Shadow Removal Using Bilateral Filtering

Qingxiong Yang, *Member, IEEE*, Kar-Han Tan, *Senior Member, IEEE*, Narendra Ahuja, *Fellow, IEEE*

Abstract—In this paper, we propose a simple but effective shadow removal method using a single input image. We first derive a 2-D intrinsic image from a single RGB camera image based solely on colors, particularly chromaticity. We next present a method to recover a 3-D intrinsic image based on bilateral filtering and the 2-D intrinsic image. The luminance contrast in regions with similar surface reflectance due to geometry and illumination variances is effectively reduced in the derived 3-D intrinsic image, while the contrast in regions with different surface reflectance is preserved. However, the intrinsic image contains incorrect luminance values. To obtain the correct luminance, we decompose the input RGB image and the intrinsic image. Each image is decomposed into a base layer and a detail layer. We obtain a shadow-free image by combining the base layer from the input RGB image and the detail layer from the intrinsic image such that the details of the intrinsic image are transferred to the input RGB image from which the correct luminance values can be obtained. Unlike previous methods, the presented technique is fully automatic and does not require shadow detection.

Index Terms—Bilateral filter, chromaticity, intrinsic image.

I. INTRODUCTION

SHADOW due to different illuminations often causes problems in computer vision and graphics task. For instance, darker regions (caused by shadows) can introduce incorrect segments in segmentation algorithms, radiance changes (caused by shadows) decrease the performance of object recognition systems, and shape from shading algorithms assume that surface radiance does not include shadow effects. In digital photography, shadows are typically considered as artifacts in the image that need to be removed, whether for enhancing certain parts of the image or simply for esthetic reasons. The problem of generating an intrinsic image that is invariant to illuminations is very important but challenging, and the recent methods can be divided into the following three categories.

1) Methods that use multiple images. Weiss [1] derived an intrinsic reflectance image of a scene given a sequence of images of the scene under a range of illumination conditions. Weiss's method implicitly assumes that the scene is composed of Lambertian surfaces. If a surface exhibits Lambertian reflectance, light falling on it is scattered such that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. Matsushita *et al.* [2] extended Weiss's method to handle non-Lambertian scenes as the Lambertian assumption does not hold for real world scenes. Matsushita *et al.* [3] later minimized a regularization function that takes advantage of the biased image derivatives to yield reflectance images less influenced by shading. These methods produce successful results. However, multiple images acquired under the required conditions are not always available in practice, e.g., for dynamic scenes and moving cameras.

2) Methods that use a single image but require human interaction. Levin and Weiss [4] proposed a user-assisted method to perform reflection separation and one of its applications is shadow removal. Wu *et al.* [5], [6] formulated the shadow formation problem as one of energy minimization guided by user-supplied hints in the form of a quad-map which can be specified easily by the user. Liu and Gleicher [7] constructed a shadow-effect free and texture-consistent gradient field between the shadow and lit area and recovered the shadow-free image from it by solving a Poisson equation. These methods produce good results with careful user-assistance, but are not of fully automatic.

3) Methods that use a single image without human interaction. Tappen *et al.* [8] proposed a learning-based approach to separate reflectance edges and illumination edges in a derivative image. Although this method successfully separates reflectance and shading for a given illumination direction used for training, it is not designed for arbitrary lighting directions. By assuming Lambertian surface, Planckian lighting[1] and narrow-band camera sensor, Finlayson and Hordley [9] derived a 1-D illuminant-invariant image. The transformation of RGB representation to a 1-D grayscale representation reduced the discrimination between every two surfaces in the image. Finlayson and Drew [10] extended this method to derive a 2-D intrinsic image using a four-band camera, and theoretically proved that 3-D invariant images can be recovered with six-band cameras. The problem with this approach is that four-band and six-band cameras are rarely used. Employing edge detection

[1]Planckian light is the radiation of a blackbody. Its color (chromaticity) depends on the temperature of the blackbody.

to the original and the 1-D illuminant-invariant image, Finlayson *et al.* [11] detected the shadow edges and re-integrated the image to obtain a 3-D shadow-free image by solving a Poisson's equation (a second-order partial differential equation). This method requires the pre-knowledge of the direction of illuminant variation (an angle ranging from 0 to 180 degree) which is obtained using a set of images of a Macbeth Color Checker[2] [12] taken at different times of the day. Assume that the shadow edge have been detected, Arbel and Hel-Or [13] find the scale factors which are then used to cancel the effect of shadows.

All above automatic methods that use a single input image require the shadow edges to be correctly detected except [8] which is invalid for arbitrary lighting directions. However, automatic edge detection is known to be non-robust (see [14] for a review), thus these methods perform poorly especially for images with multiple (unconnected) shadow regions.

In this paper, we present a new shadow removal method without shadow detection. We derive a 2-D intrinsic image from a single RGB camera image based solely on colors, particularly chromaticity. The intrinsic image is essentially 2-D because the maximum chromaticity of every pixel is the same. We next present a method to estimate a 3-D intrinsic image using bilateral filtering [15]–[19] which is used to enhance a target image using the texture of a guidance image. The guidance image has a different image modality than the target image. Joint bilateral filter has been demonstrate to be very effectively for combining the high frequencies from one image and the low frequencies from another, thus is very suitable for transferring the details from the estimated 2-D intrinsic image to the original image. Specifically, a coarse estimate of the maximum chromaticity values of the intrinsic image is computed based on the original image, and then filtered using joint bilateral filtering with the 2-D intrinsic image as the guidance image by assuming that the surface colors are locally similar. The filtered values are then used to compute a 3-D intrinsic image in which the maximum chromaticity values are different from pixel to pixel.

The luminance contrast in regions with similar surface reflectance due to geometry and illumination variances is effectively reduced in the derived 3-D intrinsic image while the contrast in regions with different surface reflectance is preserved. However, the intrinsic image contains incorrect luminance values. To obtain the correct luminance for better visualization, we convert the original RGB image and the 3-D intrinsic image to YUV format, and then decompose the Y channels into a base layer and a detail layer using cross/joint bilateral filtering [20]. We obtain a shadow-free image by combining the base layer from the original RGB image and the detail layer from the intrinsic image such that the details of the intrinsic image are transferred to the original RGB image from which the correct luminance values can be obtained.

To successfully remove the shadows, we make two main assumptions. The first assumption is that the input images have

chromatic surfaces. The proposed method requires at least two measurements (using two types of camera sensors) of the surface reflectance at each pixel to extract the intrinsic image that is independent of the illuminant. It is invalid for neutral surfaces because the response of a camera will be the same for all the sensors installed. The second assumption is that the input images are captured by narrow-band camera sensors [11] and under Planckian lighting. This assumption leads to a direct correlation between illumination temperature and the image color. The image can be then decomposed into two components using this correlation, and one of them is illuminant-invariant. The shadow removal problem is then equal to removing the other component which is related to the illumination temperature.

II. REMOVING ILLUMINANT DEPENDENCES

In this section, we give a detailed description of the derived 2-D and 3-D intrinsic images in which the illuminant dependences are removed. We start with a brief review of the adopted illumination and reflection model, then present the method for deriving the 2-D and 3-D intrinsic images, and finally consider the issue of camera calibration.

A. Illumination and Reflection Model

In this paper, we use Wien's approximation [21] to model the illumination spectrum

$$E(\lambda, T) = \frac{2hc^2}{\lambda^5} \exp\left(-\frac{hc}{\lambda k T}\right) \quad (1)$$

where $h = 6.62606896 \times 10^{-34}$ joule-second is Planck's constant, $c = 3 \times 10^8$ meter/second is the speed of light, $k = 1.3806504 \times 10^{-23}$ joule/kelvin is Boltzmann's constant, λ is the wavelength, and T is the color temperature.

For the surface reflection model, we adopt a standard diffuse + specular model (namely, that surface reflectance is a combination of a diffuse term and a specular term) commonly used in computer graphics. We assume that each image pixel can be treated independently using the method proposed below. Let $S(\lambda)$ denote surface reflectance spectrum, $Q(\lambda)$ the sensor response, and ρ_d and ρ_s geometric factors for diffuse and specular reflection, respectively. Then, the pixel intensity may be described as

$$I = I^{DIFF} + I^{SPEC}$$
$$= \rho_d \int_\lambda E(\lambda, T) S(\lambda) Q(\lambda) d\lambda + \rho_s \int_\lambda E(\lambda, T) Q(\lambda) d\lambda. \quad (2)$$

Since scene geometry may vary from point to point, the values of ρ_d and ρ_s can change from pixel to pixel.

For a RGB camera, let $Q_u(\lambda)$ denote the responses in the three color bands $u \in \{r, g, b\}$. If we assume that the camera sensitivities are Dirac delta functions, $Q_u(\lambda) = q_u \delta(\lambda - \lambda_u)$, then the sampled values become

$$I_u = \rho_d E(\lambda_u, T) S(\lambda_u) q_u + \rho_s E(\lambda_u, T) q_u. \quad (3)$$

Substituting (1) into (3), we obtain

$$I_u = (\rho_d S(\lambda_u) + \rho_s) q_u \frac{2hc^2}{\lambda_u^5} \exp\left(-\frac{hc}{\lambda_u k T}\right). \quad (4)$$

[2]A flat, physical object colored with an arrangement of standardized color samples, used for color comparisons and measurements such as checking the color reproduction of an imaging system.

B. 2-D Intrinsic Image

In this section, we compute a 2-D intrinsic image based solely on colors. Let $L_u = \log(I_u)$. Then, taking logarithm on both sides of Equation (4), we obtain

$$L_u = \log\left((\rho_d S(\lambda_u) + \rho_s) q_u \frac{2hc^2}{\lambda_u^5}\right) + \left(-\frac{hc}{\lambda_u kT}\right) \quad (5)$$

where T changes from pixel to pixel.

Let W_u be defined as

$$W_u = \log\left((\rho_d S(\lambda_u) + \rho_s) q_u \frac{2hc^2}{\lambda_u^5}\right) \quad (6)$$

we notice that W_u is invariant to T, and thus invariant to the illuminants. In this paper, we define W_u as the surface component, and $-\frac{hc}{\lambda_u kT}$ as the illuminant component. W_u can be computed from (5):

$$W_u = L_u + \frac{hc}{\lambda_u kT}. \quad (7)$$

Letting

$$W'_u = W_u \frac{\lambda_u k}{hc} \quad (8)$$

and

$$L'_u = L_u \frac{\lambda_u k}{hc} \quad (9)$$

from (7), we obtain

$$W'_u = L'_u + \frac{1}{T}. \quad (10)$$

Note that because k, h and c are constants and λ_u is the camera property which can be calibrated (Sec. II-D), L'_u can be computed from the observed images. Hence, the computation of the surface component W_u is reduced to solve for parameter T which changes from pixel to pixel.

According to the definition of chromaticity, we define

$$l_u = \frac{L'_u}{\sum_{i \in \{r,g,b\}} L'_i}$$

as the chromaticity of image L'_u, and

$$w_u = \frac{W'_u}{\sum_{i \in \{r,g,b\}} W'_i}$$

as the chromaticity of surface component W'_u. From (10), we obtain

$$w_u \left(\sum_{i \in \{r,g,b\}} W'_i\right) = l_u \left(\sum_{i \in \{r,g,b\}} L'_i\right) + \frac{1}{T}. \quad (11)$$

Summing up over every channel $u \in \{r,g,b\}$ in (11), and using $\sum_{u \in \{r,g,b\}} w_u = \sum_{u \in \{r,g,b\}} l_u = 1$,

$$\sum_{i \in \{r,g,b\}} W'_i = \left(\sum_{i \in \{r,g,b\}} L'_i\right) + \frac{3}{T}. \quad (12)$$

Substituting (12) into (11),

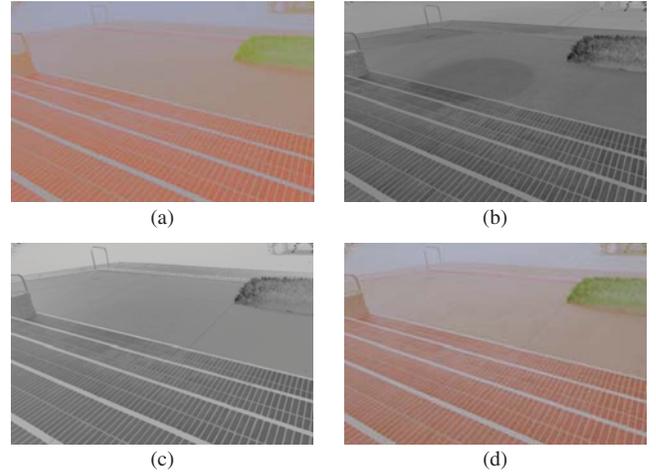$$T = \frac{1 - 3w_u}{w_u \left(\sum_{i \in \{r,g,b\}} L'_i\right) - L'_u} \quad (13)$$

Fig. 1. 3-D intrinsic image. (a) 2-D intrinsic P. (b) Coarse estimate w_m^0. (c) Refined estimate w_m^R. (d) 3-D intrinsic W.

where u can be any color channel. In this paper, we choose $u = m$, where $m = \text{argmax}_{u \in \{r,g,b\}} w_u$.

Substituting (13) into (10) and (7), we obtain

$$W'_u = L'_u + \frac{w_m \left(\sum_{i \in \{r,g,b\}} L'_i\right) - L'_m}{1 - 3w_m}, \quad (14)$$

$$W_u = \frac{hc}{\lambda_u k}\left(L'_u + \frac{w_m \left(\sum_{i \in \{r,g,b\}} L'_i\right) - L'_m}{1 - 3w_m}\right) \quad (15)$$

$$= L_u + \frac{w_m \left(\sum_{i \in \{r,g,b\}} \frac{\lambda_i}{\lambda_u} L_i\right) - \frac{\lambda_m}{\lambda_u} L_m}{1 - 3w_m}. \quad (16)$$

In general, w_m is different for every pixel. However, if we set w_m to be a constant κ ($\kappa \neq 1/3$), that is the maximum chromaticity of each pixel of the surface component image W is a constant, then an intrinsic image can be computed from Equation (16):

$$P_u = L_u + \frac{\kappa \left(\sum_{i \in \{r,g,b\}} \frac{\lambda_i}{\lambda_u} L_i\right) - \frac{\lambda_m}{\lambda_u} L_m}{1 - 3\kappa}. \quad (17)$$

Note that although the intrinsic image P is RGB image, it is essentially a 2-D image, since one of the chromaticity channel (maximum) is a constant.

1) Relationships Between P_u and W_u and a Special 2-D Intrinsic Image: The relationships between P and W are summarized in **Claim 1** and **Claim 2** (The proofs are provided in **Appendix** A and B).

Claim 1: For any two pixels $\mathbf{x}$ and $\mathbf{y}$, if $W_u(\mathbf{x}) = W_u(\mathbf{y})$, then $P_u(\mathbf{x}) = P_u(\mathbf{y})$ for $u \in \{r,g,b\}$.

Claim 2: For any two pixels $\mathbf{x}$ and $\mathbf{y}$, if $P_u(\mathbf{x}) = P_u(\mathbf{y})$, then $W_u(\mathbf{x}) = W_u(\mathbf{y})$ for $u \in \{r,g,b\}$, or the color difference of $W(\mathbf{x})$ and $W(\mathbf{y})$ is neutral

$$\frac{W_u(\mathbf{x}) - W_u(\mathbf{y})}{\sum_{i \in \{r,g,b\}} (W_i(\mathbf{x}) - W_i(\mathbf{y}))} = \frac{1}{3}.$$

Claim 1 shows that all the information included in P is also included in W, and **Claim 2** shows that some information included in W is actually lost in P. As a result, P is just an approximation of W, and we can prove that P is actually a 2-D image while W is obviously a 3-D image.
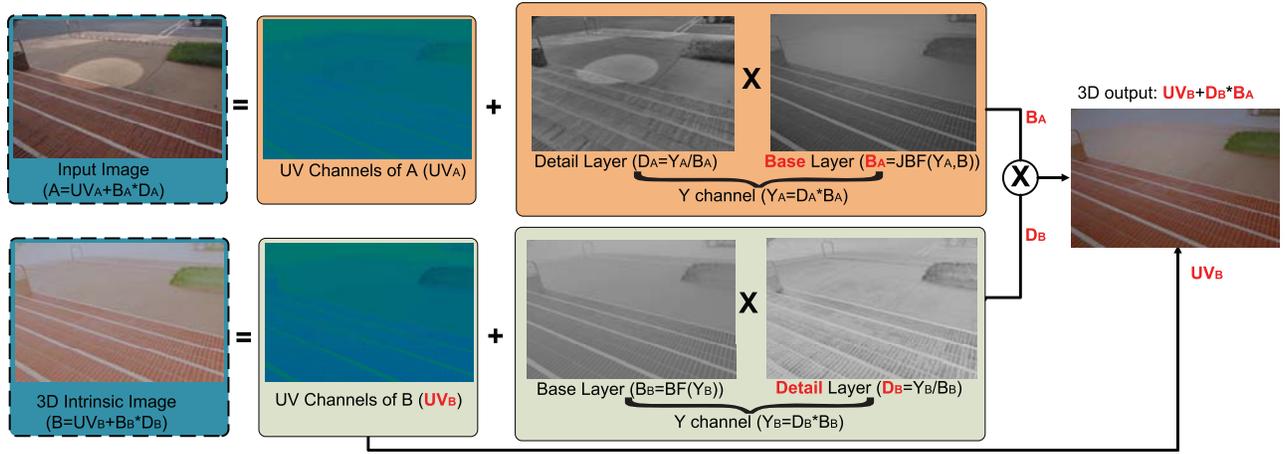
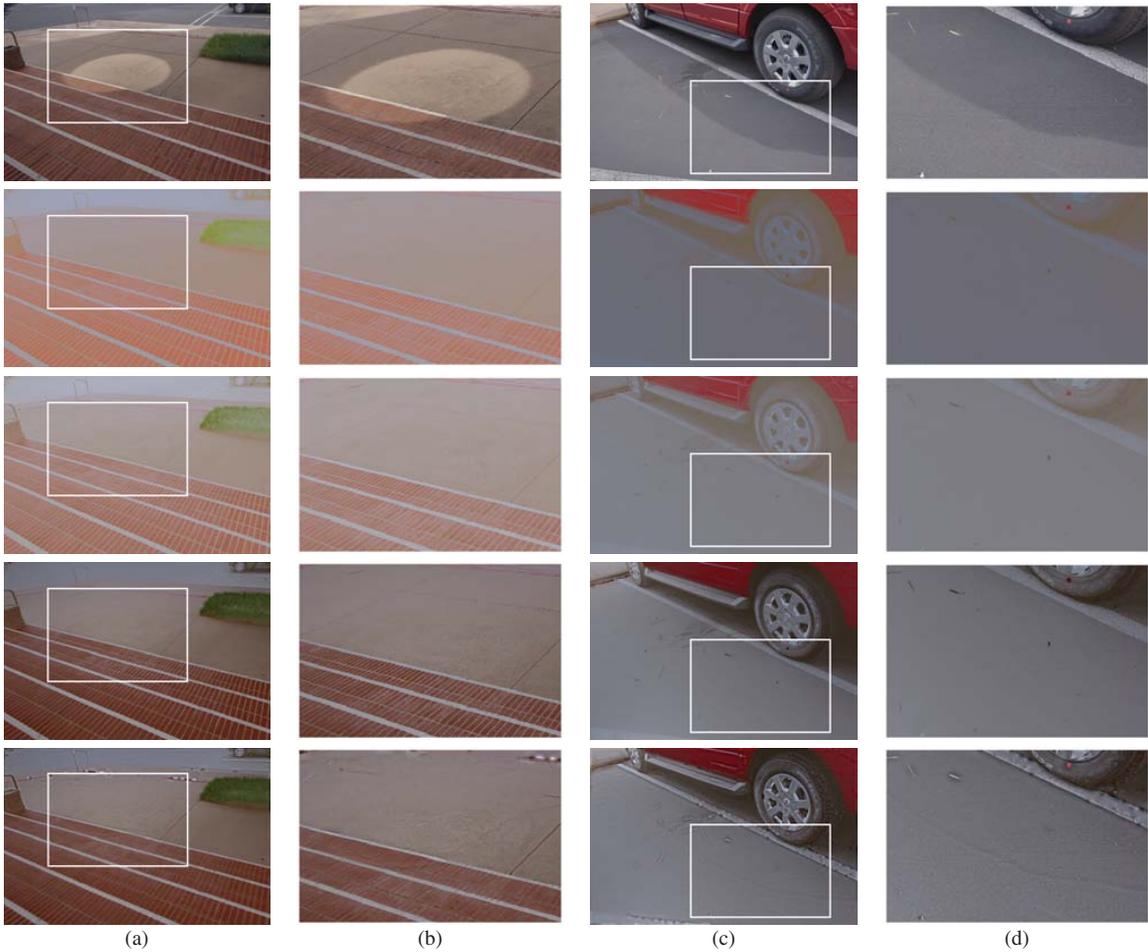Fig. 2.   Automatic luminance abstraction pipeline.

Fig. 3.   (a) and (c) Experimental result using images captured by a Canon 350-D camera. From top to bottom: original image, 2-D intrinsic image, 3-D intrinsic image, the final shadow-free image, and shadow-free image from [11]. (b) and (d) Close-up of the white rectangles in (a) and (c).

Let $\kappa = 0$, Equation (17) becomes

$$P_u = L_u - \frac{\lambda_m}{\lambda_u} L_m. \tag{18}$$

Thus $P_m = L_m - \frac{\lambda_m}{\lambda_m} L_m = 0$, which means that only two channels of the intrinsic image P contain useful information, and that the derived intrinsic image is indeed a 2-D image.

Substitute Equation (9) and (10) into Equation (18), we obtain

$$P_u = \frac{hc}{\lambda_u k} \left(L'_u - L'_m \right) \tag{19}$$

$$= \frac{hc}{\lambda_u k} \left(W'_u - W'_m \right). \tag{20}$$

Equation (20) shows that when $\kappa = 0$, the obtained intrinsic image is the weighted difference between the
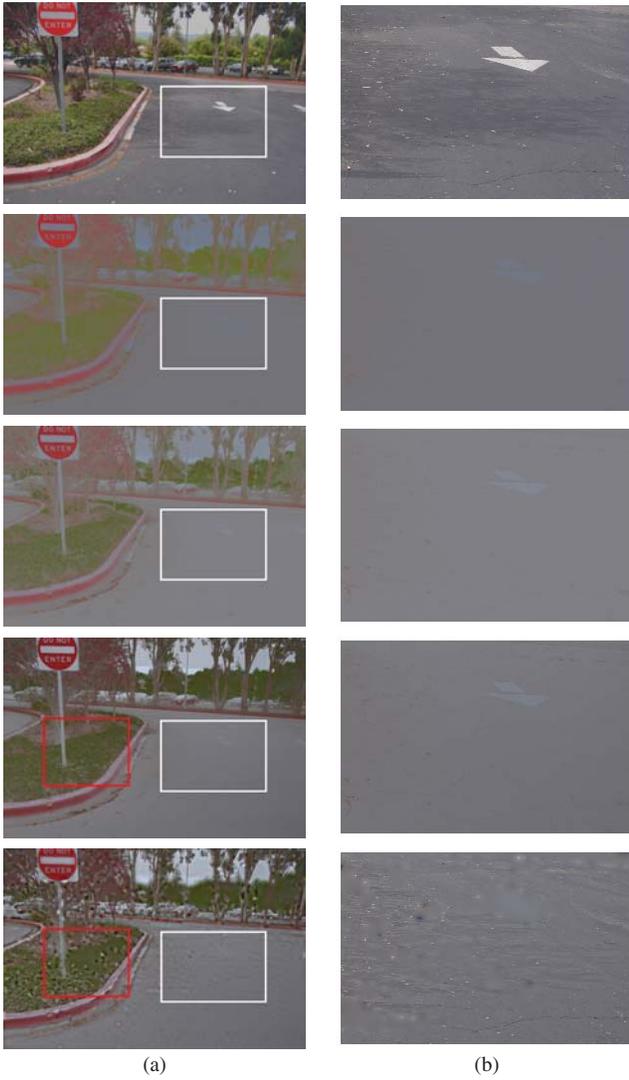
Fig. 4. (a) Failure case on neutral regions. From top to bottom: original image, 2-D intrinsic image, 3-D intrinsic image, the final shadow-free image, and shadow-free image from [11]. (b) Close-up of the white rectangles in (a). Close-up of the red rectangles in (a) is presented in Fig. 5.

Fig. 5. Close-ups of the red rectangles in Fig. 4(a). (a) Computed using our method and (b) computed using the method presented in [11]. As can be seen, false shadow edges deteriorate the performance of the method presented in [11].

normalized surface component W'_u and its maximum value $W'_m = \max(W'_r, W'_g, W'_b)$, and is thus invariant to T. Besides, Equation (19) gives a very good illustration of the intuition of the 2-D intrinsic image. According to Equation (10), each channel of L' contains a component $\frac{1}{T}$ which depends on T. However, this component is the same for every channel.

Apparently, subtracting a channel from L' will result in a 2-D image that is independent of T, and is a special case ($\kappa = 0$) of the representation in Equation (17). However, the representation in Equation (17) advances over this simple solution as it can be used to estimate a 3-D intrinsic image by assuming that w_m is locally similar. Detailed algorithm is presented in Sec. II-C.

C. 3-D Intrinsic Image

The problem of the derived 2-D intrinsic image is that its maximum chromaticity is a constant, which reduces the contrast between two surfaces with different surface reflectance. In this section, we present a joint bilateral filtering [16], [17] based approach for deriving a 3-D intrinsic image in which the maximum chromaticity is different from pixel to pixel.

In this paper, we assume that for every two pixel $\mathbf{x}$ and $\mathbf{y}$, if $P_u(\mathbf{x}) = P_u(\mathbf{y})$, then the surface maximum chromaticity $w_m(\mathbf{x}) = w_m(\mathbf{y})$. We also assume that the surface colors are locally similar, particularly, w_m is locally similar. In this case, given a coarse estimate of the surface maximum chromaticity w_m^0 where the variance within a surface with similar reflectance may be large (e.g., Figure 1(b)), we can use the texture of the 2-D intrinsic image (Figure 1(a)) to smoothen w_m^0, which is actually a joint bilateral filtering process. The filtered result is defined as the refined estimate of the surface maximum chromaticity

$$w_m^R = JBF(w_m^0, P), \tag{21}$$

$$w_m^R(\mathbf{x}) = \frac{\sum_{\mathbf{y} \in N(\mathbf{x})} (f_S(\mathbf{x}, \mathbf{y}) \cdot f_R(P(\mathbf{x}), P(\mathbf{y})) \cdot w_m^0(\mathbf{y}))}{\sum_{\mathbf{y} \in N(\mathbf{x})} (f_S(\mathbf{x}, \mathbf{y}) \cdot f_R(P(\mathbf{x}), P(\mathbf{y})))} \tag{22}$$

where $N(\mathbf{x})$ is the set of neighboring pixels around $\mathbf{x}$, and f_S and f_R are the spatial and range filter kernels. f_S and f_R are typically Gaussian in the literature [15], [20]. In this paper, we assume $w_m^0 = \operatorname{argmax}_{u \in \{r,g,b\}} l_u$, so that the surface maximum chromaticity is equal to the maximum chromaticity of the logarithm of the original image.

Substitute w_m with w_m^R, the surface component W can then be computed using Equation (16). The computed surface component W is a 3-D intrinsic image. The 3-D intrinsic image computed from w_m^R in Figure 1(c) is presented in Figure 1(d).

D. Camera Calibration

In this section, we present our camera calibration method. According to Equations (16) and (17), the computation of the intrinsic image requires the knowledge of relative wavelengths

$$\lambda'_u = \frac{\lambda_u}{\sum_{i \in \{r,g,b\}} \lambda_i}$$

of the light passing through the camera sensors. Here we investigate how it is possible to calibrate a camera given only a single image.

The calibration of λ'_u can be carried out by capturing an image of a planar scene with uniform color. Assume $\mathbf{x}$ is one of the shadow pixels, and $\mathbf{y}$ is one of the non-shadow pixels
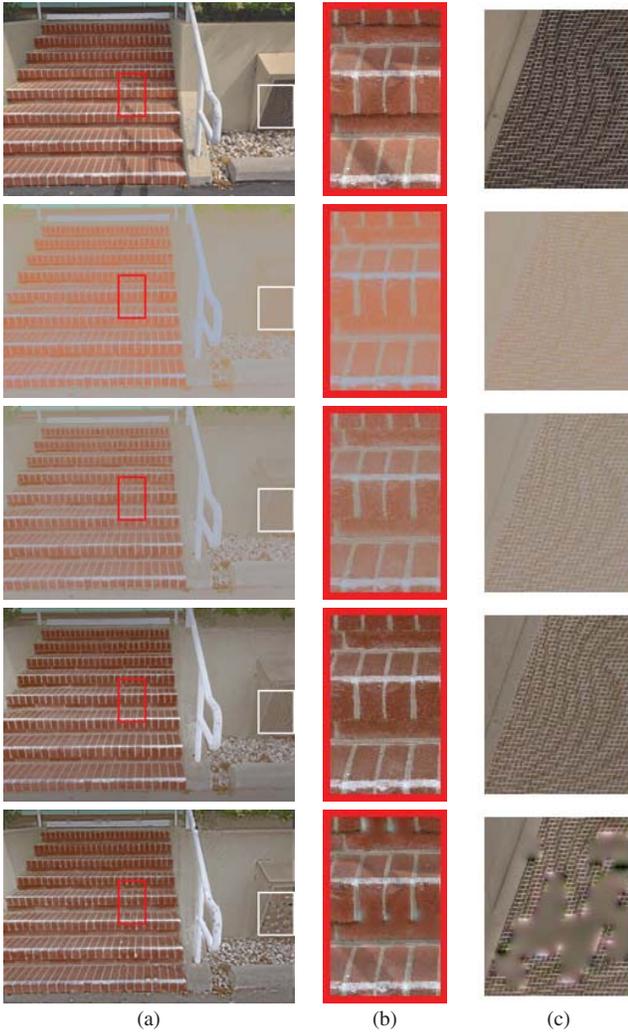
Fig. 6. (a) More experimental results using images captured by the Canon 350-D camera. From top to bottom: original image, 2-D intrinsic image, 3-D intrinsic image, the final shadow-free image, and shadow-free image from [11]. (b) and (c) are the close-ups of the red and white rectangles in (a). As can be seen, our method outperforms [11] when false shadow edges are used.

identified manually, according to Equation (5), we obtain

$$\lambda'_u = \frac{1/\left(L_u\left(\mathbf{x}\right) - L_u\left(\mathbf{y}\right)\right)}{\sum_{i\in\{r,g,b\}}\left(1/\left(L_i\left(\mathbf{x}\right) - L_i\left(\mathbf{y}\right)\right)\right)} \qquad (23)$$

$$= \frac{1/\left(\frac{hc}{\lambda_u kT(\mathbf{y})} - \frac{hc}{\lambda_u kT(\mathbf{x})}\right)}{\sum_{i\in\{r,g,b\}}\left(1/\left(\frac{hc}{\lambda_i kT(\mathbf{y})} - \frac{hc}{\lambda_i kT(\mathbf{x})}\right)\right)}. \qquad (24)$$

To reduce noise, we compute a λ'_u value using Equation (24) for every shadow and non-shadow pixel pair, and give a vote for the estimated λ'_u. Finally, the one with the largest number of votes is selected as correct.

III. RESTORING LUMINANCE

The luminance contrast in regions with similar surface reflectance due to geometry and illumination variances is effectively reduced in the derived 3-D intrinsic image (Sec. II-C) while the contrast in regions with different surface reflectance is preserved. However, the estimated 3-D intrinsic

image may contain incorrect luminance values. In this section, we present a method which abstracts the luminance of the original image and transfers it to the intrinsic image such that a shadow-free image with correct luminance can be obtained.

An overview of our luminance abstraction pipeline is presented in Figure 2. We first convert the original RGB image (A) to YUV format, and then separated the Y Component (Y_A) into two layers named base layer (B_A) and detail layer (D_A) [20] using bilateral filter [16], [17]. The YUV model defines a color space in terms of one luminance (Y, perceptual brightness) component and two chrominance (UV) components. The detail layer $D_A = Y_A/B_A$ is simply the ratio of Y Component Y_A to the base layer B_A which encodes the large-scale variations. Bilateral filter is used to prevent smoothing across surfaces with different reflectance, and the intrinsic image (B) with illuminant dependences removed is used as the filtering guidance. The filter size is the same as the image resolution, and the fast bilateral filtering method in [22] is adopted, such that the computational complexity is invariant to the filter size. The intrinsic image (B) is also converted to YUV format, and the Y Component (Y_B) is also separated into two layers using bilateral filtering. The product of its detail layer (D_B) and the based layer of the original image (B_A) is used as the Y Component of the final image

$$Y = D_B * B_A. \qquad (25)$$

Combining with the UV Components of the intrinsic image (UV_B) and converting to the RGB format, a shadow-free image with correct luminance is obtained.

IV. EXPERIMENTS

In this section, we describe the experiments we performed on real images. Results for two scenes captured by a Canon 350D camera are provided in Figure 3. As can be seen, the shadow edges in Figure 3 are relatively simple and can be detected correctly, thus both our method and the method presented in [11] work reliably on these scenes. Note that the contrast around the tire in the second and third rows of Figure 3(d) is low due to the incorrect luminance values in the intrinsic images. The fourth row of Figure 6(d) restores the luminance and thus preserves the contrast.

However, neither our method nor the method presented in [11] works for neutral regions as shown in Figure 4. The contrast between the white arrow and the ground in the original image is not preserved in the obtained shadow-free images. This is because the white arrows is neutral and the ground is also closed to neutral, thus according to **Claim 2**, the contract will be greatly reduced. However, the contrast between the white paint and the bricks/ground in Figure 3(b) is well-preserved. This is because the white paint is neutral, but the bricks and ground are not. This result shows that although our image abstraction method is based on colors, it is also valid for neutral regions, as long as the neighboring regions are not neutral.

Figure 5 visually compares the shadow-free images obtained using our method and [11]. [11] recovers incorrect texture in the shadow-free image when false shadow edges are detected

and used. Our method does not require shadow detection, the performance thus will not be affected. Another example is presented in Figure 6. As can be seen, our method visually outperforms [11].

V. CONCLUSION

We have introduced an automatic shadow removal method in this paper. Two illuminant-invariant image representations were first derived in this paper: 2-D intrinsic image computed based solely on colors, and 3-D intrinsic image estimated based on the 2-D intrinsic image and joint bilateral filtering. The details of the 3-D intrinsic image were then transferred to the original image to reduce the luminance contrast in regions with similar surface reflectance due to geometry and illumination variances. Unlike most previous methods, the presented technique does not require shadow detection.

APPENDIX A

PROOF OF CLAIM 1

For any two pixels $\mathbf{x}$ and $\mathbf{y}$, if $W_u(\mathbf{x}) = W_u(\mathbf{y})$, then $P_u(\mathbf{x}) = P_u(\mathbf{y})$.

Proof: Assume $W_u(\mathbf{x}) = W_u(\mathbf{y})$, then

$$W'_u(\mathbf{x}) = W'_u(\mathbf{y}) \qquad (26)$$

$$w_u(\mathbf{x}) = w_u(\mathbf{y}). \qquad (27)$$

Assume $L'_s = \sum_{i \in \{r,g,b\}} L'_i$, from Equation (14), (26) and (27), we obtain

$$L'_u(\mathbf{x}) + \frac{w_m(\mathbf{x})L'_s(\mathbf{x}) - L'_m(\mathbf{x})}{1 - 3w_m(\mathbf{x})} =$$
$$L'_u(\mathbf{y}) + \frac{w_m(\mathbf{x})L'_s(\mathbf{y}) - L'_m(\mathbf{y})}{1 - 3w_m(\mathbf{x})} \Rightarrow$$
$$L'_u(\mathbf{x}) - L'_m(\mathbf{x}) - L'_u(\mathbf{y}) + L'_m(\mathbf{y}) =$$
$$w_m(\mathbf{x})(L'_s(\mathbf{y}) - 3L'_u(\mathbf{y}) - L'_s(\mathbf{x}) + 3L'_u(\mathbf{x})). \qquad (28)$$

From (10) and (26), we obtain

$$L'_u(\mathbf{x}) - L'_u(\mathbf{y}) + L'_m(\mathbf{y}) - L'_m(\mathbf{x}) = 0. \qquad (29)$$

For chromatic pixels, $w_m(\mathbf{x}) \neq 0$, thus substituting (29) into (28), we have

$$L'_s(\mathbf{y}) - 3L'_u(\mathbf{y}) - L'_s(\mathbf{x}) + 3L'_u(\mathbf{x}) = 0. \qquad (30)$$

From (29) and (30),

$$(L'_u(\mathbf{x}) - L'_m(\mathbf{x})) - (L'_u(\mathbf{y}) - L'_m(\mathbf{y})) =$$
$$\kappa(L'_s(\mathbf{y}) - 3L'_u(\mathbf{y}) - L'_s(\mathbf{x}) + 3L'_u(\mathbf{x})) \Rightarrow$$
$$L'_u(\mathbf{x}) + \frac{\kappa L'_s(\mathbf{x}) - L'_m(\mathbf{x})}{1 - 3\kappa} =$$
$$L'_u(\mathbf{y}) + \frac{\kappa L'_s(\mathbf{y}) - L'_m(\mathbf{y})}{1 - 3\kappa} \Rightarrow$$
$$L_u(\mathbf{x}) + \frac{hc}{\lambda_u k}\frac{\kappa L'_s(\mathbf{x}) - L'_m(\mathbf{x})}{1 - 3\kappa} =$$
$$L_u(\mathbf{y}) + \frac{hc}{\lambda_u k}\frac{\kappa L'_s(\mathbf{y}) - L'_m(\mathbf{y})}{1 - 3\kappa} \Rightarrow$$
$$P_u(\mathbf{x}) = P_u(\mathbf{y}) \qquad (31)$$

APPENDIX B

PROOF OF CLAIM 2

For any two pixels $\mathbf{x}$ and $\mathbf{y}$, if $P_u(\mathbf{x}) = P_u(\mathbf{y})$, then $W_u(\mathbf{x}) = W_u(\mathbf{y})$, or the color difference of $W_u(\mathbf{x})$ and $W_u(\mathbf{y})$ is neutral

$$\frac{W_u(\mathbf{x}) - W_u(\mathbf{y})}{\sum_{i \in \{r,g,b\}}(W_i(\mathbf{x}) - W_i(\mathbf{y}))} = \frac{1}{3}.$$

Proof: Assume $P_u(\mathbf{x}) = P_u(\mathbf{y})$, from Equation (17), we have

$$(1 - 3\kappa)L'_u(\mathbf{x}) + \kappa L'_s(\mathbf{x}) - L'_m(\mathbf{x}) =$$
$$(1 - 3\kappa)L'_u(\mathbf{y}) + \kappa L'_s(\mathbf{y}) - L'_m(\mathbf{y}) \Rightarrow$$
$$L'_u(\mathbf{x}) - L'_m(\mathbf{x}) - L'_u(\mathbf{y}) + L'_m(\mathbf{y}) =$$
$$\kappa(L'_s(\mathbf{y}) - L'_s(\mathbf{x})) + 3\kappa(L'_u(\mathbf{x}) - L'_u(\mathbf{y})). \qquad (32)$$

Substituting (10) into (32), and assume $W'_s = \sum_{i \in \{r,g,b\}} W'_i$, we have

$$(W'_u(\mathbf{x}) - W'_u(\mathbf{y})) - (W'_m(\mathbf{x}) - W'_m(\mathbf{y})) =$$
$$\kappa(W'_s(\mathbf{y}) - W'_s(\mathbf{x})) + 3\kappa(W'_u(\mathbf{x}) - W'_u(\mathbf{y})). \qquad (33)$$

Summing up every channel $u \in \{r, g, b\}$ in (33),

$$(W'_s(\mathbf{x}) - W'_s(\mathbf{y})) - 3(W'_m(\mathbf{x}) - W'_m(\mathbf{y})) = 0 \Rightarrow$$
$$(W_s(\mathbf{x}) - W_s(\mathbf{y})) - 3(W_m(\mathbf{x}) - W_m(\mathbf{y})) = 0 \qquad (34)$$

and thus $W_m(\mathbf{x}) - W_m(\mathbf{y}) = 0$ or $W_m(\mathbf{x}) - W_m(\mathbf{y})/W_s(\mathbf{x}) - W_s(\mathbf{y}) = 1/3$. Note that m can be any channel.

REFERENCES

[1] Y. Weiss, "Deriving intrinsic images from image sequences," in *Proc. Int. Conf. Comput. Vis.*, 2001, pp. 68–75.

[2] Y. Matsushita, K. Nishino, K. Ikeuchi, and M. Sakauchi, "Illumination normalization with time-dependent intrinsic images for video surveillance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 10, pp. 1336–1347, Oct. 2004.

[3] Y. Matsushita, S. Lin, S. B. Kang, and H.-Y. Shum, "Estimating intrinsic images from image sequences with biased illumination," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 274–286.

[4] A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1647–1654, Sep. 2007.

[5] T.-P. Wu and C.-K. Tang, "A Bayesian approach for shadow extraction from a single image," in *Proc. Int. Conf. Comput. Vis.*, 2005, pp. 480–487.

[6] T.-P. Wu, C.-K. Tang, M. S. Brown, and H.-Y. Shum, "Natural shadow matting," *ACM Trans. Graph.*, vol. 26, no. 2, pp. 357–366, 2007.

[7] F. Liu and M. Gleicher, "Texture-consistent shadow removal," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 437–450.

[8] M. F. Tappen, W. T. Freeman, and H. Adelson, "Recovering intrinsic images from a single image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1459–1472, Sep. 2005.

[9] G. D. Finlayson and S. Hordley, "Color constancy at a pixel," *JOSA*, vol. 18, no. 2, pp. 253–264, 2001.

[10] G. D. Finlayson and M. S. Drew, "4-sensor camera calibration for image representation invariant to shading, shadows, lighting, and specularities," in *Proc. Int. Conf. Comput. Vis.*, 2001, pp. 473–480.

[11] G. D. Finlayson, S. D. Hordley, C. Lu, and M. S. Drew, "On the removal of shadows from images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 59–68, Jan. 2006.

[12] C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," *J. Appl. Photograph. Eng.*, vol. 2, no. 3, pp. 95–99, 1976.

[13] E. Arbel and H. Hel-Or, "Texture-preserving shadow removal in color images containing curved surfaces," in *Proc. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–8.

[14] R. Jain, R. Kasturi, and B. Schunck, *Machine Vision*. New York: McGraw-Hill, 1995.

[15] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. Int. Conf. Comput. Vis.*, 1998, pp. 839–846.

[16] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 664–672, 2004.

[17] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," in *Proc. SIGGRAPH*, 2004, pp. 673–678.

[18] Q. Yang, S. Wang, and N. Ahuja, "SVM for edge-preserving filtering," in *Proc. Comput. Vis. Pattern Recognit.*, 2010, pp. 1775–1782.

[19] Q. Yang, "Recursive bilateral filtering," in *Proc. Eur. Conf. Comput. Vis.*, 2012, to be published.

[20] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Proc. SIGGRAPH*, 2002, pp. 257–266.

[21] G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd ed. New York: Wiley, 1982.

[22] Q. Yang, K.-H. Tan, and N. Ahuja, "Real-time O(1) bilateral filtering," in *Proc. Comput. Vis. Pattern Recognit.*, 2009, pp. 557–564.

Qingxiong Yang (M'11) received the Ph.D. degree in electrical and computer engineering from the University of Illinois, Urbana-Champaign, Urbana, in 2010.

He is currently an Assistant Professor with the Computer Science Department, City University of Hong Kong, Hong Kong. His current research interests include computer vision and graphics.

Dr. Yang was a recipient of the Best Student Paper Award at MMSP in 2010 and the Best Demo at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition in 2007.

Kar-Han Tan (SM'11) received the B.Sc. degree from the National University of Singapore, Singapore, the M.S. degree from the University of California, Los Angeles, and the Ph.D. degree in computer science from the University of Illinois at Urbana-Champaign, Urbana.

He was a Beckman Graduate Fellow with the University of Illinois at Urbana-Champaign. He is currently a Senior Research Scientist with the Mobile and Immersive Experience Laboratory, Hewlett-Packard Laboratories, Palo Alto, CA, where he is researching 3-D capture and display technologies as well as next-generation remote collaboration systems. He contributes actively to the research community. Prior to joining Hewlett-Packard, he was the Manager of the Algorithms Group, EPSON Research and Development, where he led the invention of view projection, a technique that enables one-touch setup of light displays on arbitrary surfaces. He co-invented multiflash imaging at Mitsubishi Electric Research Laboratory, Cambridge, MA, and the Virtual Structures Algorithm at the University of California, Los Angeles, which is widely recognized today as one of the fundamental techniques for mobile robot formation control.

Dr. Kan was a recipient of several Best Paper Awards for his research.

Narendra Ahuja (F'92) received the B.E. degree (Hons.) in electronics engineering from the Birla Institute of Technology and Science, Pilani, India, in 1972, the M.E. degree (with distinction) in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1974, and the Ph.D. degree in computer science from the University of Maryland, College Park, in 1979.

He was a Scientific Officer with the Department of Electronics, Government of India, New Delhi, India, from 1974 to 1975. He was with the Computer Vision Laboratory, University of Maryland, from 1975 to 1979. Since 1979, he has been with the University of Illinois at Urbana-Champaign, Urbana, where he is currently a Donald Biggar Willet Professor with the Department of Electrical and Computer Engineering, the Beckman Institute, and the Coordinated Science Laboratory. He has co-authored the books *Pattern Models* (Wiley, 1983), *Motion and Structure from Image Sequences* (Springer-Verlag, 1992), and *Face and Gesture Recognition* (Kluwer, 2001) and has co-edited the book *Advances in Image Understanding* (IEEE Press, 1996). His current research interests include extraction and representation of spatial structures in images and videos, integrated uses of multiple image-based sources for scene representation and recognition, versatile sensors for computer vision, and applications including visual communication, image manipulation, and information retrieval.

Dr. Ahuja was a recipient of the Emanuel R. Piore Award from the IEEE in 1999, the Technology Achievement Award from the International Society for Optical Engineering in 1998, the TA Stewart-Dyer/Frederick Harvey Trevithick Prize from the Institution of Mechanical Engineers in 2008, and the Open Innovation Research Award from Hewlett-Packard in 2008, the Distinguished Alumnus Award from the Department of Computer Science, University of Maryland, in 2008, the Best Paper Award from the IEEE TRANSACTIONS ON MULTIMEDIA in 2006, the University Scholar Award in 1985, the Presidential Young Investigator Award in 1984, the National Scholarship from 1967 to 1972, and the President's Merit Award in 1966. He was selected by the Center for Advanced Study, University of Illinois, as an Associate in 1998–1999 and 2006–2007, and a Beckman Associate in 1990–1991. He is a fellow of the American Association for Artificial Intelligence, the International Association for Pattern Recognition, the Association for Computing Machinery, the American Association for the Advancement of Science, and the International Society for Optical Engineering. He is on the editorial boards of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, *Computer Vision, Graphics, and Image Processing*, the *Journal of Mathematical Imaging and Vision*, the *Journal of Pattern Analysis and Applications*; the *International Journal of Imaging Systems and Technology*, the *Journal of Information Science and Technology*, and the *IEE Japan Transactions on Electrical and Electronic Engineering*. He was a Guest Co-Editor of the *Artificial Intelligence Journal's* special issue on vision. He was the Founding Director of the International Institute of Information Technology, Hyderabad, where he continues to serve as the International Director.