu, Z. Xiao, Y. Fang, X. Fu, Z. Xiong, and Z.-J. Zha, "Efficient model-driven network for shadow removal," in *Proc. AAAI*, 2022.

[20] Y. Zhu, J. Huang, X. Fu, F. Zhao, Q. Sun, and Z.-J. Zha, "Bijective mapping network for shadow removal," in *Proc. CVPR*, 2022, pp. 5627–5636.

[21] N. Inoue and T. Yamasaki, "Learning from synthetic shadows for shadow detection and removal," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2020.

[22] Y. Hong, L. Niu, and J. Zhang, "Shadow generation for composite image in real-world scenes," in *Proc. AAAI*, 2022.

[23] L. Zhang, C. Long, X. Zhang, and C. Xiao, "Ris-gan: Explore residual and illumination with generative adversarial networks for shadow removal," in *Proc. AAAI*, vol. 34, no. 07, 2020, pp. 12 829–12 836.

[24] Y. Jin, A. Sharma, and R. T. Tan, "Dc-shadownet: Single-image hard and soft shadow removal using unsupervised domain-classifier guided network," in *Proc. ICCV*, 2021, pp. 5027–5036.

[25] J. Liu, Q. Wang, H. Fan, W. Li, L. Qu, and Y. Tang, "A decoupled multi-task network for shadow removal," *IEEE TMM*, pp. 1–14, 2023.

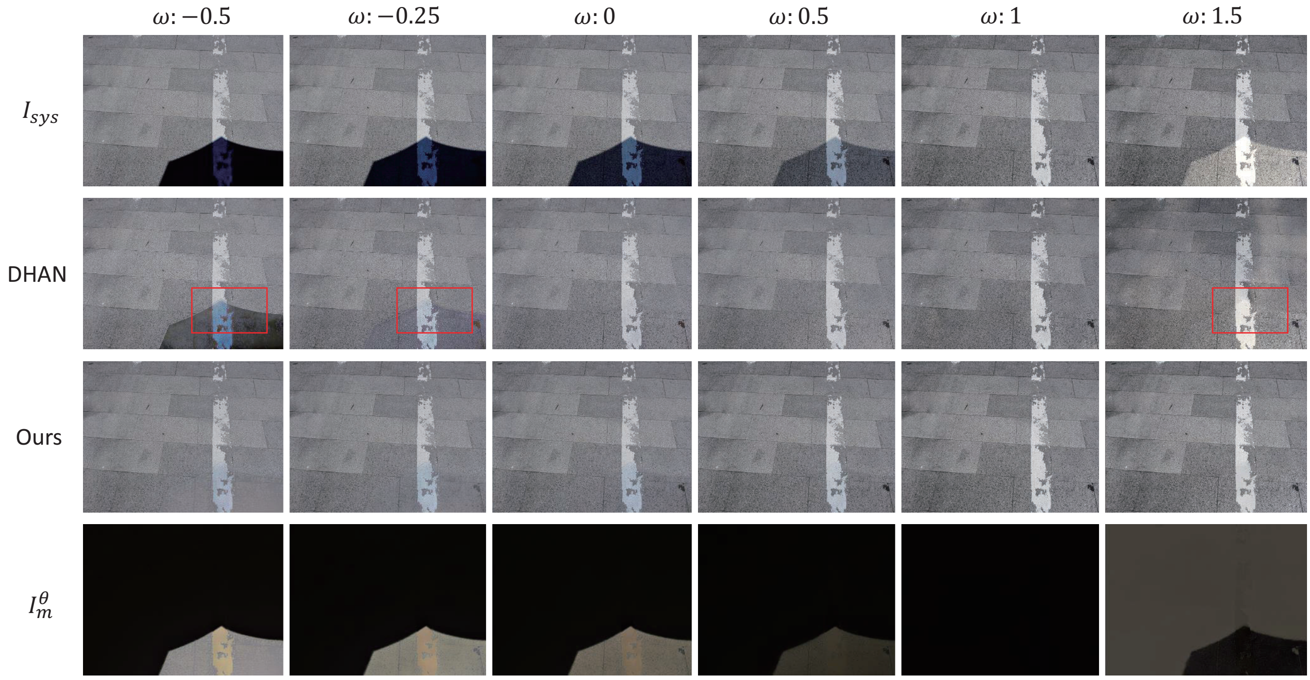

Fig. 3. Visual results of continuous intensity shadow removal. This result demonstrates the ability of our method to estimate the continuously varying illumination (I_m^θ) in shadow areas, which allows for effective and realistic removal of shadows with varying intensities.

[26] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. CVPR*, 2017, pp. 1125–1134.

[27] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. ECCV*, 2016, pp. 694–711.

[28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. ICLR*, Y. Bengio and Y. LeCun, Eds., 2015.

[29] Q. Yang, K.-H. Tan, and N. Ahuja, "Shadow removal using bilateral filtering," *IEEE TIP*, vol. 21, no. 10, pp. 4361–4368, 2012.

[30] R. Guo, Q. Dai, and D. Hoiem, "Paired regions for shadow detection and removal," *IEEE PAMI*, vol. 35, no. 12, pp. 2956–2967, 2012.

[31] H. Gong and D. Cosker, "Interactive shadow removal and ground truth for variable scene categories," in *British Machine Vision Conference*, M. F. Valstar, A. P. French, and T. P. Pridmore, Eds., 2014.

[32] X. Hu, C. W. Fu, L. Zhu, J. Qin, and P. A. Heng, "Direction-aware spatial context features for shadow detection and removal," *IEEE PAMI*, vol. 42, no. 11, pp. 2795–2808, 2020.

[33] Z. Liu, H. Yin, Y. Mi, M. Pu, and S. Wang, "Shadow removal by a lightness-guided network with training on unpaired data," *IEEE TIP*, vol. 30, pp. 1853–1865, 2021.

[34] L. Fu, C. Zhou, Q. Guo, F. Juefei-Xu, H. Yu, W. Feng, Y. Liu, and S. Wang, "Auto-exposure fusion for single-image shadow removal," in *Proc. CVPR*, 2021, pp. 10 571–10 580.

[35] Z. Chen, C. Long, L. Zhang, and C. Xiao, "Canet: A context-aware network for shadow removal," in *Proc. ICCV*, 2021, pp. 4743–4752.