www.ietdl.org

Published in IET Image Processing
Received on 11th October 2013
Accepted on 29th June 2014
doi: 10.1049/iet-ipr.2014.0242

ISSN 1751-9659

Single-image shadow detection and removal using local colour constancy computation

Xingsheng Yuan[1], *Marc Ebner*[2], *Zhengzhi Wang*[1]

[1]*College of Electronics and Automation, National University of Defense Technology, Changsha 410073, People's Republic of China*
[2]*Ernst-Moritz-Arndt-Universität Greifswald, Institut für Mathematik und Informati, Walther-Rathenau-Straße 47, D-17487 Greifswald, Germany*
E-mail: yuanxingsheng@nudt.edu.cn

Abstract: This study is concerned with the problem of shadow detection and removal from single images of natural scenes. In this work, the authors propose a shadow detection method with a surface descriptor, termed colour-shade, which allows them to include the physical considerations derived from the image formation model capturing gradual colour surface variations. The authors incorporate a colour-shade descriptor into the condition random field model to find same illumination pairs and to obtain coherent shadow regions. The authors propose a shadow removal method using an improved local colour constancy computation, which uses anisotropic diffusion to estimate the illuminant locally for each image pixel in shadow. The authors evaluate their method on two shadow detection databases. The experimental results demonstrate that their shadow detection and removal method is state of the art.

1 Introduction

A shadow occurs when an object partially or totally occludes direct light from a light source, which is an ever-present aspect of our visual experience [1]. Shadows in images have long been disruptive to computer vision problems, such as object segmentation, tracking, recognition and so on. In practice, shadows cause problems such as shape distortion, object merging and failure of object detection and segmentation. Shadows also have a double-face effect on scene understanding, depending on whether we model the shadows or ignore them.

Generally, the region where a direct light source is totally blocked is called the umbra, whereas the region where it is partially blocked is known as the penumbra. Obviously, both of them experience a change of illumination. This illumination change is often considered only as a decrease in brightness, without significant change in chromaticity. However, the assumption only works when the light sources are white and there is no colour blending among objects. This type of shadow is often called an achromatic shadow, whereas those that are not achromatic are referred to as chromatic shadows [2]. In most conditions, a shadow is really a local change in both the colour and intensity of the scene illumination [3].

Estimating the colour of prevailing scene illumination is a related problem which has received much attention [4–7]. Computing colour constancy is an ill-posed problem [8, 9]. Often, these colour constancy algorithms are derived under some restrictive conditions or assumptions. Detecting and removing chromatic shadows is also a particularly challenging task because of the fact that they are extremely difficult to distinguish from the foreground which has no clearly defined pattern. Amato *et al.* [2] proposed to use the luminance ratio to identify segments with low gradient constancy, which in turn distinguish shadows from foreground. This algorithm achieved state of the art results on challenging video conditions. Huang and Chen [10] use a Gaussian mixture model (GMM) to learn colour features that are robust to the changes of background surfaces or illuminant colour in a scene. A pixel-based GMM is created for each pixel to learn the local shadow features. The methods mentioned above detect and remove shadows by using sufficient information in image sequences; however, they cannot be applied to the single outdoor images.

Compared with shadow detection and removal from a video sequence, the shadows in single images are more difficult to deal with because of less information provided by the image. Finlayson *et al.* [3] obtained a colour shadow-free image representation by first identifying shadow edges using an illuminant invariant at each image pixel. This method can work quite well with high-quality images and calibrated sensors, but poorly for typical web-quality consumer photographs [11]. Finlayson *et al.* [12] have shown how to use entropy minimisation to seek a type of intrinsic, independent of lighting reflection-information only image with uncalibrated sensors. Recently, data-driven approaches have been proposed for single image shadow detection. These methods learn to detect shadows based on training images. Zhu *et al.* [13] classify regions based on statistics of illuminant intensity, texture and odd order derivative characteristics, and refine shadow labels using a conditional random field (CRF). However, this method was just

proposed for monochromatic images. Lalonde *et al.* [11] find shadow boundaries by using a decision tree classifier and employing CRF-based optimisation to generate coherent shadow contours. Guo *et al.* [14] use graph-cut inference to solve the labelling of shadow and non-shadow regions. Both the two methods above neglect the physical information of imaging, however, they use colour and texture features to compare different regions and discern whether they belong to the same material, or illuminated by the same illuminant. They perform well if the surface of shadowed regions is flat and less textured. However, it is difficult for these methods to extract rugged shadow regions exactly because of the changes of the surface geometry.

Our shadow detection method uses a surface descriptor, termed colour-shade [15], which allows us to include the physical considerations derived from the image formation model capturing gradual colour surface variations, and therefore can represent colour surface variations because of the geometry and lighting of a surface. To determine whether a particular region is shadowed, we first train a classifier based on the shadow sensitive features computed from each segmented region. Then, we incorporate a physical model into the data-driven method to compare the region with other regions to find same illumination pairs, that is, regions that are made of the same material and are illuminated by the same illuminant. Like Lalonde *et al.* [11], we use a CRF-based optimisation to combine detected shadow regions into a coherent shadow region. Unlike other researchers, we use colour-shade as the pairwise potentials to penalise the assignment of different labels to paired regions with different shading conditions in the same colour ridge. After detecting shadows, we apply the matting technique of Levin *et al.* [16], treating shadow pixels as foreground and non-shadow pixels as background.

The most popular approach in shadow removal was proposed in a series of papers by Finlayson and colleagues [17–19]. However, they treat the problem of shadow removal as a reintegration process based on detected shadow edges. Guo *et al.* [14] relight shadow pixels by calculating the ratio of direct to environmental light in the image.

In this work, we propose a new shadow removal method using an improved local colour constancy method – anisotropic local space average colour. This colour constancy method first estimates lines along which the illuminant in the scene is constant. Then the method estimates the illuminant locally for each image pixel by averaging non-uniformly along the iso-illumination lines. Once the colour of the illuminant for each image pixel is known, a colour corrected image which is independent of the illuminant can be computed [4]. The main contributions of this paper are: (i) A colour-shade descriptor which generates coherent shadow regions. (ii) A shadow removal procedure based on the computation of anisotropic local space average colour.

The paper is organised as follows. In Section 2, we explain about how to incorporate the colour-shade descriptor into the CRF model to obtain coherent shadow regions. In Section 3, we show how to remove shadows using the colour constancy algorithm based on local space average colour. Experiments and result are given in Section 4. Finally, we conclude and discuss our approach in Section 5.

2 Shadow region detection using the condition random field model

Shadow regions and non-shadow regions differ with respect to illuminant intensity, texture and colour appearance. We segment the image into regions using the mean shift algorithm [20], as shown in Fig. 1*b*. We assume that there are n segmented regions in the image.

To detect the shadow regions, we first construct a graph $\mathcal{G} = (\mathcal{V}, \varepsilon)$ with vertices $\mathcal{V}$ and edges ε that represents the image. The set of vertices $\mathcal{V}$ corresponds to associated label y_i, $(i \in \{1 \ldots n\})$ which represents whether the region is in shadow $(y_i = 1)$ or not $(y_i = 0)$. The set of edges ε represents different illumination pairs. We define a CRF on that graph, which expresses the log-likelihood of a particular labelling given observed data x_i as a sum of unary potential $S(x_i, y_i)$ and pairwise potential $D(y_i, y_j)$, $(i \neq j, j \in \{1 \ldots n\})$. Let $P(y/x)$ be the probability that we obtain labelling y given data x and the log-likelihood is defined as

$$-\log P(y/x_i; \lambda) = \lambda \sum_{i \in \mathcal{V}} S(x_i, y_i) + (1 - \lambda) \sum_{(i,j) \in \varepsilon} D(y_i, y_j) \tag{1}$$

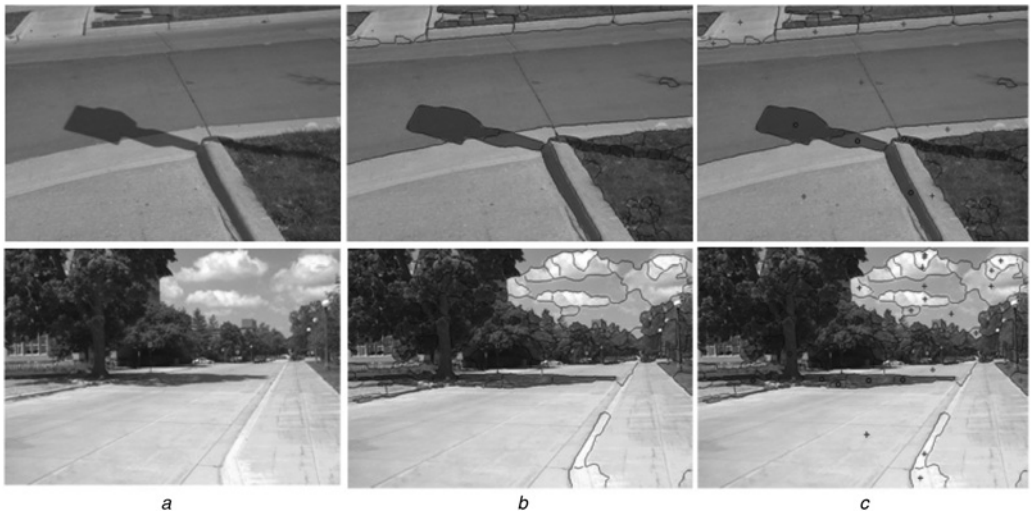

Fig. 1 *Illustration of the training processes*

a Original image
b Segmented image using mean shift algorithm
c Labelled shadow/non-shadow regions with blue or red points

where λ is a weight controlling the relative importance of the two terms, $\lambda \in [0, 1]$. In the following sections, we explain in detail how we define the potentials to integrate the information from the two terms.

2.1 Singleton potential: classifier

Shadow regions tend to be dark, with little texture. The colour appearance of shadow regions is also different because the objects in the shadow are usually illuminated by a different illuminant than the objects outside of shadow regions. Thus, we formulate features that respond to characterise colour, brightness and texture associated with shadow regions. We define a colour feature by binning kernel density estimates of the colour distribution in CIELAB space using a Gaussian kernel. For the brightness feature, we compute a histogram of values. The texture feature is represented by texton histogram as defined by Martin *et al.* [21]. We train our classifier on 100 images selected from the datasets of Guo *et al.* [14] and Zhu *et al.* [13] using a logistic regression version of Adaboost [22] with twenty 16-node decision trees as weak learner. The training set includes positive as well as negative samples. The positive training set contains the shadow and non-shadow regions were manually marked (a sample image is shown in Fig. 1*c*, the blue dots represent shadow regions, non-shadow regions are marked with a red dot).

We compute the feature histogram for each region in the image and train a classifier to estimate the probability $P(y_i/x_i)$ that region i is labelled with y_i, that is, whether region i is in the shadow or not. Intuitively, we would like the unary potentials to penalise the assignment of the 'shadow' label to regions which are not likely to be shadows according to our classifier. This can be modelled by

$$S(x_i, y_i) = -\log P(y_i/x_i) \tag{2}$$

2.2 Pairwise potentials: colour-shade descriptor

The colour-shade descriptor is based on the method proposed by Vazquez *et al.* [23]. Vazquez *et al.* describe scene reflectance by a ridge analysis of colour distribution (RAD) (RAD method). A ridge is a list of points connecting the local maxima of the colour distribution in the red, blue and green (RGB) colour space.

The RAD method is based on the dichromatic model described by Shafer in [23]. The dichromatic reflectance model maps the colour variations of a surface, including shading effects and highlights to a two-dimensional plane in the RGB colour space which is defined by two vectors: one in the direction of the surface's albedo, and the other in the direction of the illuminant [15]. In this case, the RAD method provides a single ridge for each reflectance surface of the image, as shown in Fig. 2. The RGB colour space is used for explaining the general idea of RAD method here; \ colour space do not correspond to the colour difference which are actually perceived by an human observer, we use perceptually uniform colour space $L^*a^*b^*$ in our computation in our shadow detection method.

Given an image, the RAD method returns a set of ridges $\Re = \{R_1, \ldots, R_l \ldots, R_n\}$. We use the definition of the colour-shade descriptor of a region i as

$$sd(r_i) = \arg\min_{R_l \in \Re} \text{dist}(r_i, R_l) \tag{3}$$

where $\text{dist}(r_i, R_l)$ represents the (Euclidian) distance between the $L^*a^*b^*$ value r_i and the ridge R_l. This definition was first proposed by Serra *et al.* [15]. Let $\varphi(r_i)$ be the orthogonal projection of the region value r_i onto its associated ridge $sd(r_i)$ and let σ_{ij} be the angle formed by the lines $(r_i r_j)$ and $(\varphi(r_i)\varphi(r_j))$. Based on the analysis of the colour ridges of the image, as shown in Fig. 3, the relative position between a pair of regions' $L^*a^*b^*$ values(r_i, r_j) and the set of ridges of an image can be separated into three cases:

Case A, two regions lie on two different ridges, that is, $sd(r_i) \neq sd(r_j)$; case B, two regions lie on the same ridge, that is, $sd(r_i) = sd(r_j)$, but the direction they determine is not parallel to the ridge, and the angle σ_{ij} is larger than a threshold θ_{th}, that is, $\sigma_{ij} > \theta_{th}$, where θ_{th} is a parameter fixed once for all; and case C, two regions lie on the same ridge, that is, $sd(r_i) = sd(r_j)$, and the direction they determine is parallel to the ridge, that is, the angle σ_{ij} is smaller than a threshold $\sigma_{ij} \leq \theta_{th}$.

Generally, shadow regions share different ridges with the non-shadow region, just as case A. However, shadow and non-shadow region may lie on the same ridge, but the direction they determine is parallel to the ridge, as case C.

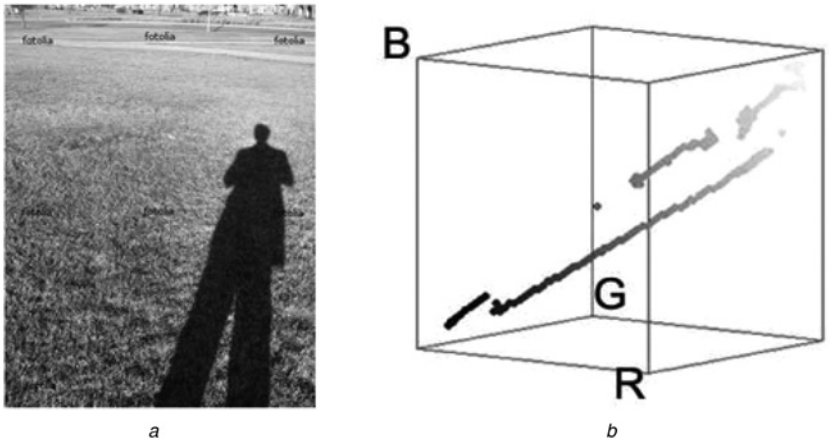

Fig. 2 *RAD method*
a Original Image
b Detected ridges

doi: 10.1049/iet-ipr.2014.0242

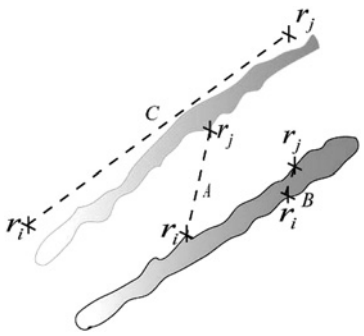

Fig. 3 *Schema of cases A, B and C for the pairwise potential computation on ridges*

We define the pairwise potential as

$$D(y_i, y_j) = \chi(i \neq j) \exp\left(-\omega_{ij} \left\| sd(r_i) - sd(r_j) \right\|_2\right) \quad (4)$$

where $\chi(\cdot)$ is a characteristic function, (y_i, y_j) are a pair label of regions of (i, j), and ω_{ij} weights the classical smoothness term according to the relative position of the $L^*a^*b^*$ values r_i and r_j of regions i, j. In this paper, we assign weights ω_{ij} with the following values for the three cases given above

$$\omega_{ij} = \begin{cases} \alpha, \text{ case } A \\ \beta, \text{ case } B \\ \gamma, \text{ case } C \end{cases} \quad (5)$$

The main idea underlying the choice of the three parameters α, β and γ is that shadow (umbra) and non-shadow region hold different ridge values in most cases and the non-uniform illuminated shadow regions caused by surface geometrical variation will belong to the same ridge with the uniform illuminated region because the pixels in the two regions belonging to the same reflectance object but with different shadings. The same reason may cause the penumbra region hold the same ridge with the non-shadow region (case C). Accordingly, (α, β, γ) should verify the inequalities $\alpha \geq \gamma \geq \beta$ and $\alpha \gg \beta$.

We incorporate the singleton potential and pairwise potentials into the CRF model. The negative likelihood in (1) can be efficiently minimised using graph cuts [24]. In order to obtain scene coherent shadow regions, we assign $\gamma = 2$, $\alpha/\gamma = 3$ and $\alpha/\beta = 100$ obtained by 2-fold cross-validation on a non-overlapping set of shadow images. The free parameter λ was assigned the value of $\lambda = 0.5$, and the threshold is $\theta_{th} = 0.05$.

3 Shadow removal based on anisotropic local colour constancy computation

Estimating and accounting for the colour of the predominant illumination has received much attention in recent years [4–8]. However, most of these colour constancy algorithms factor out the effects of shading and shadows in the scene. In this work, we consider the procedure of shadow removal as obtaining a global average illumination for each region parts. Considering shadow parts share different illumination with the non-shadow parts in shadow images, we estimate the illuminant locally and iteratively averaging neighbouring processing elements (a neighbourhood of 8 in this paper) to obtain a global average of the illumination value.

3.1 Estimating the illuminant locally

Ebner [4, 25] has developed a parallel algorithm for colour constancy which estimates the illuminant locally for each image pixel by computing local space average colour. This colour constancy based on the grey world assumption that the world is grey on average, that is, a certain standard spatial spectral average exists for the entire visual field [26].

Local space average colour can be computed using a convolution. Let $g(x, y)$ be a kernel function and let $c_i(x, y)$ be the intensity of colour channel $i \in \{R, G, B\}$ at position (x, y) of the image then local space average colour $\boldsymbol{e}_i(x, y)$ is given by

$$\boldsymbol{e}_i(x, y) = k(x, y) \int\int c_i(x, y) g(x - x', y - y') \mathrm{d}x' \mathrm{d}y' \quad (6)$$

The scale factor $k(x, y)$ is chosen such that $k(x, y) \int\int g(x', y') \mathrm{d}x' \mathrm{d}y' = 1$. We use the kernel function $g(x, y) = \mathrm{e}^{-\left(r^2/2\sigma^2\right)}$ with $r = \sqrt{x^2 + y^2}$. A suitable choice in case of a Gaussian kernel is $\sigma = 0.18s$ where $s = \max\{n_x, n_y\}/2$ and n_x and n_y is the width and height of the input image.

3.2 Shadow removal using illuminant estimation by anisotropic diffusion

Estimating the illuminant locally is possible under the grey world assumption provided that there are a sufficiently large number of differently coloured objects distributed throughout the scene. The kernel size determines the area of support. If the area of support is rather small, then we will probably have a linearly varying illuminant within this area of support. However, we need to compute local space average colour with a rather large area of support. If we have shadows in the image then we have a non-linearly varying illuminant at the border of the shadow area. A more accurate estimate of the illuminant can be computed by using anisotropic averaging [27]. This method first determines lines in the image along which the illuminant is constant (iso-illumination lines) and then averages data only along these lines.

We will treat the shadow edge as a line of constant illuminant in this section. As shown in Fig. 4, each pixel has four directions; two are perpendicular to the iso-illumination line (front/back point along the illumination gradient) and two directions point along the iso-illumination line (left/right point along the line of constant illumination). In order to compute the gradient of the illumination change, we need to estimate the illuminant first. Let $\boldsymbol{e}$ be the estimate illuminant computed using isotropic diffusion and let $[\mathrm{d}x_i, \mathrm{d}y_i]$ be the gradient of this estimate for colour channel i. The gradients of the three colour bands can be combined into a single gradient $[\mathrm{d}x, \mathrm{d}y]$ which describes the actual distribution of the illuminant [28]. We simply average the three gradients to obtain the gradient $[\mathrm{d}x, \mathrm{d}y]$ as follows

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix} = \frac{1}{3} \sum_{i \in \{r,g,b\}} \nabla \boldsymbol{e}_i \quad (7)$$

where $\nabla \boldsymbol{e}_i = \begin{bmatrix} \frac{\partial \boldsymbol{e}_i}{\partial x_i} & \frac{\partial \boldsymbol{e}_i}{\partial y_i} \end{bmatrix}^T$. We now need to calculate the

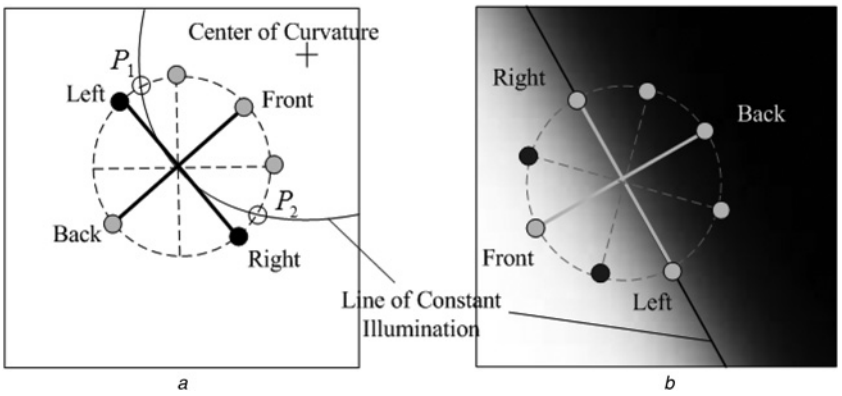

Fig. 4 *Rotated coordinate system*

a Curved line of constant illumination
b Straight line of constant illumination

local curvature of the illuminant for each pixel. Then, we can average the values along the line of constant illumination if the edge of the shadow region is curved. The curvature K of a point (x, y) on a surface $I(x, y)$ can be computed as

$$K = \frac{\begin{vmatrix} I_{xx} & I_{xy} & I_x \\ I_{yx} & I_{yy} & I_y \\ I_x & I_y & 0 \end{vmatrix}}{(I_x^2 + I_y^2)^{3/2}} \tag{8}$$

with $I_x = \partial I/\partial x$, $I_y = \partial I/\partial y$, $I_{xy} = \partial^2 I/\partial x \partial y$, $I_{yx} = \partial^2 I/\partial y \partial x$, $I_{xx} = \partial^2 I/\partial x^2$, and $I_{yy} = \partial^2 I/\partial y^2$. Thus, if we have an estimate of the illuminant, we can compute the local curvature for each pixel. Given the curvature K, we can calculate the radius R of the curve

$$R = \left|\frac{1}{K}\right| = \left|\frac{(I_x^2 + I_y^2)^{3/2}}{I_x I_{xy} I_y + I_x I_{yx} I_y - I_x^2 I_{yy} - I_y^2 I_{xx}}\right| \tag{9}$$

The curvature K tells us on which side of the curve the centre of the curvature lies. If $K > 0$, then the centre of the curvature lies on the positive side of the curve normal. If $K < 0$, then the centre lies on the negative side of the curve normal. If $K = 0$, the line of constant illumination is a straight line.

To average along the iso-illumination line, we need to determine the intersection points between the unit circle around the current pixel and the circle determined from the current estimate of the illuminant. Let P_1 and P_2 be the intersection points between the unit circle and the circle determined from the current estimate of the illuminant. Then, the two values $\boldsymbol{e}(P_1)$ and $\boldsymbol{e}(P_2)$ can be used in the anisotropic averaging operation. To compute the intersection points we assume that the centre of the curvature is located at point $(0, R)$. Then the location of the two intersection points can be computed using

$$\begin{cases} x^2 + y^2 = 1 \\ (x - R)^2 + y^2 = 1 \end{cases} \tag{10}$$

We obtain $x = 1/2R$ and $y_{(1,2)} = \pm\sqrt{1 - (1/4R^2)}$. Usually, the centre of the curvature does not lie on the X axis, and an appropriate rotation of the coordinate system has to be performed. The method estimates the illuminant by averaging the data from neighbouring elements and then adding the colour of current pixel to this average using a small percentage p. Once the data from neighbouring elements is averaged, the colour from the current element is computed as follows

$$\hat{\boldsymbol{e}}'(x, y) = \frac{1}{3}(\hat{\boldsymbol{e}}(P_1) + \hat{\boldsymbol{e}}(x, y) + \hat{\boldsymbol{e}}(P_2)) \tag{11}$$

$$\hat{\boldsymbol{e}}(x, y) = \hat{\boldsymbol{e}}'(x, y)(1 - p) + \boldsymbol{c}(x, y)p \tag{12}$$

where local space average colour at points P_1 and P_2 is obtained using bilinear interpolation. $\hat{\boldsymbol{e}}(P_1)$ and $\hat{\boldsymbol{e}}(P_2)$ is the previous estimate of local space average colour at points P_1 and P_2, $c(x, y)$ is the measured colour at location (x, y). The parameter p depends on the size of the image; has to be chosen such that the local average colour is computed over a reasonably large area of the entire image. If the parameter p is large, then local space average colour will be computed for a small neighbourhood. If the parameter p is small, the local space average colour will be computed for a large neighbourhood. We can also include the pixels along the illumination gradient using a parameter q, using

$$\begin{aligned} \hat{\boldsymbol{e}}'(x, y) = &\left(\frac{1}{3} - q\right)\hat{\boldsymbol{e}}(\text{left}) + \left(\frac{1}{3} - q\right)\hat{\boldsymbol{e}}(\text{right}) \\ &+ \frac{1}{3}\hat{\boldsymbol{e}}(x, y) + q\hat{\boldsymbol{e}}(\text{front}) + q\hat{\boldsymbol{e}}(\text{back}) \end{aligned} \tag{13}$$

where $q \in [0, 1/6]$ determines the extent of the averaging perpendicular to the iso-illumination line. We set $q = 0.05$, so that we average data including the font/back direction along the line of constant illumination.

These update equations are iterated for each pixel until colour converges. If p is large, convergence will be achieved earlier, however, the illuminant colour of the shadow image will not converge well, as shown in Fig. 5, we obtain the best result when p decreased to 0.0001. In our work, in order to make the area large enough such that the grey world assumption is valid, we have to choose small values $p = 0.0001$, such that the grey world assumption holds even if this leads to a longer running time. After convergence, we brighten the whole image using a gamma correction.

doi: 10.1049/iet-ipr.2014.0242

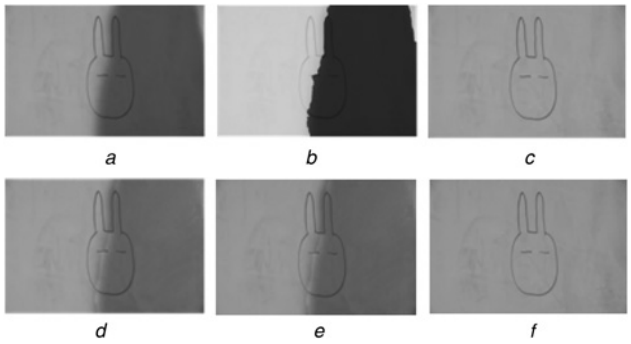

Fig. 5 *Shadow recovery with different p values*

a Original image
b Shadow mask
c Ground truth
d $p = 0.05$
e $p = 0.005$
f $p = 0.0001$

4 Experiments and results

In our experiments, we evaluate both shadow detection and shadow removal results. We evaluate our approach on two datasets: the dataset introduced by Guo *et al.* [14] and the University of Central Florida (UCF) shadow dataset introduced by Zhu *et al.* [13]. Guo *et al.* provides a dataset with 108 natural scenes, in which ground truth is determined by taking two photographs of a scene after manipulating the shadow. Zhu *et al.* made available a set of 245 images with manually labelled ground truth shadow masks. We give some comparison between our method with the ground truth images in Fig. 6.

We compare our method to a state of the art colour-based shadow detection and removal method by Zhu *et al.* [13] for single images. This method classifies regions based on statistics of intensity, gradient and texture, computed over local neighbourhoods, and refines shadow labels using a CRF. We also compare our method to the one developed by Guo *et al.*' [14], which detects shadow regions using pairwise classification based on appearances of segmented regions, and relighting each shadow pixel by their lighting model. For shadow detection, we evaluate how explicitly modelling the pairwise region relationships (colour ridge differences) affects detection results.

Table 1 Shadow detection evaluation on both dataset (per pixel accuracy)

Methods	UCF shadow dataset	Guo *et al.*'s dataset
Zhu *et al.*' method	0.848	–
Guo *et al.*' method		
unary SVM	0.871	0.817
unary SVM+ pairwise	0.902	0.891
our method		
unary potential	0.883	0.821
pairwise potential	0.923	0.897

The comparison between unary and pairwise information is given in Tables 1 and 2. Using only unary information, our performance on the UCF dataset is 88.3%, against 87.1% achieved by [13]. By combining unary information with pairwise information, we achieve an accuracy of 92.3%, better than 90.2 and 84.8% reported in [14, 13], respectively. From these quantitative results, we can see that the colour ridge relations of neighbouring regions are especially effective to detect shadow regions. The overall accuracy increases by 3% points for Guo *et al.*'s method and by 5% points for our method, as shown in Table 1.

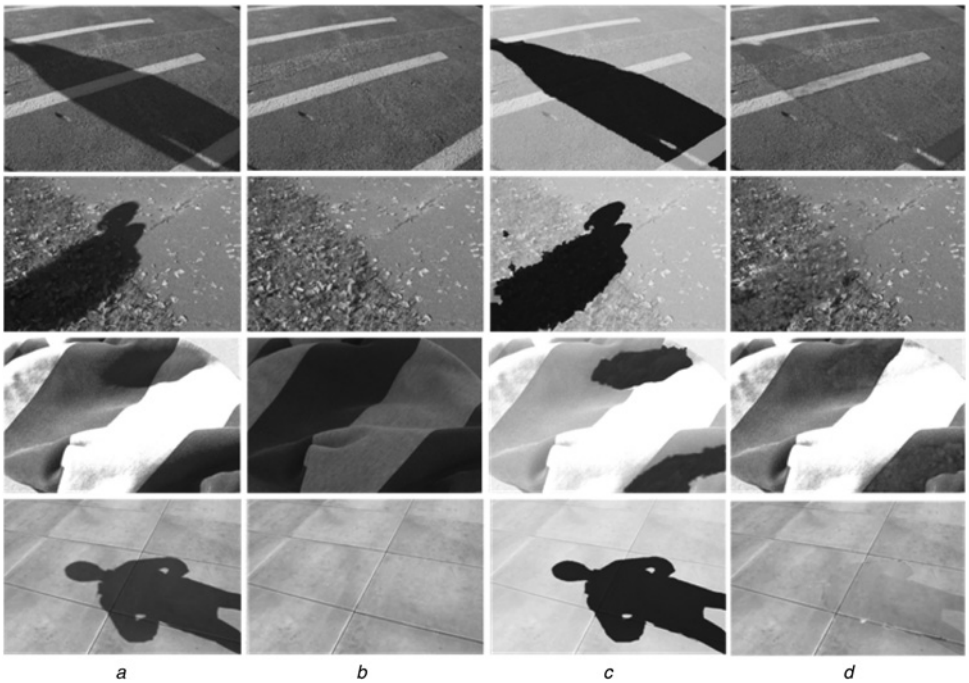

Fig. 6 *Comparisons examples between our method with ground truth images*

a Original image
b Ground truth image
c Shadow mask
d Shadow removal results with our method

Table 2 Shadow detection confusion matrices

UCF dataset (Zhu *et al.* [10])	shadow	non-shadow
shadow(GT)	0.639	0.361
non-shadow(GT)	0.063	0.937
Guo *et al.*' dataset (Guo *et al.* [11])	shadow	non-shadow
shadow(GT)	0.716	0.284
non-shadow(GT)	0.048	0.952
our method		
UCF dataset (unary)	shadow	non-shadow
shadow(GT)	0.417	0.583
non-shadow(GT)	0.057	0.943
UCF dataset (unary and pairwise)	shadow	non-shadow
shadow(GT)	0.745	0.255
non-shadow(GT)	0.019	0.981
Guo *et al.*' dataset (unary)	shadow	non-shadow
shadow(GT)	0.554	0.346
non-shadow(GT)	0.062	0.938
Guo *et al.*' dataset (unary and pairwise)	shadow	non-shadow
shadow(GT)	0.762	0.238
non-shadow(GT)	0.033	0.967

The experimental results prove that the colour ridge descriptor used in the pairwise information of the CRF model is important, which eliminates 44% of the pixel labelling errors of shadow regions on the UCF dataset and 27.3% of the errors on Guo *et al.*'s dataset, as shown in the Table 2 by confusion matrices. Using colour ridge relation as the pairwise information is better than Guo *et al.*'s pairwise illumination relations. The pairwise illumination relations eliminate 40% of the errors on their dataset and 24% of errors on UCF dataset as reported in [14].

We qualitatively evaluate the performance of our method on a subset of images from both datasets. The shadow regions are given by the shadow mask, as shown in Figs. 7*b* and *e*. Our shadow removal procedure based on local space average colour constancy computation yields results with good performance. For the colour biased shadow images, our method could correct for the colour of the illuminant, as shown in Fig. 7*c*. The shadow removal results in Fig. 7*f* demonstrate that our method works well on challenging outdoor images with varying illumination conditions, ground colours and textures.

Uneven pavement of the ground or the wall makes it is more difficult to detect the shadow regions exactly because of the changes of the surface geometry. We compare our shadow detection and removal method with Guo *et al.*'s method qualitatively, as shown in Fig. 8. The comparison results show that our method is better than their method. It detects more details in textured shadow images as shown at the top two images in Figs. 8*b* and *d*. From these results, we can see that our shadow detection method can detect the shadow of the human arm exactly without missing any small patch of the shadow region from the textured and uneven ground in the first row of Fig. 8. Our method also extracts the rugged shadow surface on the ground exactly. This is shown in the third row of Fig. 8*d*. That is because the RAD method can provide a compact representation of all the variations of a single-colour surface.

The comparison between Figs. 8*c* and *e* shows that our method is better at shadow removal, especially for the uneven reflection surface. As shown in the shadow-free images at the bottom of Figs. 8*c* and *e*, the uneven ground and the wall makes the surface reflection complex and complicate the shadow detection and removal. Our method gives better shadow detection result in as shown in Figs. 8*d* and *e*. However, Guo *et al.*'s method did not give the exact extraction of shadow region which misses the narrow long region of human arm in the first image in Fig. 8*b* and many tiny shadow part in wall image. We can say that the colour-shade descriptor incorporated in the CRF model

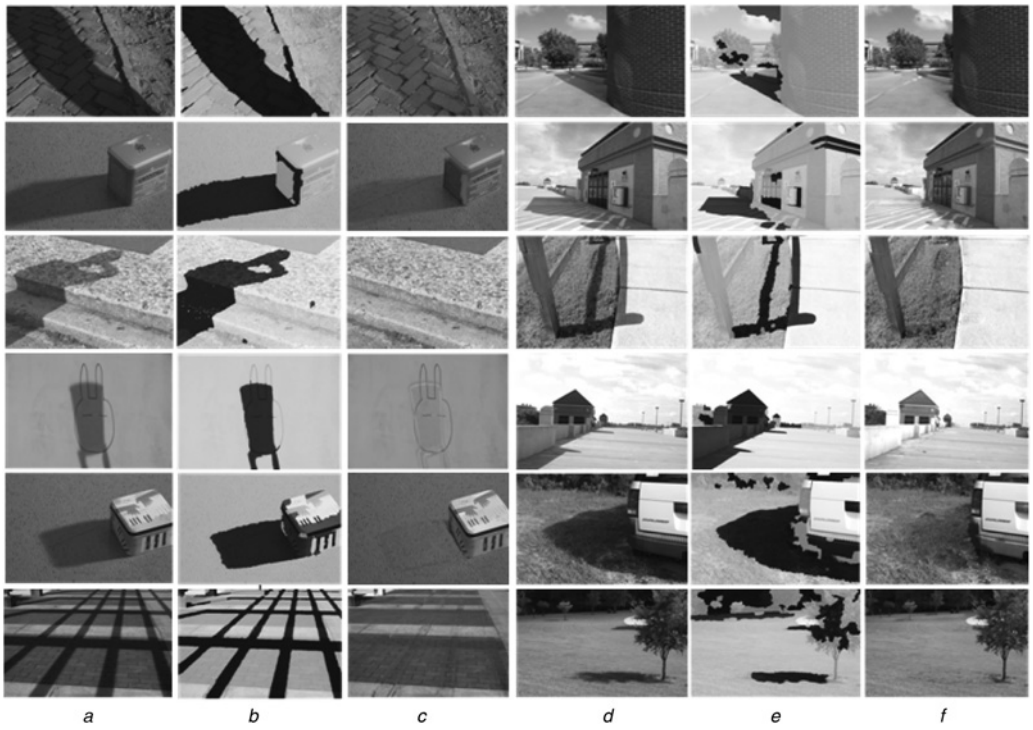

Fig. 7 *Shadow detection and removal results*

a, *d* Example of images from Guo *et al.*'s dataset and UCF dataset, respectively
b, *e* Detection results
c, *f* Shadow removal results

doi: 10.1049/iet-ipr.2014.0242

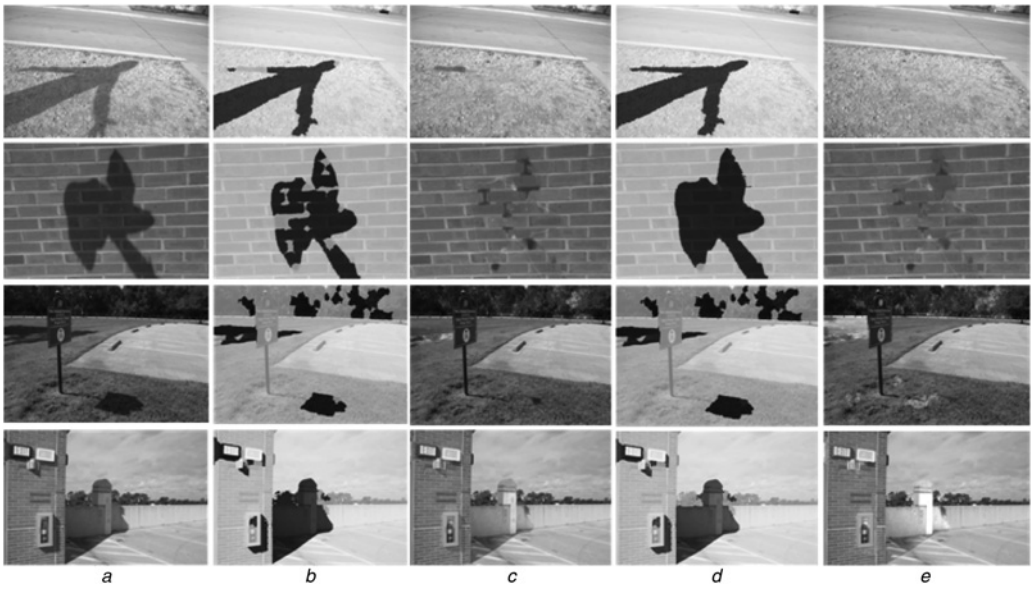

Fig. 8 *Shadow detection and removal comparisons*

a Original image
b Detection results of Guo *et al.*'s method
c Recovery results of Guo *et al.*'s method
d Detection results of our method
e Recovery results of our method

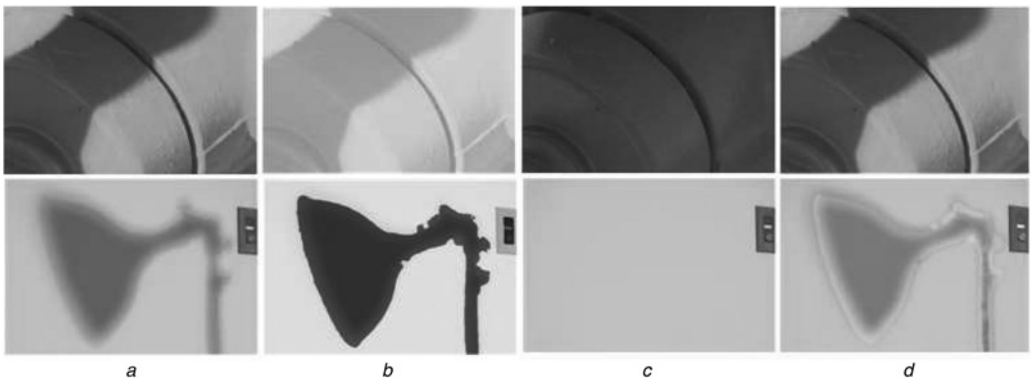

Fig. 9 *Failed examples*

a Original image
b Detection result with shadow mask
c Ground truth
d Recovery images

considering the surface reflection changes by including physical imaging information. Besides, our shadow removal method for the shadow image in the last row of Fig. 8*a* performs better than Guo *et al.*'s method. The shadow region is much weaker than Guo *et al.*'s shadow recovery result.

The first row of Fig. 9*b* shows a false result in shadow detection. The original image just has one single reflection surface without more texture and colour information. Therefore the classifier could not discriminate shadow part in the image. Our shadow removal method is based on Ebner's [27] anisotropic local average illuminant estimation which based on the grey world assumption that the average reflection in a scene under a neutral light source is achromatic. Therefore our method is sensitive to large uniformly coloured surfaces in a way. That is why our method is failed in shadow removal as shown in Fig. 9*d*.

5 Conclusions

We have created a novel method to detect and remove shadows from a single image. For shadow detection, we have proposed a new method that adds a physical model to the data-driven method by using a colour-shade descriptor as the pairwise potential of the CRF model. By using the pairwise potential on the basis of the colour-shade descriptor, the neighbouring regions of different shading will be separated exactly. It also works well for images with uneven shadow surfaces. For shadow removal, we use anisotropic local space average colour constancy computation. The experimental results on the two datasets prove the effectiveness of our method, and the comparison results show that our method performs better than recently proposed methods in most situations.

6 Acknowledgment

This work was supported by the National Science Foundation of China (No.61174206) and 973 project under Contracts 2013CB329401.

7 References

1 Klinker, G.J.: 'A physical approach to color image understanding', *Int. J. Comput. Vis.*, 1990, **4**, pp. 7–38
2 Amato, A., Mozerov, M.G., Bagdanov, A.D., Gonzalez, J.: 'Accurate moving cast shadow suppression based on local color constancy detection', *IEEE Trans. on Image Process.*, 2011, **20**, (20), pp. 2954–2966
3 Finlayson, G.D., Hordley, S.D., Lu, C., *et al.*: 'On the removal of shadows from images', *IEEE Trans. Pattern Anal. Intell.*, 2006, **28**, (1), pp. 59–68
4 Ebner, M.: 'Color constancy based on local space average color', *Mach. Vis. Appl.*, 2009, **20**, (5), pp. 283–301.
5 Gijsenij, A., Lu, R., Gevers, T.: 'Color constancy for multiple light source', *IEEE Trans. Image Process.*, 2012, **21**, (2), pp. 697–708
6 Chakrabarti, A., Hirakawa, K.: 'Computational color constancy with spatial correlation', Technical Report, TR-09-10, Harvard School of Engineering and Applied Science, 2010
7 Gijsenij, A., Gevers, T., van de Weijer, J.: 'Computational color constancy: survey and experiment', *IEEE Trans. Image Process.*, 2011, **20**, (9), pp. 2475–2489
8 Finlason, G., Hordley, S.: 'Color constancy at a pixel', *J. Opt. Soc. Am. A*, 2001, **18**, (2), pp. 253–264
9 Funt, B., Barnard, K., Martin, L.: 'Is machine colour constancy good enough?'. Fifth European Conf. on Computer Vision. Springer, June 1998, pp. 455–459
10 Huang, J.-B., Chen, C.-S.: 'Moving cast shadow detection using physics-based features'. Proc. Conf. on Computer Vision and Pattern Recognition, Miami, Florida, USA, 2009, pp. 2310–2317
11 Lalonde, J.-F., Efros, A.A., Narasimhan, S.G.: 'Detecting ground shadows in outdoor consumer photographs'. Proc. on 11th European Conf. Computer Vision, Heraklion, Crete, Greece, September 2010, pp. 322–335
12 Finlayson, G.D., Drew, M.S., Lu, C.: 'Intrinsic images by entropy minimization', Proc. of the 8th European Conf. on Computer Vision, Part III Prague, Czech Republic, May 2004, pp. 582–595
13 Zhu, J.-J., Samuel, K.-G.G., Masood, S.Z., *et al.*: 'Learning to recognize shadows in monochromatic natural images'. Proc. Conf. on Computer Vision and Pattern Recognition, San Francisco, USA, June 2010, pp. 223–230
14 Guo, R.-Q., Dai, Q.-Y., Hoiem, D.: 'Single-image shadow detection and removal using paired regions'. Proc. Conf. on Computer Vision and Pattern Recognition, Colorado Springs, USA, June 2011, pp. 2033–2040
15 Serra, M., Penacchio, O., Benavent, R., Vanrell, M.: 'Names and shades of color for intrinsic image estimation'. Proc. Conf. on Computer Vision and Pattern Recognition, 2012, Rhode502-511
16 Fredembach, C., Finlayson, G.: 'Fast re-integration of shadow free images'. Proc. on 12th Color Imaging Conf., Scottsdde, Arizona, November 2004, pp. 117–122
17 Comaniciu, D., Meer, P.: 'Mean shift: a robust approach toward feature space analysis'. *IEEE Pattern Recognition and Machine Intelligent*, 2002, **24**, (5), pp. 278–285
18 Levin, A., Lischinski, D., Weiss, Y.: 'A closed-form solution to natural image matting', *IEEE Trans. Pattern Recognit. Mach. Intell.*, 2008, **30**, (2), pp. 228–242
19 Finlayson, G.D., Hordley, S.D., Drew, M.S.: 'Removing shadows from images using retinex'. Proc. on Tenth Color Imaging Conf., November 2002, pp. 73–79
20 Fredembach, C., Finlayson, G.: 'Hamiltonian path-based shadow removal'. Proc. on 16th British Machine Vision Conf., Oxford, UK, September 2005, pp. 24, no. 5, pp. 603–619
21 Martin, D.R., Fowlkes, C.C., Malik, J.: 'Learning to detect natural image boundaries using local brightness, color, and texture cues', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2004, **26**, (5), pp. 530–549
22 Collins, M., Shapire, R., Singer, Y.: 'Logistic regression, adaboost and Bregman distances', *Mach. Learn.*, 2002, **48**, pp. 253–285
23 Vazquez, E., Baldrich, R, van de Weijer, J., Vanrell, M.: 'Describing reflectances for colour segmentation robust to shadows, highlights and textures', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2011, **33**, (5), pp. 917–930
24 Shafer, S.A.: 'Using color to separate reflection components', *Color Res. Appl.*, 1985, **10**, (4), pp. 210–218
25 Buchsbaum, G.: 'A spatial processor model for object colour perception', *J. Franklin Inst.*, 1980, **310**, (1), pp. 1–26
26 Boykov, Y., Kolmogorov, V.: 'An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2004, **26**, (9), pp. 1124–1137
27 Ebner, M.: 'A parallel algorithm for color constancy', *J. Parallel Distrib. Comput.*, 2004, **64**, (1), pp. 79–88
28 Ebner, M.: 'Estimating the color of the illuminant using anisotropic diffusion'. Proc. on 12th Int. Conf. Computing Analysis of Images and Patterns, Vienna, Austria, October 2007, pp. 441–449

doi: 10.1049/iet-ipr.2014.0242