Single Image Specular Highlight Removal on Natural Scenes

Huaian Chen[1], Chenggang Hou[1], Minghui Duan[1], Xiao Tan[1], Yi Jin[1(✉)], Panlang Lv[2], and Shaoqian Qin[2]

[1] University of Science and Technology of China, Hefei, China
{anchen,wukong,dmhustc,tx2015}@mail.ustc.edu.cn, jinyi08@ustc.edu.cn
[2] FuHuang AgileDevice Co., Ltd., Hefei, China
https://www.gaosuxiangji.com/

Abstract. Previous methods of highlight removal in image processing have exclusively addressed images taken in specific illumination environments. However, most of these methods have limitations in natural scenes and thus, introduce artifacts to nature images. In this work, we propose a specular highlight removal method that is applicable to natural scene image. Firstly, we decompose the input image into a reflectance image and an illumination image based on Retinex theory, and show that the illumination image of natural scene is obviously different from that of commonly used experimental scene. Then, the smooth features of the input image are extracted to help estimate the specular reflection coefficient in chromaticity space. Finally, the space transformation is used to calculate the specular components and the highlight removal is achieved by subtracting the specular reflection component from the original image. The experimental results show that our method outperforms most of existing methods on natural scene images, especially in some challenging scenarios with saturated pixels and complex textures.

Keywords: Specular highlight removal · Natural scenes · Reflection coefficient estimation

1 Introduction

Recovering a high quality image from a corrupted image [1–3] plays an essential role in the follow-up visual tasks [4,5]. In these techniques, specular highlight removal, which aims at separating the specular and diffuse components from the corrupted image, is one of the key problems. To address this problem, early methods [6,7] use additional hardware outside the camera. However, these approaches usually cause serious measurement distortion. Therefore, current research tends

to remove highlights directly from a single image using an image processing algorithm without the assistance of additional hardware. Typically, many works have been developed from the dichromatic reflection model that retains scene physical information for better results.

According to the physical properties of objects, the observed image intensity I(p) at pixel p can be represented by a dichromatic reflection model, and then can be decomposed into a diffuse reflection component D(p) representing object surface information and a specular reflection component S(p) representing light source information:

$$I(p) = D(p) + S(p) = m_d(p)\Lambda(p) + m_s(p)\Gamma(p), \tag{1}$$

where $m_d(p)$ and $m_s(p)$ denote the corresponding diffuse and specular reflection coefficients that represent the reflection ability of the object at position p. $\Lambda(p)$ and $\Gamma(p)$ denote the diffuse and specular chromaticity; the latter is usually regarded as the illumination chromaticity of ambient light.

Currently, many existing specular highlight remove algorithms [8–12] are derived from the artificial controlled experimental scenes, in which the background and specular highlight are usually unnatural. In contrast, the nature scenes are captured under a non-controlled environment, where objects are exposed to natural light. Therefore, the highlights usually emerge irregularly. As a result, the performance of the aforementioned methods on nature scenes dramatically decrease.

In this paper, in contrast to the aforementioned methods addressing artificial experimental scenes, we take natural scenes as our research object and propose a single natural scene image highlight removal method, focusing on the estimate of specular reflection coefficients based on the accurately estimated illumination chromaticity. The proposed method shows the different characteristics of highlight areas in natural scene images and experimental scene images often used in past methods. Our model provides the key smooth feature information, which is used in the specular reflection coefficient estimation in chromaticity space after normalization, and combines the intensity information with color information to avoid color distortion, thus achieving highlight removal with the dichromatic reflection model. The contributions of this paper are summarized as follows:

1. We propose a highlight removal method for natural scene image, which separates the specular reflection by fully considering the distribution characteristic of ambient light.
2. We explain the difference between the natural scene images and artificial experimental scene images often used in past methods, and propose using smooth feature information of nature scene images to estimate the specular reflection coefficient matrix.
3. The proposed model achieves very competitive highlight removal results in natural scenes. It greatly preserves the details and structural information of the original image in some challenging scenarios containing complex textures or saturated pixels.

2   Related Works

Specular highlight removal is a challenging problem in low level vision task. To tackle this problem, many methods have been proposed. Typically, Tan and Ikeuchi [13] proposed a method of iterating the specular component until only the diffuse component is left. They gave the concept of specular free (SF) images for the first time. The SF image contains only the diffuse component and its scene geometry is the same as the original image. By comparing pixel values of the highlight area and the adjacent regions, the maximum value of the pixel is converted to match the neighboring pixel values to achieve highlight removal. Assuming that the camera is highly sensitive, Tan and Ikeuchi [14] further discovered a method to separate the reflection components with a large number of linear equations. However, the color of objects is also affected by the material, roughness and texture. To address this problem, Shen *et al.* [15] proposed modified specular free (MSF) images. They assumed that there were only two types of pixels in MSF images, normal pixels and highlighted pixels, and then calculated the reflection components of two types of pixels using chrominance. The main disadvantage of the MSF image is that it suffers from hue-saturation ambiguity which exists in many natural images.
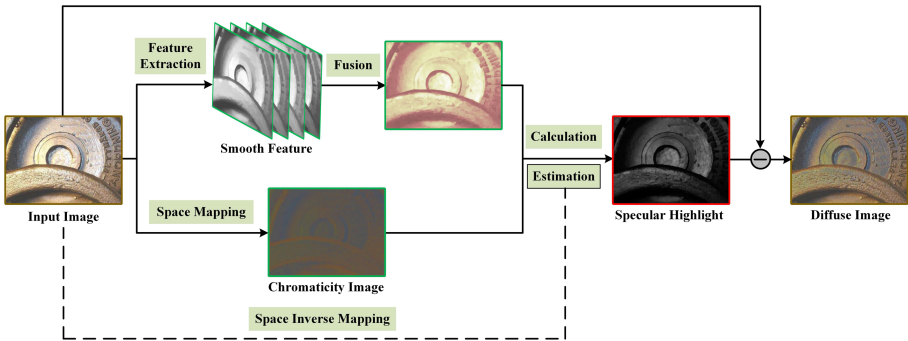

Fig. 1. Single natural scene image highlight removal pipeline.

Recently, the dichromatic reflection model, which fully considers the physical properties of the scene, has become the most widely used method. Klinker *et al.* [16] found that the diffusion and specularity showed T-shaped distribution in the RGB color space images and used them to remove highlights. However, it has been proven that this T-shaped distribution is susceptible to noise in practical applications and is prone to cause deformation in real images. Yang *et al.* [8,9] proposed a fast bilateral filter for estimating the maximum diffuse chrominance value in the local block of images, which caused the diffuse pixels to propagate to the specular pixels. Shen and Zheng [10] proposed the concept of pixel intensity ratio and constructed a pseudo-chrominance space to address the problem of texture surface, classifying pixels into clusters and robustly estimating the

intensity ratio of each cluster. Ren *et al.* [12] introduced a color-lines constraint into the dichromatic reflection model, but it is limited to objects with dark or impure color surfaces and accurate pixel clustering cannot be achieved. Jie *et al.* [17] transformed reflection separation into a solution of a low-rank and sparse matrix on the assumption that the highlight part is sparse in images. In [18], Fu *et al.* proposed a specular highlight removal method for real-world images, but it cannot achieve well result in large white specular highlight regions because this method is based on the assumption that the specular highlight is sparse in distribution and small in size.

More recently, some deep learning based methods have been proposed to tackle this problem. Wu *et al.* [19] built a dataset in a controlled lab environment and proposed a specular highlight removal network. Fu *et al.* [20] proposed a multi-task network that jointly performs specular highlight detection and removal. These methods have achieved great performance in the man-made datasets. However, deep learning based methods rely on the training data and cannot achieve general performance in the natural scene images whose distribution are different from man-made images.

3   Proposed Method

In this work, we propose a single image highlight removal method for natural scenes by estimating the specular reflection coefficients:

$$D = I - E(M(I), F(T(I_c)))\Gamma, \tag{2}$$

where I represents a highlight image, I_c is the input image from different channels of I, Γ is the specular chromaticity, and D is the output diffuse image. $T(\cdot)$ denotes the low-frequency information extraction based on image decomposition. $F(\cdot)$ is the feature fusion function of all channels. $M(\cdot)$ represents the space mapping from color space to chromaticity space. $E(\cdot)$ denotes the estimation function of specular reflection coefficients. An overview of the proposed method is shown in Fig. 1. First, we extract the smooth feature components of an image by $T(\cdot)$. Simultaneously, considering the effect of colors, the decomposition is implemented in the intensity and the three color channels. The information from the four channels is normalized and combined with $F(\cdot)$ to estimate the specular reflection coefficients by our derivation $E(\cdot)$. This process also uses the transformed input image, which is generated by transforming the input image into chromaticity space using $M(\cdot)$. Finally, highlight removal is achieved by subtracting the specular reflection component from the original image I.

3.1   Scene Illumination Evaluation

To illustrate the difference between natural scene images and experimental scene images with highlight areas, we focus on their light distribution characteristics.

According to Retinex theory [21], a digital image can be represented as the product of reflectance and illumination. The former represents the detailed information of an image, and the latter represents the ambience light. It can be simply written as:

$$I_c(x, y) = R_c(x, y) \cdot L_c(x, y), \tag{3}$$

where $I_c(x, y)$ is the pixel intensity value at (x, y) and c represents color channel R, G or B. $R_c(x, y)$ and $L_c(x, y)$ are the reflectance and illumination of channel c, respectively. Gaussian filters are usually used to calculate $L_c(x, y)$.

In Fig. 2, we present the illuminance histogram of $L_c(x, y)$ for the experimental and natural images. The illuminance histogram counts the number of pixels in each gray level. As shown in Fig. 2, the illumination distribution in the natural scene is brighter than that in the experimental scene. The ground truth illumination should be natural and provide the best visual quality for the panorama. Related study [22] has proofed that images with an average gray value of 128 are close to the best visual effects for humans. For the given images, the average values of experimental and natural scenes are 24 and 103, respectively. For the experimental scene, a large number of pixels in the captured images are in the dark interval, and their neighboring pixels are mostly dark. This distribution is usually nonexistent in natural lighting conditions.

3.2   Smooth Feature Extraction

We first start with the dichromatic reflection model presented in Eq. (1). We denote $I(p) = [I_r(p), I_g(p), I_b(p)]^T$ as the intensity value of the image pixel at $p = [x, y]$. $m_d(p)$ and $m_s(p)$ are diffuse and specular reflection coefficients, respectively, which are related to surface geometry. $\Lambda(p) = [\Lambda_r(p), \Lambda_g(p), \Lambda_b(p)]^T$ is the diffuse chromaticity, and its value usually remains the same in a continuous surface with the same color. Therefore, many studies estimate diffuse chromaticity based on the assumption that it is piecewise constant. However, the surface of the object is sometimes rough with irregularities in a nature scene, and the
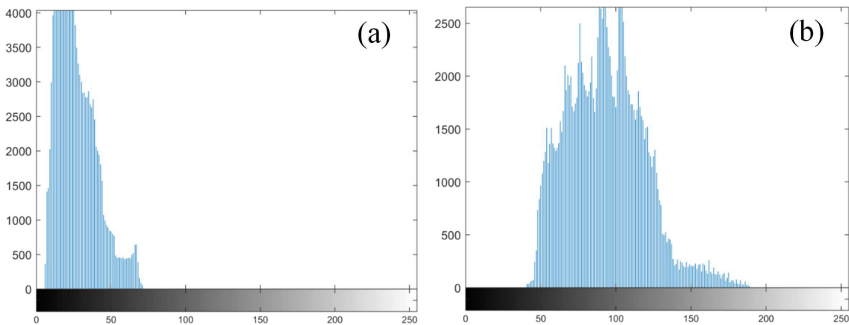
Fig. 2. Illuminance histograms of images from different surroundings. (a) presents the illuminance histogram of experimental scenes and (b) shows the illuminance histogram of natural scenes.

assumption is not suitable in all cases. By contrast, the illumination chromaticity $\Gamma(p) = [\Gamma_r(p), \Gamma_g(p), \Gamma_b(p)]^T$ can be estimated accurately via the color constancy algorithm in [23] or color-lines constraint in [12]. After normalization, the illumination is a white color, i.e., $\Gamma_r(p) = \Gamma_g(p) = \Gamma_b(p) = 1/3$. Thus, we focus on the estimate of the specular reflection coefficient $m_s(p)$.

In general, the observed image can be decomposed into two components: a low-frequency component called illumination and a high-frequency detail component called reflectance [24]. The former represents the ambient illumination and the later represents the details of objects [21]. In addition, in the dichromatic reflection model, the specular reflection component represents the information of ambient illumination and the diffuse reflection component represents the information of objects. Thus, the low-frequency information extracted from an image can truly reflect the intensity of the specular reflection at each pixel. The low-frequency information is the smooth feature. Moreover, the coefficient $m_s(p)$ precisely encodes the position and intensity of the specular reflection. For the low-frequency image at a certain position p, the larger the luminance value, the stronger the specular reflection, which accurately represents the intensity of the specular reflection. Therefore, we attempt to employ a low-frequency component to estimate the specular reflection coefficient $m_s(p)$. At the same time, we do not make any changes to the texture detail information.

To achieve accurate separation of the low- and high-frequency parts, we utilize the method proposed in [25]. The process of image robust sparse decomposition is shown in Fig. 3. The robust sparse representation improves the robustness to non-Gaussian noise. We first implement decomposition for the grayscale map. It can be written as follows:

$$I = W + H \quad \text{and} \quad W = NZ, \tag{4}$$

where I is an input image with highlight and N is the dictionary we construct, which mainly contains the extracted luminance information. Z is a sparse coefficient matrix in which very few element values are nonzero. NZ composes the low-frequency intensity information image W. H is the error matrix, containing edge and texture detail information. To construct a dictionary, we slide a small patch with a fixed step on the input image to obtain image blocks. Then, after vectorizing each block, all vectors form a vectorization matrix. Finally, the dictionary N is obtained by normalizing the vectorization matrix. We set the size of the patch to 1/10 of the image size through a large number of experiments. For an image of size 150×150, the patch size is set to 15×15. Thus, the sparse decomposition problem can be transformed into the following optimization problem with an equality constraint:

$$\min_{Z,H} \|Z\|_0 + \lambda \|H\|_{2,0} \quad \text{s.t. } I = NZ + H, \tag{5}$$

in which $\|\cdot\|_0$ denotes the l_0 norm of the matrix Z, which counts the number of nonzero entries in the matrix. $\|\cdot\|_{2,0}$ denotes the l_2 norm of the matrix H, which describes sample specific error terms and outliers. The parameter λ is used to
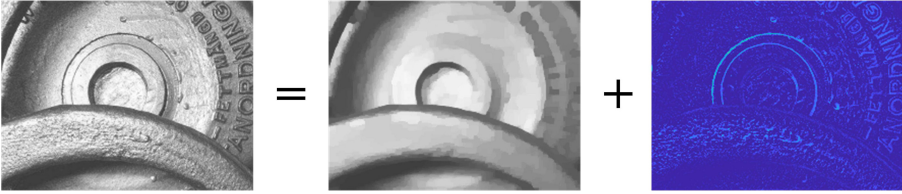
Fig. 3. The process of image robust sparse decomposition. The former denotes the low-frequency area, and the latter denotes the high-frequency area, which is rescaled for visualization.

balance the effect of different components. We can adjust the proportion of the two components by changing λ. While λ is larger, matrix H has less texture detail information, and more illuminance information is contained in matrix NZ. The appropriate value of λ is discussed in our experiment section.

However, the l_0 norm is highly nonconvex, and there is no efficient solution available. To make the optimization tractable, we relax it via replacing $\|\cdot\|_0$ with $\|\cdot\|_1$ and $\|\cdot\|_{2,0}$ with $\|\cdot\|_{2,1}$, and Eq. (5) can be formulated as the following convex optimization problem:

$$\min_{\mathrm{Z,H}} \ \|\mathrm{Z}\|_1 + \lambda\|\mathrm{H}\|_{2,1} \quad \text{s.t. } \mathrm{I} = \mathrm{NZ} + \mathrm{H}, \tag{6}$$

where $\|\cdot\|_1$ denotes the l_1 norm of the matrix, which can be calculated as follows:

$$\|\mathrm{Z}\|_1 = \sum_k \sum_j |\mathrm{Z}(j,k)|$$

$$\|\mathrm{H}\|_{2,1} = \sum_k \sqrt{\sum_j \mathrm{H}(j,k)^2}, \tag{7}$$

where (j,k) denotes the position of an element in matrix Z and H. Many efficient algorithms have been developed to solve this convex optimization problem, among which the linearized alternating direction method with the adaptive penalty (LADMAP) [26,27] is widely used. Then Z and H can be obtained, and W can be calculated by multiplying matrix N and Z.

3.3   Coefficient Estimation and Highlight Removal

We first decompose the grayscale map into the low-frequency part and the high-frequency part. However, the grayscale map contains only luminance information without any color information. It causes color distortion in the diffuse reflection part after separation when we use only the luminance to estimate the specular reflection coefficient $m_s(p)$. To preserve the original color of the input image, robust sparse decomposition is implemented in three color channels: R, G, and B. Then, the information of these channels is combined as the final low-frequency component. At the same time, we assign weights to each W of the three color channels to fully consider the contribution of different colors at each pixel, which

benefits from that our estimation of the reflection coefficient is based on global information. The weight can be calculated as:

$$\omega_c(p) = \frac{I_c(p)}{I_r(p) + I_g(p) + I_b(p)}, \tag{8}$$

where c represents color channel R or G or B. The final W can be written as follows:

$$W = \frac{W_U + \omega_R \cdot W_R + \omega_G \cdot W_G + \omega_B \cdot W_B}{4}. \tag{9}$$

where W_U, W_R, W_G, W_B denote the low-frequency information extracted in grayscale and three color channels, respectively. For the final low-frequency intensity image W at a certain position, the larger the value, the stronger the specular reflection, and the larger the value of $m_s(p)$. However, it is difficult to determine the specific interval of $m_s(p)$. Therefore, we use chromatic space to solve this problem. Chromaticity is usually defined as the function of component C(p):

$$C(p) = \frac{I(p)}{\sum_{c \in \{r,g,b\}} I_c(p)}. \tag{10}$$

Substituting (1) into (10), C(p) can be written as follows:

$$C(p) = \frac{m_d(p)}{\sum_{c \in \{r,g,b\}} I_c(p)} \Lambda(p) + \frac{m_s(p)}{\sum_{c \in \{r,g,b\}} I_c(p)} \Gamma(p). \tag{11}$$

Then, chromaticities $\Lambda(p)$ and $\Gamma(p)$ are normalized, i.e., $\sum_{c \in \{r,g,b\}} \Lambda_c(p) = 1$ and $\sum_{c \in \{r,g,b\}} \Gamma_c(p) = 1$. After that, combining with Eq. (1), the sum of the pixel intensities of the three channels can be expressed as:

$$\sum_{c \in \{r,g,b\}} I_c(p) = m_d(p) + m_s(p). \tag{12}$$

As a result, chromaticity can be written as the following:

$$C(p) = \frac{m_d(p)}{m_d(p) + m_s(p)} \Lambda(p) + \frac{m_s(p)}{m_d(p) + m_s(p)} \Gamma(p). \tag{13}$$

This process can be deemed as normalizing the reflection coefficient to the [0, 1] interval. $m_s(p)$ has strong relationship with the W. To simplify the estimation of $m_s(p)$, we directly use W to approximate $m_s(p)$. Moreover, to avoid over separation or incomplete separation of the specular reflection, the estimation is written as:

$$\frac{m_s(p)}{m_d(p) + m_s(p)} = \alpha W, \tag{14}$$

where α is an adjustable parameter, whose range is [0, 1]. In this work, we empirically set the value of α to 0.6, and present the detailed discussion in the experiment section. Thus, the specular reflection component S(p) can be obtained with Eq. (12), (13) and (14). The image after highlight removal can be obtained by subtracting the specular component from the original image.
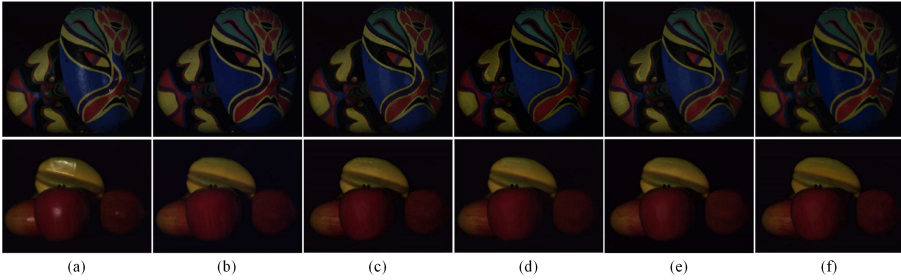
Fig. 4. Diffuse components of experimental scene images *Masks* and *Fruits*. (a) Input highlight images, (b) ground truths, (c) results of [10], (d) results of [12], (e) results of [17], (f) ours. (Color figure online)

Table 1. Quantitative comparison in terms of PSNR and SSIM for the images in Fig. 4

Image	Method							
	PSNR				SSIM			
	[10]	[12]	[17]	Ours	[10]	[12]	[17]	Ours
Masks	33.9	30.0	34.4	**34.5**	0.941	0.913	0.955	**0.968**
Fruits	**39.2**	37.5	36.4	37.1	0.960	0.952	0.930	**0.976**

**Table 2.** Quantitative comparison in terms of PSNR and SSIM for the images from SPEC database.

Image	Method							
	PSNR				SSIM			
	[9]	[10]	[12]	Ours	[9]	[10]	[12]	Ours
Woodlego	20.3	19.0	21.3	**25.9**	0.508	0.463	0.472	**0.601**
Vase	13.5	14.6	14.9	**26.0**	0.053	0.134	0.114	**0.550**
Wire	15.2	16.7	17.2	**24.2**	0.142	0.246	0.288	**0.701**
Key	14.5	12.8	12.9	**21.7**	0.169	0.176	0.175	**0.478**

4   Experiments

In this section, we evaluate the highlight removal performance of our method compared with currently effective methods. Following methods in [12,17], we first use some commonly used laboratory images to perform quantitative comparisons. Then, some typical natural images, which include some challenging scenarios with rough surfaces, saturated pixels, and complex textures, are used to perform visual comparisons. Note that, real-world natural images have no ground truth. Therefore, we do not provide the quantitative results. In addition, the related state-of-the-art methods [1,18] do not release their codes. Therefore, it is hard for us to provide the comparison results of these methods.

4.1   Quantitative Comparison on Laboratory Images

We first show the separation results of experimental scene images under a black background, which was often used in past methods. Figure 4 shows $Masks$ with pure and multicolored surfaces. The methods proposed in [10], [12], and [17] all create new artifacts while removing highlights on the yellow region of the mask on the left. The method in [12] makes the result darker due to over separation of the specular component. In contrast, our method removes highlights better in the yellow and blue regions that are closer to the ground truth. The captured surface of our result looks smoother and more continuous, which is why our method produces the highest PSNR and SSIM values as shown in Table 1. For the $Fruits$ image, our method does not achieve the best separation results but produces the highest SSIM, which demonstrates our competitiveness compared with other methods in that we retain the original structure of the image to the greatest extent.
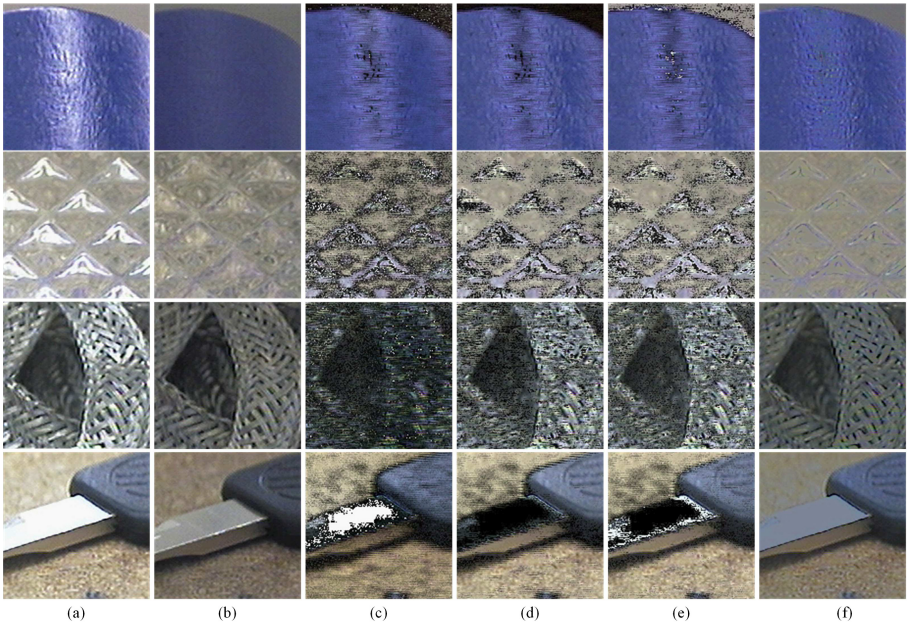
Fig. 5. Diffuse components of images close to natural scene $Woodlego$, $Vase$, $Wire$, and Key. (a) Input highlight images, (b) ground truths, (c) results of [8], (d) results of [10], (e) results of [12], (f) ours.

Moreover, we use the challenging images from the SPEC database to perform quantitative and qualitative comparisons. These images are taken under ambient light conditions created in the laboratory, which is closer to the natural scene but somewhat extreme. As shown in Fig. 5, the methods of $bilateral\ filtering$ [8]

and *intensity ratio* [10] introduce a large number of black and white noise points to the images during highlight removal, impairing the original image structure even in a single-color surface. The *color-lines constraint* method [12] tends to generate blur images because it is based on diffuse reflection pixel clustering. For scenes with unclear color or some metal surfaces, it is difficult to cluster pixels through only the chromaticity, and most of the image details are lost. In contrast, our method achieves good results in all these scenarios. Specifically, the proposed method can preserve the original structure and retain more edge details. For the severely overexposed scenes, the proposed method can still remove the specular highlight without introducing any artifacts and impairing image structure. For these challenging scenes, the proposed method produces the highest PSNR and SSIM values and is far better than other methods, as shown in Table 2.

### 4.2   Visual Effect Comparison on Natural Scene Images

Finally, we show the performance in natural scenes to further validate the superiority of our method. These images are taken under natural illumination from [12] and [17]. As ground truth results are unavailable, we provide a visual appearance comparison in Fig. 6. The method in [17] always achieves better highlight removal results than others. However, they are all inclined to degrade details and introduce additional artifact noise as shown in the red boxes. As illustrated in the first row of Fig. 6, the proposed method achieves more natural highlight removal, and the words on the *plate* are the clearest. For multicolored surfaces shown in the second row, our method produces pleasing visual result. The transition between highlight and nonhighlight regions is very natural, without the blurry edge that exists in other methods. In the third row, we not only remove highlights on the main object lock, but also on the background. The texture details in the background are optimally recovered through our method. Moreover, as shown in the fourth row, compared with other methods, the leaves recovered by the proposed method are more nature and colorful. In summary, the proposed method can produce better highlight removal results for nature scene images, which demonstrates its superiority.

4.3   Discussion of Important Parameters

The Effect of Balance Parameter λ. Parameter λ is used to balance the effect of low-frequency and high-frequency components. When λ is larger, less high-frequency information and more low-frequency information are separated. In Fig. 7(a), we observe the trend of PSNR when λ changes. Generally, PSNR shows an upward tendency with increasing λ, and there are few fluctuations between them. When $\lambda > 1.5$, the PSNR values are slightly downward. Intuitively, when λ is small, high-frequency is excessively separated, resulting in a lack of information used for estimation; the image after separation is partially blurred. However, when $\lambda > 1.5$, very few high-frequency parts are separated, and low-frequency estimation does not preserve texture details very well. The maximum PSNR
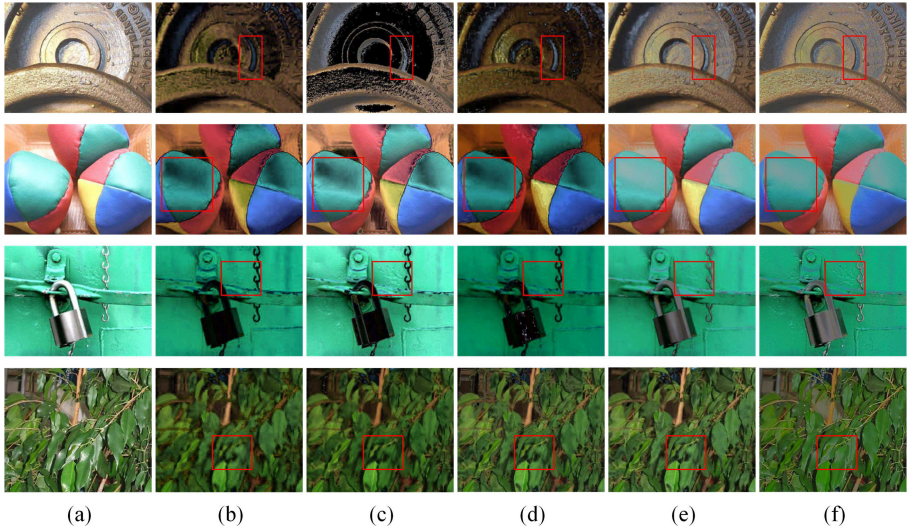
Fig. 6. Highlight removal results for natural scene image. (a) Input highlight images, (b) results of [8], (c) results of [10], (d) results of [12], (e) results of [17], (f) ours. (Color figure online)
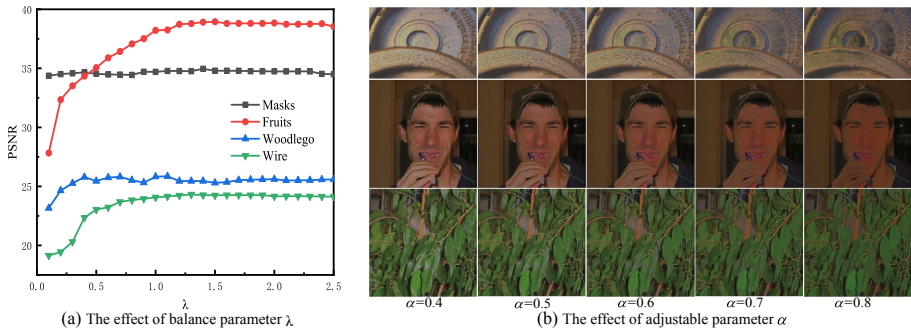
Fig. 7. Discussion of important parameters

values appear in $\lambda \in [1.1, 1.5]$. In this work, λ is set to 1.4 and this value works well for most of the highlight images.

The Effect of Adjustable Parameter α. Figure 7(b) shows the visual comparisons of natural images with different α. Larger α makes the image darker, while smaller α makes the specular reflection cannot be completely separated. When α is 0.6, the proposed method achieves relatively better visual result. Although $\alpha = 0.6$ may not be very accurate, it has little impact on the final result.

5   Conclusion

In this paper, we proposed an effective method for removing specular highlights focusing on captured natural scene images. The background and the distribution characteristics of ambient light are fully considered and the difference between the natural scene and the experimental scene is explained. We first constructed smooth feature images based on robust sparse decomposition. Then, we combined the smooth feature information with three color channels' information, and assigned different weights according to the contributions of different colors at each pixel to ensure that the color is not distorted. Finally, we converted the image into the chromaticity space, where the normalized smooth feature can be used as an accurate estimate of the specular reflection coefficient. Our method achieved very pleasing highlight removal results in natural scene images. It can preserve the original structure information and details of images. However, our method could not detect subtle bright spots or recover information that has been damaged by highlights, which will be addressed in future work.

References

1. Son, M., Lee, Y., Chang, H.S.: Toward specular removal from natural images based on statistical reflection models. IEEE Trans. Image Process. **29**, 4204–4218 (2020)
2. Chen, H., Jin, Y., Duan, M., Zhu, C., Chen, E.: DOF: a demand-oriented framework for image denoising. IEEE Trans. Industr. Inform. **17**(8), 5369–5379 (2021)
3. Chen, H., Jin, Y., Xu, K., Chen, Y., Zhu, C.: Multiframe-to-multiframe network for video denoising. IEEE Trans. Multimedia, 1–15 (2021)
4. Zhu, T., Xia, S., Bian, Z., Lu, C.: Highlight removal in facial images. In: Proceedings of the Chinese Conference on Pattern Recognition and Computer Vision (PRCV), pp. 422–433 (2020)
5. Chen, H., Jin, Y., Jin, G., Zhu, C., Chen, E.: Semisupervised semantic segmentation by improving prediction confidence. IEEE Trans. Neural Netw. Learn. Syst., 1–13 (2021)
6. Umeyama, S., Godin, G.: Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. IEEE Trans. Pattern Anal. Mach. Intell. **26**(5), 639–647 (2004)
7. Wang, F., Ainouz, S., Petitjean, C., Bensrhair, A.: Specularity removal: a global energy minimization approach based on polarization imaging. Comput. Vis. Image Underst. **158**, 31–39 (2017)
8. Yang, Q., Wang, S., Ahuja, N.: Real-time specular highlight removal using bilateral filtering. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6314, pp. 87–100. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15561-1_7
9. Yang, Q., Tang, J., Ahuja, N.: Efficient and robust specular highlight removal. IEEE Trans. Pattern Anal. Mach. Intell. **37**(6), 1304–1311 (2015)
10. Shen, H.-L., Zheng, Z.-H.: Real-time highlight removal using intensity ratio. Appl. Opt. **52**(19), 4483–4493 (2013)
11. Suo, J., An, D., Ji, X., Wang, H., Dai, Q.: Fast and high quality highlight removal from a single image. IEEE Trans. Image Process. **25**(11), 5441–5454 (2016)

12. Ren, W., Tian, J., Tang, Y.: Specular reflection separation with color-lines constraint. IEEE Trans. Image Process. **26**(5), 2327–2337 (2017)
13. Tan, R., Ikeuchi, K.: Separating reflection components of textured surfaces using a single image. IEEE Trans. Pattern Anal. Mach. Intell. **27**(2), 178–193 (2005)
14. Tan, R., Ikeuchi, K.: Reflection components decomposition of textured surfaces using linear basis functions. In: Proceedings of the IEEE Conference on Computer Visual Pattern Recognition (CVPR), vol. 1, pp. 125–131 (2005)
15. Shen, H.-L., Zhang, H.-G., Shao, S.-J., Xin, J.H.: Chromaticity-based separation of reflection components in a single image. Pattern Recognit. **41**(8), 2461–2469 (2008)
16. Klinker, G.J., Shafer, S.A., Kanade, T.: The measurement of highlights in color images. Int. J. Comput. Vis. **2**(1), 309–334 (1992)
17. Guo, J., Zhou, Z., Wang, L.: Single image highlight removal with a sparse and low-rank reflection model. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11208, pp. 282–298. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01225-0_17
18. Fu, G., Zhang, Q., Song, C., Lin, Q., Xiao, C.: Specular highlight removal for real-world images. Comput. Graph. Forum. **38**(7), 253–263 (2019)
19. Wu, Z., Zhuang, C., Shi, J., Xiao, J., Guo, J.: Deep specular highlight removal for single real-world image. In: SIGGRAPH Asia, pp. 1–2. ACM (2020)
20. Fu, G., Zhang, Q., Zhu, L., Li, P., Xiao, C.: A multi-task network for joint specular highlight detection and removal. In: Proceedings of the IEEE/CVF Conference on Computer Visual Pattern Recognition (CVPR), pp. 7752–7761 (2021)
21. Wang, S., Zheng, J., Hu, H.-M., Li, B.: Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Trans. Image Process. **22**(9), 3538–3548 (2013)
22. Yefeng, H., Bo, L., Jin, Z., Bai, Y.: An adaptive image enhancement based on the vector closed operations. In: Proceedings of the International Conference on Image and Graphics (ICIG), pp. 75–80 (2007)
23. Tan, T., Nishino, K., Ikeuchi, K.: Illumination chromaticity estimation using inverse-intensity chromaticity space. In: Proceedings of the IEEE Conference on Computer Visual Pattern Recognition (CVPR), vol. 1, pp. 673–680 (2003)
24. Cai, J., Gu, S., Zhang, L.: Learning a deep single image contrast enhancer from multi-exposure images. IEEE Trans. Image Process. **27**(4), 2049–2062 (2018)
25. Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 210–227 (2009)
26. Lin, Z., Liu, R., Su, Z.: Linearized alternating direction method with adaptive penalty for low-rank representation. In: Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), pp. 612–620 (2011)
27. Zhang, Y., Jiang, Z., Davis, L.S.: Learning structured low-rank representations for image classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 676–683 (2013)