Efficient and Robust Specular Highlight Removal

Qingxiong Yang, *Member, IEEE*, Jinhui Tang, *Senior Member, IEEE*, and Narendra Ahuja, *Fellow, IEEE*

Abstract—A robust and effective specular highlight removal method is proposed in this paper. It is based on a key observation—the maximum fraction of the diffuse colour component in diffuse local patches in colour images changes smoothly. The specular pixels can thus be treated as noise in this case. This property allows the specular highlights to be removed in an image denoising fashion: an edge-preserving low-pass filter (e.g., the bilateral filter) can be used to smooth the maximum fraction of the colour components of the original image to remove the noise contributed by the specular pixels. Recent developments in fast bilateral filtering techniques enable the proposed method to run over $200\times$ faster than state-of-the-art techniques on a standard CPU and differentiates it from previous work.

Index Terms—Specular reflection separation, highlight, bilateral filter

1 Introduction

The spectral energy distribution of the light reflected from an object is the product of the spectral energy distribution of the illumination and the surface reflectance. Using the dichromatic reflection model [19], the reflected light can be separated into two components, due to specular and diffuse reflections, respectively. Specular reflection presents difficulties for many computer vision tasks, such as segmentation, detection and matching, since it captures light source characteristics, creating a discontinuity in the omnipresent, object-determined diffuse part. For simplification, specularities are usually disregarded as outliers by methods that are based on diffuse component analysis. Since the presence of specular reflection is inevitable in the real world, and they do capture important scene information, e.g., surface shape [2], [4] and light source characteristics, incorporation of specularity regions in the analysis is important.

1.1 Related Work

Previous methods for separating reflection components can be separated into two categories by the number of images used. The first category uses multiple images taken under specific conditions (e.g.,viewpoint, lighting direction, etc.) or using a single image. Nayar et al. [12] used multiple images captured from different polarization angles. Sato and Ikeuchi [18] employed the dichromatic model for separation by analyzing colour signatures in many images captured with a moving light source. Lin and Shum [9] also changed the light source direction to produce two photometric images and used linear basis functions to separate the specular components. Park and Tou [16] again requires different illumination directions. These approaches are of restricted use in a general setting since the light source is usually fixed in the real world. A feasible solution is to change the viewpoint instead of changing the illumination direction. Using multiple images taken from different viewing directions, Lee [7] presented a method for specular region detection and Lin [8] removed the highlights by treating the specular pixels as outliers, and matching the remaining diffuse parts in other views. However, this method may fail if the size of the highlight region is large, because then the large number of pixels involved cannot be considered as outliers. These methods are moderately practical, since it may not always be possible to meet the required conditions in practice.

Highlight removal using a single image, as in the other category, is generally much more challenging. When dealing with multi-coloured images, most single-image-based methods require colour segmentation (e.g., [6], [1]) which is known to be non-robust for complex textured images or requires user assistance for highlight detection [20]. By relaxing general surfaces to chromatic surfaces, the highlight removal problem can be greatly simplified. With this additional assumption, Tan and Ikeuchi [24] demonstrated that highlights from textured objects with complex multi-coloured scenes can be effectively removed without explicit colour segmentation. This method removes highlights by iteratively shifting chromaticity values towards those of the neighbouring pixel having the maximum chromaticity in the neighbourhood. The neighbourhood is determined using a "pseudo-coded" diffuse image which has exactly the same geometrical profile as the diffuse component of the input image and can be generated by shifting each pixel's intensity and maximum chromaticity nonlinearly. Similar to the "pseudo-coded" diffuse image presented in [24], Mallick et al. [11] proposed an SUV colour space which separated the specular and diffuse components into an S channel and UV channels. This SUV space was further used for highlight removal by iteratively eroding the specular channel using either a single image or video sequences [10]. This type of approach may encounter problems due to discontinuities in surface colours, across which diffuse information cannot be accurately propagated. Kim et al. [5] propose a more stable solution based on the observation that for most natural images the dark channel of an image provides a pseudo specular reflection result. However, its computational complexity is relatively high due to the use of TV-L1 and TV-L2 optimization.

Other approaches for single-image highlight removal analyze the distributions of image colours within a colour space. Tan and Ikeuchi [23] related the specular pixels to diffuse pixels for every surface colour by projecting image colours along the illumination colour direction to a point of lowest observed intensity. As a result, the decomposition can be expressed in a close form, and can be solved directly for every pixel. Their experimental results contained noise caused by a number of factors, including image noise, colour blending at edges and multiple surface colours. By integrating the texture from outside the highlight to determine the candidate diffuse colour for traditional colour-space techniques, that is for each pixel, a set of candidate diffuse colours is obtained from a texture scale of 1×1, and is iteratively pruned as the texture scale increases, Tan et al. [21] showed appreciable improvements in diffuse-highlight separation. This method requires that there are enough repetitive textures locally and the highlight does not have a similar colour to the surface.

1.2 Our Contribution

All the above methods that use a single input image share the same problem that they are not capable of real-time applications, e.g., stereo matching for specular surfaces, and generally result in noticeable artifacts. In this paper, we propose a simple but effective specular highlight reduction method using a single input image. Our method is closely related to [24], in which, the diffuse colour of a specular pixel is derived as a nonlinear function of its colour (from input image) and the maximum fraction of the diffuse colour components which is denoted as maximum diffuse chromaticity in

- *Q. Yang is with the Department of Computer Science, City University of Hong Kong, Hong Kong Special Administrative Region, China. E-mail: qiyang@cityu.edu.hk.*
- *J. Tang is with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, China. E-mail: jinhuitang@njust.edu.cn.*
- *N. Ahuja is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL. E-mail: n-ahuja@illinois.edu.*

Manuscript received 13 May 2014; revised 30 Aug. 2014; accepted 18 Sept. 2014. Date of publication 1 Oct. 2014; date of current version 8 May 2015.
Recommended for acceptance by M.S. Brown.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2014.2360402

the paper. The final step is estimating the maximum diffuse chromaticity value for every pixel which is a non-trivial problem, and a method which iteratively shifts chromaticity values towards those of the neighbouring pixel having the maximum chromaticity in the neighbourhood was proposed. Our method is the same as [24] except for the way of estimating the maximum diffuse chromaticity. Based on a key observation—the maximum diffuse chromaticity in local patches in colourful images generally changes smoothly—we estimate the maximum diffuse chromaticity values of the specular pixels by directly applying a low-pass filter to the maximum fraction of the colour components of the original image, such that the maximum diffuse chromaticity values can be propagated from the diffuse pixels to the specular pixels. Our method can be directly extended for multi-colour surfaces if edge-preserving filters (e.g., the bilateral filter [15], [27] and the local Laplacian filter [14]) are used such that the smoothing can be guided by the maximum diffuse chromaticity. In practice, maximum diffuse chromaticity is unknown and is to be estimated. We thus present an approximation and demonstrate its effectiveness.

The main computational complexity of the proposed method resides in the employed bilateral filter, and the recent development in fast bilateral filtering techniques [3], [13], [17], [28] enables the proposed method to run 200× faster than [24][1] on average. Another advantage of our method is that image pixels are processed independently, allowing for parallel implementation. Our GPU implementation shows that our highlight removal method can process 1 MB images at video rate on an NVIDIA Geforce 8800 GTX GPU.

Besides having the speed advantage, our method does not have the non-converged artifacts due to discontinuities in surface colours as presented in [24]. The use of a low-pass filter guarantees that the estimated maximum diffuse chromaticity will be locally smooth, as are the estimated diffuse reflections. Nevertheless, since the theory of our method is heavily built upon [24], *it shares most of the limitations of [24], e.g., the input images have chromatic surfaces, the output of the camera is linear to the flux of the incident light and the illumination chromaticity can be correctly measured/estimated*.

2 Algorithm

In this section, we first briefly review the adopted reflection model (in Section 2.1) and then present a real-time highlight removal method using bilateral filter (in Section 2.2). Finally, the sensitivity to the bilateral filter parameters and the illumination chromaticity estimation is discussed (in Sections 2.3 and 2.4).

2.1 Reflection Model

Using standard diffuse + specular reflection models commonly used in computer graphics, the reflected light colour ($\vec{J}$) captured by an RGB camera can be represented as a linear combination of diffuse ($\vec{J^D}$) and specular ($\vec{J^S}$) colours:

$$\vec{J} = \vec{J^D} + \vec{J^S}. \tag{1}$$

Let chromaticity be defined as the fraction of colour component c:

$$\sigma_c = \frac{J_c}{\sum_{c\in\{r,g,b\}} J_c}, \tag{2}$$

where $c \in \{r, g, b\}$, we define diffuse chromaticity Λ_c and illumination chromaticity Γ_c as follows:

$$\Lambda_c = \frac{J_c^D}{\sum_{c\in\{r,g,b\}} J_c^D}, \tag{3}$$

$$\Gamma_c = \frac{J_c^S}{\sum_{c\in\{r,g,b\}} J_c^S}. \tag{4}$$

1. The source code is available on its author's homepage [22].

Following the chromaticity definition in Eqs. (2), (3), and (4), we express the reflected light colour J_c as:

$$J_c = \Lambda_c \sum_{u\in\{r,g,b\}} J_u^D + \Gamma_c \sum_{u\in\{r,g,b\}} J_u^S. \tag{5}$$

Assuming that the illumination chromaticity can be measured (with a white reference) or estimated [26], using which the input image can be normalized such that $\Gamma_r = \Gamma_g = \Gamma_b = 1/3$ and $J_r^S = J_g^S = J_b^S = J^S$. Then the diffuse component can be written as:

$$J_c^D = J_c - J^S, \tag{6}$$

according to Eq. (5).

Following the chromaticity definition in Eqs. (2) and (3), we define maximum chromaticity as:

$$\sigma_{\max} = \max(\sigma_r, \sigma_g, \sigma_b), \tag{7}$$

and maximum diffuse chromaticity be:

$$\Lambda_{\max} = \max(\Lambda_r, \Lambda_g, \Lambda_b). \tag{8}$$

Tan [24] shows that the diffuse component can be represented as a function of $\Lambda_{\max}$

$$J_c^D(\Lambda_{\max}) = J_c - \frac{\max_{u\in\{r,g,b\}} J_u - \Lambda_{\max} \sum_{u\in\{r,g,b\}} J_u}{1 - 3\Lambda_{\max}}. \tag{9}$$

Since surface materials may vary from point to point, $\Lambda_{\max}$ changes from pixel to pixel in real images but is limited from $\frac{1}{3}$ to 1.

Estimating the maximum diffuse chromaticity $\Lambda_{\max}$ for every pixel from a single image is a non-trivial problem. However, if it is set to a constant, then a "pseudo-coded" diffuse image which has exactly the same geometrical profile as the diffuse component of the input image can be obtained. In this case, the saturation values of all pixels are made constant and this "pseudo-coded" diffuse image is essentially a two-channel image, while the ground-truth diffuse image is a three-channel image. The "pseudo-coded" diffuse image is just an approximation of ground truth, which will fail to preserve the feature discriminability for surfaces having the same hue but different saturation. However, it is the best estimate we can get, and has been demonstrated to be effective for solving the highlight removal problem in [24].

2.2 Highlight Removal Using Bilateral Filter

According to Eq. (9), the highlight removal problem can be reduced to the searching for the maximum diffuse chromaticity $\Lambda_{\max}$ which changes from pixel to pixel. However, $\Lambda_{\max}$ is related to the surface material and is locally similar. The maximum chromaticity $\sigma_{\max}$ is the same as the maximum diffuse chromaticity $\Lambda_{\max}$ except for specular pixels, which cause the intensity/colour discontinuities within local patches of the same surface colour. Intuitively, applying low-pass filtering to the maximum chromaticity $\sigma_{\max}$ will smooth out the variances due to specular highlights. However, there are two issues:

1) The smoothing filter should be edge-aware, such that the $\sigma_{\max}$ values of two pixels associated with different surface materials ($\Lambda_{\max}$ values are different) will not be blended together.
2) The diffuse pixels will be affected by the specular pixels after smoothing.

As a popular edge-aware operator, a joint bilateral filter can be employed to smooth the maximum chromaticity $\sigma_{\max}$ using the maximum diffuse chromaticity $\Lambda_{\max}$ as the smoothing guidance. But $\Lambda_{\max}$ is to be estimated; thus we need to find a substitution or an approximation. Although the "pseudo-coded" diffuse image

presented in [24] is free of specularity, it cannot be *directly* used for this problem because its colour depends on both the surface geometry and material, while Λ_c is invariant to the surface geometry. Let

$$\sigma_{min} = \min(\sigma_r, \sigma_g, \sigma_b), \tag{10}$$

we approximate Λ_c using λ_c computed as follows:

$$\lambda_c = \frac{\sigma_c - \sigma_{min}}{1 - 3\sigma_{min}}. \tag{11}$$

The relationship between the approximate diffuse chromaticity λ_c and the real diffuse chromaticity Λ_c is captured in Claim 1 and Claim 2.[2]

Claim 1. For any two pixels $\vec{p}$ and $\vec{q}$, if $\Lambda_c(\vec{p}) = \Lambda_c(\vec{q})$, then $\lambda_c(\vec{p}) = \lambda_c(\vec{q})$.

Claim 2. For any two pixels $\vec{p}$ and $\vec{q}$, if $\lambda_c(\vec{p}) = \lambda_c(\vec{q})$ and $\Lambda_{min}(\vec{p}) = \Lambda_{min}(\vec{q})$, then $\Lambda_c(\vec{p}) = \Lambda_c(\vec{q})$.

Note that λ_c is just an approximation of Λ_c, which will fail for the specific case specified in Claim 2. However, it is the best estimate we can get. It is actually strongly related to the chromaticity of the "pseudo-coded" diffuse image presented in [24]. According to [24], this chromaticity that be computed as follows:

$$\beta_c = \frac{\sigma_c - \sigma_{max}}{1 - 3\sigma_{max}} + \alpha \frac{1 - \sigma_c}{1 - 3\sigma_{max}}, \tag{12}$$

where α is a constant value used to approximate Λ_{max} and σ_{max} is the maximum component of the image chromaticity. Let σ_{min} denote is the corresponding minimum component, Eq. (12) can be rewritten as

$$\beta_c = \frac{\sigma_c - \sigma_{min}}{1 - 3\sigma_{min}} + \alpha \frac{1 - \sigma_c}{1 - 3\sigma_{min}} = \lambda_c + \alpha \frac{1 - \sigma_c}{1 - 3\sigma_{min}}, \tag{13}$$

where α is now a constant value used to approximate Λ_{min} which is the minimum component of the diffuse chromaticity. As can be seen, λ_c is indeed a special case of β_c (when $\alpha = 0$). As a result, λ_c is indeed the chromaticity of the "pseudo-coded" diffuse image presented in [24].

In practice, only the maximum component of the diffuse chromaticity Λ_{max} is required to compute the diffuse reflection (as shown in Eq. (9)). This means that the estimated maximum diffuse chromaticity values of two pixels should be enforced to be the same, as long as the *maximum* component of the approximate diffuse chromaticity values (λ of the two pixels) are the same (even if the other components of the chromaticity are different or the chromaticity is completely different in each channel). As a result, the maximum component of the approximate diffuse chromaticity λ is computed and used as the smoothing guidance:

$$\begin{aligned} \lambda_{\max} &= \max(\lambda_r, \lambda_g, \lambda_b) \\ &= \max\left(\frac{\sigma_r - \sigma_{min}}{1 - 3\sigma_{min}}, \frac{\sigma_g - \sigma_{min}}{1 - 3\sigma_{min}}, \frac{\sigma_b - \sigma_{min}}{1 - 3\sigma_{min}}\right). \end{aligned} \tag{14}$$

Using the approximate maximum diffuse chromaticity defined in Eq. (14) to guide the smoothing, the filtered maximum chromaticity $\sigma_{\max}$ can be computed as follows:

$$\sigma^F_{\max}(\vec{p}) = \frac{\sum_{\vec{q}\in\Omega} \mathcal{F}(\vec{p}, \vec{q}) \mathcal{G}(\lambda_{\max}(\vec{p}), \lambda_{\max}(\vec{q})) \sigma_{\max}(\vec{q})}{\sum_{\vec{q}\in\Omega} \mathcal{F}(\vec{p}, \vec{q}) \mathcal{G}(\lambda_{\max}(\vec{p}), \lambda_{\max}(\vec{q}))}, \tag{15}$$

where

$$\mathcal{F}(\vec{p}, \vec{q}) = exp\left(-\frac{||\vec{p} - \vec{q}||^2}{2\sigma_S^2}\right) \tag{16}$$

and

$$\mathcal{G}(\lambda_{\max}(\vec{p}), \lambda_{\max}(\vec{q})) = exp\left(-\frac{|\lambda_{\max}(\vec{p}), \lambda_{\max}(\vec{q})|^2}{2\sigma_R^2}\right) \tag{17}$$

are spatial and range weighting functions which are typically Gaussian in the literature [27], [3], and σ_S and σ_R are the geometric and photometric spread used to adjust the spatial similarity and the range similarity, respectively.

The variances of the maximum chromaticity $\sigma_{\max}$ due to specular highlights will be reduced after filtering, and the filtered maximum chromaticity $\sigma^F_{\max}$ will be closer to $\Lambda_{\max}$ than $\sigma_{\max}$ for the specular pixels. However, after smoothing, the diffuse pixels will be affected by the specular pixels too. According to [25], the filtered maximum chromaticity values $\sigma^F_{\max}$ of the diffuse pixels will be lower than the un-filtered values $\sigma_{\max}$. As a result, to exclude the contribution of the specular pixels, we compare $\sigma^F_{\max}$ and $\sigma_{\max}$ and take the maximum value:

$$\sigma_{\max}(\vec{p}) = \max\big(\sigma_{\max}, \sigma^F_{\max}(\vec{p})\big). \tag{18}$$

We then iteratively apply a joint bilateral filter to $\sigma_{\max}$ such that the maximum diffuse chromaticity values can be gradually propagated from the diffuse pixels to the specular pixels. In practice, we compare the filtered values $\sigma^F_{\max}$ with $\sigma_{\max}$ after every iteration. The algorithm is believed to converge when their difference is smaller than a threshold (set at 0.03 in our experiments) at every pixel. Our method generally converges after two-three iterations. The proposed highlight removal algorithm is summarized in Algorithm 1.

Algorithm 1. Highlight Removal Using a Single RGB Image

1: Compute $\sigma_{\max}$ according to Eq. (7) at every pixel using the input image and store it as a grayscale image.
2: Compute $\lambda_{\max}$ according to Eq. (14) at every pixel using the input image and store it as a grayscale image.
3: **repeat**
4: -Apply joint bilateral filter to image $\sigma_{\max}$ using $\lambda_{\max}$ as the guidance image (Eq. 18), store the filtered image as $\sigma^F_{\max}$;
5: -For each pixel $\vec{p}$, $\sigma_{\max}(\vec{p}) = \max(\sigma_{\max}(\vec{p}), \sigma^F_{\max}(\vec{p}))$;
6: **until** $\sigma^F_{\max} - \sigma_{\max} < 0.03$ at every pixel.

2.3 Sensitivity to the Bilateral Filter Parameters

As shown in Algorithm 1, the parameters involved in the proposed highlight removal method are mainly the parameters related to the joint bilateral filter. According to Eqs. (15), (16), and (17), there will be two parameters σ_S and σ_R[3] which are used to adjust the spatial similarity and the range similarity, respectively.

To evaluate the performance of the proposed method with respect to parameter σ_S and σ_R, we conducted experiments on a synthetic image with ground-truth diffuse reflection available as presented in Fig. 1. Fig. 1a1 is the diffuse image and (a2) is the combination of (a1) and the specular highlight and is used as the input image for the proposed method; (b1) and (b2) are the maximum chromaticity $\sigma_{\max}$ values computed from (a1) and (a2), respectively; the upper rows of (c1) and (c2) are the close up of (b1) and (b2), and the bottom the colour-mapped values. Red means a larger and blue a smaller value. As can be seen, the specular highlights decrease the maximum chromaticity $\sigma_{\max}$ values. (d1) is the

2. The proof is presented in our ECCV paper [29].

3. σ_S controls the size of filter kernel and σ_R adjusts the range (intensity/color) similarity, see [27] for details. A normalized image coordinate is used so that σ_S resides in [0, 1]. Image intensity is also normalized and thus σ_R also ranges from 0 to 1.

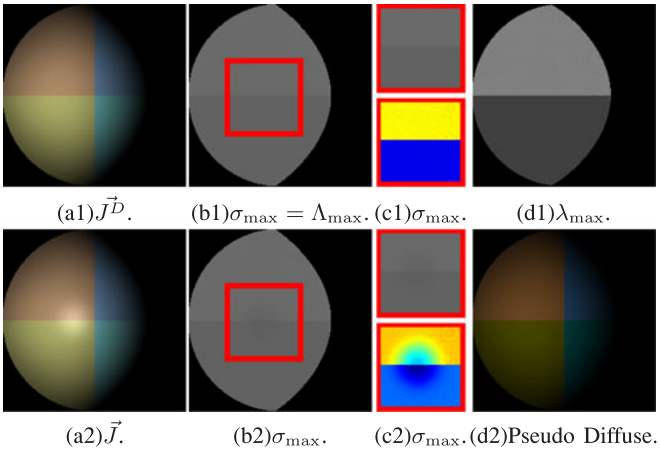


Fig. 1. Highlight removal on a synthetic image.

approximate maximum diffuse chromaticity $\lambda_{\max}$ and (d2) is the pseudo-code diffuse image proposed in [24], respectively. The basic idea of the proposed method is using (d1) as the guidance image to iteratively smooth (b2) to approximate (b1).

As can be seen in Figs. 1c1 and 1c2, the specular highlights always decrease the value of the maximum chromaticity $\sigma_{\max}$. The basic idea is then to use the approximate maximum diffuse chromaticity $\lambda_{\max}$ in Fig. 1d as the guidance image to propagate the maximum chromaticity $\sigma_{\max}$ in Fig. 1b2 from diffuse pixels to specular pixels. This can be achieved via an image smoother by treating the $\sigma_{\max}$ values computed from the specular pixels as image noise. The smoother should be edge-aware; thus the joint bilateral filter is a good choice. Initially, $\sigma_{\max}$ is computed from the input image to approximate the ground-truth maximum diffuse chromaticity in Fig. 1b1.

The performance of the joint bilateral filter *w.r.t* to its two parameters σ_S and σ_R is presented in Figs. 2a1, 2a2. We use the peak signal-to-noise ratio (PSNR) to evaluate numerical accuracy. For two intensity images I, $J \in [0,1]$, this ratio is defined as $10\log_{10}((h \cdot w)/\sum_{\mathbf{x}} |I(\mathbf{x}) - J(\mathbf{x})|^2)$, where h and w are the height and width of image I and J, and $\mathbf{x}$ is one of the pixels. It is assumed [13] the PSNR values above 40 dB often correspond to almost invisible differences. Fig. 2a1 shows that the performance of the proposed method is robust to σ_S as long as the joint bilateral filter is large enough to cover the specular highlight regions and (b) shows that σ_R should be relatively small to enforce edge preservation. The highest performance is achieved when $\sigma_R = 0.01$ and $\sigma_S = 0.65$. Part (c) presents the performance *w.r.t* σ_S when $\sigma_R = \{0.01, 0.02, 0.03\}$, and (a2) presents the performance *w.r.t* σ_R when $\sigma_S = \{0.60, 0.65, 0.70\}$. *Quantitative evaluation in Figs. 2a1, 2a2 concludes that the proposed method is robust to the two parameters as long as the size of filter kernel (represented by σ_S) is large enough to cover the specular highlight regions and σ_R is small enough (smaller than 0.06) to reject noise from other surfaces.*

The synthetic image presented in Fig. 1 contains only two maximum approximate diffuse chromaticity $\lambda_{\max}$ values (excluding the dark area) as can be seen in Fig. 1d; thus there is no noise for this data set, and the performance of the joint bilateral filter increases as the photometric spread σ_R decreases. This is not true for real images with sensor noise. A diffuse image was captured and then combined with synthetic specular reflection to produce a specular image for numerical comparison as shown in Fig. 3. Note that the captured diffuse image in Fig. 3a is highly textured and the specular region in the synthetic image in Fig. 3b is relatively large (which occupies *a quarter of the image*); both are difficult with regard to the highlight removal problem. Fig. 3c is a closeup of (b), and specular region covers almost the whole image. Fig. 3d is the diffuse extracted from (c) using the proposed method. As can be seen, the proposed method is valid even when the specular highlight is relatively large.

We use PSNR for numerical evaluation. The performance of the proposed method with a joint bilateral filter is presented in Figs. 2b1 and 2b2. As can be seen from Figs. 2b1 and 2b2, the performance is robust to σ_S and σ_R when $\sigma_S > 0.2$ and $\sigma_R < 0.08$. The grey planes/lines in (b1)-(b2) correspond to the PSNR values computed from the captured diffuse image in Fig. 3a and the synthetic input image in Fig. 3b. Performance lower than this threshold is meaningless. The highest performance is achieved when $\sigma_R = 0.02$ and $\sigma_S = 0.25$. Unlike the synthetic data set in Fig. 1, the

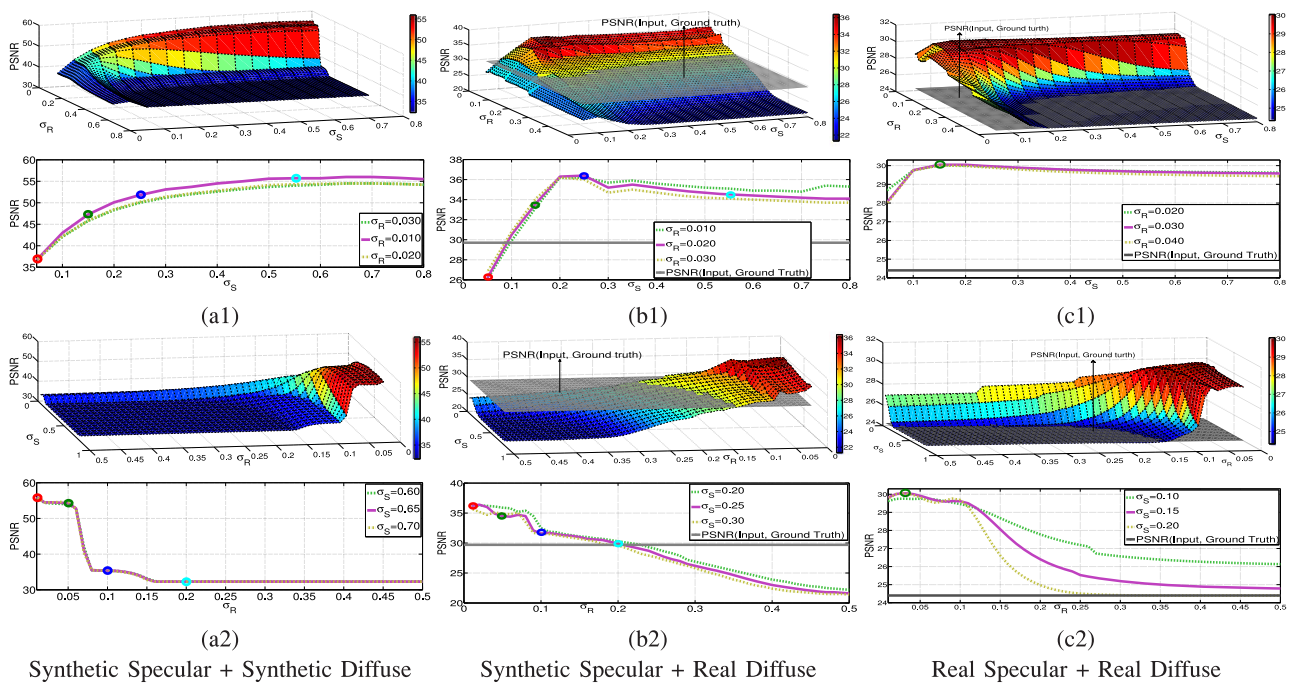
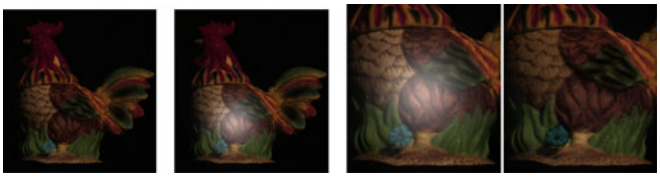


Fig. 2. Performance *w.r.t* to the bilateral filter parameters σ_S and σ_R.



(a) Captured diffuse image. (b) Synthetic specular reflection. (c) Closeup of (b). (d)Diffuse from (c).

Fig. 3. A real diffuse image and the corresponding synthetic specular intensity and specular image.

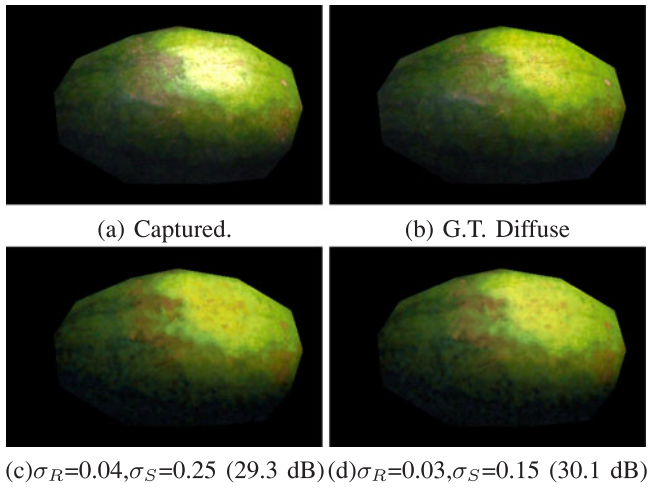
performance actually decreases when the photometric spread σ_R drops from 0.02 to 0.01 due to sensor noise.

The conducted experiments show that the range filter kernel (controlled by σ_R) should be small enough to reject noise from other surfaces but cannot be too small to account for sensor noise in real images. In practice, we set σ_R to 0.04 ($= 10/255$). The spatial filter kernel should be large enough to cover the specular highlight. According to Figs. 2a1 and 2a2 and (b1)-(b2), we set σ_S to 0.25 so that the filter kernel is as large as the test image, and rely on a small range filter kernel to reject noise from other surfaces.

This conclusion is demonstrated using a real scene in Fig. 4a. The ground-truth diffuse reflection in Fig. 4b is computed based on polarizing filter [12]. Figs. 4c and 4d are diffuse reflections estimated using default and optimal parameters, respectively. They are visually similar and the PSNR difference is also small, which demonstrates the robustness of the default parameter setting. The detailed evaluation *w.r.t.* the bilateral filter parameters σ_S and σ_R is presented in Figs. 2c1 and 2c2.

2.4 Sensitivity to Illumination Chromaticity Estimation

The proposed algorithm assumes that illumination chromaticity is known as it can be measured with a white reference or estimated using an existing algorithm like [26]. In practice, there will be estimation error. For instance, the estimation error is around $0.01 \sim 0.03$ in [26]; and thus a highlight removal algorithm should be robust to a certain amount of estimation error. Fig. 5 quantitatively evaluates the performance of the proposed algorithm with



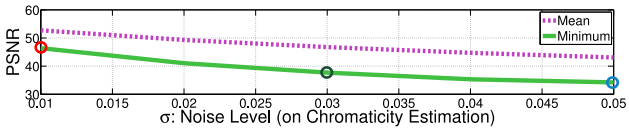
(a) Captured. (b) G.T. Diffuse (c)σ_R=0.04,σ_S=0.25 (29.3 dB) (d)σ_R=0.03,σ_S=0.15 (30.1 dB)

Fig. 4. Evaluation on a real scene. (c) and (d) are diffuse reflections estimated using constant and optimal parameters, respectively. Note that they are visually similar and the PSNR difference is also small. The images are gamma corrected for better illustration.

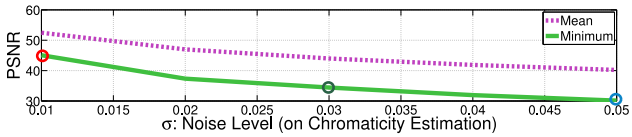
respect to different levels of noise (in illumination chromaticity estimation). A total of $1{,}000$ images were synthesized under each noise level as follows:

$$J_c^{noisy} = J_c^D + (1/3 + \mathbf{randc}() * \sigma)(3J^S), \tag{19}$$

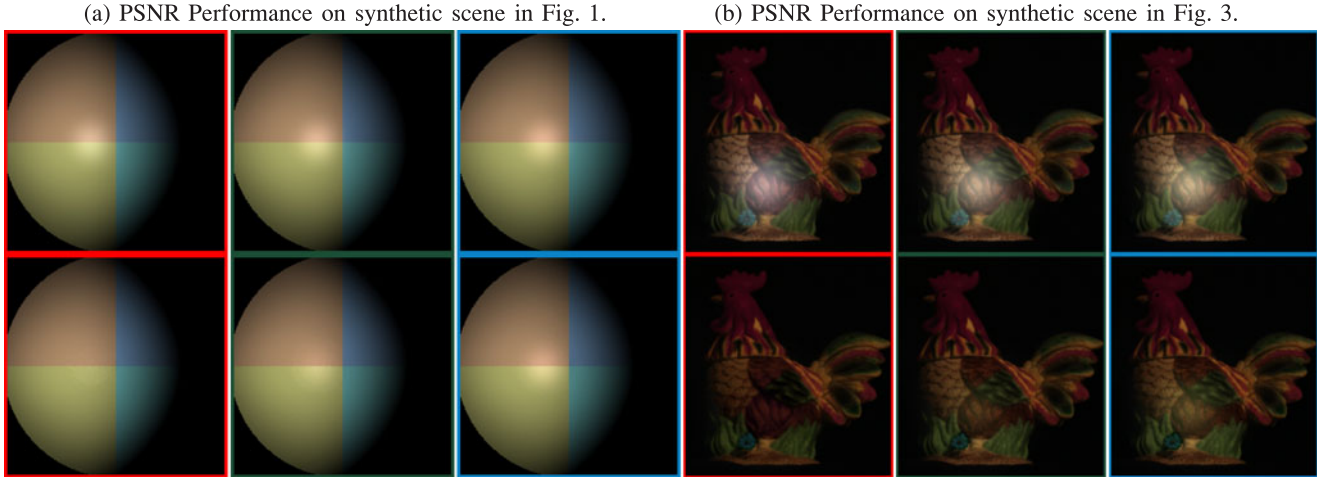
where function $\mathbf{randc}()$ generates a random value between -1 and 1, and $\sigma = \{0.01,0.02,0.03,0.04,0.05\}$ correspond to five different noise levels. The ground-truth diffuse and specular components (J_c^D and J^S) in Figs. 1a2 and 3b were used in Eq. (19), which results in a total of $10,000$ test images. The PSNR value computed from two diffuse images estimated with and without noise is used as the evaluation criteria. The mean and minimum PSNR values estimated under different noise levels are presented in Figs. 5a and 5b. As can be seen, the mean PSNR values are always higher than 40 dB (which corresponds to almost no visible difference). This shows that the proposed algorithm is robust to standard illumination



(a) PSNR Performance on synthetic scene in Fig. 1.



(b) PSNR Performance on synthetic scene in Fig. 3.



(a1) σ = 0.01 (46.4 dB) (a2) σ = 0.03 (37.7 dB) (a3) σ = 0.05 (34.2 dB) (b1) σ = 0.01 (45.1 dB) (b2) σ = 0.03 (34.4 dB) (b3) σ = 0.05 (30.1 dB)

Fig. 5. Sensitivity to illumination chromaticity.

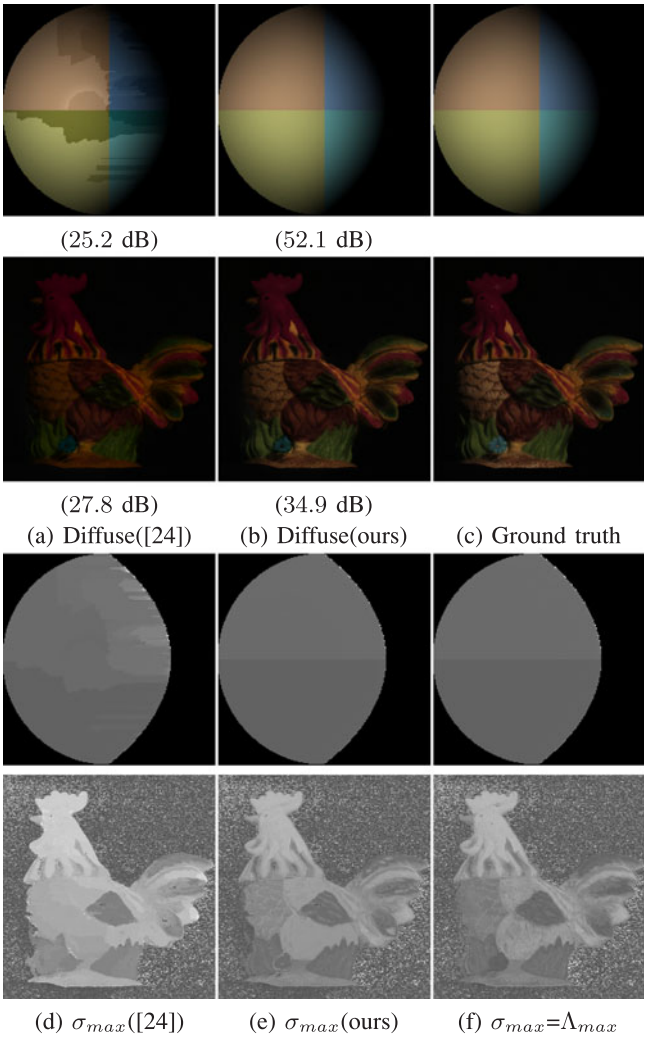


Fig. 6. Highlight removal on synthetic images in Figs. 1a2 and 3b.

chromaticity estimation error which is around $0.01 \sim 0.03$. Figs. 5a1, 5a2, and 5a3 and (b1)-(b3) are the estimated diffuse images corresponding to the red, green and blue circles in (a) and (b), respectively. They represent the worst case under three different noise levels. Note that there are very few visible errors when $\sigma \leq 0.03$.

3 Experimental Results

This section presents evaluation on both synthetic and real images. According to the discussion presented in Section 2.3, *we set* $\sigma_S = 0.25$ *(so that the filter covers the whole image) and* $\sigma_R = 0.04$ *(to reject noise from other surfaces) for all the experiments conducted in this section*.

We first compare the proposed method with [24] in Fig. 6 using the synthetic images in Figs. 1a2 and 3b. Figs. 6a, 6b are the diffuse images extracted using [24] and the proposed algorithm, respectively. (c) is the ground truth. (d)-(f) present the maximum chromaticity σ_{max} computed from (a)-(c), respectively. The numbers under the images are PSNR values computed based on the ground truth presented in (c). Both visual and quantitative evaluation show improvement in the robustness with respect to [24].

We next conducted experiments on several real images either used in the previous work [24] or captured by a Sony DFW-X700 camera with gamma correction off. For the images captured by ourselves, we compared our results with images captured with polarizing filters over the camera and the light source. Comparison with the highlight removal method presented in [24] is also provided. Fig. 7 compare our method with [24] using the images provided by the author of [24]. From left to right are the input image, diffuse image extracted using [24] and our method, respectively. Visual comparison shows that our method is comparable to [24] for these data sets but much faster.

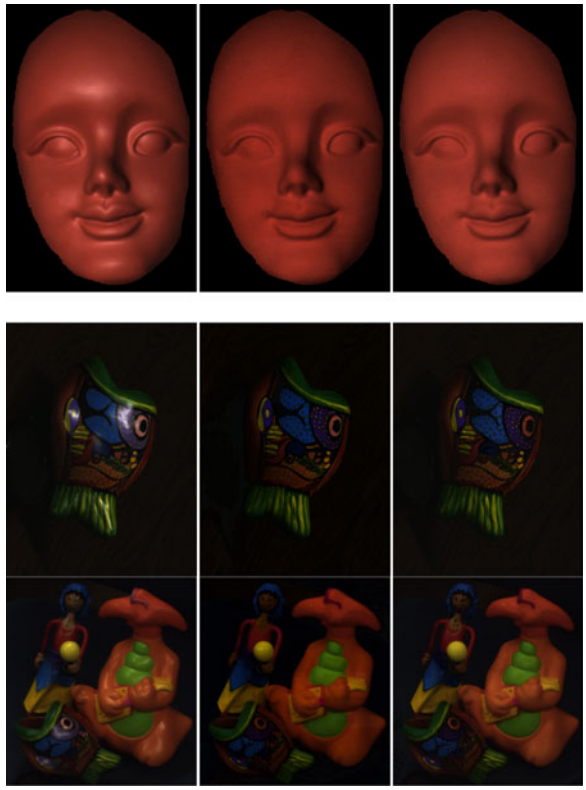


Fig. 7. Highlight removal on the images provided in [24].

Finally, we conducted experiments on two real images with diffuse reflections captured by polarizing filters over the camera and the light source. The experimental results are presented in Fig. 8. The measured diffuse reflections are presented in Fig. 8d. Fig. 8e presents our approximations (Eq. (14)) of the maximum diffuse chromaticities Λ_{max} which are used as the guidance image for the joint bilateral filtering process in Algorithm 1. We next computed the maximum chromaticity σ_{max} from Figs. 8b, 8c, and 8d, and presented the results in Figs. 8f, 8g, and 8h, respectively. Visual comparison between Figs. 8b, 8c, and 8d and (f)-(h) shows that both methods are suitable for low-textured images. Fig. 8 also shows our method is a bit better than [24] for highly-textured image but neither is of high quality as both methods are invalid for grayscale surfaces as shown in the PSNR values under (f) and (g) which are computed based on the ground-truth Λ_{max} in (h).

Fig. 9 compares the runtime of the method presented in [24] and ours. Note that the speedup factor of our method is generally over 200 except for the second data set, which is the synthetic image presented in Fig. 1. But note that [24] obtains incorrect diffuse reflection from this data set; thus probably, the iterative algorithm presented in [24] stops before convergence.

4 Conclusions

We have proposed a new highlight removal model in this paper. Using a single colour image, the highlight removal problem is formulated as an iterative bilateral filtering process which normally

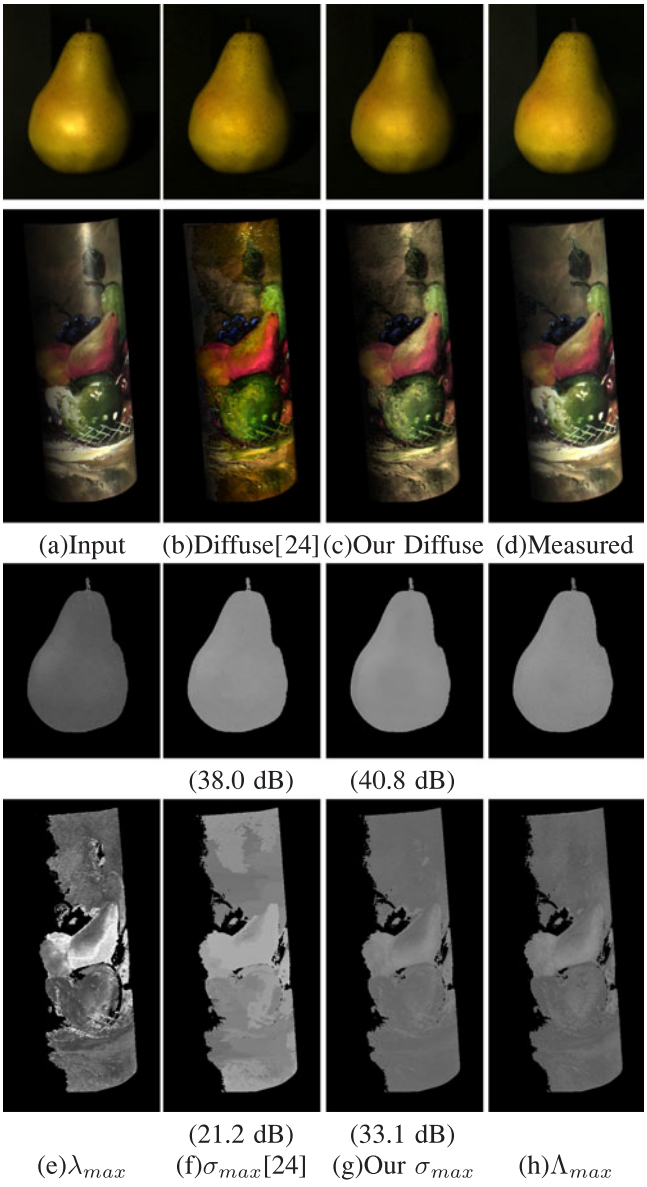


Fig. 8. Highlight removal on both low-textured and highly-textured images.

converges in two to three iterations. Unlike previous methods, the presented technique can process high-resolution images at video rate and thus is suitable for real-time applications, *e.g.*, stereo matching for specular surfaces. Besides, this technique does not result in noticeable artifacts, and guarantees that the estimated diffuse reflections will be locally smooth.

However, our algorithm highly depends on the approximate maximum diffuse chromaticity λ_{max}, which however has limitation as shown in Claim 2. Fig. 10 shows a failure case of our algorithm. Figs. 10a, 10b, and 10c are the input image, ground-truth diffuse image, and the maximum diffuse chromaticity Λ_{max},

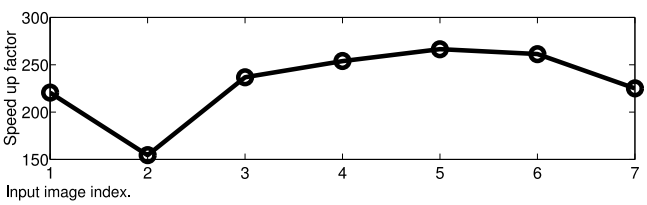


Fig. 9. Speed comparison. This figure shows that our method runs over $200\times$ faster than [24] on a standard CPU.

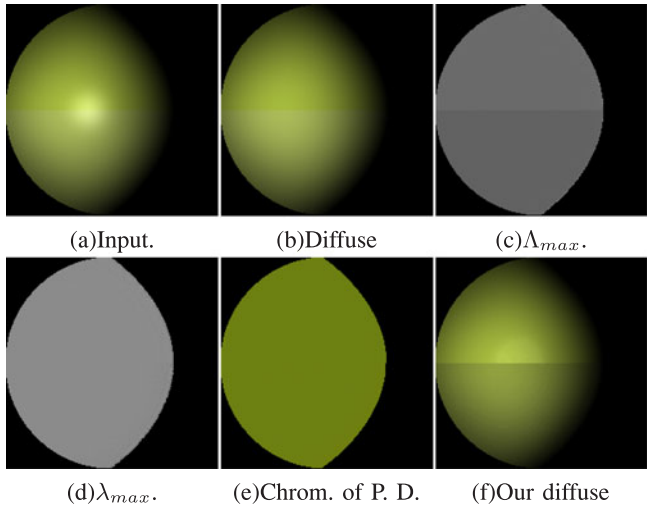


Fig. 10. A failure case.

respectively. Note that unlike Λ_{max}, λ_{max} and the chromaticity of the pseudo-code diffuse image (proposed in [24]) in Figs. 10d and 10e are uniformly distributed on every pixel. The use of λ_{max} as smoothing guidance is thus incorrect and results in visible errors in Fig. 10f.

Acknowledgments

This work was supported by a grant from HP Labs and a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21201914).

References

[1] R. Bajcsy, S. W. Lee, and A. Leonardis, "Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation," *Int. J. Comput. Vis.*, vol. 17, no. 3, pp. 241–272, 1996.
[2] A. Blake, "Specular stereo," in *Proc. 9th Int. Joint Conf. Artif. Intell.*, 1985, pp. 973–976.
[3] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Proc. 29th Annu. Conf. Comput. Graphics Interactive Tech.*, 2002, vol. 21, pp. 257–266.
[4] R. W. Fleming, A. Torralba, and E. H. Adelson, "Specular reflections and the perception of shape," *J. Vis.*, vol. 4, pp. 798–820, 2004.
[5] H. Kim, H. Jin, S. Hadap, and I. Kweon, "Specular reflection separation using dark channel prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 1460–1467.
[6] G. Klinker, S. Shafer, and T. Kanade, "The measurement of highlights in color images," *Int. J. Comput. Vis.*, vol. 2, no. 1, pp. 7–32, 1988.
[7] S. W. Lee and R. Bajcsy, "Detection of specularity using color and multiple views," in *Proc. 2nd Eur. Conf. Comput. Vis.*, 1992, pp. 99–114.
[8] S. Lin, Y. Li, S. B. Kang, X. Tong, and H.-Y. Shum, "Diffuse-specular separation and depth recovery from image sequences," in *Proc. Eur. Conf. Comput. Vis.*, 2002, pp. 210–224.
[9] S. Lin and H.-Y. Shum, "Separation of diffuse and specular reflection in color images," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2001, pp. 341–346.
[10] S. P. Mallick, T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Specularity removal in images and videos: A PDE approach," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 550–563.
[11] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur, "Beyond lambert: Reconstructing specular surfaces using color, " in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2005, vol. pp. II, pp. 619–626.
[12] S. K. Nayar, X. S. Fang, and T. Boult, "Separation of reflection components using color and polarization," *Int. J. Comput. Vis.*, vol. 21, no. 3, pp. 163–186, 1996.
[13] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," *Int. J. Comput. Vis.*, vol. 81, pp. 24–52, 2009.
[14] S. Paris, S. Hasinoff, and J. Kautz, "Local Laplacian filters: Edge-aware image processing with a Laplacian pyramid," *Trans. Graph.*, vol. 30, pp. 68:1–68:12, 2011.
[15] S. Paris, P. Kornprobst, and J. Tumblin, *Bilateral Filtering: Theory and Applications*. Delft, The Netherlands: Now Publishers, 2009.
[16] J. S. Park and J. T. Tou, "Highlight separation and surface orientation for 3-d specular objects," in *Proc. 10th Int. Conf. Pattern Recog.*, 1990, vol. pp. I, pp. 331–335.

[17] F. Porikli, "Constant time O(1) bilateral filtering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2008, pp. 1–8.
[18] Y. Sato and K. Ikeuchi, "Temporal-color space analysis of reflection," *J. Opt. Soc. Amer.*, vol. 11, no. 11, pp. 2990–3002, 1994.
[19] S. A. Shafer, "Using color to separate reflection components," *Color Res. App.*, vol. 10, no. 4, pp. 210–218, 1985.
[20] P. Tan, S. Lin, L. Quan, and H.-Y. Shum, "Highlight removal by illumination-constrained inpainting," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, 2003, p. 164.
[21] P. Tan, L. Quan, and S. Lin, "Separation of highlight reflections on textured surfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2006, pp. 1855–1860.
[22] R. Tan, (2005). Highlightremoval from a single image [Online]. Available: http://php-robbytan.rhcloud.com/code/highlight.zip.
[23] R. Tan and K. Ikeuchi, "Reflection components decomposition of textured surfaces using linear basis functions," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2005, vol. pp. I, pp. 125–131.
[24] R. Tan and K. Ikeuchi, "Separating reflection components of textured surfaces using a single image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 178–193, Feb. 2005.
[25] R. Tan, K. Nishino, and K. Ikeuchi, "Separating reflection components based on chromaticity and noise analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 10, pp. 1373–1379, Oct. 2004.
[26] R. T. Tan, K. Nishino, and K. Ikeuchi, "Illumination chromaticity estimation using inverse-intensity chromaticity space," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2003, pp. 673–680.
[27] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.
[28] Q. Yang, K.-H. Tan, and N. Ahuja, "Real-time o(1) bilateral filtering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2009, pp. 557–564.
[29] Q. Yang, S. Wang, and N. Ahuja, "Real-time specular highlight removal using bilateral filtering," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 87–100.