Single image shadow removal by optimization using non-shadow anchor values

Saritha Murali¹ (✉), V. K. Govindan¹, and Saidalavi Kalady¹

© The Author(s) 2019.

Abstract Shadow removal has evolved as a pre-processing step for various computer vision tasks. Several studies have been carried out over the past two decades to eliminate shadows from videos and images. Accurate shadow detection is an open problem because it is often considered difficult to interpret whether the darkness of a surface is contributed by a shadow incident on it or not. This paper introduces a color-model based technique to remove shadows from images. We formulate shadow removal as an optimization problem that minimizes the dissimilarities between a shadow area and its non-shadow counterpart. To achieve this, we map each shadow region to a set of non-shadow pixels, and compute an anchor value from the non-shadow pixels. The shadow region is then modified using a factor computed from the anchor value using particle swarm optimization. We demonstrate the efficiency of our technique on indoor shadows, outdoor shadows, soft shadows, and document shadows, both qualitatively and quantitatively.

Keywords mean-shift segmentation; particle swarm optimization; HSV; YCbCr; anchor values

1 Introduction

Shadowing is a physical phenomenon arising in areas of a surface where the illumination is lower than adjacent areas due to an obstructing opaque object present between the surface and the light source. The geometry and density of shadows vary depending on the number and position of light sources, the

geometry of the obstructing object, and the distance between the object and the surface. For instance, at different time of day, the shadow cast by the Sun on the ground changes dramatically depending on the elevation angle between the Sun and the horizon. Shadows are shorter at noon when the elevation angle is large, and are larger early or late in the day when the elevation angle is small.

The presence of shadows in images and videos reduces the success rate of machine vision applications such as edge extraction, object identification, object counting, and image matching. Some scenarios in which shadows affect different applications are demonstrated in Fig. 1. Image segmentation techniques may incorrectly classify a shadow region as an object or as a portion of an object. This misclassification may in turn affect the detection of objects, as shown in Fig. 1(a) where the part of the pavement in shadow is misclassified as a portion of the brick seat. Shadows may also cause object merging in video tracking systems. Figure 1(b) illustrates a case in which two vehicles are merged and counted as a single object due to the connecting shadow region. Shadows can also obstruct image interpretation in remote-sensing images, and thus prevent target detection (e.g., locating roads), as demonstrated in Fig. 1(c). The challenge posed by illumination inconsistencies, especially shadows, in the lane detection module of driver assistance systems is mentioned in Ref. [1].

Based on their density (i.e., darkness), shadows can be classified into hard shadows and soft shadows. As the name indicates, hard shadows have high density, and the surface texture is nearly destroyed. On the other hand, soft shadows have low density, and their boundaries are usually diffused to the non-shadow surroundings. Furthermore, the shadow generated by

¹ Department of Computer Science and Engineering, National Institute of Technology Calicut, Kerala 673601, India. E-mail: S. Murali, saritha.mkv@gmail.com (✉); V. K. Govindan, vkg@nitc.ac.in; S. Kalady, said@nitc.ac.in.

Manuscript received: 2019-01-29; accepted: 2019-05-18

Fig. 1 Shadows causing problems in (a) object detection, (b) traffic monitoring, and (c) interpretation of remote-sensing images.

obstruction of a non-point light source by an object has two distinct regions, the umbra and penumbra. The umbra is the high density region, lying towards the inner shadow area, while the penumbra is the low density region, lying towards the outer shadow area. Each of these shadow areas is depicted in Fig. 2.

Over the past two decades, numerous studies have been conducted on the detection and removal of shadows appearing in images [2, 3] and videos [4]. Shadow elimination from a single image is more challenging since the only source of information available is that single image, and there is no context. Removal of shadows is often considered to be a relighting task in which the brightness of shadow pixels is increased to make them as well illuminated as the non-shadow surroundings. In the survey [5], shadow removal techniques are categorized into reintegration methods [6], relighting methods [7], patch-based methods [8], and color transfer methods [9]. Furthermore, these techniques are either automatic [7] or interactive [10], depending on whether the user is provided an interface to incorporate his knowledge into the system. Various works related to shadow removal are presented in the next section.

The rest of this paper is structured as follows. A concise view of the state-of-the-art shadow removal techniques is presented in Section 2. Section 3 briefly discusses background information related to the proposed work. Section 4 details our proposed shadow elimination framework. Experimental results and comparison with other shadow removal algorithms are included in Section 5. Section 6 discusses applications and limitations of the proposed method, and Section 7 concludes the paper with some observations on possible future enhancements to the proposed method.

2 Previous work

Several methods are available for shadow elimination from indoor images, natural images, satellite images, and videos. Early works on shadow removal from an image are based on the assumption that camera calibration is necessary to generate a shadow-free representation, called an invariant image. Finlayson et al. [6] introduced a method to develop a shadow invariant by projecting the 2D chromaticity perpendicular to the illumination-varying direction. Another way to construct the invariant image is

Fig. 2 Types of shadow: (a) hard shadow, (b) soft shadow, (c) shadow regions, umbra and penumbra.

by minimizing entropy [11], which does not need calibrated cameras. Shadow edges are determined using edge detection on both original and invariant images. Image gradient along the shadow edges is then replaced by zeroes, followed by integration to generate the shadow-free image. These techniques require high-quality input images; however. Yang et al. [12] derived a 3D illumination invariant image from a 2D illumination invariant image and joint bilateral filtering. The detail layer from this 3D illumination invariant image is then transferred to the input shadow image to obtain the shadow-free image.

Baba et al. [13] used brightness information to segment an image into lit, umbra, and penumbra regions. Shadow regions are modified by adjusting the brightness and color. This method applies only to images of single textures. Features that differentiate shadow regions from their non-shadow counterparts are used for shadow detection in monochromatic images by Zhu et al. [14]. A classifier trained using features such as intensity, local maximum, smoothness, skewness, gradient and texture measures, entropy, and edge response is used for shadow identification. In Ref. [15], monochromatic shadows are corrected in the gradient domain by fitting a Gaussian function for each vertical and horizontal line segment across the shadow boundary. Arbel and Hel-Or [16] showed that shadows could be cleared by adding them to a constant value in log space. Guo et al. [7] also used a scaling constant to relight shadow pixels. They used a region-based technique to detect shadows in natural images. A single-region classifier and a pairwise classifier are trained using support vector machines (SVMs). An energy function combining these classifier predictions is then minimized using graph cut. Such region-based shadow removal is shown to provide better results than pixel-based methods. In Ref. [8], the best matching shadow matte is estimated for each shadow patch using a Markov random field. This matte is separated from the image, leading to a shadow-free output. The technique proposed by Sasi and Govindan [17] splits and merges adjacent homogeneous quadtree blocks in the image based on a fuzzy predicate built using entropy, edge response, standard deviation, and mean of quadtree blocks.

Interactive shadow detection techniques allow users

to incorporate their knowledge into the detection task. The inputs may be a quad map [9], or rough strokes [18]. The work proposed by Gong and Cosker [10] requires rough scribbles in the shadow and non-shadow regions. They unwrap the penumbra boundary and perform multi-scale smoothing to derive shadow scales. A recent method by Yu et al. [19] requires the user to provide strokes on the shadow and non-shadow regions. The shadow scales are estimated using statistics of shadow-free regions, followed by illumination optimization. Certain techniques use color models which can separate illumination and reflectance components of the image, such as normalized RGB, c1c2c3, CIE $L^*a^*b^*$, and HSV to detect shadows [20]. Detection using these methods mostly leads to false labeling since a pixel is classified as shadow or non-shadow without considering the neighbors.

Vicente et al. [21] mapped each shadow region to the corresponding lit region using a trained probabilistic SVM classifier. Each shadow region is then enhanced by matching the histogram of the shadow region with the lit region. Recent techniques for shadow detection use deep learning to learn shadow features automatically. Khan et al. [22] used deep learning for local and across-boundary features using two convolutional deep neural networks (CNNs). The most significant structural features of shadow boundaries are learned using a CNN in Ref. [23]. Although the accuracy of shadow detection using feature learning is better than for non-learning approaches, the time and memory requirements for training the classifier are usually high. Shadow removal results may show more noise in recovered areas than surrounding non-shadow areas. In such cases, noise removal [24] is introduced as a post-processing step to provide a seamless transition between shadow and lit areas.

This paper presents a simple, color-model based technique to remove shadows from an image. Shadow removal is modeled as an optimization task. Every pixel in a shadow segment is modified based on an anchor value computed from a set of non-shadow pixels mapped to the segment. In this way, each shadow segment is processed to acquire the final shadow-free image. The method is suitable for both hard shadows and soft shadows. Also, it recovers the umbra and penumbra regions by finding different

scale factors for these regions, and there is no need for a separate method to classify shadow pixels to umbra pixels or penumbra pixels.

The next section briefly introduces the color spaces used in the proposed method, the mean-shift segmentation algorithm, and the particle swarm optimization (PSO) technique.

3 Background

3.1 Color spaces

Shadow regions are usually darker than their surroundings. Therefore, the creation of shadows is often attributed to the change in illumination. In RGB color space, a pixel is represented using its intensity in three different color channels, red, green, and blue, whereas in YCbCr color space, a pixel is represented in terms of its brightness and two color-difference values. To extract the illumination information from the color information, we transform the RGB image into YCbCr color space. The Y channel in YCbCr is the illumination, Cb is the blue difference, and Cr is the red difference.

In order to map shadow regions to non-shadow regions, we use HSV (hue, saturation, and value) color space. For any pixel, hue represents the basic color of the pixel, saturation represents the grayness, and value represents the brightness. Since hue is unaffected by changes in illumination, we assume that shadows, which are caused by changes in illumination, do not alter the hues of pixels.

3.2 Mean-shift segmentation

Our shadow removal system requires the input shadow image to be segmented into constituent regions. For this purpose, we use a non-parametric clustering technique called mean-shift clustering, introduced by Fukunaga and Hostetler [25]. In this, we cluster pixels in RGB space. Pixels with similar colors in RGB color space are gathered into the same cluster. Pixels are recolored according to their cluster, resulting in the segmented image.

The mean-shift method models feature points using a probability density function where local maxima or modes correspond to dense areas. Gradient ascent is performed on the estimated density of each data point until the algorithm converges. All data points belonging to a particular stationary point are grouped into a single cluster. The mean-shift algorithm can

be summarized as follows:

- For each sample t_i , find the mean-shift vector $M(t_i)$.
- Shift the density estimation window from t_i to $t_i + M(t_i)$.
- Repeat until the samples reach equilibrium (i.e., convergence).

We use the Edge Detection and Image Segmentation System (EDISON) developed in the Robust Image Understanding Laboratory at Rutgers University [26–28] to perform mean-shift segmentation.

3.3 Particle swarm optimization

In our work, the removal of shadows is achieved by solving an optimization problem. The objective of this optimization problem is to maximize the similarity between a shadow area and its non-shadow counterpart. It is solved using particle swarm optimization (PSO) [29]. PSO was developed to model the social behavior of an animal in its group, for instance, the movement of each bird within a flock. The algorithm starts with a group of possible solutions called particles and moves these particles in the space based on their velocity and position. The position of a particle depends on the best position of the individual particle, and of the group. This is repeated until a solution is obtained.

4 Shadow removal by optimization

The central objective of shadow removal techniques is to achieve a seamless transition between the shadow and its surroundings by reducing the difference in intensities of the shadow area and its surroundings. Our novel approach to removing shadows by minimizing this difference is detailed in this section. Figure 3 depicts the workflow of the proposed method for shadow removal. The input to our system is a color image containing shadowed as well as non-shadowed areas. Shadow regions are detected using the method in Ref. [7]. Image segmentation is done using mean-shift segmentation [27], and each segment is given a unique segment identifier (*SID*).

Our shadow removal algorithm is divided into four stages: (i) decomposition of the image into frameworks, (ii) initial shadow enhancement, (iii) shadow to non-shadow mapping, and (iv) shadow

Fig. 3 Overall workflow of our proposed shadow removal framework: (a) input shadow image, (b) input image in YCbCr space, (c) mean-shift segmentation of input image, (d) image decomposed into frameworks, (e) binary shadow mask with shadow regions in white, (f) shadow segments mapped to non-shadow regions, (g) YCbCr image after shadow removal, and (h) final shadow-free output in RGB space.

correction. Each of these steps is detailed in the subsections that follow.

4.1 Splitting the image into frameworks

Barrow and Tenenbaum [30] proposed that the formation of any Lambertian image (I) can be modeled using:

$$I = R \circ L \quad (1)$$

I is the Hadamard product ($\circ$) of the reflectance component (R) and the illumination component (L). Since shadows occur in less illuminated areas, we segment the image into different frameworks based on the pixel illumination, so that each framework contains pixels with similar illumination. For instance, pixels belonging to hard shadow on a surface would constitute a framework. The concept of splitting the image into frameworks is adopted from the tone-mapping method for high dynamic range (HDR) images proposed in Ref. [31].

Initially, the $\log_{10}$ of the input image illumination is clustered using the k -means algorithm. For this purpose, the centroids are initialized to values starting from the lowest to the highest illumination of the intensity image with a step size of ten units. The

final set of centroids obtained after convergence of the algorithm represents the frameworks. These frameworks are further refined by discarding the centroids with zero pixel membership. To restrict distinct frameworks with very close illumination, centroid pairs with a difference of less than 0.05 are iteratively merged. If C_i and C_{i+1} are two adjacent centroids to be merged, with N_i and N_{i+1} pixels in each respectively, the new centroid value C_j on merging C_i and C_{i+1} is given by

$$C_j = \frac{C_i N_i + C_{i+1} N_{i+1}}{N_i + N_{i+1}} \quad (2)$$

For a pixel at position i , the probability of its membership of each framework is computed as the Gaussian of the difference of pixel illumination (Y_i) and the centroid representing the framework (C_j). A pixel is classified into the framework for which its probability is maximum. Let F_i denote the framework to which the i th pixel belongs. Then the probability of the i th pixel belonging to the j th framework is given by

$$Pr\{F_i = j\} = e^{-(C_j - Y_i)^2 / 2\sigma^2} \quad (3)$$

where σ denotes the maximum distance between any two adjacent centroids. Frameworks that contain

no pixels with a probability greater than 0.3 are removed by combining them with adjacent frameworks. Finally, the framework to which each pixel belongs is recomputed as mentioned above.

4.2 Initial shadow correction

An initial enhancement of the shadow regions is performed as a pre-processing step to aid the shadow to non-shadow mapping done in the next phase of our algorithm. In this step, the shadow regions are roughly enhanced to make them better match the non-shadow equivalents. At first, unwanted texture details are removed by applying a mean filter of window size 11×11 to the input shadow image. Further, each shadow pixel is enhanced by using a factor computed from the pixels constituting the dominant framework (the one with the most pixels). Let P indicate the pixels belonging to this framework, and $\bar{R}_P$, $\bar{G}_P$, and $\bar{B}_P$ indicate the mean values of these pixels in each color plane in the filtered image Z . Let X denote the color plane for which the mean value is the lowest. Then the factor f is computed as

$$f = 2 - \frac{\bar{Z}_X^{\text{shadow}}}{\min(\bar{R}_P, \bar{G}_P, \bar{B}_P)} \quad (4)$$

Shadow pixels are then modified using:

$$Z^{\text{shadow}} = f \cdot Z^{\text{shadow}} \quad (5)$$

The output image from this stage is used for mapping shadow regions to corresponding non-

shadow regions in the next step.

4.3 Shadow to non-shadow mapping

The objective of this step is to find the matching set of non-shadow pixels for each shadow segment. An image can contain several disconnected shadow regions, and a single shadow region can contain multiple segments in the segmentation result. Let SID_S and SID_N indicate the segment IDs of all shadow pixels, and the segment IDs of all non-shadow pixels, respectively. Each shadow segment ($unique(SID_S)$) is processed separately in our method. For this purpose, the image obtained from the initial shadow correction in the previous stage is converted into HSV color space. The shadow segments are mapped to a set of non-shadow pixels based on hue-matching. The hue-matching procedure is described in Algorithm 1.

4.4 Shadow removal

Once the shadow to non-shadow mappings are available, shadow removal is done by modifying the Y, Cb, and Cr planes of the input image. For each mapping, we select the framework to which the most non-shadow pixels in this mapping belongs. All non-shadow pixels in this framework are selected as candidate pixels. This is followed by computation of an anchor value (A) from the candidate pixels by selecting the highest intensity value after discarding

Algorithm 1 Hue-matching for shadow to non-shadow mapping

Input: initial shadow-corrected image in HSV, segmented image

Output: shadow to non-shadow mapping

- 1: Cluster the pixels of HSV image into 37 bins of equal width ($B^i (i = 1, \dots, 37)$) based on their hue values (0 to 360).

$$B^i = B_S^i \cup B_N^i \quad (6)$$

Here, B_S^i : shadow pixels in B^i , and B_N^i : non-shadow pixels in B^i . B^i , B_S^i , and B_N^i can be $\emptyset$.

- 2: **for** each $y \in unique(SID_S)$ **do**
 - 3: Sort the bins in descending order of number of shadow pixels with $SID = y$ in each bin, i.e., $n(y, SID(B_S^i))$
 - 4: **repeat**
 - 5: Select a bin B^i from the sorted list.
 - 6: Compute a weight W_x for each $x \in unique(SID(B_N^i))$.

$$W_x = n(x, SID(B_N^i)) + n(x, SID(surr_S)); \quad (7)$$

$n(x, SID(B_N^i))$: number of non-shadow pixels with $SID = x$ in bin B^i ;
 $n(x, SID(surr_S))$: number of non-shadow pixels with $SID = x$ in a dilated neighborhood of shadow segment S .
 - 7: Sort the weights in descending order.
 - 8: Iteratively add the non-shadow pixels belonging to the segment with maximum weights into the set N_y until the number of selected non-shadow pixels is at least 10% the size of the shadow segment.
 - 9: **until** $n(N_y) \geq 0.1 \cdot n(y, SID_S)$
 - 10: **end for**
 - 11: N_y is the final mapped set of non-shadow pixels for each shadow segment y .
-

the top 5% intensities in each plane [32]. In this way, separate anchor values are computed for the Y, Cb, and Cr planes. These anchor values are the non-shadow intensities to which the shadow intensities can be related. The shadow pixels in each segment are modified in the Y, Cb, and Cr planes as follows:

$$I_i^{\text{new}} = I_i^{\text{old}} + \frac{A}{I_i^{\text{old}}} M \quad (8)$$

where I_i^{old} and I_i^{new} indicate the old and modified value of the i th shadow pixel, respectively. M is a scale factor computed using particle swarm optimization by minimizing the deviation of each shadow pixel from the mean candidate pixel value, subject to the following constraints:

$$|\mu_c - \mu_s| \leq 0 \quad (9)$$

$$|\sigma_c - \sigma_s| \leq 0.4 \quad (10)$$

where s and c indicate the shadow segment and the mapped set of candidate pixels, respectively. The threshold of 0.4 for Eq. (10) was computed by including all images present in the dataset given by Guo et al. [7]. These constraints ensure that shadow segments are modified such that their mean values become closer to those of the non-shadow regions. At this stage, certain pixels, particularly those near the shadow boundary, may be over-corrected. To address this issue, the shadow pixels for which the modified value exceeds the anchor value in the illumination plane are considered as a separate segment and are processed separately. The modified image in log space is finally converted into intensity space, and then to RGB.

5 Experimental results

This section presents results of our shadow removal technique and a comparison with some of the available methods for shadow removal from images. Implementation of the algorithm and all the experimentation was done in MATLAB R2017a on an Intel Core i5-4570 CPU @ 3.20 GHz with 8 GB RAM. For evaluation, the 76 images from the UIUC dataset [7] were divided into two classes based on the method used by the authors to generate the shadowless ground truth images. Set A consists of 46 images for which the ground truth is created by removing the object causing shadow. Set B consists of 30 images for which the ground truth is captured by removing the light source responsible for the appearance of

shadow. We also evaluated the performance of the proposed shadow removal method on images from the ISTD [33] and LRSS datasets [8].

5.1 Qualitative analysis

To illustrate the visual quality of the output of the proposed method, various examples are included in Fig. 4 and Fig. 5. The figures also include the outputs obtained using other shadow removal methods. In Fig. 4, the results displayed are for five images from the UIUC dataset; the first three images are from Set A, and the next two images are from Set B. For each example, the input shadow image and the corresponding shadow-free ground truth are included in rows (a) and (i), respectively. The shadow region of the input image in Fig. 4(i)(a) has two different segments—one with a yellow background, and the other with a gray background. Most existing methods fail to handle the individual shadow segments, which is clear from the images in Figs. 4(i)(c–f). Our proposed method handles each shadow segment separately, and the result is better than those of other methods. Figure 4(ii) also illustrates a similar case. Guo et al. [7] and Yu et al. [19] were unable to handle the shadow on the yellow lines. Zhang et al. [34] and Gryka et al. [8] failed to remove the shadow. The outputs obtained by Su and Chen [35] and Gong and Cosker [10] have artifacts near the lower region of the shadow.

Although the results in Fig. 4(iii) appear to be good for most of the methods, several visual artifacts can be seen on close observation. Figure 4(iv)(b) demonstrates the scenario in which dark regions in an object are misclassified as shadows by Guo et al. [7], resulting in undesirable modification of these non-shadow regions. Gryka et al. [8] was unable to present a visually pleasing result for the same image. Even though the shadow appears to be removed in the image in Fig. 4(v)(b), the text on the box is lost. Also, the attempt to remove shadows using Ref. [34] results in the false coloring of the shadow region, as shown in Fig. 4(v)(d). The shadow-removal results of the state-of-the-art methods and the proposed method for the 76 images in the UIUC dataset are included in the Electronic Supplementary Material (ESM).

Figure 5 compares the shadow removal results of various methods on 450 images from the ISTD dataset. Row (a) shows the input shadow images, while row (f) shows the corresponding shadow-free ground truths.

	(i)	(ii)	(iii)	(iv)	(v)
(a) Input shadow image					
(b) Guo et al. [7]					
(c) Su and Chen [35]					
(d) Zhang et al. [34]					
(e) Gryka et al. [8]					
(f) Yu et al. [19]					
(g) Gong and Cosker [10]					
(h) Proposed method					
(i) Ground Truth					

Fig. 4 Comparison of shadow removal techniques using images from the UIUC dataset. Row (a) shows the original shadow images. Row (i) shows the corresponding shadow-free ground truths. Rows (b–h) contain results using the approach named in the first column.

Figure 5(i) illustrates the efficiency of the proposed method in removing hard shadows. The image in Fig. 5(ii)(a) has a dark region which is misclassified as shadow region by the methods in rows (b–d). The soft shadow in Fig. 5(v)(a) is efficiently removed by our method. For all the input shadow images, our proposed method gives better results than other state-of-the-art shadow removal techniques.

5.2 Quantitative analysis

This section analyzes the accuracy of the proposed method in removing shadows from images using the

three different metrics described below.

5.2.1 Root mean-squared error (RMSE)

RMSE is the square root of the mean of the squared error. The error for an image is computed as the per-pixel difference between the shadow-free ground truth and the shadow removal output. Let G_i and O_i indicate the intensity of the i th pixel in the ground truth and the output images, respectively. Then the RMSE of an n -pixel image is computed by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (G_i - O_i)^2} \quad (11)$$

Fig. 5 Comparison of shadow removal techniques using images from the ISTD dataset. Row (a) has the original shadow images. Row (f) has the corresponding shadow-free ground truths. Rows (b–e) contain results using the approach named in the first column.

Since RMSE is an error measure, a lower value indicates that the output image is more similar to the ground truth.

5.2.2 Peak signal to noise ratio (PSNR)

PSNR is used to find peak error. It computes the peak value of SNR, in decibels, between two images. The value of PSNR ranges between 0 and infinity. A higher PSNR value signifies better quality of the output image. Initially, we compute the mean-squared error (MSE) between the shadow-free ground truth and the shadow removal output as below:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (G_i - O_i)^2 \quad (12)$$

where G_i and O_i indicate the intensity of the i th pixel in the ground truth and output images, respectively. Then the PSNR is computed using:

$$\text{PSNR} = 10 \log\left(\frac{\eta^2}{\text{MSE}}\right) \quad (13)$$

where η denotes the maximum possible pixel value in the images.

5.2.3 Structural similarity index (SSIM)

SSIM is a metric used to compute the correlation between two images. It is considered to be a measure that is closest to the human visual system. The SSIM formula given in Eq. (14) is a weighted combination of comparisons of three elements between the images, namely luminance, contrast, and structure:

$$\text{SSIM}(G, O) = \frac{(2\mu_G\mu_O + c_1)(2\sigma_{GO} + c_2)}{(\mu_G^2 + \mu_O^2 + c_1)(\sigma_G^2 + \sigma_O^2 + c_2)} \quad (14)$$

where G and O are the shadow-free ground truth and the shadow removal output, respectively, μ_G and μ_O are the mean values of G and O , σ_G^2 and σ_O^2 are the variances of G and O , and σ_{GO} is the covariance of G and O . c_1 and c_2 are variables computed using the dynamic range of pixels.

The average values of RMSE, PSNR, and SSIM for the two sets of images in the UIUC dataset are tabulated in Table 1. We have computed separate RMSE for shadow areas, non-shadow areas, and

the entire image. The average values of RMSE, PSNR, and SSIM of the original shadow image with the corresponding ground truths are included in the first row of each table. Imperfect shadow detection has resulted in non-zero RMSE in the non-shadow regions of the original shadow images and the corresponding ground truths. Another reason for this is the misalignment of the input images and the ground truths. Variations may also arise due to changes in the direct and ambient light reaching the non-shadow surface on removal of the shadow source. The values in Table 1 indicate that our results are superior to those of other methods.

The reason for partitioning the images in the UIUC dataset into Set A and Set B can be explained by comparing the first rows in Table 1(a) and Table 1(b). The shadow-free ground truths for Set B images are darker than their corresponding shadow images since these ground truths are generated by eliminating the

light sources causing the shadow (e.g., see Fig. 4(v)). Hence, the shadow region of the ground truth is more similar to the shadow region of the original image than the results of various shadow removal techniques. This explains the lower shadow region RMSE, and the higher non-shadow region RMSE of images in Set B for the original shadow images. This reason also explains the higher PSNR and SSIM values.

Quantitative results on 450 images from the ISTD dataset [33] are included in Table 2. The values in the table indicate that the proposed method performs better than other state-of-the-art shadow removal techniques.

Also, to analyze the effectiveness of our shadow removal approach, we conducted experiments on images with soft shadows. A method to remove soft shadows from images was proposed in Ref. [8]. A quantitative analysis done on 15 images from their dataset is included in Table 3. The shadow-free

Table 1 Quantitative results on the images from the UIUC dataset [7]. (a) Images for which the ground truth is created by removing the object causing shadow. (b) Images for which the ground truth is created by removing the light source responsible for the appearance of the shadow. Values are averages computed over the set of images

(a) 46 images (Set A) in UIUC dataset [7]					
Method	All region RMSE	Shadow RMSE	Non-shadow RMSE	PSNR	SSIM
Original shadow image	8.8173	17.3812	2.8577	17.7652	0.9030
Su and Chen [35]	4.9956	6.1427	4.5471	22.7521	0.9692
Guo et al. [7]	4.7058	6.5041	3.4174	24.7683	0.9694
Zhang et al. [34]	4.8693	6.6691	4.1972	23.9528	0.9678
Gong and Cosker [10]	4.1367	5.9933	3.0721	25.2210	0.9769
Gong and Cosker [18]	4.1087	5.2053	3.5888	25.4569	0.9775
Yu et al. [19]	4.6588	5.9763	3.9566	24.5882	0.9730
Proposed method	4.0399	5.1585	2.9574	25.6460	0.9871
(b) 30 images (Set B) in UIUC dataset [7]					
Method	All region RMSE	Shadow RMSE	Non-shadow RMSE	PSNR	SSIM
Original shadow image	13.8101	3.1606	16.3750	14.4300	0.7512
Su and Chen [35]	18.7457	15.0770	19.8418	11.6994	0.6342
Guo et al. [7]	16.3436	12.5870	17.8478	13.0993	0.6857
Zhang et al. [34]	18.1892	16.9061	18.4132	12.0023	0.6363
Gong and Cosker [10]	15.9808	13.1515	16.8974	13.2526	0.6845
Gong and Cosker [18]	16.0947	13.4348	17.1333	13.2527	0.6832
Yu et al. [19]	16.8660	15.9963	17.2683	12.8850	0.6692
Proposed method	16.1408	11.9644	16.5986	13.2843	0.6859

Table 2 Quantitative results for 450 images from the ISTD dataset [33]. The values are averages computed over the set of images

Method	All region RMSE	Shadow RMSE	Non-shadow RMSE	PSNR	SSIM
Original shadow image	6.1766	13.7189	3.4580	6.5062	0.9518
Yang et al. [12]	7.2868	7.7342	7.1477	6.5163	0.9438
Yu et al. [19]	4.9312	5.9729	4.6248	5.9612	0.9710
Wang et al. [33]	3.3029	3.7225	3.1293	6.6833	0.9832
Proposed method	3.0317	3.2869	3.0466	7.0744	0.9859

Table 3 Quantitative results on 15 images from the soft shadows dataset given by Gryka et al. [8]. The values are averages computed on the set of images

Method	All region RMSE	Shadow RMSE	Non-shadow RMSE	PSNR	SSIM
Original shadow image	9.5544	17.5722	3.6759	17.2982	0.8757
Arbel and Hel-Or [16]	4.8108	7.3165	3.3282	25.1363	0.9715
Guo et al. [7]	5.2438	9.6300	3.3381	23.8182	0.9609
Gryka et al. [8]	3.0229	4.1910	2.1896	27.5319	0.9877
Yu et al. [19]	3.8492	4.5063	3.3401	25.5477	0.9791
Proposed method	2.9399	4.0471	2.1600	27.7551	0.9894

ground truths and the results using other methods were made available by the authors. The average values of RMSE, PSNR, and SSIM in the table indicate that the proposed method can be used to remove soft shadows from images efficiently.

6 Discussion

6.1 Need for initial shadow correction

Figure 6 demonstrates the need for initial shadow correction as a pre-processing step to improve the accuracy of shadow to non-shadow mapping. The left column (Figs. 6(a)–6(c)) illustrates the case in which the input image itself is used for hue-matching

to find the shadow to non-shadow mapping. The second column (Figs. 6(d)–6(f)) uses the initial shadow corrected image for shadow to non-shadow mapping. It is evident from the figure that the shadow region is mapped to the wrong non-shadow region in the first case. This is because the hue values of shadow pixels are a closer match to those of the dark non-shadow region in the image. In the shadow corrected image, the illumination of the shadow region is roughly enhanced and hence the hue value of the shadow region is matched to the lighter non-shadow surroundings. Also, smoothing the unwanted texture details restricts the hue values in the shadow region from being distributed over a wide range of bins.

6.2 Applications

The experimental results in the previous section show the efficiency of the proposed method in removing shadows from indoor and outdoor images. Figure 7 demonstrates the capability of our shadow removal algorithm in removing shadows from document images. Oliveira et al. [36] and Bako et al. [37] developed shadow removal methods explicitly for document images, and Gong and Cosker [18] for general images. The results of the methods mentioned in the figure were obtained from Ref. [37]. It is evident from Fig. 7(e) that our proposed system can process shadows on document images, and the result is comparable with that of Bako et al. [37].

6.3 Limitations

Some limitations of the proposed shadow removal method are shown in Fig. 8. The input shadow image in Fig. 8(a) has an object that is completely covered by shadow. In this case, the proposed method fails to find a matching non-shadow region for that object. This results in the object being mapped to the road and hence the result in Fig. 8(a)(iv).

Figure 8(b) shows an image with shadow on a complex textured surface. The segmentation of this

Fig. 6 Need for initial shadow correction.

Fig. 7 Shadow removal from document images.

image produces an image in which very small objects on the ground are clustered into the same segment even though their colors (hue) are different. Also, hue-matching leads to an incorrect shadow to non-shadow mapping in this case, thereby resulting in the output shown in Fig. 8(b)(iv).

Detection of soft shadows with a highly diffused penumbra is a challenging task. Figure 8(c) illustrates the effect of the shadow detection and image segmentation results of the proposed method on soft shadow images with a highly diffused penumbra. The penumbra region is not properly detected leading to improper shadow correction output.

(i) Input shadow image			
(ii) Segmented image			
(iii) Shadow detection input			
(iv) Output of proposed method			
	(a)	(b)	(c)

Fig. 8 Failures of the proposed shadow removal method for (a) object completely in shadow, (b) shadow on a complex textured surface, and (c) soft shadow with highly diffuse penumbra.

7 Conclusions

Shadows are natural phenomenon that may lead to the incorrect execution of certain tasks in computer vision, such as object detection and counting. We have presented a simple, color-model based method to

remove shadows from an image. Shadow regions are initially mapped to corresponding non-shadow regions using hue-matching. This is followed by removal of shadows using a scale factor computed using particle swarm optimization and an anchor value from the non-shadow region. This method does not require prior training, and works on both hard shadows and soft shadows. Moreover, it recovers the umbra and penumbra regions by finding separate scaling factors for these regions. The umbra and penumbra regions are handled separately, without requiring a classifier to distinguish them. Furthermore, unlike many other existing methods, our method does not alter pixels in non-shadow areas.

Our method was evaluated on various datasets, and the results quantitatively analyzed. Also, the subjective quality of the output images was illustrated using selected examples. The method works for both indoor and outdoor images, and for images under various lighting conditions. This method does not require high-quality inputs or accurate image segmentation. However, the time taken to process an image grows with the number of segments and the time needed for the PSO to converge. Also, the shadow to non-shadow mapping may lead to incorrect results if a non-shadow equivalent of a shadow segment does not exist.

Electronic Supplementary Material Supplementary material is available in the online version of this article at <https://doi.org/10.1007/s41095-019-0148-x>.

References

- [1] Niu, J. W.; Lu, J.; Xu, M. L.; Lv, P.; Zhao, X. K. Robust lane detection using two-stage feature extraction with curve fitting. *Pattern Recognition* Vol. 59, 225–233, 2016.
- [2] Sasi, R. K.; Govindan, V. K. Shadow detection and removal from real images. In: *Proceedings of the 3rd*

- International Symposium on Women in Computing and Informatics, 309–317, 2015.
- [3] Murali, S.; Govindan, V. K.; Kalady, S. A survey on shadow detection techniques in a single image. *Information Technology and Control* Vol. 47, No. 1, 75–92, 2018.
 - [4] Sanin, A.; Sanderson, C.; Lovell, B. C. Shadow detection: A survey and comparative evaluation of recent methods. *Pattern Recognition* Vol. 45, No. 4, 1684–1695, 2012.
 - [5] Murali, S.; Govindan, V. K.; Kalady, S. A survey on shadow removal techniques for single image. *International Journal of Image, Graphics and Signal Processing* Vol. 8, No. 12, 38–46, 2016.
 - [6] Finlayson, G. D.; Hordley, S. D.; Drew, M. S. Removing shadows from images. In: *Computer Vision — ECCV 2002. Lecture Notes in Computer Science, Vol. 2353*. Heyden, A.; Sparr, G.; Nielsen, M.; Johansen, P. Eds. Springer Berlin Heidelberg, 823–836, 2002.
 - [7] Guo, R. Q.; Dai, Q. Y.; Hoiem, D. Paired regions for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 35, No. 12, 2956–2967, 2013.
 - [8] Gryka, M.; Terry, M.; Brostow, G. J. Learning to remove soft shadows. *ACM Transactions on Graphics* Vol. 34, No. 5, Article No. 153, 2015.
 - [9] Wu, T. P.; Tang, C. K. A Bayesian approach for shadow extraction from a single image. In: *Proceedings of the 10th IEEE International Conference on Computer Vision*, Vol. 1, 480–487, 2005.
 - [10] Gong, H.; Cosker, D. Interactive removal and ground truth for difficult shadow scenes. *Journal of the Optical Society of America A* Vol. 33, No. 9, 1798–1811, 2016.
 - [11] Finlayson, G. D.; Drew, M. S.; Lu, C. Entropy minimization for shadow removal. *International Journal of Computer Vision* Vol. 85, No. 1, 35–57, 2009.
 - [12] Yang, Q. X.; Tan, K. H.; Ahuja, N. Shadow removal using bilateral filtering. *IEEE Transactions on Image Processing* Vol. 21, No. 10, 4361–4368, 2012.
 - [13] Baba, M.; Mukunoki, M.; Asada, N. Shadow removal from a real image based on shadow density. In: *Proceedings of the ACM SIGGRAPH 2004 Posters*, 60, 2004.
 - [14] Zhu, J. J.; Samuel, K. G. G.; Masood, S. Z.; Tappen, M. F. Learning to recognize shadows in monochromatic natural images. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 223–230, 2010.
 - [15] Xu, M. L.; Zhu, J. J.; Lv, P.; Zhou, B.; Tappen, M. F.; Ji, R. R. Learning-based shadow recognition and removal from monochromatic natural images. *IEEE Transactions on Image Processing* Vol. 26, No. 12, 5811–5824, 2017.
 - [16] Arbel, E.; Hel-Or, H. Shadow removal using intensity surfaces and texture anchor points. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 33, No. 6, 1202–1216, 2011.
 - [17] Sasi, R. K.; Govindan, V. K. Fuzzy split and merge for shadow detection. *Egyptian Informatics Journal* Vol. 16, No. 1, 29–35, 2015.
 - [18] Gong, H.; Cosker, D. User-assisted image shadow removal. *Image and Vision Computing* Vol. 62, 19–27, 2017.
 - [19] Yu, X. M.; Li, G.; Ying, Z. Q.; Guo, X. Q. A new shadow removal method using color-lines. In: *Computer Analysis of Images and Patterns. Lecture Notes in Computer Science, Vol. 10425*. Felsberg, M.; Heyden, A.; Krüger, N. Eds. Springer Cham, 307–319, 2017.
 - [20] Murali, S.; Govindan, V. K. Shadow detection and removal from a single image using LAB color space. *Cybernetics and Information Technologies* Vol. 13, No. 1, 95–103, 2013.
 - [21] Vicente, T. F. Y.; Hoai, M.; Samaras, D. Leave-one-out kernel optimization for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 40, No. 3, 682–695, 2018.
 - [22] Khan, S. H.; Bennamoun, M.; Soheli, F.; Togneri, R. Automatic shadow detection and removal from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 38, No. 3, 431–446, 2016.
 - [23] Shen, L.; Chua, T. W.; Leman, K. Shadow optimization from structured deep edge detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2067–2074, 2015.
 - [24] Xu, M.; Lv, P.; Li, M.; Fang, H.; Zhao, H.; Zhou, B.; Lin, Y.; Zhou, L. Medical image denoising by parallel non-local means. *Neurocomputing* Vol. 195, 117–122, 2016.
 - [25] Fukunaga, K.; Hostetler, L. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory* Vol. 21, No. 1, 32–40, 1975.
 - [26] Christoudias, C. M.; Georgescu, B.; Meer, P. Synergism in low level vision. In: *Proceedings of the 16th International Conference on Pattern Recognition*, Vol. 4, 40150, 2002.
 - [27] Comaniciu, D.; Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 24, No. 5, 603–619, 2002.
 - [28] Meer, P.; Georgescu, B. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 23, No. 12, 1351–1365, 2001.

- [29] Eberhart, R.; Kennedy, J. A new optimizer using particle swarm theory. In: *Proceedings of the 6th International Symposium on Micro Machine and Human Science*, 39–43, 1995.
- [30] Barrow, H.; Tenenbaum, J. M. Recovering intrinsic scene characteristics. In: *Computer Vision Systems*. Hanson, A.; Riseman, E. Eds. New York: Academic Press, 3–26, 1978.
- [31] Krawczyk, G.; Myszkowski, K.; Seidel, H. P. Lightness perception in tone reproduction for high dynamic range images. *Computer Graphics Forum* Vol. 24, No. 3, 635–645, 2005.
- [32] Gilchrist, A.; Kossyfidis, C.; Bonato, F.; Agostini, T.; Cataliotti, J.; Li, X. J.; Spehar, B.; Annan, V.; Economou, E. An anchoring theory of lightness perception. *Psychological Review* Vol. 106, No. 4, 795–834, 1999.
- [33] Wang, J. F.; Li, X.; Yang, J. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1788–1797, 2018.
- [34] Zhang, L.; Zhang, Q.; Xiao, C. X. Shadow remover: Image shadow removal based on illumination recovering optimization. *IEEE Transactions on Image Processing* Vol. 24, No. 11, 4623–4636, 2015.
- [35] Su, Y. F.; Chen, H. H. A three-stage approach to shadow field estimation from partial boundary information. *IEEE Transactions on Image Processing* Vol. 19, No. 10, 2749–2760, 2010.
- [36] Oliveira, D. M.; Lins, R. D.; de França Pereira e Silva, G. Shading removal of illustrated documents. In: *Image Analysis and Recognition. Lecture Notes in Computer Science, Vol. 7950*. Kamel, M.; Campilho, A. Eds. Springer Berlin Heidelberg, 308–317, 2013.
- [37] Bako, S.; Darabi, S.; Shechtman, E.; Wang, J.; Sunkavalli, K.; Sen, P. Removing shadows from images of documents. In: *Computer Vision – ACCV 2016. Lecture Notes in Computer Science, Vol. 10113*. Lai, S. H.; Lepetit, V.; Nishino, K.; Sato, Y. Eds. Springer Cham, 173–183, 2017.

Saritha Murali is currently pursuing her Ph.D. degree in image processing at the Department of Computer Science and Engineering, National Institute of Technology Calicut, India. She holds her B.Tech. degree in computer science and engineering from Kannur University, India, and M.Tech. degree in computer

science (information security) from the National Institute of Technology Calicut, India. Her research interests are in the areas of computer vision and image processing. She has a few research publications to her credit.

V. K. Govindan is an Emeritus Professor in computer science and engineering. He served in the Department of Computer Science and Engineering of National Institute of Technology Calicut from 1982 to 2015. He has also worked as professor in computer science and engineering at the Indian Institute of

Information Technology, Kottayam. He completed his bachelor and master degrees in electrical engineering in the National Institute of Technology Calicut, and obtained his Ph.D. degree in the area of character recognition from the Indian Institute of Science, Bangalore. He has more than 40 years of teaching and research experience and has served as head of the Department of Computer Science and Engineering, and academic dean at the National Institute of Technology Calicut. His research interests include image processing, pattern recognition, machine learning, and operating systems. He has more than 180 research publications, completed several sponsored research projects, authored 20 books, produced 15 Ph.D.s, and is currently guiding two Ph.D. scholars.

Saidalavi Kalady is an associate professor in the Department of Computer Science and Engineering at the National Institute of Technology Calicut. He received his M.E. degree in computer science from the Indian Institute of Science, Bangalore, and Ph.D. degree in agent-based systems from the National

Institute of Technology Calicut. He has served as head of the Department of Computer Science and Engineering at the National Institute of Technology Calicut. His research interests include computational intelligence and operating systems.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Other papers from this open access journal are available free of charge from <http://www.springer.com/journal/41095>. To submit a manuscript, please go to <https://www.editorialmanager.com/cvmj>.