

Color from Black and White

BRIAN FUNT and JIAN HO

*Department of Computing Science, Simon Fraser University, Vancouver,
British Columbia, Canada V5A 1S6*

Abstract

Color constancy can be achieved by analyzing the chromatic aberration in an image. Chromatic aberration spatially separates light of different wavelengths and this allows the spectral power distribution of the light to be extracted. This is more information about the light than is registered by the cones of the human visual system or by a color television camera; and, using it, we show how color constancy, the separation of reflectance from illumination, can be achieved. As examples, we consider grey-level images of (a) a colored dot under unknown illumination, and (b) an edge between two differently colored regions under unknown illumination. Our first result is that in principle we can determine completely the spectral power distribution of the reflected light from the dot or, in the case of the color edge, the difference in the spectral power distributions of the light from the two regions. By employing a finite-dimensional linear model of illumination and surface reflectance, we obtain our second result, which is that the spectrum of the reflected light can be uniquely decomposed into a component due to the illuminant and another component due to the surface reflectance. This decomposition provides the complete spectral reflectance function, and hence color, of the surface as well as the spectral power distribution of the illuminant. Up to the limit of the accuracy of the finite-dimensional model, this effectively solves the color constancy problem.

1. Introduction

A central problem with Maloney and Wandell's recent color constancy algorithm is that it requires more than three sensor classes for adequate recovery of surface spectral reflectance [14, 18]. Since it is generally agreed that the human visual system has only three types of receptors active at photopic light levels, some other mechanism for obtaining information about the spectrum of the incoming light must be present if the theory is to hold. In this paper, we show how chromatic aberration can be used to obtain the necessary extra information about the input spectrum, and furthermore, how this information can be applied to the color constancy problem.

Chromatic aberration arises from the fact that the refractive index of a medium depends on wavelength. The focal length of a lens therefore also varies with wavelength, with the result that a single-element lens can only truly be in focus for

one wavelength at a time. Other, out-of-focus wavelengths are almost always present and lead to distortion. In images with serious chromatic aberration, the distortion appears as color bands around the sharp edges between differently colored regions. Usually, chromatic aberration is viewed as a negative feature of an optical system and much effort has been expended on the design of lenses corrected to minimize its effect.

For the human eye, the amount of chromatic aberration has been measured to be approximately 1.82 diopters [9] and so, from a negative standpoint, we can say that the human visual system suffers substantially from its distorting effect. But is this negative view appropriate? Could chromatic aberration, because of its prismatic action that separates light on the basis of wavelength, perhaps in fact benefit the visual system in some way? While we have as yet no evidence to support such a conjecture for human vision, we can at least explore what interesting information

might possibly be extracted from chromatic aberration for use in a visual system, whether human or machine.

Because chromatic aberration is a wavelength-dependent phenomenon, we expect it to yield information concerning the spectrum of the incoming light. This, it turns out, is exactly the kind of information needed to successfully address the color constancy problem. Color constancy generally refers to the ability of humans to make consistent judgements of surface color under a large variation in the spectral character of the incident illumination. For our purposes, the color constancy problem is that of determining from an image of not more than three bands a surface's spectral reflectance function (percentage reflectance as a function of wavelength) to within a multiplicative constant when the spectral energy distribution of the illuminant is unknown.

The main aim of this paper, therefore, is to show how color constancy can be achieved by using spectral information derived from the distortion caused by chromatic aberration. As an example of how this can work, we first examine the problem of determining the spectral reflectance function of a colored spot against a black background under unknown illumination from a single-band, greyscale image. We then generalize to the case of chromatic aberration at the edge between two regions of different color.

2. Extracting the SPD of the Reflected Light

The fact that chromatic aberration can yield information about the spectral power distribution (SPD) of light can be understood intuitively by analogy to regular spectroscopy in which a prism is used to bend different wavelengths of light through different angles. By measuring the intensity of light exiting the prism as a function of angle (or position, if the light is cast onto film), we obtain the complete spectral power distribution function of the light. Both chromatic aberration and prismatic spectral separation arise as a result of the fact that the refractive index of a medium (in this case glass) varies as a function of wavelength. The "image" formed by a prism is just an extreme case of chromatic aberration.

While the situation created by chromatic

aberration is similar to that of the prism, it is by no means as simple. The focused image of a point of monochromatic light is a point; however, only one wavelength can be in focus at once—all other wavelengths will be out of focus. The "focused" image of a point of mixed wavelengths is an amalgam of many, slightly out-of-focus points.

The out-of-focus image of a point is a disk whose diameter varies with the amount the point is out of focus. Because the different wavelengths of the light from the point will each be out of focus to a different degree, the image of a colored point formed under chromatic aberration can be understood as the superposition of a (possibly infinite) set of disk images. Each disk is the out-of-focus image of the point for one particular wavelength of the incoming light. Even though the information from different wavelengths overlaps in the chromatic-aberration image of a point, the regularity of the superposition means the SPD of the incoming light can be extracted nonetheless.

We will first consider the image formation process in detail, deriving an expression for the point-spread function of the optical system possessing chromatic aberration as a function of the SPD of the incoming light. This will enable us to invert the image formation process to obtain the incoming light's SPD.

For monochromatic light, the point-spread function can be calculated as a function of wavelength when the imaging parameters are known. To do so, we employ two well-known, thin-lens results from optics [11, 16]:

$$1/S + 1/S' = 1/f \quad (1)$$

$$1/f = (n'/n - 1)(1/R1 - 1/R2) \quad (2)$$

where

f = focal length of the lens

S = distance from the imaged object to the lens

S' = distance from the lens to the object's image

n = refractive index of external medium

n' = refractive index of lens medium

$R1, R2$ = radii of curvature of the first (the incident) and second lens surfaces

of wavelength λ to the image intensity distribution, $\rho(x)$, is simply $I(\lambda)R(x, \lambda)$. $R(x, \lambda)$ is the one-dimensional linespread function of the isotropic point-spread function, with its wavelength dependence made explicit. Integrating over all wavelengths in the visible spectrum $[\lambda_{\min}, \lambda_{\max}]$ we obtain the intensity at x :

$$\rho(x) = \int_{\lambda_{\min}}^{\lambda_{\max}} I(\lambda)R(x, \lambda) d\lambda \quad (5)$$

Since (5) is an integral equation in which $R(x, \lambda)$ is known and $\rho(x)$ can be measured, we can solve for $I(\lambda)$, the SPD of the incoming light. First we note that $D(\lambda)$ increases monotonically with λ . For a sequence of decreasing wavelengths, $\lambda_{\max} = \lambda_0 > \lambda_1 > \dots > \lambda_n = \lambda_{\min}$, the corresponding sequence of disk diameters decreases, $D(\lambda_0) > D(\lambda_1) > \dots > D(\lambda_n)$. With this subdivision, the integral in (5) becomes a summation which can be written recursively in terms of $I(\lambda)$ as

$$\rho(D(\lambda_0)) = I(\lambda_0) R(D(\lambda_0), \lambda_0) (\lambda_0 - \lambda_1) \quad (6)$$

$$\rho(D(\lambda_k)) = \sum_{i=0}^k I(\lambda_i) R(D(\lambda_k), \lambda_i) (\lambda_i - \lambda_{i+1})$$

The initial condition holds because $R(D(\lambda_a), \lambda_b) = 0$ for all $\lambda_a > \lambda_b$. Rearranging for $I(\lambda)$ we obtain

$$I(\lambda_0) = \rho(D(\lambda_0)) / \{R(D(\lambda_0), \lambda_0) (\lambda_0 - \lambda_1)\} \quad (7)$$

$$I(\lambda_k) = \left\{ \rho(D(\lambda_k)) - \sum_{i=0}^{k-1} I(\lambda_i) R(D(\lambda_k), \lambda_i) (\lambda_i - \lambda_{i+1}) \right\} \div \{R(D(\lambda_k), \lambda_k) [\lambda_k - \lambda_{k+1}]\}$$

This equation can be solved numerically given the image intensity distribution $\rho(x)$. Hence, when light is reflected from a colored dot superimposed on a black background, we can recover the complete SPD, $I(\lambda)$, of the light reflected from it by analyzing the grey-level image intensity distribution, $\rho(x)$, in terms of chromatic aberration.

3. Extracting the SPD Difference Across a Color Edge

Chromatic aberration at a color edge causes the colors to bleed into one another. The mixture of the light from different locations and of different wavelengths is more complex than in the dot case. For brevity, we analyze the case of one-dimen-

sional, thin-line regions, which illustrates all the relevant issues, and will present the two-dimensional case elsewhere [7]. The situation is as shown in figure 3.

Let points above the axis be positive and points below negative. The intensity at a point on the image plane depends on the light emitted from a range of points in the scene. The effect of a scene point in the image is limited by the size of the point-spread function (we assume a spatially invariant psf) for the longest wavelength, $D(\lambda_{\max})$. Therefore, points in region 1 can at the furthest affect points on the image plane for $x \leq D(\lambda_{\max})$. Bear in mind that the lens inverts the image. Similarly, region 2 only contributes to image points for $x \geq -D(\lambda_{\max})$.

Let the SPD of the light reflected from region 1 be $I_1(\lambda)$ and that from region 2 be $I_2(\lambda)$. Let $D_{\max} \equiv D(\lambda_{\max})$. Analogous to the point source case, the following equations describe the image irradiance $\rho(x)$.

$$\text{For } x > D_{\max}, \quad \rho(x) = \int_{\lambda_{\min}}^{\lambda_{\max}} \int_{-D_{\max}}^{D_{\max}} I_2(\lambda) R(s, \lambda) ds d\lambda \quad (8)$$

$$\text{For } x < -D_{\max}, \quad \rho(x) = \int_{\lambda_{\min}}^{\lambda_{\max}} \int_{-D_{\max}}^{D_{\max}} I_1(\lambda) R(s, \lambda) ds d\lambda \quad (9)$$

$$\text{For } -D_{\max} \leq x \leq D_{\max}, \quad \rho(x) = \int_{\lambda_{\min}}^{\lambda_{\max}} \int_{-x}^{D_{\max}} I_2(\lambda) R(s, \lambda) ds d\lambda + \int_{\lambda_{\min}}^{\lambda_{\max}} \int_{-D_{\max}}^{-x} I_1(\lambda) R(s, \lambda) ds d\lambda \quad (10)$$

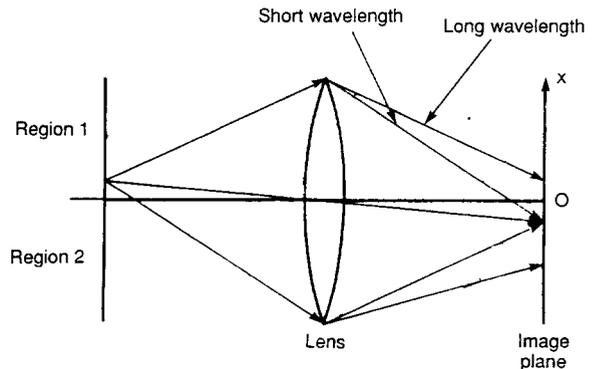


Fig. 3. Image formation for an off-axis point near a color edge.

Equation (8) simply states that for a point above the optical axis and at a distance from the axis greater than D_{\max} , the intensity is due solely to light from region 2. Similarly with (9), for a point $x < -D_{\max}$ the intensity at that point is due solely to light from region 1. For image points within the range $[-D_{\max}, D_{\max}]$, both regions contribute so the integration is broken into two parts. The first integral in (10) represents the light coming from region 2 and the second represents the light coming from region 1.

Let us now consider the simplest case, in which the spread function, $R(x, \lambda)$, is uniform. Since it becomes zero at distance $D(\lambda)$ from its center, its width at wavelength λ is $2D(\lambda)$. Letting $\lambda(D)$ be the inverse function of $D(\lambda)$, the following expression describes the spread function:

$$R(x, \lambda) = \begin{cases} 1/(2D(\lambda)), & \lambda \geq \lambda(x) \\ 0 & \lambda < \lambda(x) \end{cases}$$

Let $U(\lambda) = 1/(2D(\lambda))$ and assume that $I(\lambda)R(x, \lambda)$ is such that the order of integration can be changed, in which case (10) can be rewritten as follows.

$$\begin{aligned} \text{For } -D_{\max} \leq x \leq D_{\max}, \\ \rho(x) &= \int_{-x}^{D_{\max}} \int_{\lambda_{\min}}^{\lambda_{\max}} I_2(\lambda) R(s, \lambda) d\lambda ds \\ &+ \int_{-D_{\max}}^{-x} \int_{\lambda_{\min}}^{\lambda_{\max}} I_1(\lambda) R(s, \lambda) d\lambda ds \quad (11) \\ &= \int_{-x}^{D_{\max}} \int_{\lambda(s)}^{\lambda_{\max}} I_2(\lambda) U(\lambda) d\lambda ds \\ &+ \int_{-D_{\max}}^{-x} \int_{\lambda(s)}^{\lambda_{\max}} I_1(\lambda) U(\lambda) d\lambda ds \end{aligned}$$

Since $\lambda(s) = \lambda(-s)$, the first derivative yields

$$\begin{aligned} \rho'(x) &= \int_{\lambda(x)}^{\lambda_{\max}} I_2(\lambda) U(\lambda) d\lambda \\ &- \int_{\lambda(x)}^{\lambda_{\max}} I_1(\lambda) U(\lambda) d\lambda \quad (12) \end{aligned}$$

Taking the second derivative we have

$$\begin{aligned} \rho''(x) &= -I_2(\lambda(x))U(\lambda(x))\lambda'(x) \\ &+ I_1(\lambda(x))U(\lambda(x))\lambda'(x) \end{aligned}$$

or

$$\begin{aligned} -I_2(\lambda(x)) + I_1(\lambda(x)) = \\ \rho''(x)/[U(\lambda(x))\lambda'(x)] \quad (13) \end{aligned}$$

Substitution of $D(\lambda)$ for x yields

$$\begin{aligned} I_1(\lambda) - I_2(\lambda) = \\ \rho''(D(\lambda))/[U(\lambda)\lambda'(D(\lambda))] \quad (14) \\ = 2\rho''(D(\lambda))D(\lambda)/\lambda'(D(\lambda)) \end{aligned}$$

Chromatic aberration at a color edge, therefore, spatially separates the light in such a way that (14) can be used to determine the *difference* between the spectral power distributions of the light reflected from two regions that meet at the edge. We can also obtain the difference of the spectral power distributions when the spread function is not uniform; however, in that case there is no longer a closed-form solution and the equations must be solved numerically.

As an aside, we note that the properties of the derivatives of an edge's intensity profile as predicted by the chromatic aberration equations provide some further justification for the Marr-Hildreth (15) edge-detection operator. Both the first and second derivatives are symmetric about the edge:

$$\rho'(x) = \rho'(-x)$$

and

$$\rho''(x) = -\rho''(-x) \quad (x \text{ in } [-D_{\max}, D_{\max}])$$

Therefore, $\rho''(0) = 0$, which is consistent with the use of zero-crossings of the second derivative for edge detection.

4. Chromatic Aberration Experiments

A C program was written which numerically solves the chromatic aberration integral equations. As long as the edge intensity profiles are as expected, the calculation of the SPD difference will be correct.

We use the case of a black-to-white edge to illustrate the intensity profiles because then the SPD difference is simply the SPD of the light reflected from the white side. The predicted and experimentally measured intensity profiles for a black-to-white edge under different illuminants are plotted in figures 4 and 5. The three illuminants are tungsten light of 2805° K passed

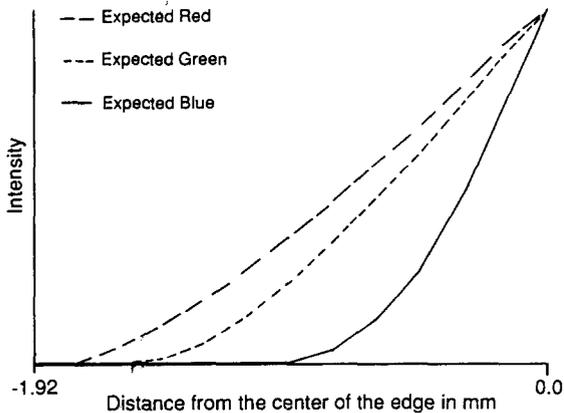


Fig. 4. The expected intensity profiles for a black-to-white edge illuminated with 2805° K light through Kodak filters #25 (red), #58 (green), and #47B (blue).

through red, green, and blue Kodak filters (numbers 25, 58, and 47B).

The experimental curves were obtained by photographing the edge with ILFORD PAN F black and white film and then measuring the film density with a microdensitometer. Generally speaking the predicted and measured curves are quite similar. The slight differences most likely are due to unaccounted for nonlinearities in the film response, error in image-plane distance measurement and imperfections in the lens.

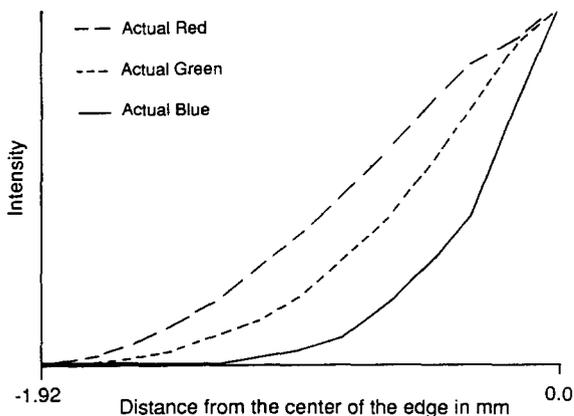


Fig. 5. The actual black-to-white edge profiles of 2805° K light through Kodak filters #25 (red), #58 (green), and #47B (blue).

Clearly, the chromatic aberration effect is present and significant enough to be measured using a good, calibrated camera rather than a microdensitometer since it extends over a distance of almost 2 mm.

5. Extracting Illuminant and Surface Colors

Having shown how spectral information can be obtained from chromatic aberration effects, we must now consider how that information can be used to solve the color constancy problem. Essentially, this means separating the SPD of the illuminant from the surface spectral reflectances. These two components are multiplicatively combined in the light reflected from a surface, but from chromatic aberration we now at least have more information about that light's spectrum than a simple trichromatic signal. We employ finite-dimensional linear models of light and reflectance in the solution.

Maloney [14], following on the work of Cohen [3], showed that the surface spectral reflectances of the large set of natural objects measured by Krinov [12] can be adequately modeled by a finite-dimensional linear model using a set of three basis functions. Similarly, Judd's analysis [10] shows that three to five basis functions suffice in modeling most natural illuminants. Such models work adequately because the spectral power distributions and spectral reflectance functions of natural lights and objects are relatively smooth over the visible spectrum.

Given the SPD of light reflected from a surface where a finite-dimensional linear model adequately characterizes the surface reflectance and the spectrum of the light illuminating it, we show how it can be separated into reflectance and illuminant components. The decomposition is unique up to a multiplicative scale factor.

For the finite-dimensional linear model, we have n basis functions $S_1(\lambda), \dots, S_n(\lambda)$ for reflectance and m basis functions $E_1(\lambda), \dots, E_m(\lambda)$ for illumination. The surface spectral reflectance at a point is expressed as

$$S(\lambda) = \sum_{j=1}^n \sigma_j S_j(\lambda) \quad (15)$$

Similarly, the spectral power distribution of the illumination is:

$$E(\lambda) = \sum_{i=1}^m \epsilon_i E_i(\lambda) \tag{16}$$

For the case of a dot or a black-to-white edge, chromatic aberration provides the SPD, $I(\lambda)$ of the incoming light. Since we are assuming that the light has been reflected from a surface, $I(\lambda) = E(\lambda)S(\lambda)$. In terms of the finite-dimensional linear model,

$$I(\lambda) = \left[\sum_{i=1}^m \epsilon_i E_i(\lambda) \right] \left[\sum_{j=1}^n \sigma_j S_j(\lambda) \right] \tag{17}$$

$I(\lambda)$ is known, so (17) can be used to generate as many equations for the ϵ and σ variables as necessary simply by substituting specific values for λ . The actual number of these that will be independent, however, depends on the dimension of the set of functions $E_i(\lambda)S_j(\lambda)$.

In the optimal situation, all the functions $E_i(\lambda)S_j(\lambda)$ are linearly independent. While it might seem surprising, the first three of Cohen's reflectance-basis functions [3] combined with the first three of Judd's illuminant basis functions [10] form just such an independent set. In this case, $m = n = 3$ and (17) yields 9 independent linear equations in 9 $\epsilon_i \sigma_j$ product-pair unknowns.

Since the ϵ_i and σ_j unknowns appear in products, we can solve for them only up to a multiplicative constant. To do so, we constrain one component of the ϵ vector to unit length. In other words, the colors of the surface and the illuminant will be uniquely determined, but not their relative brightness. Letting $c_{ij} \equiv \epsilon_i \sigma_j$, we have $\epsilon_i = (c_{ij}/c_{1j}) \epsilon_1$, for $i > 1$. Adding the illumination brightness constraint, $\epsilon_1 = 1$, yields the ϵ_i 's and hence also the σ_j 's.

Errors can arise when the finite-dimensional model does not model the illumination and reflectance well. In this case, a minimum-error solution can be obtained. The details of this method will be described elsewhere [6], but figures 6, 7, and 8 show an example of the decomposition obtained when the incoming light is a combination of the SPD of daylight and the spectral reflectance of heather (dense growth before flowering) as measured by Krinov [12]. The spectrum in figure 6 was generated by multiplying the

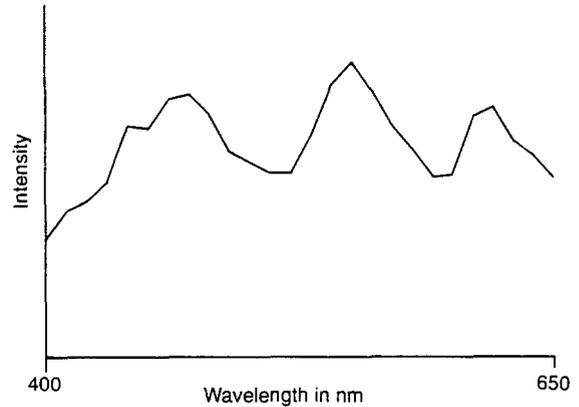


Fig. 6. Spectral power distribution of light reflected from heather illuminated by 10000° K daylight.

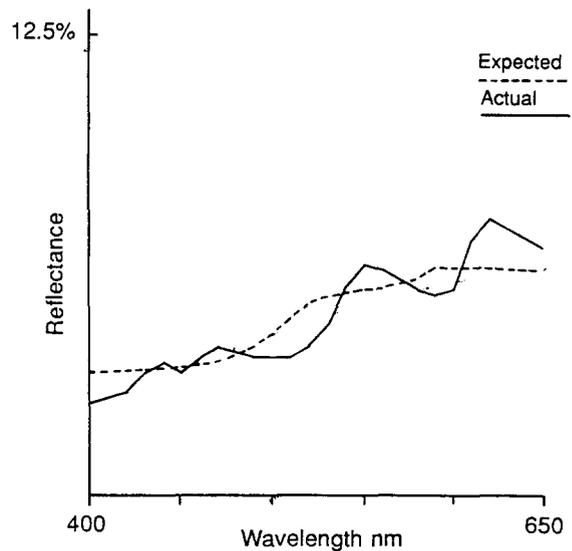


Fig. 7. Krinov's actual surface spectral reflectance for heather and the estimated reflectance found by decomposing the reflected light into reflectance and illumination components.

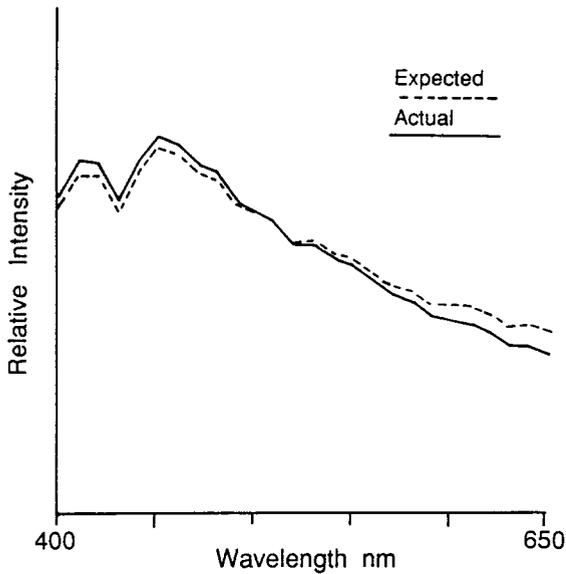


Fig. 8. Judd's actual 10000° K daylight spectrum and that estimated by separating the reflected light into reflectance and illumination components.

actual spectral reflectance of heather given in figure 7 by the actual SPD of daylight given in 8. It is the only input to the program that separates the light into illumination and reflectance components.

6. Conclusion

We have shown that spectral information can be extracted from chromatic aberration effects. This information can be used to address the most serious issue in color constancy, what might be called the "white-patch problem." This is the problem of determining either the illumination or reflectance properties of the scene for at least one image point. Land's [13] "retinex" theory, for example, simply assumes that the whitest point is in fact white, which dictates our choice of a name, but this heuristic is often violated [2]. Maloney [14] and Gershon [5] propose more sophisticated solutions. Once the color of one scene point is established, the colors of all other points can be calibrated in terms of it [4].

For the case of a colored spot under unknown illumination or a scene in which there is one

black-to-color edge, we have shown that the white-patch problem can be solved completely, since the illumination and reflectance can be separated out given the SPD of the reflected light and this in turn can be extracted using chromatic aberration. Surprisingly, the method relies only on broad-band intensity data and does not require a multiband color image. For the case in which two regions of different color (neither being black) meet at an edge, a generalization of the above method [7] that uses multiband data in addition to the chromatic aberration information provides color constancy for both regions.

Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council of Canada grant 4322. The second author is also grateful for the financial assistance provided by his C.D. Nelson Memorial Scholarship.

References

1. Max Born and Emil Wolf, *Principles of Optics*. 6th Pergamon Press: Elmsford, NY, 1980.
2. D. Brainard and B. Wandell, "An analysis of the retinex theory of color vision," *J. Opt. Soc. Am. A* 3:1651-1661, 1986.
3. J. Cohen, "Dependency of the spectral reflectance curves of the Munsell color Chips," *Psychon. Sci.* 1:369-370, 1964.
4. B. Funt and M. Drew, "Color constancy computation in near-mondrian scenes using a finite dimensional linear model." *Proc. IEEE Conf. Comp. Vision Pattern Recog.*, Ann Arbor, June 1988.
5. R. Gershon, A.D. Jepson, and J. Tsotsos, "From {R,G,B} to surface reflectance: computing color constant descriptors in images," *Proc 10th Intern. Joint Conf. Artif. Intell.* Milan, 1987.
6. J. Ho, B. Funt, and M. Drew, "Disambiguation of illumination and surface reflectance from spectral power distribution of color signal: Theory and applications." Tech. Rep. CSS/LCCR TR 88-18, Simon Fraser University, 1988.
7. J. Ho and B. Funt, "Color constancy from chromatic aberration," Tech. Rep. CSS/LCCR TR 88-17, Simon Fraser University, 1988.
8. H.H. Hopkins, "The frequency response of a defocused optical system," *Proc. Roy. Soc. (London)* A231:91-103, 1955.

9. Peter Alan Howarth and Arthur Bradley, "The longitudinal chromatic aberration of the human eye, and its correction," *Vision Research* 26:361-366, 1986.
10. Deane B. Judd, David L. MacAdam, and Gunter Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," *J. Opt. Soc. Am.* 54(8): 1031-1040, 1964.
11. Miles V. Klein, *Optics*. Wiley: New York, 1970.
12. E.L. Krinov, "Spectral reflectance properties of natural formations," *Tech. Trans.* TT-439, National Research Council of Canada, 1947.
13. E.H. Land, "The retinex theory of color vision," *Proc. Roy. Soc. (London)* B207, 1980.
14. Laurence T. Maloney, "Computational approaches to color constancy." Ph.D. thesis, Stanford University, 1985.
15. D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. (London)* B207, 1980.
16. Jurgen R. Meyer-Arendt, *Introduction to Classical and Modern Optics*. Prentice-Hall: Englewood Cliffs, NJ, 1972.
17. P.A. Stokseth, "Properties of a defocused optical system," *J. Opt. Soc. Am.* 59:1314-1321, 1969.
18. Brian A. Wandell, "The synthesis and analysis of color images," *IEEE Trans. PAMI* 9(1):2-13, 1987.