A QUADRATIC OPTIMISATION APPROACH FOR SHADING AND SPECULARITY RECOVERY FROM A SINGLE IMAGE

Lin Gu [1] *Antonio Robles-Kelly*[2,3]

[1]Bioinformatics Institute, A*STAR Singapore, Singapore
[2]NICTA†, Locked Bag 8001, Canberra ACT 2601, Australia
[3]Research School of Engineering, Australian National University, Canberra ACT 0200, Australia

ABSTRACT

In this paper we present a method to recover the shading and specularities in the scene from a single image. The method presented here is based on the dichromatic model and enforces a local smoothness assumption over the object surfaces in the scene. This naturally leads to a setting where the estimate of the shading at a particular pixel can be expressed in terms of its neighbours up to a pair of Gaussian kernels accounting for the irradiance similarity between pixels and their spatial proximity on the image plane. This yields a quadratic cost function for both, the specular coefficient and the shading factor of the dicromatic model which can be solved using gradient descent. We show results for both, specular highlight recovery and shading estimation and compare them against a number of alternatives.

Index Terms— Specularity removal, shape-from-shading, photometric parameter recovery

1. INTRODUCTION

Shading and specularity recovery are important problems in computer vision that has attracted ample attention in the research community. This stems from the notion that, if the specular highlights and the shading are in hand, other photometric parameters can be recovered in a straightforward manner. Futher, the modelling and recovery of photometric parameters are also a topic of pivotal importance for purposes of surface analysis and image understanding. For instance, Nayar and Bolle [1] have used photometric invariants to recognise objects with different reflectance properties. Dror *et al.* [2] have shown how surfaces may be classified from single images through the use of reflectance properties.

Attempts to recover the shape of the object through its shading include the classic approaches of Ikeuchi and Horn [3], and Horn and Brooks [4]. Since the shape-from-shading process is under constrained, methods elsewhere in the literature attempt to solve the problem making use of additional hardware [1], requiring color segmentation as a post-processing step [5] or using reflectance models to account for the distribution of image brightness [6]. Other methods assume additional constraints. For instance, Zheng and Chellappa [7] have imposed a gradient consistency constraint that penalises differences between the image intensity gradient and the brightness gradient for the recovered surface. Worthington and Hancock [8] impose the Lambertian radiance constraint in a hard manner by demanding that the recovered surface normals lie on cones whose axis is in the light source direction and whose apex angle is the inverse cosine of the normalised image radiance.

Regarding specularities, there have been several attempts to remove specular highlights from images of non-Lambertian objects. For instance Brelstaff and Blake [9] used a thresholding strategy to identify specularities on moving curved objects. Narasimhan *et al.* [10] have formulated a scene radiance model for the class of "separable" Bidirectional Reflectance Distribution Functions (BRDFs). More recently, Zickler *et al.* [11] introduced a method for transforming the original RGB colour space into an illuminant-dependent colour space to obtain photometric invariants.

Here, we depart from the dichromatic model so as to describe the image radiance as a combination of shading, specular highlights, surface reflectance and the illuminant power spectrum. We note that, if the illuminant power spectrum is in hand, the estimation of the shading and specularities from a single colour image can be posed as a quadratic optimisation problem. We do this by recovering the shading making use of a smoothness constraint on the scene albedo. With the shading in hand, we proceed to recover the specular highlights making use of an exponential mapping so as to obtain yet another quadratic cost function. In both cases, the optimisation scheme used here is convex, *i.e.* quadratic, and can be solved using gradient descent.

2. SHADING AND SPECULARITY

Recall that, for the pixel u and colour channel $\ell = \{R, G, B\}$, the dichromatic model proposed in [12] is given by

$$I(\ell, u) = g(u)L(\ell)S(\ell, u) + k(u)L(\ell) \quad (1)$$

where $L(\cdot)$ is the power spectrum of the illuminant, $S(\cdot, u)$ is the albedo, the shading factor $g(u)$ governs the proportion of diffuse light reflected from the object and depends solely on the surface geometry and the coefficient $k(u)$ models surface irregularities that cause specularities in the scene.

To recover the shading factor $g(u)$ and specular coefficient $k(u)$, we note that the power spectrum of the illuminant may be obtained using methods elsewhere in the literature, such as White-Patch [13], Grey-World [14] or Grey-Edge [15]. With

†NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

the illuminant power spectrum in hand, we can write

$$R(\ell,u) = \frac{I(\ell,u)}{L(\ell)} = g(u)S(\ell,u) + k(u) \tag{2}$$

2.1. Shading Estimation

For estimating the shading, we make use of the fact that, by subtracting the average $\hat{R}(u) = \sum_\ell \frac{1}{3} R(\cdot,u)$ of $R(\ell,u)$ over the colour channels we can arrive to the expression

$$R(\ell,u) - \hat{R}(u) = \frac{I(\ell,u)}{L(\ell)} - \frac{1}{3}\sum_\ell \frac{I(\ell,u)}{L(\ell)} \tag{3}$$

Moreover, if the pixels u and v have the same albedo, we have that

$$\frac{g(u)}{g(v)} = \frac{R(\ell,u) - \hat{R}(u)}{R(\ell,v) - \hat{R}(v)} \tag{4}$$

this suggests the use of a shading smoothness measure over the neighbourhood $\mathcal{N}_u$ of u. Thus, if two pixels have the same albedo and are close to each other, it is expected that their estimated shading should also be close to one another. As a result, the estimated shading factor is given by

$$g^*(u) = (1-\varrho)g(u) + \varrho\left(\frac{\sum_{v\in\mathcal{N}_u}\varphi(u,v)g(v)}{\sum_{v\in\mathcal{N}_u}\varphi(u,v)}\right) \tag{5}$$

where $0 \leq \varrho \leq 1$ is a pixel-dependent weight that measures the albedo similarity between the pixel u and its neighbours given by

$$\varrho = \exp\left(\frac{-\left|\left|\frac{I(u,\cdot)}{\sum_\ell I(u,\ell)} - \frac{I(u,\cdot)}{\sum_\ell I(v,\ell)}\right|\right|^2}{q}\right) \tag{6}$$

and

$$\varphi(u,v) = \exp\left(\frac{-d^2(u,v)}{r}\right) \tag{7}$$

is a scalar that quantifies the contribution of each neighbour of u to the weighted average on the right-hand side of Equation 5.

Note that, here, both ϱ and $\varphi(u,v)$ are given by Gaussian kernels, where $d^2(u,v)$ is the squared distance between the pixels u and v on the image lattice, $\frac{I(u,\cdot)}{\sum_\ell I(u,\ell)}$ is the chromaticity at pixel u, $||\cdot||^2$ is the vector norm and q, r are bandwidth parameters.

To take our analysis further, we rearrange terms in Equation 4 so as to write

$$g(u)\sum_\ell \frac{1}{R(\ell,u) - \hat{R}(u)} = g(v)\sum_\ell \frac{1}{R(\ell,v) - \hat{R}(v)} \tag{8}$$

and follow Equation 5 in order to arrive to the cost function

$$\min_{\mathbf{g}}\left\{\left| g(u)\sum_\ell (1-\varrho)\frac{1}{R(\ell,u)-\hat{R}(u)} - \varrho\frac{\sum_{v\in\mathcal{N}_u} g(v)\varphi(u,v)\sum_\ell \frac{1}{R(\ell,v)-\hat{R}(v)}}{\sum_{v\in\mathcal{N}_u}\varphi(u,v)}\right|^2\right\} \tag{9}$$

Note that the equation above can be written as a maximisation in a compact form using matrix notation as follows

$$\max_{\mathbf{g}}\{-\mathbf{g}^T[(\mathbf{C}-\mathbf{Q})(\mathbf{C}-\mathbf{Q})^T]\mathbf{g}\} \tag{10}$$

where $\mathbf{g}$ is a vector whose j^{th} entry is given by the shading factor for pixel v, $\mathbf{C}$ is a diagonal matrix whose i^{th} diagonal element is given by $(1-\varrho)\sum_\ell \frac{1}{R(\ell,u)-\hat{R}(u)}$ and $\mathbf{Q}$ is a matrix whose entries indexed i,j are given by $\varrho\frac{\sum_{v\in\mathcal{N}_u}\varphi(u,v)\sum_\ell \frac{1}{R(\ell,v)-\hat{R}(v)}}{\sum_{v\in\mathcal{N}_u}\varphi(u,v)}$.

This notation is important since we can now view the shading factor recovery as that of estimating the entries of the vector $\mathbf{g}$. Further, we can introduce boundary conditions to the process by noting that, for dark pixels, *i.e.* those whose brightness is almost zero, the shading factor should be also zero. To this end, we introduce a Lagrange multiplier λ and optimise the cost function

$$\max_{\mathbf{g}}\{-\mathbf{g}^T[(\mathbf{C}-\mathbf{Q})(\mathbf{C}-\mathbf{Q})^T]\mathbf{g} + \lambda\mathbf{b}^T\mathbf{I}\mathbf{g}\} \tag{11}$$

as an alternative to that in Equation 10. In the equation above, $\mathbf{I}$ is the identity matrix and $\mathbf{b}$ is a vector whose i^{th} entry is zero if the pixel u is dark and unity otherwise.

2.2. Specularity Computation

With the shading factors in hand, we can use a similar formalism to recover the specular coefficient $k(u)$. To do this, we note that, if $g(u)$ is known, we can reorder the terms in Equation 2 and apply an exponential to both sides of the resulting expression so as to write

$$\exp\left(\frac{R(\ell,u)}{g(u)}\right) = \exp\left(S(\ell,u)\right)\exp\left(\frac{k(u)}{g(u)}\right) \tag{12}$$

where by following a rationale akin to that in the previous section, we can then use the shorthand $h(u) = \exp\left(\frac{k(u)}{g(u)}\right)$ so as to arrive to the expression

$$h(u)\sum_\ell \frac{1}{\exp\left(\frac{R(\ell,u)}{g(u)}\right)} = h(v)\sum_\ell \frac{1}{\exp\left(\frac{R(\ell,v)}{g(v)}\right)} \tag{13}$$

which is reminiscent of Equation 8.

Moreover, the assumption over the dark pixels used in the previous section also holds for the specular coefficient $k(u)$, *i.e.* if the pixel is dark, $k(u) \to 0$. Thus, we solve the optimisation problem

$$\max_{\mathbf{h}}\{-\mathbf{h}^T[(\mathbf{D}-\mathbf{T})(\mathbf{D}-\mathbf{T})^T]\mathbf{h} + \lambda\mathbf{b}^T\mathbf{I}\mathbf{h}\} \tag{14}$$

where, similarly to Equation 11, $\mathbf{h}$ is a vector whose i^{th} entry is $h(u)$, $\mathbf{D}$ is a diagonal matrix whose i^{th} diagonal element is given by $(1-\varrho)\sum_\ell \exp\left(\frac{R(\ell,u)}{g(u)}\right)^{-1}$ and $\mathbf{T}$ is a matrix whose entries indexed i,j are given by

$$\varrho\left(\sum_{v\in\mathcal{N}_u}\varphi(u,v)\sum_\ell \exp\left(\frac{R(\ell,v)}{g(v)}\right)^{-1}\right)\left(\sum_{v\in\mathcal{N}_u}\varphi(u,v)\right)^{-1}$$

.

Once the optimal vector $\mathbf{h}$ is in hand, computing the specular coefficient becomes a straightforward matter using the formula

$$k(u) = g(u)\log\left(h(u)\right) \tag{15}$$

which follows from the shorthand used to arrive to Equation 13.

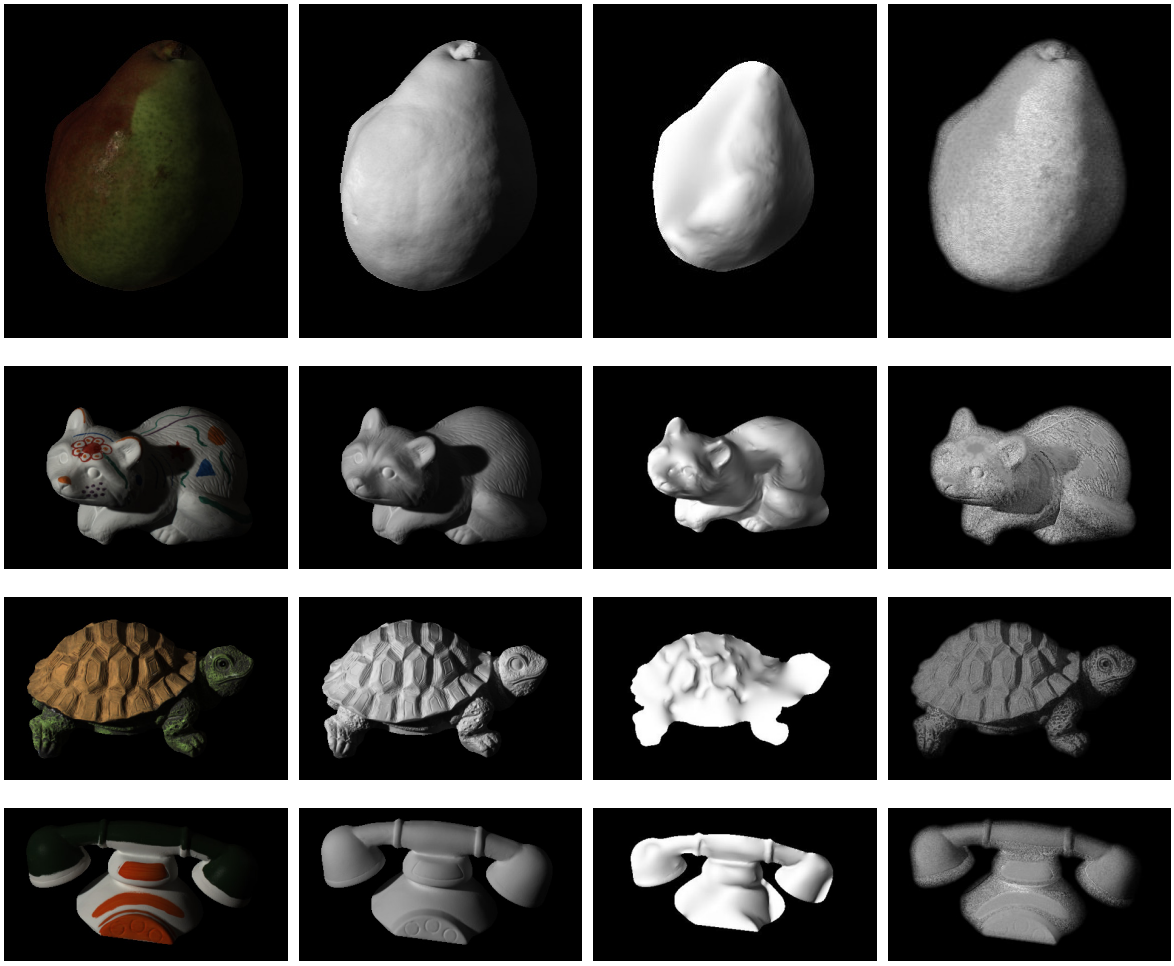

Fig. 1. Shading factor recovery results. From left-to-right: Sample images from the MIT Intrinsic Images dataset, their ground truth, the results yielded by the approach in [16] and our method.

3. RESULTS

Now we turn our attention to the results obtained by our shading and specularity estimation method on both, synthetic and real-world colour imagery. For our shading factor recovery experiments we have used the MIT Intrinsic Images dataset [19] whereas for the specular coefficient computation we have employed imagery from [17]. For all our experiments, we have used 20×20 pixel neighbourhoods, used the Lagrange multiplier $\lambda = 0.5$ and set the bandwidths $q = r = 2$. Here we have used the Grey-World method [14] to recover the illuminant power spectrum and considered a pixel to be a dark one if its normalised brightness is less or equal to 0.1. For the optimisation of the cost functions involved in our computations we use the Biconjugate gradient optimisation method in [20].

Also, note that, in practice, and as a result of the squared distance on the image lattice used in Equation 7, the matrices $\mathbf{Q}$ in Equation 10 and $\mathbf{T}$ in Equation 14 can be rendered sparse by using a threshold τ. This has the benefit of greatly improving the efficiency of the gradient descent method in [20] and can be done without a loss in performance due to the fact that a large number of entries in both matrices will exhibit quite small values. Thus, we have used $\tau = 0.001$ and set all the entries in both matrices which are less or equal to τ to zero.

In Figure 1 we show, from left-to-right, sample input images from the MIT Intrinsic Images dataset, their corresponding ground truth, the results yielded by the approach of Barron and Malik [16] and our method. Note that the alternative tends to oversmooth the object's shading, whereas the detail is better preserved by our method. Moreover, from inspection, we can conclude that the overall structure of the object shape is better conveyed by the shading yielded by our approach. This is particularly evident for the turtle, where the shading delivered by the alternative is devoid of the detail on the shell.

In Figure 2, we illustrate the utility of the specular coefficient computation presented earlier for reflection component separation. In the figure, we show the input real-world images on the top row, whereas the specularity-free images and specular maps are shown in the bottom rows. Here, we have compared our results to those yielded by the method in [17]

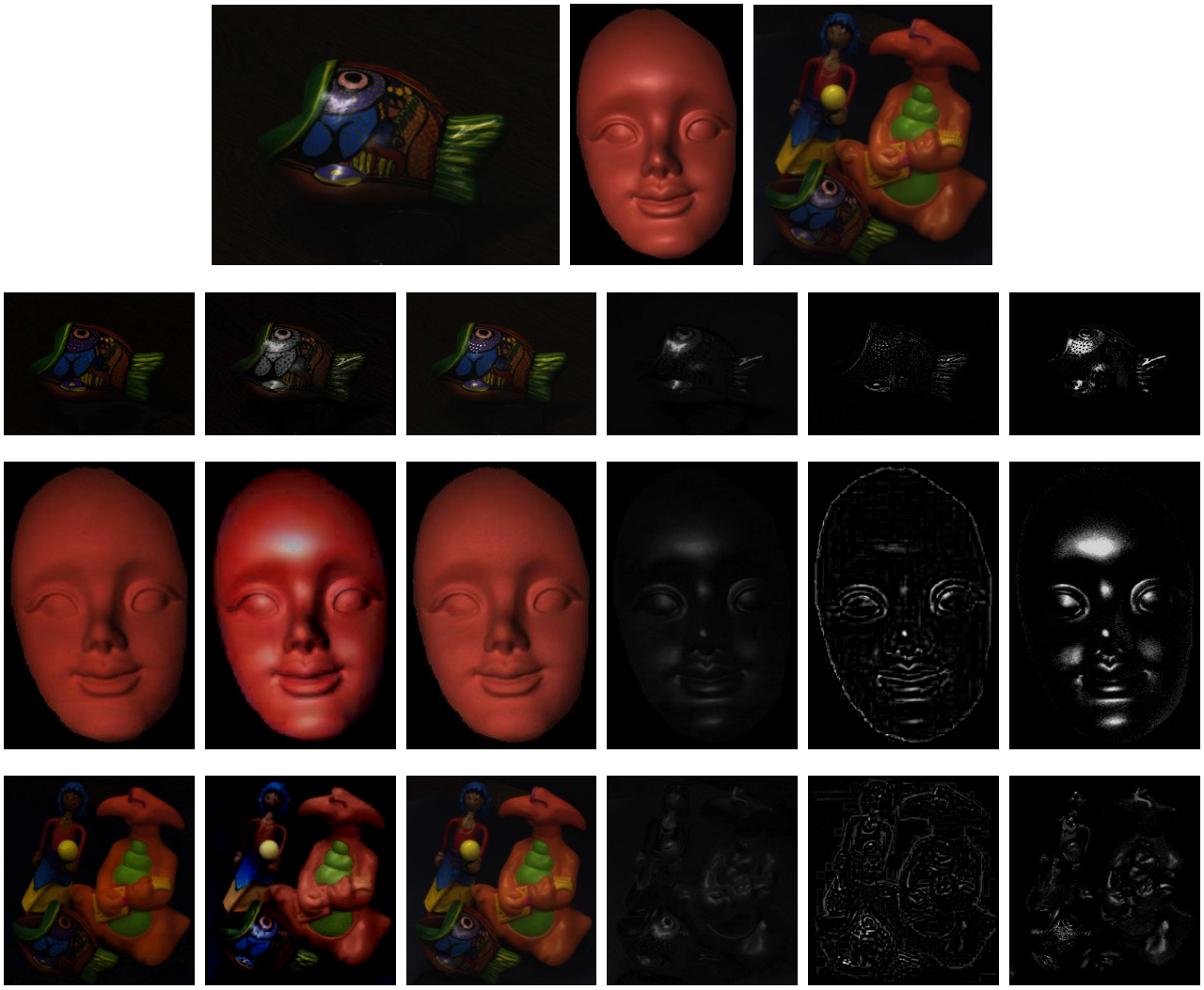

Fig. 2. reflection component separation results. Top row: Input imagery; Bottom rows, from left-to-right: Diffuse component delivered by the method in [17], that yielded by the algorithm in [18], shading recovered by our approach, specular map delivered by the alternatives in [17] and [18] and our algorithm, respectively.

and that in [18]. Note that the specular maps yielded by our approach are in better accordance with the detail of the image than those delivered by the alternatives. This is particularly evident for the mask and the fish. The same can be said of the shading maps. Note that, our method recovers a better estimate of the albedo. This is more evident on the region around the eyes of the fish.

4. CONLUSIONS

In this paper, we have presented a method to estimate the shading factor and specularity components of the dichromatic model from a single colour image. The method presented here benefits from a local smoothness assumption so as to pose the problem in hand in a quadratic optimisation setting. Here, we have tackled the resulting optimisation task using a gradient descent method and have illustrated the use of the developments presented here for purposes of shading and specularity recovery. We have also compared our results against those yielded by alternatives elsewhere in the literature.

5. REFERENCES

[1] S. Nayar and R. Bolle, "Reflectance based object recognition," *International Journal of Computer Vision*, vol. 17, no. 3, pp. 219–240, 1996.

[2] R. O. Dror, E. H. Adelson, and A. S. Willsky, "Recognition of surface reflectance properties from a single image under unknown real-world illumination," in *Proc. of the IEEE Workshop on Identifying Objects Across Variations in Lighting*, 2001.

[3] K. Ikeuchi and B.K.P. Horn, "Numerical shape from

shading and occluding boundaries," *Artificial Intelligence*, vol. 17, no. 1-3, pp. 141–184, 1981.

[4] B. K. P. Horn and M. J. Brooks, "The variational approach to shape from shading," *CVGIP*, vol. 33, no. 2, pp. 174–208, 1986.

[5] G. Klinker, S. Shafer, and T. Kanade, "A physical approach to color image understanding," *Intl. Journal of Computer Vision*, vol. 4, no. 1, pp. 7–38, 1990.

[6] H. Ragheb and E. R. Hancock, "A probabilistic framework for specular shape-from-shading," *Pattern Recognition*, vol. 36, no. 2, pp. 407–427, 2003.

[7] Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo, and shape from shading," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 680–702, 1991.

[8] P. L. Worthington and E. R. Hancock, "New constraints on data-closeness and needle map consistency for shape-from-shading," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1250–1267, 1999.

[9] G. Brelstaff and A. Blake, "Detecting specular reflection using lambertian constraints," in *Int. Conference on Computer Vision*, 1988, pp. 297–302.

[10] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 713–724, 2003.

[11] T. Zickler, S. P. Mallick, D. J. Kriegman, and P. N. Belhumeur, "Color subspaces as photometric invariants," *International Journal of Computer Vision*, vol. 79, no. 1, pp. 13–30, 2008.

[12] Steven A. Shafer, "Using color to separate reflection components," *Color Research & Applications*, vol. 10, no. 4, pp. 210–218, 1985.

[13] E. H. Land and J. J. Mccann, "Lightness and retinex theory," *Journal of the Optical Society of America*, vol. 61, pp. 1–11, 1971.

[14] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 337–350, 1980.

[15] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. on Image Processing*, vol. 16, no. 9, pp. 2207–2214, 2007.

[16] Jonathan T. Barron and Jitendra Malik, "Color constancy, intrinsic images, and shape estimation," *European Conf. on Computer Vision*, 2012.

[17] R. T. Tan, K. Nishino, and K. Ikeuchi, "Separating reflection components based on chromaticity and noise analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1373–1379, 2004.

[18] C. P. Huynh and A. Robles-Kelly, "A solution of the dichromatic model for multispectral photometric invariance," *International Journal of Computer Vision*, vol. 90, no. 1, pp. 1–27, 2010.

[19] R. Grosse, M. K. Johnson, E. H. Adelson, and W. T. Freeman, "Ground truth dataset and baseline evaluations for intrinsic image algorithms," in *Int. Conference on Computer Vision*, 2009.

[20] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, 1994.