A Bayesian Approach for Shadow Extraction from a Single Image [*]

Tai-Pang Wu and Chi-Keung Tang

Vision and Graphics Group
The Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Abstract

This paper addresses the problem of shadow extraction from a single image of a complex natural scene. No simplifying assumption on the camera and the light source other than the Lambertian assumption is used. Our method is unique because it is capable of translating very rough user-supplied hints into the effective likelihood and prior functions for our Bayesian optimization. The likelihood function requires a decent estimation of the shadowless image, which is obtained by solving the associated Poisson equation. Our Bayesian framework allows for the optimal extraction of smooth shadows while preserving texture appearance under the extracted shadow. Thus our technique can be applied to shadow removal, producing some best results to date compared with the current state-of-the-art techniques using a single input image. We propose related applications in shadow compositing and image repair using our Bayesian technique.

1 Introduction

Shadows are problematic to many computer vision algorithms that assume intensity or color constancy. In image segmentation, unwanted shadow boundaries may be detected together with the true object boundaries. In motion estimation, the computation of optical flows are susceptible to illumination changes caused by cast shadows. The problem of shadow removal becomes particularly challenging if only a single and uncalibrated image of a complex scene is given. While previous works using a single image (e.g., [7, 6]) have produced very good results on automatic shadow detection and removal, several assumptions were made about the light source and the camera. If complex textures are present in the image, they may be blurred after shadow removal.

In this paper, we present an alternative approach for shadow extraction from a single image. Our system accepts a few simple and rough hints supplied by the user. The rest of the processing is automatic: the user-supplied hints are first translated into pertinent color statistics, which will be used for the effective encoding of the likelihood and prior probabilities in our Bayesian optimization framework. Using our image model, texture gradients and shadow smoothness are simultaneously optimized. Thus, our method not only produces a shadow-free image where textures are preserved, but also extracts a smooth shadow from the image. Besides the Lambertian assumption, we make no other simplifying assumption on the light source and the camera.

1.1 Related work

Shadow removal In computer vision, shadow and shading removal is often addressed using intrinsic image estimation. For example, Weiss [15] used a sequence of N images to derive one reflectance image and N illumination images. Matsushita et al [10] estimated intrinsic images from image sequences with biased illumination. In [14], Tappen et al separated intrinsic images from a single image by classifying image derivatives as changes due to either reflectance or shading, followed by belief propagation to correct ambiguous regions. Recently, Finlayson et al [6] described an automatic shadow removal algorithm using one image only. The algorithm uses entropy minimization to derive an illumination invariant grayscale image for shadow removal, without resorting to any calibration, an important advancement over an earlier work [7]. The technique makes several assumptions, such as Planckian lighting (e.g., sunlight) and narrow-band cameras [7, 6]. While the above approaches are capable of identifying shadow and producing shadow-free images, they are not designed to *extract* the shadow from the image. The shadows obtained using these shadow removal algorithms may not be usable. More importantly, if the image contains complex textures and patterns under the extracted shadow, the above methods may blur them after shadow removal.

Intensity transfer We believe that user interaction is necessary for general shadow extraction or removal from an uncalibrated image, because without prior knowledge it is difficult for even the human eyes to distinguish whether
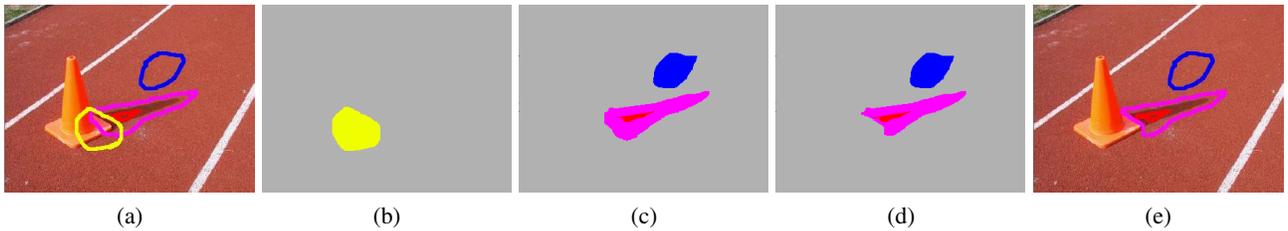
Figure 1: (a) The quadmap [17]: $\mathcal{S}$ = shadowed (red), $\overline{\mathcal{S}}$ = nonshadowed (blue), $\mathcal{U}$ = uncertain (pink), and $\mathcal{E}$ = excluded (yellow) regions marked by the user. (b) shows $\mathcal{E}$ (the yellow region) indicated in (a). (c) shows $\mathcal{S}$(red region), $\overline{\mathcal{S}}$(blue region) and $\mathcal{U}$(pink region) indicated in (a). (d) and (e) show the refined quadmap, where the cone is automatically excluded from all shadow calculations.

an image was captured under dim environment or under shadows if no shadow-free cues are present in the image. Given a few rough user-supplied hints, it may be possible to perform color transfer to remove shadows. For example, in [12, 16], the colors or intensities are assumed to follow a single Gaussian distribution. In that case, it is possible to transfer to the shadowed region the mean and standard deviation of the shadow-free colors. If the single Gaussian assumption is violated, however, the user needs to divide the shadow-free/shadowed regions into smaller patches manually to satisfy the assumption. However, even with additional user interaction, the color transfer approach may still produce unacceptable results because there is no provision for maintaining spatial coherence or handling textures or scene structure, especially along hard shadow boundary.

Reflection separation In [9], a user-assisted method was proposed to perform reflection separation and one of its applications is shadow removal. Because it is a general separation approach, it requires users to indicate shadow boundaries and depends on higher order information such as the gradient of shadow and shadowless image. Soft shadow cannot be handled.

Image matting It is possible to use natural image matting techniques [2, 13] to extract shadows from a single image. However, these approaches are designed to extract foreground objects. The underlying texture and structure under the shadow may therefore be extracted together with the shadow. They use a user-specified trimap, which is insufficient in distinguishing the shadow from the background when the shadow-casting object is present. The only approach that directly addressed shadow matting [3] makes several assumptions about the scene lighting (e.g., a primary point light source) and requires a video sequence to solve for the shadow matte. Only hard shadows are extracted in [3].

1.2 Overview of our approach

We analyze and formulate the shadow extraction problem by using the Bayesian framework. Our system requires very simple user input (quadmap) as opposed to the

trimap used in the majority of image matting system, where a somewhat accurate marking step is required. The user-supplied hints are employed to effectively impose useful constraints for solving the difficult and ill-posed shadow extraction problem from one image. The quadmap [17] indicates to the system useful hints on the shadowed and shadowless regions, making it possible to simplify and solve the problem in two steps. The first step computes a shadowless image estimate by making use of the hints only. A Poisson equation is formulated and solved. The second step performs shadow optimization by using the initial shadowless image and a smooth shadow prior. Using our image model, the final shadowless image is computed by dividing the input image with the estimated shadow image. Note that except for the user-supplied quadmap, all the processing are performed automatically.

1.3 Outline of the paper

Our paper is organized as follows: Section 2 defines the image model. Section 3 formulates our Bayesian shadow extraction, where the likelihood and the prior will be defined. The shadowless image estimate employed in our likelihood function will be detailed, together with the definition of the smooth shadow prior. Section 4 compares shadow removal results generated by our method and the method in [6] using a single image, and presents applications using our Bayesian shadow extraction. The applications include lens flare removal and image compositing with the extracted shadow. Section 5 discusses the limitation of our method and proposes future work.

2 Image model

We start from the image formation equation. In [1], an image of a Lambertian scene is the pixelwise product of its two corresponding intrinsic images defined as:

$$I = \mathcal{R}\mathcal{L} \tag{1}$$

where $\mathcal{R}$ is the albedo image and $\mathcal{L}$ is the illumination image. $\mathcal{L}$ encapsulates illumination, including both shading and shadows. From Eqn. 1, we can further decompose $\mathcal{L}$ into a multi-channel shadow image β, and another image

$\mathcal{L}'$, such that $\mathcal{L} = \beta\mathcal{L}'$. A multi-channel shadow image used in our equation allows the three color channels to respond differently to shadows. We can express Eqn. 1 as $I = \beta\mathcal{R}\mathcal{L}'$ or simply

$$I = \beta\mathcal{F} \qquad (2)$$

where $\mathcal{F}$ is the shadowless image. This image model simply means that the observed shadowed image I is the shadowless image $\mathcal{F}$ scaled by the shadow image β. I is expressed as a pixelwise product of $\mathcal{F}$ and β in Eqn. 2 and thus they are interdependent. This means that when one term is optimized, the other term can be optimized implicitly.

3 Bayesian shadow extraction

Since $I = \beta\mathcal{F}$, we focus on how to find β because $\mathcal{F} = \frac{I}{\beta}$. Note that a shadow responds differently for different channels and so our shadow extraction is performed on each channel independently. So, starting from section 3.2, we shall assume a single-channel image.

3.1 Quadmap

It is insufficient to find the β image given I only, because we need to provide color samples for the shadowed and non-shadowed colors. We use the quadmap introduced in [17] to sample colors. The idea is in fact very neat and simple. The user-specified quadmap consists of four regions: "definitely shadowed regions" $\mathcal{S}$, "definitely nonshadowed regions" $\overline{\mathcal{S}}$, "uncertain regions" $\mathcal{U}$, and "excluded regions" $\mathcal{E}$ when the shadow-casting object is present, as shown in Fig. 1. $\overline{\mathcal{S}}$ indicates shadow-free samples of the shadowed region. $\mathcal{S}$ is the darkest shadow region where β is roughly constant spatially. $\mathcal{S}$ and $\overline{\mathcal{S}}$ should therefore be similar in texture in order to produce reasonable result. All these regions are easily marked by the user.

While the quadmap is similar in concept to the trimap in Bayesian matting [2], there are two major differences. First, because the shadow casting object typically has very different color statistic from the shadow and shadow-free background, it should be excluded from consideration. Second, the processing of $\mathcal{U}$, as we shall see, is very different from that of the uncertain region in a trimap. In both cases, although the uncertain region indicates where the model parameters are estimated, a typical uncertain region $\mathcal{U}$ in a quadmap is much larger than one in a trimap. Consequently, the user can mark a quadmap more easily. This is because our Bayesian shadow optimization only requires color statistics information, as opposed to [2] where the trimap needs to supply spatial cues as well.

In implementation, $\mathcal{S} \cup \mathcal{U}$ will be processed to extract the shadow. $\mathcal{E}$, the region where some color static will be excluded, will be refined automatically by identifying pixels
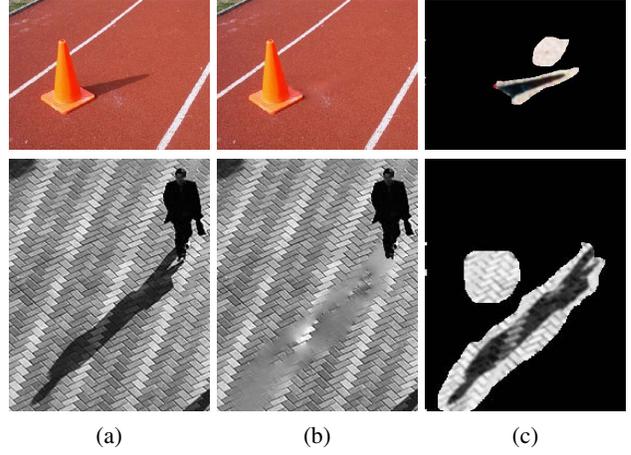

Figure 2: Running example: (a) input, (b) $\hat{\mathcal{F}}^*$, which is used to encode the likelihood $L(\hat{\mathcal{F}}^*|\beta)$, (c) the normalized affinity map, which is used as a rough β estimate for computing the $\hat{\mathcal{F}}^*$. Note that the region of black pixels in (c) are excluded from all calculations.

whose shadow and shadow-free probability density functions (pdfs) are near zero. The shadow and shadow-free pdfs are estimated as Gaussian Mixture Models (GMM), which are derived from the histogram of $\mathcal{S}$ and $\overline{\mathcal{S}}$. Fig. 1(d) and (e) illustrates the refined quadmap where $\mathcal{E}$ refines $\mathcal{U}$ to exclude irrelevant statistics of the shadow casting object.

3.2 The Bayesian formulation

Given these four regions and the input image I, we solve for β by using the Bayesian framework. In other words, we solve:

$$\begin{aligned} \beta^* &= \arg\max_\beta P(\beta|\mathcal{S},\overline{\mathcal{S}},\mathcal{U},\mathcal{E}) \\ &= \arg\max_\beta P(\mathcal{S},\overline{\mathcal{S}},\mathcal{U},\mathcal{E}|\beta)\mathcal{P}(\beta) \qquad (3) \end{aligned}$$

To solve Eqn. 3 once and for all is very difficult. Thus, we propose a two-step approach. By considering the four regions only, we analyze and gather pertinent information from the user-supplied cues.

Since $\mathcal{E}$ excludes irrelevant pixels, which do not provide any useful cue to solve for β or $\mathcal{F}$, so we focus on the other three regions. $\mathcal{S}$ and $\overline{\mathcal{S}}$ are two so-called "corresponding" regions, as they are marked because they have similar texture patterns but mostly differs in intensity. It is therefore possible to estimate β or $\mathcal{F}$ inside $\mathcal{S}$ by considering $\mathcal{S}$ and $\overline{\mathcal{S}}$ only. For $\mathcal{U}$, it can be thought as the mixture of the pdfs given by $\mathcal{S}$ and $\overline{\mathcal{S}}$. So, if the separation of the mixture can be performed, it is possible to estimate β or $\mathcal{F}$.

Thus, given the four regions, it is possible to find a rough estimation $\hat{\mathcal{F}}$ of $\mathcal{F}$ by solving:

$$\hat{\mathcal{F}}^* = \arg\max_{\hat{\mathcal{F}}} P(\hat{\mathcal{F}}|\mathcal{S},\overline{\mathcal{S}},\mathcal{U},\mathcal{E}) \qquad (4)$$

And consequently, Eqn. 3 is rewritten into:

$$\beta^* = \arg\max_{\beta} P(\hat{\mathcal{F}}^*|\beta)P(\beta)$$

$$= \arg\min_{\beta} L(\hat{\mathcal{F}}^*|\beta) + L(\beta) \quad (5)$$

where $L(\cdot) = \log P(\cdot)$.

3.3 $\hat{\mathcal{F}}^*$ estimation

Solving Eqn. 4 for an optimal solution is still difficult. We propose a fast method based on solving a Poisson equation to estimate a decent $\hat{\mathcal{F}}^*$ (Fig. 2(b)), which is optimal subject to the guidance field to be introduced. While $\hat{\mathcal{F}}^*$ is insufficient by itself as the output shadowless image $\mathcal{F}$, it provides sufficient guidance for solving Eqn. 5 when cooperated with our β prior.

The basic idea on estimating $\hat{\mathcal{F}}^*$ is that we maintain the image gradient during intensity transfer, which is guided by certain confidence measurement. Before describing how to use a Poisson equation to estimate the parameters inside $\mathcal{U}$, we describe how to estimate the parameters inside $\mathcal{S}$ first, because it is a much easier problem and the parameters are required for solving the Poisson equation.

Estimating β inside $\mathcal{S}$ As we mentioned, the β inside $\mathcal{S}$ is roughly constant. Since $\mathcal{S}$ and $\overline{\mathcal{S}}$ have similar textures, using our image model defined in Eqn. 2, the mean intensity of the pixels inside $\overline{\mathcal{S}}$ and the mean intensity of $\mathcal{F}$ inside $\mathcal{S}$ should be very similar. We let $I_{m\mathcal{S}}$ and $I_{m\overline{\mathcal{S}}}$ be the mean intensities of region $\mathcal{S}$ and $\overline{\mathcal{S}}$ respectively, thus the β inside $\mathcal{S}$ is given by $\beta = I_{m\overline{\mathcal{S}}}/I_{m\mathcal{S}}$.

Estimating β inside $\mathcal{U}$ Here, we suggest a method to approximate the β inside $\mathcal{U}$. We consider the *model affinity* of each pixel inside $\mathcal{U}$ to the respective probability distribution of $\mathcal{S}$ and $\overline{\mathcal{S}}$. This roughly tells us the respective contribution of the two distributions that results in the observed pixel's intensity. As we shall see, it also provides the approximate β inside $\mathcal{U}$ (Fig. 2(c)).

1. Let the shadow and shadow-free pdfs be the GMMs of $\mathcal{S}$ and $\overline{\mathcal{S}}$ respectively: $\{G_{\overline{\mathcal{S}}}(i; \mu_{\overline{\mathcal{S}}}(i), \sigma_{\overline{\mathcal{S}}}(i))\}$ and $\{G_{\mathcal{S}}(i; \mu_{\mathcal{S}}(i), \sigma_{\mathcal{S}}(i))\}$, $i = 1, 2, \cdots$. These GMMs are obtained by considering the respective color histograms in $\mathcal{S}$ and $\overline{\mathcal{S}}$. Typically, the number of Gaussians seldom exceeds three because of the small size of $\mathcal{S}$ and $\overline{\mathcal{S}}$. If the image can be roughly modelled by Eqn. 2, the $\{G_{\overline{\mathcal{S}}}\}$ and $\{G_{\mathcal{S}}\}$ that are respectively sorted by μ's are corresponding because Eqn. 2 describes a linear transformation.

2. Then, we compute the *affinity map* $\wp_{\mathcal{S}}(x)$ for encoding the model affinity of pixel x to $\mathcal{S}$'s pdf. Define $\wp_{\mathcal{S}}(x) = \frac{1}{2}\sum_{i=1}^{K}(\pi_{\mathcal{S}}^i + \pi_{\overline{\mathcal{S}}}^i)\wp_{\mathcal{S}}^i(x)$, where $\pi_{\mathcal{S}}^i, \pi_{\overline{\mathcal{S}}}^i$ are the respective weights of the Gaussian, with each

$\wp_{\mathcal{S}}^i(x)$ encoding the shadow probability of x by the normalized color distance: $\wp_{\mathcal{S}}^i(x) = \frac{d_{\mathcal{S}}^i(x)}{d_{\mathcal{S}}^i(x)+d_{\overline{\mathcal{S}}}^i(x)}$ where $d_{\mathcal{S}}^i(x) = (I(x)-\mu_{\mathcal{S}}^i)/\sigma_{\mathcal{S}}^i$ and $d_{\overline{\mathcal{S}}}^i(x) = (I(x)-\mu_{\overline{\mathcal{S}}}^i)/\sigma_{\overline{\mathcal{S}}}^i$. A Gaussian filter is applied to the final affinity map $\wp_{\mathcal{S}}(x)$ to reduce noise.

3. Finally, to obtain the approximated β inside $\mathcal{U}$, we normalize the affinity map $\wp_{\mathcal{S}}(x)$ to the range between one and the β found in $\mathcal{S}$ (inclusive).

Estimating $\hat{\mathcal{F}}^*$ inside $\mathcal{U} \cup \mathcal{S}$ Note that $\hat{\mathcal{F}}^*$ inside $\mathcal{S}$ can be estimated by the β found inside $\mathcal{S}$ (i.e. $I(I_{m\mathcal{S}}/I_{m\overline{\mathcal{S}}})$). Therefore, in the following, we describe the estimation of $\hat{\mathcal{F}}^*$ inside $\mathcal{U}$. During the estimation, we maintain the image gradient of the shadowless component in $\mathcal{U}$ while adjusting the intensity. Given the β found in $\mathcal{U}$, we want:

$$\nabla\hat{\mathcal{F}} = \nabla\mathcal{F}_{est} \quad (6)$$

where $\mathcal{F}_{est}$ is the estimation of $\mathcal{F}$ given the β image inside $\mathcal{U}$ ($\mathcal{F} = \frac{I}{\beta}$). Our goal now is to solve the following equation:

$$\hat{\mathcal{F}}^* = \arg\min_{\hat{\mathcal{F}}} \int\int_{x\in\mathcal{U}} ||\nabla\hat{\mathcal{F}}(x) - \nabla\mathcal{F}_{est}(x)||^2 dx \quad (7)$$

with the Dirichlet boundary condition $\hat{\mathcal{F}}|_{\partial\mathcal{U}} = \hat{\mathcal{F}}'|_{\partial\mathcal{U}}$, where $\hat{\mathcal{F}}'|_{\partial\mathcal{U}}$ equals to the $\hat{\mathcal{F}}^*$ estimated in $\mathcal{S}$ if $x \in \mathcal{S}$, otherwise, $\hat{\mathcal{F}}'|_{\partial\mathcal{U}}$ equals to $I(x)$. Note that $\partial\mathcal{U}$ inside $\mathcal{E}$ does not need the Dirichlet boundary condition in order to exclude the shadow-casting object. Here, $\nabla\mathcal{F}_{est}$ is in fact analogous to the guidance field introduced in [11].

However, the result produced by Eqn. 7 has a problem. Note that the β image approximated in step (3) is not exact especially along hard shadow boundary. Therefore, the guidance field along the shadow boundary is not trustable. In order to reduce the effect of the guidance field along the shadow boundary, we associate a confidence measurement function with the guidance field:

$$C = \lambda_c[1 - \wp_{\mathcal{S}}(1 - \wp_{\mathcal{S}})] \quad (8)$$

where λ_c is the normalization factor. For hard shadow boundary, C is small and vice versa. Using the confidence function C, Eqn. 7 is now rewritten as:

$$\hat{\mathcal{F}}^* = \arg\min_{\hat{\mathcal{F}}} \int\int_{x\in\mathcal{U}} ||\nabla\hat{\mathcal{F}}(x) - C(x)\nabla\mathcal{F}_{est}(x)||^2 dx$$
$$(9)$$

and the associated Poisson equation with same boundary condition becomes

$$\triangle\hat{\mathcal{F}} = div(C\nabla\mathcal{F}_{est}) \quad (10)$$

To solve this Poisson equation, the Gauss-Seidel method with successive over-relaxation, used in [11], is employed.

Note that although the solution $\hat{\mathcal{F}}^*$ inside $\mathcal{U} \cup \mathcal{S}$ is quite good (as shown in Fig. 2), it is still a very rough approximation and cannot be output as the final shadowless image. Despite that, $\hat{\mathcal{F}}^*$ was experimentally shown to be adequate in guiding the Bayesian optimization represented by Eqn. 5. This equation is very important since the β prior is used together to achieve a significantly better result. Our prior considers shadow smoothness and implicitly texture gradient at the same time. In the following subsections, we define the terms in Eqn. 5.

3.4 Definition of $P(\hat{\mathcal{F}}^*|\beta)$

Here we define the likelihood $P(\hat{\mathcal{F}}^*|\beta)$. Since $\hat{\mathcal{F}}^*$ is a rough estimation of shadow-free image, the likelihood models the error of the shadow-free image (implicitly by β because they are interdependent) in the measurement of $\hat{\mathcal{F}}^*$. The error distribution is assumed to be Gaussian. Therefore, $P(\hat{\mathcal{F}}^*|\beta)$ is written as:

$$P(\hat{\mathcal{F}}^*|\beta) = \exp(-\frac{\sum_{x \in \mathcal{U} \cup \mathcal{S}}[I(x) - \beta(x)\hat{\mathcal{F}}^*(x)]^2}{2\sigma_c^2}) \quad (11)$$

where σ_c^2 is the standard deviation of the measurement error.

3.5 Definition of $P(\beta)$

To design the smooth β prior, we observe that for most shadow casting objects, the cast shadow is spatially smooth and contains low frequency components only, except along shadow boundary. So, a good estimation for β should be one that is smooth inside the umbra and shadow-free regions. With this prior knowledge, we define the smoothness constraint as follow:

$$\exp(-\frac{\sum_{\{x,y\} \in \mathcal{N}}[\beta(x) - \beta(y)]^2 k(x,y)}{2\sigma_1^2}) \quad (12)$$

where σ_1^2 is the uncertainty in the smoothness, $\mathcal{N}$ is the set of pixel pairs in $\mathcal{S} \cup \mathcal{U}$ and $k = \exp(-\frac{(\wp_{\mathcal{S}}(x) - \wp_{\mathcal{S}}(y))^2}{2\sigma_p^2})$, which is used to weaken the effect of the smoothness prior on high frequency component of the shadow.

In the presence of very hard shadow boundaries and relatively smooth textures, we can impose more smoothness on the texture along a hard shadow boundary. The following expression models this constraint:

$$\exp(-\frac{\sum_{\{x,y\} \in \mathcal{N}} \wp_{\mathcal{S}}(x)\wp_{\overline{\mathcal{S}}}(y)(\beta(y)I(x) - \beta(x)I(y))^2}{2\sigma_2^2}) \quad (13)$$

where σ_2^2 is the uncertainty in the texture smoothness and $\wp_{\overline{\mathcal{S}}}$ is the model affinity to $\overline{\mathcal{S}}$'s pdf, which can be found in a similar way for $\wp_{\mathcal{S}}$. The anisotropic smoothing effect imposed by Eqn. 13 operates by identifying large $\wp_{\mathcal{S}}(x)$ and large $\wp_{\overline{\mathcal{S}}}(y)$, which together indicates hard shadow boundary. The shadow border will be smoothed by minimizing $(\mathcal{F}(x) - \mathcal{F}(y))^2$, or equivalently $(\beta(y)I(x) - \beta(x)I(y))^2$.

Figure 3: Brick wall: the input image and result of shadow removal by our method.

Thus, the complete β prior, $P(\beta)$, is the multiplication of Eqn. 12 and Eqn. 13. Under the Bayesian optimization and our image model $\mathcal{F} = I/\beta$, $P(\beta)$ is maximized and therefore maintains shadow smoothness in β and texture gradients in $\mathcal{F}$ (if any) at the same time.

3.6 Optimization

In order to estimate β, for each pixel in each iteration, we solve Eqn. 5 by taking the first derivative w.r.t β and setting the whole equation to zero. Optimization is performed in $\mathcal{S} \cup \mathcal{U}$. To start the process, β is initialized into 1. The optimization stops when β enters a stationary state.

There is only one crucial parameter, σ_1, which explicitly controls the smoothness of the shadow. In most of the examples shown, $\frac{1}{2\sigma_1^2} = 1000$. σ_2 is set to a very large value (infinity) unless in the presence of hard boundary. σ_c and σ_p are always set to 1 and 0.0005, respectively.

4 Results

This section is organized into three subsections. We first apply our method in shadow removal. We then compare our shadow removal results with [6] which is an automatic method based on several assumptions about the light source and the camera. The test images were obtained from the authors' website. Our results show that we can remove soft and hard shadows, and preserve the textures under the extracted shadow. Next, we present applications on shadow extraction. We show improved result on segmentation by [4] using our output shadowless image. Some examples on shadow compositing by using the extracted shadows are shown. We also extend our method without any modification to repair images, and show an application on lens flare removal.

4.1 Shadow removal

We show the results produced by our method under different scenarios. Fig. 3 shows an example where structure and texture are present under the shadowed area. The shadow is smooth and soft. Fig. 4 shows another shadow removal result where textures and strong patterns are under the hard shadow. Both results are very reasonable considering the minimum user assistance we need.

Figure 4: Long Shadow: the input image and result of shadow removal by our method.

Figure 5: Flower: the input image and result of shadow removal by our method.

Fig. 5 shows an example where complex textures are present under the shadow, while the shadow is smooth. Note that the unique texture pattern of the ground is maintained and the shadowless image is visually acceptable.

Fig. 6 shows an example where the shape of the shadow is very complex and soft while the background contains some wood-like texture. Our method can remove all the shadow while preserving the texture under the shadow.

4.2 Comparison

We select some complex results produced in [6] for comparison. Note that [6] is an automatic method assuming Planckian light source and narrow-band camera, while ours is a user-assisted approach without any assumption on the light source and the camera. Besides the quadmap input, the rest of processing is left to the computer.

The first example in Fig. 7 contains three noticeable shadows from the man, the plant and the lamp (on the left hand side of the input). The output by [6] smooths out the textures while our method faithfully preserves them. Note, however, we cannot remove the lamp shadow well due to
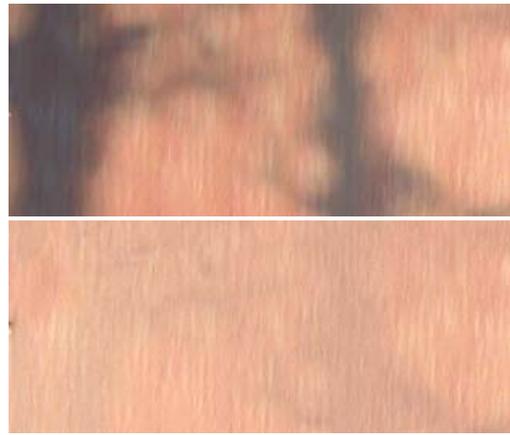

Figure 6: Wood: the input image and result of shadow removal by our method.

the dominant image noises, possibly caused by the inadequate resolution of the image we obtained.

The second example depicted in Fig. 7 contains complex textures under hard shadows. Our method produces visually better results along the shadow boundary while the textures are not smoothed out.

4.3 Applications

In this section, we first present an example in segmentation using the shadowless image output by our Bayesian method. Then, we propose potential applications. In image matting, shadow compositing is important because the composited image will look unrealistic without the shadow. In image restoration, we apply our technique as is to repair images corrupted by lens flare.

Image segmentation   We used the Meanshift color segmentation [4] to perform image segmentation on an image before and after shadow removal. Fig. 8 shows the color segments we obtained. In fact, by adjusting the bandwidth, the Meanshift method can eliminate the effect of shadow in some cases. However, in this complex example, the hard shadow cannot be merged with the wooden floor using a large bandwidth parameter, while the pot, the ground and the parapet are undesirably merged. The result produced with our shadowless image is significantly better because while a smaller bandwidth is used, the pot, the ground and the parapet are considered as individual segments.

Shadow compositing   Fig. 10 (top two rows) shows two shadows we extracted. Note that, in the case of the bird, the background under the shadow contain significant textures. The extracted shadows shown in (b) are smooth and contain no texture contribution from the grounds, so they can be used as shadow mattes for shadow compositing. In (c), the image composites look unrealistic without the shadows. In (d), we composite the extracted shadows and the
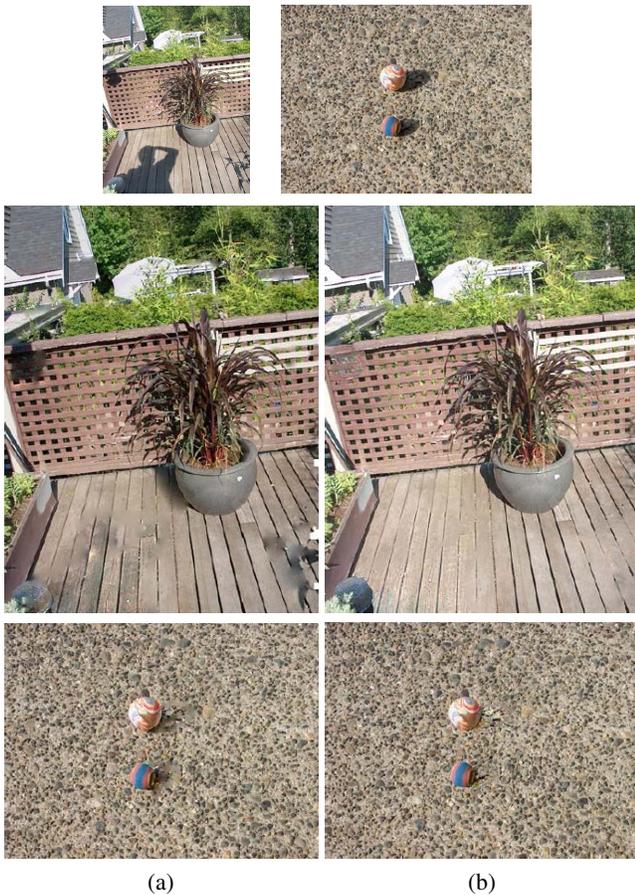
Figure 7: Shadow removal in complex scene. Top: input images from [6]. (a) result of shadow removal by [6], (b) result of shadow removal by our method.

foregrounds onto new background images to create visually plausible image composite. Fig. 10 (last row) composites the shadow extracted from Fig. 6 into a new background.

Image repair Here, an application in image repair is demonstrated: lens flare removal. In fact, our method can be applied to remove this artifact, which can also be modeled by the image model $I = \beta\mathcal{F}$, where $\mathcal{F}$ is the original image to be recovered, and β is the image transfer due to the artifact, analogous to the effect of a shadow. Fig. 9 shows the result of lens flare removal. The structure of the scene is preserved while the lens flare can be removed by our method. Note that our proposed method is useful because [8, 5] are not applicable here: As shown in this example, our method is capable of repairing the valley where there are no appropriate samples anywhere in the image. The absence of samples will hamper non-parametric techniques which strongly rely on patch sampling.

5 Discussion and Conclusion

The most important limitation of our single-image and image-based approach is that shading and shadows cannot
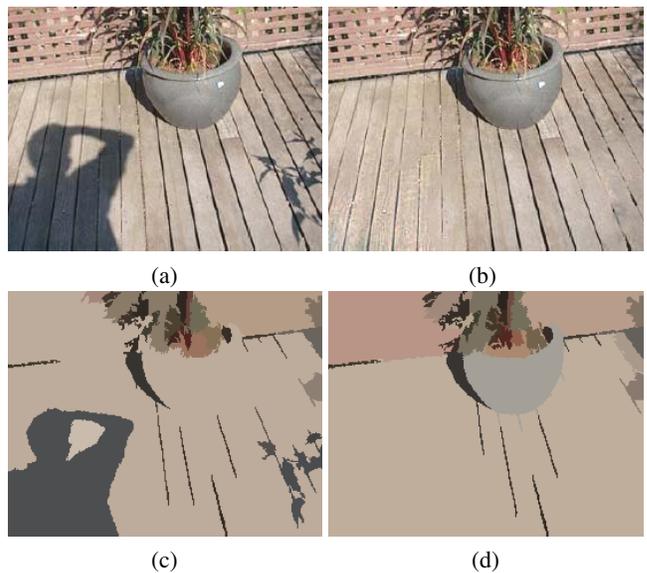

Figure 8: Segmentation. (a) Original image, (b) shadowless image after our shadow removal. (c) Meanshift segmentation result with input (a). The shadow remains as a region while the pot, the ground and the parapet are undesirably merged into one region when a large bandwidth is used. (d) Meanshift segmentation result with input (b). The pot, the ground and the parapet are separately considered as three homogeneous color regions.

Figure 9: Lens flare removal.

be distinguished. Let $\gamma = N \cdot L = \gamma_1\gamma_2$, where N and L are the normal and light directions, then Eqn. 2 can be expressed as $I = (\gamma_1\beta_1)(\gamma_2\mathcal{F}_2) = \beta\mathcal{F}$. Therefore, in practice, the effect of shading can be "distributed" to the estimated β and $\mathcal{F}$ so long as the objective functional is minimized. This inherent limitation of the image-based approach can be alleviated to a certain extent by prior user knowledge. For example, the effect of the smooth β prior may be increased if the shadow is known to be smooth, thus distributing more shading effect to the estimated $\mathcal{F}$.

From the image formation model and using the Bayesian formulation, we have defined and presented a novel technique for shadow extraction from a single image. Our approach is user-assisted and no simplifying assumption is needed on the camera or the light source, other than the Lambertian assumption. First, the user supplies a quadmap which can be marked roughly because only color statistics
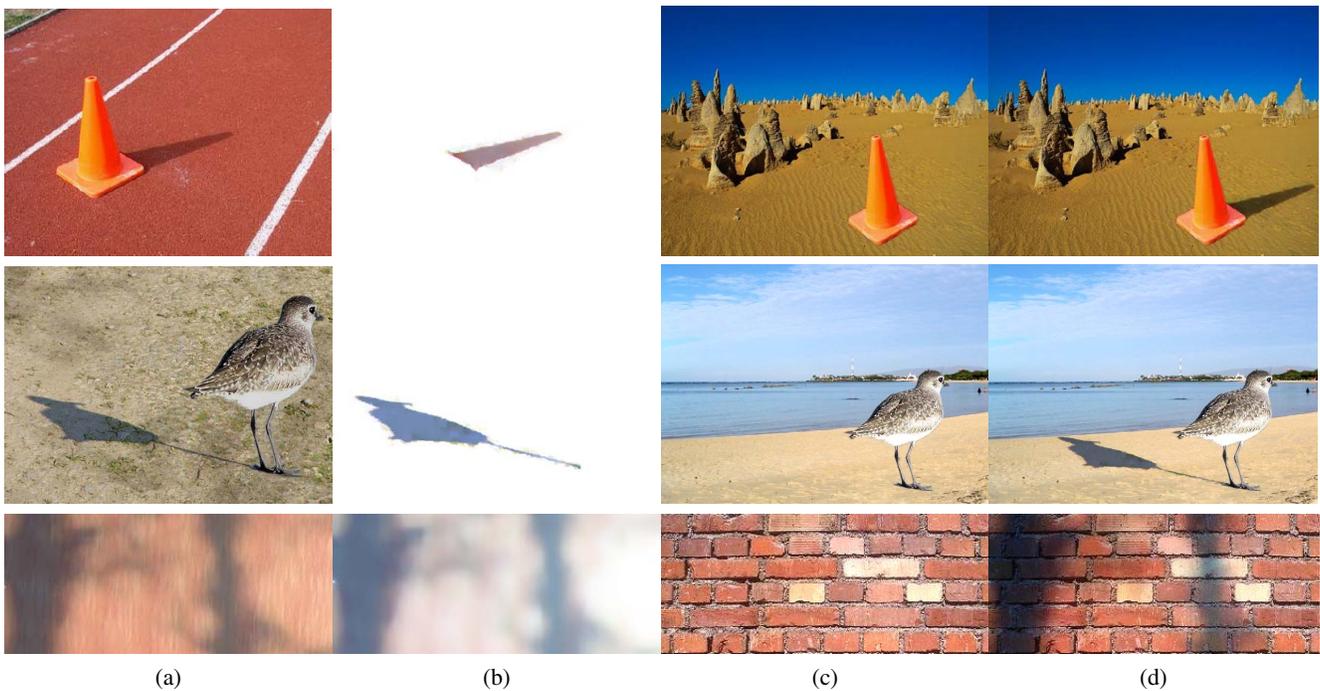
| | | | |
| (a) | (b) | (c) | (d) |

Figure 10: Application in image composition: (a) input image, (b) shadow image β extracted by our method, (c) image composite without shadow/background, (d) image composite with shadow.

but not spatial cues are needed in our system. The technical contribution of our technique lies in the way the user-supplied hints are encoded and employed in the Bayesian optimization which makes the rest of the processing fully automatic. Our method translates the very rough user-supplied hints into the effective likelihood and prior functions for the Bayesian optimization. To solve the optimization, a decent estimation of the shadowless image is obtained by solving the associated Poisson equation with the necessary Dirichlet boundary conditions. Using our image model, our Bayesian method simultaneously extracts smooth shadows while preserving texture appearance under the extracted shadow. Our technique can be applied to improve results on image segmentation, in shadow removal, in shadow extraction and compositing. New and potential applications such as lens flare removal is illustrated with encouraging result. Our future work will address the issue of shading and shadows, and investigate the extraction of shadows from videos.

References

[1] H.G. Barrow and J.M. Tenenbaum. Recovering intrinsic scene characteristics from images. In *CVS78*, pages 3–26, 1978.

[2] Y.-Y. Chuang, B. Curless, D.H. Salesin, and R. Szeliski. A bayesian approach to digital matting. *Proceedings of CVPR'01, Vol. II, 264-271*, 2001.

[3] Y.-Y. Chuang, D.B. Goldman, B. Curless, D.H. Salesin, and R. Szeliski. Shadow matting and compositing. *ACM Trans. Graph.*, 22(3):494–500, 2003.

[4] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-24(5):603–619, May 2002.

[5] I. Drori, D. Cohen-Or, and H. Yeshurun. Fragment-based image completion. *ACM Trans Graph.*, 22:303–312, 2003.

[6] G.D. Finlayson, M.S. Drew, and C. Lu. Intrinsic images by entropy minimization. In *ECCV04*, pages Vol III: 582–595, 2004.

[7] G.D. Finlayson, S.D. Hordley, and M.S. Drew. Removing shadows from images. *Proceedings of ECCV'02, Vol. IV, 823-836*, 2002.

[8] J. Jia and C.K. Tang. Inference of segmented color and texture description by tensor voting. *PAMI*, 26(6):771–786, June 2004.

[9] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *Proceedings of ECCV'04, Vol. I, 602-613*, 2004.

[10] Y. Matsushita, S. Lin, S.B. Kang, and H.Y. Shum. Estimating intrinsic images from image sequences with biased illumination. In *ECCV04*, pages Vol II: 274–286, 2004.

[11] P. Perez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Trans. Graph.*, 22(3):313–318, 2003.

[12] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE CG&A*, 21:34–41, 2001.

[13] J. Sun, J. Jia, C.-K. Tang, and H.-Y. Shum. Poisson matting. *ACM Trans. Graph.*, 23(3):351–321, 2004.

[14] M.F. Tappen, W.T. Freeman, and E.H. Adelson. Recovering intrinsic images from a single image. In *MIT AIM*, 2002.

[15] Y. Weiss. Deriving intrinsic images from image sequences. In *ICCV01*, pages II: 68–75, 2001.

[16] T. Welsh, M. Ashikhmin, and K. Mueller. Transferring color to greyscale images. *ACM Trans. Graph.*, 21:277–280, 2002.

[17] T.-P. Wu, C.-K. Tang, M.S. Brown, and H.-Y. Shum. Natual shadow matting. *ACM Trans. Graph.*, (Accepted).