Real-Time Specular Highlight Removal Using Bilateral Filtering⋆

Qingxiong Yang, Shengnan Wang, and Narendra Ahuja

University of Illinos, Urbana Champaign
http://vision.ai.uiuc.edu/~qyang6/

Abstract. In this paper, we propose a simple but effective specular highlight removal method using a single input image. Our method is based on a key observation - the maximum fraction of the diffuse color component (so called maximum diffuse chromaticity in the literature) in local patches in color images changes smoothly. Using this property, we can estimate the maximum diffuse chromaticity values of the specular pixels by directly applying low-pass filter to the maximum fraction of the color components of the original image, such that the maximum diffuse chromaticity values can be propagated from the diffuse pixels to the specular pixels. The diffuse color at each pixel can then be computed as a nonlinear function of the estimated maximum diffuse chromaticity. Our method can be directly extended for multi-color surfaces if edge-preserving filters (e.g., bilateral filter) are used such that the smoothing can be guided by the maximum diffuse chromaticity. But maximum diffuse chromaticity is to be estimated. We thus present an approximation and demonstrate its effectiveness. Recent development in fast bilateral filtering techniques enables our method to run over **200×** faster than the state-of-the-art on a standard CPU and differentiates our method from previous work.

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 [14], 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 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 the diffuse component analysis. Since the presence of specular reflection is inevitable in real world, and they do capture important scene information, e.g., surface shape and source characteristics, incorporation of specularity regions in the analysis is important.

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 single. [9] used multiple images captured from different polarization angles. Sato and Ikeuchi [13] employed the dichromatic model for separation by analyzing color signatures in many images captured with a moving light source. Lin and Shum [6] also changed the light source direction to produce two photometric images and used linear basis functions to separate the specular components. [11] 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 view point instead of changing the illumination direction. Using multiple images taken from different viewing directions, Lee [4] presented a method for specular region detection and Lin [5] 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 can not 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-colored images, most single-image-based methods require color segmentation (*e.g.*, [3],[1]) which is known to be non-robust for complex textured images or requires user assistance for highlight detection [16]. It was therefore a significant advance when Tan and Ikeuchi [20] demonstrated that highlights from textured objects with complex multi-colored scenes can be effectively removed without explicit color segmentation. This method removes highlights by iteratively shifting chromaticity values towards those of the neighboring pixel having the maximum chromaticity in the neighborhood. The neighborhood 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. Assuming that the specular intensity is either zero (for diffuse pixels) or a constant, Shen and Cai [15] introduced a fast highlight removal method using a modified "pseudo-coded" diffuse image, but this method only works for multi-color surfaces when the dominant highlight region is approximately uniform. Similar to the "pseudo-coded" diffuse image presented in [20], Mallick *et. al.* [8] proposed an SUV color space which separated the specular and diffuse components into 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 [7]. This type of approaches may encounter problems due to discontinuities in surface colors, across which diffuse information cannot be accurately propagated.

Other approaches for single-image highlight removal analyze the distributions of image colors within a color space. Tan and Ikeuchi [19] related the specular pixels to diffuse pixels for every surface color by projecting image colors along the illumination color 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, color blending at edges, and multiple surface colors. By integrating the texture from outside the highlight to determine the candidate diffuse color for traditional

color-space technique, that is for each pixel, a set of candidate diffuse colors is obtained from a texture scale of 1×1, and is iteratively pruned as the texture scale increase, Tan et al. [17] showed appreciable improvements in diffuse-highlight separation. This method requires that there are enough repetitive textures locally and the highlight does not have similar color to the surface.

All above methods that use a single input image share the same problem that they are not capable for 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 [20], in which, the diffuse color of a specular pixel is derived as a nonlinear function of its color (from input image) and the maximum fraction of the diffuse color components which is denoted as maximum diffuse chromaticity in 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 shifting chromaticity values towards those of the neighboring pixel having the maximum chromaticity in the neighborhood was proposed. Our method is the same as [20] except for the way of estimating the maximum diffuse chromaticity. Based on a key observation - the maximum diffuse chromaticity in local patches in colorful images generally changes smoothly, we estimate the maximum diffuse chromaticity values of the specular pixels by directly applying low-pass filter to the maximum fraction of the color 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-color surfaces if edge-preserving filters (e.g., bilateral filter) 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.

Our method have benefited a lot from the recent development in fast bilateral filtering techniques [12], [23], [24], which enables our method to run **200$\times$** faster than [20][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 1MB 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 colors as presented in [20]. The use of low-pass filter guarantees that the estimated maximum diffuse chromaticities will be locally smooth, so are the estimated diffuse reflections. Nevertheless, since the theory of our method is heavily built upon [20], it shares most of the limitations with [20], *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 (Sec. 2.1), and then present a real-time highlight removal method (Sec. 2.2).

---

[1] The source code is available on its author's homepage [18].

2.1   Reflection Model

Using standard diffuse+specular reflection models commonly used in computer graphics, the reflected light color (J) captured by a RGB camera can be represented as a linear combination of diffuse (J^D) and specular (J^S) colors:

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

Let chromaticity be defined as the fraction of color 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:

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

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

Following the chromaticity definition in Eqn. (2), (3) and (4), we express the reflected light color J_c as

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

Assumed that the illumination chromaticity can be measured (with a white reference) or estimated [21], 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 Eqn. 5.

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

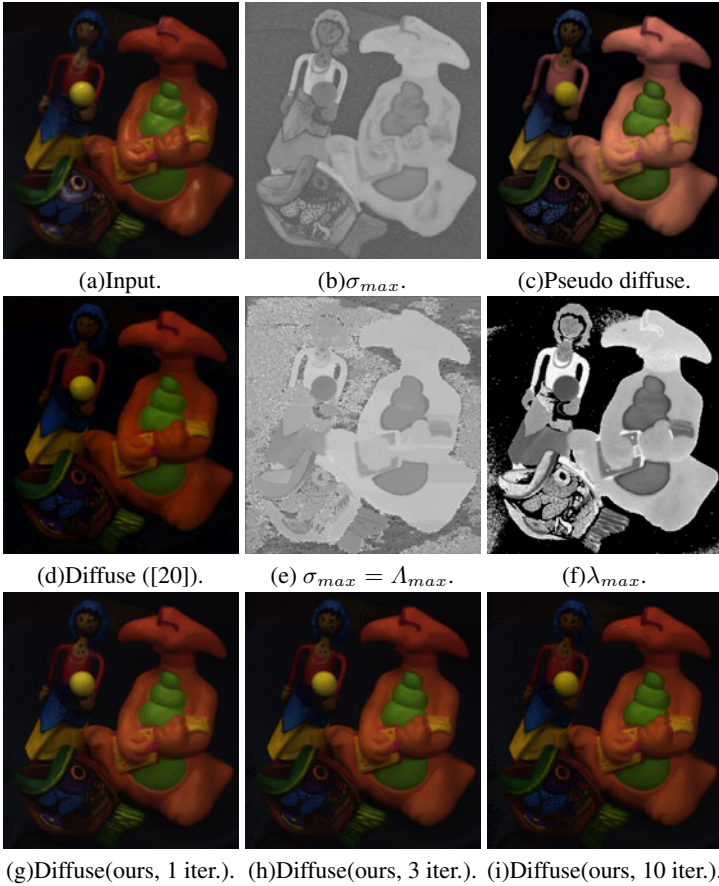$$\sigma_{max} = \max(\sigma_r, \sigma_g, \sigma_b) \tag{7}$$

and maximum diffuse chromaticity be

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

Tan [20] shows that the diffuse component can be represented as a function of Λ_{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, Λ_{max} changes from pixel to pixel in real images but is limited from $\frac{1}{3}$ to 1.

(a)Input.   (b)σ_{max}.   (c)Pseudo diffuse.

(d)Diffuse ([20]).   (e) $\sigma_{max} = \Lambda_{max}$.   (f)$\lambda_{max}$.

(g)Diffuse(ours, 1 iter.). (h)Diffuse(ours, 3 iter.). (i)Diffuse(ours, 10 iter.).

Fig. 1. Highlight removal on real data set. (a) is the input image; (b) is the maximum chromaticity σ_{max} (Eqn. 2) computed from (a); (c) is the "pseudo-coded" diffuse image computed using Eqn. (9) by setting Λ_{max} to a constant 0.5; (d) is the extracted diffuse reflection using the method presented in [20][1]; (e) is the maximum chromaticity σ_{max} computed from (d) using Eqn. (2) and (7). Assuming (d) is the ground-truth diffuse reflection, then (e) is also the maximum diffuse chromaticity Λ_{max}; (f) is our approximation (λ_{max}, Eqn. 12) of the maximum diffuse chromaticity Λ_{max}; (g)-(i) are the extracted diffuse reflections using our method after 1, 3 and 10 iterations, respectively. As can be seen, our method generally converges after 3 iterations, and (h) and (i) are visually closed to (d).

Estimating the maximum diffuse chromaticity Λ_{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 2D image, while the ground-truth diffuse image is a 3D image. The 2D "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 [20]. Fig. 1 presents such an example by setting Λ_{max} to a constant 0.5. Fig. 1 (a) is the input image, (b) is the maximum chromaticity values σ_{max} computed from the input image (a), (c) is the "pseudo-coded" diffuse image, (d) presents diffuse reflection extracted using the method presented in [20] and (e) presents the maximum chromaticity values σ_{max} computed from (d) using Eqn. (2) and (7). Assume that the specular highlights are correctly removed from (d), (e) is also presented as the maximum diffuse chromaticity Λ_{max}.

2.2   Highlight Removal Using Bilateral Filter

According to Eqn. (9), the highlight removal problem can be reduced as the searching for the maximum diffuse chromaticity Λ_{max} which changes from pixel to pixel. However, as shown in Fig. 1 (d) and (e), the variance of Λ_{max} is very small in local patches when the surface colors are consistent. The maximum chromaticity σ_{max} in Fig. 1 (b) is the same as the maximum diffuse chromaticity Λ_{max} except for specular pixels, which cause the intensity/color discontinuities within local patches of the same surface color. Intuitively, applying low-pass filtering to the maximum chromaticity σ_{max} in Fig. 1 (b) 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 σ_{max} values of two pixels associated with different surface materials (Λ_{max} values are different) won't be blend together.
2. The diffuse pixels will be affected by the specular pixels after smoothing.

As a popular edge-aware operator, joint bilateral filter can be employed to smooth the maximum chromaticity σ_{max} using the maximum diffuse chromaticity Λ_{max} as the smoothing guidance. But Λ_{max} is to be estimated, thus we need to find a substitution or an approximation. Although the "pseudo-coded" diffuse image presented in [20] is free of specularity, it is not a good substitution for this problem because its color 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 approximiate 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 approximated diffuse chromaticity λ_c and the real diffuse chromaticity Λ_c is captured in Theorem 1 and Theorem 2.

Theorem 1. *For any two pixels p and q, if $\Lambda_c(p) = \Lambda_c(q)$, then $\lambda_c(p) = \lambda_c(q)$.*

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

Note that λ_c is just an approximation of Λ_c, which will fail for the specific case specified in Theorem 2. However, it is the best estimate we can get. Fig. 1 (f) presents the maximum values of the approximated diffuse chromaticity

$$\lambda_{max} = \max(\lambda_r, \lambda_g, \lambda_b) = \max(\frac{\sigma_r - \sigma_{min}}{1 - 3\sigma_{min}}, \frac{\sigma_g - \sigma_{min}}{1 - 3\sigma_{min}}, \frac{\sigma_b - \sigma_{min}}{1 - 3\sigma_{min}}) \qquad (12)$$

computed from the input image presented in Fig. 1 (a).

Using the approximated maximum diffuse chromaticity defined in Eqn. (12) to guide the smoothing, the filtered maximum chromaticity σ_{max} can be computed as follows

$$\sigma_{max}^F(\boldsymbol{p}) = \frac{\sum_{\boldsymbol{q} \in \Omega} \mathcal{F}(\boldsymbol{p}, \boldsymbol{q}) \mathcal{G}(\lambda_{max}(\boldsymbol{p}), \lambda_{max}(\boldsymbol{q})) \sigma_{max}(\boldsymbol{q})}{\sum_{\boldsymbol{q} \in \Omega} \mathcal{F}(\boldsymbol{p}, \boldsymbol{q}) \mathcal{G}(\lambda_{max}(\boldsymbol{p}), \lambda_{max}(\boldsymbol{q}))}, \qquad (13)$$

where $\mathcal{F}$ and $\mathcal{G}$ are spatial and range weighting functions which are typically Gaussian in the literature [22], [2].

The variances of the maximum chromaticity σ_{max} due to specular highlights will be reduced after filtering, and the filtered maximum chromaticity σ_{max}^F will be more closed to Λ_{max} than σ_{max} for the specular pixels. However, after smoothing, the diffuse pixels will be affected by the specular pixels too. According to Theorem 3, the filtered maximum chromaticity values σ_{max}^F of the diffuse pixels will be lower than the unfiltered values σ_{max}. As a result, to exclude the contribution of the specular pixels, we compare σ_{max}^F and σ_{max} and take the maximum value:

$$\sigma_{max}(\boldsymbol{p}) = \max(\sigma_{max}, \sigma_{max}^F(\boldsymbol{p})). \qquad (14)$$

Theorem 3. *Assume $\Gamma_c = \frac{1}{3}$, then $\Lambda_{max} \geq \sigma_{max}$. Equality holds when the $\Lambda_{max} = \frac{1}{3}$.*

We then iteratively apply joint bilateral filter to σ_{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 σ_{max}^F with σ_{max} after every iteration. The algorithm is believed to converge when their difference is smaller than a threshold (set to 0.03 in our experiments) at every pixel. Our method generally converges after $2 - 3$ iterations. Fig. 1 (g)-(i) present the extracted diffuse reflections after 1, 3 and 10 iterations, respectively. The proposed highlight removal algorithm is summarized in Algorithm 1.
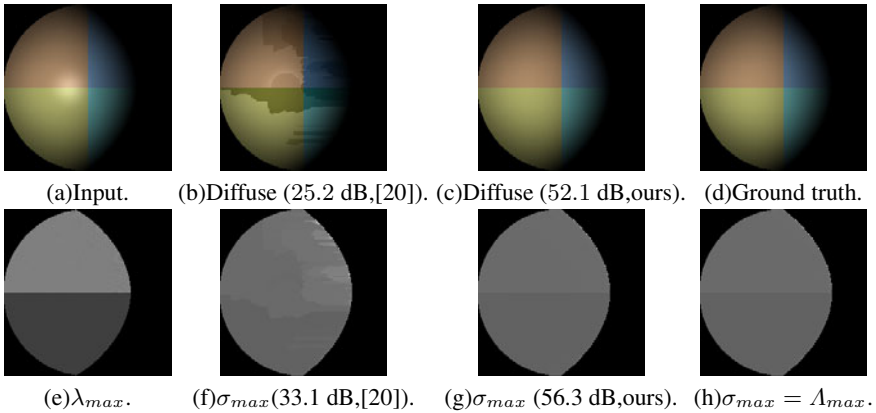

Algorithm 1. Highlight removal using a single RGB image

1: Compute σ_{max} at every pixel using the input image and store it as a grayscale image.
2: Compute λ_{max} at every pixel using the input image and store it as a grayscale image.
3: **repeat**
4:    -Apply joint bilateral filter to image σ_{max} using λ_{max} as the guidance image (Eqn. 14), store the filtered image as σ_{max}^F;
5:    -For each pixel $\boldsymbol{p}$, $\sigma_{max}(\boldsymbol{p}) = \max(\sigma_{max}(\boldsymbol{p}), \sigma_{max}^F(\boldsymbol{p}))$;
6: **until** $\sigma_{max}^F - \sigma_{max} < 0.03$ at every pixel.
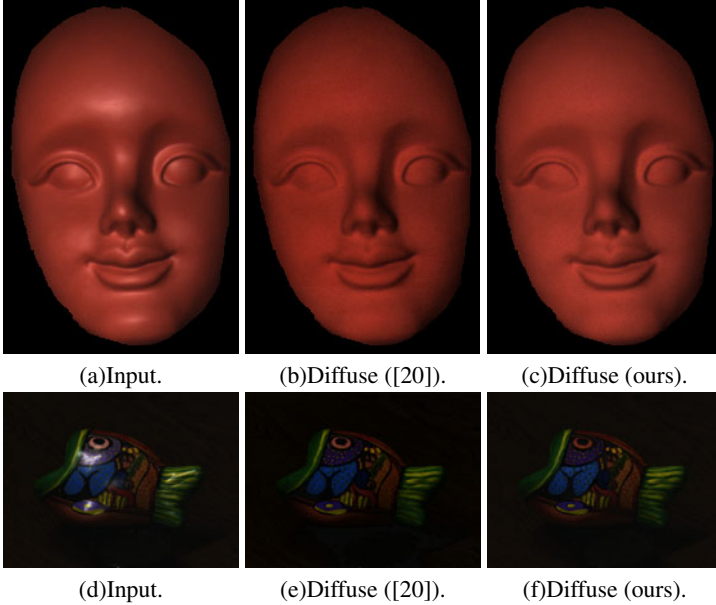
3   Experimental Results

To evaluate our method, we conducted experiments on a synthetic data set and several real images either used in the previous work [20] or captured by a Sony DFW-X700 camera with gamma correction off. Quantitative evaluation can only be performed on the synthetic image. For real images, 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 [20] is also provided.

We first numerically compare our method with the method presented in [20] using a synthetic image (used in [6]) presented in Fig. 2 (a). It turns out that [20] performs poorly on this dataset. [20] propagates the maximum chromaticities from diffuse pixels to specular pixels by comparing the maximum chromaticities of two pixels each time, and the greater value is adopted. However, the propagation fails to stop at color edges, and separates incorrect diffuse components for this image. Our method uses a local patch instead of only two pixels, and it is thus more robust and propagates the diffuse chromaticity faster. For numerically evaluation, we compute peak signal-to-noise ratio (PSNR) from the extracted diffuse reflections and the ground-truth presented in Fig. 2 (d). PNSR larger than 40 dB often corresponds to almost invisible differences

(a)Input.   (b)Diffuse (25.2 dB,[20]).   (c)Diffuse (52.1 dB,ours).   (d)Ground truth.

(e)λ_{max}.   (f)σ_{max}(33.1 dB,[20]).   (g)σ_{max} (56.3 dB,ours).   (h)$\sigma_{max} = \Lambda_{max}$.

Fig. 2. Highlight removal on a synthetic image. From (a)-(d): input image, diffuse reflection extracted using [20][1], diffuse reflection extracted using our method and the ground truth. (e) is our approximation (Eqn. 12) of the maximum diffuse chromaticity Λ_{max}. Unlike the "pseudo-coded" diffuse image presented in [20], our approximation is invariant to surface geometry. (e) is used as the guidance image for the bilateral filtering process in Algorithm 1. (f)-(h) present the maximum chromaticity computed from (b)-(d), respectively. As can be seen in (b) and (f), [20] has two main problems for this data set: (i) pixels around highlight edges remain specular although there is no visible discontinuities around the highlight regions in the estimated maximum diffuse chromaticity as presented in (f); (ii) [20] propagates the maximum chromaticities from diffuse pixels to specular pixels. However, the fact that the propagation does not stop at color edges results in extracting incorrect diffuse reflections. Our method does not have these problems as can be seen in (c) and (g). The PSNR value computed from (b) and (d) is 25.2 dB, and 52.1 dB from (c) and (d).

(a)Input.   (b)Diffuse ([20]).   (c)Diffuse (ours).

(d)Input.   (e)Diffuse ([20]).   (f)Diffuse (ours).

Fig. 3. Highlight removal on the images provided in [20][1]. From left to right: input image, diffuse reflection extracted using [20], diffuse reflection extracted using our method. Visual comparison shows that our method is comparable to [20] for these data sets but much faster.
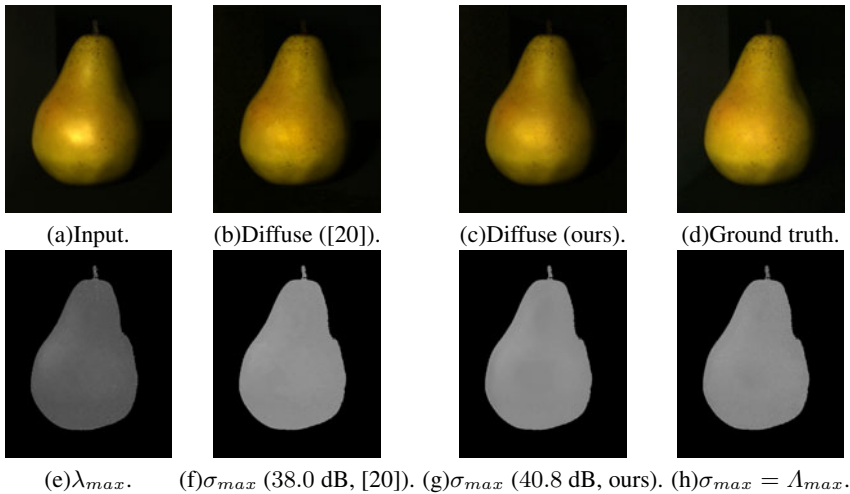
as suggested in [10]. The PSNR values obtained using [20] and our method are 25.2 and 52.1 dB, respectively, which agree with our subjective evaluation that there is no visible difference between the extracted diffuse reflection using our method (Fig. 2 (c)) and the ground truth (Fig. 2 (d)). Fig. 2 (e) presents our approximation (Eqn. 12) of the maximum diffuse chromaticity Λ_{max} which is used as the guidance image for the bilateral filtering process in Algorithm 1. Fig. 2 (f)-(h) present the maximum chromaticity computed from (b)-(d), respectively.

Fig. 1 and 3 compare our method with [20] using the images provided by the author of [20]. Visual comparison shows that the performance of the two methods is similar while no ground truth is available for quantitative comparison.
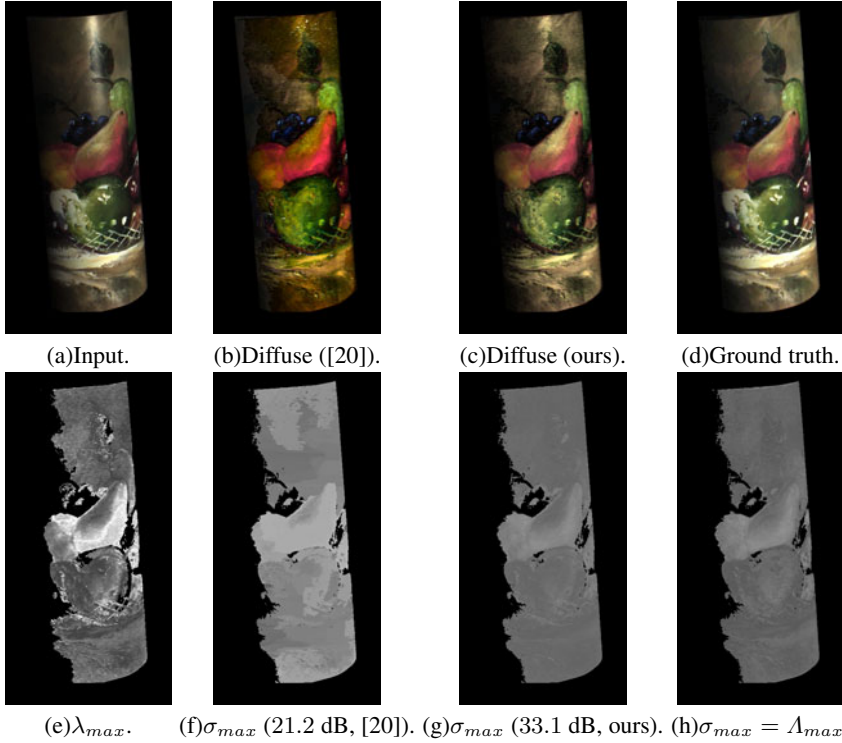
Finally, we conducted experiments on two real images with ground-truth diffuse reflections captured by polarizing filters over the camera and the light source. The experimental results are presented in Fig. 4 and 5. The ground-truth diffuse reflections are presented in Fig. 4 (d) and 5 (d). Fig. 4 (e) and 5 (e) present our approximations (Eqn. 12) of the maximum diffuse chromaticities Λ_{max} which are used as the guidance image for the bilateral filtering process in Algorithm 1. We next computed the maximum chromaticity σ_{max} using Fig. 4 (b)-(d) and 5 (b)-(d), respectively, and presented the results in Fig. 4 (f)-(h) and 5 (f)-(h). The dark pixels are excluded to avoid quantization noise. We next compared (f) and (g) in Fig. 4 and 5 with the ground-truth maximum diffuse chromaticity images presented in Fig. 4 (h) and Fig. 5 (h) since the diffuse reflection is represented as a function of the maximum diffuse chromaticity Λ_{max}. Accurate Λ_{max}

results in accurate diffuse reflection. We don't use the PSNR values computed from the ground-truth diffuse reflection in Fig. 4 (d) and Fig. 5 (d) because the amount of light captured with and without the polarizing filters are different. The PSNR values computed from Fig. 4 (f) and (g) are 38.0 and 40.8 dB, which shows that both [20] and our method are suitable for single-color objects. The PSNR values computed from Fig. 5 (f) and (g) are 21.2 and 33.1 dB. Hence, the estimated diffuse reflection using our method (Fig. 5 (g)) is not very accurate although it is a bit better than [20] (Fig. 5 (f)). However, if we exclude pixels with maximum chromaticity less than or equal to 0.37, which means that most of the pixels near neutral (our method is limited to chromatic surfaces) are excluded, the PSNR values obtained using [20] and our method are raised to 24.7 and 40.7 dB, respectively. In this case, the diffuse reflection computed using our method is comparable to the ground truth as PSNR larger than 40 dB often corresponds to almost invisible differences. However, the diffuse reflection extracted using the method presented in [20] is still of low quality as can be seen from the obtained PSNR values. Both visual and numerical comparison on real images show that our method is a bit more robust/accurate than [20], but both methods are invalid for grayscale surfaces.

Fig. 6 compares the runtime of the method presented in [20] and ours. Note that the speeeup factor of our method is generally over 200.

(a)Input.   (b)Diffuse ([20]).   (c)Diffuse (ours).   (d)Ground truth.

(e)λ_{max}.   (f)σ_{max} (38.0 dB, [20]). (g)σ_{max} (40.8 dB, ours). (h)$\sigma_{max} = \Lambda_{max}$.

Fig. 4. Highlight removal on a low-textured image. From (a)-(d): input image, diffuse reflection extracted using [20][1], diffuse reflection extracted using our method and the ground truth. (e) is our approximation (Eqn. 12) of the maximum diffuse chromaticity Λ_{max}. (e) is used as the guidance image for the bilateral filtering process in Algorithm 1. (f)-(h) are the maximum chromaticity computed from (b)-(d), respectively. The numbers under (f)-(g) are PSNR values computed by comparing with the ground truth in (h), which numerically proves both [20] and our method are suitable for low-textured scenes. The images are gamma corrected for better illustration.

(a)Input.   (b)Diffuse ([20]).   (c)Diffuse (ours).   (d)Ground truth.

(e)λ_{max}.   (f)σ_{max} (21.2 dB, [20]). (g)σ_{max} (33.1 dB, ours). (h)$\sigma_{max} = \Lambda_{max}$.

Fig. 5. Highlight removal on a heavy-textured image. From (a)-(d): input image, diffuse reflection extracted using [20][1], diffuse reflection extracted using our method and the ground truth. (e) is our approximation (Eqn. 12) of the maximum diffuse chromaticity Λ_{max}. (e) is used as the guidance image for the bilateral filtering process in Algorithm 1. (f)-(h) are the maximum chromaticity computed from (b)-(d), respectively. The numbers under (f)-(g) are PSNR values computed by comparing with the ground truth in (h). Note that neither our method nor [20] is of high quality when comparable to the ground truth as many pixels are close to neutral. Nevertheless, the PSNR value computed from our method is still a bit higher than the value computed from [20]. More over, visually comparison shows that the color of extracted diffuse reflection using [20] is a bit less accurate. The images are gamma corrected for better illustration.

Input image index (correspond to Fig. 1(a), 2(a), 3(a), 3(d), 4(a), 5(a), respectively).
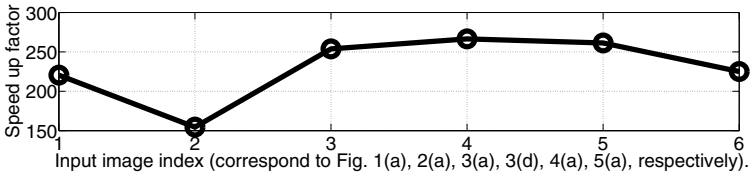
Fig. 6. Speed comparison. This figure shows that our method runs over $200\times$ faster than [20] on a standard CPU except for the second data set, which is the synthetic image presented in Fig. 2 (a). But note that [20] obtains incorrect diffuse reflection from this data set, thus probably, the iterative algorithm presented in [20] stops before convergence.

4   Conclusions

We have proposed a new highlight removal model in the paper. Using a single color image, the highlight removal problem is formulated as an iterative bilateral filtering process which normally converges in 2 to 3 iterations. Unlike previous methods, the presented technique can process high-resolution images at video rate 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.

References

1. Bajcsy, R., Lee, S., Leonardis, A.: Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation. IJCV 17(3), 241–272 (1996)
2. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. In: Siggraph, vol. 21 (2002)
3. Klinker, G., Shafer, S., Kanade, T.: The measurement of highlights in color images. IJCV 2(1), 7–32 (1988)
4. Lee, S., Bajcsy, R.: Detection of specularity using color and multiple views. In: Sandini, G. (ed.) ECCV 1992. LNCS, vol. 588, pp. 99–114. Springer, Heidelberg (1992)
5. Lin, S., Li, Y., Kang, S., Tong, X., Shum, H.Y.: Diffuse-specular separation and depth recovery from image sequences. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2352, pp. 210–224. Springer, Heidelberg (2002)
6. Lin, S., Shum, H.Y.: Separation of diffuse and specular reflection in color images. In: CVPR, pp. 341–346 (2001)
7. Mallick, S.P., Zickler, T., Belhumeur, P.N., Kriegman, D.J.: Specularity removal in images and videos: A pde approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 550–563. Springer, Heidelberg (2006)
8. Mallick, S.P., Zickler, T.E., Kriegman, D.J., Belhumeur, P.N.: Beyond lambert: Reconstructing specular surfaces using color. In: CVPR, pp. II619–II626 (2005)
9. Nayar, S., Fang, X., Boult, T.: Separation of reflection components using color and polarization. IJCV 21(3) (1996)
10. Paris, S., Durand, F.: A fast approximation of the bilateral filter using a signal processing approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3954, pp. 568–580. Springer, Heidelberg (2006)
11. Park, J., Tou, J.: Highlight separation and surface orientation for 3-d specular objects. In: ICPR, pp. I331–I335 (1990)
12. Porikli, F.: Constant time o(1) bilateral filtering. In: CVPR (2008)
13. Sato, Y., Ikeuchi, K.: Temporal-color space analysis of reflection. JOSA 11(11), 2990–3002 (1994)
14. Shafer, S.: Using color to separate reflection components. Color Res. App. 10(4), 210–218 (1985)
15. Shen, H.L., Cai, Q.Y.: Simple and efficient method for specularity removal in an image. Applied Optics 48(14), 2711–2719 (2009)
16. Tan, P., Lin, S., Quan, L., Shum, H.Y.: Highlight removal by illumination-constrained inpainting. In: ICCV, p. 164 (2003)
17. Tan, P., Quan, L., Lin, S.: Separation of highlight reflections on textured surfaces. In: CVPR, pp. 1855–1860 (2006)

18. Tan, R.: Highlight removal from a single image,
    `http://www.commsp.ee.ic.ac.uk/~rtan/code.html`
19. Tan, R., Ikeuchi, K.: Reflection components decomposition of textured surfaces using linear basis functions. In: CVPR, pp. I125–I131 (2005)
20. Tan, R., Ikeuchi, K.: Separating reflection components of textured surfaces using a single image. PAMI 27(2), 178–193 (2005)
21. Tan, R.T., Nishino, K., Ikeuchi, K.: Illumination chromaticity estimation using inverse-intensity chromaticity space. In: CVPR, pp. 673–680 (2003)
22. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: ICCV, pp. 839–846 (1998)
23. Yang, Q., Tan, K.H., Ahuja, N.: Real-time o(1) bilateral filtering. In: CVPR (2009)
24. Yang, Q., Wang, S., Ahuja, N.: SVM for Edge-Preserving Filtering. In: CVPR (2010)

A Proof of Theorem 1

For any two pixels p and q, let

$$J_{min} = \min(J_r, J_g, J_b), \tag{15}$$
$$J_{min}^D = \min(J_r^D, J_g^D, J_b^D), \tag{16}$$
$$\tag{17}$$

according to the definition of λ_c in Eqn. 11, we obtain

$$\lambda_c = (\sigma_c - \sigma_{min})/(1 - 3\sigma_{min}) \tag{18}$$

$$= \left(\frac{J_c}{\sum_{u \in \{r,g,b\}} J_u} - \frac{J_{min}}{\sum_{u \in \{r,g,b\}} J_u} \right)$$

$$/ \left(\frac{\sum_{u \in \{r,g,b\}} J_u}{\sum_{u \in \{r,g,b\}} J_u} - \frac{3 \cdot J_{min}}{\sum_{u \in \{r,g,b\}} J_u} \right)$$

$$= \left(\frac{J_c - J_{min}}{\sum_{u \in \{r,g,b\}} J_u} \right) / \left(\frac{\sum_{u \in \{r,g,b\}} (J_u - J_{min})}{\sum_{u \in \{r,g,b\}} J_u} \right)$$

$$= \frac{J_c - J_{min}}{\sum_{u \in \{r,g,b\}} (J_u - J_{min})}$$

$$= \frac{(J_c^D + J^S) - (J_{min}^D + J^S)}{\sum_{u \in \{r,g,b\}} ((J_u^D + J^S) - (J_{min}^D + J^S))}$$

$$= \frac{J_c^D - J_{min}^D}{\sum_{u \in \{r,g,b\}} (J_u^D - J_{min}^D)}$$

$$= \left(\frac{J_c^D - J_{min}^D}{\sum_{u \in \{r,g,b\}} J_u^D} \right) / \left(\frac{\sum_{u \in \{r,g,b\}} (J_u^D - J_{min}^D)}{\sum_{u \in \{r,g,b\}} J_u^D} \right)$$

$$= (\Lambda_c - \Lambda_{min}) / \left(\sum_{u \in \{r,g,b\}} (\Lambda_u - \Lambda_{min}) \right). \tag{19}$$

For any two pixels p and q, if $\Lambda_c(p) = \Lambda_c(q)$, Eqn. (19) ensures that $\lambda_c(p) = \lambda_c(q)$.

B   Proof of Theorem 2

For any two pixels p and q, assume $\lambda_c(p) = \lambda_c(q)$, from Equation (19) we obtain

$$\frac{\Lambda_c(p) - \Lambda_{min}(p)}{1 - 3\Lambda_{min}(p)} = \frac{\Lambda_c(q) - \Lambda_{min}(q)}{1 - 3\Lambda_{min}(q)}. \tag{20}$$

If $\Lambda_{min}(p) = \Lambda_{min}(q)$, from (20) we obtain $\Lambda_c(p) = \Lambda_c(q)$.

C   Proof of Theorem 3

let $J_{max} = \max(J_r, J_g, J_b)$, according to the definition of σ_{max},

$$\sigma_{max} = \frac{J_{max}}{\sum_{u \in \{r,g,b\}} J_u}, \tag{21}$$

$$= \frac{J_{max}^D + J^S}{\sum_{u \in \{r,g,b\}} J_u^D + 3J^S} \tag{22}$$

$$= (\Lambda_{max} + \frac{J^S}{\sum_{u \in \{r,g,b\}} J_u^D}) / (1 + \frac{3J^S}{\sum_{u \in \{r,g,b\}} J_u^D}) \tag{23}$$

$$= \Lambda_{max} + \frac{J^S}{\sum_{u \in \{r,g,b\}} J_u^D} - \sigma_{max} \frac{3J^S}{\sum_{u \in \{r,g,b\}} J_u^D} \tag{24}$$

$$= \Lambda_{max} + (1 - 3\sigma_{max}) \frac{J^S}{\sum_{u \in \{r,g,b\}} J_u^D}. \tag{25}$$

According to its definition, $\sigma_{max} \geq 1$, thus

$$\sigma_{max} - \Lambda_{max} \leq 0, \tag{26}$$

and the quality holds when the $\Lambda_{max} = \sigma_{max} = \frac{1}{3}$.