Separation of specular and diffuse components using tensor voting in color images

Tam Nguyen, Quang Nhat Vo, Hyung-Jeong Yang, Soo-Hyung Kim, and Guee-Sang Lee*
Department of Electronics and Computer Engineering, Chonnam National University, Gwangju, South Korea
*Corresponding author: gslee@jnu.ac.kr

Received 2 July 2014; revised 22 September 2014; accepted 15 October 2014; posted 16 October 2014 (Doc. ID 215137); published 17 November 2014

Most methods for the detection and removal of specular reflections suffer from nonuniform highlight regions and/or nonconverged artifacts induced by discontinuities in the surface colors, especially when dealing with highly textured, multicolored images. In this paper, a novel noniterative and predefined constraint-free method based on tensor voting is proposed to detect and remove the highlight components of a single color image. The distribution of diffuse and specular pixels in the original image is determined using tensors' saliency analysis, instead of comparing color information among neighbor pixels. The achieved diffuse reflectance distribution is used to remove specularity components. The proposed method is evaluated quantitatively and qualitatively over a dataset of highly textured, multicolor images. The experimental results show that our result outperforms other state-of-the-art techniques. © 2014 Optical Society of America

OCIS codes: (100.3020) Image reconstruction-restoration; (100.2960) Image analysis; (330.1690) Color; (120.5700) Reflection.

http://dx.doi.org/10.1364/AO.53.007924

1. Introduction
Specular reflection detection and removal are important problems that affect image perception and analysis. By resolving these issues, a surface reflectance model and other intrinsic characteristics, such as illumination, shading, and surface geometry, can be achieved [1]. Such knowledge would be useful for all topics related to image processing and can be used to address three-dimensional (3D) reconstruction problems.

Many computer vision tasks, such as segmentation, pattern recognition, and tracking, have difficulties if specular reflections exist in the input images. Almost all of these tasks consider specular regions to be outliers and accept the effects of the specular component in the results of the analysis. Therefore, robustness and accuracy are not ensured. Meanwhile, the appearance of specular reflections is inevitable, and, in fact, the information obtained in these regions is important.

Many studies relating to the detection and removal of specular reflection have been proposed [2]. These studies have adopted different techniques to overcome the problems of separating diffuse and specular components from an input image. These methods can either use a single image or multiple images as input [2].

When dealing with multiple images, the configuration of the input images can be different, such as a different point of view of the same scene [3], multi-baseline stereo [4], a chain of images or videos of texture scenes [5], and different polarization angles [6]. These methods are restrictive in that having many input images with different specific conditions creates problems with the hardware and software setup.

The approach using a single image is even more challenging to design properly. There are two types of single-image techniques: color space analysis...
and neighborhood analysis [2]. There are different methods to use color space analysis to separate the specular component from a single input image [7,8]. In more detail, Klinker [7] mapped the image color space to the dichromatic reflection model and extracted the highlight positions from the color information [7]. A color reflection model was also presented which was a combination of the dichromatic model for dielectric materials and a 3D spectral space constructed from three orthogonal basic functions, called S space. In this space, each color measurement is represented by a 3D vector. From a set of basic functions, the first axis in the S space can be determined such that it best approximates a flat spectrum in the least-squares sense in the visible range of light. Then two other axes orthogonal to the first are determined through Gram–Schmidt orthogonalization [8]. These approaches are not robust for use with highly textured surfaces and need support from users.

The former group can overcome the limitations of the latter because the method is based on the relationship between the adjacency pixels for separating specular positions without segmentation. There have been several studies concerning the specular-free (SF) image that is a pseudo-diffuse component [9–11,12]. The SF image is determined by obtaining the specularity-invariant value for each pixel by subtracting the minimum RGB value from each channel [11]. This method is based on the local ratios between the adjacency pixels that are used to detect specular reflection. The SF image discussed by Shen and co-workers [9,10] is similar to that presented by Yoon et al. [11] but differs in the specular removal step. The authors introduced a modified SF image achieved by adding an offset. Shen et al. [9] used a constant scalar value and later [10] considered the scalar value as a threshold to modify only the pixels that have a minimum value larger than that of the threshold. Tan et al. [12] calculated the SF image by shifting the intensity and chromaticity of a pixel nonlinearly while retaining its hue. The intensity of the logarithmic differentiation on the SF image and the normalized image were applied in an iterative framework to remove specular components.

These approaches may encounter difficulties when a discontinuity appears in the surface color and the input images contain highly textured surfaces. Moreover, the original object color that is not preserved can cause problems in cases where it is necessary to recover color. A decomposition technique based on a mean-shift decomposition, an eigen-decomposition, can be used to separate two reflection components in the textured surfaces [13]. Then, specular components are shifted to the diffuse regions. In this method, the shifting process can cause the incomplete removal of the specular component.

Yang et al. [14] proposed a method which is similar to Tan et al.’s except in the way of estimating the maximum diffuse chromaticity using a bilateral filter. The speed improved for real-time applications.

Shen proposed a real-time method [15] similar to Yang’s using the intensity ratios between the maximum and range values of the pixels. Both of the two real-time algorithms mentioned above focused on speed and still face up to highly textured surfaces.

In this paper, we introduce a novel, effective method for specular reflection detection and removal in a single input image using tensor voting. By using this technique, our method can obtain the diffuse reflection distribution and dominant color of diffuse regions in the original image by analyzing the information of the eigenvalues from tensors, regardless of noise. Unlike previous methods, tensor voting is an effective and accurate method to find the dominant colors in color images [16]. During the voting process, tensors are voted by neighbors, and they change both in orientation and magnitude. This procedure forms tensor clusters which correspond to color regions in the original image. These clusters as well as their distribution are clearly visible in the tensor space. And in each cluster, tensors which have the largest magnitude represent the structures and shapes of the clusters. The color value corresponding to this tensor is considered as the dominant color. In addition, this is a noniterative approach. Therefore, our proposed method does not need to be concerned with convergence problems. Besides, specular components are removed using information from the diffuse ones directly, so the tensor voting method could overcome the limitations and offer an improvement over current methods.

In summary, the contributions of our proposed method are as follows:

- The data in a single input image is analyzed to obtain the diffuse reflectance distribution using tensor voting. The voting process creates tensors with different saliency through the eigenvalues. Noise and specular pixels in small regions, denoted by small tensors, are isolated and removed using cosine similarity and saturation information. The diffuse tensors with significant saliency present a diffuse reflectance distribution.

- Diffuse pixels are determined from the original image based on the diffuse reflection distribution. This distribution includes tensor groups. These groups illustrate the diffuse color regions in the original image. By finding the dominant color in each group, diffuse pixels are extracted. The remaining part is considered as a specular candidate.

- A specular reflection removal process is formulated using the information of the diffuse pixels and reflection model. For each nondiffuse pixel, the specular component is removed based on the values of diffuse pixels around it. Therefore, the color of the diffuse image which is obtained after specular removal as well as the smooth characteristic is maintained in the original image.

We evaluate our proposed algorithm with some standard and some challenging images that contain multicolor, highly textured surfaces. We perform a
comparision of the proposed method with two well-known techniques proposed by Shen et al. [9] and Tan et al. [12].

The rest of this paper is organized as follows. Section 2 presents the concept of our proposed method. The experimental results and discussion are provided in Section 3, and Section 4 is the conclusion of this paper.

2. Proposed Method

We present an overview of our proposed method for specular reflection detection and removal using tensor voting (Fig. 1).

Given an original image, we first encode pixels into second-order, symmetric, nonnegative definite tensors in color space. The voting process is implemented among these tensors. This process creates a saliency map which presents the color data structure in the original image. Then, by analysis of this map, we extract saliency areas corresponding to diffuse color regions in the original image, which we call the diffuse reflection distribution. With this distribution, diffuse and specular candidates are classified. Finally, we remove the specular component from specular candidates using the reflection model’s principle and the diffuse candidates’ information.

A. Tensor Voting Process

1. Encoding

To minimize the effect of lighting in the highlight detection and removal process, we convert original image into $L*a*b*$ color space [17]. In this space, the $L*$-component is the lightness of each color vector in the image. The $a*$-component presents the color from green to red, and the $b*$-component presents the color from blue to yellow. Since the $L*$-component contains all of the information of the luminance level without any data of the real color, only the $a*$-component and $b*$-component are used as a feature vector of each pixel in the image to encode into the tensor.

Tensors are mathematical entities that are introduced to extend the notion of scalars, vectors, and matrices [16]. In this study, we mention second-order, symmetric, nonnegative definite tensors which can encode saliency and orientation at the same time. In general, they can be represented geometrically as ellipses in $2-D$. The shapes of these ellipses indicate the orientation, and their sizes specify saliency.

Let us denote $T$ as an arbitrary second-order, symmetric, nonnegative definite tensor with eigenvectors $e_1$ and $e_2$ and corresponding eigenvalues $\lambda_1$ and $\lambda_2$ in a decreasing order. The tensor $T$ is represented in Eq. (1), which in turn can be rewritten as Eq. (2). This tensor $T$ can be decomposed into stick components and ball components by using Eq. (3), as shown in Fig. 2(a):

$$T = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \hat{e}_1^T \\ \hat{e}_2^T \end{bmatrix},$$

$$T = \lambda_1 \hat{e}_1 \hat{e}_1^T + \lambda_2 \hat{e}_2 \hat{e}_2^T,$$

$$T = (\lambda_1 - \lambda_2) \hat{e}_1 \hat{e}_1^T + \lambda_2 (\hat{e}_1 \hat{e}_1^T + \hat{e}_2 \hat{e}_2^T).$$
Initially, pixels in the input image are isotropic. Therefore, in our method, ball tensors are selected to encode these pixels. They are special cases of general tensors when eigenvalues are equal together.

2. Tensor Voting

After the encoding process is complete, we achieve a tensor space in which each tensor corresponds to a feature vector in the image. The voting process is implemented in this space. The tensors are gradually deformed as a result of the accumulation of votes cast from other neighboring tensors in a definite range, \( \sigma \), by a ball voting field, as seen in Fig. 3. \( \sigma \) presents the region where the voter can cast a vote onto the votee.

Figure 2(b) illustrates the tensor voting process from voter \( O \) to votee \( P \). The region from which \( P \) can receive vote information from \( O \) depends on the inscribed circle with the center \( C \) determined by the value of sigma with

\[
c = \frac{-16 \log(0.1) \times (\sigma - 1)}{\pi^2},
\]

where \( \sigma \) is the factor used to measure the degree of smoothness. In Eq. (4), if the value of \( \sigma \) is small, the value of \( c \) will be small. This means that the number of voting neighbors is small, and the votes are few and inverse. This value can provide control over the ability to remove noise from the image.

The information that \( P \) catches from \( O \) is calculated by

\[
DK(s, k, \sigma) = e^{-\left(\frac{s^2 + k^2}{\sigma^2}\right)},
\]

where \( DK \) is the magnitude of the vote, \( s = \frac{\cos \theta}{\sin \theta} \) is the arc length, \( k = \frac{2 \sin \theta}{l} \) is the curvature, and \( \theta \) is the angle in Fig. 2(b).

![Fig. 3. (a) Original image; (b) saliency map; (c) angles between peaks in the saliency map; (d), (e) diffuse and specular pixels in the input image and their corresponding saturation values, respectively; (f) saliency map without noise and highlight tensors (diffuse reflectance distribution); (g) diffuse (white color) and specular (black color) candidates image.](image-url)
B. Saliency Information Analysis

As mentioned in Section 1, each tensor is represented by eigenvectors \((e_1, e_2)\) and eigenvalues \((\lambda_1, \lambda_2)\) which control the orientation and saliency of the tensor, respectively. Therefore, the saliency map can be achieved as the total of all eigenvalues \((\lambda_1 + \lambda_2)\). Figure 3(b) shows the 3D saliency map of the input image in Fig. 3(a). Two axes on the horizontal plane illustrate the color space \((256 \times 256)\). The vertical axis shows the value of \(\lambda_1 + \lambda_2\). Small pink circles describe local maxima of the saliency map.

We release that saliency of tensors in the color space groups into areas. Each area and local maximum presents a corresponding color region and its centroid, respectively, in the input image. Moreover, in the voting process, the saliency of tensors increases proportionally to the number of neighbor tensors inside voting range. Therefore, the bigger the local maximum is in the saliency map, the larger the color region is in the input image.

Understanding this rule, we see that noise and highlight regions in the input image correspond to saliency areas in the saliency map with local maxima being small. By removing small saliency tensors, we could achieve diffuse tensors in the saliency map which we call the diffuse reflection distribution.

In this paper, the cosine similarity and saturation information are used to remove highlight tensors based on two characteristics of diffuse and specular pixels:

- Similar colors have parallel orientations even when degraded with uneven lighting or highlight.
- For the specular-only or monochromatic pixels, their saturations are identical to zero.

Using the extreme algorithm, all peaks in the saliency map are collected. The peaks with cosine similarity values within a threshold, which is an empirical value of 0.95 in this approach, are grouped together pairwise. Now, the peak with smaller saturation and saliency values is considered as indicating the highlight and the other one as indicating the diffuse component. In this paper, the cosine similarity and the saturation values are computed according to \[8\] and \[18\], respectively. The tensors belonging to specular peaks are also removed. The saliency map with only diffuse candidate tensors is called the diffuse reflection distribution.

Figure 3 illustrates the steps for achieving the diffuse reflection distribution. From an input image [Fig. 3(a)], the saliency map [Fig. 3(b)] is extracted after the tensor voting process with small pink circles being peaks of regions in this map. The angles between peaks are shown in Fig. 3(c). As can be seen, there are three very small angles which are created by peaks of specular pixels and diffuse pixels in the same color regions in the input image (yellow, red, and green color regions, respectively). Figure 3(d) shows the diffuse reflection distribution of the input image after removing highlight peaks. A comparison of diffuse pixels and specular pixels in hue and saturation is shown in Fig. 3(d) and Fig. 3(e). The specular pixels (A and C) have a smaller saturation value than corresponding diffuse pixels (B and D).

C. Diffuse and Specular Candidates Determination

Diffuse and specular candidates in the input image could be discovered based on the local maxima (peaks) in the diffuse reflection distribution. These local maxima are the color centroids of the diffuse regions in the input image, called dominant colors.

Let \(S\) be the set of dominant color \(sS = \{S_u\}\), with \(u = [1, n], n: \) number of dominant colors. For each pixel \(x\) in the \(L \ast a \ast b \ast \) image, we calculate the dissimilarity to all dominant colors using the Euclidean distance:

\[
d_{x,S_u} = \sqrt{(S_{u,a} - x_{i,a})^2 + (S_{u,b} - x_{i,b})^2}. \tag{6}
\]

where \(S_{u,a}\) and \(S_{u,b}\) are the a-component and b-component colors of \(S_u\), respectively. \(x_{i,a}\) and \(x_{i,b}\) are the a-component and b-component colors of \(x_i\), respectively.

Let \(D\) be the distance matrix from pixels in the image to all dominant colors. We have

\[
D = \begin{bmatrix}
d_{x_1,S_1} & d_{x_1,S_2} & \cdots & d_{x_1,S_n} \\
\vdots & \ddots & \ddots & \vdots \\
& \cdots & \ddots & d_{x_1,S_n} \\
& & \cdots & d_{x_k,S_n}
\end{bmatrix}, \tag{7}
\]

where \(k\) is the number of pixels in the input image.

While the dissimilarity of the specular pixels to the dominant colors is very high, that of the diffuse pixels is low. Therefore, the diffuse pixels corresponding to each dominant color are determined by \(D_{\text{diffuse,S}} = \min(\text{col}_u(D))\).

To determine the diffuse and specular candidates for each dominant color, first, we calculate the average value of the minima of the RGB components of the color vectors of diffuse pixels achieved above. Second, excepting diffuse pixels, the rest of the pixels are assigned to corresponding regions of dominant colors by finding the minimum value of these pixels’ dissimilarities (min(\(\text{col}_u(D^T)\))).

In a pixel, if the minimum of RGB components is smaller than the average value, it is considered as a diffuse pixel. Otherwise, it is a specular pixel. Figure 3(g) shows the separation of diffuse and specular candidates of a train image [in Fig. 3(a)]. The white and black pixels illustrate the diffuse and specular candidate respectively as the result of tensor voting.

D. Specular Reflection Removal

1. Reflection Model

In this paper, we assume that the light source is white, the specular pixels are not saturated, and the color camera behaves in a linear manner, which means that the RGB response values are proportional to the intensity of the light entering the sensor.
Besides, the dichromatic reflection model is used to
describe the reflection of most inhomogeneous materi-
als [19]. The model assumes that the light reflected
from an object is a linear combination of diffuse and
specular colors, which can be presented as

\[ J = J^D + J^S, \]  

(8)

where \( J \) is the reflected light color captured by an
RGB camera and \( J^D \) and \( J^S \) are the diffuse and
specular components, respectively.

We assume that the RGB response values are pro-
portional to the intensity of the light incident on the
sensor. Let \( J_c(x) \) be the color of channel \( c \) \( (c = 1, 2, 3, \)
or equivalently, red, green, blue) at pixel \( x \), and \( r(\lambda, x) \)
is the spectral reflectance of wavelength \( \lambda \) at pixel \( x \).
Then let \( l(\lambda) \) be the spectral power distribution of the
illumination, and \( s_c(\lambda) \) be the spectral sensitivity of
the channel \( c \). The component color value at \( x \) can be
formulated as

\[ J_c(x) = \int r(\lambda, x)l(\lambda)s_c(\lambda)d\lambda. \]  

(9)

According to the reflection model, the spectral re-
fectance can be decomposed into two independent com-
ponents:

\[ r(\lambda, x) = a(x)r_b(\lambda) + \beta(x)r_s(\lambda), \]  

(10)

where \( r_b(\lambda) \) and \( r_s(\lambda) \) are the wavelength com-
position of the diffuse and the specular reflectances,
respectively, and \( a(x) \) and \( \beta(x) \) denote the geometrical
factors of these two reflections at pixel \( x \), respectively.
According to Shen et al. [9], the specular reflection
component is similar to that of the illumination,
where \( r_s(\lambda) \) is independent of wavelength \( \lambda \). Equa-
tion (10) then becomes

\[ r(\lambda, x) = a(x)r_b(\lambda) + \beta(x)r_s(\lambda). \]  

(11)

Substituting the value of \( r(\lambda, x) \) from Eq. (11) into
Eq. (9), we have

\[ J_c(x) = a(x)r_b(\lambda)l(\lambda)s_c(\lambda)d\lambda + \beta(x)rs \int l(\lambda)s_c(\lambda)d\lambda \]
\[ = a(x)J_{b,c} + \beta(x)J_{s,c}, \]  

(12)

where \( J_{b,c} = \int r_b(\lambda)l(\lambda)s_c(\lambda)d\lambda \) denote the intrinsic
body color of the material, and \( J_{s,c} = \int r_s(\lambda)s_c(\lambda)d\lambda \)
is the illumination color. The illumination color can be
obtained by imaging a white object surface.
Then the color of each pixel is normalized with re-
spect to the illuminant color and is rescaled to a
range from 0 to 255 [11]. Therefore, the surface color
becomes pure white, \( J_{s,c} = 255 \) for each channel.
Equation (12) can then be written as

\[ J_c(x) = a(x)J_{b,c} + 255\beta(x). \]  

(13)

Compared to Eq. (8), we have \( J = J_c(x), J^D =
a(x)J_{b,c} \) and \( J^S = 255\beta(x) \).

2. Specular Reflection Removal Process

As mentioned in Section 2.A, each pixel \( x \) is a linear
combination of the diffuse and specular reflection.
According to Shen et al. [9], a SF image was intro-
duced to remove the specular component. This image
could be obtained simply by subtracting the mini-
um of the RGB components of the color \( I(x) \), and
all specular components are eliminated while the
geometry information is reserved in the SF image.
However, the color of the SF image is always darker
than that of the original image because at least one
element of each pixel’s color is 0. Therefore, a scalar
value is added into the SF image to obtain a result
which is close to that of the original image.

This means that the specular reflection removal pro-
cess was applied on the entire image, including
both the diffuse and specular pixels. The color of the
diffuse pixels could change unnecessarily. Moreover,
by adding a scalar value into the SF image, the
specular regions could not be removed when there
were highly textured surfaces and multicolor images.
Therefore, in this paper we consider only the specu-
lar pixels for the removal process instead of the
entire image.

We let \( I_{s,c}(x) \) be the color vector of the specular pixel
\( x \) assigned to a region having corresponding domi-
nant color \( S_w \). The color \( I_{s,c}(x) \) of the corresponding
SF pixel is calculated by

\[ I_{s,c,u}(x) = I_u(x) - \min(I_r(x), I_g(x), I_b(x)) \]
\[ = I_u(x) - I_{\min}(x). \]  

(14)

where \( I_r(x), I_g(x), I_b(x) \) are the three color values of
red, green, and blue of pixel \( x \):

\[ I_{\min}(x) = \min(I_r(x), I_g(x), I_b(x)) \]
\[ = a(x) \min(I_{b,r}, I_{b,g}, I_{b,b}) + 255\beta(x) \]
\[ = a(x)I_{b,\min} + 255\beta(x). \]  

(15)

By combining Eqs. (14) and (15), we have

\[ I_{s,c,u}(x) = a(x)(I_{b,u} - I_{b,\min}). \]  

(16)

Let \( I_d, c(x) \) be the actual diffuse reflection of specu-
lar pixel \( x \) in the original image. We then have

\[ I_{d,u}(x) = I_u(x) - 255\beta(x) \]
\[ = I_{s,c,u}(x) + I_{\min}(x) - 255\beta(x) \]
\[ = I_{s,c,u}(x) + a(x)I_{b,\min}(x) + 255\beta(x) - 255\beta(x) \]
\[ = I_{s,c,u}(x) + a(x)I_{b,\min}(x). \]  

(17)

where \( I_{b,\min}(x) \) is the diffuse amount of the minimum
color component of the pixel \( x \).
Since the specular candidates in the region having corresponding dominant color $S_u$ and the diffuse pixels in this region have the same body color (the same color on the same object), the diffuse component of specular pixels is similar to the diffuse component of diffuse pixels. Therefore, the diffuse amount of the minimum color component of specular pixels may be considered to be the average diffuse amount of the minimum color component of diffuse pixels, $I_{tb,b}\min$:

$$I_{b,\text{min}} \approx I_{tb,b,\text{min}}. \quad (18)$$

Let $I_{tb,\text{min}}$ be the average value of the minimum RGB component of all diffuse candidates within this region. We have

$$I_{tb,\text{min}} = \frac{\sum I_{\text{min},i}}{n}, \quad i \in n, \quad (19)$$

where $n$ and $i$ are the number of diffuse pixels and the index of these pixels, respectively. $I_{\text{min},i}$ is the minimum of the RGB components of the $i$th diffuse pixel. Based on the reflection model, we have

$$I_{tb,\text{min}} = \alpha(x)I_{tb,b,\text{min}} + 255\beta(x). \quad (20)$$

Since the diffuse pixels contain only the diffuse component, the specular component equals zero $[255\beta(x) = 0]$. Equation (20) can be rewritten as

$$I_{tb,\text{min}} = \alpha(x)I_{tb,b,\text{min}} \approx \alpha(x)I_{b,\text{min}}. \quad (21)$$

In addition, to maintain the 3D structure of objects after removing highlight, we give a factor called the natural coefficient (coef). This factor controls the percentage of $I_{tb,\text{min}}$ to be added to SF pixels. If specular pixels stay in the center of highlight regions, the specular component is very high. Therefore, the natural coefficient is approximately equal to 1. Its value reduces to 0 when specular pixels are far from the center of highlight regions according to the following equation:

Fig. 4. Separation results of the toys image. (a) Original image; (b), (c) reflection components by the proposed method; (d), (e) reflection components by Tan et al.’s method [12]; (f), (g) reflection components by Shen et al.’s method [9].
\[
\text{coef} = \frac{I_{b,\text{min}} - I_{tb,\text{min}}}{255 - I_{tb,\text{min}}} .
\]

From Eqs. (17), (21), and (22) the diffuse component from specular pixels can be calculated as

\[
I_{d,u}(x) = I_{sf,u}(x) + \text{coef}I_{tb,\text{min}} .
\]

3. Experimental Results and Analysis

A. Dataset

For our method, we evaluate the performance by undertaking experiments on all images which were used in previous studies [9,12]. Additionally, we also apply the experiment on multicolor, multiobject images with highly textured surfaces captured on camera by ourselves (25 images). In these images, the specular component is located in the area between the two color regions and is widely distributed. In practice, we set up the parameter as follows: sigma \( \sigma = 20 \) for all investigated cases. This value is good enough to ensure the smoothness degree. Shen et al.’s method [9] and Tan et al.’s method [12] are the two methods from previous studies that are used here to compare against our proposed method. The source code of both is freely available on the authors’ websites (Shen et al.: http://www.ivlab.org/publications.html; Tan et al.: http://php-robbytan.rhcloud.com/code.html).

We use qualitative visual perception and the quantitative mean squared error (MSE) to compare the proposed method with existing methods:

\[
\text{MSE} = \frac{1}{k} \sum_{i=1}^{k} (\hat{X}_i - X_i)^2 .
\]

where \( \hat{X}_i, X_i \) are the specular removed image from our proposed method and the ground truth image, respectively. \( k \) is the number of pixels in the image. The human visual system is sensitive to color changes at edges, and MSE measures image differences on a pixel-wise basis. Therefore, two evaluation

Fig. 5. Separation results of the fish image. (a) Original image; (b), (c) reflection components obtained by the proposed method; (d), (e) reflection components obtained by Tan et al.’s method [12]; (f), (g) reflection components obtained by Shen et al.’s method [9].
methods becomes complementary to each other in image evaluation.

B. Experiments on Natural Scene Images

Figure 4 shows the diffuse and specular component separation results of three methods for a standard image in the specular reflection removal field. This is the textured image of toys in which some objects have multiple colors. In general, all of three methods get sufficiently good results for this case. But in detail, in Shen’s method, the color of the specular component removed regions in the diffuse image slightly differs from the surroundings, especially on the right thigh and arm of dinosaur. This issue is caused by a constant value being added to all pixel values in the free specular image [9]. About the color information, the result in Tan et al.’s method is significantly darker than the original image. Meanwhile, by subtracting each channel at each pixel from its minimum RGB value before adding a constant value, the color in Shen et al.’s method seem fade lightly. Our method gets the diffuse image with the color which seems nearest to the original image.

A toy fish image which has various colors and wide specular regions is illustrated in Figure 5. This is a well-known image which is used in most previous researches in the specular removal field. The result of our method is close to Tan et al.’s and Shen et al.’s methods with respect to diffuse component separation.

Figure 6 includes another case where it is difficult to execute the specular component removal task. The input image is a multicolored train image that contains strong and wide specular regions. In particular, the specular components on the right wheel are very difficult to separate. Both Shen et al.’s and Tan et al.’s methods are not very good in this case. The specular regions in the resulting image appear with different colors in Shen et al.’s method (both of the specular positions near the left and right wheel). In the right wheel, there exists a scratch according to the orientation of the specular region.

Fig. 6. Separation results of the train image. (a) Original image; (b), (c) reflection components by using the proposed method; (d), (e) reflection components by using Tan et al.’s method [12]; (f), (g) reflection components by using Shen et al.’s method [9].
The specular components in some positions, which are marked in the red rectangle, still have not been removed. With Tan et al.’s method, the resulting image still appears to have a scratch on the right wheel. The tail of the train in the image disappears in Tan et al.’s diffuse image. Our method obtains better, smoother results. Moreover, the color appearance of the diffuse component of our proposed method approximates that of the original image more closely while Tan et al.’s [12] is darker and Shen et al.’s [9] is lighter than the original image. The specular component removal process is applied on only specular pixels as a result of the diffuse reflectance distribution information which was obtained after tensor voting. Therefore, the specular component image in our proposed method includes specular regions instead of redundant portions, as do Tan et al.’s and Shen et al.’s methods.

Table 1. Global and Regional MSE Values of Reflection Separation Methods, with the Lowest Global and Regional MSEs in Bold

<table>
<thead>
<tr>
<th>Image</th>
<th>Global MSE</th>
<th>Regional MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tan</td>
<td>Shen</td>
</tr>
<tr>
<td>Animals</td>
<td>65</td>
<td>24</td>
</tr>
<tr>
<td>Cups</td>
<td>61</td>
<td>11</td>
</tr>
<tr>
<td>Masks</td>
<td>180</td>
<td>40</td>
</tr>
<tr>
<td>Circle</td>
<td>3124</td>
<td>43</td>
</tr>
</tbody>
</table>

The global and regional MSE values, respectively, are shown in parentheses.

Fig. 7. Separated diffuse reflection components of two images named animal and cups (number 1 and number 2) using three methods. The global and regional MSE values, respectively, are shown in parentheses.
The first image in the Fig. 7 shows the result of diffuse reflections of rabbits (Fig. 7.1). As can be seen, the highlight on the rabbit’s ears is quite difficult to remove. Visually, Tan et al.’s method produces the diffuse component darkly. Meanwhile, Shen et al.’s method and the proposed method get good results. But in term of global and regional MSE, the proposed method performs best (the lowest MSE values). The second image includes two cups containing girl and monkey drawings. Tan et al.’s result creates two regions on the girl’s face, and the color of the diffuse image in Tan et al.’s method still darker than the ground truth image. Although the global MSE of Shen et al.’s method is equal to the proposed method, the regional MSE still a little higher than the proposed method.

Figure 8 shows two very difficult cases to detect and remove highlight regions. The first image (Fig. 8.3) contains heavy textures. Tan et al.’s method removes much specular reflection and thus produces quite high MSE (180 for global MSE and 177 for regional MSE). There is a difference in color between highlight regions and objects in Shen et al.’s diffuse image. The color of highlight regions after removing the specular component is quite faded compared to the real color of the objects. The proposed method

![Fig. 8](image-url) Separated diffuse reflection components of two images named mask and circle (numers 3 and number 4, respectively) using three methods. The global and regional MSE values, respectively, are shown in parentheses.
gets the best result quantitatively and qualitatively. The rest is an image with a circle including four quadrants with difference colors. The main issue here is that the highlight region is across two quadrants. This region is strong and wide. Shen et al.’s method is not very good, as can be seen. The diffuse image appears very rough around the specular region, and the color in the output image seems to fade slightly. With respect to Tan et al.’s method, the specular region appears similar in color to the original image, but all the rest have changed color, with the diffuse image being very dark. Both global and regional MSE in Tan et al.’s result are much too high. The color of the diffuse image is quite different from the ground truth. In this case, the proposed method still produces acceptable results with the lowest MSE in both global and regional MSE.

Table 1 summarizes the global and regional MSE values of reflection separation methods for all images from Fig. 7 and Fig. 8. The lowest global and regional MSEs are given in bold. As can be seen, for the proposed method, the MSE values are lowest in all cases.

Figure 9 compares the results of our proposed method across different images. The first row is a complex surface with different textures and colors which contains some plastic objects. The wide highlight region on the yellow cup and the strong specularity on the blue box are a challenge for the specular component removal task. In the results, all highlight regions are separated and eliminated successfully. This is evidence of the efficacy of our proposed method. Another good result is shown in the second row. The strong and weak highlight regions of the two green apples are satisfactorily decomposed.

In addition, the running times are mentioned in this paper to compare to the other two methods of Tan et al. and Shen et al. In Tan et al.’s method, the computation is quite heavy as the loop of specularity reduction ends only when the maximum chromaticities of all pixels in a single-colored surface are the same. With Shen et al.’s method, an iterative framework continues until all the nondiffuse pixels are dealt with. In comparison, although our proposed method does not need any iterative loop, we trade speed to obtain accuracy using tensor voting. Therefore, the running times of our proposed method are relatively slower than Shen et al.’s method but still faster than Tan et al.’s method. Table 2 show the running time of nine well-known images with different resolution. Three methods are run on a laptop five-core CPU, 2.4 GHz, 8192 MB of memory.

### Table 2. Comparison of Running Times

<table>
<thead>
<tr>
<th>Image</th>
<th>Resolution</th>
<th>Tan's Method</th>
<th>Shen's Method</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toys</td>
<td>353 x 387</td>
<td>19.715</td>
<td>0.581</td>
<td>16.980</td>
</tr>
<tr>
<td>Fish</td>
<td>642 x 480</td>
<td>29.954</td>
<td>1.782</td>
<td>10.256</td>
</tr>
<tr>
<td>Train</td>
<td>683 x 640</td>
<td>54.823</td>
<td>0.924</td>
<td>11.752</td>
</tr>
<tr>
<td>Animal</td>
<td>396 x 321</td>
<td>8.291</td>
<td>0.322</td>
<td>3.345</td>
</tr>
<tr>
<td>Cups</td>
<td>640 x 480</td>
<td>25.160</td>
<td>1.764</td>
<td>10.155</td>
</tr>
<tr>
<td>Circle</td>
<td>334 x 335</td>
<td>5.964</td>
<td>0.139</td>
<td>4.751</td>
</tr>
<tr>
<td>Helmet</td>
<td>632 x 466</td>
<td>54.27</td>
<td>1.34</td>
<td>10.896</td>
</tr>
<tr>
<td>Apple</td>
<td>633 x 465</td>
<td>31.45</td>
<td>0.415</td>
<td>6.412</td>
</tr>
<tr>
<td>Mask</td>
<td>500 x 450</td>
<td>18.992</td>
<td>1.017</td>
<td>9.957</td>
</tr>
</tbody>
</table>

4. Conclusions

In this paper, we introduced a novel and effective method to detect and remove the specular component in a single image. By performing an analysis of the data with tensor voting, the diffuse reflectance distribution can be obtained after removing noise, and specular tensors are denoted as those belonging to a small saliency. With a noniterative algorithm, nonpredefined constraints, and without being based on the local relationship between neighboring pixels, our proposed method can overcome the limitations of existing algorithms, especially with multicolored, highly textured input images. The experimental
results showed that the performance of the proposed method is promising. In following studies, we want to improve our method to implement it as a real-time application.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2014-024950 and 2013-006535) and the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NUPA-2014-H0301-14-1014) supervised by the NIPA (National IT Industry Promotion Agency).

References


7936 APPLIED OPTICS / Vol. 53, No. 33 / 20 November 2014