

Updated Paired Regions for Shadow Detection from Single Image

Xiao Wang¹³
wangxiao@iie.ac.cn
Siyuan Yao¹³
yaosiyuan@iie.ac.cn
Pengwen Dai¹³
daipengwen@iie.ac.cn
Rui Wang¹²³
wangrui@iie.ac.cn
Xiaochun Cao¹³
caoxiaochun@iie.ac.cn

¹ State Key Laboratory of Information Security
Institute of Information Engineering,
Chinese Academy of Sciences
Beijing, China
² Zhejiang Lab
Hangzhou, China
³ School of Cyber Security
University of Chinese Academy of
Sciences
Beijing, China

Abstract

In recent years, deep neural network based shadow detection approaches achieve high performance on benchmarks, but they require large amounts of labeled data for training to learn the statistical attributes of the shadowed and unshadowed regions. The physical relationship between non-grid regions is ignored due to the character of deep convolution network. In this paper, we seek to analyze the physical principle of the shadowed region and present a paired region based algorithm for shadow detection without extra model training. To be specific, we first segment the image via a region growing based approach to maintain the character of shadowed region. Then the penumbra is detected and used as an important cue. After that we adopt three physics based confidence coefficients when comparing the color of two regions. At last, we design a novel objective function to find the best paired strategy and detect the shadowed region. The proposed approach is tested on three public datasets. The comparative results show that the proposed non-learning approach performs favorably against the state-of-the-art approaches.

1 Introduction

Shadow detection is a fundamental task in the computer vision community. Its result can provide important cues for calculating light direction [14], scene geometry [22] and the parameters of camera location [12] from image. It can also benefit some high-level computer vision tasks such as image segmentation [5], object detection [4], visual tracking [20] and so on. Despite the importance and long researched history of shadow detection, it is still a challenging and ill-posed task, particularly when only a single image is available.

In the past few years, deep neural network based approaches are prevalent in shadow detection [29] [11] [32]. These approaches train a deep neural network to learn the feature

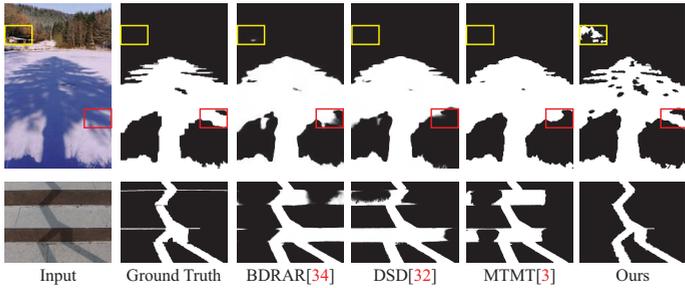


Figure 1: Some incorrect results of the state-of-the-art deep shadow detection approaches. The first row is caused by penumbra region and inaccurate ground truth. The second row is because the relationship between non-grid regions is not clearly explored.

representation of shadowed and unshadowed regions under the guidance of the ground truth data. Although these detectors have achieved remarkable performance on the benchmarks, the training process requires large amounts of expensive and time-consuming labeled data. Besides, state-of-the-art deep approaches may fail to deal with pictures that have lots of penumbra regions precisely as shown in the first row of Figure 1. Furthermore, the deep neural network based approaches adopt the convolution operation to generate convolutional features, only the local region around each pixel is involved. The relationship between regions in non-grid structure is not clearly analyzed, which may hamper the final detection results as shown in the second row of Figure 1. The data distribution between training data and testing data is also required to be consistent. If a new dataset is adopted for testing, the deep network needs to be retrained [32] [3].

To address these problems, we propose a non-learning approach for shadow detection following the paired regions theory [9] [31], which detects shadowed region by finding a paired unshadowed region that has the same spectral reflectance. First, the input image is segmented using a region growing approach to maintain the character of shadowed region, especially the penumbra information. Then we locate the candidate paired regions for each region using mathematical morphology operations. Next, when comparing the color of two regions, we adopt physics based confidence coefficients to evaluate whether the darker region belongs to the shadowed region. Finally, we acquire the best paired strategy by maximizing an objective function that combines the relationship between regions, confidence coefficients, penumbra, and other information. Different from [9], the proposed approach uses physical confidence coefficients to improve the reliability when comparing two regions.

The contribution of this paper can be summarized as i), we adopt three non-learning physics based confidence coefficients when evaluating the relationship between two regions. ii), we propose a feasible way to detect the penumbra region around the shadow’s edge and take it as an important cue for shadow detection. iii), we design a novel objective function to evaluate the paired strategy and achieve good performance.

2 Related Work

Traditional approaches The traditional approaches explore the illumination models like [24] [19], then design different hand-craft features to identify shadowed regions. For exam-

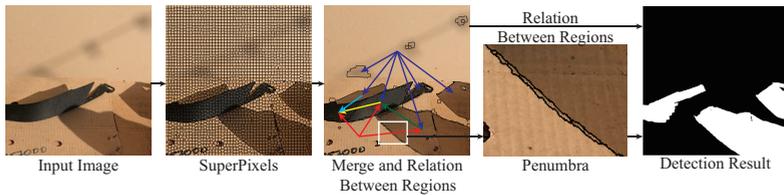


Figure 2: The illustration of our framework. The input image is first segmented using region growing approach. After that we model the relation between regions and detect penumbra region. Finally, the relation between regions and penumbra are used as cues to detect shadowed region.

ple, Zhu et al. [33] classified regions using a conditional random field (CRF). The feature was built based on statistics of intensity, gradient, and texture, computed over local neighborhoods. Huang et al. [30] collected and analyzed the color and intensity change across the edge of shadowed region, then they used the statistical observation to determine whether a detected edge is a shadowed region’s edge. Guo et al. [9] and Vicente et al. [28] combined different kinds of hand-craft features into a histogram as the feature of a candidate region for shadow detection. This kind of approaches rely on the reliability of hand-craft features and traditional discriminators.

Deep neural network based approaches Data-driven deep based approaches use the deep neural network for shadow detection. The large amount of labeled training data assure the performance of these approaches. Nguyen et al. [21] designed a discriminator network to determine whether the generated image is consistent with ground truth. Le et al. [16] designed adversarial training based framework to detect shadowed region. Wang et al. [29] used a generative network to directly produce the shadow detection result under the constrain of both shadow detection and shadow removal. Hu et al. [11] designed a direction-aware module for the backbone network. This module can provide neural network the ability to capture the direction information between pixels, which is very useful for shadow detection. Zhu et al. [34] designed a recurrent attention residual (RAR) module to combine the contexts of adjacent layers to iteratively integrate spatial contexts. Zheng et al. [32] presented a distraction-aware shadow (DS) module to predict hard examples, and fused the obtained distraction features for shadow detection. Chen et al. [3] adopted Multi-Task Mean Teacher(MTMT) structure to reduce the requirement for labeled data when training shadow detection network. Deep neural network based approaches have the advantage of fully-automatic and end-to-end pipeline. But their performance heavily relies on the training data and the consistency between training data and testing data. The convolutional operator also can not effectively model the non-grid relationship between different regions in detail.

In this paper, based on the illumination model [24], we utilize physical equations to analyze the relationship between shadowed region and paired unshadowed region. We also propose an effective strategy to model the relationship between regions for shadow detection.

3 Methodology

According to the paired regions based theory [9], a region can be determined as shadowed region only when a paired unshadowed region is found. To detect shadowed regions, we

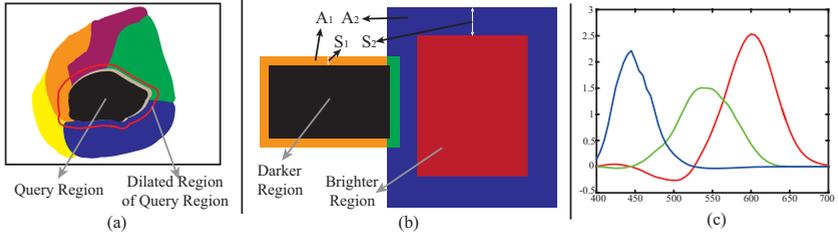


Figure 3: (a) The mathematical morphology operation on how to locate candidate paired regions. The black region is the query region. The red line is the edge of dilated query region. All the regions shown in this figure are the candidate paired regions of query regions. (b) The illustration of getting candidate penumbra region. For two input regions (marked in black and red), the dilated operation is used to get their surrounding regions (marked in yellow and blue). Their intersection region (marked in green) is the candidate penumbra region. (c) The color matching functions of sRGB color space proposed by [25].

need to segment the image properly and find the most reasonable paired strategy between shadowed regions and unshadowed regions.

The illustration of our framework is shown in Figure 2. The input image is first segmented into superpixels using SLIC [1]. Then the superpixels are merged into regions using a region growing image segmentation approach like [24]. We adopt the region growing approach based on the following reasons. (i) The shadowed region only covers some parts of an object in most cases, semantic segmentation [23] and instance segmentation [15] approach may fail to segment the shadowed region correctly. (ii) As a paired regions based approach, the color and texture inside each segmented region should be consistent. (iii) The penumbra region is an important cue for shadow detection and the region growing approach can segment it.

After image segmentation, we model the relation between regions using relative position and confidence coefficients. The penumbra region is detected at the same time. Finally, we evaluate all possible paired strategies that are built based on the relation between regions to find the best one through an objective function.

3.1 Candidate Paired Regions and Penumbra Regions

The step after segmentation is to find candidate paired regions for each region. Because the existence of penumbra region, using adjacent regions as candidates may lead to incorrect results. So we define the regions whose shortest distance to the query region is smaller than a threshold τ as the query region's candidate paired regions. In practice, the query region is dilated by size τ . All the regions that have some parts in the dilated region P are defined as its candidate paired region as shown in Figure 3 (a).

For each region, to detect whether the penumbra region exists between it and one of its candidate paired regions, we design the following operations. First, the darker region is dilated by a small size S_1 , then minus the original region from dilated region to get its surrounding region, denoted by A_1 . S_1 is set to be a small number (5 pixels in the proposed approach) to ignore the involvement of irrelevant regions. The same morphology operation is done to the brighter region to get its surrounding region A_2 . The dilation size S_2 for

the brighter region is τ because the penumbra region sometimes can be very large. The intersection region of A_1 and A_2 is the candidate penumbra region. The illustration of this process is shown in Figure 3 (b). Next, all regions that have some parts covered by candidate penumbra region are explored. If the mean color of a region is bigger than darker region and smaller than brighter region in three channels, and all its pixels lie in the P region of the darker region, we determine there exists penumbra region between these two regions.

3.2 Confidence Coefficients of Candidate Paired Regions

When comparing a candidate paired region with its query region, we adopt physics based confidence coefficients using their mean color to ignore the error caused by feature extraction. According to the illumination model of shadowed region proposed by [24], the only difference between the shadowed region and its paired unshadowed region is that they have different lighting conditions. The color and intensity of shadowed region lack the effect of shielded light source since it is only lighted by the remaining light source. To separate the effect of shielded light source and remaining light source, for each candidate paired regions, we use the following equation:

$$I^d = I - I^s \quad (1)$$

where I refers to the mean RGB color of the brighter region. I^s is the mean color of the darker region. If any channel of I^d is negative, this candidate paired region can not be a paired region of its query region. Because if a light source is added to light a region, each channel of the new color should be at least the same as its original value. This character is set as the first confidence coefficient. So if negative value appears, this region will be eliminated from the candidate paired regions of the query region.

For the remaining candidate paired regions, we need to calculate their relationship between I^d and I^s . To represent these two colors using physical equation, we follow the color estimation proposed by [6]. The three channels of I^d and I^s can be expressed as:

$$I_i^j = L^j \int_{400}^{700} E(\lambda, T^j) R(\lambda) Q_i(\lambda) d\lambda \quad (2)$$

In this equation, j is used to represent d and s . i refers to the i -th channel of the three channels in sRGB color space. L^j refers to the intensity of shielded light source and remaining light source. The spectral distribution function of light source E is simulated by Planck's Law. T^j refers to the color temperature of shielded light source and remaining light source. $R(\lambda)$ refers to the spectral reflectance. $Q_i(\lambda)$ is the parameter of i -th channel in sRGB color matching functions (CMFs), which is shown in Figure 3 (c). To explore the relationship between I^d and I^s , we compare the ratio of their the corresponding channels using Equation 2.

The ratio calculated by Equation 2 can not be used directly, so we need to simplify this equation. The authors in [8] assumed the CMFs is a Dirac delta function, $Q_i = q_i \delta(\lambda - \lambda_i)$, for simplify, then the relationship of I^d and I^s can be expressed as:

$$cc_1(d, s) = \left| \left(\frac{\lambda_r \lambda_b}{\lambda_r - \lambda_b} \right) \log \left(\frac{I_r^d / I_b^d}{I_r^s / I_b^s} \right) - \left(\frac{\lambda_g \lambda_b}{\lambda_g - \lambda_b} \right) \log \left(\frac{I_g^d / I_b^d}{I_g^s / I_b^s} \right) \right| \quad (3)$$

Equation 3 is used to decide whether the shadowed and unshadowed regions belong to the same material in [18], it is also used for intrinsic color calculation in [13]. Here it is set as the

second confidence coefficient. During experiment, we realize that there exists a problem in this equation. According to the CMFs shown in Figure 3 (c), many visible wavelengths have influence on the three channels of RGB color. The assumption like [8] lost many details of E and R . To reinforce this confidence coefficient, we design a new one. Based on the conclusion that naturally occurring spectral reflectance are relatively smooth functions with respect to wavelength [7] [2], we assume the reflectance $R(\lambda)$ to be a constant. Then the ratio of I^d and I^s can be expressed as:

$$\frac{I_i^d}{I_i^s} = \frac{L^d P_i(T^d)}{L^s P_i(T^s)}, \text{ where } P_i(T) = \int_{400}^{700} E(\lambda, T) Q_i(\lambda) d\lambda \quad (4)$$

From Equation 4, we can see that if two regions have the same spectral reflectance but different lighting conditions, the ratio of each corresponding channel is mainly affected by the color of light sources in the scene. So we define the third confidence coefficient as follow:

$$cc_2(d, s) = \arg \min_{T^d \in U^d, T^s \in U^s} \sum_{i=1}^3 |C_i^d / C_i^s - C_i(T^d) / C_i(T^s)| \quad (5)$$

U^d and U^s are the possible color temperature set of remaining light source and shielded light source. In the proposed approach, $U^d \in [3500\text{K}, 6500\text{K}]$ and $U^s \in [6500\text{K}, 12000\text{K}]$. C_i refers to the value of i -th channel after the intensity of color is regularized to one. The overall confidence coefficient of candidate paired regions is defined as:

$$cc(d, s) = \alpha_1 cc_1(d, s) + \alpha_2 cc_2(d, s) \quad (6)$$

3.3 Paired Strategy and Objective Function

According to the result of candidate paired regions, we can build the candidate paired strategies set. First, for all the regions with candidate paired shadowed regions, their indexes are listed as a vector $M^{n \times 1}$. For each element M_i , it has a corresponding vector V_i that contains the index of its candidate paired shadowed regions. The regions in V_i are divided into g_i groups by putting the regions that have similar color and intensity into the same group. Then the paired strategy can be symbolized by a vector $Y^{n \times 1}$. The element $Y_i = \{0, 1, \dots, g_i\}$ refers to the index of shadowed region group under the current paired strategy, and 0 means it has no paired shadowed regions. For the candidate paired regions with penumbra region, corresponding Y_i is fixed. The candidate paired strategies set is formed by all possible values of Y . According to the physical characteristics of the scene, in one paired strategy Y each shadowed region can only have one paired unshadowed region and one region can't be shadowed and unshadowed at the same time. If a candidate paired strategy disobeys this, it is removed from the candidate set.

To find the best paired strategy, we design an objective function using the pixel number, intensity of region and confidence coefficients, which is defined as following:

$$\hat{Y} = \arg \max_Y \sigma_1 \frac{N^d + N^s}{N} + \sum_i^P (\sigma_2 \frac{l_i}{l_j P} - \sigma_3 \max(cc(i, j) - \delta, 0)) \quad (7)$$

In the equation, N^d refers to the pixel number of the unshadowed region while N^s is the pixel number of the shadowed region under the current paired strategy. N is the pixel number of the whole image. i is the index of unshadowed region and j refers to its paired shadowed regions. l is the intensity of corresponding region. P is the number of unshadowed regions

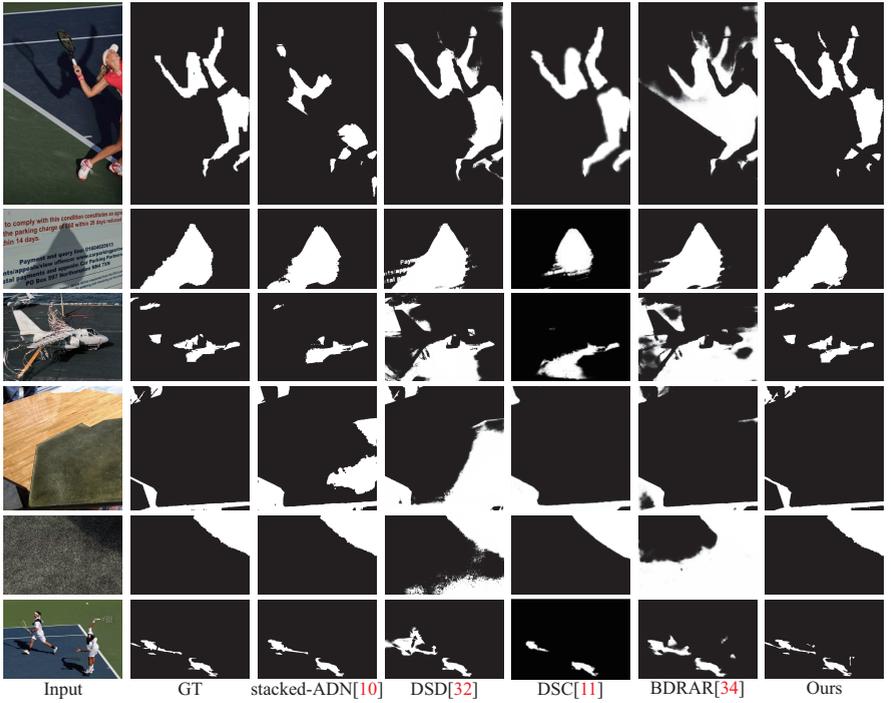


Figure 4: Visual comparison of shadow detection results between different approaches.

in the current paired strategy. δ is a hyperparameter to relieve the error caused by different factors such as noise and secondary reflected light.

4 Experiment

4.1 Datasets and Metrics

We adopt three widely-used shadow detection datasets to evaluate the proposed approach. The first dataset is SBU dataset [27]. It contains 4089 training images and 638 testing images. The second dataset is UCF dataset [33], which has 221 images. The authors in [17] divide this dataset into 111 training images and 110 testing images. The third dataset is ISTD dataset [29], containing 1870 image triplets of shadow image, shadow mask, and shadow-free images. 1330 of them are used as the training set.

To quantitatively evaluate the performance of an approach, we employ the common used shadow detection metrics, the balance error rate (BER) and average precision (AP). They are calculated by:

$$BER = (1 - \frac{1}{2}(\frac{TP}{N_p} + \frac{TN}{N_n})) \times 100, AP = \frac{TP + TN}{N_p + N_n} \quad (8)$$

In this equation, TP and TN are true positives pixels and true negatives pixels in the result of the evaluated approach. N_p and N_n are the number of shadow pixels and the number of non-shadow pixels in the ground truth data.

		SBU [27]		UCF [33]		ISTD [29]	
		AP	BER	AP	BER	AP	BER
Deep Method	scGAN [21]	0.94	9.1	0.92	11.5	0.97	4.7
	stacked-CNN [26]	0.92	11.00	0.89	13.00	0.95	8.6
	ST-CGAN [29]	0.89	13.56	0.87	17.69	0.97	3.85
	A+D Net [16]	0.95	7.67	0.92	11.05	0.97	2.97
	BDRAR [34]	0.96	6.61	0.93	9.45	0.98	2.69
	DSC [11]	0.97	5.59	0.93	10.38	0.97	3.42
	DSD [32]	0.97	4.45	0.95	<u>7.59</u>	0.98	<u>2.31</u>
	stacked-ADN [10]	0.97	4.7	0.94	9.4	0.98	2.72
Traditional Method	Guo[9]	0.76	25.03	0.72	28.32	0.73	27.16
	Zhang[31]	0.95	7.13	0.91	9.21	0.92	8.56
	Ours	0.97	<u>4.65</u>	0.95	7.56	0.98	2.21

Table 1: Quantitative comparison results on shadow detection with AP and BER metric. Bigger AP and lower BER is better.

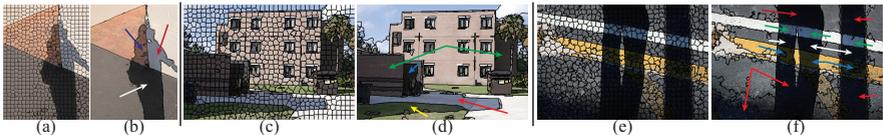


Figure 5: Some intermediate results of the proposed approach. (a)(c)(e) represent the results of image segmentation after SLIC. (b)(d)(f) represent the results after merging superpixels and the arrows represent the paired results.

4.2 Shadow Detection Results and Comparison

Comparison Approaches. Since there are no non-learning shadow detection approaches available, we compare the proposed approach with some classical shadow detection approaches. These approaches include two traditional paired regions based approaches, Guo [9] and Zhang[31], and several state-of-the-art deep based approaches, stacked-CNN [26], scGAN [21], ST-CGAN [29], A+D Net [16], BDRAR [34], DSC [11], DSD [32], stacked-ADN [10]. For deep based approaches, we obtain their evaluation metrics by directly taking the results from the authors if their shadow detection results are not publicly available.

Quantitative comparison. Table 1 represents the quantitative comparison results. It can be seen that although the proposed paired regions based approach is a non-learning approach, it outperforms the best traditional paired based approach [31] by 34.8 %, 17.9 % and 74.1 % on SBU, UCF and ISTD under BER metric, respectively. Early deep based approaches also have a lower performance comparing with the proposed approach. As to the latest deep based approaches, the proposed approach performs favorably against them. The unique deep algorithm that outperforms our approach is DSD [32], which obtains slight performance gains on SBU dataset resorting to data training.

Visual comparison. Some shadow detection results produced by the proposed approach and several state-of-the-art deep neural network based approaches are visualized and compared in Figure 4. From these results, we can see that the involvement of relationship be-

	SBU [27]	ISTD [29]
DSD [32] trained on SBU	4.45	4.80
DSD [32] trained on ISTD	8.51	2.31
Ours	4.65	2.21

Table 2: The model adaptability experiment with BER metric. The numbers in red color refer to the results when the data distribution of training data and testing data are not consistent.

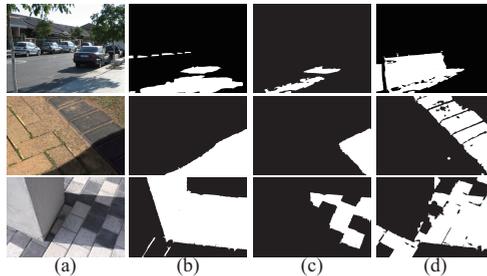


Figure 6: Some shadow detection results using different objective term. (a) are input images. (b) are the results using all objective terms. (c) are the results when the pixel number term is not used. (d) are the results when the intensity term is not used.

tween non-grid regions makes the proposed approach successfully distinguish the dark unshadowed region, while the deep based approaches sometimes generate wrong results.

Intermediate Results. We present some intermediate results in Figure 5. From these results, it can be seen that the merging of superpixels can segment the shadowed region and penumbra region. The paired results satisfy the requirement that the shadowed region and unshadowed region have the same spectral reflectance. These intermediate results show that our approach works well.

4.3 Experiment on Adaptability

We design an experiment to evaluate the adaptability of learning based approaches in different domains. In this experiment, we choose the deep based approach DSD [32] as a comparison. The DSD is trained on SBU dataset and tested on ISTD dataset first. Then, it is trained on ISTD dataset and tested on SBU dataset. The quantitative results are shown in Table 2. It can be seen that if the data distribution of training data and testing data are not consistent, the performance of deep based approaches degrades obviously. Since our approach does not require model training, it is more flexible and robust to handle different testing data in various scenes.

4.4 Objective Function and Failure Cases

Objective function term discussion. We discuss the influences of different objective terms on the final results of shadow detection. We randomly select some pictures from the dataset and show the result of shadow detection using different term combinations as the objective function. The results are shown in Figure 6. From these results, we can see that if the

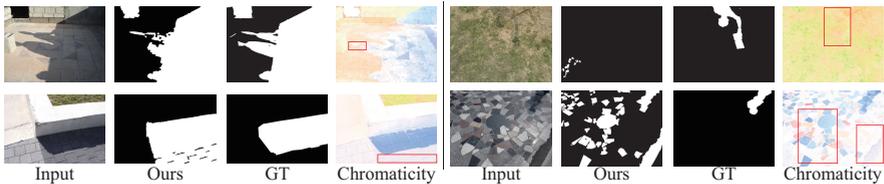


Figure 7: Some failure cases of the proposed approach. The red boxes in chromaticity represent the failure regions.

pixels term is not involved, only the paired region that has the largest ratio in intensity is chosen as the result of optimization. If the intensity ratio term is not involved, the paired result will be very easy to make a mistake, which makes the unshadowed region paired with unshadowed region. If either of these two terms is removed, the proposed approach won't be able to detect shadowed regions anymore. So we don't explore the quantitative metric on benchmark datasets if one of these two terms is removed. The confidence coefficient term is the key part of the proposed approach. So the effect of this term is not explored.

Failure cases. We represent some failure cases in Figure 7. The red boxes in chromaticity images represent the failure regions. In the upper left image, it can be seen the shadowed region in the red box has a different color with the main part of shadowed region, which leads to incorrect segmentation result. In the upper right image, the shadowed region is too shallow to be segmented. These two failure cases are caused by incorrect segmentation results. The remaining two images have dark unshadowed regions and these regions do not have paired shadowed regions, which lead to failure cases.

5 Conclusion and Future Work

In this paper, we propose an updated paired region based approach for shadow detection. First, we design a region growing image segmentation approach to maintain the character of shadowed region. Second, we use physics based confidence coefficients and penumbra to evaluate the relationship between two regions. Then we handle the shadow detection task by designing and maximizing an objective function. The quantitative result proves that these updated operations help the proposed approach outperforms the traditional paired based approaches, and it performs favorably against the state-of-the-art deep neural network based approaches.

Although the proposed approach achieves remarkable shadow detection results, its weakness is also obvious. When a shadowed region doesn't have paired unshadowed region in the picture, the proposed approach will have difficulty in finding it. If two unshadowed regions accidentally meet the confidence coefficient and don't have paired shadowed regions, the proposed approach may fail too. More information such as global features will be explored to solve these problems in the future. We will also try to combine the relationship between non-grid regions with deep neural network to improve performance.

Acknowledgement. This work is supported in part by the National Natural Science Foundation of China Under Grants No. U20B2066, the Open Research Projects of Zhejiang Lab (Grant No. 2021KB0AB01), and the National Key R&D Program of China (Grant No. 2020AAA0109304).

References

- [1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine SÄijsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012. doi: 10.1109/TPAMI.2012.120.
- [2] Jose M. Bioucas-Dias, Antonio Plaza, Gustavo Camps-Valls, Paul Scheunders, Nasser Nasrabadi, and Jocelyn Chanussot. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, 1(2):6–36, 2013. doi: 10.1109/MGRS.2013.2244672.
- [3] Zhihao Chen, Lei Zhu, Liang Wan, Song Wang, Wei Feng, and Pheng-Ann Heng. A multi-task mean teacher for semi-supervised shadow detection. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5610–5619, 2020. doi: 10.1109/CVPR42600.2020.00565.
- [4] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1337–1342, 2003. doi: 10.1109/TPAMI.2003.1233909.
- [5] Aleksandrs Ecins, Cornelia FermÄijller, and Yiannis Aloimonos. Shadow free segmentation in still images using local density measure. In *2014 IEEE International Conference on Computational Photography (ICCP)*, pages 1–8, 2014. doi: 10.1109/ICCPHOT.2014.6831803.
- [6] G.D. Finlayson, S.D. Hordley, Cheng Lu, and M.S. Drew. On the removal of shadows from images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1): 59–68, 2006. doi: 10.1109/TPAMI.2006.18.
- [7] David A. Forsyth and Jean Ponce. *Computer Vision - A Modern Approach, Second Edition*. Pitman, 2012.
- [8] B.V. Funt and G.D. Finlayson. Color constant color indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):522–529, 1995. doi: 10.1109/34.391390.
- [9] Ruiqi Guo, Qieyun Dai, and Derek Hoiem. Paired regions for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12): 2956–2967, 2013. doi: 10.1109/TPAMI.2012.214.
- [10] Le Hou, TomÄas F. Yago Vicente, Minh Hoai, and Dimitris Samaras. Large scale shadow annotation and detection using lazy annotation and stacked cnns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4):1337–1351, 2021. doi: 10.1109/TPAMI.2019.2948011.
- [11] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, Jing Qin, and Pheng-Ann Heng. Direction-aware spatial context features for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(11):2795–2808, 2020. doi: 10.1109/TPAMI.2019.2919616.

- [12] Imran N. Junejo and Hassan Foroosh. Estimating geo-temporal location of stationary cameras using shadow trajectories. In *European Conference on Computer Vision*, pages 318–331, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. doi: 10.1007/978-3-540-88682-2_25.
- [13] R. Kawakami, K. Ikeuchi, and R.T. Tan. Consistent surface color for texturing large objects in outdoor scenes. In *The IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1200–1207 Vol. 2, 2005. doi: 10.1109/ICCV.2005.62.
- [14] Jean-François Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. Estimating natural illumination from a single outdoor image. In *2009 IEEE 12th International Conference on Computer Vision*, pages 183–190, 2009. doi: 10.1109/ICCV.2009.5459163.
- [15] Issam Laradji, David Vázquez, and Mark Schmidt. Where are the masks: Instance segmentation with image-level supervision. In *2019 British Machine Vision Virtual Conference (BMVC)*, 2019.
- [16] Hieu Le, Tomas F. Yago Vicente, Vu Nguyen, Minh Hoai, and Dimitris Samaras. A+net: Training a shadow detector with adversarial shadow attenuation. In *European Conference on Computer Vision*, pages 680–696, Cham, 2018. Springer International Publishing. doi: 10.1007/978-3-030-01216-8.
- [17] Li Shen, Teck Wee Chua, and K. Leman. Shadow optimization from structured deep edge detection. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2067–2074, 2015. doi: 10.1109/CVPR.2015.7298818.
- [18] John A. Marchant and Christine M. Onyango. Shadow-invariant classification for scenes illuminated by daylight. *J. Opt. Soc. Am. A*, 17(11):1952–1961, Nov 2000. doi: 10.1364/JOSAA.17.001952.
- [19] Bruce A. Maxwell, Richard M. Friedhoff, and Casey A. Smith. A bi-illuminant dichromatic reflection model for understanding images. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008. doi: 10.1109/CVPR.2008.4587491.
- [20] S. Nadimi and B. Bhanu. Physical models for moving shadow and object detection in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8):1079–1087, 2004. doi: 10.1109/TPAMI.2004.51.
- [21] Vu Nguyen, Tomas F. Yago Vicente, Maozheng Zhao, Minh Hoai, and Dimitris Samaras. Shadow detection with conditional generative adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 4520–4528, 2017. doi: 10.1109/ICCV.2017.483.
- [22] Takahiro Okabe, Imari Sato, and Yoichi Sato. Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1693–1700, 2009. doi: 10.1109/ICCV.2009.5459381.

- [23] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):640–651, 2017. doi: 10.1109/TPAMI.2016.2572683.
- [24] Yael Shor and Dani Lischinski. The shadow meets the mask: Pyramid-based shadow removal. *Computer Graphics Forum*, 27(2):577–586, 2008. doi: 10.1111/j.1467-8659.2008.01155.x.
- [25] W. S. Stiles and J. M. Burch. Npl colour-matching investigation: Final report (1958). *Optica Acta*, 6(1):1–26, 1963.
- [26] Tomás F. Yago Vicente, Le Hou, Chen-Ping Yu, Minh Hoai, and Dimitris Samaras. Large-scale training of shadow detectors with noisily-annotated shadow examples. In *European Conference on Computer Vision*, pages 816–832, Cham, 2016. Springer International Publishing. doi: 10.1007/978-3-319-46466-4.
- [27] Tomás F. Yago Vicente, Le Hou, Chen-Ping Yu, Minh Hoai, and Dimitris Samaras. Large-scale training of shadow detectors with noisily-annotated shadow examples. In *European Conference on Computer Vision*, pages 816–832, Cham, 2016. Springer International Publishing. doi: 10.1007/978-3-319-46466-4.
- [28] Tomãas F. Yago Vicente, Minh Hoai, and Dimitris Samaras. Leave-one-out kernel optimization for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3):682–695, 2018. doi: 10.1109/TPAMI.2017.2691703.
- [29] Jifeng Wang, Xiang Li, and Jian Yang. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1788–1797, 2018. doi: 10.1109/CVPR.2018.00192.
- [30] Xiang Huang, Gang Hua, J. Tumblin, and L. Williams. What characterizes a shadow boundary under the sun and sky? In *2011 International Conference on Computer Vision*, pages 898–905, 2011. doi: 10.1109/ICCV.2011.6126331.
- [31] Ling Zhang, Qing Zhang, and Chunxia Xiao. Shadow remover: Image shadow removal based on illumination recovering optimization. *IEEE Transactions on Image Processing*, 24(11):4623–4636, 2015. doi: 10.1109/TIP.2015.2465159.
- [32] Quanlong Zheng, Xiaotian Qiao, Ying Cao, and Rynson W.H. Lau. Distraction-aware shadow detection. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.
- [33] Jiejie Zhu, Kegan G. G. Samuel, Syed Z. Masood, and Marshall F. Tappen. Learning to recognize shadows in monochromatic natural images. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 223–230, 2010. doi: 10.1109/CVPR.2010.5540209.
- [34] Lei Zhu, Zijun Deng, Xiaowei Hu, Chi-Wing Fu, Xuemiao Xu, Jing Qin, and Pheng-Ann Heng. Bidirectional feature pyramid network with recurrent attention residual modules for shadow detection. In *European Conference on Computer Vision*, pages 122–137, Cham, 2018. Springer International Publishing. doi: 10.1007/978-3-030-01231-1_8.