Volume 0 (1981), Number 0 pp. 1–12 COMPUTER GRAPHICS *forum*

Fast Shadow Removal Using Adaptive Multi-scale Illumination Transfer

Chunxia Xiao[†1] Ruiyun She [1] Donglin Xiao [1] and Kwan-Liu Ma[2]

[1]School of Computer, Wuhan University, Wuhan, China
[2]Department of Computer Science, University of California-Davis, CA, USA

Abstract

In this paper, we present a new method for removing shadows from images. First, shadows are detected by interactive brushing assisted with a Gaussian Mixture Model. Second, the detected shadows are removed using an adaptive illumination transfer approach that accounts for the reflectance variation of the image texture. The contrast and noise levels of the result are then improved with a multi-scale illumination transfer technique. Finally, any visible shadow boundaries in the image can be eliminated based on our Bayesian framework. We also extend our method to video data and achieve temporally consistent shadow free results. We show that our method is fast and can generate satisfactory results for images with complex shadows.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

Although shadows provide important visual cues for perception of shape, occlusion, etc, shadow-free images will help improve the performance of computerized tasks such as object recognition, image enhancement, and video object tracking [BTSD78, XG12, FHLD06]. Shadow removal is also important in image editing and processing tasks. Although shadow removal becomes increasingly needed, it is still a challenging research problem. Shadows in a scene can be quite complex, and the changes of the occluding objects, inter-reflections, reflectance variation, as well as the intensity and colour variation of the illumination, all have effects on the shadowed scene. Thus, shadow removal using a single shadow removal model is a difficult task. In addition, the computational complexity of shadow removal calculations can be prohibitively high for high resolution images.

In this paper, we address the problem of shadow detection and removal from a single image. To determine whether a pixel is dark due to shading or the albedo (reflectance) involves some degree of image understanding. For natural image with complex scenes, automatic shadow detection is a difficult problem. In computer graphics community, one feasible approach to detect the shadow for shadow removal and editing is to involve appropriate user interaction. We identify shadow areas using a Gaussian Mixture Model (GMM) brush, which finds pixels that are similar to the user-specified shadowed regions according to the appearance similarity. This technique applies only minimal user assistance and achieves satisfactory shadow detection results.

With the detected shadowed regions and selected lit regions, we carry out shadow removal from a single image using an adaptive multi-scale illumination transfer method. Inspired by the color transfer technique [RAGS01] and the affine shadow formation model [SL08], we apply the illumination of lit regions as the guided information to recover the illumination information in the shadowed regions. Instead of using a uniform parameter for the shadow removal model [SL08], we introduce a novel illumination transfer technique using adaptive recovery parameters, which takes the image reflectance variation into account. As different image textures have different reflectance, our method recovers the edge information and texture detail much better than [SL08]. To handle extremely complicated shadowed scenes, based on a multi-scale image decomposition, we developed a multi-scale illumination restoration model, which can further improve the detail recovering and reduce the noise in the recovered regions.

† Corresponding to Chunxia Xiao, Email: cxxiao@whu.edu.cn

At the shadow boundaries, the shadow intensity usually changes more rapidly than in the interior of the shadow. Furthermore, complicated shadow boundaries are hard to be accurately detected in a complex scene. Thus, the boundaries of the shadow require special treatment to achieve continuous illumination effects along the boundaries regions. To produce a seamless transition between the original and the recovered regions, we propose a Bayesian shadow boundaries estimation approach to eliminate the shadows and illumination discontinuity along the boundaries, which produces faithful results and avoids modifying the image content of the boundary regions.

The presented methods can be easily extended to video shadow removal. The GMM brush based shadow detection, multi-scale illumination transfer and shadow boundaries processing can be applied to video data with little modification. For video shadow removal, besides of removing the detected shadow of each frame, we also present techniques to preserve the shadow free video temporally consistent, which makes our method a potential applicable video shadow removal tool.

The main contributions of our work are as follows: 1) Present a shadow detection method using GMM brush, which can efficiently detect the complex shadow scenes with minimal user interaction. 2) Introduce a multi-scale illumination restoration model for shadow removal which efficiently derives a shadow-free image and avoids the loss of texture detail. 3) Solve the illumination discontinuity problem along the shadow boundaries using a Bayesian shadow estimation approach.

2. Related work

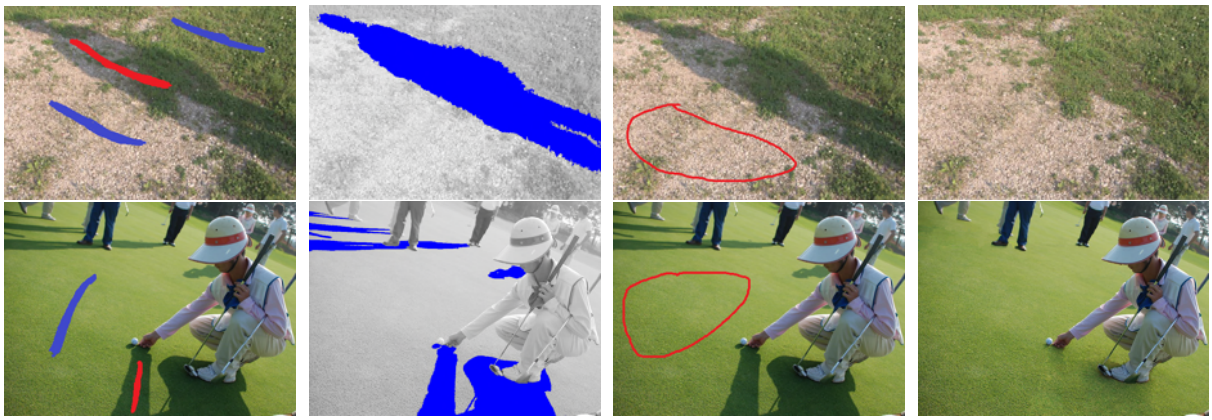
Many shadow detection and removal methods have been developed, a complete review of existing works is beyond the scope of this paper, we refer readers to [LG08, GDH11, AHO11] for excellent overviews on these methods.

Shadow detection: Many automatic shadow detection methods have been presented. Many existing methods are motivated by physical formulation models of illumination [MFS08,FHLD06]. For example, Finlayson et al. [FHLD06] compared edges in the original RGB image to edges found in an illuminant-invariant image. This method can work quite well with high-quality images and calibrated sensors, but often performs poorly for consumer photographs [LEN10]. Several data-driven approaches also have been proposed that learn to detect shadows based on training images [ZSMT10, LEN10, GDH11]. Although these methods work well for shadow detection, but still performs poorly for detecting complex shadow scenes automatically, for example, the scenes with soft shadows and complex texture structure. More automatic shadow detection algorithms refer to the survey paper [ANBSE11].

Semi-automatic shadow detection methods also have been proposed, these method involve the appropriate user assistance [WT05,WTBS07,LG08,SL08]. Wu and his colleagues [WT05, WTBS07] applied the user-specified quadmap to divide the input image into four regions, definitely shadowed regions, definitely nonshadowed regions, uncertain regions, and excluded regions (where the shadow-casting object is present). Liu and Gleicher [LG08] used a brush tool to specify the shadow boundary, thus the image is divided into three areas: definite umbra area, definite lit ares, and boundary. Shor et al. [SL08] applied iterative region growing process from the given shadow seed to detect shadowed regions. This method applied the alpha matting approach [LLW08] to compute a complete binary mask of the shadow regions, which was time-consuming for processing high resolution images. The region growing method is also not convenient for selecting disconnected similar shadow areas. Our GMM shadow detection brush can select disconnected similar shadow areas which is common in complex shadow scenes.

Shadow removal: Many shadow removal methods have been proposed. One of the most popular approaches in shadow removal is proposed in a series of papers by Finlayson and colleagues [FHLD06, FHD06, FDL04, FF05]. Finlayson and colleagues treated shadow removal as a reintegration problem based on detected shadow edge and produced some impressive results. However, the reintegration method depends on precise detection of shadow edges, without careful parameter tuning, inaccurate shadow edges may produce unpleasing results. Furthermore, these methods do not consider the self-shadowing and the effect of the ambient illumination, which makes these methods not work well for recovering the fine texture detail. Finlayson et al. [FF06] also presented a further simplification method, where they replaced the integration by scaling the shadow with a single constant factor. Once shadowed areas are detected, the shadow can be removed by multiplying a suitable scalar to the shadowed pixels to cancel the effect of the illumination change. Arbel and Hel-Or [AHO07] took surface geometry into account when computing the scale factors, which helped to preserving texture in both umbra and penumbra areas. This method works well for processing uniform shadows. More recently, Arbel and Hel-Or [AHO07] further improved their approach to process nonuniform shadows, curved and textured surfaces.

Several shadow removal methods based on color transfer technique [RAGS01] are presented [LG08, SL08]. Liu and Gleicher [LG08] constructed a shadow-effect free and texture-consistent gradient field for the shadow and lit area, respectively, with the mean and deviation of the gradients in the shadow and lit regions, they recovered the shadow-free image for the shadow area by solving a Poisson equation similar to [FHLD06]. Although solving Poisson equation can be accelerated [Aga07], it is still not convenient to process image with many disconnected shadow regions, especially for narrow shadow regions. Shor et al. [SL08] first identified shadowed and lit areas on the same surface in the



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

Figure 1: *Shadow regions detection and lit region specification. (a) Scribbles on the shadowed regions (F) and lit regions (B), F and B are used as training data for shadow detection, (b) detected shadow, (c) specified lit regions (L), L is used to estimate the adaptive parameters in the shadow removal approach, (d) shadow removal results.*

scene, and used these areas to estimate the parameters of an affine shadow formation model, then by applying the color transfer technique [RAGS01] to produce a shadow-free image. This paper produces high quality shadow-free image, however, as pointed in [SL08], when the shadow is cast onto a surface consisting of complicated texture with very strong self shadowing effects, and strong variations in color, one single set of affine parameters used in [SL08] are unable to remove the shadow in a satisfactory manner.

Shadow removal also can be considered as a matting problem [WT05, WTBS07, GDH11]. Both [WT05] and [WTBS07] applied user-specified quadmap to simplify the shadow removal optimization. For example, Wu et al. [WTBS07] performed the shadow removal in two steps. They first computed an approximate shadowless image based on color transfer techniques [RAGS01] with the hints in quadmap, then they used the shadowless image to define an optimization function (consisted of a color term and a smoothness term) to receive the final shadow removal and extraction results. Both [WT05] and [WTBS07] can receive robust shadow matte extraction while maintaining texture in the shadowed region. To process complex scenes, these two methods still may not recover the image detail in the shadow areas. After detecting shadows, Guo et al [GDH11] treated shadowed pixels as foreground and nonshadowed pixels as background, and applied matting technique of Levin et al. [LLW08] to recover the shadow coefficients. Then they calculated the ratio between direct light and environment light and generated the recovered image by relighting each pixel. This approach can produce high-quality shadow free image, even for some soft shadows, this method also works well. However, this method also does not take the reflectance variation into account, and it also can not recover the texture detail well.

Shadow editing: Several shadow editing methods also have been proposed. Mohan et al. [MCT07] proposed a shadow editing tool for a single-input image by fitting a gradient domain shadow edge model to the shadows. Besides enabling users to remove the shadows, this method also can simulate a variety of lighting conditions, such as ambient-only lighting and shadows from synthetic occluders, for shadow manipulating. Chuang et al. [CGC*03] introduced a shadow matting and compositing equation which considered the input image as a linear combination of a shadow-free image and a shadowed image. This method works well for seamless shadow composition.

3. Shadow detection

Image segmentation methods such as Grabcut [RKB04] also can be applied for interactive shadowed regions detection. However, in each segmentation processing, Grabcut [RKB04] produces a closed segmentation contour, thus, only one closed shadowed region is detected. For image with many disconnected shadow areas, this method is tedious. We would like to develop an algorithm to detect disconnected shadow areas simultaneously with little user interaction.

We present a new GMM brush for shadow detection. GMM is a probability density function composed of a set of Gaussian models, and is a widely used model in image and video analysis for the description of color distribution [RKB04]. The GMM is modeled using a full-covariance Gaussian mixture with K components (typically $K = 5$): $G(\vec{x}) = \Sigma_{k=1}^{K} p_k g_k(\vec{x})$, where $\vec{x}$ is a D dimension vector, p_k is the mixing coefficients satisfying $\Sigma_{k=1}^{K} p_k = 1$, and $g_k(\vec{x})$ is the Gaussian density function. Using the GMM model as the image representation, we present a GMM brush for detecting shadow interactively.

(a) (b) (c)

Figure 2: *(a) Input image with shadow, (b) the shadow detection result without using multi-scale GMM brush, (c) the shadow detection result using multi-scale GMM brush.*

We first cluster the input image using the mean-shift clustering algorithm to create small regions (on average 80 pixels in size) on the input image. The image is clustered based on pixels similarity measure defined in the feature space of input data set. We define the similarity between the pixels using both spatial locality (x,y) and color value (r,g,b) of pixel. Each pixel p is represented as a five-dimensional feature vector $\vec{p} = (r,g,b,x,y)$, which constitutes the feature space of the input data set. The similarity measure between two pixels p_i and p_j is defined as $z_{ij} = \exp(-\|\vec{p}_i - \vec{p}_j\|^2)$. The mean-shift clustering can be accelerated [XL10], and it requires about 1 second for an image with size of 720×480 to complete the clustering. With the clustering as a preprocess, some abandon scattered dark points can be clustered into nearest cluster, these points would otherwise be selected incorrectly as shadowed regions.

As global brightness of the image is an effective cue for detecting shadow [BA03], with the image brightness we calculate the shadow density ς to show the degree of the light's effect [BA03]. The density ς and color space channels (r,g,b) constitute a four-dimensional feature space. For each cluster, we compute the average density $\bar{\varsigma}$ and average channels $(\bar{r},\bar{g},\bar{b})$ of the pixels contained in the cluster. We define a GMM brush on the four-dimensional feature vector $(\bar{\varsigma},\bar{r},\bar{g},\bar{b})$, which handles image similarity measure between the user-selected shadowed pixels and the rest pixels to identify the shadowed regions.

We use strokes to specify both some shadowed regions F and some lit regions B from the image as the training data, as illustrated in Figure 1. Then we build GMM models $GMM(F)$ and $GMM(B)$ for both shadowed regions F and lit regions B, respectively. Let μ_F and μ_B be the mean GMM probability value of all pixels in F and B. Let $\omega_F(c)$ and $\omega_B(c)$ be the probability value that cluster c belongs to F and B, respectively. The confidence that cluter c being F can be evaluated as $H(c) = \phi \cdot |\omega_F(c) - \mu_F| + \varphi \cdot |\omega_B(c) - \mu_B|$. The weights ϕ and φ are defined as the weighted sum of its similarity between cluter c and each cluster in the stroked cluster set in similar way as [XY*11]. Using these techniques, cluster c can be efficiently classified as shadowed regions or lit regions by evaluating $H(c)$: if $H(c) > \delta$, c belongs to the F; otherwise, c belongs to B. The parameter δ can be tuned to control the amount of selected shadowed regions.

The image noise may cause problems in shadow selection. Since progressively coarser image levels increase the degree of abstraction in the resulting image [FAR07], we can build a more robust GMM brush on the coarse image with detail layers attenuated, and to provide a more consistent shadow regions selection. We construct M progressively coarser images $I_1,\cdots,I_M$ for the input image I using an edge preserving filter [FAR07, FFLS08, TM98]. Then we build a GMM brush on the coarsest image level I_M for shadowed region selection. From our experiments, we find out that by setting $M = 3$ generates good results. We call this method multi-scale shadow detection. In the top row of Figure 1, our multi-scale GMM brush eliminates the appearance noise interfering, and detect the shadow in complex background effectively. In Figure 2, the shadow detection result using multi-scale GMM brush is more accurate than the results without using multi-scale GMM brush.

The presented GMM brush is capable of achieving consistent shadow detection results, that is, the GMM brush supports selection of disconnected shadow areas, while these are common situations in complex shadow images. As illustrated in bottom row in Figure 1, using minimal user interactivity, with the similarity propagation property of the GMM brush, the disconnected shadow areas are selected effectively. Note that the interactive image segmentation method Grabcut [RKB04] can also detect the shadowed regions in the bottom row in Figure 1, however, more user interaction is required. Note that for some complex images, we perform the shadow region detection in an iterative process, and the user can refine initial estimates with additional strokes.

4. Shadow removal model

In image formation equation [BTSD78], an image $I(x)$ at the point x in the scene is the pixelwise product of its two corresponding intrinsic images:

$$I(x) = R(x)L(x)$$

where $L(x)$ and $R(x)$ are the illumination and the reflectance (albedo) at the same point x.

Assuming a scene whose shadows are cast due to a single primary source of illumination, if a point x in the scene is unshadowed (lit), the illumination can be described as a sum of two terms, $L(x) = L^d(x) + L^a(x)$, where $L^d(x)$ is the direct illumination and $L^a(x)$ is the indirect (ambient) illumination. Thus the intensity on point x is: $I^{lit}(x) = L^d(x)R(x) + L^a(x)R(x)$. If some object occludes the primary light source, it will cast a shadow on point x. As this occluder would also block some of the ambient illumination, thus, the reflected intensity on point x is $I^{shadow}(x) = \eta(x)L^a(x)R(x)$, where $\eta(x)$ is the variant attenuation factor of the ambient illumination inside the shadowed area due to the occluder.

With above observation, the lit intensity at x can be expressed as an affine function between the illuminated and

shadowed intensities [SL08]:

$$I^{lit}(x) = L^d(x)R(x) + \frac{1}{\eta(x)} I^{shadow}(x)$$

This affine function can be reformulated as:

$$I_k^{lit}(p) = \alpha_k(p) + \gamma(p) I_k^{shadow}(p) \quad (1)$$

where α_k, $k \in \{R,G,B\}$, is the reflected direct illumination in the three *RGB* color channels, and $\gamma(p) = \frac{1}{\eta(x)}$ is the inverse of the ambient attenuation factor.

Based on the affine model (1), the illuminated pixel color may be recovered from its shadowed color by estimating those four affine parameters α_k and $\gamma(p)$. Using the general affine relationship between the shadowed and the lit regions, Shor et al. [SL08] estimated four uniform parameters for the affine model, and applied the color transfer techniques [RAGS01] to recover the illuminated intensity for each pixel in the shadowed regions. However, uniform parameters for the shadowed regions may not work well. One reason is that the shadowed regions usually contain surfaces with different reflectance, especially for images with complex scenes and texture structures. Another reason is that strong self shadowing effects and strong variations in color also present in some texture images. For example, each stone in Figure 3 may have different reflectance, self shadowing also exists in this image.

Ideally, we should recover illuminated intensity at a shadowed pixel from lit regions that come from the same surface (thus with same materials), and then estimate the four adaptive parameters of the affine model (1). However, the shadowed regions usually contain different kinds of surfaces with different materials, thus, for each of these surfaces it is not feasible to find a corresponding region with the same material in the lit regions. For example, sometimes there is no corresponding region in the lit regions. We would like to present a more general and adaptive illumination transfer method for shadow removal that respecting to the reflectance variation.

We present an adaptive illumination transfer approach for shadow removal. With the selected lit region L, and the detected shadowed regions S, we estimate adaptive parameters for illumination transfer method. Let $\mu(S)$ and $\mu(L)$ be the mean colors of the pixels in S and L, and $\sigma(S)$ and $\sigma(L)$ be the standard deviation of their illuminance. According to color transfer technique [RAGS01], by setting

$$\gamma = \frac{\sigma(L)}{\sigma(S)} \quad (2)$$

$$\alpha_k = \mu_k(L) - \gamma\mu_k(S), k \in \{R,G,B\} \quad (3)$$

and applying these four uniform parameters to each of pixels in S, the mean and the standard deviation of the resulting set S' would match those of L, which makes S' a shadow-free image like the lit region L. This is also the basic shadow removal idea of [SL08].

We observe that in the shadowed regions, the variation of attenuation factor γ is usually small, thus, we can use uniform value γ in shadowed areas in the model (1). As for $\alpha = L^d \cdot R$, the intensity parameter of reflected direct illumination, is the product of the direct illumination $L^d((x)$ and the reflectance $R(x)$. As the illumination $L^d(x)$ may be considered as uniform for the scenes in the image, however, the $R(x)$ may differ significantly for different materials. Thus, parameter α_k should be estimated accounting for the different materials with different reflectances $R(x)$. As shown in Figure 3, the uniform parameters α_k in [SL08] make the results covered by a layer of veil, since they do not account for the reflectance variation of the different materials.

Now we come to find an adaptive α for each pixel x in shadowed region S. let I_{avg} be shadowed pixel representing shadowed region S, with intensity $\mu(S)$, reflectance R_{avg}, the direct and ambient illumination at this pixel is L^d and L^a, and γ is the occluding parameter. Let the recovered intensity of I_{avg} is $\mu(L)$ (mean colors of the pixels in L). As $\mu(S)$ and $\mu(L)$ are the intensity values of I_{avg} under the shadowed and lit case, and since in the lit case, there are no occluder, thus there are no attenuation, so

$$\begin{aligned} I_{avg} &= \mu(S) = \tfrac{1}{\gamma} \cdot L^a \cdot R_{avg} \\ I_{recov} &= \mu(L) = (L^d + L^a) \cdot R_{avg} \end{aligned} \quad (4)$$

With Eq.(1), Eq.(2), Eq.(3) and Eq.(4), the recovered intensity under the reflected direct illumination is:

$$\begin{aligned} \alpha_{avg} &= \mu(L) - \gamma \cdot \mu(S) \\ &= (L^d + L^a) \cdot R_{avg} - L^a \cdot R_{avg} = L^d \cdot R_{avg} \end{aligned} \quad (5)$$

For each pixel x with intensity I and reflectance R in the shadowed regions S, under the reflected direct illumination L^d, its recovered intensity α under the reflected direct illumination is:

$$\alpha = L^d \cdot R \quad (6)$$

As we assume in the shadowed regions, the ambient illumination L^a is uniform, so for each pixel x, its intensity I is generated due to the ambient illumination:

$$I = \frac{1}{\gamma} \cdot L^a \cdot R \quad (7)$$

With Eq. (4), (5), (6), (7), we come to following results

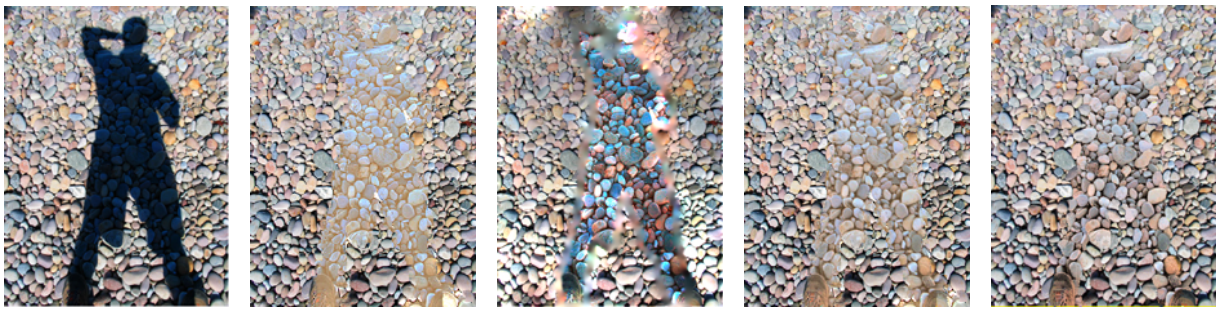
$$\frac{\alpha_{avg}}{\alpha} = \frac{L^d \cdot R_{avg}}{L^d \cdot R} = \frac{R_{avg}}{R} \quad (8)$$

$$\frac{I_{avg}}{I} = \frac{\frac{1}{\gamma} \cdot L^a \cdot R_{avg}}{\frac{1}{\gamma} \cdot L^a \cdot R} = \frac{R_{avg}}{R} \quad (9)$$

With Eq.(8) and Eq.(9), we receive

$$\alpha = \alpha_{avg} \cdot \frac{I}{I_{avg}} \quad (10)$$

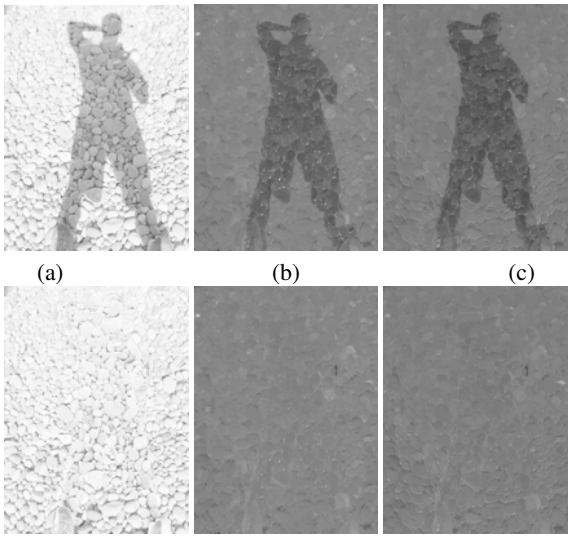
So for each pixel x in the shadowed regions S with the intensity value I, and with the $\alpha_{avg} = \mu(L) - \gamma \cdot \mu(S)$, $I_{avg} =$



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

Figure 3: *Shadow removal results comparison.(a) Input image, (b)shadow removal result of [SL08] (This picture is taken from the manuscript [SL08]), (c) result of [FHLD06], (d) result of [WTBS07], (e) our result.*

$\mu(S)$, its recovered intensity α under the reflected direct illumination can be estimated. Thus, although we do not know the reflectance R of pixel x, we can estimate its recovered intensity α under the reflected direct illumination using the information from both S and L.



(a) (b) (c)

(d) (e) (f)

Figure 4: *Shadow removal in la^*b^* color space. (a) (b) (c) are the l,a^*,b^* channel of the input image (Figure 3(a)), respectively, (d) (e) (f) are the shadow removal results of (a) (b) (c), respectively.*

We perform proposed shadow removal operation in the la^*b^* color space. As pointed out in [RCC98], compared with RGB color space, each component (L, a^* and b^* channel) of the la^*b^* color space is more uncorrelated for natural image. For example, in natural image, L channel varies widely from areas of bright illumination to those in shadow. As we remove the shadow in each channel dependently, thus, la^*b^* color space works better in our natural image shadow processing. We first transform the *RGB* color space into the la^*b^* color space, and perform shadow removal in the la^*b^* color space, and finally transform the shadow removal results back to the *RGB* color space. The main steps for the proposed shadow removal approach are as follows:
(1) Detecting the shadowed regions S and selecting a lit region L;
(2) With S and L, estimating the uniform parameters $I_{avg} = \mu_k(S), \mu_k(L), \sigma_k(S), \sigma_k(L), \alpha_{avg}$, and $\gamma = \sigma_k(L)/\sigma_k(S)$ based on global illumination transfer, $k \in \{l, a^*, b^*\}$;
(3) Compute the intensity α_k under the direct illumination for each pixel x using Eq.(10), $k \in \{l, a^*, b^*\}$;
(4) Recover the final intensity for pixel x: $I_k^{lit}(p) = \alpha_k(p) + \gamma(p) I_k^{shadow}(p)$, $k \in \{l, a^*, b^*\}$.

In Figure 4, we give shadow removal in la^*b^* color space for each channel of the input image, as illustrated in Figure 4, the shadow is removed in each channel. In Figure 3, using the proposed method, the shadow is removed much better than the previous method [SL08, WTBS07, FHLD06]. The texture structures and detail are better recovered.

5. Multi-scale shadow removal

Although our previous shadow removal method can recover the texture detail well in most cases, for some extremely complicated cases, where the shadowed regions are too dark and have heavy noise, or the edges information and texture detail in the shadowed regions have been weakened seriously due to the illumination occluding, our previous method may not work well, as illustrated in Figure 5. We present a multi-scale shadow removal method to address these problems. We decompose the image in a multi-scale detail levels using an edge-preserving filter, and remove the shadow in each level, then combine the shadow removal results into the final results in a spatially varying manner using the weighted average. This approach can receive more faithful results.

Let I be the input image. Using an edge preserving filter [FAR07, TM98], we compute M progressively coarser

version $I_1, \cdots, I_M$ of I, which suppress the noise while preserving the strong features in all subsequent decomposition iterations. We compute M subsequent image detail levels: $\{D_j\}_{j=1}^{M}$, where $D_j = I_j - I_{j-1}$. We then compute $\sigma_j(S)$, $\mu_j(S)$ and $\sigma_j(L)$, $\mu_j(L)$ for the S and L regions on each image detail level D_j, respectively, and recover the shadowless image D_j^{SL} for each D_j using the method described in previous section. Similarly, we also compute a shadowless image I_M^{SL} for the base image I_M. The final shadowless image of the input image I is defined as sum of I_M^{SL} and the weighted average of the shadowless image detail levels $\{D_j^{SL}\}_{j=1}^{M}$:

$$I^{sl} = I_M^{SL} + \omega \cdot \sum_{j=1}^{m} W_j D_j^{sl} / \sum_{j=1}^{m} W_j, 0 < \omega \leq 1 \qquad (11)$$

where the weight W_j is defined as:

$$W_j = g_\sigma * (j \cdot e^{-\frac{1}{n}\Sigma|\nabla I_j|})$$

n is the number of the image pixels, the Gaussian convolution g_σ is used to locally smooth the weight. This weighted average method prefers coarser detail level by giving them larger weight, because the color and luminance noise are smoothed more, and the prominent edges are better preserved in the coarser images. These properties help to recover the edge information and suppress the noise in the shadowed regions. In our experiments, we find out that setting $M = 3$ works well for generating good results.

Using this multi-scale shadow removal technique, both the prominent edges and color detail in the shadowed regions can be better recovered, and furthermore, the luminance noise has less influence on the final results. As shown in Figure 5, the results of multi-scale shadow removal are better than the results produced using a single level image technique.

6. Shadow boundaries processing

As the shadow boundaries usually change more rapidly than interior regions of the shadow, and the uniform direct illumination assumption may not work well to produce faithful results along the boundaries. In addition, for complicated shadows, to accurately detect the shadow boundaries is also a challenging work. Thus, we have to remove the shadow and illumination inconsistencies around the boundaries regions. We present a Bayesian shadow estimation approach to eliminate the shadows around the boundaries, and produce a consistent luminance transition between the recovered regions and the original lit regions.

According to [WT05,WTBS07], an image I of a scene can be defined as the following natural shadow equation $I = F\beta$, where F is the shadowless image, and β is a multi-channel shadow matte which is a factor to describe the attenuating effect of the illumination at each pixel. As we have known the F and I in the lit regions (F can be considered to be equal to I) and interior regions of the shadowed regions, according to $\beta = I/F$, the β in these regions can be computed (in the shadowed regions is below 1, while in the lit region equals to 1). Then with known F, I, β in the neighborhood of the boundaries, and the I value on the boundaries regions, we can estimate the F, β on the boundaries regions.

Inspired by [CCSS01], we present a Bayesian shadow edge estimation framework. By constructing oriented Gaussian distributions in the neighborhood containing computed β and F, we apply a continuously sliding neighborhood window that marches from known recovered regions and original regions inward the interior of the boundaries regions. We formulate the problem of estimating β and F in a Bayesian framework and solve it using the maximum a posteriori (MAP) technique. We find the most likely estimates for F, β given the existing intensity I:

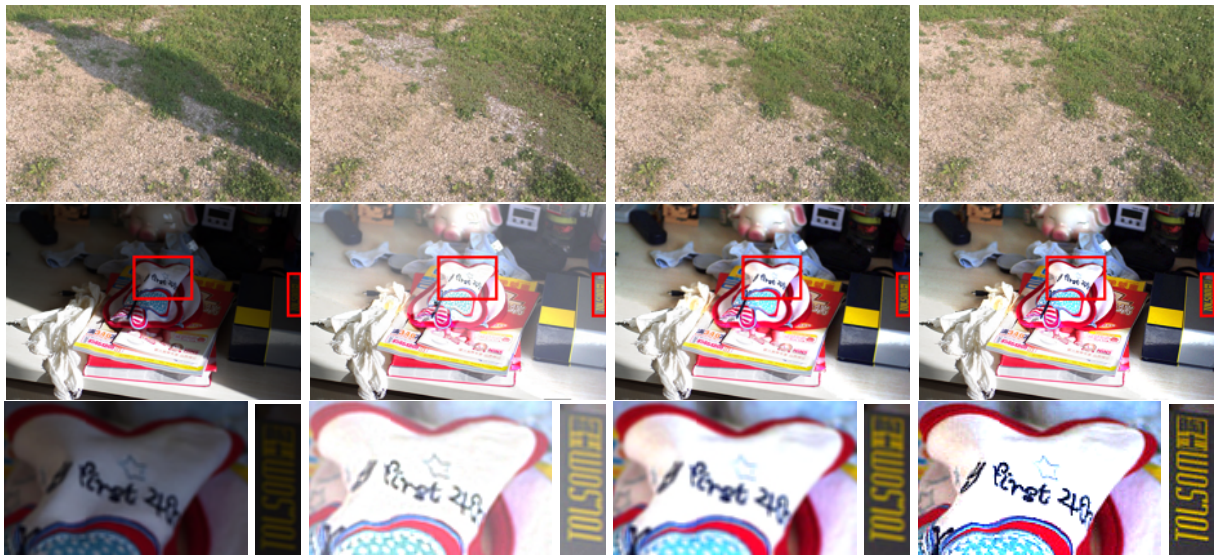
$$\begin{aligned} & argmax_{F,\beta} P(F,\beta|I) \\ = \; & argmax_{F,\beta} P(I|F,\beta)P(F)P(\beta) \\ = \; & argmax_{F,\beta} L(I|F,\beta) + L(F) + L(\beta) \end{aligned} \qquad (12)$$

where $L(\cdot) = \log P(\cdot)$. The log likelihoods $L(I|F,\beta)$ is defined as the difference between the observed color I and the color that would be predicted by the estimated F and β: $L(I|F,\beta) = -\parallel I - F\beta \parallel^2 / \sigma_I^2$, with standard deviation σ_I. With similar definition to [CCSS01], $L(F)$ is defined from an oriented elliptical Gaussian distribution. We also assume that the log likelihood for the opacity $L(\beta)$is constant and is omitted from the maximization procedure. The overall equation (12) can be efficiently optimized, and with the continuously sliding strategy, we can efficiently receive the final results for F and β. Compared with Bayesian shadow extraction approach [WT05], our method is more simplified and efficient, however, we found that the proposed method works well for our aim.

As shown in Figure 6, our method efficiently removes the shadows around the boundaries, and produces continuous illumination transition between the recovered regions and the original lit regions. Note that many image matting methods have been proposed, see [WC07] and more recent methods, some of these matting ideas can be applied to further improve the shadow boundaries processing.

7. Extension to Video

Our algorithm can be efficiently extended to video shadow removal. To perform video shadow detection, one method is to firstly detect the shadow regions in the selected key frames, then use the optical flow to track the spatiotemporal shadowed regions. However, as the video shadow is usually complex, automatically propagating shadow regions from one frame to others is difficult and unreliable, and the approach may not work well for video with sparkled shadowed regions and soft regions. As an alternative method, we consider the input video as the spatiotemporal video volume. By applying the user interface presented in [WBC*05] to paint spatiotemporal scribbles on the definitely shadowed regions



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

Figure 5: *Multi-scale shadow removal. (a) Input images, (b) shadow removal results of [SL08], (c) results of our single level shadow removal, (d) results of our multi-scale shadow removal.*

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

Figure 6: *Boundary processing. (a) Input image, (b) our result without using shadow boundary processing, (c) our result using shadow boundary processing, (d) shadow removal result of [SL08].*

and lit regions in the video volume, the GMM brush used in image case can be easily extended to video volume to select the shadowed spatiotemporal regions S_v.

For video shadow removal, it is important to keep the shadow removed video temporally consistent, if we remove shadow frame by frame, and only using the information of the current frame, the resulted video may have such artifacts as flickering. To address this issue, we estimate the uniform parameters $I_{avg} = \mu_k(S), \sigma_k(S)$, for each frame i in S_v using the color information of the neighboring frames of frame i, in our experiments, we use four neighboring frames. In a similar way, the $\mu_k(L)$ and $\sigma_k(L)$ of frame i are also computed using the information of neighboring frames in lit spatiotemporal regions L_v. Then the intensity $\alpha_k^{i,x}$ for each pixel x of the frame i is computed using the weighted average of the intensity α_k of the spatio-temporal neighboring pixels $\alpha_k^{i,x} = \sum_{y \in N(x)} \omega_y \alpha_k^y$, and $\omega_y = e^{-|y-x|^2/2\sigma^2}$. Finally we recover the intensity for pixel x in the same way as image case. For video shadow boundaries processing, all the probability distribution P, log likelihood $L(I|F,\beta)$ and $L(F)$ can be

defined in spatio-temporal regions. We apply the Bayesian shadow estimation in the spatio-temporal regions for each pixel in the boundary regions, and remove the shadow artifacts in the boundary regions.

In Figure 7, we present video shadow removal result. The size for the videos is $538 \times 303 \times 50$. As shown in the supplemental videos, we receive shadow free and temporal consistent results. As the shadow region in the video is relative small, in our experiment, for each frame, we spend less than one second to remove shadow. Note that in this video, the shadow regions do not show rapid motion, thus, we need not to take motion compensation into condition. In the future, we will work on video shadow removal with more complex shadow regions.

8. Results and discussion

In this section, we show the efficiency of the proposed method by processing a variety of images. We also present comparison results on both shadow removal performance and quality with some of the most related methods. Limitations of the proposed method are also presented. All the results are implemented using C++ on a machine equipped with Pentium(R) Dual-Core CPU E5200@2.50GHz with 2GB RAM.

In Figure 5, we compare the results using and without using multi-scale shadow removal technique. In our experiments, we find that in most cases, our single scale shadow removal method can receive satisfactory results. When the shadowed regions are too dark (heavily shadowed) or contain very fine texture structures, we have to resort to the multi-scale strategy to receive better results. As shown in Figure 5, for these images with complicated texture detail, the results of our single scale shadow removal method are better than [SL08], while with the multi-scale strategy, the fine texture detail recovering are further improved. The noise in the recovered regions is also efficiently suppressed. Using the accelerated multi-scale image decomposition [FAR07], multi-scale strategy can be performed efficiently.

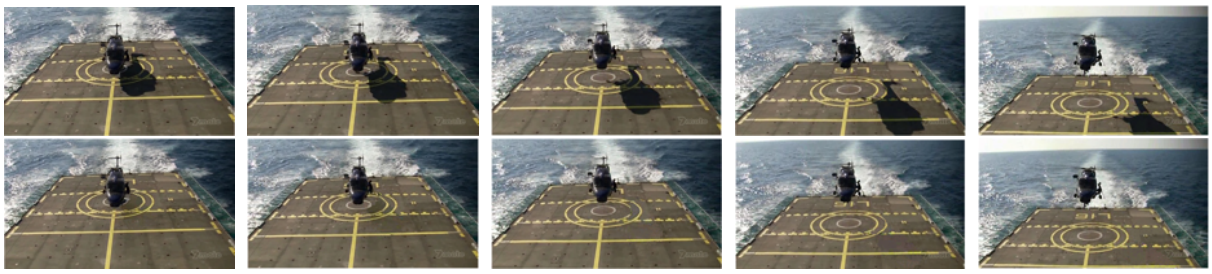
In Figure 6, we give the comparison results on the shadow boundaries processing. As shadow boundaries usually change rapidly and nonlinearly, thus, direct interpolation between the original and the recovered regions cannot produce faithful and visually natural results. Inpainting based border processing techniques [SL08] can produce visually pleasing results for images with abundant self-similar texture information, while this method cannot receive faithful shadow-free boundaries since inpainting techniques modify the image boundaries content. Especially for shadow boundaries with sharp structures, the artifacts are more noticeable, as shown in Figure 6. Our method efficiently removes the shadow around the boundaries while not modifying the content of the image.

In Figure 3 and Figure 8, we compare with several most related methods [FHLD06, WTBS07, SL08, GDH11]. Finlayson and colleagues [FHLD06] treated shadow removal as a reintegration problem based on precise shadow edge detection. As shown in Figure 3 and Figure 8, the unpleasing results happen on the boundaries. Furthermore, the fine texture detail is also not recovered well. As Shor et al. [SL08] applied uniform affine parameters in illumination transfer to remove the shadows, even they applied the pyramid-based restoration for further enhancing texture contrast and and reducing noise, sometimes the fine texture detail still cannot be efficiently recovered. For example, for some complex scenes such as the stone texture image in Figure 3, as each stone may have different reflectance, uniform reflectance parameter setting may make the results covered by a layer of by a layer of veil. To process shadowed scenes with weak texture and structures, the approach [WTBS07] can receive pleasing results, as shown in Figure 8(e). To process complex scenes, this method [WTBS07] still cannot receive satisfactory resultsčňas illustrated in Figure 3 (d) and Figure 8 (a)(b)(c)(d). Guo et al. [GDH11] relighted the whole image based on the shadow coefficient value estimated using matting technique [LLW08]. This method works well for seamless transition between the recovered regions and the original lit regions, but still cannot recover the texture detail efficiently in the interior regions, as illustrated in Figure 8.

In Figure 3 and Figure 8, to make a fair comparison, all the methods work on the same shadowed regions with the same boundaries. When the shadowed image has several parts with different reflectance, as shown in some images in Figure 8 (b), one plausible way to receive satisfactory results is to find a corresponding lit region for each different part in the shadowed regions, and then estimate the parameters, respectively. While this method is tedious for complex scene, in addition, it may not produce continuous results between adjacent parts. In our method, in all the examples, we use only a single lit region for the shadow removal parameters estimation. Also note that in Figure 3 and Figure 8, our results is derived using a single level shadow removal method, not using multi-scale strategy. In shadow detection, we perform the GMM brush on the coarse image with level $M = 3$. In Figure 8, we also present ground truth shadow-free images for some of the shadowed image. The shadow-free images come from the manuscript [GDH11]. Note that for comparison with [GDH11], we use the shadow removal results from the manuscript [GDH11].

Figure 9 shows the shadow removal result of a fresco from Mogao Caves of Dunhuang. The frescos of Mogao Caves are precious materials for researching on the Chinese art history. In this Figure, we successfully remove the shadow of the fresco and the image details are also well recovered.

The time consumption of the proposed method is consisted of three stages: shadow detection, shadow removal, and boundaries processing. Using our method, for an image with size of 800×600, it usually takes 0.4 to 0.6 seconds for shad-



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

Figure 7: *Video shadow removal results. Top row, the five frames of the input video; bottom row, the corresponding shadow removal results.*

ow detection, 0.3 to 0.5 seconds in shadow removal stage for one single scale, and 1 to 2 seconds for boundaries processing. Given an input image, it should be noted that the computational time also depends on the size and the shape of the shadowed regions needed to be processed.

To remove the shadow regions for an image with size of 800×600, it takes about 60s to 80s, 1s to 2s, 90s to 120s, 60s to 70s and 0.3s to 0.5s for [WTBS07], [SL08], [FHLD06], [GDH11], and our method, respectively. As method [FHLD06] has to process a reintegration method, which may be little time-consuming. Minimizing an energy minimization [WTBS07] also has relative high computational complexity. For method [GDH11], to estimate a fractional shadow coefficient value using closed-from matting [LLW08] is time and memory consuming for solving a linear system. Our method and [SL08] are much faster by estimating shadow removal parameters. In most cases, Shor et al. [SL08] have to apply the pyramid-based restoration to recover the image detail, which makes it a little more time-consuming than our method. While our method can produce satisfactory results using one single scale in most cases, and without resorting to multi-scale techniques.

(a) (b)

Figure 9: *(a) Input image with shadow, (b) the shadow removal result using the proposed method.*

Limitations: Although our method can produce impressive results, for some complex images, to derive satisfactory results, the selection of the lit regions is important for our method. As shown in Figure 10, the texture materials of the wood block are complex, and the illumination conditions are also complex, there are umbra, penumbra and other complex soft shadows presented in the images. Our method is based on the illumination transfer technique, which transfers the illumination information of the lit regions to the shadowed regions to perform shadow removal. Thus, in these complex cases, to receive satisfactory results, we have to select appropriate lit regions as samples, and there must be a coherent lit region in the image that shows all materials existed in the shadow region. As illustrated in Figure 10, appropriate lit regions produce much better result.

Figure 10: *Limitation. Left column, the selected different lit regions on the input image. Right column, corresponding shadow removal results.*

9. Conclusion and future work

We present a fast image removal approach from a single image using an adaptive multi-scale illumination transfer method. By estimating different reflectance for different scene materials from the the lit regions, we can obtain a high-quality shadow-free image. Our method is simple but effective, and is easy for implementing. We demonstrate that our approach can produce results that are in most cases superior in quality to those of previous shadow removal methods.

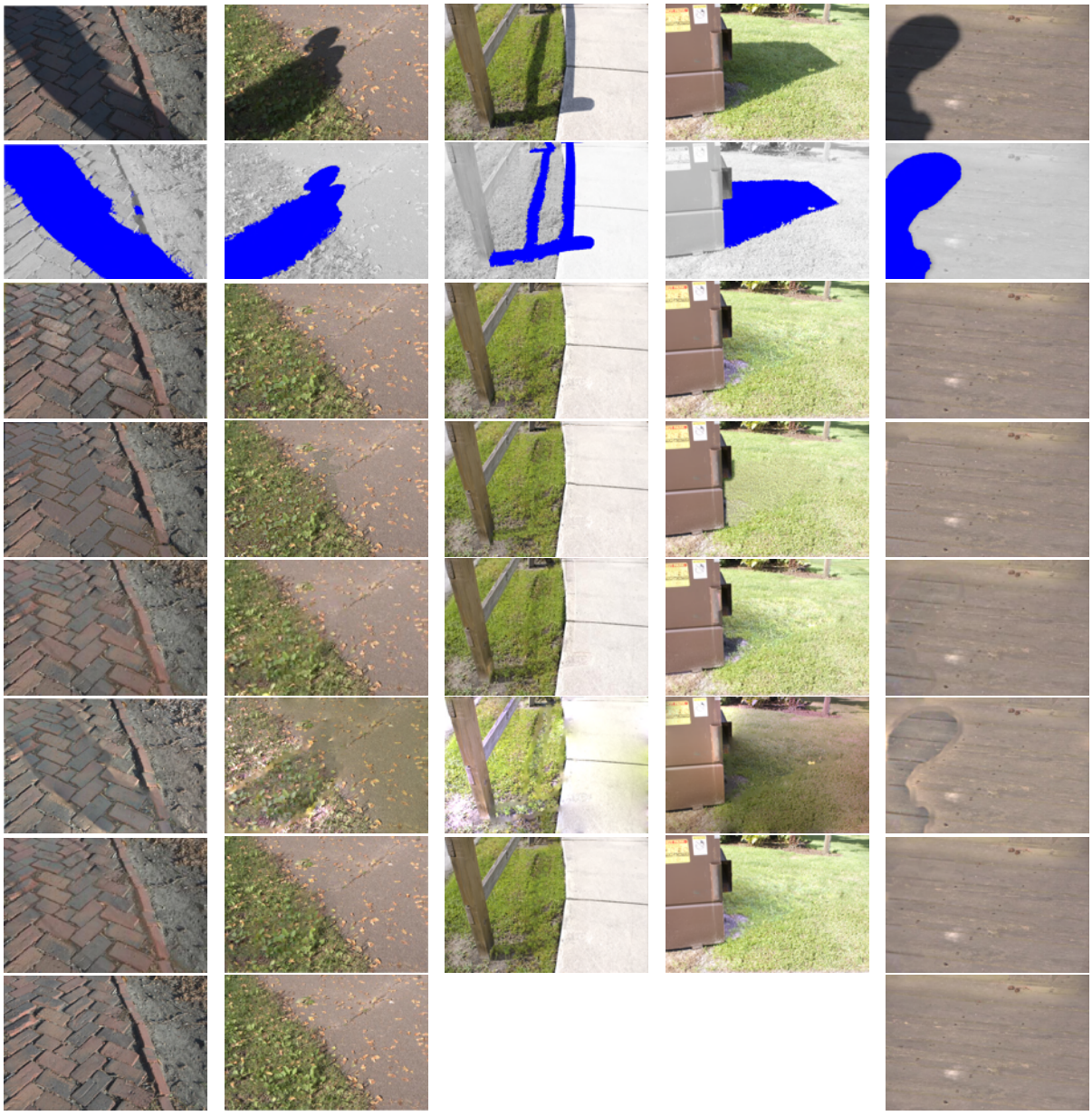


Figure 8: *Shadow removal results comparison: The first row: input shadowed images; The second row: detected shadow; the third row: our results; the fourth row: the results of [SL08]; the fifth row: the results of [GDH11]; the sixth row: the results of [FHLD06]; the seventh row: results of [WTBS07]; the eighth row: ground truth shadow-free images. Please zoom in the images for displaying the difference. Note that for both the results of [GDH11] and ground truth results, we take the pictures directly from the manuscript [GDH11].*

In the future, we will work on image shadow removal under several different light sources, which is more challenging while can make the proposed method more general and applicable. As shadow removal and modification help to produce more visually pleasing results in image editing propagation, material editing, recoloring, in the future, we would like to take the shadow processing involved for more pleasing image editing results.

Acknowledgment

We would like to thank the anonymous reviewers for their valuable comments. This work was partly supported by the National Basic Research Program of China (No. 2012CB725303), the NSFC (No. 61070081, and No. 41271431), the Open Project Program of the State Key Lab of CAD&CG(GrantNo. A1208), the Luojia Outstanding Young Scholar Program of Wuhan University, the Project of the Science and Technology Plan for Zhejiang Province (Grant No.2012C21004), the Fundamental Research Funds for the Central Universities.

References

[Aga07] AGARWALA A.: Efficient gradient-domain compositing using quadtrees. In *ACM Transactions on Graphics (TOG)* (2007), vol. 26, ACM, p. 94. 2

[AHO07] ARBEL E., HEL-OR H.: Texture-preserving shadow removal in color images containing curved surfaces. In *CVPR'07* (2007), IEEE, pp. 1–8. 2

[AHO11] ARBEL E., HEL-OR H.: Shadow removal using intensity surfaces and texture anchor points. *PAMI*, 99 (2011). 2

[ANBSE11] AL-NAJDAWI N., BEZ H., SINGHAI J., EDIRISINGHE E.: A survey of cast shadow detection algorithms. *Pattern Recognition Letters* (2011). 2

[BA03] BABA M., ASADA N.: Shadow removal from a real picture. In *ACM SIGGRAPH 2003 Sketches & Applications* (2003), ACM, pp. 1–1. 4

[BTSD78] BARROW H., TENENBAUM J., SCIENCE S. I. A. I. C. C., DIVISION T.: *Recovering intrinsic scene characteristics from images*. Artificial Intelligence Center, SRI International, 1978. 1, 4

[CCSS01] CHUANG Y., CURLESS B., SALESIN D., SZELISKI R.: A bayesian approach to digital matting. In *CVPR 2001* (2001), vol. 2, IEEE, pp. II–264. 7

[CGC*03] CHUANG Y., GOLDMAN D., CURLESS B., SALESIN D., SZELISKI R.: Shadow matting and compositing. *ACM Transactions on Graphics 22*, 3 (2003), 494–500. 3

[FAR07] FATTAL R., AGRAWALA M., RUSINKIEWICZ S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM Transactions on Graphics 26*, 3 (2007), 51. 4, 6, 9

[FDL04] FINLAYSON G., DREW M., LU C.: Intrinsic images by entropy minimization. *ECCV 2004* (2004), 582–595. 2

[FF05] FREDEMBACH C., FINLAYSON G.: Hamiltonian path based shadow removal. In *16th British Machine Vision Conf, BMVC2005* (2005), pp. 970–980. 2

[FF06] FREDEMBACH C., FINLAYSON G.: Simple shadow removal. In *ICPR 2006* (2006), vol. 1, IEEE, pp. 832–835. 2

[FFLS08] FARBMAN Z., FATTAL R., LISCHINSKI D., SZELISKI R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. In *ACM Transactions on Graphics (TOG)* (2008), vol. 27, ACM, p. 67. 4

[FHD06] FINLAYSON G., HORDLEY S., DREW M.: Removing shadows from images. *ECCV 2006* (2006), 129–132. 2

[FHLD06] FINLAYSON G., HORDLEY S., LU C., DREW M.: On the removal of shadows from images. *PAMI 28*, 1 (2006), 59–68. 1, 2, 6, 9, 10, 11

[GDH11] GUO R., DAI Q., HOIEM D.: Single-image shadow detection and removal using paired regions. In *CVPR 2011* (2011), IEEE, pp. 2033–2040. 2, 3, 9, 10, 11

[LEN10] LALONDE J., EFROS A., NARASIMHAN S.: Detecting ground shadows in outdoor consumer photographs. *Computer Vision–ECCV 2010* (2010), 322–335. 2

[LG08] LIU F., GLEICHER M.: Texture-consistent shadow removal. *Computer Vision–ECCV 2008* (2008), 437–450. 2

[LLW08] LEVIN A., LISCHINSKI D., WEISS Y.: A closed-form solution to natural image matting. *PAMI 30*, 2 (2008), 228–242. 2, 3, 9, 10

[MCT07] MOHAN A., CHOUDHURY P., TUMBLIN J.: Editing soft shadows in a digital photograph. *IEEE Computer Graphics and Applications* (2007), 23–31. 3

[MFS08] MAXWELL B., FRIEDHOFF R., SMITH C.: A bi-illuminant dichromatic reflection model for understanding images. In *CVPR 2008* (2008), IEEE, pp. 1–8. 2

[RAGS01] REINHARD E., ADHIKHMIN M., GOOCH B., SHIRLEY P.: Color transfer between images. *Computer Graphics and Applications, IEEE 21*, 5 (2001), 34–41. 1, 2, 3, 5

[RCC98] RUDERMAN D. L., CRONIN T. W., CHIAO C.-C.: Statistics of cone responses to natural images: implications for visual coding. *JOSA A 15*, 8 (1998), 2036–2045. 6

[RKB04] ROTHER C., KOLMOGOROV V., BLAKE A.: Grabcut: Interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics (TOG)* (2004), vol. 23, ACM, pp. 309–314. 3, 4

[SL08] SHOR Y., LISCHINSKI D.: The shadow meets the mask: Pyramid-based shadow removal. In *Computer Graphics Forum* (2008), vol. 27, Wiley Online Library, pp. 577–586. 1, 2, 3, 5, 6, 8, 9, 10, 11

[TM98] TOMASI C., MANDUCHI R.: Bilateral filtering for gray and color images. In *ICCV 1998* (1998), pp. 839–846. 4, 6

[WBC*05] WANG J., BHAT P., COLBURN R., AGRAWALA M., COHEN M.: Interactive video cutout. In *ACM Transactions on Graphics (TOG)* (2005), vol. 24, ACM, pp. 585–594. 7

[WC07] WANG J., COHEN M.: Image and video matting: a survey. *Foundations and Trends® in Computer Graphics and Vision 3*, 2 (2007), 97–175. 7

[WT05] WU T., TANG C.: A bayesian approach for shadow extraction from a single image. In *ICCV 2005* (2005), vol. 1, IEEE, pp. 480–487. 2, 3, 7

[WTBS07] WU T., TANG C., BROWN M., SHUM H.: Natural shadow matting. *ACM Transactions on Graphics (TOG) 26*, 2 (2007), 8. 2, 3, 6, 7, 9, 10, 11

[XG12] XIAO C., GAN J.: Fast image dehazing using guided joint bilateral filter. *The Visual Computer 28*, 6-8 (2012), 713–721. 1

[XL10] XIAO C., LIU M.: Efficient mean-shift clustering using gaussian kd-tree. In *Computer Graphics Forum (Pacific Graphics 2010)* (2010), vol. 29, Wiley Online Library, pp. 2065–2073. 4

[XY*11] XIAO C., YONGWEI N., ET AL.: Efficient edit propagation using hierarchical data structure. *Visualization and Computer Graphics, IEEE Transactions on*, 99 (2011), 1135–1147. 4

[ZSMT10] ZHU J., SAMUEL K., MASOOD S., TAPPEN M.: Learning to recognize shadows in monochromatic natural images. In *CVPR 2010* (2010), IEEE, pp. 223–230. 2