Component estimation of surface spectral reflectance

Shoji Tominaga

Faculty of Engineering, Osaka Electro-Communication University, Neyagawa, Osaka, 572 Japan

Brian A. Wandell

Department of Psychology, Stanford University, Stanford, California, 94305

Received February 6, 1989; accepted September 27, 1989

For inhomogeneous materials, the standard reflectance model suggests that under all viewing geometries surface reflectance functions can be described as the sum of a constant function of wavelength (specular) and a diffuse function that is characteristic of the material. As the viewing geometry varies, the relative contribution of these two terms varies. In a previous study [J. Opt. Soc. Am. A **6**, 576 (1989)] we described how to use light reflected from inhomogeneous materials, measured in different viewing geometries, to estimate the relative spectral power distribution of the ambient light. Here we show that two restrictions, that (a) surface reflectance functions are all nonnegative and (b) surface reflectance functions are the positive weighted sum of subsurface (diffuse) and interface (specular) components, may be used to estimate the subsurface component of the surface reflectance function. A band of surface spectral reflectances is recovered, as possible solutions for the subsurface estimates.

1. INTRODUCTION

The surface reflectance function of an object varies with the viewing geometry. A simple model of surface reflectance is proposed for inhomogeneous materials (see, e.g., Refs. 1 and 2). Inhomogeneous materials are composed of different component materials, such as a vehicle at the surface layer and embedded pigments at the colorant layer. Plastics and paints are inhomogeneous. The model suggests that to a good approximation the reflectance function can be described as the weighted sum of two functions. One function represents the interface (specular) reflection, and the second function represents the subsurface (diffuse) reflection.

Let $S_I(\lambda)$ and $S_S(\lambda)$ be the surface spectral reflectances for the two components of interface and subsurface reflections, and let $E(\lambda)$ be the spectral power distribution of the incident light. Then the reflected light is

$$Y(\theta, \lambda) = c_I(\theta)S_I(\lambda)E(\lambda) + c_S(\theta)S_S(\lambda)E(\lambda), \quad (1)$$

where $c_I(\theta)$ and $c_S(\theta)$ are the geometric scale factors. The parameters θ vary with the viewing geometry such as the incident angle of light on the surface, the viewing angle, and the phase angle. In a previous paper[2] we evaluated this model and showed how to use light reflected from pairs of surfaces to estimate the spectral power distribution of the illuminant (see also Refs. 3 and 4). From a component analysis of the reflected light, we concluded that the model is correct and that the spectral composition of the specular component of the reflected light is the same as the spectral composition of the incident light. The estimation problem of the illuminant spectrum can then be reduced to finding the common spectral information from the measurements from two or more surfaces. We presented an algorithm to obtain an illuminant estimate without using a reference white standard.

The subsurface reflectance function $S_S(\lambda)$ is characteristic of the surface and therefore is an important clue to the surface's identity. The interface term $S_I(\lambda)$, which is present in all homogeneous materials, does not contain much information about the surface's identity, although it may be useful in determining the geometry of the objects in the image.[5]

The methods in our previous paper provided a unique estimate of the illuminant, but the method did not provide an estimate of the subsurface spectral reflectance term, $S_S(\lambda)$. The estimate of the subsurface reflectance function was restricted to be a function of the form $\alpha + \beta S(\lambda)$, where $S(\lambda)$ is derived from the measurements and α and β are free parameters.

In our previous study we did not take advantage of two important constraints on the surface reflectance functions. First, the subsurface reflectance function $S_S(\lambda)$ must be nonnegative. Second, for all viewing geometries the parameters c_I and c_S must both be nonnegative. These constraints were identified by Lawton and Sylvestre,[6] who also offered an elegant mathematical analysis of how the constraints might be applied. Our methods are also related to recent studies in multispectral image analysis by Kawata *et al.*[7] In Section 2 we state these constraints more precisely.

2. PHYSICAL CONSTRAINTS

The standard reflectance model assumes that the color signals $Y(\theta, \lambda)$ can be expressed as a linear combination of the two component vectors of light reflection,

$$Y(\theta, \lambda) = c_I(\theta)L_I(\lambda) + c_S(\theta)L_S(\lambda), \quad (2)$$

where $L_I(\lambda) = S_I(\lambda)E(\lambda)$ and $L_S(\lambda) = S_S(\lambda)E(\lambda)$. The two vectors $L_I(\lambda)$ and $L_S(\lambda)$ span a two-dimensional space of the possible observations. All the color signals observed from an inhomogeneous surface fall within this two-dimensional subspace.

Using our estimate of the ambient light, we can define the

0740-3232/90/020312-06$02.00

two-dimensional space of surface reflectances by dividing the two-dimensional space of color signals by $E(\lambda)$:

$$\frac{Y(\theta, \lambda)}{E(\lambda)} = S(\theta, \lambda) = c_I(\theta)S_I(\lambda) + c_S(\theta)S_S(\lambda). \quad (3)$$

Note that this surface reflectance plane defined by $S_I(\lambda)$ and $S_S(\lambda)$ does not change with different illuminants. The surface reflectance plane is inherent in the object surface, unlike the color signal plane defined by $L_I(\lambda)$ and $L_S(\lambda)$. By hypothesis, the interface component $S_I(\lambda)$ of the two dimensions is a constant specular function. The subsurface function $S_S(\lambda)$ must (a) be all nonnegative over the visible wavelength as $S_S(\lambda) \geq 0$ and (b) contribute only nonnegative weights for all of the measured color signals as $c_S(\theta) \geq 0$.

Lawton and Sylvestre treated the general problem of determining the shapes of two functions with such physical constraints from an observed set of additive mixtures of the two functions. Their solution method was based on principal component analysis. Their method produced two bands of estimated functions. In our application, one function, the interface function $S_I(\lambda)$, is fixed. Thus our analysis is a special case of their general two-dimensional analysis. It should be noted that our problem is the estimation of a band of the subsurface reflectance functions rather than the estimation of a unique $S_S(\lambda)$; that is, we can recover only a range of subsurface estimates, not a unique estimate. In Section 3 we introduce the special case of the Lawton–Sylvestre analysis that we call the quarter-circle analysis.

3. QUARTER-CIRCLE ANALYSIS

A. Observation of Surface Reflectances

First, assume that we measure m color signals from an object surface under various geometric conditions $\theta_1, \theta_2, \ldots, \theta_m$. All the measurements are made with a single fixed illuminant. Each color signal that we measure, $Y(\theta_i, \lambda)(i = 1, 2, \ldots, m)$, is sampled at n points $\lambda_1, \lambda_2, \ldots, \lambda_n$ in the visible wavelength region. Next let $\hat{E}(\lambda_j)(j = 1, 2, \ldots, n)$ be an estimate of the spectral power distribution of the illuminant. Then the observations of m surface spectral reflectances are obtained by $S(\theta_i, \lambda_j) = Y(\theta_i, \lambda_j)/\hat{E}(\lambda_j)$. We represent these reflectances by n-dimensional column vectors $\mathbf{s}_i$ $(i = 1, 2, \ldots, m)$. These vectors are normalized so that $\|\mathbf{s}_i\|^2 = 1$, where the notation $\|\mathbf{s}\|^2$ is defined as $\mathbf{s}^T\mathbf{s}$.

Since the surface spectral reflectances are two dimensional, the normalized reflectances $\mathbf{s}_i$ can be represented in terms of two orthonormal basis vectors as

$$\mathbf{s}_i = p_{i1}\mathbf{u}_1 + p_{i2}\mathbf{u}_2 \qquad (i = 1, 2, \ldots, m). \quad (4)$$

This orthogonal expansion can be determined by using the singular value decomposition of an $n \times m$ matrix of the observed reflectances (see, e.g., Ref. 2). The vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are the first two of the n-dimensional left singular vectors. The coefficients p_{i1} and p_{i2} are specified from the singular values σ_j and the elements of the m-dimensional right singular vectors $\mathbf{v}_j$ as $p_{ij} = \sigma_j v_{ij}$ $(j = 1, 2)$. The first basis vector $\mathbf{u}_1$ has all nonnegative elements, so that $\mathbf{u}_1 \geq 0$ $(u_{j1} \geq 0, j = 1, 2, \ldots, n)$ and $p_{i1} \geq 0$. However, because of the orthogonality condition imposed on eigenvectors, the second $\mathbf{u}_2$ has some negative elements. Because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2$, it follows that

$$\begin{aligned} 1 = \|s_i\|^2 &= \|p_{i1}\mathbf{u}_1 + p_{i2}\mathbf{u}_2\|^2 \\ &= \|p_{i1}\mathbf{u}_1\|^2 + \|p_{i2}\mathbf{u}_2\|^2 \\ &= {p_{i1}}^2 + {p_{i2}}^2. \end{aligned} \quad (5)$$

Thus the representation of the normalized surface reflectance measurements will have unit length in the principal component coordinate frame $(\mathbf{u}_1, \mathbf{u}_2)$.

B. Basis Transformations

The unit-length basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are computed from the normalized observations of surface reflectance functions. These vectors depend on experimental factors such as the choice of measurement angles θ and the number of measurements at each of the measurement angles. These vectors do not characterize the surface properties. Our goal is to find a method of transforming the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ to a new representation that estimates a range of the subsurface reflectance vector $\mathbf{S}_S$ for $S_S(\lambda)$. This vector $\mathbf{S}_S$ characterizes the object surface and is independent of the experimental parameters. The transformation that we apply to the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ to estimate $\mathbf{S}_S$ is based on three premises. First, according to theory (see, e.g., Ref. 1) and confirmation by our experimental measurements,[2] the unit vector $\mathbf{S}_I$ in the direction of interface reflectance, $[S_I(\lambda) = 1/\sqrt{n}$ $(\lambda = \lambda_1, \lambda_2, \ldots, \lambda_n)$, $\sum_{i=1}^{n} {S_I}^2(\lambda_i) = 1]$, is within a linear transformation of the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Second, the subsurface reflectance $S_S(\lambda)$ is nonnegative. Third, the surface reflectances measured under different angles θ are positive mixtures of the subsurface and the interface reflectances; that is, all measurements are of the form $c_I(\theta)S_I(\lambda) + c_S(\theta)S_S(\lambda)$, where $c_I(\theta), c_S(\theta) \geq 0$. On the basis of these three premises, we can estimate a band functions that must include the subsurface function $S_S(\lambda)$.

For conceptual simplicity, we begin by transforming the data representation from the principal component coordinates (p_1, p_2) with respect to the basis functions $\mathbf{u}_1$ and $\mathbf{u}_2$ into a new coordinate frame with respect to the coordinate system defined by the vector S_I and a unit-length vector perpendicular to $\mathbf{S}_I$ that we call $\mathbf{S}_I{}^{\perp}$. From our experimental measurements we know that $\mathbf{S}_I$ can be obtained by a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$. In other words, $\mathbf{S}_I$ is within the surface reflectance plane spanned by $(\mathbf{u}_1, \mathbf{u}_2)$. Then we can find the unique unitary transformation $\mathbf{T}$ such that

$$[\mathbf{S}_I, \mathbf{S}_I{}^{\perp}] = [\mathbf{u}_1, \mathbf{u}_2]\mathbf{T}. \quad (6)$$

The unitary transformations $\mathbf{T}$ rotates $\mathbf{u}_1$ into $\mathbf{S}_I$. The second vector $\mathbf{S}_I{}^{\perp}$ is the unique unit-length vector that is orthogonal to $\mathbf{S}_I$ within the surface reflectance plane.

In practice, because of measurement error and small failures of the model, the interface reflectance component $\mathbf{S}_I$ cannot be assumed to be precisely in the plane of the first two principal components $(\mathbf{u}_1, \mathbf{u}_2)$ of pixel data. Moreover the same situation may occur when more than three objects are used to infer the illuminant spectrum and the interface reflectance $\mathbf{S}_I$. The interface component is then obtained as the projection of a unit vector onto the $(\mathbf{u}_1, \mathbf{u}_2)$. The unitary matrix $\mathbf{T}$ can be found by solving the least-squares problem that minimized $\|\mathbf{i} - t_1\mathbf{u}_1 - t_2\mathbf{u}_2\|^2$, where $\mathbf{i}$ is a unit vector with entries of $1/\sqrt{n}$. By convention we choose $\mathbf{u}_2$ so that t_2 is positive. (If the original choice of $\mathbf{u}_2$ leads to a negative solution, then simply replace $\mathbf{u}_2$ with $-\mathbf{u}_2$, since this sign is

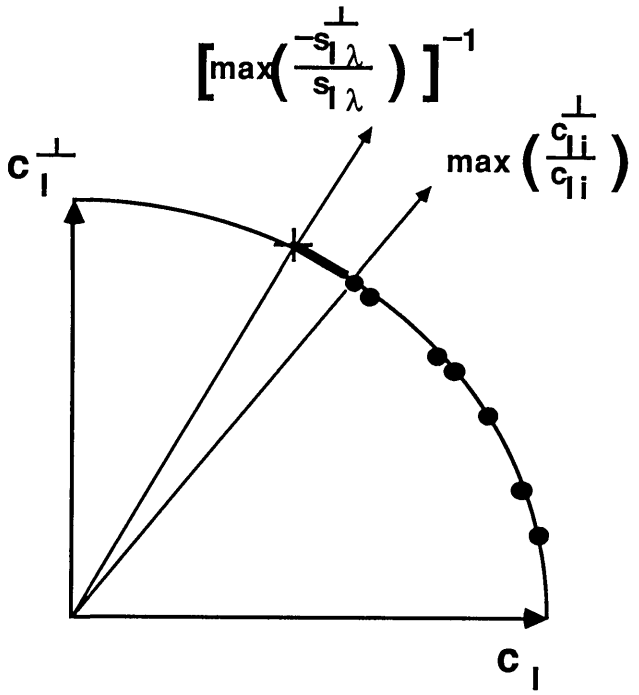

Fig. 1. Pictorial diagram of the $(c_I, c_I^{\perp})$ coordinate system in which the observed surface spectral reflectances are projected onto the quarter of the unit circle within the positive quadrant. Each dot indicates the observation point. A + on the quarter circle indicates the upper limit of the permissible directions for describing the surface reflectance vectors. The possible solution band is shown by the bold arc on the circle.

arbitrary in the construction of the principal components.) The matrix $\mathbf{T}$ is

$$\mathbf{T} = \frac{1}{(t_1^2 + t_2^2)^{1/2}} \begin{bmatrix} t_1 & t_2 \\ t_2 & -t_1 \end{bmatrix}. \tag{7}$$

An observation that is represented by (p_{i1}, p_{i2}) in the principal component coordinate frame will be represented by $[c_{Ii}, c_{Ii}^{\perp}] = [p_{i1}, p_{i2}]\mathbf{T}$ in the transformed coordinate frame.

There are several simplifying features of this coordinate frame that are illustrated in Fig. 1. First, because the surface reflectances observed have been normalized to unit length, and $\mathbf{T}$ is a unitary transformation, the coordinates $(c_I, c_I^{\perp})$ satisfy $c_{Ii}^2 + c_{Ii}^{\perp 2} = 1$. Thus the representation of the measurements falls on the unit circle. Next, the value of c_I must exceed 0 because this component describes the contribution from the interface reflectance, which is always nonnegative. Furthermore, the value of $c_I^{\perp}$ is proportional (but not equal) to the subsurface contribution, c_S, which is always positive as well. By our choice of transformation $\mathbf{T}$, we have assured ourselves that the constant of proportionality is positive. Thus the observations $(c_{Ii}, c_{Ii}^{\perp})$ must lie on the quarter of the unit circle within the positive quadrant.

The point (1, 0) is consistent with a perfectly specular interface reflection. The point (0, 1) cannot be observed because $\mathbf{S}_I^{\perp}$ has negative elements.

C. Solution Band

The observation of the surface reflectances is totally ordered along the quarter circle. As the contribution of the subsurface component increases, the observation point moves counterclockwise on the circle. The point for the pure subsurface reflectance function lies somewhere on the quarter circle. The only issue that remains, then, is to determine the direction of the vector $\mathbf{S}_S$ representing the subsurface reflectance $S_S(\lambda)$.

Lawton and Sylvestre[6] made the following observations. First, the requirement that both c_S and c_I be positive implies that the observations must fall within a cone defined by the direction of $\mathbf{S}_S$ and the horizontal axis c_I. To understand this point, notice that, if a data point falls counterclockwise with respect to the vector $\mathbf{S}_S$, then the data point would be represented with a negative value of c_I; this is impermissible. It follows that directions that qualify as candidates for $\mathbf{S}_S$ must fall counterclockwise with respect to all the data points. The data point closest to the $c_I^{\perp}$ axis (y axis) defines this limit.

Second, the vector $\mathbf{S}_S$ representing the subsurface reflectance $S_S(\lambda)$ must have all nonnegative elements. Not all vector directions satisfy this requirement. For example, the vector defined by the $c_I^{\perp}$ axis contains negative elements. As we consider unit vectors in the clockwise direction from the $c_I^{\perp}$ axis, there will be a first direction in which all the entries of the reflectance vectors are nonnegative. This direction is a second bound on the region of candidate directions for $\mathbf{S}_S$. The true vector direction representing $\mathbf{S}_S$ must fall between this bound (marked by a + on the quarter circle in Fig. 1) and the data point at the farthest counterclockwise direction. The set of permissible directions for describing the subsurface reflectance vector $\mathbf{S}_S$ is given by this band.

Both of these constraints may be stated by using our notation. Suppose that we represent the vector direction of the subsurface reflectance as (r_1, r_2); that is, the shape of the subsurface reflectance is represented in the form $S_S(\lambda) = r_1 S_I(\lambda) + r_2 S_I^{\perp}(\lambda)$. The requirement that the direction of $\mathbf{S}_S$ be counterclockwise to the last data point can be written as

$$\frac{r_2}{r_1} \geq \max_{1 \leq i \leq m} \left(\frac{c_{Ii}^{\perp}}{c_{Ii}} \right), \tag{8}$$

where the subscript i ranges over the m observations. The second requirement, that all the entries of $\mathbf{S}_S$ be nonnegative, can be expressed as $r_1 S_I(\lambda) + r_2 S_I^{\perp}(\lambda) \geq 0$ for all visible wavelengths $\lambda_1, \lambda_2, \ldots, \lambda_n$. This can be written as

$$\frac{r_1}{r_2} \geq \max_{1 \leq \lambda \leq n} \left(\frac{-s_{I\lambda}^{\perp}}{s_{I\lambda}} \right), \tag{9}$$

where $s_{I\lambda}$ and $s_{I\lambda}^{\perp}$ are the elements of the vectors $\mathbf{S}_I$ and $\mathbf{S}_I^{\perp}$, respectively. Both r_1/r_2 and the maximum value are always positive, so we can rewrite this relation as

$$\frac{r_2}{r_1} \leq \left[\max_{1 \leq \lambda \leq n} \left(\frac{-s_{I\lambda}^{\perp}}{s_{I\lambda}} \right) \right]^{-1}. \tag{10}$$

We can summarize the bounds on the direction of the vector $\mathbf{S}_S$ in a single equation as

$$\max_{1 \leq i \leq m} \left(\frac{c_{Ii}^{\perp}}{c_{Ii}} \right) \leq \frac{r_2}{r_1} \leq \left[\max_{1 \leq \lambda \leq n} \left(\frac{-s_{I\lambda}^{\perp}}{s_{I\lambda}} \right) \right]^{-1}. \tag{11}$$

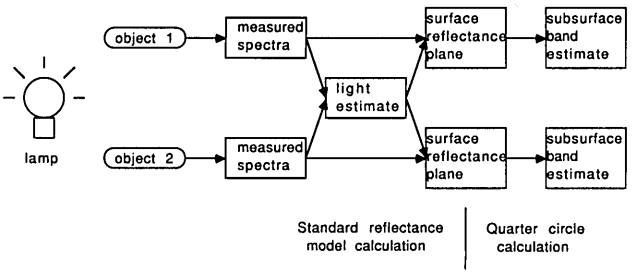

Fig. 2. Schematic representation of the experimental procedure.

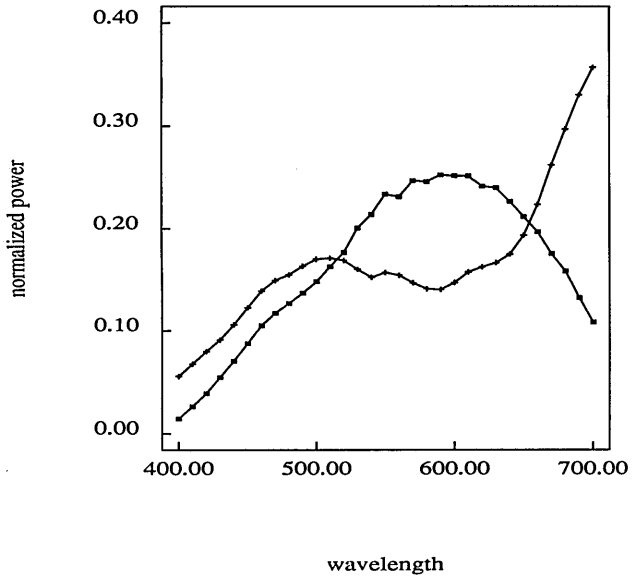

Fig. 3. Estimation results of the illuminant spectral power distributions of two light sources.[2] Crosses represent the estimate of the flood lamp, and filled squares represent that of the tungsten halogen lamp of a slide projector.

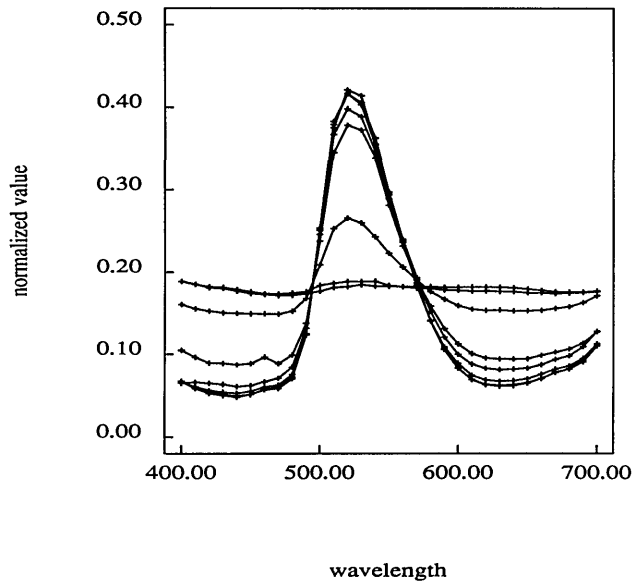

Fig. 4. Normalized curves of the observed surface spectral reflectances of the green plastic ashtray.

Any estimate of $S_S(\lambda)$ with this bounded region satisfies all the nonnegativity constraints and provides an equally satisfactory explanation of the data.

4. EXPERIMENTS

We performed several experiments to evaluate the quarter-circle analysis method of estimating the subsurface reflectance function. The flow chart in Fig. 2 describes the procedure. Two object surfaces are measured with different geometric conditions under a light source. The illuminant spectrum is estimated from the measured color signals by using the method described by Tominaga and Wandell.[2] The observations of the surface spectral reflectances are then obtained by dividing the measured color signals by the estimated ambient light. The subsurface reflectance functions are estimated by the quarter-circle analysis. Some typical results are shown in the following subsections.

A. Plastics

A red cup and a green ashtray were measured under a flood lamp for daylight photograph. The measured spectral data were shown in the previous paper.[2] The estimate of the illuminant spectral power distribution is shown as the curve with crosses in Fig. 3. Figure 4 shows a set of the normalized curves of the surface spectral reflectances of the green ashtray $S(\theta_i, \lambda)$ $(i = 1, 2, \ldots, 8)$. Two principal components are extracted from the singular value decomposition, and the 2×2 transformation matrix is determined on the least-squares fitting to the unit vector with entries of $1/\sqrt{n}$. Figure 5 shows the estimated basis curves of $S_I(\lambda)$ and $S_I^{\perp}(\lambda)$. The nearly straight line indicates a constant spectrum of the interface reflectance function $S_I(\lambda)$. Figure 6 shows the projection of the observed reflectances into the weighting coefficient coordinates $(c_I, c_I^{\perp})$. Note that the observed points (filled squares) lie on the unit quarter circle. The bounded region for $S_S(\lambda)$ is given as the interval $0.842 \leq r_2/r_1$

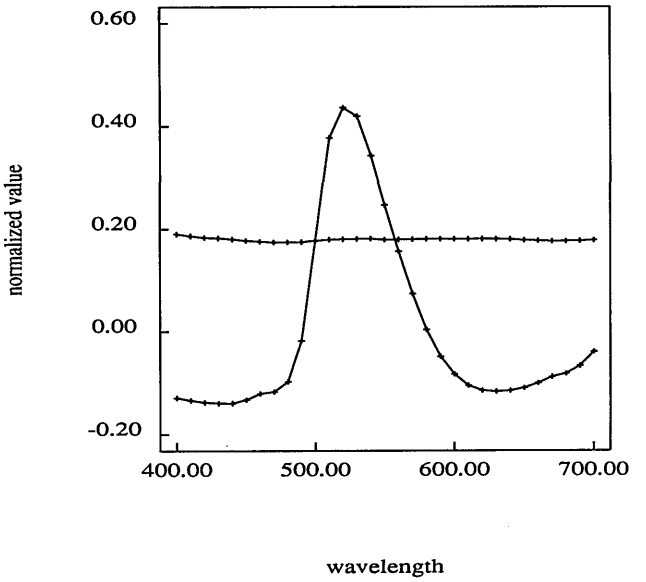

Fig. 5. Basis curves of the observed surface reflectances. The nearly straight line indicates a constant spectrum of the interface reflectance function.

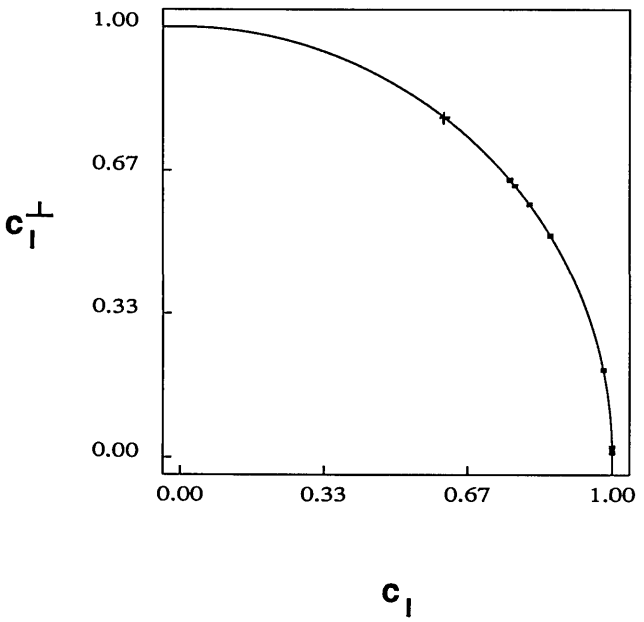

Fig. 6. Projection points of the surface reflectances. The eight points are indicated with filled squares. The upper limit of the solution band is marked with a +.

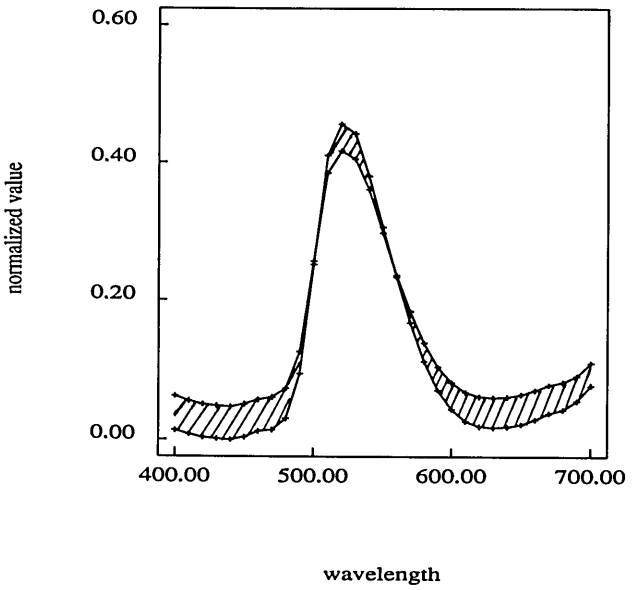

Fig. 7. An estimate of the band of the subsurface reflectance functions of the green plastic ashtray.

≤ 1.293 on the circle. A + indicates the upper limit and corresponds to the limit of physically realizable reflectance curves. An estimate of the band of the subsurface reflectance function is shown in Fig. 7, in which a range of the subsurface estimates is given as the hatched band. The curve that has a point of tangency to the wavelength axis is the theoretically limiting reflectance from nonnegativity. The other curve is the most extreme reflectance of real measurements. The width of the band represents the uncertainty of the estimates.

B. Fruits

An apple and a lemon were measured under the tungsten halogen lamp of a slide projector. The illuminant estimate is shown in a curve with filled squares in Fig. 3 (also refer to the previous paper[2]). Figure 8 shows the normalized curves of the surface spectral reflectances of the lemon. The same computational procedure has been applied. Figure 9 shows an estimated band of the subsurface reflectance function. We have the interval $0.841 \leq r_2/r_1 \leq 0.851$. In this case, the uncertainty of the estimates is small. There is little discrepancy between the theoretically limiting reflectance and the most extreme real reflectance.

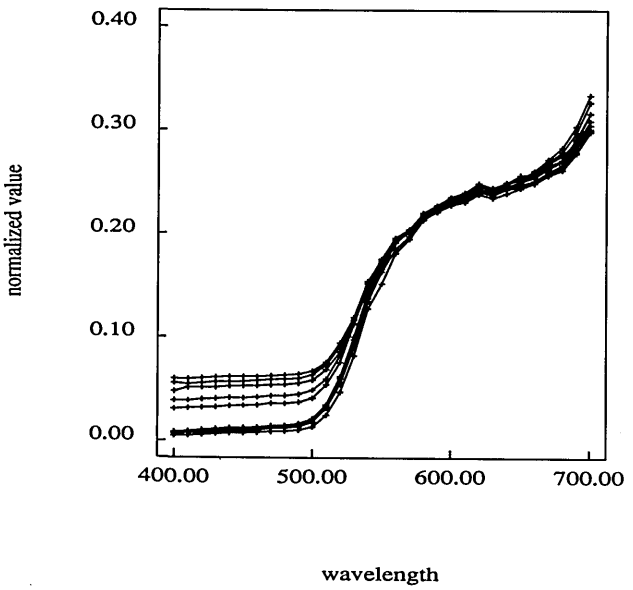

Fig. 8. Normalized curves of the observed surface spectral reflectances of the lemon.

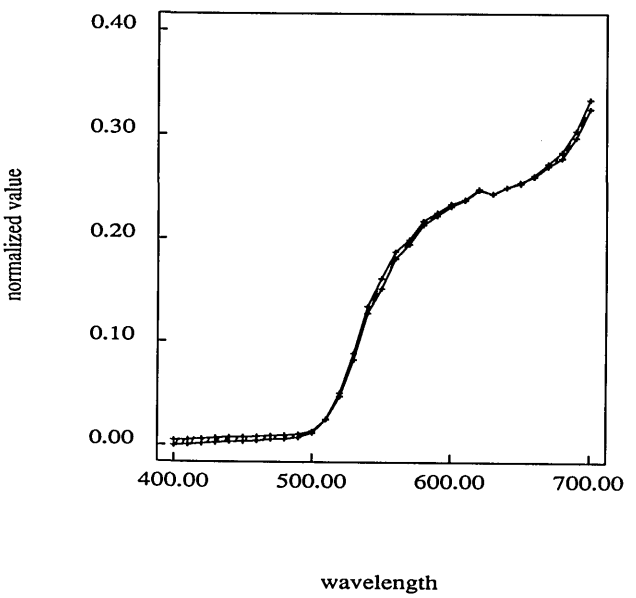

Fig. 9. Estimate of the band of the subsurface reflectance functions of the lemon.

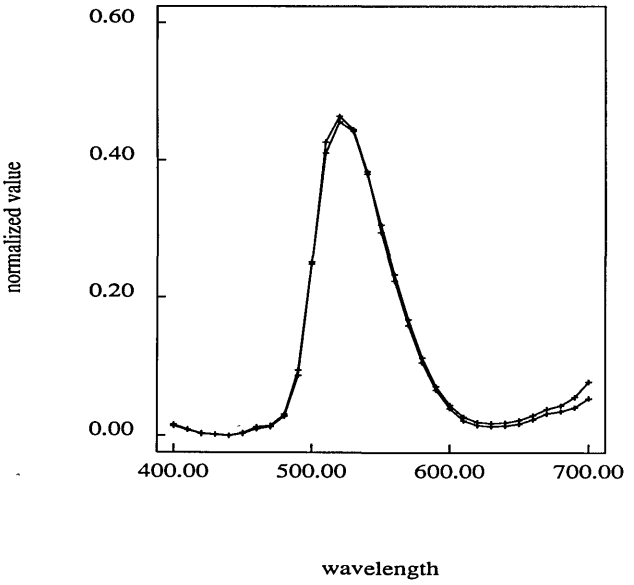

Fig. 10. Comparison between the estimated physical limits of the subsurface reflectance function of the green ashtray in two cases, in which a flood lamp and a slide projector were used as the light sources.

C. Accuracy Check

We have repeated the experiment described in Subsection 4.A, using plastics under a different light source. Because we have two light sources, as is shown in Fig. 3, this time we have used the slide projector to illuminate the two plastic surfaces. The illuminant spectrum is first estimated on these reflected lights. We have confirmed that the shape of the estimated illuminant spectrum is almost the same as the curve from fruits (filled squares in Fig. 3). Next the observed surface reflectances of the green ashtray are analyzed in the same way. The resulting estimate of the subsurface reflectance function is compared with the result in Fig. 7, which was obtained under the flood lamp. Note that the theoretical limits of the subsurface reflectances should be the same in the two experiments with different light sources, although the reflectances of real measurements are different. Figure 10 shows a comparison of the physical limit curve of the subsurface reflectance function between the two experiments. The two curves are coincident to within 2%. These results show the reliability of the whole process of the illuminant spectral estimation and the subsurface reflectance estimation.

5. CONCLUSION

We have described a method for estimating the subsurface (diffuse) component of the surface reflectance function for inhomogeneous materials. The method is based on several physical constraints that apply to the surface reflectance functions. The method yields an estimated band of the spectral reflectance functions as possible solutions for the subsurface reflectance component.

REFERENCES

1. G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," Int. J. Comput. Vision **2**, 7–32 (1988).
2. S. Tominaga and B. A. Wandell, "The standard surface reflectance model and illuminant estimation," J. Opt. Soc. Am. A **6**, 576–584 (1989).
3. M. D'Zmura and P. Lennie, "Mechanisms of color constancy," J. Opt. Soc. Am. A **3**, 1662–1672 (1986).
4. H. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," J. Opt. Soc. Am. A **3**, 1694–1699 (1986).
5. B. K. P. Horn and R. W. Sjoberg, "Calculating the reflectance map," Appl. Opt. **4**, 1770–1779 (1979).
6. W. H. Lawton and E. A. Sylvestre, "Self modeling curve resolution," Technometrics **13**, 617–633 (1971).
7. S. Kawata, K. Sasaki, and S. Minami, "Component analysis of spatial and spectral patterns in multispectral images. I. Basis," J. Opt. Soc. Am. A **4**, 2101–2106 (1987).