MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I.Memo No. 1068 August, 1988

Estimating the Illuminant Color from the Shading of a Smooth Surface

Hsien-Che Lee

ABSTRACT

The image of a uniform wall illuminated by a spot light often gives a strong impression of the illuminant color. How can it be possible to know if it is a white wall illuminated by yellow light or a yellow wall illuminated by white light? If the wall is a Lambertian reflector, it would not be possible to tell the difference. However, in the real world, some amount of specular reflection is almost always present. In this memo, it is shown that the computation is possible in most practical cases.

Light reflection from a surface is usually modeled as having two components: the interface (specular) reflection and the body (diffuse) reflection. For a surface of inhomogeneous material, the spectral composition of the interface reflection component is often similar to that of the illuminant. The problem of computing the illuminant chromaticity from the shading of a single smooth surface is to separate these two components. An image of an illuminated uniform wall, according to the above model, gives only one physical constraint about the illuminant chromaticity, not enough to determine a unique solution. However, since the spatial scale over which the interface reflection changes significantly is much smaller than that of the body reflection, it can be shown that one can effectively exploit the scale difference to find a unique solution, which is often very accurate. The method can also be generalized to compute the illuminant chromaticity for a nonuniform smooth surface.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124. The author was supported by Eastman Kodak Company.

1 Introduction

When we walk through many rooms inside a building with different lightings in various locations, we are always aware of the change of the illuminant color. Although there are few quantitative measurements available, our experiences tell us that the human visual system seems to be able to perceive qualitatively the scene illuminant quite well. Even when we have difficulty judging the "true" color of a piece of fabrics under certain indoor lighting (an illustration of the breakdown of color constancy), we seldom fail to tell the color of the illuminant even without looking at it directly. This effortless perception of scene illumination does not depend on stereo or motion, as our experience in watching projection slides of still natural scenes can tell us. It does seem to depend on the presence of gradual shadings on the object surfaces. It is well known that color patches of uniform chromaticity and luminance do not give the perception of illumination. This is called the aperture (or film) mode of color perception. In order to give an impression of illumination, the surface shading has to have gradual variations. This is called the surface mode of color perception, that is, the perceived color seems to belong to the object surface [7].

Computing the scene illuminant color from a given color image is not a simple problem. The difficulty is that the recorded color image irradiances are functions of the illuminant, the surface shape, and the surface reflectances. Without knowing any two of them, there are infinite possible solutions (a mathematically ill-posed problem [9]). Recent work [5] [12] suggests that the specular (interface reflection) component of surface reflection can be used to compute the illuminant chromaticity. Since this method for computing the illuminant color is based on the idea of finding the converging point of the surface chromaticity loci, we will call it the chromaticity convergence method [5], which is explained in the next section.

2 The Reflection Model and the Chromaticity Diagram

The general light reflection model of a uniform surface for a three color imaging system used in the chromaticity convergence method is as follows (see [6] for details). Assume that the incident radiance on a surface can be written as:

$$L(\theta_i, \phi_i, \lambda) = c(\lambda) L_0(\theta_i, \phi_i) \tag{1}$$

with $c(\lambda)$ normalized to one at its maximum, and θ_i and ϕ_i are the incident angles. Let $R_r(\lambda)$, $R_g(\lambda)$, and $R_b(\lambda)$ be the spectral responsivity functions of the color imaging system, and

$$\begin{aligned} L_r &= \int c(\lambda) R_r(\lambda) d\lambda \\ L_g &= \int c(\lambda) R_g(\lambda) d\lambda \\ L_b &= \int c(\lambda) R_b(\lambda) d\lambda, \end{aligned} \tag{2}$$

then

$$E_r(x,y) = k L_r (\rho_r f(x,y) + \rho_s h(x,y))$$

$$
\begin{aligned}
E_g(x,y) &= kL_g(\rho_g f(x,y) + \rho_s h(x,y)) \\
E_b(x,y) &= kL_b(\rho_b f(x,y) + \rho_s h(x,y)),
\end{aligned} \tag{3}
$$

where x and y are the spatial coordinates on the image plane, and r, g, and b indicate the conventional red, green, and blue channels or any other combinations of them. E_r, E_g, and E_b are the red, green, and blue image irradiances, ρ_r, ρ_g, and ρ_b are the diffuse (body) reflectance factors, ρ_s represents the specular (interface) reflectance factor, and $f(x,y)$ and $h(x,y)$ represent the factors dependent on the imaging geometry and surface shapes. This reflection model makes a strong assumption that the reflectance factor of the interface (specular) reflection component is independent of the imaging geometry as well as the spectral responsivity of each color channel. We will call this reflection model the neutral interface reflection (NIR) model. This assumption, of course, is not strictly true in practice (especially for homogeneous material, as discussed in detail by Cook and Torrance [1]), but there are experimental data showing that it is a reasonable approximation for many types of surface material [6]. Although the NIR model is a special case of a more general model, the dichromatic reflection model, described by Shafer [10], which does not assume that the interface reflection is neutral, it is often used in real applications, because of its good approximation and simplicity.

In general, one can not recover the absolute magnitudes of both the light source intensity and the reflectance factor at the same time from the image irradiance signal alone, because if one raises the intensity by a factor of 2 and reduces the reflectance factor by a factor of 2, the image irradiance will remain the same. Therefore, it is useful to define quantities which would specify the color of the light, independent of its intensity. For this purpose, the chromaticity of a given beam of light irradiating the image plane is defined by the following ratios [11] (The standard notations for CIE chromaticity coordinates are x, y, and z. To avoid confusion with the spatial coordinates, we will use u, v, and w in the text, and use x, y, z in the figures where CIE chromaticity diagrams are used.):

$$
\begin{aligned}
u &= E_r/(E_r + E_g + E_b), \\
v &= E_g/(E_r + E_g + E_b), \\
w &= E_b/(E_r + E_g + E_b).
\end{aligned} \tag{4}
$$

Since $u + v + w = 1$, one needs only u and v to specify the chromaticity coordinates of the light. It is also easy to see that they are independent of the light intensity. One useful property of the chromaticity coordinates is that if one additively mixes two lights to produce the third light, then the chromaticity coordinates of the third light is a linear combination of the chromaticity coordinates of the first two lights. Let (u_1, v_1) and (u_2, v_2) be the chromaticity coordinates of the first two lights, and let the "total irradiance" $E = E_r + E_g + E_b$ of the first light be E_1 and that of the second light be E_2, then it can be easily shown that the chromaticity coordinates of the mixed light are:

$$
\begin{aligned}
u_3 &= \left(\frac{E_1}{E_1 + E_2}\right)u_1 + \left(\frac{E_2}{E_1 + E_2}\right)u_2, \\
v_3 &= \left(\frac{E_1}{E_1 + E_2}\right)v_1 + \left(\frac{E_2}{E_1 + E_2}\right)v_2,
\end{aligned} \tag{5}
$$

The important consequence of this property for the NIR model is that *since the light reflected from a uniform surface is the additive mixture of the interface reflection component and the body reflection component, its chromaticity locus will fall on a straight line segment connecting the chromaticity of the interface reflection and the chromaticity of the body reflection.* Therefore, each uniform surface will show up in the chromaticity space as a

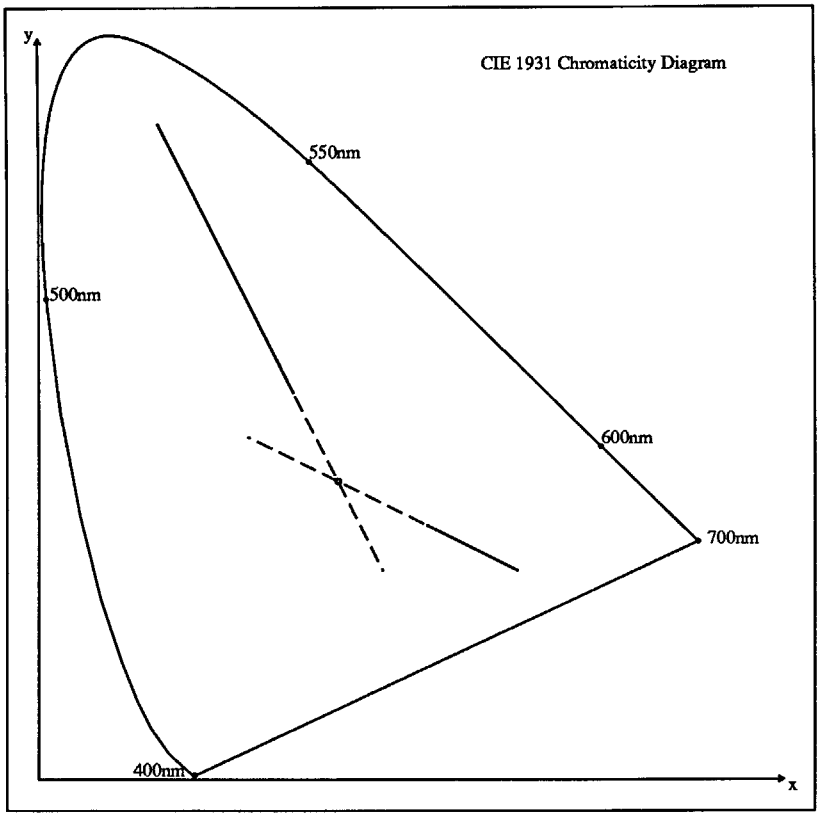

Figure 1: The chromaticity convergence method requires two surfaces of different colors.

short line segment. If there are two uniform surfaces of different colors, the intersection point is the chromaticity locus of the illuminant (see Figure 1). This is the essence of the chromaticity convergence method for determining the illuminant color.

3 Can One Compute the Illuminant Color From A Single Surface?

The theory behind the chromaticity convergence method requires at least two surfaces of different colors be present in the scene in order to compute the illuminant chromaticity.

The two surfaces have to reflect the illuminating light as mixtures of the specular and the diffuse components in various proportions.

However, if one looks at a uniform wall illuminated by a spot light, one has a strong impression not only of the shading variations, but also of the color of the light source. How can it be possible for us to perceive that it is a white wall illuminated by yellow light, but not a yellow wall illuminated by white light? Since there is only one surface in the image, the theory of the chromaticity convergence method tells us that the light source chromaticity is constrained to be along a straight line in the chromaticity diagram, having infinitely many possible solutions.

Some thought will lead us to the following two possible solutions:

1. In the evolution process, the visual system had a long period of time to learn about the regularity of the color variations in the natural illuminants, e.g., sunlight, skylight, and fire. From early morning to late afternoon, daylight (sunlight plus skylight) continues to change its color, depending on the solar angle, the clouds, and the water vapor content of the air mass. The chromaticities of daylight in various phases of the day and various amount of cloudiness had been systematically measured [3]. Their loci on the CIE 1931 chromaticity diagram is a fairly smooth curve, almost parallel to the chromaticity loci of blackbody radiation at various temperatures. Furthermore, many man-made light sources have similar regularity in their chromaticity distributions. Partially due to this reason, color temperature has been used to specify the light source color [11]. Figure 2 shows the CIE daylight locus and the chromaticity locus of the CIE standard illuminant A. If this chromaticity curve is used as a constraint, then the straight line chromaticity loci of the light reflected from a single surface can be extended to intersect with this illuminant curve, yielding a unique solution. If the true illuminant chromaticity is not on the illuminant curve, the computed solution will be incorrect. Therefore, this solution, although biologically feasible, is not physically a good solution.

2. The chromaticity convergence method needs two surfaces because it does not assume any knowledge about surface shape, lighting geometry, or physical characteristics other than those explicitly expressed in the NIR model. For example, the specular reflection is more angle dependent than the diffuse reflection. If our visual system is aware of the difference, it may be able to compute the solution uniquely. As we will show later, this is indeed possible.

4 The Solution from the Smoothness Constraint

If we inspect the image irradiance equations 3 for a uniform surface, we will discover that by subtracting a properly scaled green irradiance signal from the red irradiance signal we can completely cancel spatially either the interface reflection component or the body reflection component. For example,

$$E_r(x,y) - (\frac{L_r}{L_g})E_g(x,y) \quad = \quad kL_r(\rho_r - \rho_g)f(x,y) \tag{6}$$

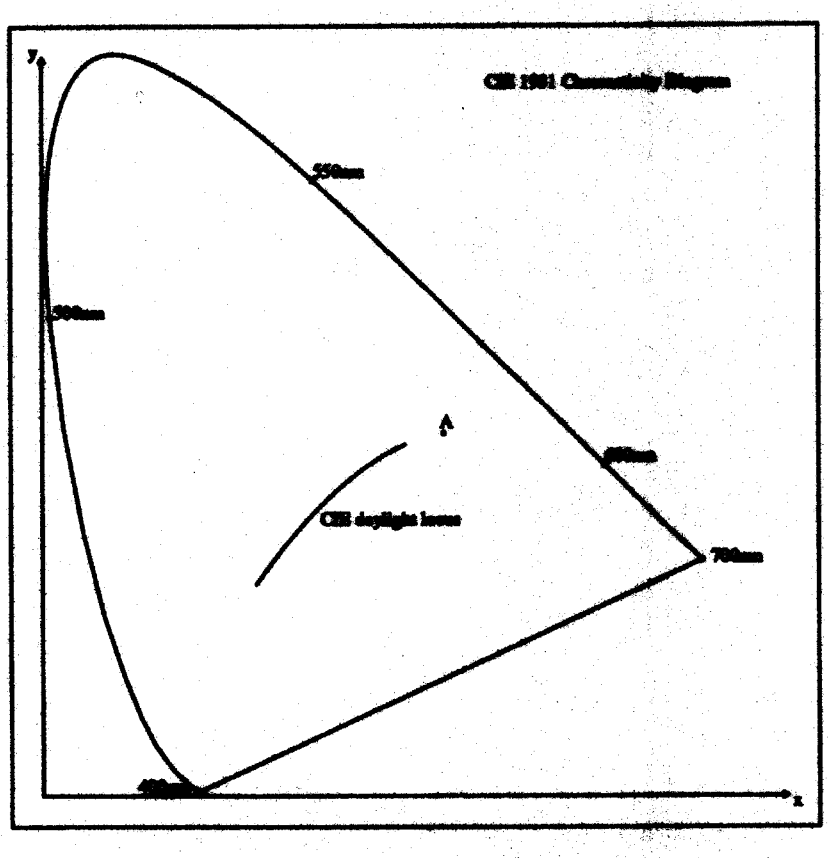

Figure 2: The chromaticity curve of CIE daylight locus.

$$E_r(x,y) - (\frac{L_r\rho_r}{L_g\rho_g})E_g(x,y) = kL_r\rho_s(1-\frac{\rho_r}{\rho_g})h(x,y). \quad (7)$$

The important implication is that if we can find the proper factors, we can completely recover the shapes of the two reflection components, and the proper factors happen to determine the chromaticity of the illuminant. It is also true that the difference signal need not be from the red and the green signals. Any other combination of the three image irradiance signals will also serve well. If we want the factor to come out as the chromaticity of the illuminant, we can define:

$$\begin{aligned} E(x,y) &= E_r(x,y) + E_g(x,y) + E_b(x,y) \\ &= kL(\rho f(x,y) + \rho_s h(x,y)), \end{aligned} \quad (8)$$

where

$$\begin{aligned} L &= L_r + L_g + L_b, \\ \rho &= (L_r\rho_r + L_g\rho_g + L_b\rho_b)/L. \end{aligned} \quad (9)$$

Now, the difference signals become:

$$\begin{aligned} E_r(x,y) - (\frac{L_r}{L})E(x,y) &= kL_r(\rho_r - \rho)f(x,y) \\ E_g(x,y) - (\frac{L_g}{L})E(x,y) &= kL_g(\rho_g - \rho)f(x,y) \\ E_b(x,y) - (\frac{L_b}{L})E(x,y) &= kL_b(\rho_b - \rho)f(x,y) \end{aligned} \quad (10)$$

and the factors $\frac{L_r}{L}$, $\frac{L_g}{L}$, and $\frac{L_b}{L}$ are precisely the chromaticity coordinates of the illuminant. Similar expressions can be written for the $h(x,y)$ component:

$$\begin{aligned} E_r(x,y) - (\frac{L_r\rho_r}{L\rho})E(x,y) &= kL_r\rho_s(1-\frac{\rho_r}{\rho})h(x,y) \\ E_g(x,y) - (\frac{L_g\rho_g}{L\rho})E(x,y) &= kL_g\rho_s(1-\frac{\rho_g}{\rho})h(x,y) \\ E_b(x,y) - (\frac{L_b\rho_b}{L\rho})E(x,y) &= kL_b\rho_s(1-\frac{\rho_b}{\rho})h(x,y) \end{aligned} \quad (11)$$

Now we know that we can recover the shape of one of the reflection components by computing $E_r - s_rE$, if we can select s_r "properly". The remaining question is how to define what we mean by "properly". If we have some prior knowledge about the functions $f(x,y)$ or $h(x,y)$, then we can continue changing the factor s_r until the resulting difference signal $E_r - s_rE$ behaves the way we know it should. The knowledge need not be precise or even quantitative. It serves only as a criterion for selecting the right s_r. One obvious choice for the needed constraint is to require that the reflection component $f(x,y)$ be the smoothest function among all the possible alternatives in the family of functions generated by $E_r - s_rE$. The choice does have a good physical basis, because when the two reflection components, each having its own shape, are mixed together, the combined function is not as smooth as the smoother component alone. This is true because of the following two reasons:

1. The spatial scale at which the irradiance change occurs is, in general, much smaller for the specular (interface) component than for the diffuse (body) component. The diffuse component as modeled by a Lambertian surface varies as a function of cosine, while the specular component is typically modeled as the cosine function raised to the 20th or 40th power. Therefore, of the two reflection components the diffuse component $f(x, y)$ is almost always the smoother one.

2. The peak of the diffuse component rarely occurs at the same place as that of the specular component, because the former is at the place where the surface normal is pointing to to the light source, while the latter occurs at the place where the surface normal is approximately the bisector of the angle between the source vector and the viewing vector. Any slight mixture of the specular component $h(x, y)$ with the diffuse component $f(x, y)$ will create an extra "bump" in the irradiance signal. Therefore, if the factor s_r is not selected properly, the difference signal $E_r - s_r E$ will not be as smooth as when it is.

One possible violation of the first condition is when the light source is very close to the illuminated surface, and therefore creates a sharply changing $f(x, y)$ because of the inverse square fall-off of the light intensity, and because of the large change of the incident angle within a short distance. Also, a surface with large curvatures has similar effect. Another possibility is a microscopically (in terms of the resolution of the imaging system) very rough surface, making the specular component very non-directional. Violation of the second condition can happen under the single light source assumption only when the light source is located on the optical axis of the imaging system, a physically unlikely situation. With multiple light sources and mutual illumination among surfaces, the second condition can be violated, but not frequently. In most practical cases, the two types of violation are rare, and the smoothness constraint should give a fairly reasonable answer.

It is important to understand that the smoothness constraint here refers to the assumption that the body reflection function, $f(x, y)$, is much smoother than the interface reflection, $h(x, y)$. It does not imply that the underlying surface has to be macroscopically smooth. Because the relations between the surface shape and the shape of the image irradiance can be very complicated, smoothness in the shape of the image irradiance signal does not imply smoothness in the surface shape, and vice versa.

It seems that equations 10 and 11 would give two "good" solutions for selecting the "proper" s_r. However, as a consequence of the smoothness constraint, the solution in equation 10 is chosen most of the times, because the body reflection component $f(x, y)$ is almost always smoother than the interface reflection component $h(x, y)$.

The next question is how to define smoothness. For computational reasons, we choose the laplacian operator for its symmetry. It is very likely that another choice of measure of smoothness could give equally good or better results. We can now put the problem in the following mathematical form: select s_r and s_g such that the following are minimized:

$$\int [\nabla^2 (E_r(x, y) - s_r E(x, y))]^2 dx dy,$$
$$\int [\nabla^2 (E_g(x, y) - s_g E(x, y))]^2 dx dy. \tag{12}$$

The solution turns out to be quite simple:

$$\begin{aligned} s_r &= \frac{\int \nabla^2 E_r(x,y) \nabla^2 E(x,y) dxdy}{\int \nabla^2 E(x,y) \nabla^2 E(x,y) dxdy} \\ s_g &= \frac{\int \nabla^2 E_g(x,y) \nabla^2 E(x,y) dxdy}{\int \nabla^2 E(x,y) \nabla^2 E(x,y) dxdy}. \end{aligned} \tag{13}$$

The solution given by 13 has the following nice property. As we have seen in Section 2, the chromaticity of the light reflected from a uniform surface falls on a straight line pointing toward the illuminant chromaticity. Let the straight line equation be: $au + bv = c$, i.e.,

$$a(\frac{E_r}{E}) + b(\frac{E_g}{E}) = c, \tag{14}$$

or,

$$a(E_r) + b(E_g) = c(E). \tag{15}$$

It can be shown that the solution given by 13 is also on the same line, and therefore it satisfies the (only) constraint which can be derived from physics without making assumptions about the surface shape and imaging geometry. That means the solution is guaranteed to be a feasible solution. This is true because the solution in 13 is still a linear combination of the irradiance signals. The proof is as follows:

$$\begin{aligned} a(s_r) + b(s_g) &= a(\frac{\int \nabla^2 E_r \nabla^2 E dxdy}{\int \nabla^2 E \nabla^2 E dxdy}) + b(\frac{\int \nabla^2 E_g \nabla^2 E dxdy}{\int \nabla^2 E \nabla^2 E dxdy}) \\ &= \frac{\int \nabla^2 (aE_r + bE_g) \nabla^2 E dxdy}{\int \nabla^2 E \nabla^2 E dxdy} \\ &= \frac{\int \nabla^2 (cE) \nabla^2 E dxdy}{\int \nabla^2 E \nabla^2 E dxdy} \\ &= c. \end{aligned}$$

If we expand the solution 13 in terms of the function $f(x,y)$ and $h(x,y)$, we get the following expressions:

$$\begin{aligned} s_r &= (\frac{L_r}{L}) \frac{\int (\rho_s^2 (\nabla^2 h)^2 + \rho_r \rho (\nabla^2 f)^2 + \rho_s (\rho_r + \rho) \nabla^2 h \nabla^2 f) dxdy}{\int (\rho_s^2 (\nabla^2 h)^2 + \rho^2 (\nabla^2 f)^2 + 2\rho_s \rho \nabla^2 h \nabla^2 f) dxdy}, \\ s_g &= (\frac{L_g}{L}) \frac{\int (\rho_s^2 (\nabla^2 h)^2 + \rho_g \rho (\nabla^2 f)^2 + \rho_s (\rho_g + \rho) \nabla^2 h \nabla^2 f) dxdy}{\int (\rho_s^2 (\nabla^2 h)^2 + \rho^2 (\nabla^2 f)^2 + 2\rho_s \rho \nabla^2 h \nabla^2 f) dxdy}. \end{aligned} \tag{16}$$

They show that the chromaticity estimation is accurate only when the term $\rho_s^2 (\nabla^2 h(x,y))^2$ is much larger than the sum of the other terms, and this suggests methods to improve the accuracy of the estimate. For example, we can give more weight to the regions of large irradiance because the specular reflection region is often brighter than the neighboring regions. Another way is to compute the rate of chromaticity change at each pixel and use it as a weighting function, because only when the specular reflection is significant in magnitude relative to the diffuse reflection would it cause a noticeable chromaticity change. The following section will show how well the estimator works, and when it could fail.

5 Estimating the Illuminant Chromaticity

5.1 From the Shading of A Uniform Surface

The estimator derived in the last section allows us to estimate only the illuminant chromaticity. In order to show how well the estimator works, we also estimate the chromaticity of the body reflection component, so that we can plot both components at the same time. For a uniform surface, the algorithm computes the chromaticity distribution of all the pixels. Because this is a uniform surface, the chromaticity loci is a straight line segment in the chromaticity space. To compute the chromaticity of the body reflection component, the algorithm simply takes the end point of the line segment that is farthest away from the estimated illuminant chromaticity. This is equivalent to the assumption that the pixel which has the most "saturated" color has only a negligible amount of interface reflection, compared with its body reflection. Knowing the chromaticities of both reflection components, we can separate the two components by simply projecting the red, green, and blue irradiance signals to the two vectors representing the reflection components in the (E_r, E_g, E_b) space.

Separation of reflection components had been studied by by Klinker, Shafer, and Kanade [4], and by Gershon, Jepson, and Tsotsos [2]. The former group searches for a "skewed T" signature, while the latter searches for a "dog leg" distribution in three-dimensional color space. Both methods require that the specular reflection component rise and fall in a very short spatial distance within which the diffuse component is essentially constant. Although the illuminant chromaticity estimator developed here is based on the same idea that the rate of change in the specular reflection component is larger than the diffuse component, the requirement is less stringent because the estimator does not explicitly search for signatures or patterns.

Let us start with experiments on one dimensional signals. A point light source is positioned in front of a plane. The plane is tilted at an angle of θ with respect to the optical axis of the camera (see Fig. 3). The spatial coordinate system has the origin at the center of the image plane. The z axis, $x = 0$, corresponds to the optical axis.The focal point is located at $(x = 0, z = f)$. The object plane intersects with the z axis at $z = -d$. The light source position (xs, zs) and the angle θ are changed to produce different irradiance signals on the image plane. The focal length f is chosen to be 50mm. The distance from the image plane to the center of the surface is 395cm. The light source chromaticity is (0.3333, 0.3333), and the surface reflectance factors (ρ_r, ρ_g, ρ_b) are (1.0, 0.8, 0.4). The specular reflection is calculated according to Phong's model [8] with the cosine raised to the 20th power. The specular reflectance factor ρ_s is 2.0. Two values for the angle θ are used: 50 and 38 degrees.

Figures 4, 5, 6, and 7 show the results of the estimation for different light source positions. The top graphs show the input image irradiance signal from the red channel. The dotted curves are the true reflection components, $f(x)$, and $h(x)$. The bottom graphs show the estimated illuminant chromaticity, and the two estimated reflection components.

As can be seen, the estimate is, in general, very accurate. The performance begins to deteriorate when the light source gets very close to the camera (see Fig. 5) or when the

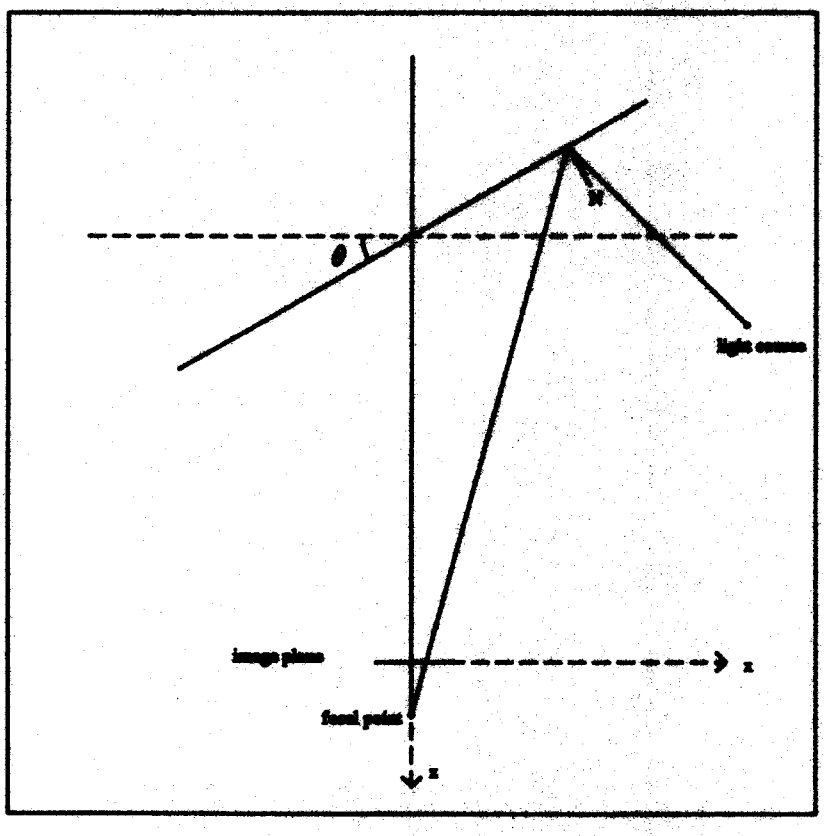

Figure 3: A point light source illuminating a tilted planar surface.

Light source at (xs, zs): 128cm, -325cm. Illuminant chromaticity: (0.3333, 0.3333). Angle : 50 degrees.

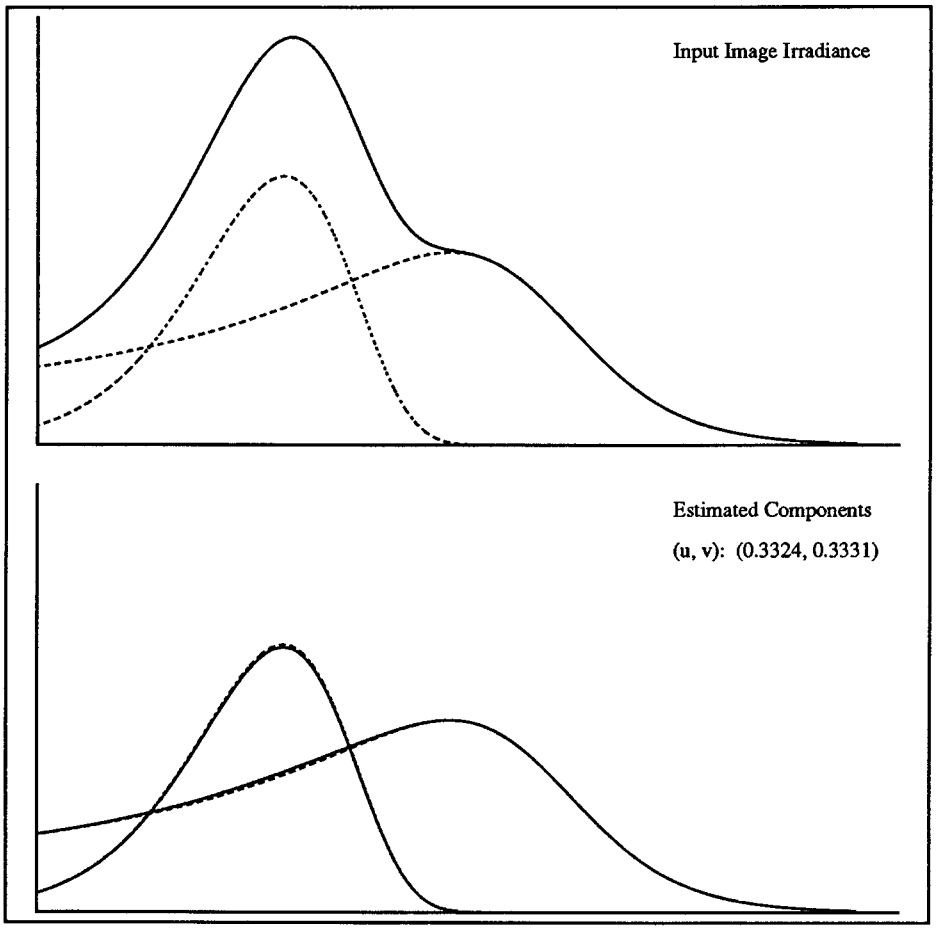

Figure 4: Separation of reflection components. Example I.

Light source at (xs, zs): 111cm, -425cm. Illuminant chromaticity: (0.3333, 0.3333). Angle : 50 degrees.

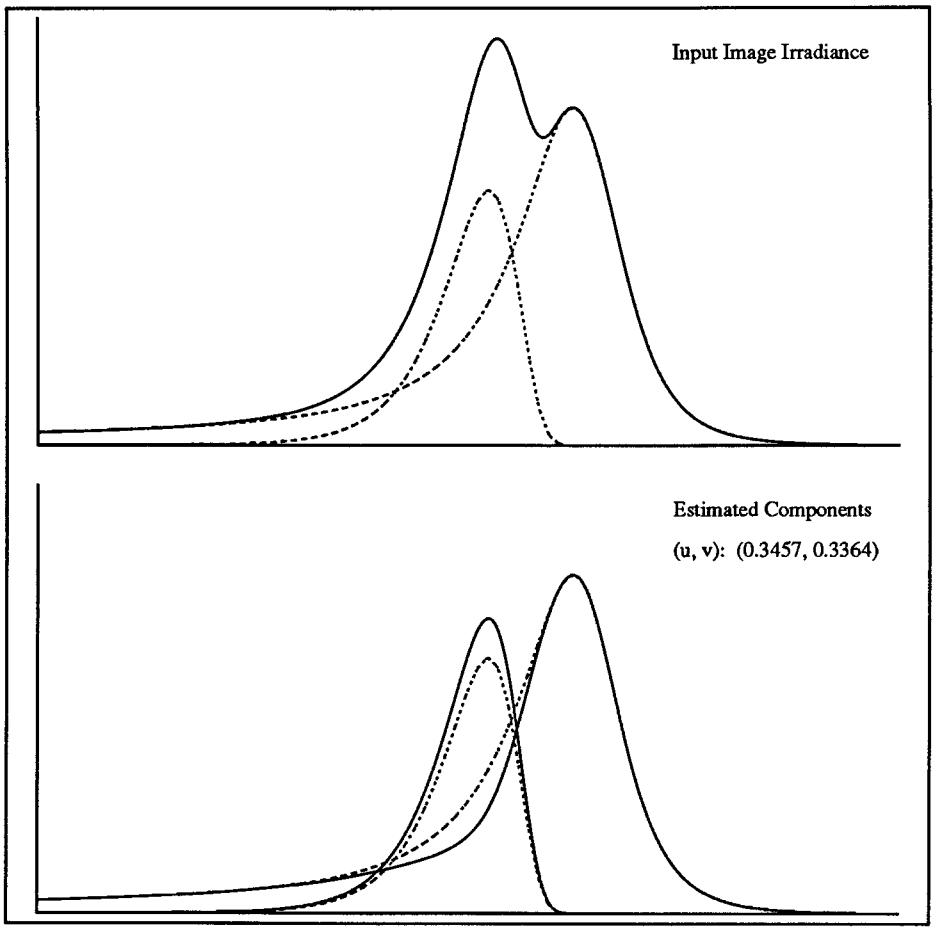

Figure 5: Separation of reflection components. Example II.

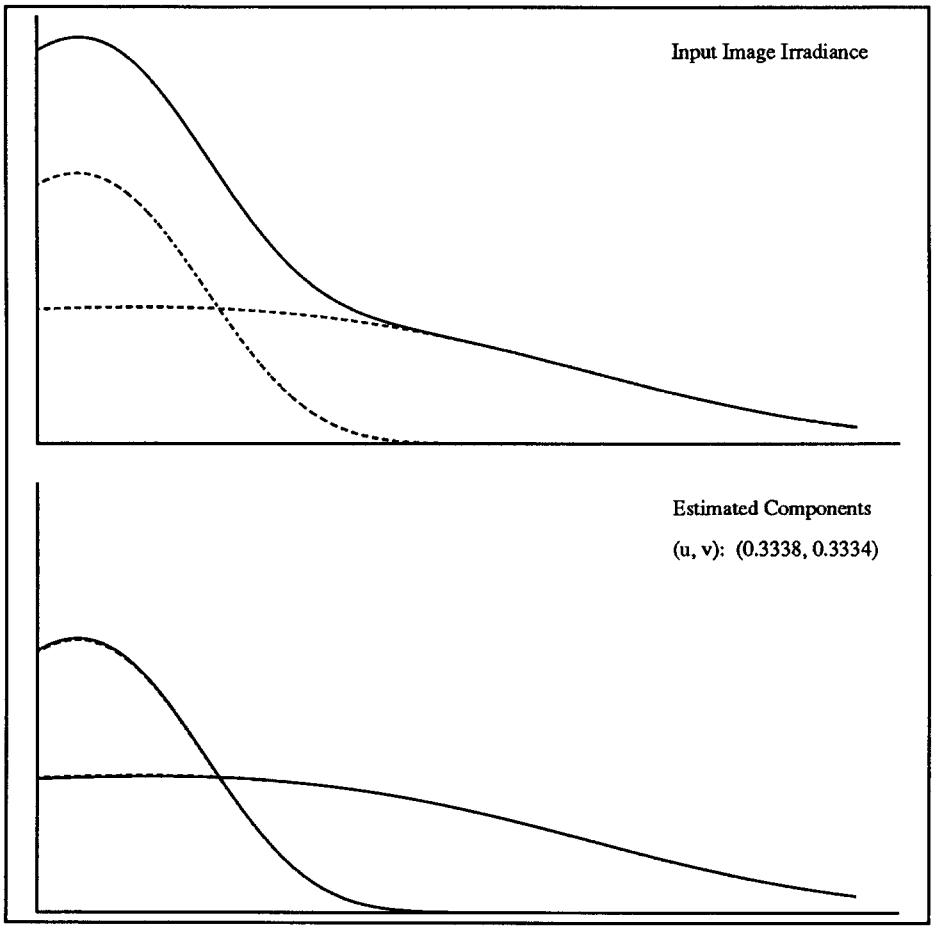

Figure 6: Separation of reflection components. Example III.

Light source at (xs, zs): -35cm, -91cm. Illuminant chromaticity: (0.3333, 0.3333). Angle : 38 degrees.

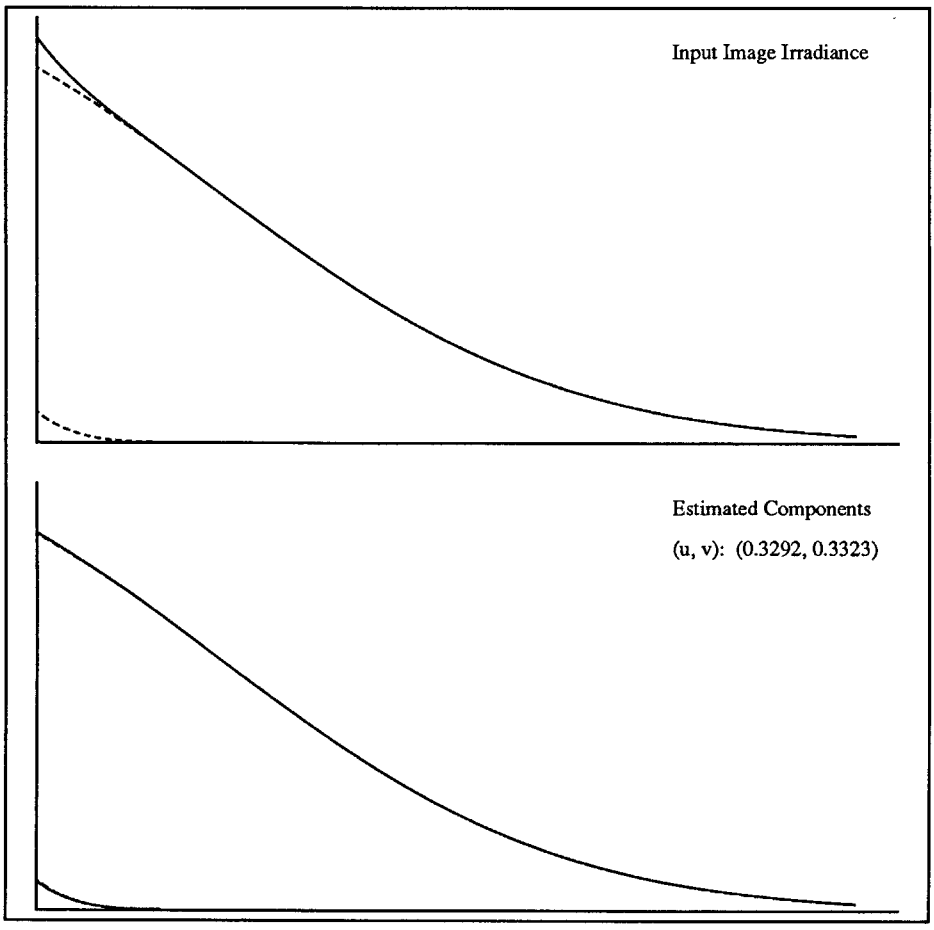

Figure 7: Separation of reflection components. Example IV.

specular component is very weak (see Fig. 7). When the light source gets very close to the camera, the light incident angle changes so fast from one point to another that the diffuse component is no longer "much smoother" than the specular component. In the case of Fig. 5, the light source is only 65.75 cm away from the object plane surface, and neither one of the two components looks much smoother than the other.

To test the estimator on two-dimensional images, we replace the plane with a sphere whose center is 395cm away from the image plane. The radius of the sphere is 120cm. The point light source is located at x = 300cm, y = 300cm, and z = 400cm. The specular reflectance factor is 2.0, and it is modeled as the cosine raised to the 40th power. In this experiment, the illuminant chromaticity is changed from image to image, and the estimator is applied to each image to estimate the illuminant color. The following table shows the results:

image	ρ_r	ρ_g	ρ_b	true illuminant (u,v)	estimated illuminant (u,v)
w000	1.0	0.4	0.6	0.5027 0.3995	0.5757 0.3326
w001	1.0	0.4	0.6	0.2631 0.4245	0.3128 0.3780
w002	0.4	0.6	1.0	0.5027 0.3995	0.4702 0.4109
w003	1.0	0.4	0.6	0.1566 0.3586	0.1878 0.3253

As we discussed in the last section that there are several ways to improve the estimation results by selectively weighting each pixel differently, i.e., to select s_r and s_g such that the following are minimized:

$$\int (W(x,y))[\nabla^2(E_r(x,y) - s_r E(x,y))]^2 dxdy,$$
$$\int (W(x,y))[\nabla^2(E_g(x,y) - s_g E(x,y))]^2 dxdy. \tag{17}$$

The solution becomes:

$$s_r = \frac{\int W(x,y)\nabla^2 E_r(x,y)\nabla^2 E(x,y)dxdy}{\int W(x,y)\nabla^2 E(x,y)\nabla^2 E(x,y)dxdy}$$
$$s_g = \frac{\int W(x,y)\nabla^2 E_g(x,y)\nabla^2 E(x,y)dxdy}{\int W(x,y)\nabla^2 E(x,y)\nabla^2 E(x,y)dxdy}. \tag{18}$$

It can easily be shown that the modified solution still satisfies the chromaticity collinearity constraint. So the improvement incurs some cost for computation, but does not sacrifice the original nice property. Since the highlight is always brighter than its surrounds, a simple weighting scheme is to make $W(x,y)$ a function of the image irradiance. The following table shows how much the estimation accuracy can be improved by a simple weighting function $W(x,y) = (E(x,y))^4$.

image	ρ_r	ρ_g	ρ_b	true illuminant (u,v)	estimated illuminant (u,v)
w000	1.0	0.4	0.6	0.5027 0.3995	0.5037 0.3985
w001	1.0	0.4	0.6	0.2631 0.4245	0.2638 0.4238
w002	0.4	0.6	1.0	0.5027 0.3995	0.5022 0.3996
w003	1.0	0.4	0.6	0.1566 0.3586	0.1570 0.3581

Fig. 8 shows the red record of one of the images (w000). Fig. 9 shows its estimated specular component, and Fig. 10 its estimated diffuse component.

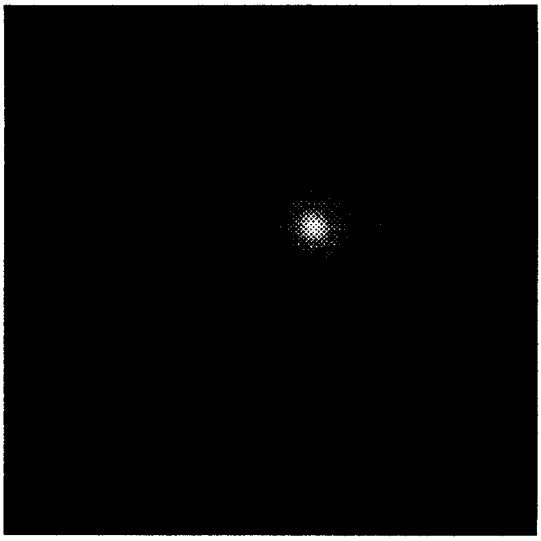

Figure 8: An image of a uniform spherical surface.

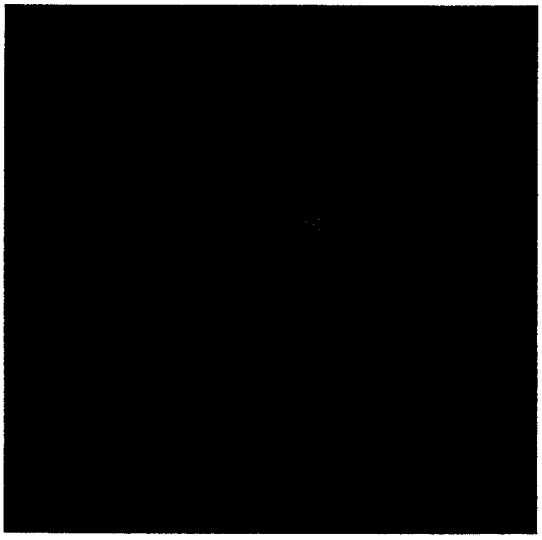

Figure 9: The estimated interface reflection component of Figure 8.

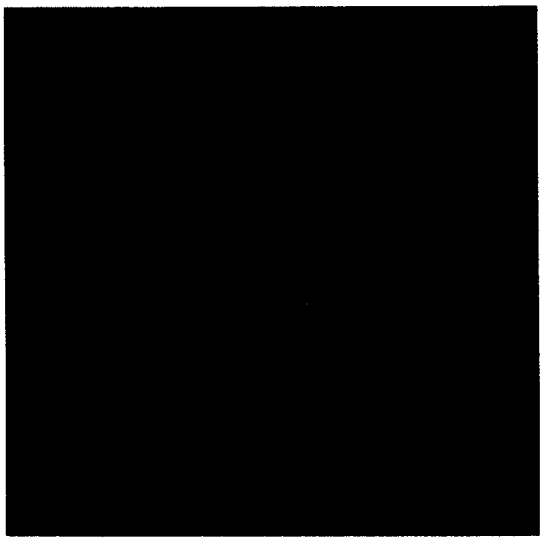

Figure 10: The estimated body reflection component of Figure 8.

5.2 From the Shading of A Non-uniform Surface

Extension of the above illuminant estimator to a non-uniform surface is conceptually straightforward. If we examine equations 6 and 7, we find that the derivation of the estimator based on the smoothness constraint still applies to the case of a non-uniform surface, provided that we can detect and exclude, from our computation of laplacians, the edge pixels between regions of different body relectance factors. Although it is well known that edge detection itself is difficult and sensitive to noise, the situation here is not as bad, because we can afford to exclude the false edges caused by noise from our computation without losing much of the essential information for illuminant color estimation.

If the non-uniform surface consists of two differently colored uniform regions i and j, then

$$\begin{aligned} s_r &= \frac{\int_{i,j} \nabla^2 E_r \nabla^2 E dx dy}{\int_{i,j} \nabla^2 E \nabla^2 E dx dy} \\ &= (\alpha_i) \frac{\int_i \nabla^2 E_r \nabla^2 E dx dy}{\int_i \nabla^2 E \nabla^2 E dx dy} + (\alpha_j) \frac{\int_j \nabla^2 E_r \nabla^2 E dx dy}{\int_j \nabla^2 E \nabla^2 E dx dy}. \end{aligned} \tag{19}$$

where,

$$\begin{aligned} \alpha_i &= \frac{\int_i \nabla^2 E \nabla^2 E dx dy}{\int_{i,j} \nabla^2 E \nabla^2 E dx dy} \\ \alpha_j &= \frac{\int_j \nabla^2 E \nabla^2 E dx dy}{\int_{i,j} \nabla^2 E \nabla^2 E dx dy}. \end{aligned}$$

The illuminant estimator extended this way therefore has a very good physical interpretation: it is the weighted average of the estimates made by each uniform region on the non-uniform surface. This is true even if the uniformly colored regions are not connected together, because the relation between f and h should hold in each region. For example, on a tiled surface, all the areas covered by the same colored tiles are combined together to make one estimate, which is in turn combined with estimates from other colors. Therefore, in principle it does not matter how many regions any color is broken into. Of course, in practice, the boundary pixels have to be excluded, and therefore there is additional loss of information for every breakup.

To generate images of non-uniform surface, random rectangles of different sizes and colors are mapped to the surface of a sphere. The center of the sphere is placed again at 395cm away from the image plane. The position of the light source, and the distribution of the colors of the random rectangles are changed to generate images with different shading and color distributions. To exclude the boundary pixels between regions of different colors, two approaches are tried: (1) exclude pixels whose laplacian magnitudes are above a predetermined threshold; (2) exclude pixels which are explicitly detected as boundary pixels. The idea behind the first approach is to treat any pixel with an unusually large laplacian magnitude as coming from physical events not related to the scale difference of interest to the estimator. It requires the determination of a threshold, and the practical question is how sensitive is the estimate to the exact value of the threshold. We pick two images, compute their histograms of laplacian values, and select the threshold as 0.75 times the average value of the non-zero laplacians. We use this threshold to process 12 images. The first four images, r000 to r003, are generated with the red, green, and blue reflectance factor distributions having equal means, and the light source located at x = 300cm, y = 300cm, and z = 400cm. We are particularly interested in comparing our estimate with the estimate given by the gray-world assumption, which takes the average image irradiances of the red, green, and blue records as the illuminant color. The results are shown as follows:

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
r000	0.5027 0.3995	0.4552 0.3991	0.4849 0.4176
r001	0.2631 0.4245	0.2080 0.4864	0.2552 0.4192
r002	0.5027 0.3995	0.5046 0.3904	0.4956 0.4009
r003	0.1566 0.3586	0.0999 0.2762	0.1495 0.3718

As can be seen in this case, the gray-world estimates are in general more accurate. We then generate two more sets of images (s000 - s003, and f000 - f003) which do not have equal means in the red, green, and blue reflectance distributions. The s000 - s003 images have the light source located at the same position as that of the r000 - r003 images, while the light source of the f000 - f003 images is located at x = 300cm, y = 300cm, and z = -100cm. The results are shown as follows:

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
s000	0.5027 0.3995	0.5196 0.3542	0.6761 0.2622
s001	0.2631 0.4245	0.2780 0.4445	0.4435 0.3141
s002	0.5027 0.3995	0.5269 0.3775	0.6918 0.2453
s003	0.1566 0.3586	0.1509 0.2705	0.2830 0.3122

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
f000	0.5027 0.3995	0.5034 0.3857	0.6700 0.2707
f001	0.2631 0.4245	0.2897 0.3668	0.4550 0.2815
f002	0.5027 0.3995	0.5322 0.3395	0.7096 0.2218
f003	0.1566 0.3586	0.1358 0.2679	0.2836 0.3250

This time the failure of the gray-world estimation is serious indeed. Comparatively, our illuminant estimator is more accurate. If we raise the threshold by 20%, the results remain about the same, the reason being that most edges are very sharp in these synthetic images. There are edges in the very dark shadings which are included in the computation of the estimate. They are not separable from the shading simply by thresholding, and are the major sources of the error.

The second approach is to use edge detection to locate the boundary pixels. A simple gradient edge detector with non-maximum suppression is applied to the individual red, green, and blue records. Any pixels detected as edges in one of the color records are declared as color edges. To ensure that all region boundaries are excluded properly, pixels which are less than two pixel away from the detected color edges are also excluded from the laplacian computation. Fig. 11 shows the red record of the image r000, and Fig. 12

Figure 11: An image of a non-uniform spherical surface.

its excluded pixels from edge detection. The following table shows the estimation results:

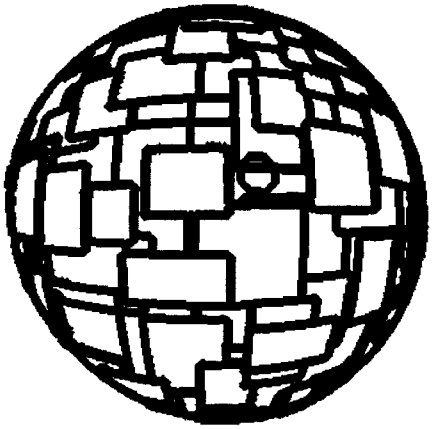

Figure 12: The excluded edge pixels of Figure 11.

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
f000	0.5027 0.3995	0.5129 0.3881	0.6700 0.2707
f001	0.2631 0.4245	0.2872 0.4239	0.4550 0.2815
f002	0.5027 0.3995	0.5118 0.3942	0.7096 0.2218
f003	0.1566 0.3586	0.1710 0.3487	0.2836 0.3250

We now apply the simple weighting function $W(x,y) = (E(x,y))^4$ as we did in the uniform surface cases to improve the estimates. The following tables show the results:

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
r000	0.5027 0.3995	0.5032 0.3992	0.4849 0.4176
r001	0.2631 0.4245	0.2634 0.4252	0.2552 0.4192
r002	0.5027 0.3995	0.5001 0.4019	0.4956 0.4009
r003	0.1566 0.3586	0.1571 0.3597	0.1495 0.3718

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
f000	0.5027 0.3995	0.5040 0.3986	0.6700 0.2707
f001	0.2631 0.4245	0.2651 0.4219	0.4550 0.2815
f002	0.5027 0.3995	0.5043 0.3978	0.7096 0.2218
f003	0.1566 0.3586	0.1589 0.3580	0.2836 0.3250

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
s000	0.5027 0.3995	0.5040 0.3986	0.6761 0.2622
s001	0.2631 0.4245	0.2658 0.4231	0.4435 0.3141
s002	0.5027 0.3995	0.5045 0.3981	0.6918 0.2453
s003	0.1566 0.3586	0.1584 0.3589	0.2830 0.3122

If we move the light source to $x = 300$cm, $y = 300$cm, and $z = 40000$cm, we have the following results:

image	true illuminant (u,v)	estimated illuminant (u,v)	gray-world estimate (u,v)
g000	0.5027 0.3995	0.5039 0.3987	0.6799 0.2576
g001	0.2631 0.4245	0.2659 0.4219	0.4380 0.3294
g002	0.5027 0.3995	0.5046 0.3977	0.6830 0.2566
g003	0.1566 0.3586	0.1581 0.3591	0.2840 0.3059

As can be seen, all the estimates are consistently very accurate. They are now better than the gray-world estimates even when the images are generated by gray-world statistics (images r000 to r003).

It should be pointed out that the laplacian signal of the interface reflection component $h(x,y)$ is relatively small compared with that of a reflectance edge. If any reflectance edge pixel is not excluded from the estimation, the result can be very wrong. For example, the gradient edge detector sometimes computes the location of the true edge incorrectly. If we have not had excluded pixels which are neighbors of the detected edges, the estimation results will be useless for some images, as can be seen in the following tables. The columns marked as 1-pixel and 2-pixel are the estimation results from excluding pixels which are less than or equal to one and two pixel distance away from the detected edges.

image	true (u,v)	edge alone (u,v)	1-pixel (u,v)	2-pixel (u,v)
f000	0.5027 0.3995	0.5044 0.3982	0.5041 0.3985	0.5040 0.3986
f001	0.2631 0.4245	0.4059 0.3412	0.2653 0.4216	0.2651 0.4219
f002	0.5027 0.3995	0.5487 0.3622	0.5044 0.3977	0.5043 0.3978
f003	0.1566 0.3586	0.1621 0.3537	0.1598 0.3569	0.1589 0.3580

6 Concluding Remarks

Light reflection components usually have different spatial scales. The interface (specular) reflection component typically has a much finer scale than the body (diffuse) component. Because of this scale difference, one can effectively separate them by imposing a smoothness constraint to extract the shape of the diffuse component. In doing so, an estimate of the illuminant color can be obtained. This illuminant estimator happens to satisfy the chromaticity collinearity constraint automatically, which makes it particularly attractive for applications. When it is generalized to a non-uniform surface, it essentially computes the weighted average of the individual estimates made by each of the uniformly colored regions, provided that one can eliminate the edge pixels.

The estimator has been tested on computer generated images with reasonably good performance. Unlike the gray-world estimator, it is quite immune to heavy bias in the reflectance statistics of the scene. However, in order to use it for a non-uniform surface, explicit exclusion of edges is necessary. Fortunately, one need not worry about false edges because they can be excluded from the estimator without much loss of information. On the other hand, one has to worry about undetected edges because their inclusion in computing the estimator can introduce a large error in the estimation.

The estimation accuracy can be greatly improved by weighting each pixel with a function of its irradiance. Other types of weighting function and definition of smoothness can be used, but are not pursued in this work. One final question one might ask is: Can this estimator be applied to images with many surfaces ? The answer is yes if all the surfaces are spatially near each other in depth. For surfaces far away from each other, one has to distinguish the scale difference due to depth from the scale difference between the reflection components. Compared with the chromaticity convergence method, the illuminant-color estimator developed here is simpler in implementation. It works well for a single surface, and does not explicitly search for straightline signatures in the chromaticity space. Furthermore, the computation is more local and therefore potentially can take better care of images with multiple light sources. However, the chromaticity convergence method does not use the difference-in-spatial-scale assumption, and would not have any problem when several surfaces are at different depths. Combination of the two methods to compute a robust estimation of the scene-illuminant chromaticity is a subject of future research.

7 Acknowledgement

This work was done while the author was on leave from Kodak Research Laboratories. He would like to thank Professor Tomaso Poggio of MIT, Mr. Terence Lund, and Mr. William Hutchinson of Kodak for making this possible, people in the Artifical Intelligence Laboratory and the Center for Biological Information Processing for making the visit so enjoyable, and Eastman Kodak Company for the financial support. Anya Hurlbert, Tomaso Poggio, and James Little provided many good suggestions and comments, making the final form more readable.

References

[1] R.L. Cook and K.E. Torrance, "A reflectance model for computer graphics," *ACM Transactions on Graphics*, 1, 1, January 1982, pp. 7-24.

[2] R. Gershon, A.D. Jepson, and J.K. Tsotsos, "The use of color in highlight identification," *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, Milan, Italy, August 1987, pp. 752-754.

[3] D.B. Judd, D.L. MacAdam, and G.W. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," *Journal of the Optical Society of America*, 54, 1031, 1964,

[4] G.J. Klinker, S.A. Shafer, and T. Kanade, "Image segmentation and reflection analysis through color," *Proceedings of Image Understanding Workshop*, Cambridge, Massachusetts, April 1988, pp. 838-853.

[5] H.-C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *Journal of the Optical Society of America*, A, 3, 10, October 1986, pp.1694-1699.

[6] H.-C. Lee, E.J. Breneman, and C.P. Schulte, "Modeling light reflection for computer color vision," Submitted to IEEE PAMI.

[7] *The Science of Color*, Committee on Colorimetry, Optical Society of America, published by Optical Society of America, Washington, D.C., 1963, 145-152.

[8] B.T. Phong, "Illumination for computer generated pictures," *Communications of the ACM*, 18, 1975, 311-317.

[9] T. Poggio and C. Koch, "Ill-posed problems in early vision: from computational theory to analog networks, *Proceedings of Royal Society*, London, B, 226, 1985, pp. 303-323.

[10] S.A. Shafer, "Using color to separate reflection components", *COLOR research and application*, 10, 4, Winter 1985, pp. 210-218.

[11] G. Wyszecki and W.S. Stiles, *Color Science*, 2nd edition, John Wiley and Sons, Inc., 1982.

[12] M. D'Zmura and P. Lennie, "Mechanisms of color constancy", *Journal of the Optical Society of America*, A, 3, 10, October 1986, pp.1662-1672.

CS-TR Scanning Project
Document Control Form

Date : 4/27/95

Report # AIM-1068

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☐ Technical Report (TR) ☒ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 24 (30-IMAGES)
Not to include DOD forms, printer instructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other:_____

Check each if included with document:

☒ DOD Form (2) ☐ Funding Agent Form ☐ Cover Page
☐ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages (by page number):_____

Photographs/Tonal Material (by page number): 16, 17, 19

Other (note description/page number):

Description :	Page Number:
IMAGE MAP (1) UNNUMBERED TITLE PAGE	
(2-24) PAGES #'ED 1-23	
(25-27) SCANCONTROL, DOD (2)	
(28-30) TRGTS	

Scanning Agent Signoff:
Date Received: 4/27/95 Date Scanned: 5/1/95 Date Returned: 5/1/95

Scanning Agent Signature: Michael W. Cook

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Completed 1/13/89

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AIM 1068	2. GOVT ACCESSION NO.: A 201 333
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): Estimating the Illuminant Color from the Shading of a Smooth Surface	5. TYPE OF REPORT & PERIOD COVERED: memorandum
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Hsien-Che Lee	8. CONTRACT OR GRANT NUMBER(s): N00014-85-K-0124
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209	12. REPORT DATE: August 1988
	13. NUMBER OF PAGES: 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, VA 22217	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Unlimited

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

color vision
specular reflection
shading
illuminant chromaticity

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The image of a uniform wall illuminated by a spot light often gives a strong impression of the illuminant color. How can it be possible to know if it is a white wall illuminated by yellow light or a yellow wall illuminated by white light? If the wall is a Lambertian reflector, it would not be possible to tell the difference. However, in the real world, some amount of specular reflection is almost always present. In this memo, it is shown that the computation is possible in most practical cases.

See back

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 continued

Light reflection from a surface is usually modeled as having two components: the interface (specular) reflection and the body (diffuse) reflection. For a surface of inhomogeneous material, the spectral composition of the interface reflection component is often similar to that of the illuminant. The problem of computing the illuminant chromaticity from the shading of a single smooth surface is to separate these two components. An image of an illuminated uniform wall, according to the above model, gives only one physical constraint about the illuminant chromaticity, not enough to determine a unique solution. However, since the spatial scale over which the interface reflection changes significantly is much smaller than that of the body reflection, it can be shown that one can effectively exploit the scale difference to find a unique solution, which is often very accurate. The method can also be generalized to compute the illuminant chromaticity for a nonuniform smooth surface.

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**