



Handling Specularity in Intrinsic Image Decomposition

Siraj Muhammad¹(✉), Matthew N. Dailey¹, Imari Sato²(✉),
and Muhammad F. Majeed³(✉)

¹ Asian Institute of Technology (AIT), Khlong Luang, Thailand
{Siraj.Muhammad,mdailey}@ait.asia

² National Institute of Informatics (NII), Tokyo, Japan
imarik@nii.ac.jp

³ Shaheed Benazir Bhutto University, Sheringal (SBBU Sheringal),
Sheringal, Pakistan
m.faran.majeed@ieee.org

Abstract. Intrinsic image decomposition (IID), the process of separating an image into reflectance and shading components, is one of the fundamental problems in computer vision. Various approaches for IID have been proposed, but most assume Lambertian surfaces. In this paper, we propose a method that handles specularity while decomposing an input image into reflectance and shading components. The method first removes specularities from the image, and then it decomposes the image into reflectance and shading components. We propose a new algorithm for reconstruction of an image's diffuse component and demonstrate the effectiveness of the method under specularity based on the extracted reflectance and shading images. Future work will focus on a more extensive empirical evaluation against ground truth and handling of shadows.

Keywords: Di-chromatic reflection model · Specularity · Reflectance Shading · Intrinsic image decomposition

1 Introduction

In many computer vision applications, steps such as segmentation, object detection or recognition, three-dimensional modeling, and albedo and shading analysis are adversely affected by lighting conditions. In particular, specularity can disrupt a variety of different visual processing steps.

Intrinsic image decomposition (IID) is one such step that aims to decompose an image into reflectance, shading, and illumination components, simplifying the robustness of downstream processes. We find that, like other vision processes, IID algorithms are disrupted by the kinds of specularities seen in practical applications such as video surveillance. If IID is to be useful across a wide spectrum of practical applications, then we require methods that are robust to specularities. Therefore, in this paper, we develop a method for handling specularity prior to decomposing an image into reflectance and shading with IID.

Many algorithms for IID have been proposed. Land and McCann’s classic retinex method [1] was among the first to decompose images into illumination and reflectance components and is used as a basic approach in many of today’s methods. Rother et al. [2] address the problem of decoupling material-dependent properties (reflectance, albedo, and diffuse reflectance) and light-dependent properties (specular reflectance, shading, and inter-object reflection) from a single image using a prior on the reflectance component. Similarly, Barron et al. [3] decompose images into albedo and illumination components. Barron and Malik [4] recover shape, surface color, and illumination color from a single image. Laffont et al. [5] propose a method that decomposes an input image into intrinsic images from multiple view angles of a scene, without a priori knowledge of scene geometry. The method uses multi-view stereo to reconstruct a 3D point cloud. Although the cloud is sparse and incomplete, it provides the necessary information about illumination at each point. The authors optimize the point cloud and estimate sun visibility, i.e., decomposing illumination into sun, sky, and indirect illumination layers. Shi et al. [6] present an approach that decomposes input RGBD images into reflectance and shading images. The method assumes piecewise constant reflectance, allowing superpixel clustering of the reflectance image. Pixels lying in the same superpixel are assumed to have identical reflectance; any color variation is assumed to arise from shading. Kang et al. [7] perform IID on hyper-spectral images and extract features from the image set thus obtained. They use a method by Shen et al. [8] for the IID step.

Despite the richness of this literature, most IID methods assume Lambertian surfaces and diffuse lighting and are thus not robust under more adverse conditions, especially specular reflections common in practical applications with less-than-ideal illumination, such as video surveillance.

Therefore, in this paper, we propose to handle specularities while decomposing an image into reflectance and shading components. Our method comprises two steps. In the first step, we decompose the image into a specular component and a diffuse component with a new method that improves upon the state of the art. In the second step, the resulting image is decomposed into its specular-free reflectance and shading components. In the future we will add shadow handling and compare our pipeline with that of Fan et al [9]. We evaluate our method qualitatively by conducting experiments on synthetic and real images. The results show that our approach is successful at producing specular-free reflectance and shading components. Our method may help computer vision applications such as object detection, segmentation, and tracking under adverse lighting conditions.

In the next section, we present the method in detail. In Sect. 3, we discuss our experimental setup and results, and in Sect. 4, we conclude and discuss future work.

2 Proposed Model

A schematic of our method is shown in Fig. 1. It consists of two components: specularity removal and IID. We describe each component separately in Sects. 2.1 and 2.2.

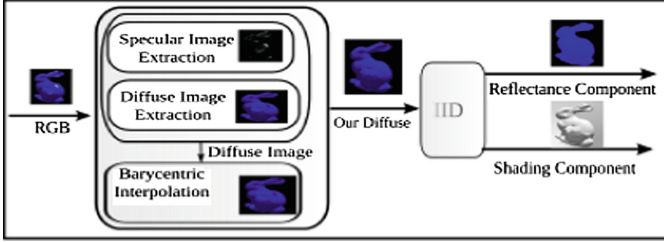


Fig. 1. Proposed model for specularity-aware intrinsic image decomposition (IID).

2.1 Specularity Removal

In the specularity removal component, we begin with the well-known dichromatic reflection model for image formation [10]. This model has been used in many researchers' work [11–14]. According to the dichromatic reflection model, the RGB color $\mathbf{V}(p)$ of observed pixel with index p is the combination of diffuse and specular reflection

$$\mathbf{V}(p) = \alpha(p)\mathbf{V}_d(p) + \beta(p)\mathbf{V}_s(p), \quad (1)$$

where $\mathbf{V}_d(p)$ is the (unknown) underlying diffuse color value of pixel p , $\mathbf{V}_s(p)$ is the (unknown) underlying specular color value of pixel p , and $\alpha(p)$ and $\beta(p)$ represent the (unknown) contributions of diffuse and specular reflection to the observed pixel p , respectively. We use row vectors to denote RGB intensities for a pixel.

Clearly, the problem of finding $\alpha(p)$, $\mathbf{V}_d(p)$, $\beta(p)$, and $\mathbf{V}_s(p)$ consistent with $\mathbf{V}(p)$ is underconstrained. Our method extends that of H.L. Shen et al. [14]. First, following H.L. Shen et al., we categorize each pixel p as either diffuse or specular based on whether all three color components are above (specular) a threshold or not (diffuse). Next, let p be the diffuse pixel with the highest value of R, G, or B. Set $\mathbf{V}_d(p) = \mathbf{V}(p)$ and $\mathbf{V}_s(p) = [0, 0, 0]$. After that, for each diffuse pixel q whose chromaticity (normalized RGB vector) is within a threshold distance of that of p , the method calculates $\alpha(q)$ and $\beta(q)$ as

$$\begin{bmatrix} \alpha(q) \\ \beta(q) \end{bmatrix} = [\mathbf{V}_d(p)^T \quad \mathbf{L}^T]^+ \mathbf{V}(q)^T, \quad (2)$$

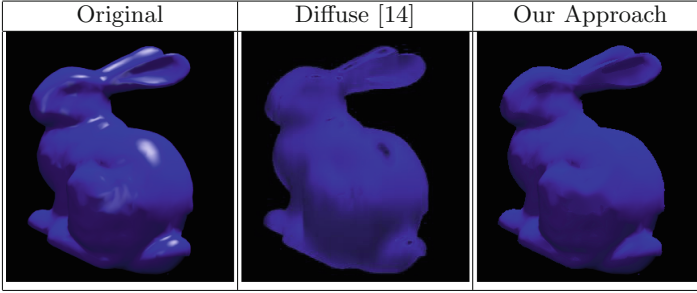


Fig. 2. Diffuse component of image with large bright specularities using H.L. Shen et al. method [14] and our method.

where $\mathbf{L} = [255 \ 255 \ 255]$. By assuming $\mathbf{V}_d(q) = \mathbf{V}_d(p)$ for each pixel q similar to p , we are left with two simple linear equations with one unknown each, allowing direct solution for $\mathbf{V}_s(q)$. The process is repeated with another p not yet considered until every diffuse pixel is resolved.

In H.L. Shen et al.’s method, after each diffuse pixel q has been assigned values for $\mathbf{V}_d(q)$ and $\mathbf{V}_s(q)$, a similar procedure is applied to the specular pixels, assuming that the diffuse component for each specular pixel q is the same as the diffuse pixel p with the most similar chromaticity.

This method works well in some cases, but it does not work well in the case of large, bright specularities, which are common in real world situations. Bright specularity dominates the diffuse component of the light to the point that the original diffuse chromaticity cannot be recovered, and the assumption that the specular pixel’s diffuse component should be same as that of the diffuse pixel with the most similar chromaticity introduces artifacts in the estimated diffuse image. An example is shown in Fig. 2.

To handle this issue and provide more accurate reconstruction of the diffuse component of the image under such conditions, we propose a simple interpolation method able to accurately reconstruct the diffuse component of bright specular regions. Our method uses barycentric interpolation of the RGB components of the diffuse image based on the boundary of the specular region, as determined in the first step. An example is shown in Fig. 2. Diffuse pixel value $\mathbf{V}_d(q)$ for pixel q within a specular region is reconstructed as

$$\mathbf{V}_d(q) = \sum_{j \in B(q)} [\omega_{jq} \mathbf{V}_d(j)], \quad (3)$$

where $B(q)$ represents the set of pixels on the boundary of the specular region containing pixel q , $\mathbf{V}_d(j)$ represents the RGB diffuse component of boundary pixel j , and ω_{jq} is a barycentric weight calculated as

$$\omega_{jq} = \frac{\exp(-\gamma \|j - q\|)}{\sum_{j \in B(q)} \exp(-\gamma \|j - q\|)}. \quad (4)$$

Here $\|j - q\|$ represents the spatial Euclidean distance between specular pixel q and boundary pixel j . $\gamma > 0$ is a normalization constant. Large values of γ give high relative weights to the closest boundary points, and lower values of γ give more uniform weights over boundary points. It would be possible to make γ adaptive to the size of the specular region in question, but we find that $\gamma = 0.25$ gives good weight distributions for most specular region size. We pass the diffuse component of the input image to the next step, intrinsic image decomposition.

2.2 Intrinsic Image Decomposition (IID)

After removing specularities, the next step is to decompose the image into reflectance and shading components. We use the well-known intrinsic image decomposition model of Rother et al. [2], which has been used by many other researchers [5, 8, 15, 16]. We adopt their basic approach following Shen et al. [8], assuming that after specularity removal,

$$\mathbf{V}_d(p) = \mathbf{L}(p)\mathbf{R}(p), \quad (5)$$

where $\mathbf{V}_d(p)$ represents an input diffuse image pixel, $\mathbf{L}(p)$ represents the diffuse illumination, and $\mathbf{R}(p)$ represents the diffuse reflectance of pixel p . J. Shen et al. [8] assume that neighbor pixels having similar intensity will have similar reflectance, leading them to minimize the energy function

$$E(\mathbf{R}, \mathbf{L}) = \sum_p \left(\mathbf{R}(p) - \sum_{q \in \mathcal{N}(p)} w_{pq} \mathbf{R}(q) \right)^2 + \sum_p (\mathbf{V}_d(p)/\mathbf{L}(p) - \mathbf{R}(p))^2, \quad (6)$$

where \mathbf{R} is the reflectance image, \mathbf{L} is the shading image, $\mathcal{N}(p)$ is a local neighborhood around pixel p (for example a 3×3 window), and w_{pq} is a weight indicating the similarity of pixels p and q :

$$w_{pq} = e^{-[(\tilde{\mathbf{V}}_d(p), \tilde{\mathbf{V}}_d(q))^2 / \rho_{pT}^2 + (Y(p) - Y(q))^2 / \rho_{pY}^2]}, \quad (7)$$

where $(\tilde{\mathbf{V}}_d(p), \tilde{\mathbf{V}}_d(q))$ is the cosine of the angle between normalized RGB vectors $\tilde{\mathbf{V}}_d(p)$ and $\tilde{\mathbf{V}}_d(q)$, and $Y(p)$ denotes the intensity or luminosity of pixel p in \mathbf{V} . ρ_{pT}^2 and ρ_{pY}^2 represent the variance of the angles and intensities, respectively, around pixel p .

3 Empirical Evaluation

To evaluate the method described in the previous section, we use two datasets. First, we generated a new dataset in which well-known 3D models (monkey, bunny, and teapot) are rendered with specularity using the Blender software

package. For purposes of quantitative analysis using Local Mean Squared Error (LMSE) [17], we also rendered corresponding ground-truth reflectance and shading images without specularity. Second, we use a dataset of images provided by H.L. Shen et al. [14]. No ground truth is available for these images, so the results can only be evaluated qualitatively. We performed two sets of experiments. In Experiment I, we ran dataset 1 through the pipeline of Fig. 1, and in Experiment II, we ran dataset 2 through the pipeline of Fig. 1. We report local mean square error and show qualitative results for the case (dataset 1) where we have ground truth, and we show qualitative results for dataset 2, where we have no ground truth.

3.1 Experiment I: Images with Specularity

As dataset 1 (consisting of synthetic images rendered by Blender) has ground truth, we compare the results of specularity removal and IID with the ground truth reflectance and shading components obtained from Blender. A quantitative comparison is shown in Fig. 3. The evaluation shows that our proposed framework produces better results than those of J. Shen et al. [8].

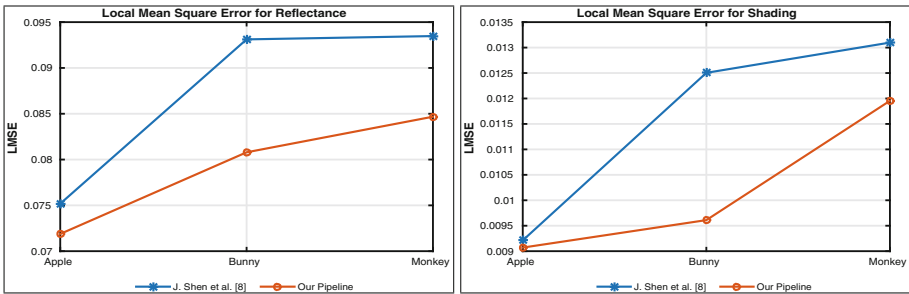


Fig. 3. Quantitative analysis of proposed framework using local mean squared error (LMSE)

Qualitative results are shown in Fig. 4. It can be seen that our proposed framework produces better specular-free reflectance and shading images than those provided by Shen et al. [8].

3.2 Experiment II: Specularity from [14]

This experiment uses images provided by Shen et al. [14]. Qualitative experimental results are shown in Fig. 4. By visual inspection, it is clear that our proposed model produces better specular-free reflectance and shading images than produced by the method of Shen et al. [8].



























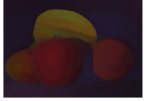



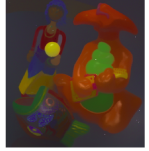



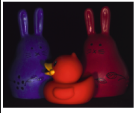




Experiment I		Image	J. Shen et al. [8]	Proposed	Ground Truth	
		Reflectance Results				
						
						
		Shading Results				
						
						
Experiment II		Image	J. Shen et al. [8] Reflectance	Proposed Reflectance	J. Shen et al. [8] Shading	Proposed Shading
						
						
						

Fig. 4. Experiment I & II results. Proposed method produces reflectance images closer to the ground truth than the original J. Shen et al. method.

4 Conclusion and Future Work

Adverse lighting conditions, particularly specular reflections, negatively affect important computer vision tasks such as feature detection, tracking, image segmentation, and object detection. In this paper, we propose a pipeline for removal of specularity from images prior to decomposition into reflectance and shading components. We find that the proposed framework is capable in many cases of producing accurate specularity-free reflectance and shading images from single input images. In future work, we plan to improve the specularity removal method and extend our pipeline to handle shadows. Finally, we will explore the use of these methods in the context of real-world video processing systems.

References

1. Land, E.H., McCann, J.J.: Lightness and retinex theory. *J. Opt. Soc. Am.* **61**(1), 1–11 (1971)
2. Rother, C., Kiefel, M., Zhang, L., Schölkopf, B., Gehler, P.V.: Recovering intrinsic images with a global sparsity prior on reflectance. In: *Advances in Neural Information Processing Systems (NIPS)*, pp. 765–773 (2011)
3. Barron, J.T., Malik, J.: Shape, albedo, and illumination from a single image of an unknown object. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 334–341. IEEE (2012)
4. Barron, J.T., Malik, J.: Color constancy, intrinsic images, and shape estimation. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *ECCV 2012*. LNCS, vol. 7575, pp. 57–70. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33765-9_5
5. Laffont, P.Y., Bousseau, A., Drettakis, G.: Rich intrinsic image decomposition of outdoor scenes from multiple views. *IEEE Trans. Vis. Comput. Graph.* **19**(2), 210–224 (2013)
6. Shi, J., Dong, Y., Tong, X., Chen, Y.: Efficient intrinsic image decomposition for RGBD images. In: *ACM Symposium on Virtual Reality Software and Technology*, pp. 17–25. ACM (2015)
7. Kang, X., Li, S., Fang, L., Benediktsson, J.A.: Intrinsic image decomposition for feature extraction of hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **53**(4), 2241–2253 (2015)
8. Shen, J., Yang, X., Li, X., Jia, Y.: Intrinsic image decomposition using optimization and user scribbles. *IEEE Trans. Cybern.* **43**(2), 425–436 (2013)
9. Fan, C., Zhu, H., Lin, G., Cao, L.: Deriving reflectance and shading components from a single image. In: *Fourth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 139–143. IEEE (2013)
10. Tan, R.T., Nishino, K., Ikeuchi, K.: Illumination chromaticity estimation using inverse-intensity chromaticity space. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, p. I-673 (2003)
11. Tan, R.T., Nishino, K., Ikeuchi, K.: Color constancy through inverse-intensity chromaticity space. *JOSA A* **21**(3), 321–334 (2004)
12. Tan, R.T., Ikeuchi, K.: Separating reflection components of textured surfaces using a single image. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(2), 178–193 (2005)

13. Mallick, S.P., Zickler, T., Belhumeur, P., Kriegman, D.: Dichromatic separation: specularity removal and editing. In: ACM SIGGRAPH 2006 Sketches, p. 166. ACM (2006)
14. 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)
15. Carroll, R., Ramamoorthi, R., Agrawala, M.: Illumination decomposition for material recoloring with consistent interreflections. *ACM Trans. Graph.* **30**(4), 43 (2011)
16. Chen, Q., Koltun, V.: A simple model for intrinsic image decomposition with depth cues. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 241–248 (2013)
17. Grosse, R., Johnson, M.K., Adelson, E.H., Freeman, W.T.: Ground truth dataset and baseline evaluations for intrinsic image algorithms. In: IEEE International Conference on Computer Vision (ICCV), pp. 2335–2342. IEEE (2009)