

Illumination estimation from specular highlight in a multi-spectral image

Dongsheng An,^{1,2} Jinli Suo,¹
Haoqian Wang,^{2,3} and Qionghai Dai^{1,2*}

¹Department of Automation, Tsinghua University, Beijing 100084, China

²Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China

³haoqianwang@sz.tsinghua.edu.cn

^{*}qh dai@tsinghua.edu.cn

Abstract: The reflection spectrum of an object characterizes its surface material, but for non-Lambertian scenes, the recorded spectrum often deviates owing to specular contamination. To compensate for this deviation, the illumination spectrum is required, and it can be estimated from specularity. However, existing illumination-estimation methods often degenerate in challenging cases, especially when only weak specularity exists. By adopting the dichromatic reflection model, which formulates a specular-influenced image as a linear combination of diffuse and specular components, this paper explores two individual priors and one mutual prior upon these two components: (i) The chromaticity of a specular component is identical over all the pixels. (ii) The diffuse component of a specular-contaminated pixel can be reconstructed using its specular-free counterpart describing the same material. (iii) The spectrum of illumination usually has low correlation with that of diffuse reflection. A general optimization framework is proposed to estimate the illumination spectrum from the specular component robustly and accurately. The results of both simulation and real experiments demonstrate the robustness and accuracy of our method.

© 2015 Optical Society of America

OCIS codes: (150.2950) Illumination; (330.6180) Spectral discrimination; (120.5700) Reflection; (150.0150) Machine vision.

References and links

1. C. P. Huynh and A. Robles-Kelly, "A solution of the dichromatic model for multi-spectral photometric invariance," *Int. J. Comput. Vision* **90**(1), 1–27 (2010).
2. A. Gijsenij, T. Gevers, and J. van De Weijer, "Computational color constancy: Survey and experiments," *IEEE Trans. Image Process.* **20**(9), 2475–2489 (2011).
3. T. Zickler, S. P. Mallick, D. J. Kriegman, and P. N. Belhumeur, "Color subspaces as photometric invariants," *Int. J. Comput. Vision* **79**(1), 13–30 (2008).
4. J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts, "Color invariance," *IEEE Trans. Pattern Anal.* **23**(12), 1338–1350 (2001).
5. A. Artusi, F. Banterle, and D. Chetverikov, "A survey of specular removal methods," *Comput. Graph. Forum* **30**(8), 2208–2230 (2011).
6. P. Koirala, P. Pant, M. Hauta-Kasari, and J. Parkkinen, "Highlight detection and removal from spectral image," *J. Opt. Soc. Am. A* **28**(11), 2284–2291 (2011).
7. Q. Yang, S. Wang, and S. Ahuja, "Real-time specular highlight removal using bilateral filtering," In *Proceedings of European Conference on Computer Vision* (Springer, 2010), pp. 87–100.

8. R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Am. A* **21**(3), 321–334 (2004).
9. K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data," *IEEE Trans. Image Process.* **11**(9), 972–984 (2002).
10. K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy Algorithms. II. Experiments with image data," *IEEE Trans. Image Process.* **11**(9), 985–996 (2002).
11. J. V. D. Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. Image Process.* **16**(9), 2207–2214 (2007).
12. L. Shi and B. Funt, "MaxRGB reconsidered," *J. Imaging Sci. Technol.* **56**(2), 20501-1–20501-10(10) (2012).
13. M. P. Lucassen, T. Gevers, A. Gijsenij, and N. Dekker, "Effects of chromatic image statistics on illumination induced color differences," *J. Opt. Soc. Am. A* **30**(9), 1871–1884 (2013).
14. D. Cheng, D. K. Prasad, and M. S. Brown, "Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution," *J. Opt. Soc. Am. A* **31**(5), 1049–1058 (2014).
15. K. Barnard, "Improvements to gamut mapping colour constancy algorithms," in *Proceedings of European Conference on Computer Vision* (Springer, 2000), pp. 390–403.
16. D. A. Forsyth, "A novel algorithm for color constancy," *Int. J. Comput. Vision* **5**(1), 5–35 (1990).
17. L. Shi, W. Xiong, and B. Funt, "Illumination estimation via thin-plate spline interpolation," *J. Opt. Soc. Am. A* **28**(5), 940–948 (2011).
18. P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," in *Proceedings of International Conference on Computer Vision and Pattern Recognition* (IEEE, 2008), pp. 1–8.
19. S. M. Newhall, R. W. Burnham, and R. M. Evans, "Color constancy in shadows," *J. Opt. Soc. Am. A* **48**(12), 976–984 (1958).
20. R. Kawakami, J. Takamatsu, and K. Ikeuchi, "Color constancy from black body illumination," *J. Opt. Soc. Am. A* **24**(7), 1886–1893 (2007).
21. M. S. Drew and B. V. Funt, "Variational approach to interreflection in color images," *J. Opt. Soc. Am. A* **9**(8), 1255–1265 (1992).
22. S. A. Shafer, "Using color to separate reflection components," *Color Res. Appl.* **10**(4), 210–218 (1985).
23. G. D. Finlayson and G. Schaefer, "Convex and non-convex illuminant constraints for dichromatic colour constancy," in *Proceedings of International Conference on Computer Vision and Pattern Recognition* (IEEE, 2001), pp. 598–604.
24. H. C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *J. Opt. Soc. Am. A* **3**(10), 1694–1699 (1986).
25. T. M. Lehmann and C. Palm, "Color line search for illuminant estimation in real-world scenes," *J. Opt. Soc. Am. A* **18**(11), 2679–2691 (2001).
26. L. Shi and B. Funt, "Dichromatic illumination estimation via Hough transforms in 3D" in *European Conference on Colour in Graphics, Imaging, and Vision* (IS&T, 2008), pp. 259–262.
27. J. Toro and B. Funt, "A multilinear constraint on dichromatic planes for illumination estimation," *IEEE Trans. Image Process.* **16**(1), 92–97 (2007).
28. J. Toro, "Dichromatic illumination estimation without pre-segmentation," *Pattern Recogn. Lett.* **29**(7), 871–877 (2008).
29. K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal.* **33**(12), 2341–2353 (2011).
30. M. S. Drew, H. R. V. Joze, and G. D. Finlayson, "Specularity, the zeta-image, and information-theoretic illuminant estimation," in *Proceedings of European Conference on Computer Vision Workshops and Demonstrations* (Springer, 2012), pp. 411–420.
31. M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Proces.* **54**(11), 4311–4322 (2006).
32. Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," in *Technical Report UILU-ENG-09-2215* (University of Illinois Urbana-Champaign, 2009).
33. R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," in *CS Technical Report* (Technion–Israel Institute of Technology, 2008).
34. J. Suo, L. Bian, F. Chen, and Q. Dai, "Bispectral coding: compressive and high-quality acquisition of fluorescence and reflectance," *Opt. Express* **22**(2), 1697–1712 (2014).
35. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Multispectral Image Database," <http://www.cs.columbia.edu/CAVE/databases/multispectral/>.
36. J. Jun and J. Gu, "Recovering spectral reflectance under commonly available lighting conditions," in *Proceedings of International Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, 2012), pp. 1–8.

1. Introduction

A multi-spectral (or hyperspectral) image characterizes the material of a scene surface. The accurate measurement of spectra is very helpful for the studies of color constancy and photometric invariance [1–4]. However, the recorded spectrum of a non-Lambertian surface is influenced by specularities and deviates from its true spectrum; e.g., the highlight regions of the glossy color plates in Fig. 1(a). Many computer-vision tasks, such as spectrum clustering (i.e., grouping the pixels according to their chromaticity) demonstrated in Fig. 1(b), would fail in such non-Lambertian scenes. To address this problem, researchers have put much effort into recovering the diffuse reflectance [5] before spectrum clustering in the recovered specular-free image. Most of these highlight-removal approaches are implemented with a known illumination spectrum, such as in [6] and [7]. The performance of spectrum clustering on the data recovered from inaccurate illumination are significantly degenerated, as shown in Fig. 1(c), compared to that from accurate illumination shown in Fig. 1(d). Therefore, the accuracy of illumination estimation is crucial for identifying the final spectrum.

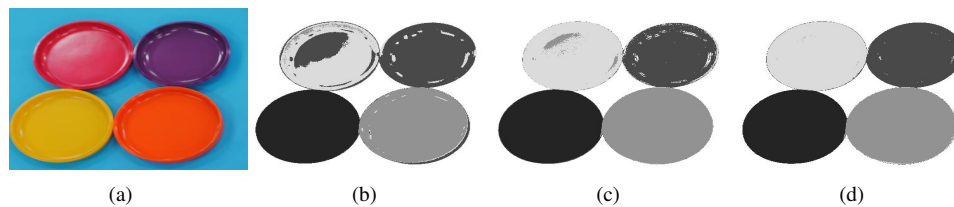


Fig. 1. An example showing the influence of illumination estimation accuracy on spectrum clustering. (a) Multi-spectral image visualized by integrating with the RGB response curves of Canon 20D. (b), (c), and (d), respectively, illustrate the spectrum clustering results of the original image, the recovered specular-free image with inaccurate illumination (estimated by [8]) and the recovered specular-free image with correct illumination (calibrated with color checker). The specular-free images in (c) and (d) are obtained using the method proposed in [7].

The estimation of illumination spectra has been extensively studied, but it is still an open problem. As reviewed in [2, 9, 10], there is extensive literature on this topic, including statistics-based methods [11–14], gamut-based methods [15, 16], and learning-based methods [17, 18]. In addition, there are methods to estimate illumination based on various physical properties, such as shadows [19], black-body radiation [20], and inter-reflections [21]. However, all these methods are limited to Lambertian surfaces, and the estimation performance would degenerate in specular-influenced images.

Physically, the reflectance of non-Lambertian materials can be described using the dichromatic reflection model [22], which assumes that the surface reflectance is a linear combination of diffuse and specular components. For scenes containing non-Lambertian reflections, the illumination spectrum can be estimated from specularities itself; this method has been demonstrated to be of high accuracy and has attracted wide attention [23–30].

According to the dichromatic model, the reflected radiance of pixels belong to the same material falls on a hyperplane spanned by the reflectance spectrum and illumination spectrum of the material. Assuming identical illumination chromaticity over a scene, Finlayson et al. [23] calculated the illumination spectrum as an intersection of dichromatic hyperplanes defined by different materials. In CIE XYZ color space, intersecting hyperplanes become intersecting dichromatic lines [24, 25]. These methods need to distinguish surface colors beneath specular highlight and are inapplicable to cases in which only one material is affected by specularities. In addition, it is difficult to apply these methods to multi-channel images because the

number of required material types increases linearly with the number of spectral channels. To avoid explicit material discrimination, one can also build a linear model of the illumination spectrum according to the statistics and fit the model parameters from the specular pixels. By adopting such a strategy, Tan et al. [8] and Shi&Funt [26] used the Hough transform to estimate the parameter, and Toro et al. [27, 28] exhaustively searched for a plausible solution from an illumination database. Although specularity can be detected using the dark channel prior of nature images [29], a sufficient number of specular pixels and a large strength range are required for robust linear modeling. Some other researchers estimated illumination by utilizing the consistency of illumination chromaticity (i.e., normalized spectrum) and the diversity of diffuse reflection spectra over the image. Huynh and Robles-Kelly [1] applied singular value decomposition (SVD) to the spectrum of highlight pixels and assumed the first component to be the illumination vector. They obtained satisfying performance only on monochromatic patches. Drew et al. [30] treated the geometric mean of pure specular pixels or near white materials as the estimated illumination. This assumption performs well in cases with strong specularity or white objects, but it does not hold in cases with only one or two non-white materials contaminated by weak specularity. In summary, owing to the complexity of material composition of natural non-Lambertian scenes and the variety of specularity strength, thus far, there has been no illumination-estimation method robust enough for such diverse cases.

In this study, we develop a generally applicable and robust illumination-estimation method. By adopting the dichromatic reflection model [22], a general illumination-estimation approach utilizing different priors is proposed. Specifically, we formulate the illumination estimate as a signal-separation problem. Though this problem is obviously ill-posed, we can solve it by introducing two individual priors and one mutual prior:

1. The chromaticity of specular component is consistent over all the pixels. Therefore, we can mathematically formulate this component to be a low-rank matrix.
2. One can match pixels with the same diffuse reflection spectrum but different specularity strengths. If we use the specular-free counterpart to reconstruct the spectrum of a specular pixel, the residue depends only on illumination. Thus, we separate the input image into two regions: a specular-free region and a specular region. A dictionary of the scene materials can be learned from the specular-free region, while the diffuse component of the specular region can be reconstructed upon this dictionary with sparse coefficients. It is worth noting that this prior also holds for the pixels between stronger and weaker specular regions. In other words, we do not need the accurate detection of specular regions, but rather a discrimination between stronger and weaker specular levels.
3. We notice that the correlation between material chromaticity and illumination is relatively low. Here, a low-correlation prior is defined to penalize the solutions deviating towards the material chromaticity. Moreover, the low-correlation prior is scene dependent; thus, we propose to assign a scene-adaptive weighting factor to it.

By extensively exploring and modeling these priors, an optimization-based algorithm is derived for robust illumination estimation. Experiments on varying scenes, illuminations, and specularity strengths are conducted, and the results show that our approach can provide consistently superior performance compared to state-of-the-art methods.

The remainder of this paper is organized as follows. We present the adopted dichromatic reflection model and the problem formulation in Sec. 2. Then, the solution of our optimization problem is introduced in Sec. 3. Sec. 4 presents the experiments on various scenes and the superiority of our method over state-of-the-art methods. Finally, we conclude our paper in Sec. 5 with extensive discussions.

2. Formulation

We adopt the widely used dichromatic reflection model [22] to describe the reflection properties of non-Lambertian surfaces. By assuming identical illumination chromaticity over the entire scene, the multi-spectral image \mathbf{I} taken by a camera can be formulated as

$$\mathbf{I}_c(\mathbf{x}) = \omega_r(\mathbf{x}) \int_{\Lambda} r(\mathbf{x}, \lambda) e(\lambda) q_c(\lambda) d\lambda + \omega_l(\mathbf{x}) \int_{\Lambda} e(\lambda) q_c(\lambda) d\lambda. \quad (1)$$

Here, $\mathbf{I}_c(\mathbf{x})$ is the intensity of channel c at pixel \mathbf{x} , with c indexing the camera channel and $\mathbf{x} = \{x, y\}$ representing the 2D location. The two terms on the right-hand side of the equation are, respectively, the diffuse and specular components, with $\omega_r(\mathbf{x})$ and $\omega_l(\mathbf{x})$ denoting the corresponding strength factors. In each term, λ denotes the wavelength with its range being Λ , while $r(\cdot, \lambda)$, $e(\lambda)$, and $q_c(\lambda)$ represent the surface reflectance, the illumination spectrum, and the camera's spectral response for channel c , respectively. For convenience, the above equation can be simplified as

$$\mathbf{I}(\mathbf{x}) = \mathbf{I}_d(\mathbf{x}) + \mathbf{I}_s(\mathbf{x}), \quad (2)$$

where $\mathbf{I}(\mathbf{x})$ is a vector describing the spectrum at pixel location \mathbf{x} , the c th entry of which is $\mathbf{I}_c(\mathbf{x})$ in Eq. (1); $\mathbf{I}_d(\mathbf{x})$ is the diffuse component; and $\mathbf{I}_s(\mathbf{x})$ is the specular component. Equation (2) can be rewritten in the matrix notation as follows:

$$\mathbf{I} = \mathbf{I}_d + \mathbf{I}_s, \quad (3)$$

with each column of \mathbf{I} , \mathbf{I}_d , and \mathbf{I}_s describing the spectrum at a certain position. Thus far, illumination estimation entails estimating \mathbf{I}_s from a given specular-contaminated image \mathbf{I} . Obviously, this is ill-posed, and we need to introduce priors to solve it.

2.1. Individual prior: low-rank specular highlight

According to Eq. (1), in addition to the spectral response of the camera sensor that can be calibrated beforehand, the specular component \mathbf{I}_s only reflects the property of illumination, while the diffuse component \mathbf{I}_d is affected by both the illumination and the surface reflectance. Therefore, one can first decompose the multi-spectral image of a non-Lambertian scene into diffuse and specular components and then use the latter to estimate the illumination. In the present study, we assume that there is only one light source in the scene, as shown in Fig. 2(a), or multiple light sources with the same normalized spectrum, as shown in Fig. 2(b). In such cases, the specular components of different pixels can form a low-rank matrix \mathbf{I}_s , the rank of which is ideally 1. We can utilize this low-rank property of \mathbf{I}_s to help estimate the illumination. When a scene is illuminated by several light sources with different normalized spectra, such as the scene shown in Fig. 2(c), the rank of matrix \mathbf{I}_s is higher than 1. The illumination estimation of such scenes is quite challenging and beyond the scope of this paper.

2.2. Individual prior: sparse diffuse reflectance

In the case of specular-influenced color pixels that are not white or grey, it is reasonable to suppose that their counterparts with the same material exist in specular-free regions. Supposing that the specular regions Ω_l can be correctly separated from non-specular regions Ω_r , we can use the pixels in Ω_r to build a dictionary \mathbf{D} to represent the materials in the scene and then force high-quality reconstructions of $\{\mathbf{I}_d(\mathbf{x}), \mathbf{x} \in \Omega_l\}$ with sparse coefficients (ideally only one non-zero entry) upon \mathbf{D} . From Eq. (3), for pixels in the region Ω_l , we have

$$\mathbf{I} = \mathbf{D}\mathbf{C} + \mathbf{I}_s. \quad (4)$$

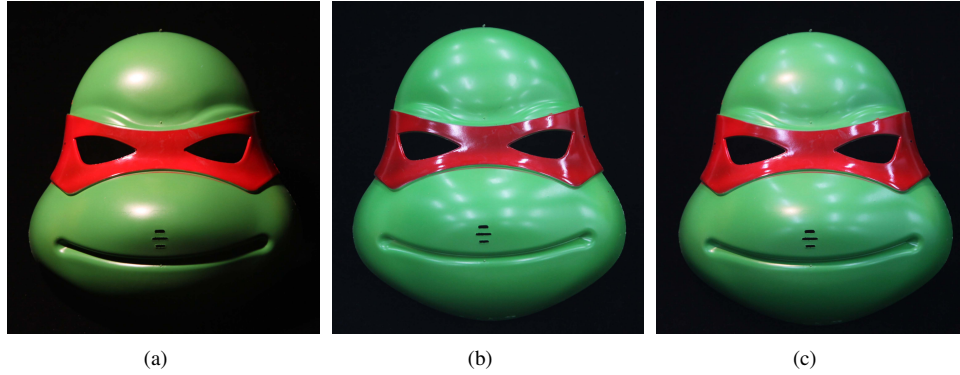


Fig. 2. A glossy mask illuminated by three different light sources: (a) a single incandescent lamp, (b) a set of fluorescent tubes with the same normalized spectrum, and (c) an incandescent lamp and multiple fluorescent tubes.

In this equation, $\mathbf{I} \in R^{k \times n}$ represents the specular-influenced pixels, with k being the number of channels and n being the number of pixels in the region Ω_l ; $\mathbf{D} \in R^{k \times p}$ is the dictionary learned from the region Ω_r and each column of \mathbf{D} describes the chromaticity of a specific material in the scene; p denotes the number of material types in the scene; $\mathbf{C} \in R^{p \times n}$ is a matrix denoting the reconstruction coefficients of the diffuse component, in which only one nonzero entry exists in each column, i.e., each pixel in the region Ω_l corresponds to a specific material in \mathbf{D} ; $\mathbf{I}_s \in R^{k \times n}$ represents the specular component.

The accurate separation of regions with and without specularity is nontrivial. Fortunately, we do not need an exact separation because a pixel with strong specularity can be reconstructed using the spectrum describing the same material but with weaker specular contaminations, in which the residue is still material-independent, i.e., reflects only the illumination spectrum. Therefore, the separation of pixels with and without specularity entails discriminating pixels with strong specularity from those with weak/no specularity.

Statistically, the spectrum of a purely diffuse reflection (neither white nor bright grey) is most likely to have at least one channel with extremely low intensity, i.e., the dark-channel image [29], but this does not hold for pixels affected by specular highlight. Hence, the bright pixels in the dark-channel image tend to contain specularity, as shown in the upper subfigure of Fig. 3(a). The dark-channel image can be calculated as follows:

$$\mathbf{I}_{dark}(\mathbf{x}) = \min_c(\mathbf{I}_c(\mathbf{x})). \quad (5)$$

where $\mathbf{I}_{dark}(\mathbf{x})$ is the strength of the dark-channel image \mathbf{I}_{dark} at location \mathbf{x} , and $\min_c(\mathbf{I}_c(\mathbf{x}))$ represents the lowest intensity within the spectrum at location \mathbf{x} in the multi-spectral image \mathbf{I} .

We first compute the dark-channel image and apply thresholding to decompose the dark-channel image into two regions: Ω_l with strong specularity and Ω_r with weak or no specularity, as shown in the lower subfigure of Fig. 3(a). In this study, we choose Ω_l as the pixel set with the top 5% intensities of the nonzero elements in the dark-channel image. According to our model, the final estimation is insensitive to this percentage, and one can choose from a large range of values. After the decomposition into regions, we use the K-SVD algorithm proposed by Aharon et al. [31] to learn an over-complete dictionary from the pixels in Ω_r . The dictionary entries for the image in Fig. 1(a) are plotted in Fig. 3(b) with color-coded lines.

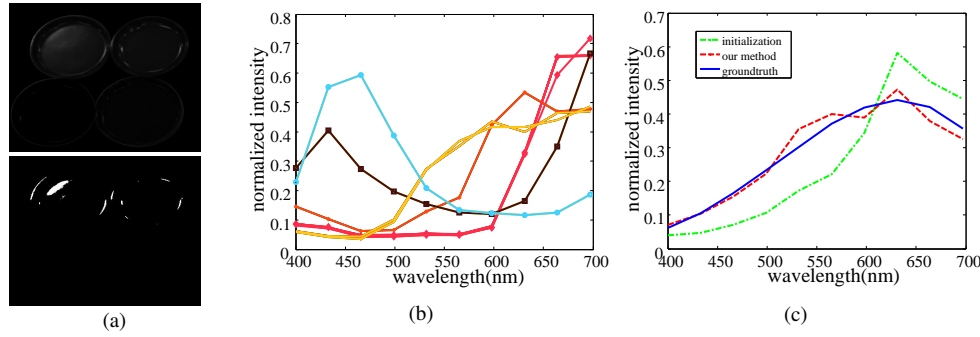


Fig. 3. Demonstration of our illumination-estimation method on the image in Fig. 1(a). (a) Top: dark-channel image. Bottom: a binary image denoting pixel grouping. The white region Ω_l includes the pixels used in illumination optimization, and the dark region Ω_r is composed of the pixels used to learn the over-complete dictionary. (b) The learned dictionary. We integrate the chromaticity of each base with the RGB response curves of Canon 20D to obtain the corresponding line color. (c) Our initial and final estimation of the illumination spectrum in comparison with the ground truth. The definition of chromaticity in (b) and (c) is described in Eq. (6).

2.3. Mutual prior: low correlation between diffuse reflectance and illumination

The above prior is insufficient in some challenging cases, such as scenes with only one material contaminated by specular highlight. In Fig. 4(a), we render an image consisting of a glossy sphere illuminated by a point light source (top) and display its diffuse (middle) and specular components (bottom).

For a clear explanation, we define the chromaticity of the diffuse component $\mathbf{I}_d(\mathbf{x})$ and specular component $\mathbf{I}_s(\mathbf{x})$ in terms of the l_2 norm as

$$\tilde{\mathbf{I}}_d(\mathbf{x}) = \frac{\mathbf{I}_d(\mathbf{x})}{\|\mathbf{I}_d(\mathbf{x})\|_2} \quad \text{and} \quad \tilde{\mathbf{I}}_s(\mathbf{x}) = \frac{\mathbf{I}_s(\mathbf{x})}{\|\mathbf{I}_s(\mathbf{x})\|_2}. \quad (6)$$

The former equation describes the inherent surface material and the latter indicates the illumination spectrum. Because we assume that the illumination chromaticity $\tilde{\mathbf{I}}_s(\mathbf{x})$ is uniform in the entire scene (i.e., location independent), $\tilde{\mathbf{I}}_s(\mathbf{x})$ can be simplified as $\tilde{\mathbf{I}}_s$ and Eq. (2) can be rewritten as

$$\mathbf{I}(\mathbf{x}) = r(\mathbf{x})\tilde{\mathbf{I}}_d(\mathbf{x}) + l(\mathbf{x})\tilde{\mathbf{I}}_s, \quad (7)$$

where the coefficients $r(\mathbf{x})$ and $l(\mathbf{x})$ denote the corresponding strength factors and are location dependent.

Because all pixels are of the same diffuse chromaticity $\tilde{\mathbf{I}}_d$ and specular chromaticity $\tilde{\mathbf{I}}_s$ in Fig. 4(a), the entire spectra in this scene lie on the hyperplane spanned by $\tilde{\mathbf{I}}_d$ and $\tilde{\mathbf{I}}_s$ in Fig. 4(b). Supposing that point A has the strongest specularity, the nonnegativity constraint on the decomposition coefficients enforces all the spectrums to lie within the yellow region. Correspondingly, the estimated illumination spectrum will lie outside this region. Taking location \mathbf{x}_B as an example, the sparseness constraint (i.e., minimizing $r(\mathbf{x}_B)$) tends to give a trivial solution $r(\mathbf{x}_B) = 0$, as illustrated by the blue arrow (#1). Based on the parallelogram law, the sparsity prior will bias the estimated illumination towards $\mathbf{I}(\mathbf{x}_B)$ (but will not surpass $\mathbf{I}(\mathbf{x}_A)$ owing to the nonnegativity of the coefficients), as visualized by the red arrow (#2). To address this problem, we introduce a new constraint $\|\mathbf{D}'\mathbf{I}_s\|_2^2$ (where \mathbf{D}' is the transpose of \mathbf{D}) called the low-correlation prior based on the observation that the diffuse reflectance and illumination usually have low correlation.

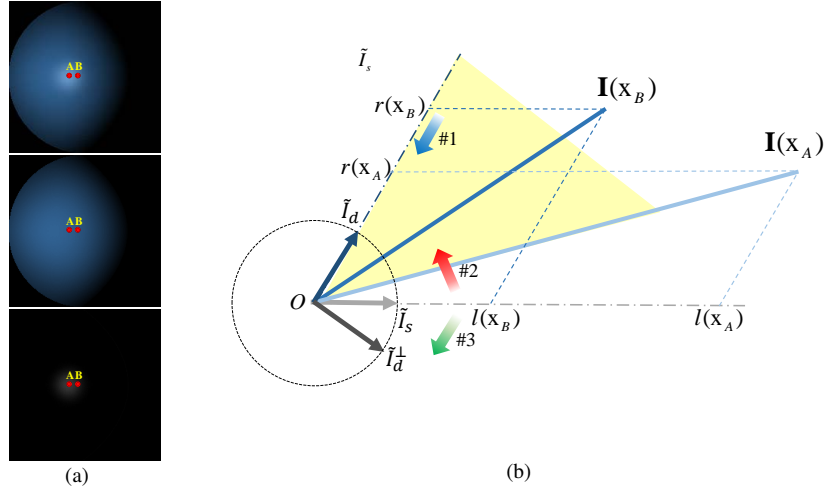


Fig. 4. Geometric interpretation of our mutual prior. (a) Top: synthetic image of a specular sphere. Middle: pure diffuse component. Bottom: specular component. (b) Illustration of the adaptive balance of the scene between the low-correlation prior and the sparsity constraint. $\tilde{\mathbf{I}}_d$ and $\tilde{\mathbf{I}}_s$ are the ground-truth chromaticity of the diffuse and specular reflection, respectively. The yellow area is the region of all the spectra in the scene. We label two locations A and B with image coordinates \mathbf{x}_A and \mathbf{x}_B , respectively, where the former has the strongest specularity. $\mathbf{I}(\mathbf{x}_A) = r(\mathbf{x}_A)\tilde{\mathbf{I}}_d + l(\mathbf{x}_A)\tilde{\mathbf{I}}_s$ and $\mathbf{I}(\mathbf{x}_B) = r(\mathbf{x}_B)\tilde{\mathbf{I}}_d + l(\mathbf{x}_B)\tilde{\mathbf{I}}_s$ describe the dichromatic model of points A and B , respectively. $\tilde{\mathbf{I}}_d^\perp$ denotes the direction orthogonal to the spectrum $\tilde{\mathbf{I}}_d$ in the hyperplane spanned by $\tilde{\mathbf{I}}_d$ and $\tilde{\mathbf{I}}_s$. The blue arrow (#1) and red arrow (#2) denote the effects of the sparsity constraint, and the green arrow (#3) denotes that of the low-correlation prior.

Intuitively, minimizing $\|\mathbf{D}'\mathbf{I}_s\|_2^2$ will draw the candidate solution along the direction orthogonal to the diffuse spectrum, i.e., $\tilde{\mathbf{I}}_d^\perp$, as visualized by the green arrow (#3). We balance these two constraints to obtain an accurate illumination spectrum.

2.4. Objective definition and parameter settings

By combining the above-mentioned individual and mutual priors, the illumination component \mathbf{I}_s in the region Ω_l can be estimated by solving the following optimization relation:

$$\begin{aligned} \arg \min_{\mathbf{I}_s} \quad & \|\mathbf{I}_s\|_* + \alpha_1 \|\mathbf{C}\|_1 + \alpha_2 \|\mathbf{N}\|_2^2 + \alpha_3 \|\mathbf{D}'\mathbf{I}_s\|_2^2 \\ \text{subject to} \quad & \mathbf{I} = \mathbf{D}\mathbf{C} + \mathbf{I}_s + \mathbf{N} \\ & \mathbf{I}_s \geq 0. \end{aligned} \quad (8)$$

In this objective function, $\|\cdot\|_*$ denotes the nuclear norm, which can be used to force a matrix to be low rank; $\|\cdot\|_1$ describes the l_1 norm of the coefficients; \mathbf{N} is the imaging noise; and minimizing the l_2 norm of $\mathbf{D}'\mathbf{I}_s$ describes the low-correlation prior favouring low correlation between reflectance and illumination. The first constraint formulates the fidelity of the data to the dichromatic reflection model, and the second one enforces nonnegativity of the illumination spectrum. With the reconstruction result, the illumination chromaticity $\tilde{\mathbf{I}}_s$ could be calculated by applying normalization according to Eq. (6).

In order to achieve high estimation accuracy, a good balance among these energy terms and constraints is crucial. The constraints of the sparsity coefficient and of noise are scene

independent, and we can statistically set their weighting factors. In implementation, we set $\alpha_1 = 1$ and $\alpha_2 = 5 \cdot 10^4$ to ensure that these two energy terms have the same magnitude, i.e., we attach the same importance to these two priors. Experiments show that our algorithm can converge with a large range of values of these two parameters. In contrast, the correlation between diffuse reflection and illumination may significantly vary across different images. Hence, an adaptive parameter setting for α_3 is crucial. If the pixel with the highest specular strength in the dark-channel image (e.g., A in Fig. 4) is sufficiently strong, i.e., the spectrum of this pixel (\mathbf{I}_A) is very close to the true illumination $\tilde{\mathbf{I}}_s$, we need only a smaller α_3 to avoid deviation towards it (\mathbf{I}_A). Inspired by this observation, we propose to set α_3 according to the specular strength, which can be measured using the ratio between the average strength in region Ω_r to that in Ω_l . Specifically, $\alpha_3 = \frac{\mathbf{I}(\mathbf{x} \in \Omega_r)}{\mathbf{I}(\mathbf{x} \in \Omega_l)} * 5 \cdot 10^{-3}$ here.

3. Numerical solution via optimization

The problem defined in Eq. (8) is a typical optimization problem with equality constraints. It has been proven that the augmented Lagrange multiplier method (ALM) with the alternating direction minimizing (ADM) strategy proposed by Lin et al. [32] is suitable to solve this optimization. Specifically, we minimize the following Lagrangian function with an auxiliary variable $\mathbf{S} = \mathbf{I}_s$:

$$\begin{aligned} \text{Lag} &= \|\mathbf{S}\|_* + \alpha_1 \|\mathbf{C}\|_1 + \alpha_2 \|\mathbf{N}\|_2^2 + \alpha_3 \|\mathbf{D}'\mathbf{I}_s\|_2^2 \\ &+ \frac{\beta}{2} \|\mathbf{I} - \mathbf{DC} - \mathbf{I}_s - \mathbf{N}\|_2^2 + \langle \mathbf{Y}_1, \mathbf{I} - \mathbf{DC} - \mathbf{I}_s - \mathbf{N} \rangle \\ &+ \frac{\beta}{2} \|\mathbf{S} - \mathbf{I}_s\|_2^2 + \langle \mathbf{Y}_2, \mathbf{S} - \mathbf{I}_s \rangle, \end{aligned} \quad (9)$$

where $\langle \cdot \rangle$ denotes the inner product, \mathbf{Y}_1 and \mathbf{Y}_2 are two Lagrange multipliers, and β is a parameter balancing the constraints. To solve this problem, we start from a coarse initialization and update iteratively. Specifically, this optimization can be sequentially decomposed into several sub-optimization problems with respect to \mathbf{S} , \mathbf{N} , \mathbf{C} , and \mathbf{I}_s .

• **Initialization:** We set the initial illumination spectrum $\tilde{\mathbf{I}}_s^0$ using the weighted SVD algorithm. Based on the observation that a higher intensity in the dark-channel image indicates stronger specular strength, we regard the top principal component of the weighted spectra as the illumination spectrum. Here, the weights are set as the intensities of the dark-channel image. Intuitively, the precision of the initialization is determined by the strength of specular strength. The initial illumination spectrum of the image in Fig. 1(a) is plotted with a green dash-dot curve in Fig. 3(c), which shows that some deviation exists. Next, the initial diffuse component $\tilde{\mathbf{I}}_d^0$ is set by subtracting the maximum projection of \mathbf{I} along $\tilde{\mathbf{I}}_s^0$ satisfying the nonnegative constraint on the residue. We calculate \mathbf{C}^0 by projecting $\tilde{\mathbf{I}}_d^0$ onto the the most correlated entry in \mathbf{D} . In addition, the Lagrange multipliers are initialized as $\mathbf{Y}_1^0 = \mathbf{Y}_2^0 = 0$, and the balancing parameter $\beta^0 = 0.1$.

• **S-subproblem:** By preserving the items related to \mathbf{S} in Eq. (9), we can update \mathbf{S} following

$$\begin{aligned} \mathbf{S}^{t+1} &= \arg \min_{\mathbf{S}} \left\{ \|\mathbf{S}\|_* + \frac{\beta^t}{2} \|\mathbf{S} - \mathbf{I}_s^t + \frac{\mathbf{Y}_2^t}{\beta^t}\|_2^2 \right\} \\ &= \mathbf{U}\Sigma_{[1]}\mathbf{V}^T, \end{aligned} \quad (10)$$

where $\mathbf{U}\Sigma\mathbf{V}^T$ denotes the singular value decomposition of $(\mathbf{I}_s^t - \frac{\mathbf{Y}_2^t}{\beta^t})$, $\Sigma_{[1]}$ denotes the preservation of only the largest entry of matrix Σ , and t indexes the iteration.

Algorithm 1: Algorithm for illumination-spectrum estimation

Input : A multichannel image \mathbf{I} contaminated by specularity

Output : Estimated illumination spectrum $\tilde{\mathbf{I}}_s$

- 1 Get the dark-channel image to obtain regions Ω_r and Ω_l and calculate α_3 ;
 - 2 Build the over-complete dictionary \mathbf{D} using the pixels in Ω_r ;
 - 3 Initialize the illumination spectrum $\tilde{\mathbf{I}}_s^0$ using the pixels in Ω_l through weighted SVD;
 - 4 Initialize \mathbf{I}_d^0 by subtracting \mathbf{I} 's component along $\tilde{\mathbf{I}}_s^0$ and set $\mathbf{C}^0 = \text{OMP}(\mathbf{D}, \mathbf{I}_d^0, 1)$;
 - 5 Initialize $\mathbf{Y}_1^0 = \mathbf{Y}_2^0 = 0$, $\beta^0 = 0.1$, and set $t = 0$;
 - 6 **while** *not converged* **do**
 - Update \mathbf{S}^{t+1} according to Eq. (10);
 - Update \mathbf{N}^{t+1} according to Eq. (11);
 - Update \mathbf{C}^{t+1} according to Eq. (12);
 - Update \mathbf{I}_s^{t+1} according to Eq. (13);
 - Update \mathbf{Y}_1^{t+1} , \mathbf{Y}_2^{t+1} and β^{t+1} according to Eq. (14);
 - $t := t + 1$;
 - 7 **end**
 - 8 Obtain the illumination chromaticity according to Eq. (6).
-

• **N-subproblem:** The items related to \mathbf{N} compose a typical quadratic energy function and can be minimized in a closed-form manner as

$$\begin{aligned}\mathbf{N}^{t+1} &= \underset{\mathbf{N}}{\operatorname{argmin}} \left\{ \alpha_2 \|\mathbf{N}\|_2^2 + \frac{\beta^t}{2} \|\mathbf{I} - \mathbf{D}\mathbf{C}^t - \mathbf{N} - \mathbf{I}_s^t + \frac{\mathbf{Y}_1^t}{\beta^t}\|_2^2 \right\} \\ &= \frac{\beta^t}{2\alpha_2 + \beta^t} (\mathbf{I} - \mathbf{D}\mathbf{C}^t - \mathbf{I}_s^t + \frac{\mathbf{Y}_1^t}{\beta^t}).\end{aligned}\quad (11)$$

• **C-subproblem:** The coefficients \mathbf{C} of the diffuse reflection spectrum is updated based on the assumption that the diffuse component \mathbf{I}_d can be exactly reconstructed by \mathbf{C} , i.e., $\mathbf{I}_d = \mathbf{D}\mathbf{C}$. Hence, we use the orthogonal matching pursuit algorithm (OMP) proposed in [33] to find the entry with the highest correlation to $\mathbf{I} - \mathbf{I}_s^t - \mathbf{N}^{t+1}$:

$$\mathbf{C}^{t+1} = \text{OMP}(\mathbf{D}, \mathbf{I} - \mathbf{I}_s^t - \mathbf{N}^{t+1}, 1). \quad (12)$$

Here, 1 denotes the number of matching bases.

• **I_s-subproblem:** Similarly to the optimization of \mathbf{N} , we can derive a closed-form updating rule of \mathbf{I}_s as

$$\begin{aligned}\mathbf{I}_s^{t+1} &= \max\{0, \underset{\mathbf{I}_s}{\operatorname{argmin}} \alpha_3 \|\mathbf{D}'\mathbf{I}_s\|_2^2 + \frac{\beta^t}{2} (\|\mathbf{I} - \mathbf{D}\mathbf{C}^{t+1} - \mathbf{N}^{t+1} - \mathbf{I}_s + \frac{\mathbf{Y}_1^t}{\beta^t}\|_2^2 + \|\mathbf{S}^{t+1} - \mathbf{I}_s + \frac{\mathbf{Y}_2^t}{\beta^t}\|_2^2)\} \\ &= \max\{0, (2\mathbf{E} + \frac{\alpha_3}{\beta^t} \mathbf{D}\mathbf{D}')(\mathbf{S}^{t+1} + \mathbf{I} - \mathbf{D}\mathbf{C}^{t+1} - \mathbf{N}^{t+1} + \frac{\mathbf{Y}_1^t + \mathbf{Y}_2^t}{\beta^t})\},\end{aligned}\quad (13)$$

where $\mathbf{E} \in \mathbb{R}^{k \times k}$ is the unit matrix.

In addition, two Lagrangian Multipliers \mathbf{Y}_1 and \mathbf{Y}_2 and the weighting factor β are updated

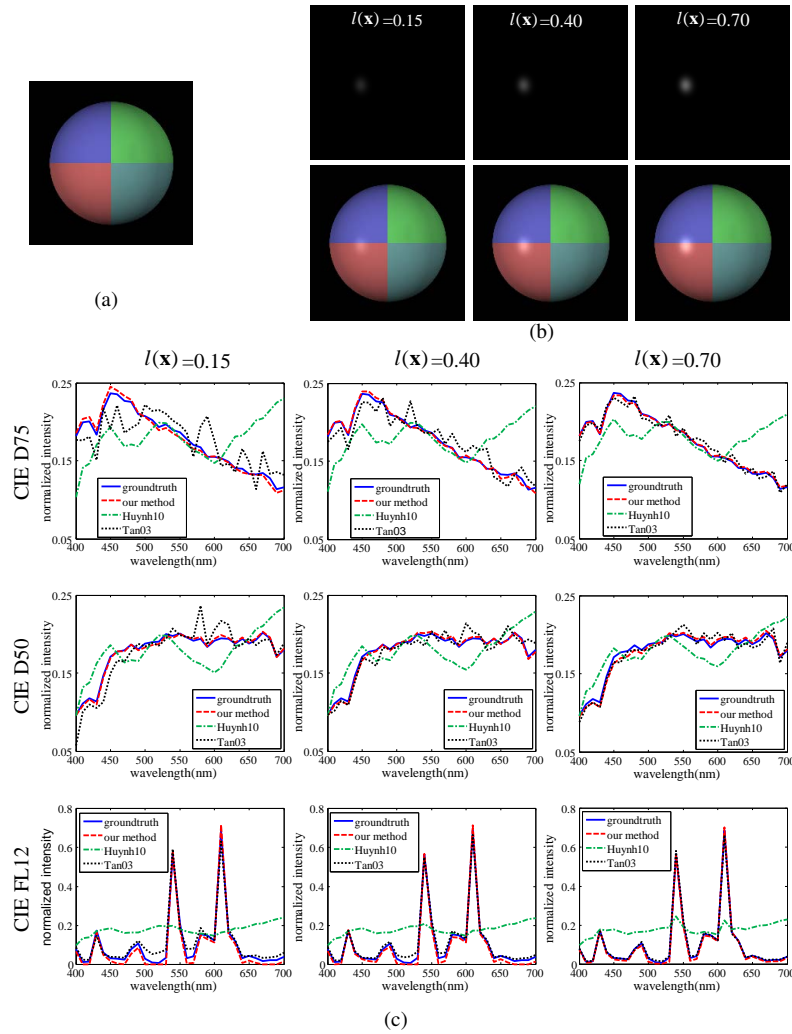


Fig. 5. Illumination-estimation results for synthetic multi-spectral data with one specular pattern. (a) Diffuse component. (b) Three specular components of the same distribution but with different strengths (top) and corresponding synthetic specularity-contaminated images (bottom). Here we only show data synthesized from white illumination, instead of three different color illuminations, because of the limited space. (c) Performance comparison of the simulated data with different illuminations and the varying specularity strengths in (b).

according to the following rules:

$$\begin{aligned}
 \mathbf{Y}_1^{t+1} &= \mathbf{Y}_1^t + \beta^t (\mathbf{I} - \mathbf{D}\mathbf{C}^{t+1} - \mathbf{I}_s^{t+1} - \mathbf{N}^{t+1}), \\
 \mathbf{Y}_2^{t+1} &= \mathbf{Y}_2^t + \beta^t (\mathbf{S}^{t+1} - \mathbf{I}_s^{t+1}), \\
 \beta^{t+1} &= \min\{\rho\beta^t, \beta_{\max}\},
 \end{aligned} \tag{14}$$

where ρ and β_{\max} are constants set empirically in this paper as $\rho = 1.1$ and $\beta_{\max} = 2 \cdot 10^6$, respectively. The parameters are the same as in [34], and the experiments show that slight parameter changes will not affect the results.

The main steps of the numeric algorithm are summarized in Alg. 1. After iterative optimization, the estimated illumination spectrum approximately converges to the ground truth, as plotted in Fig. 3(c).

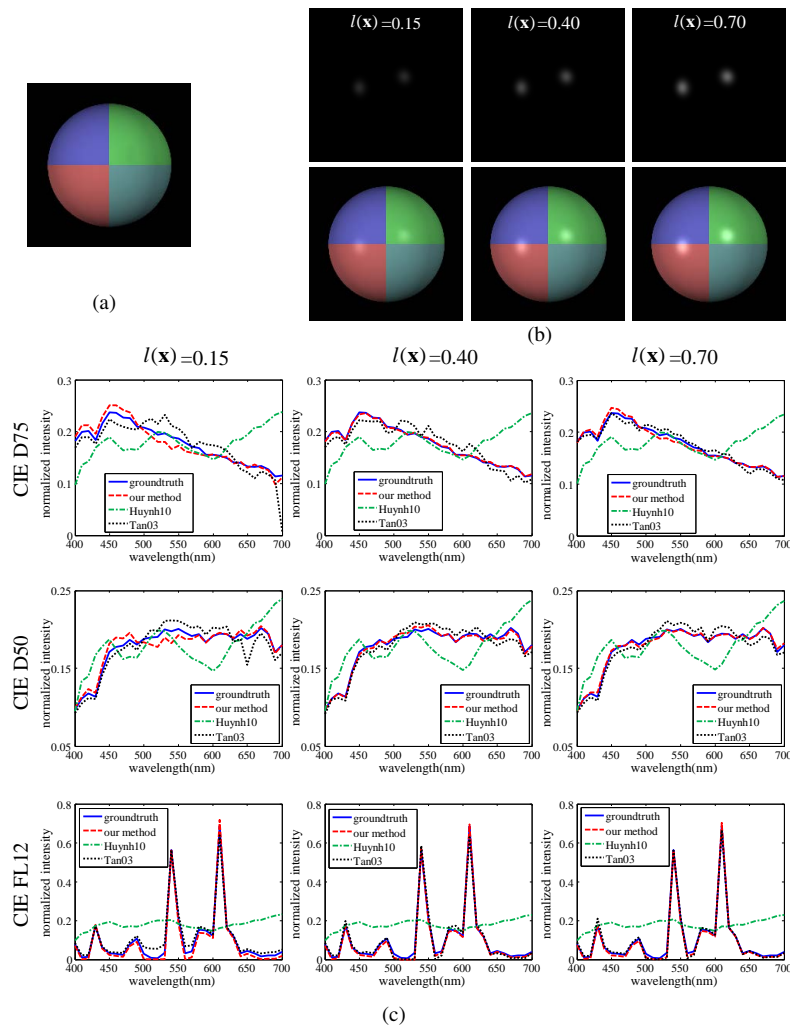


Fig. 6. Illumination-estimation results for synthetic multi-spectral data with two specular patterns. (a) Diffuse component. (b) Three specular components with different strengths (top) and corresponding specularity-contaminated images (bottom) with two specular regions. Again, we only show white illumination here. (c) Performance comparison of the simulated data with different illuminations and the three specularity strengths in (b).

4. Experiments

In this section, we conduct a series of experiments to evaluate our algorithm quantitatively. In the simulation, specular multi-spectral images are generated from a purely diffuse image, given an illumination spectrum and specular distribution pattern according to the dichromatic reflection model. We also compare the accuracy between our algorithm and two previous methods—the latest [1] and the most cited [8]—under different illuminations and specularity strengths. In real captured data, we obtain the ground truth of the illumination with a color checker and demonstrate the superiority and robustness of the proposed method in both simple and textured surfaces. We test the running time on an Intel Xeon 2.27 GHz CPU workstation with a 64-bit Windows 7 operating system. On average, the algorithm converges within 150 loops, and processing an image of 500×500 pixels takes 5.0 s. For high-resolution images, we can first down-sample the images into low-resolution ones and then run the proposed algorithm for acceleration.

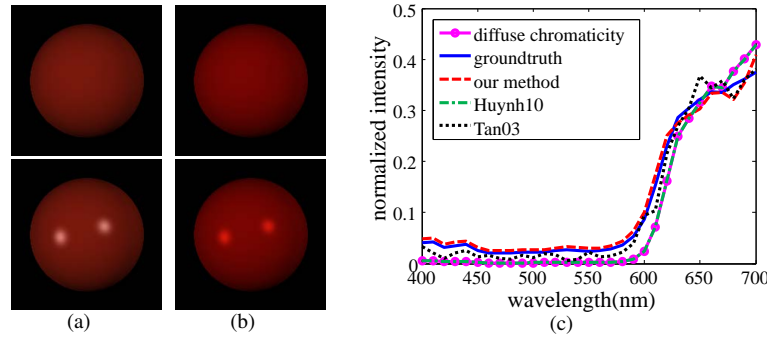


Fig. 7. Performance for an example with similar diffuse and specular chromaticity. (a) A red sphere under white illumination. (b) The same red sphere of (a) under red illumination. (c) Diffuse chromaticity under red illumination and the estimated illuminations.

4.1. Synthetic analysis

In this experiment, we synthesize a 31-channel image with one or two Gaussian-shaped specular region(s). The specular-free image is visualized in Fig. 5(a) and Fig. 6(a) by integrating the multi-spectral data with the RGB response curves of Canon 20D. The synthesized 31-channel image is a linear combination of the specular-free component with the reflectance factor $r(\mathbf{x}) = 1$ and the specular component, the weighting factor $l(\mathbf{x})$ of which varies from 0.15 to 0.7, as shown in Fig. 5(b) and Fig. 6(b). Without loss of generalization, the 31-channel specular components are set by integrating the specular strength pattern generating the data in the upper row of Fig. 5(b) and Fig. 6(b) with the spectrum curves of CIE-standard illuminants: CIE D75, CIE D50, and CIE FL12. The results of two state-of-the-art specular-based illumination-estimation methods [1, 8] are also shown for comparison. Without making any assumption on the illumination spectrum, our method can estimate both smooth (the first and second rows of Fig. 5(c) and Fig. 6(c)) and non-smooth illumination spectra (the third row of Fig. 5(c) and Fig. 6(c)) with high accuracy. The results for three different illuminations exhibit a consistent trend: the accuracy of the previous methods is satisfactory only in the cases with strong specularity, while our method can achieve high accuracy in all the cases. Overall, our approach performs better than the two previous methods in all the cases with different illuminations and specularity strengths.

To test the wide applicability of the proposed approach, we run our algorithm on a red sphere illuminated by two red light sources, the chromaticity of which is the same as the sphere's diffuse reflection under white illumination. In Fig. 7(a) and Fig. 7(b), the appearance of the scene illuminated by white and red illumination is shown, respectively. Although both Eq. (1) and the images indicate that the sphere under red illuminations appears redder compared to that under white illumination, the correlation between illumination and diffuse reflection is still very high, as shown in Fig. 7(c). From the plot in Fig. 7(c), we can see that in this case, the performance of our method is still higher than those of previous methods. This is mainly due to the scene-adaptive setting of the weighting factor α_3 .

4.2. Real captured multi-channel images

We test the performance of our illumination-estimation approach for real captured data. The multichannel images in this experiment are taken from two databases. One is the widely used multi-spectral data set collected by Yasuma et al. [35], as exemplified in Fig. 8(a). They captured the data by adding a liquid-crystal tunable filter (VariSpec) in front of the lens of a cooled CCD camera (Apogee Alta U260). The spectrum covers 400nm~700nm and includes 31 non-overlapping channels. To demonstrate the performance in cases with weak specularity, we col-

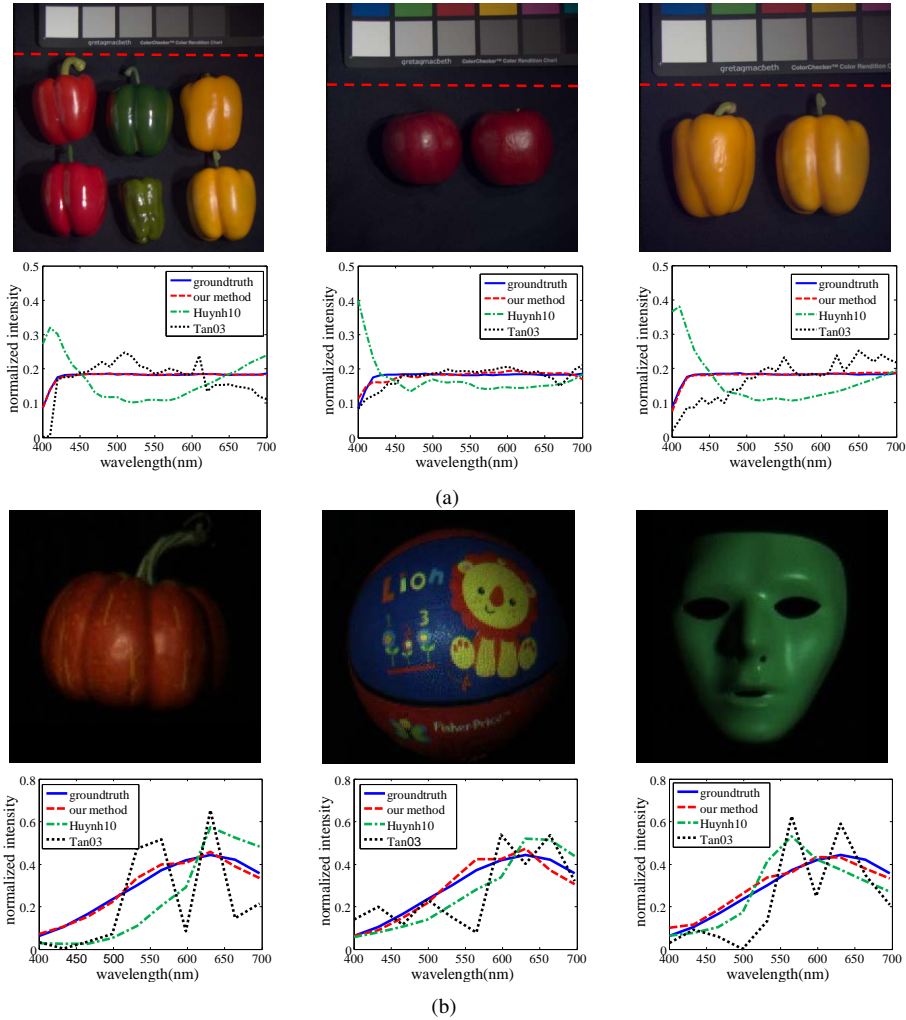


Fig. 8. Illumination-estimation results for real captured data. (a) Results for data with strong specularities from [35]. We use the pixels below the red line to estimate illumination and the color checker above this line to obtain the ground truth. (b) Results for examples with weak specularities and the true illumination spectrum are obtained by using a color checker.

lect a second database by placing a set of narrow-band filters manufactured by Thorlabs Inc. in front of a monochrome camera PointGrey GRAS-50S5M. The light source in this experiment is a halogen lamp with a very smooth spectrum, as shown in Fig. 9(a). The spectrum has the same range as in [35] and is decomposed into 11 non-overlapping bands, with the central wavelengths and bandwidths plotted in Fig. 9(c). The spectra in Fig. 9(a) and Fig. 9(c) are measured using the Maya2000 Pro Spectrometer manufactured by Ocean Optics. Fig. 8(b) displays three example scenes. For images in our database, we integrate the multi-spectral images with the RGB response curves of Canon 20D shown in Fig. 9(b) to generate the RGB images for visualization [36].

For the images in Fig. 8(a), we use the regions below the red dashed line to estimate illumination and use the gray swatches of the color checker in the upper region to obtain the ground truth. The ground-truth illumination spectrum in Fig. 8(b) is obtained using a similar color checker. One can see that our algorithm achieves high estimation accuracy for both

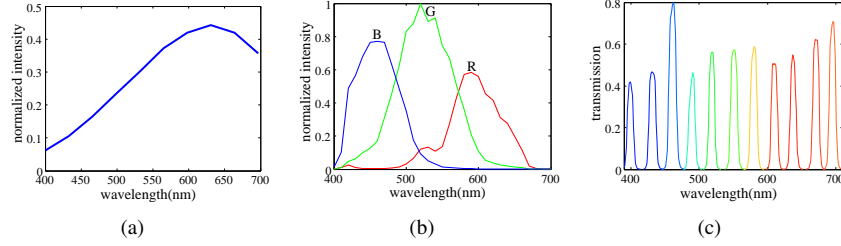


Fig. 9. (a) Normalized spectrum of the halogen lamp. (b) Sensor spectral response of Canon 20D. (c) Transmission curves of the narrow-band filters.

data sets. On comparing Fig. 8(a) and Fig. 8(b), the performance of our algorithm is found to drop by a small amount when high correlation exists between the illumination chromaticity and diffuse reflectance, such as in the images of Fig. 8(b), where the scene materials appear reddish/greenish and the illumination appears yellowish. However, our algorithm still gives promising performance because of the scene-adaptive setting of the weighting factor α_3 on the low-correlation prior $\|\mathbf{D}'\mathbf{I}_s\|$. The plots also show that our algorithm can be used to obtain an estimation closer to the ground truth than that obtained with state-of-the-art methods.

For the scenes including white or grey materials, the chromaticity of diffuse and specular components are identical at these locations. We specially design a series of experiments to demonstrate the applicability of our method in such cases, as shown in Fig. 10. We test the performance and analyze the behaviors of our algorithm in three different cases: the white/grey materials wholly separated into Ω_l , wholly separated into Ω_r , and partially separated into Ω_l and Ω_r . In the scene displayed in the top row of Fig. 10(a), pixels corresponding to the white material (characters on the top) are wholly separated into Ω_l in the binary map, as shown in Fig. 10(b). The reflectance of these white materials can be regarded as pure specular highlight with the diffuse component being zero, i.e., $r(\mathbf{x}) = 0$. Therefore, our approach can result in a good estimation in this case, as shown in Fig. 10(d). In the scene shown in the second row, the white and grey patches on the checkerboard are all separated into Ω_r ; therefore, a base exists (the black solid curve in Fig. 10(c)) with a spectrum similar to the illumination spectrum. In spite of the existence of this “illumination entry,” our optimization model chooses the correct entry to fit the diffuse component for specular-influenced pixels (we use only one dictionary entry to reconstruct the diffuse component of each pixel); otherwise, the low-rank constraint on the residual specular component will be violated. Thus, the estimated illumination is still accurate. The third row shows a scene in which the white material is separated into two regions: Ω_r and Ω_l . In the optimization, the reflectance of the white material in Ω_l can be reconstructed using the “illumination entry” with the residual being zero, while the reflectance of colored materials is fitted in the same manner as that in the second example is. Overall, our illumination-estimation result still shows only a small deviation from the true value. From the above exhaustive listing of the cases containing white materials, we can see that our approach has wide applicability and high accuracy. Though the correlation between the dictionary and illumination increases in the latter two cases, α_3 would automatically decrease.

The comparison with previously published methods leads to a similar conclusion to that from the synthetic experiment: (i) The proposed approach shows the best performance among the three algorithms. (ii) The superiority of our approach is especially prominent in cases with weak specularities for scenes with both simple and rich textures. In summary, the effectiveness and robustness of our approach is further validated with real captured data.

5. Conclusions and discussions

We introduced a specular-based illumination-estimation approach with an optimization framework. Our approach benefits from the extensive utilization of multiple priors, scene-adaptive

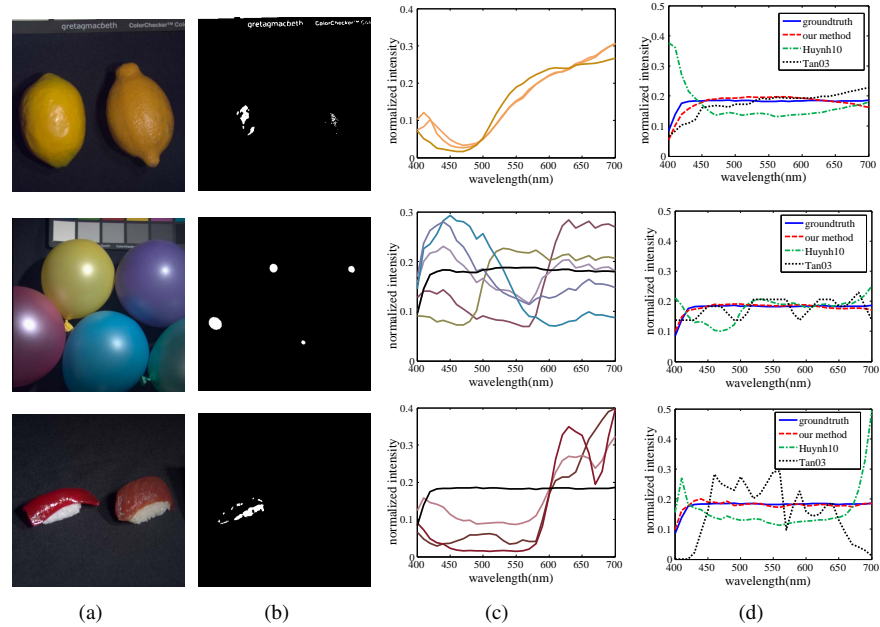


Fig. 10. Illumination estimation from non-Lambertian scenes containing white and grey colors, using the data from [35]. (a) The source images. (b) The threshold binary map. (c) The learned dictionary. (d) The ground-truth illumination and estimation results.

parameter setting, and effective numerical solution to achieve superior performance compared to those of previous methods. Experiments demonstrate that the proposed approach is robust to the diversity of nature images, including illumination types, strength of specular, and structure of scene surfaces.

The proposed approach has wide applicability because it imposes only a weak assumption on the target scene: each pixel within Ω_l has a counterpart describing the same material in Ω_r . This assumption is violated when the pixels describing one specific material are completely covered with strong specular. However, because such pixels are usually small in number, our model would treat them as noise and thus still has high robustness. The number of channels is a factor affecting the final performance. Mathematically, the number of channels is mainly related to the low-rank prior, which plays a larger role with more color channels. Experimentally, we need no less than approximately 5 channels to obtain high performance. The proposed method cannot handle situations in which only white or grey materials are contaminated by specular, because in this case, the specular and diffuse components have exactly the same chromaticity and are thus inseparable. This is a limitation shared by all illumination-estimation methods based on the dichromatic reflection model.

The performance of our method is quite promising but can be improved further. One possible improvement entails incorporating the priors of illumination chromaticity. We can either resort to a data-driven strategy or physical knowledge. In the future, we plan to extend the current approach to address cases with several different illuminants.

Acknowledgments

This work was supported by projects of the National Science Foundation of China (No. 61171119 and 61120106003). The research was also funded by the Beijing Key Laboratory of Multi-dimension & Multi-scale Computational Photography (MMCP), Tsinghua University.