CGI2019 manuscript No.
(will be inserted by the editor)

Effective Shadow Removal via Multi-scale Image Decomposition

Ling Zhang · Qingan Yan · Yao Zhu · Xiaolong Zhang · Chunxia Xiao*

Abstract Shadow removal is a fundamental and challenging problem in image processing field. Current approaches can only process shadows with simple scenes. For complex texture and illumination, the performance is less impressive. In this paper, we propose a novel shadow removal algorithm based on multi-scale image decomposition, which can recover the illumination for complex shadows with inconsistent illumination and different surface materials. Independent of shadow detection, our algorithm only requires a rough boundary distinguishing shadow regions from non-shadow regions. It first performs a multi-scale decomposition for the input image based on an illumination-sensitive smoothing process and then removes shadows in the basic layer using a local-to-global optimization strategy, which fuses all local shadow-free results in a global manner. Finally, we recover the texture details for the shadow-free basic layer and obtain the final shadow-free image. We validate the performance of the proposed method under various lighting and texture conditions and show consistent illumination between shadow and surrounding regions in the shadow removal results.

Ling Zhang · Xiaolong Zhang
Hubei Key Laboratory of Intelligent Information Processing and Realtime Industrial System, School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China, 430081. Email: lingzhang@whu.edu.cn, xiaolong.zhang@wust.edu.cn.

Qingan Yan
JD.com American Technologies Corporation, CA, 94043. Email: qingan.yan@jd.com.

Yao Zhu · Chunxia Xiao
School of Computer Science, Wuhan University, Wuhan, China, 430072. Email: juryoo@qq.com, cxxiao@whu.edu.cn.
*Corresponding to Chunxia Xiao.

Keywords Shadow removal · image processing · complex shadow · image decomposition

1 Introduction

Shadow is ubiquitous in our daily life. However, its low brightness in shadow regions always gives rise to interference for vision tasks, such as object recognition, image retrieval and target tracking. Shadow removal is thus a fundamental topic in computer vision and image processing communities, especially for complex shadows. While there have been many works in this field [7,9,10,12,27,32,35], most of them are aimed at processing hard shadows under simple reality conditions. Rare methods can actually handle shadows with inconsistent illumination and different surface materials.

In order to achieve desired shadow-free images, it involves three main challenges. First, complex shadows are usually hard to be detected accurately, for example, the soft shadow with unconspicuous shadow boundary. Second, it is difficult to obtain successful shadow removal results when the illumination in shadow regions is non-uniform, especially in shadow boundaries. Third, the rapid illumination change in shadow regions is easy to cause color distortion or detail loss (artifacts) in the shadow removal result.

One typical group of traditional shadow removal methods are to recover the illumination in shadow regions using illumination transfer [21,28,33–35], which borrow the illumination from non-shadow regions to shadow regions. The recovered illumination may be different according to different non-shadow sampling regions. Another typical group of shadow removal approaches remove shadows based on gradient domain manipulation [7,15,17]. A common idea in these tech-

niques is to nullify the gradient on shadow boundaries and reconstruct the shadow-free result utilizing the gradient information on shadow regions. But the invalid gradient will lead to texture loss on shadow boundaries. Moreover, there may be error propagation in the process of illumination reconstruction, which produces shadow removal result with color distortion.

More recently, deep neural network is introduced for shadow removal [11,19,25]. This technology is very effective when there is a large number dataset with different shadow and shadow-free image pairs. In order to activate this technique, it requires that the dataset should contain a full range of shadow types and a wide variety of surface materials in shadow regions. But current publicly available shadow datasets are relatively small, as shown in Table 1. The limited amount of training data could cause unsatisfied artifacts for some shadow images. Besides, such learning-based methods usually compress the image resolution before processing, which is always accompanied by certain quality reduction for the shadow removal result.

To overcome the limitations of aforementioned methods, we present a novel shadow removal algorithm which can recover illumination for complex shadow images. Different from the most existing shadow removal methods, our approach needs not to detect the accurate shadow regions. We distinguish the shadow and non-shadow regions using a rough shadow mask provided by users. To remove the shadow, we first perform a multi-scale image decomposition for the shadow image based on illumination-sensitive smoothing method. Then, we remove shadows in the basic layer through local-to-global optimization strategy and produce the final shadow-free result by reintegrating the detail information into the shadow-free basic layer. Since the rapid illumination change may cause texture detail loss which leads to distortion in the shadow removal result, to acquire high-quality outputs, we refine the details within these distortion regions via texture detail synthesis. Figure 1 shows the overview of the proposed shadow removal algorithm.

The main contributions and advantages of the proposed method are two aspects:

- Introduce a novel shadow removal method using the local-to-global optimization strategy, which not only works well in shadow image with different materials but also can handle shadows with inconsistent illumination.
- Develop an illumination-sensitive smoothing method for shadow images, which can extract the texture details with less or no illumination information.

2 Related work

Several shadow removal methods are proposed based on gradient domain manipulation [5–7]. Finlayson et al. [7] removed the shadows by performing gradient operations for non-shadow regions. This method depends on accurate shadow edges detection and may not produce satisfactory results due to the inaccurate shadow edges detection. Mohan et al. [17] removed shadows using gradient domain manipulation. This method requires much user interaction to specify the shadow boundary. Liu et al. [15] removed shadow by solving a Poisson equation, which constructed a shadow-free and texture-consistent gradient field between the shadow and lit area.

Shadow matting is also exploited in shadow detection and removal [13,14]. Chuang et al. [16] proposed a method for shadow extracting and editing which considered the input image as a linear combination of a shadow-free image and a shadow matte image. Wu et al. [26,27] supposed shadow effect as a light attenuation problem. Methods [26] and [27] applied user-supplied hints to identify shadow and non-shadow regions. Although these two methods tried to preserve the texture appearance under the extracted shadow, they still do not effectively recover the image detail in the shadow areas. Gryka et al. [9] removed soft shadows applying a data-driven method. They estimated the shadow mattes by learning a regression function from the shadowed image regions and their shadow mattes. But this method is limited by the training set. Xiao et al. [30] applied depth information provided by the depth sensor to remove shadows in RGB-D images.

Several shadow removal methods are proposed based on illumination or color transferring. Inspired by the color transfer theory [20], Shor et al. [21] performed pyramid-based restoration process to remove shadows in the image. This method requires that the shadow regions and the sample region share similar texture to produce satisfied results. Moreover, this method can only handle the shadow regions with uniform texture and illumination. By improving [21], Xiao et al. [28] recovered the illumination under the shadow regions using adaptive multi-scale illumination transfer. Later, Xiao et al. [29] completed shadow removal by performing illumination transfer from the non-shadow regions to the matched shadow regions. But this method may fail to ensure smooth transition between sub-regions. Guo et al. [10] presented a method to detect as well as removing shadows based on paired regions. Zhang et al. [35] removed the shadows in image by using a coarse-to-fine illumination optimization strategy. Due to the size limitation of local patches, this method is difficult to

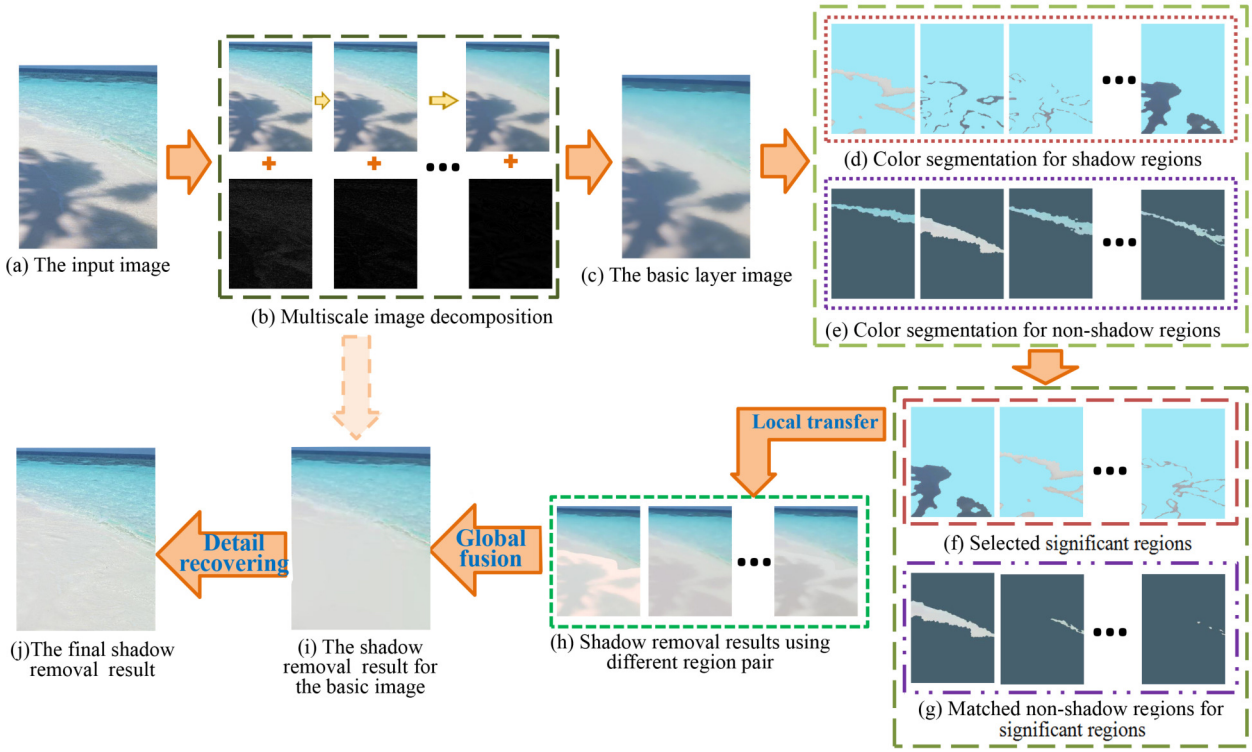


Fig. 1 The overview of the proposed shadow removal method. We remove shadows in the basic layer (g) using the local-to-global strategy (d-i) and obtain the final shadow removal result after detail recovering.

provide satisfying results for complicated shadows and shadows with large illumination variances.

Recently, deep learning networks have been applied to the task of shadow detection and removal [11,12,19, 25]. Khan et al. [12] used multiple convolutional neural networks to learn useful feature representations for shadow detection. With the shadow detection results, integrating multi-level color transfer they further proposed Bayesian formulation for the shadow removal. Vicente et al. [23] considered shadow detection as a problem of labeling image regions and trained a kernel least-squares support vector machine for labeling shadow regions. Qu et al. [19] proposed an end-to-end DeshadowNet to recover illumination in shadow regions. These methods may create visible artifacts in the shadow regions if the shadow type and surface material are not well represented in the training database. In particular, Wang et al. [25] learned shadows from a single image and proposed a novel STacked Conditional Generative Adversarial Network (ST-CGAN) to perform the two tasks of shadow detection and shadow removal. Different from the commonly used multi-branch paradigm, they stacked all the tasks in a perspective for multi-task learning. Additionally, such deep learning methods require access to the training database, which greatly restricts possible application scenarios.

3 Shadow removal

In order to process complex shadow, such as shadows with inconsistent illumination and different surface materials, we introduce a local-to-global strategy to perform the task of shadow removal. Since smooth image is piecewise continuous and beneficial to region division, we first perform multi-scale image decomposition and then do shadow removal in the basic layer image.

Figure 1 shows the overview of the proposed shadow removal method. We first decompose the input image using an illumination-sensitive filtering method to get a smooth base layer (Fig. 1(c)). Then, we remove shadows in the basic layer using the local-to-global strategy. Specifically, we divide the basic layer into different regions (Fig. 1(d)) and select the significant regions in shadow regions (Fig. 1(f)), which have high occurrence probability or contain key material and illumination information in shadow regions. For each significant region, we find its matched region (Fig. 1(g)) in the divided non-shadow regions (Fig. 1(e)). Next, we remove shadows separately based on the different match pairs (Fig. 1(h)). Subsequently, we blend the different shadow removal results and obtain the global shadow-free result for the basic layer (Fig. 1(i)). Finally, we integrate the details back to the shadow-free basic layer and obtain the shadow removal results (Fig. 1(j)).

(a) (b) (c) (d) (e) (f)

Fig. 2 Image segmentation process. (a) is input image. (b) is shadow mask. (c) is the basic layer after multi-scale image decomposition. (d) is the superpixel segmentation result using SLIC method. The superpixel number is set to 1500 and compactness is 10. (e) is the clustering result for shadow regions. (f) is the clustering result for non-shadow regions.

Different from most existing shadow removal methods, we only need a rough shadow mask, covering all shadows, to distinguish shadow regions from non-shadow regions, as shown in Fig. 2(b). In the shadow mask, the black part is the non-shadow region, whereas the white indicates the shadows.

3.1 Multi-scale image decomposition

Multi-scale image decomposition [3,18,22,31] is to decompose an image into one smooth basic layer and several detail layers. The basic layer contains the main color information of the image, while the detail layers represent the coarse shape information for the image. Let I be the input image, the multi-scale decomposition for I can be represented as:

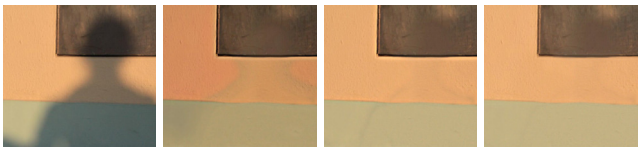
$$I = b + \sum_{i=1}^{N} L_i, \tag{1}$$

where N is the number of layers, b is the basic layer image, L_i represents detail layer information at the i-th smoothing operation. The basic layer b is the image after the N-th smoothing operation on the input image. Let S_i be the smooth result after the i-th smoothing, L_i is defined as:

$$L_i = S_{i-1} - S_i, \tag{2}$$

where $i \in \{1, 2, \cdots, N\}$ and $S_0 = I$.

While there are many filter algorithms, such as Gaussian smoothing and bilateral filtering, they do not take the illumination variation into consideration, and the detail layer would exhibit details caused by illumination changing, such as details from shadow boundaries. This may lead to severe artifacts in the shadow removal result, as shown in Fig. 3(b, c). Therefore, to handle shadows with inconsistent illuminance, we present an illumination-sensitive smoothing algorithm for image multi-scale decomposition.



(a) (b) (c) (d)

Fig. 3 Shadow removal results with different smoothing methods. (a) is the input image. (b) is shadow removal result using Gaussian filter. (c) is shadow removal result using bilateral filter. (d) is our shadow removal result using the proposed smoothing method.

It is notable that while the color, illumination and tone of one pixel point vary with different light condition, its material keeps the same. Therefore, we find related points for each point based on the relevance of the color, illumination and tone between point i and its neighbor j.

Color relevance The color for two points with the same illumination and reflectance are similar. So we use color relevance D_c to identify whether two points have the similar illumination and material. The formula is:

$$D_c = exp(-\frac{||I_i - I_j||^2}{2\sigma_c^2}), \tag{3}$$

where I_i is the intensity at point i in RGB space and I_j is the intensity at point j. σ_c is color deviation in its local region.

Illumination relevance As the light source is obscured, the illumination in shadow region is often lower than that in non-shadow region. The illumination relevance D_l between two points is defined as:

$$D_l = exp(-\frac{(L_i - L_j)^2}{2\sigma_l^2}), \tag{4}$$

where L_i and L_j are the illumination at point i and j, which are represented by the L channel in Lab color space. σ_l is the illumination deviation in its local region.

Tone relevance Generally, shadow and non-shadow regions vary in tone. Thus it can be used as a feature

to distinguish shadow from non-shadow areas. The tone relevance D_h is defined as:

$$D_h = exp(-\frac{(H_i - H_j)^2}{2\sigma_h^2}), \tag{5}$$

where H_i and H_j are the tone value at point i and j, which are represented by the H channel in HIS color space. σ_h is the deviation of tone in its local region.

In our experiments, if the overall relevance value $D = D_c \times D_l \times D_h$ is above 0.8 (the value of 0.8 is determined by lots of experiments, which could produce good performance for most scenes), we regard point i and j as the relevant points. If one point has no relevant point, we regard it as an isolated point and remain its illumination unchanged; otherwise, we use the average intensity of all relevant points as the new illumination for this point.

Let R_i be the set of relevant points at point i. i is a pixel point in shadow regions. The new illumination at point i is computed using the relevant point $j \in R_i$:

$$I_i^{smooth} = \frac{\Sigma_{j \in R_i} w_{ij} I_i}{\Sigma_{j \in R_i} w_{ij}}, \tag{6}$$

where I_i is the intensity at point i. w_{ij} is the similarity between point i and j, and $w_{ij} = exp(-\frac{(I_i - I_j)^2}{2\sigma^2})$.

Figure 3(d) is the shadow removal result using the proposed smoothing methods, which shows that the proposed illumination-sensitive smoothing method works better in the shadow images. Note that, in the process of image smoothing, we process the shadow regions and non-shadow regions separately.

3.2 Shadow removal for the basic layer

Our smoothing method is illumination-sensitive. The detail layers contain less or no illumination details. Therefore, the shadow removal task can be transferred from the input image to the basic layer. On the other hand, to deal with shadows with inconsistent illumination, we propose a local-to-global strategy to recover the illumination in shadow regions for the basic layer.

Local transfer Considering that the illumination in shadow regions may be different, we first process the basic layer in local manner, which divide the basic layer into different regions and remove their shadows respectively. Since the basic layer is a smooth image and a superpixel is a small part with similar color, we cluster the basic layer using superpixels instead of pixels. In our experiment, we use method [1] to perform superpixel segmentation. Note that, we process the shadow and non-shadow regions separately. To increase the accuracy of superpixel clustering, we adopt a small size for superpixels, as shown in Fig. 2(d). We cluster the superpixels by comparing the color and illumination in the superpixels.

Suppose P_1 and P_2 are two superpixels with feature vectors (c_1, l_1) and (c_2, l_2) respectively. The clustering similarity between them is:

$$s = exp(\frac{-(c_1 - c_2)^2}{2\sigma_c^2}) + exp(\frac{-(l_1 - l_2)^2}{2\sigma_l^2}), \tag{7}$$

where c_1, c_2 are the corresponding color value and l_1, l_2 are the corresponding illumination value. If $s < T$, we consider P_1 and P_2 locating in the same sub-region (in our experiments, T is set to 1.2).

After clustering, the superpixels with similar color and illumination are divided into a common region. The shadow and non-shadow regions are divided into different regions based on different colors and illumination, as shown in Fig. 2(e, f). Because the number of the divided regions may be large, we extract the significant regions from the divided shadow regions, as shown in Fig. 1(f). The significant regions consist of two parts: main regions and key regions. The main regions are those with large size in the divided shadow regions and have high occurrence probability in the image. The key regions have low probability but containing key material or illumination information.

The selection for significant regions follows the following steps:

(1) Select 5 main regions from the divided regions which have the largest sizes. If the average intensity difference between two main regions is lower than 10 in all three channels, the smaller one will be deleted from the main regions.

(2) Find several key regions based on surface material and illumination in shadow regions. If the average intensity difference between current region and the main regions selected in (1) is larger than 20, we consider it as the key region.

(3) Label significant regions. If the key regions and main regions contain a common region, the common region will be labeled once.

To remove the shadows, we find the matched region in the divided non-shadow regions for each significant region using method based on covariance matrix [35]. Since the basic layer has poor texture information, we use the information of the input image for region matching. With the matched region pair, we can remove shadows in the significant region using illumination transfer equation in [21].

Suppose there are N significant regions in shadow regions, labeled as $S_{color}^r, r \in \{1, 2, \cdots, N\}$. Using the method based on covariance matrix [35], we find the matched region L_{color}^r for each significant region S_{color}^r.

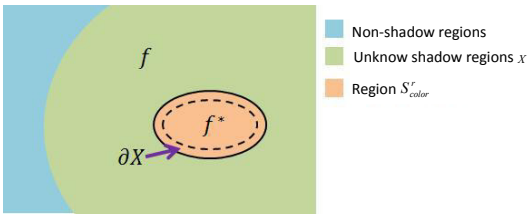


Fig. 4 Shadow interpolation process.

S^r_{avg} and $\sigma(S)$ are the mean value and standard deviation of S^r_{color}, respectively. The mean value and standard deviation of L^r_{color} are labeled as L^r_{avg} and $\sigma(L)$. The intensity at shadow point i in the basic layer is b_i. B^r_i is the shadow-free value at point i, which is estimated as:

$$B^r_i = \frac{\sigma(L)}{\sigma(S)}(b_i - S^r_{avg}) + L^r_{avg}. \tag{8}$$

Equation 8 is used to remove shadows in the corresponding significant regions, but other shadow regions still need further processing. We suppose that illumination in shadow regions is constant (the inconsistent illumination in shadows will be discussed in the following), so the gradient information does not change after shadow removal. Using the idea of Poisson equation interpolation [8], we consider the shadow-free significant regions as known areas and reconstruct the shadow-free illumination of other regions using the gradient information in shadow regions.

As shown in Fig. 4, X represents the unknown shadow regions which is shadow regions besides significant region S^r_{color}. ∂X is the boundary of S^r_{color}. Let f be the objective function for the shadow points in X and f^* be the known function defined over S^r_{color}. The vector field V is guidance field, which is the gradient of the basic layer. The objective function f can be obtained by solving the following Poisson equation:

$$\min_f \iint_X |\nabla f - V|^2, f|_{\partial x} = f^*|_{\partial x}. \tag{9}$$

For $i \in X$, N_i is the set of its 4-connected neighbors. Let $< i, j >$ denote point pair and $j \in N_i$. The boundary ∂X can be denoted that $\partial X = \{i \in S^r_{color} | X : N_i \cap X \neq \emptyset\}$. Then, Eq. 9 can be written as:

$$\min_{f|X} \sum_{<i,j> \bigcap X \neq \emptyset} (f_i - f_j - v_{ij}), f_i = f^*_i, i \in \partial X. \tag{10}$$

According to the optimality condition, the solution for Eq. 10 satisfies the following simultaneous linear equation:

$$|N_i| f_i - \sum_{j \in N_i \cap X} f_j = \sum_{j \in N_i \cap \partial X} f^*_j + \sum_{j \in N_i} v_{ij}, \tag{11}$$

where $|N_i|$ is the number of N_i. v_{ij} is the gradient on oriented edge $[i, j]$ in the basic layer. We use Gauss-Seidel iterative method to solve this linear equation. In the iterative optimization process, we use illumination in S^r_{color} to estimate the shadow-free values in X.

Global fusion Using a region pair, we can get a shadow-free result for the basic layer. The significant region is local area and the shadow removal result could have better effect on the local area. But the results may not be satisfactory in other regions, as shown in Fig. 1(h). The reason for the unsatisfied results is that the local transfer part supposes illumination in shadow regions is consistent, but the illumination is inconsistent in these regions actually. Given a different region pair, it produces another different shadow removal result. Therefore, to get an optimal shadow removing result for the basic layer, we fuse all the candidate shadow-free results in a global manner.

Let B^r be the candidate shadow-free result corresponding to region S^r_{color}. The fusion operation at point i is defined as:

$$B_i = \frac{\sum_r \lambda^r_i d^r_i B^r_i}{\sum_r \lambda^r_i d^r_i}, \tag{12}$$

where B_i is the intensity at point i after fusion, λ^r_i and d^r_i measure the color difference and distance difference between point i and S^r_{color}, respectively. B^r_i is the intensity at point i in candidate result B^r.

Let B^r_{avg} be the average intensity on S^r_{color} in B^r, b_i is the intensity at point i in the basic layer and S^r_{avg} is the average intensity on S^r_{color} in the basic layer. The color measurement λ^r_i between point i and S^r_{color} is computed using the following equation:

$$\lambda^r_i = 1 - \frac{B^{r\text{-}est}_i - B^r_i}{B^r_{avg}}, \tag{13}$$

where $B^{r\text{-}est}_i = \frac{b_i B^r_{avg}}{S^r_{avg}}$.

The position measurement d^r_i between point i and S^r_{color} is computed as:

$$d^r_i = 1 - \sqrt{\frac{(d_{ix} - d^r_{ix})^2 + (d_{iy} - d^r_{iy})^2}{D^2_x + D^2_y}}, \tag{14}$$

where D_x and D_y are the width and height of the input image respectively. (d_{ix}, d_{iy}) denotes the coordinate position of point i in the image, (d^r_{ix}, d^r_{iy}) is the coordinate position of point that is nearest point to i in region S^r_{color}. The fusion result (Fig. 1(i)) illustrates the benefit of the local-to-global strategy.

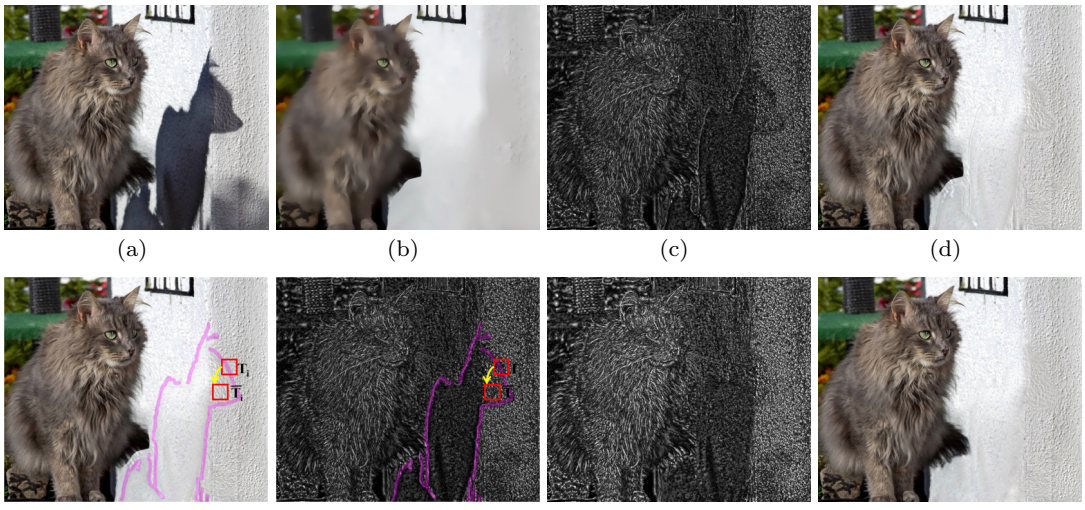


Fig. 5 Shadow boundary processing. (a) is the input image. (b) is shadow-free result for the basic layer. (c) is the detail image, which are magnified 8 times. (d) is the shadow removal result incorporating detail image (c). The purple region in (e) and (f) are the distortion shadow regions. (g) is the repaired detail image on shadow boundaries. (h) is the corresponding shadow removal result using the repaired detail image.

Fig. 6 Detail recovering and enhancement. (a) is the input image. (b) is the basic layer, which is a smooth image. (c) is the detail image by magnifying 2 times. (d) is the shadow-free result for basic layer. (e) is the shadow removal result with $\beta = 1$. (f) is the detail enhancement result in shadow regions with the estimated $\beta = 2.03$.

Fig. 7 Shadow removal result using our algorithm. (a) is the input image. (b) is the shadow removal result.

3.3 Detail recovering

Detail recovering includes two stages: detail recovering for the basic layer and detail recovering for the distortion regions on shadow boundaries.

Detail recovering for the basic layer The shadow removal result for the basic layer is a smooth image. We should reconstruct the details of the image. We use the detail information on the multi-scale decomposition to recover the detail. Let I^{free} be the shadow removing result with detail recovering, and B be the shadow removing result for the basic layer. Then the detail recovering result can be expressed as:

$$I^{free} = B + \sum_{i=1}^{N} L_i. \tag{15}$$

As low brightness of shadow regions usually weakens the detail information, to effective recovery the detail in shadow regions, we add a weight coefficient β in Eq. 15 for detail enhancement, which is that:

$$I^{free} = B + \beta \sum_{i=1}^{N} L_i. \tag{16}$$

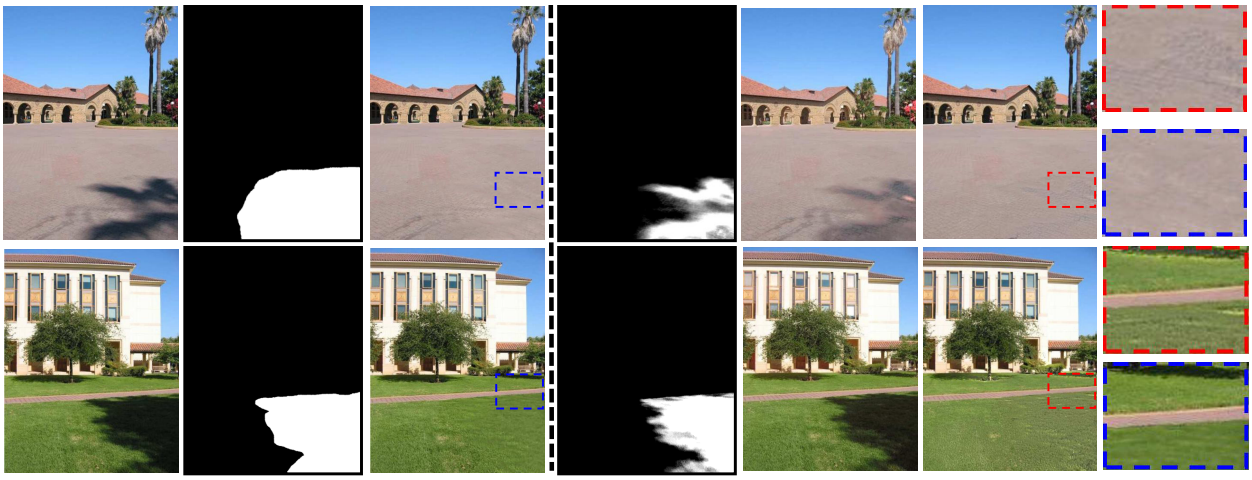


Fig. 8 Shadow removal results for shadows with inconsistent illumination. (a) Input images. (b) Shadow mask used in our method. (c) Our shadow removal results. (d) Shadow mask used in methods of [10] and [30]. (e) Shadow removal results of [10]. (f) Shadow removal results of [30]. (e) Close-ups for the color boxes in (c) and (f)

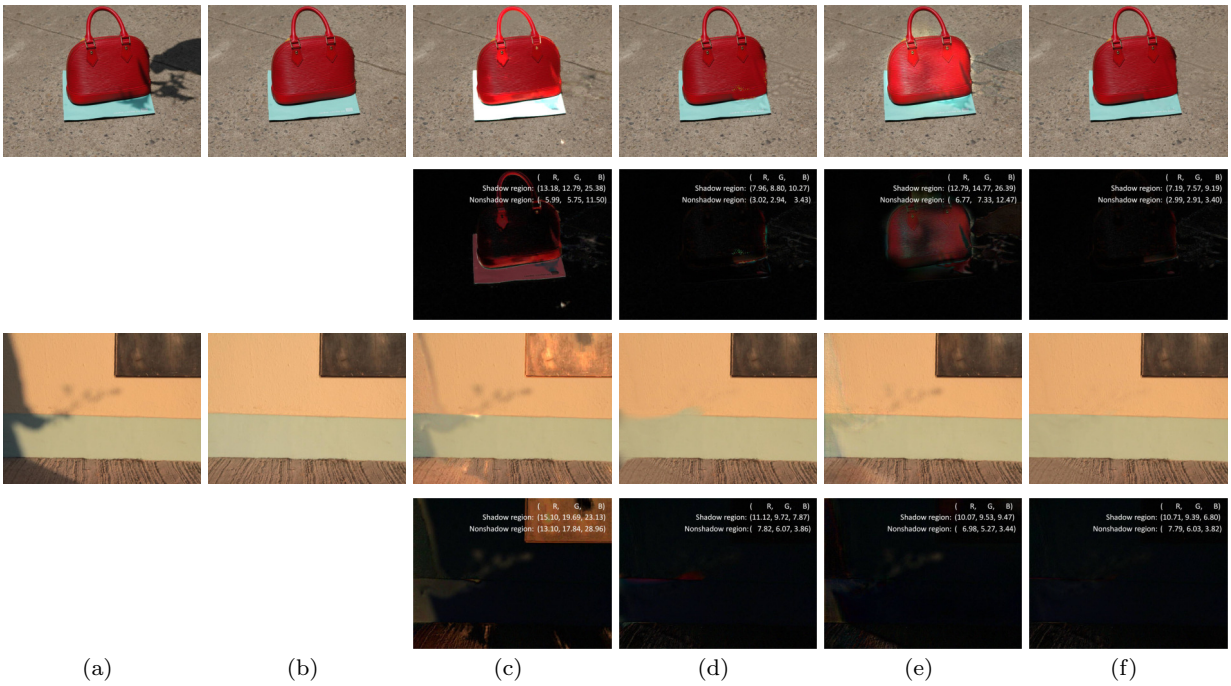
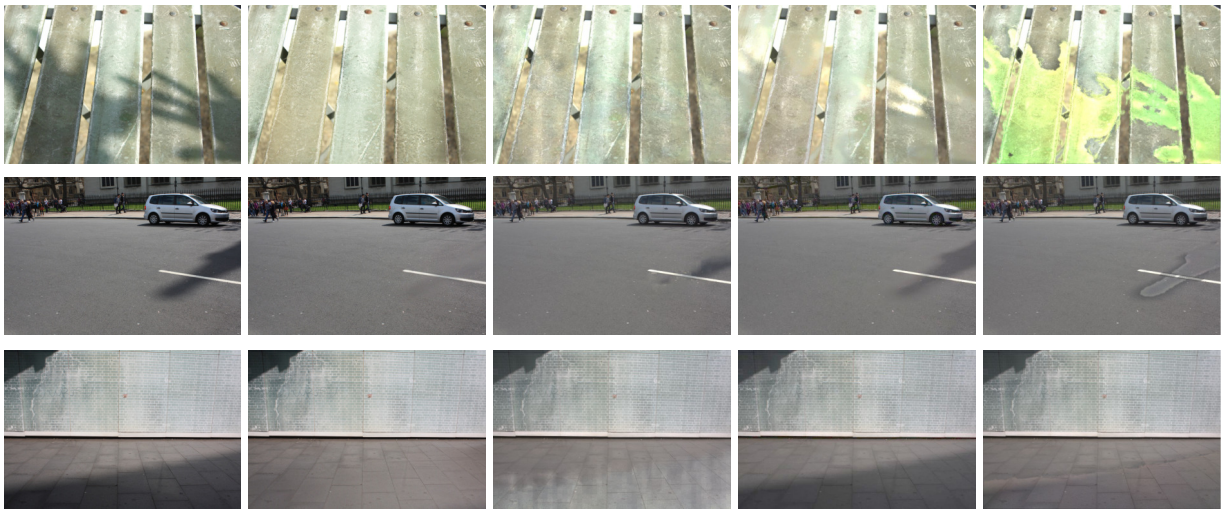


Fig. 9 Shadow removal comparison results. The first and third row are the shadow removal results. The second and fourth roware the pixelwise differences between the corresponding shadow removal result and the ground truth. We also present the average difference values in shadow and nonshadow region separately, which are calculated in the RGB color space. (a) Input images from SRD dataset [19]. (b) Ground truth. (c) Results of [10]. (d) Results of [35]. (e) Results of [19]. (f) Our results.

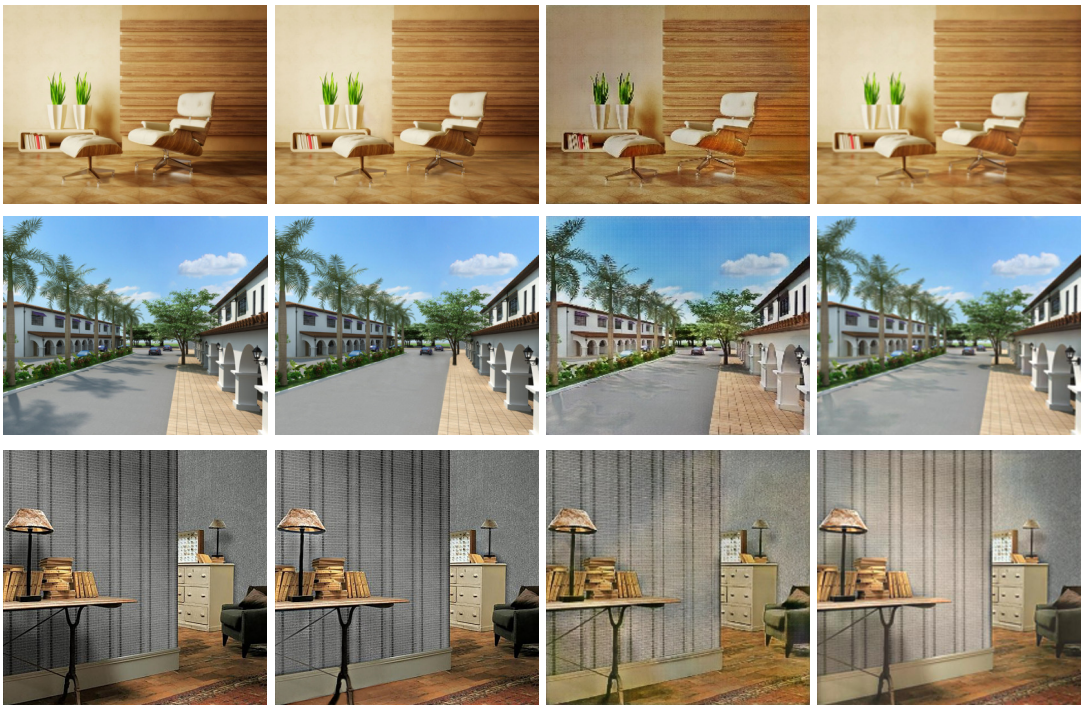
Parameter β is controlled by the detail information between shadow and non-shadow regions, and $\beta = \frac{D_L}{D_S}$. D_L is the mean value of the detail image in non-shadow regions and D_S is the mean value in shadow regions. The value of β is different for different image. Figure 6 shows an example of the effectiveness of detail measurement.

Detail recovering for shadow boundaries The rapid illumination changing on shadow boundaries usually leads to detail loss (Fig. 5(c)), which results in brightness or color distortion in the shadow removal result (Fig. 5(d)). So we need to repair the distortion regions. In the basic layer, there is no distortion in its shadow removal result, so we only repair details in the distortion regions.



(a) (b) (c) (d) (e)

Fig. 10 Shadow removal results. (a) Input images from LRSS dataset [9]. (b) Our results. (c) Results of [9]. (d) Results of [10]. (e) Results of [2].



(a) (b) (c) (d)

Fig. 11 Shadow removal results compared with methods using deep neural network. (a) Input images. (b) Our results. (c) Results of [25]. (d) Results of [11].

We use the local patch optimization method proposed in [4] to repair the lost details. We first specify the distortion regions and find a suitable matching patch from the surrounding regions for each patch in the distortion regions, as shown the purple regions in Fig. 5(e). Then we fill details into the distortion regions by using details in the matched patch. We use patch matching method [4] because the patch operations in this method such as translation, rotation and scaling effectively increase the accuracy of patch matching. Let T_i be a patch centered in the input image at point i in the distortion regions and $\overline{T_i}$ be the matching patch for T_i. The details W_i for patch $\overline{T_i}$ can be expressed as:

$$W_i = \overline{T_i} - \overline{b_i}, \tag{17}$$

where $\overline{b_i}$ is the patch in the basic layer which has the same position with $\overline{T_i}$. As shown in Fig. 5(f), we use the details from 8 neighbors of the corresponding point to fill the distortion detail in T_i. When all points in distortion regions are processed, the lost details in distortion boundaries would be repaired.

The shadow removing result with boundary processing is more natural, as shown in Fig. 5(f). Note that, there is rare detail losing and shadow boundary distortion for soft shadows using our method, as shown in Figs. 6 and 7.

4 Experiments

To illustrate the effectiveness of our approach, we present various experimental results for shadow removing and compare with several state-of-the-art shadow removal methods. Our approach was implemented in C++ on a computer with Pentium Dual-Core 2.50GHz CPU and 2GB RAM. The parameters σ_c , σ_l and σ_h are set to 50 in our experiments.

In Fig. 7 we present a shadow removal result for soft shadow image. It can be seen that our approach gets good result in both umbra and penumbra areas. The recovered illumination in shadow region is consistent with surrounding environment and the textures in shadow regions are well retained. Please refer to our supplementary material for more results.

Experiments with traditional methods Guo et al. [10] divide the image into many irregular regions and use the linear mapping model between a matched pair to remove shadows in the image. Different matched pairs use a unified linear mapping model, which produces unnatural shadow-free results for soft shadows, as shown in Fig. 8(e). Moreover, large regions may contain several different kinds of colors and textures, which leads to calculation error for the ratio between direct light and environment light using a unified linear mapping model. This calculation error would generate unsatisfactory shadow-free results, as shown in Fig. 9(c). The same situation presents in Fig. 10(d). But for simple shadows with consistent illumination and unique material in shadow regions, this method can get a good result, as shown the second row in Fig. 10(d). Compared with the limitation of simple shadow scene for method [10], our method can recover the illumination in shadow regions for shadows with inconsistent illumination and multiple materials using the proposed local-to-global strategy.

Xiao et al. [30] apply depth information to remove shadows in a single image. But the inaccurate depth may cause calculation error in the process for shadow detection and shadow removal, which will lead to color and illumination distortion in the shadow removal results, as shown in Fig. 8(f). Moreover, the overall color as shown in the second row of Fig. 8(f) is also changed which is not desirable in the process of shadow removal. But our method can better keep the original tone in the shadow-free result not only in non-shadow regions but also in shadow regions, as shown in Fig. 8(c). Besides, method [30] needs the accurate shadow detection result, while our method only need a rough shadow mask, as shown in Fig. 8(b, d).

Table 1 Available shadow related datasets.

Dataset	Num.	Content of Images	Purpose
SRD[19]	3088	shadow and shadow-free	shadow removal
UIUC[10]	76	shadow and shadow-free	shadow removal
LRSS[9]	37	shadow and shadow-free	shadow removal
SBU[24]	4727	shadow and shadow mask	shadow detection
UCF[36]	245	shadow and shadow mask	shadow detection
ISTD[25]	1870	shadow, shadow mask and shadow-free	shadow detection and removal

Zhang et al. [35] remove the shadows using local patches, which can deal with some soft shadow. But it is limited by the size of the patches, which will lead to unsatisfactory result, especially for shadows with narrow width, as shown in Fig. 9(d). This method also requires an accurate shadow detection result. Our method does not have these restrictions. Moreover, our method can get more visually consistent results, as shown in Fig. 9(f). Both the illumination and texture of our results are very close to those of the ground truth images (Fig. 9(b)).

Arbel et al. [2] consider each image channel as an intensity surface. They fit an intensity surface to the shadow-free surface and obtain an approximate shadow removal result. The smooth intensity surface requires that the shadows are projected into an smooth surface and the surface has uniform material. Otherwise, there may be some artifacts in the shadow-free result, as shown in Fig. 10(e). Nevertheless, the local-to-global strategy used in our method can avoid these problems and obtain high-quality results, as shown in Fig. 10(b).

Experiments with data-driven methods Recently, there are many shadow removal methods based on data-driven fashion. These methods share a common constraint that they depend heavily on the provided training data. These methods would fail when the test shadow image is not well represented in the dataset. Table 1 shows the available shadow datasets. For example, Gryka et al. [9] deal with soft shadows by using a shadow image dataset for soft shadow understanding. With the limitation of dataset, there may be artifacts in some shadow removal results, as shown in Fig. 10(c).

Qu et al. [19] remove shadows using an automatic and end-to-end deep neural network (DeshadowNet).

The shadow-free result for this method is also limited by the training sets. Moreover, their training sets are usually images with simple shadows, for complex shadows, this method may not obtain satisfying shadow-free results, as shown in Fig. 9(e). Some shadows are not removed completely. The difference values presented in Fig. 9 show that our shadow removal results have the smallest differences from ground truth images, compared with [10,19,35].

Images in Fig. 11 show complex shadow images in outdoor and indoor scene. Currently, there is rare such kind of shadow samples in datasets, because the ground truth is hard to acquire. Thus methods [25,11] do not produce desirable results for such complex shadows, as shown in Fig. 11(c, d). Instead, our method can effectively remove shadows in these images and get more visually natural results, as shown in Fig. 11(b).

Quantitative evaluation To evaluate the performance of our method and the related shadow removal methods, we use the shadow and shadow-free image pairs in LRSS and SRD dataset for testing. LRSS dataset contains 37 image pairs. Only a test set of 408 image pairs is publicly available for SRD dataset, so we use the 408 image pairs for testing. We compute the root-mean-square error (RMSE) between recovered result and ground truth in Lab color space as evaluation metric, which directly measures the pixelwise error between the two images. Table 2 reports the RMSE values of different shadow removal methods on LRSS and SRD dataset. We compute RMSE values in shadow and non-shadow regions respectively. The results in Table 2 demonstrate that our shadow removal results have the smallest differences from ground truth images.

The time consumption of the proposed method depends on the size of the shadow regions and the number of the selecting feature sub-regions. For example, the size of the input image in Fig. 5 is 805×684. It takes about 56 seconds for performing multi-scale image decomposition (9 detail levels). We select 5 feature sub-regions in shadow regions. The process for shadow removal takes about 105 seconds, including 21 seconds for feature sub-region matching, 80 seconds for shadow removing in the basic layer image and 4 seconds for the detail repairing.

Limitation Our method also suffers from several limitations. First, the heavy noises may reduce the accuracy of the sub-region matching in the step of shadow removal, which cause the unnatural shadow removal result. In addition, if details in shadow regions are lost, it is difficult for our method to recover detail in the shadow removed results, as shown in Fig. 12. Third, computational cost is currently a computational bottleneck to our algorithm.

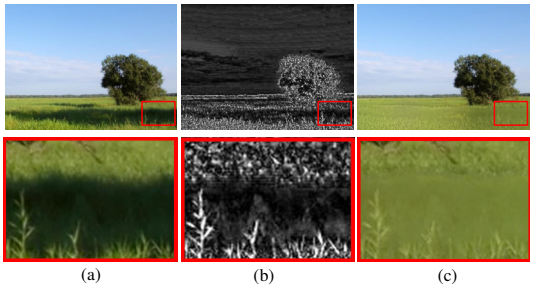


Fig. 12 Failure example with detail lost in shadow regions. (a) is input image. (b) is the detail image, which is magnified 8 times. (c) is the shadow-free result. The second row is the close-up corresponding to the above image.

5 Conclusion

In this paper we have proposed a method to recover the illumination in shadow regions using a local-to-global optimization strategy. We first introduce an illumination-sensitive smoothing method, which works well on shadow image decomposition. Then we extract the significant regions in shadow regions and remove the shadows in the basic layer. By fusing all candidate shadow removal results, we obtain the global optimal result for the basic layer. At last, we reconstruct the details for the shadow-free basic layer image and obtain the final shadow removal result. The experiments demonstrate the effectiveness of the proposed method in indoor and outdoor complex environments.

Shadow detection is an important and challenging problem, especially for large outdoor scenarios with complex soft shadows. Our method distinguishes the shadows from non-shadow regions with rough user interaction. In the future, we could extend our framework to other interesting applications, such as image harmonization and style transfer. In addition, we would facilitate the proposed method for real-time video shadow removal and editing.

Acknowledgment

This work was partly supported by The National Key Research and Development Program of China (2017YFB1002600), the NSFC (No. 61672390), Wuhan Science and Technology Plan Project (No. 2017010201010109), Key Technological Innovation Projects of Hubei Province (2018AAA062), and China Postdoctoral Science Found (No. 070307).

Compliance with Ethical Standards

Conflicts of Interest: The authors have no conflict of interest.

Table 2 RMSE statistics for shadow removal results on LRSS and SRD dataset (smaller is better).

Dataset		Original	Guo[10]	Xiao[29]	Gryka[9]	Qu[19]	Wang[25]	Our
LRSS[9]	Shadow	44.45	31.58	14.94	13.67	14.21	13.09	7.77
	Nonshadow	4.10	13.89	5.07	8.40	4.17	8.29	4.18
SRD[19]	Shadow	42.38	29.89	12.85	–	11.78	12.51	9.02
	Non-shadow	4.56	6.47	5.93	–	4.84	7.33	4.66

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Ssstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on PAMI **34**(11), 2274–2282 (2012)
2. Arbel, E., Hel-Or, H.: Shadow removal using intensity surfaces and texture anchor points. IEEE Transactions on PAMI **33**(6), 1202–1216 (2011)
3. Clarenz, U., Griebel, M., Rumpf, M., Schweitzer, M.A., Telea, A.: Feature sensitive multiscale editing on surfaces. The Visual Computer **20**(5), 329–343 (2004)
4. Darabi, S., Shechtman, E., Barnes, C., Dan, B.G., Sen, P.: Image melding. ACM TOG **31**(4), 1–10 (2012)
5. Finlayson, G.D., Drew, M.S., Lu, C.: Intrinsic images by entropy minimization. In: ECCV, pp. 582–595 (2004)
6. Finlayson, G.D., Hordley, S.D., Drew, M.S.: Removing shadows from images. In: ECCV (4), vol. 2353, pp. 823–836 (2002)
7. Finlayson, G.D., Hordley, S.D., Lu, C., Drew, M.S.: On the removal of shadows from images. IEEE Transactions on PAMI **28**(1), 59–68 (2005)
8. Gangnet, M., Blake, A.: Poisson image editing. In: ACM SIGGRAPH, pp. 313–318 (2003)
9. Gryka, M., Terry, M., Brostow, G.J.: Learning to Remove Soft Shadows. ACM TOG (2015)
10. Guo, R., Dai, Q., Hoiem, D.: Single-image shadow detection and removal using paired regions. In: CVPR, pp. 2033–2040 (2011)
11. Hu, X., Fu, C.W., Zhu, L., Qin, J., Heng, P.A.: Direction-aware spatial context features for shadow detection and removal. In: CVPR (2018)
12. Khan, S.H., Bennamoun, M., Sohel, F., Togneri, R.: Automatic shadow detection and removal from a single image. IEEE Trans PAMI **38**(3), 431–446 (2016)
13. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. IEEE Transactions on PAMI **30**(2), 228–242 (2008)
14. Li, H., Zhang, L., Shen, H.: An adaptive nonlocal regularized shadow removal method for aerial remote sensing images. IEEE Transactions on Geoscience and Remote Sensing **52**(1), 106–120 (2014)
15. Liu, F., Gleicher, M.: Texture-consistent shadow removal. In: ECCV, pp. 437–450 (2008)
16. Matting, S., Chuang, Y.Y., Dan, B.G., Curless, B., Salesin, D.H., Szeliski, R.: Shadow matting and compositing. ACM TOG **22**(3), 494–500 (2003)
17. Mohan, A., Tumblin, J., Choudhury, P.: Editing soft shadows in a digital photograph. IEEE Computer Graphics Applications **27**(2), 23–31 (2007)
18. Pajak, D., Čadík, M., Aydın, T.O., Okabe, M., Myszkowski, K., Seidel, H.P.: Contrast prescription for multiscale image editing. The Visual Computer **26**(6-8), 739–748 (2010)
19. Qu, L., Tian, J., He, S., Tang, Y., Lau, R.W.H.: Deshadownet: A multi-context embedding deep network for shadow removal. In: CVPR, pp. 2308–2316 (2017)
20. Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Computer Graphics Applications **21**(5), 34–41 (2001)
21. Shor, Y., Lischinski, D.: The shadow meets the mask: Pyramid-based shadow removal. In: Computer Graphics Forum, pp. 577–586 (2008)
22. Subr, K., Soler, C.: Edge-preserving multiscale image decomposition based on local extrema. ACM TOG **28**(5), 1–9 (2009)
23. Vicente, T.F.Y., Hoai, M., Samaras, D.: Leave-one-out kernel optimization for shadow detection and removal. IEEE Transactions on PAMI **PP**(99), 1–1 (2018)
24. Vicente, T.F.Y., Hou, L., Yu, C.P., Hoai, M., Samaras, D.: Large-Scale Training of Shadow Detectors with Noisily-Annotated Shadow Examples. Springer International Publishing (2016)
25. Wang, J., Li, X., Hui, L., Yang, J.: Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In: CVPR (2018)
26. Wu, T.P., Tang, C.K.: A bayesian approach for shadow extraction from a single image. In: ICCV, pp. 480–487 (2005)
27. Wu, T.P., Tang, C.K., Brown, M.S., Shum, H.Y.: Natural shadow matting. ACM TOG **26**(2), 8 (2007)
28. Xiao, C., She, R., Xiao, D., Ma, K.L.: Fast shadow removal using adaptive multi-scale illumination transfer. Computer Graphics Forum **32**(8), 207–218 (2013)
29. Xiao, C., Xiao, D., Zhang, L., Chen, L.: Efficient shadow removal using subregion matching illumination transfer. Computer Graphics Forum **32**(7), 421–430 (2013)
30. Xiao, Y., Tsougenis, E., Tang, C.: Shadow removal from single rgb-d images. In: CVPR, pp. 3011–3018 (2014)
31. Yagyu, S., Sakiyama, A., Tanaka, Y.: Edge preserving multiscale image decomposition with customized domain transform filters. In: Signal and Information Processing, pp. 458–462 (2016)
32. Yang, Q., Tan, K.H., Ahuja, N.: Shadow removal using bilateral filtering. IEEE TIP **21**(10), 4361–4368 (2012)
33. Yanli, L., Xavier, G.: Online tracking of outdoor lighting variations for augmented reality with moving cameras. IEEE Transactions on Visualization Computer Graphics **18**(4), 573–580 (2012)
34. Zhang, L., Yan, Q., Liu, Z., Zou, H., Xiao, C.: Illumination decomposition for photograph with multiple light sources. IEEE Transactions on Image Processing **26**(9), 4114–4127 (2017)
35. Zhang, L., Zhang, Q., Xiao, C.: Shadow remover: Image shadow removal based on illumination recovering optimization. IEEE TIP **24**(11), 4623–36 (2015)
36. Zhu, J., Samuel, K.G.G., Masood, S.Z., Tappen, M.F.: Learning to recognize shadows in monochromatic natural images. In: CVPR, pp. 223–230 (2010)