Separating Specular, Diffuse, and Subsurface Scattering Reflectances from Photometric Images*

Tai-Pang Wu and Chi-Keung Tang

Vision and Graphics Group
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Abstract. While subsurface scattering is common in many real objects, almost all separation algorithms focus on extracting specular and diffuse components from real images. In this paper, we propose an appearance-based approach to separate non-directional subsurface scattering reflectance from photometric images, in addition to the separation of the off-specular and non-Lambertian diffuse components. Our mathematical model sufficiently accounts for the photometric response due to non-directional subsurface scattering, and allows for a practical image acquisition system to capture its contribution. Relighting the scene is possible by employing the separated reflectances. We argue that it is sometimes necessary to separate subsurface scattering component, which is essential to highlight removal, when the object reflectance cannot be modeled by specular and diffuse components alone.

1 Introduction

Photometric appearance of an object depends on surface geometry, material property, illumination and viewing direction of the camera. The ubiquitous Lambertian assumption used in computer vision algorithms is seldom satisfied, making them error prone to specular highlight, off-specular reflection, and subsurface scattering. Almost all reflectance separation algorithms separate specular and Lambertian diffuse reflectances. Lin and Lee [8] made use of the Lafortune model [5] to separate off-specular and non-Lambertian diffuse components. The Lafortune model assumes the bidirectional reflectance distribution function (BRDF), which describes the ratio of outgoing radiance to incoming irradiance at the *same* surface point locally.

However, many real objects such as wax, paper, and objects with translucent protective coating do not belong to this category. Subsurface scattering is common, where the outgoing radiance observed at a surface point may be due to the incoming irradiances at *different* surface points. Fig. 1 illustrates a common phenomenon. To account for subsurface scattering, the more general bi-directional surface scattering reflectance distribution function (BSSRDF) should be used [4]. However, BSSRDF is very difficult to capture without knowledge of object geometry and material property.

In this paper, we propose an appearance-based method, and derive the mathematical model for reflectance separation that accounts for non-directional subsurface scattering,

* This work is supported by the Research Grant Council of Hong Kong Special Administration Region, China: HKUST6193/02E.

T. Pajdla and J. Matas (Eds.): ECCV 2004, LNCS 3022, pp. 419–433, 2004.

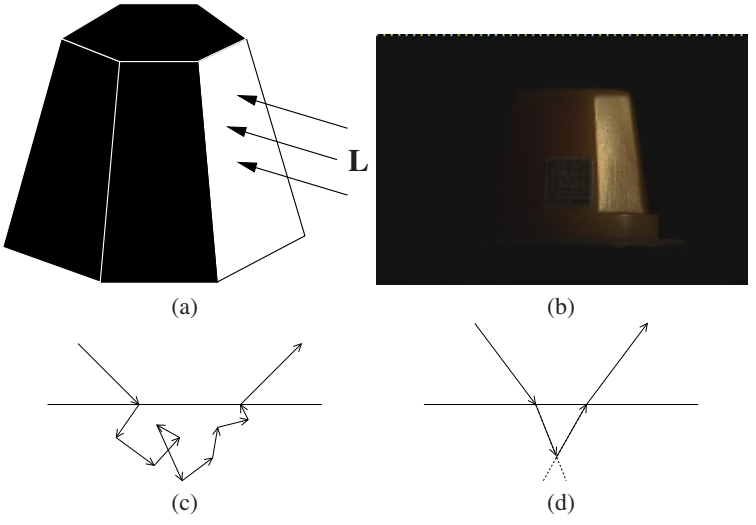


Fig. 1. Ideally, only the face visible to the directional light **L** can be seen (a), if the object reflectance is explained by BRDF. In reality, due to subsurface scattering, significant radiance is observed on other faces (b). In this paper, we model non-directional subsurface scattering (c), but not single scattering (d) where a dominant outgoing radiance direction exists.

Fig. 1. We do not model single scattering [2], which has a dominant direction in the outgoing radiance. Given images only, we decompose the input into three photometric components: *off-specular* I_s, non-Lambertian or *local diffuse* D, and *non-directional subsurface scattering* I_{scat}. BSSRDF can describe I_s, D and I_{scat}, but BRDF can only describe the local I_s and D. Our approach avoids the difficult problem of capturing/recovering the BRSSDF. Yet, it is versatile enough to describe a wide range of lighting conditions, re-light the scene and remove specular highlight. Transparency and single scattering are the topics of our ongoing research.

2 Related Work

We review some representative reflectance models used in computer vision and computer graphics. In computer vision, given images, reflectance models are often used in parameter fitting for seperating specular and diffuse components. In computer graphics, given parameters to the model, photorealistic images are rendered to simulate specular and diffuse effects.

Forward approach: graphics rendering. Cook and Torrance [1] introduced a reflectance model to describe the distribution of the reflected light and color shift as the reflectance changes with incidence angle. Later, Poulin and Fournier [10] introduced another model to handle anisotropic reflection. This model considers the micro-structure of material surface as small cylinders. Oren and Nayar [9] proposed a model to describe complex

geometry and radiometric phenomena to account for non-Lambertian diffuse reflection. They assume or approximate light transport by bi-directional reflection distribution function (BRDF). To account for reflectance due to subsurface scattering, Hanrahan and Krueger [2] introduced a model that considers the subsurface reflectance due to back-scattering in a layered medium, by using transport theory. However, this model cannot capture the diffusion of light across shadow boundaries. Jensen et al. [4] introduced another model to overcome this problem. This model approximates the bi-directional surface scattering reflectance distribution function (BSSRDF), which does not restrict the light ray to enter and exit the surface at the same location. In fact, BRDF is the special case of BSSRDF.

Although the rendering method by Jensen et al. [4] predicts the appearance of translucent materials very well, complete knowledge on the scene (the geometry of the scene, object materials, lighting condition and camera configuration) should be known. Shape-from-shading method may be used to estimate the surface geometry information. For example, Ragheb and Hancock [11] used an iterated conditional modes algorithm to estimate the Lambertian and specular reflectance components and recovering local surface normals. Hertzmann and Seitz [3] introduced a method to recover the local surface normals with a reference object whose geometry is known. However, current methods use BRDF, but not the more general BSSRDF, which is capable of explaining subsurface scattering.

Backward approach: reflectance separation. When accurate geometry and reflectance information are not available, image-based methods are used. The camera position is usually fixed. Shashua [12] proved that, if Lambertian diffusion is assumed, images illuminated under any lighting direction can be produced by linear combination of three photometric images captured under three linearly independent lighting vectors.

Lin and Lee [7] derived an approximate specular reflectance model based on the Torrance-Sparrow model, where the local specular effect can be expressed by logarithms of three intensity-normalized photometric images under certain illumination and surface conditions. In [6], they introduced a method for representing diffuse and specular reflections for objects of uniform surface roughness using four photometric images. To relight the image, only four images are needed without the knowledge of surface geometry. By extending the same idea to the Lafortune model [5], Lin and Lee [8] introduced a method capable of separating any number of reflectances into distinct, manageable components such that the model parameters can be estimated readily. Any novel lighting condition can be simulated to relight images, by non-linear combination of the photometric images of the separated reflectances and the model parameters.

Our contribution. We show that by transforming light vectors in the set of equations, Lin and Lee's method can be generalized to include a non-directional subsurface scattering term I_{scat} in the diffuse term. By ensuring invariance in the resulting appearance and preserving outgoing radiance energy, the proposed transformation of light vectors is legitimate.

In implementation, our key mathematical derivation is translated into the estimation of point spread functions (PSF) to estimate the I_{scat} term. A new and practical image

acquisition system is built to capture subsurface scattering reflectance from real objects of unknown geometry.

Potential application is the removal of the specular highlight term I_s in the presence of subsurface scattering I_{scat} (e.g., paper, wax). Both are problematic to many computer vision algorithms. Alternatively, I_s, D, and I_{scat} can be combined non-linearly to simulate a novel image under a different lighting condition.

The rest of our paper is organized as follows. Section 3 discusses different choices of reflectance models. In section 4, we derive our new reflectance representation for photometric images. Section 5 discusses how the parameters of our reflectance representation can be estimated. The new image acquisition method is introduced in Section 6. Section 7 presents the implementation and shows our results. We conclude our work in Section 8.

3 Reflectance Models

Except for mirror-like surface material such as metal, most material exhibits certain extent of translucence which is resulted by subsurface scattering. This phenomenon is described by BSSRDF. Mathematically, BSSRDF S is defined as:

$$dL_o(x_o, \overrightarrow{\omega_o}) = S(x_i, \overrightarrow{\omega_i}; x_o, \overrightarrow{\omega_o}) d\Phi_i(x_i, \overrightarrow{\omega_i}) \tag{1}$$

where $L_o(x_o, \overrightarrow{\omega_o})$ is the outgoing radiance at point x_o in direction $\overrightarrow{\omega_o}$, $\Phi_i(x_i, \overrightarrow{\omega_i})$ is the incident flux at the point x_i from direction ω_i. When $i = o$, S becomes BRDF. By integrating the incident irradiance over all incoming directions and area A on the surface, the total outgoing radiance at x_o is:

$$L_o(x_o, \overrightarrow{\omega_o}) = \int_A \int_{2\pi} S(x_i, \overrightarrow{\omega_i}; x_o, \overrightarrow{\omega_o}) L_i(x_i, \overrightarrow{\omega_i}) (\overrightarrow{n_i} \cdot \overrightarrow{\omega_i}) d\overrightarrow{\omega_i} dA(x_i) \tag{2}$$

Actually, $L_i(x_i, \overrightarrow{\omega_i})(\overrightarrow{n_i} \cdot \overrightarrow{\omega_i}) d\overrightarrow{\omega_i}$ is the energy of the incident flux defined by Cook and Torrance [1]. In other words, this equation describes how the incident flux (energy) is distributed from location x_i to x_o. With this equation, a good prediction on the appearance of translucent objects is possible. However, S should be known *a priori* for all x and ω. Recovering the BSSRDF S is difficult without special equipment. It is unknown if a gonioreflectometer (used to capture BRDF) can be used to measure BSSRDF.

An alternative is to approximate S by fitting analytical functions. However, almost all reflectance models approximate BRDF, except a few on BRSSDF [4]. Our appearance-based model is capable of representing a class of BRSSDF, by generalizing the Lafortune model [5]. Lin and Lee [8] investigated the Lafortune model, which considers each type of reflectance as a non-linear parametric primitive function. The final object appearance is the linear combination of these primitives. This model provides high flexibility. The other possible choice is the Cook and Torrance model [1]. This model gives a good prediction on specular effect and color shift at grazing angle. Although Lin and Lee [6] showed that the parameters of Cook and Torrance model can be estimated under limited viewing conditions, it is still too complex to use. As a result, we chose the Lafortune model. The Lafortune model [5] is defined as

$$R(\mathbf{L}, \mathbf{V}) = \sum_i [C_{x,i} L_x V_x + C_{y,i} L_y V_y + C_{z,i} L_z V_z]^{n_i} \tag{3}$$

where i is the index of primitives, $\mathbf{L} = (L_x, L_y, L_z)^T$ is light direction, $\mathbf{V} = (V_x, V_y, V_z)^T$ is viewing direction, C is weighting coefficient and x, y and z index a local coordinate system with the z-axis aligned with the surface normal, and the x-axis and y-axis aligned with the principal directions of anisotropy [5], except for unusual type of anisotropy. According to [8], since the z-axis is aligned with a surface normal, we have

$$L_z V_z = (\mathbf{N} \cdot \mathbf{L})(\mathbf{N} \cdot \mathbf{V}) \tag{4}$$

$$L_x V_x + L_y V_y = (\mathbf{L} \cdot \mathbf{V}) - (\mathbf{N} \cdot \mathbf{L})(\mathbf{N} \cdot \mathbf{V}) \tag{5}$$

With this relationship, Lin and Lee [8] showed that the non-Lambertian diffuse and off-specular components, $I_d(x)$ and $I_s(x)$, at pixel x on image I can be expressed as:

$$I_d(x) = \rho(x) \frac{n_d + 2}{2\pi} [(\mathbf{N}(x) \cdot \mathbf{L})(\mathbf{N}(x) \cdot \mathbf{V})]^{n_d} \tag{6}$$

$$I_s(x) = [C_1(\mathbf{L} \cdot \mathbf{V}) + C_2(\mathbf{N}(x) \cdot \mathbf{L})(\mathbf{N}(x) \cdot \mathbf{V})]^{n_s} \tag{7}$$

where $\rho(x)$ is surface albedo, $\mathbf{N}(x)$ is normal at x, n_d and n_s are the exponents for non-Lambertian diffuse and off-specular components respectively, $C_1 = C_x$ and $C_2 = C_z - C_x$.

In [1], it is stated that the specular component is the result of reflection from the surface of material, and the diffuse component is that of internal scattering or multiple surface reflections. Internal scattering is due to the penetration of incident light beneath the material surface. With this definition, we embed the effect of subsurface scattering into the diffuse term in the equations.

4 Reflectance Representation

We assume directional light source in our derivation. The following notations are useful in the discrete formulation for our implementation: a small surface patch on a 3D object is represented by $\overrightarrow{v}$. x is the (quantized) image point of $\overrightarrow{v}$. $\mathbf{L}$ as the directional lighting vector, which encapsulates both the direction and intensity, $\mathbf{N}$ is an overloaded function that returns the normal direction at $\overrightarrow{v}$, or the normal direction at x ($\mathbf{N}(\overrightarrow{v}) = \mathbf{N}(x)$ if the image of $\overrightarrow{v}$ is x). $m_{\overrightarrow{v_0}}(\overrightarrow{v_i})$ is the discrete representation for the BSSRDF S between $\overrightarrow{v_0}$ and $\overrightarrow{v_i}$, which specifies how the energy of incoming ray distributes from $\overrightarrow{v_i}$ to $\overrightarrow{v_0}$.

Using the above notations, the discrete version of Eqn. (2) is re-written as:

$$energy(\overrightarrow{v_0}, \mathbf{L}) = \sum_{i=0}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i}) \mathbf{N}(\overrightarrow{v_i}) \cdot \mathbf{L} \tag{8}$$

Without loss of generality, $energy(\overrightarrow{v_0}, \mathbf{L})$ indicates the total outgoing radiance at $\overrightarrow{v_0}$, n is the total number of surface patch on the surface, and i is the index of the subsurface patch. It is important to note that, for m, the only input parameter is a surface patch $\overrightarrow{v}$, which is different from the BSSRDF S described in section 3 requiring both the position and direction of incident and outgoing rays. The main reason is that we do not model the effect of single scattering, which is our future work. It requires the knowledge of light ray directions.

From Eqn. (8), we can consider that m is a PSF which models the diffusion of light due to *local diffuse* D and *subsurface scattering* I_{scat}. When a light ray enters a highly scattering medium from an arbitrary direction at $\overrightarrow{v}$, the light distribution is non-directional, since the light is scattered and absorbed randomly inside the medium. Since the PSF depends on surface geometry and material property of the object, the PSF at each $\overrightarrow{v}$ is different in general. Therefore, the subscript of $m_{\overrightarrow{v_0}}(\overrightarrow{v_i})$ reflects that it is a local variable, and surface patch $\overrightarrow{v_i}$ is the only input parameter.

By rotating the local coordinate system of each surface patch $\overrightarrow{v_i}$ with rotation matrix ψ_i such that $\mathbf{N}(\overrightarrow{v_i})$ aligns with $\mathbf{N}(\overrightarrow{v_0})$, Eqn. (8) becomes:

$$\begin{aligned} energy(\overrightarrow{v_0}, \mathbf{L}) &= \sum_{i=0}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i})\mathbf{N}(\overrightarrow{v_0}) \cdot (\psi_i \mathbf{L}) \\ &= \mathbf{N}(\overrightarrow{v_0}) \cdot [\sum_{i=0}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i})(\psi_i \mathbf{L})] \qquad (9) \\ &= \mathbf{N}(\overrightarrow{v_0}) \cdot \mathbf{L}_{final} \qquad (10) \end{aligned}$$

where $\mathbf{L}_{final} = \sum_{i=0}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i})(\psi_i \mathbf{L})$. From Eqn. (10), the total outgoing radiance at $\overrightarrow{v_0}$ can be considered as resizing and jittering of the original $\mathbf{L}$ to produce $\mathbf{L}_{final}$. By putting Eqn. (10) into (6), we express $I_d(x)$ as:

$$\begin{aligned} &\rho(x)\frac{n_d+2}{2\pi}[(\mathbf{N}(\overrightarrow{v_0}) \cdot \mathbf{L}_{final})(\mathbf{N}(x) \cdot \mathbf{V})]^{n_d} \\ &= \rho(x)\frac{n_d+2}{2\pi}[(\mathbf{N}(x) \cdot \mathbf{L}_{final})(\mathbf{N}(x) \cdot \mathbf{V})]^{n_d} \\ &= \rho(x)\frac{n_d+2}{2\pi}\{[m_{\overrightarrow{v_0}}(\overrightarrow{v_0})\mathbf{N}(x) \cdot (\psi_0 \mathbf{L}) + \mathbf{N}(x) \cdot [\sum_{i=1}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i})(\psi_i \mathbf{L})]][\mathbf{N}(x) \cdot \mathbf{V}]\}^{n_d} \qquad (11) \\ &= \rho(x)\frac{n_d+2}{2\pi}\{[m_{\overrightarrow{v_0}}(\overrightarrow{v_0})\mathbf{N}(x) \cdot \mathbf{L} + \mathbf{N}(x) \cdot \mathbf{L}_{scat}][\mathbf{N}(x) \cdot \mathbf{V}]\}^{n_d} \qquad (12) \end{aligned}$$

where ψ_0 is defined to be an identity matrix, and $\mathbf{L}_{scat} = \sum_{i=1}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i})(\psi_i \mathbf{L})$. Then, the n_d-th root of Eqn. (12) is

$$\begin{aligned} I_d(x)^{\frac{1}{n_d}} &= [\rho(x)\frac{n_d+2}{2\pi}]^{\frac{1}{n_d}}[m_{\overrightarrow{v_0}}(\overrightarrow{v_0})\mathbf{N}(x) \cdot \mathbf{L}][\mathbf{N}(x) \cdot \mathbf{V}] \\ &+ [\rho(x)\frac{n_d+2}{2\pi}]^{\frac{1}{n_d}}[\mathbf{N}(x) \cdot \mathbf{L}_{scat}][\mathbf{N}(x) \cdot \mathbf{V}] \\ &= m_{\overrightarrow{v_0}}(\overrightarrow{v_0})I_{local}(x)^{\frac{1}{n_d}} + I_{scat}(x)^{\frac{1}{n_d}} \qquad (13) \end{aligned}$$

In Eqn. (13), I_{local} and I_{scat} have the same form as Eqn. (6), which contribute to the diffuse component of the outgoing radiance I_d. I_{local} accounts for the outgoing radiance locally reflected at $\overrightarrow{v_i}$ without any subsurface scattering (since ψ_0 is an identity matrix). I_{scat} accounts for the outgoing radiance resulting from incident irradiances at different patches. Using the PSF m, it can be thought that appropriate portions of $\mathbf{L}$ are "distributed" to I_{local} and I_{scat}. Putting together, a captured image $I(x) = I_d(x) + I_s(x)$ is given by:

$$I(x) = [m_{\overrightarrow{v_0}}(\overrightarrow{v_0})I_{local}(x)^{\frac{1}{n_d}} + I_{scat}(x)^{\frac{1}{n_d}}]^{n_d} + I_s(x) \qquad (14)$$

Since any lighting vector $\mathbf{L}$ could be written as the linear combination of three linear independent lighting vectors $\mathbf{L_1}, \mathbf{L_2}$ and $\mathbf{L_3}$ such that

$$\mathbf{L} = \alpha_1 \mathbf{L_1} + \alpha_2 \mathbf{L_2} + \alpha_3 \mathbf{L_3}, \tag{15}$$

it is possible to represent the n_d-th root of the diffuse component as expressed in Eqn. (13) by combining three photometric images that are illuminated by $\mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3}$. By putting Eqn. (15) into (11) and taking the n_d-th root in both sides,

$$\begin{aligned}
I_d(x)^{\frac{1}{n_d}} &= [\rho(x)\frac{n_d+2}{2\pi}]^{\frac{1}{n_d}} [m_{\overrightarrow{v_0}}(\overrightarrow{v_0})\mathbf{N}(x)\cdot(\alpha_1\mathbf{L_1}+\alpha_2\mathbf{L_2}+\alpha_3\mathbf{L_3})][\mathbf{N}(x)\cdot\mathbf{V}] \\
&+ [\rho(x)\frac{n_d+2}{2\pi}]^{\frac{1}{n_d}} [\mathbf{N}(x)\cdot[\sum_{i=1}^{n} m_{\overrightarrow{v_0}}(\overrightarrow{v_i})(\psi_i(\alpha_1\mathbf{L_1}+\alpha_2\mathbf{L_2}+\alpha_3\mathbf{L_3}))]][\mathbf{N}(x)\cdot\mathbf{V}] \\
&= m_{\overrightarrow{v_0}}(\overrightarrow{v_0})[\alpha_1 I_{local,1}(x)^{\frac{1}{n_d}} + \alpha_2 I_{local,2}(x)^{\frac{1}{n_d}} + \alpha_3 I_{local,3}(x)^{\frac{1}{n_d}}] \\
&+ [\rho(x)\frac{n_d+2}{2\pi}]^{\frac{1}{n_d}} [\mathbf{N}(x)\cdot(\alpha_1\mathbf{L}_{scat,1}+\alpha_2\mathbf{L}_{scat,2}+\alpha_3\mathbf{L}_{scat,3})][\mathbf{N}(x)\cdot\mathbf{V}] \\
&= m_{\overrightarrow{v_0}}(\overrightarrow{v_0})[\alpha_1 I_{local,1}(x)^{\frac{1}{n_d}} + \alpha_2 I_{local,2}(x)^{\frac{1}{n_d}} + \alpha_3 I_{local,3}(x)^{\frac{1}{n_d}}] \\
&+ [\alpha_1 I_{scat,1}(x)^{\frac{1}{n_d}} + \alpha_2 I_{scat,2}(x)^{\frac{1}{n_d}} + \alpha_3 I_{scat,3}(x)^{\frac{1}{n_d}}]
\end{aligned} \tag{16}$$

Similarly, by putting Eqn. (15) into (7) and taking the n_s-th root, the specular component can be represented by:

$$I_s(x)^{\frac{1}{n_s}} = \alpha_1 I_{s,1}(x)^{\frac{1}{n_s}} + \alpha_2 I_{s,2}(x)^{\frac{1}{n_s}} + \alpha_3 I_{s,3}(x)^{\frac{1}{n_s}} \tag{17}$$

Eqns (16) and (17) show that any photometric image can be represented by combining three images captured at independent lighting directions $\mathbf{L_1}, \mathbf{L_2}$ and $\mathbf{L_3}$.

5 Issues in Parameter Estimation

After deriving our reflectance separation, next we need to estimate the associated parameters. The desired method is to estimate the parameters directly from images. To achieve this, the method suggested by Lin and Lee [8] is a possible choice. First, the diffuse and specular components are related. Second, six images, acquired with different lighting vectors, are divided into two sets with three images each. Images in one set are described by the combination of the other set. Finally, the parameters of their representation are estimated by optimizing an objective function. Here, we need to find the relationship between the specular and our diffuse component that includes a subsurface scattering term.

If an object does not exhibit any light scattering effect, the term $I_{scat}(x)$ is zero for all the pixels on an captured image. Then, our derivation will be the same as [8]. Thus, the formulation of Lin and Lee [8] is a special case of ours. From [8], the relationship between local diffuse $I_{local,k}$ and specular component $I_{s,k}$ is:

$$I_{s,k}(x)^{\frac{1}{n_d}} = a_k + b(x) I_{local,k}(x)^{\frac{1}{n_d}} \tag{18}$$

where $k = 1, 2, 3$ are the indices of images, $a_k = C_1 \mathbf{L_k}\cdot\mathbf{V}$ and $b(x) = C_2/[\rho(x)\frac{n_d+2}{2\pi}]^{\frac{1}{n_d}}$. With this relationship, the image I_k acquired can be expressed as:

$$I_k(x) = [m_{\overrightarrow{v_0}}(\overrightarrow{v_0}) I_{local,k}(x)^{\frac{1}{n_d}} + I_{scat,k}(x)^{\frac{1}{n_d}}]^{n_d} + [a_k + b(x) I_{local,k}(x)^{\frac{1}{n_d}}]^{n_s} \quad (19)$$

In order to simplify Eqn. (19), we define $D_k(x) = m_{\overrightarrow{v_0}}(\overrightarrow{v_0})^{n_d} I_{local,k}(x)$ and $B(x) = C_2 / [m_{\overrightarrow{v_0}}(\overrightarrow{v_0})^{n_d} \rho(x) \frac{n_d+2}{2\pi}]^{\frac{1}{n_d}}$. Eqn. (19) becomes:

$$I_k(x) = [D_k(x)^{\frac{1}{n_d}} + I_{scat,k}(x)^{\frac{1}{n_d}}]^{n_d} + [a_k + B(x) D_k(x)^{\frac{1}{n_d}}]^{n_s} \quad (20)$$

Similarly, adding the powers of n_d and n_s of Eqns (16) and (17) respectively:

$$\begin{aligned} I_k(x) = & [(\alpha_{1,k} D_1(x)^{\frac{1}{n_d}} + \alpha_{2,k} D_2(x)^{\frac{1}{n_d}} + \alpha_{3,k} D_3(x)^{\frac{1}{n_d}}) \\ & + (\alpha_{1,k} I_{scat,1}(x)^{\frac{1}{n_d}} + \alpha_{2,k} I_{scat,2}(x)^{\frac{1}{n_d}} + \alpha_{3,k} I_{scat,3}(x)^{\frac{1}{n_d}})]^{n_d} \\ & + \{\alpha_{1,k}[a_1 + B(x) D_1(x)^{\frac{1}{n_d}}] \\ & + \alpha_{2,k}[a_2 + B(x) D_2(x)^{\frac{1}{n_d}}] \\ & + \alpha_{3,k}[a_3 + B(x) D_3(x)^{\frac{1}{n_d}}]\}^{n_s} \end{aligned} \quad (21)$$

From the above two equations, we can see that the parameters needed to be estimated consist of five global parameters $\{a_1, a_2, a_3, n_d, n_s\}$ and seven local parameters $\{B(x), D_1(x),\ D_2(x), D_3(x), I_{scat,1}(x), I_{scat,2}(x), I_{scat,3}\}$. By using the framework of Lin and Lee [8], we have to capture n images that are illuminated by n different lighting vectors. Three of them ($k = 1, 2, 3$) are represented by Eqn. (20), and the others ($k \geq 4$) are represented by Eqn (21). Since we have $7p + 5$ unknowns, where p is the total number of pixels in image I, $n \geq 8$ images are needed.

The values $\alpha_{i,k}$ for each of the image $k \geq 4$ may be approximated from pixels whose intensity does not exceed a threshold by the method suggested in [12]. In order to estimate the global and local parameters, we minimize the following error function:

$$Err = \sum_x [\lambda_1 \sum_{k=1}^{3} e(k,x)^2 + \lambda_2 \sum_{k=4}^{n} e(k,x)^2] \quad (22)$$

where $e(k,x)$ is an error function for $I_k(x)$ and $\lambda_1 = 1$ and $\lambda_2 = 2$ are weighting coefficients (as in [8]). In our implementation, $e(k,x)$ is image difference. The function *Err* is then minimized by some optimization algorithm such as Levenberg-Marquardt algorithm. It is suggested that the global parameters are estimated first before the local parameters are estimated pixel-by-pixel in order to reduce the computation load. The former can be done by selecting $p \geq \lceil (n-7)p + 5 \rceil$ pixels from the image and perform optimization with Eqn. (22).

Although everything above seems to be fine so far, there are some reasons that the above framework cannot be used directly for solving our reflectance Eqns (20) and (21). One problem is that there are too many local parameters, leading to a lot of local minima. Besides, too many images ($n - 3$, where $n \geq 8$) are required for the approximation of $\alpha_{i,k}$. Both problems complicate the optimization process and make it less stable. Worst, the local variable $I_{scat,k}(x)$ is *independent* of $D_k(x)$ and $I_s(x)$ in Eqn. (20). Given a pixel x, during the optimization, if specular component does not exist, the value of diffuse term $I_d(x)$ can be "distributed" to $D_k(x)$ and $I_{scat,k}(x)$ arbitrarily to produce (local) minima.

To solve these problems, some parameters should preferably be estimated first. Fixing them before the optimization process will make it more stable. Suppose that the value of $I_{scat,k}(x)$, where $k = 1, 2, 3$, can be found before the optimization stage, we shall have five global and four local parameters to estimate, and the number of images required becomes $n \geq 6$. Then, the number of parameters and the number of required images are the same as in [8].

In the following section, a new image acquisition system is built to estimate $I_{scat,k}(x)$.

6 Image Acquisition for Estimating I_{scat}

Except at the end of this section, we drop the subscript k since the same image acquisition method is used for estimating all $I_{scat,k}$. Recall from Eqn. (13) that the PSF term m is included into the subsurface scattering term I_{scat}.

If there is a *single* light ray incident at a surface patch at $\overrightarrow{v_0}$, we can warp the PSF on the object surface centered at $\overrightarrow{v_0}$. Refer to Eqn. (20), the scattering term I_{scat} is completely independent of the D and I_s terms. Therefore, given a single ray incident at $\overrightarrow{v_0}$, except at $\overrightarrow{v_0}$, the outgoing radiance at other surface patches $\overrightarrow{v_i}, i \neq 0$, is due to subsurface scattering *only*. For these patches at $\overrightarrow{v_i}$, the corresponding images D and I_s are zero. Therefore, if we could produce such a single light ray incident at each $\overrightarrow{v}$ on the object surface, it would be possible to recover I_{scat}. However, it is difficult to build high precision equipment capable of shooting a single ray of light. A light slab produced by a ray box used in optics experiments covers more than one surface patches, as shown in Fig. 2(a). Following, we show how we deploy a light slab to estimate I_{scat}.

Analysis. Consider $I_{scat}(x)$ in Eqn. (13). By expanding $\mathbf{L}_{scat}$, which is the linear combination of rotated and resized $\mathbf{L}$ from n surface patches, the relationship between energy contribution from each of the surface patch and $I_{scat}(x)$ can be written as:

$$I_{scat}(x)^{\frac{1}{n_d}} = I_{v,1}(x)^{\frac{1}{n_d}} + I_{v,2}(x)^{\frac{1}{n_d}} + ... + I_{v,n}(x)^{\frac{1}{n_d}} \tag{23}$$

where each $I_{v,i}(x)$ represents the images of light diffusion resulting from surface patch i. It shows that $I_{scat}(x)^{1/n_d}$ is just the linear combination of $I_{v,i}(x)^{1/n_d}$. Consider that a light slab of width q is incident to more than one surface patches. Due to subsurface scattering, other surface patches that are not directly illuminated also produces non-zero reflectance. Suppose we sweep the light slab by translation so that it covers all visible patches on the surface. Let us assume for now that the effect of local diffuse D and specular I_s are absent, the summation of the n_d-th root of all captured images is:

$$qI_{v,1}(x)^{\frac{1}{n_d}} + qI_{v,2}(x)^{\frac{1}{n_d}} + ... + qI_{v,n}(x)^{\frac{1}{n_d}} = qI_{scat}(x)^{\frac{1}{n_d}} \tag{24}$$

which means that the image sum resulting by the sweeping light slab is simply a scaled-up version of the original $I_{scat}(x)$ of the images. Thus, as long as we can remove the contribution of D and I_s, Eqn. (24) allows us to use a light slab to estimate I_{scat}, which is easier to produce in practice.

Steps. Our image acquisition procedure consists of the following steps.

1. First, we capture an image $I(x)$, which is illuminated by $\mathbf{L}$.
2. Then, we sweep a slab of rays parallel to $\mathbf{L}$ in constant speed, and capture a sequence of images $I_{seq,i}(x)$. $I(x)$ and $I_{seq,i}(x)$ are captured by the same digital video camera, using the same, fixed viewpoint.

After that, we set up the following equation:

$$CI(x)^{\frac{1}{n_d}} = D(x)^{\frac{1}{n_d}} + I_{seq,1}(x)^{\frac{1}{n_d}} + I_{seq,2}(x)^{\frac{1}{n_d}} + ... + I_{seq,p}(x)^{\frac{1}{n_d}} \qquad (25)$$

where p is total number of frames and $C = q \times g$. Variable g is the velocity of the sweeping plane of rays. We have to consider this term because the speed of the plane may not be one patch per frame. Since Eqn. (25) only contains two unknowns, n_d and C, we can choose $p \geq 2$ pixels x from a region where specular effect is not significant to solve it, by using some standard optimization algorithm.

Then, we use simple thresholding to remove the contributions of D and I_s for each $I_{seq,i}$ since, by observation, the local diffuse and specular terms in the illuminated area are often much larger than the non-local subsurface scattering term in other area (Fig. 2(a)). The resulting hole after thresholding is shown in Fig. 2(b).

However, for $q > 1$, thresholding removes not only $D(x)$ and $I_s(x)$, but also some $I_{v,i}(x)$. Therefore, the missing $I_{v,i}(x)$ should be found. Since the response of the PSF for each $I_{seq,i}$ is small, quantization must produce some error. One suggestion to fill the hole is to treat each $I_{seq,i}$ as a 3D surface point with color values as z-coordinates. Then, the holes are filled by interpolation. Another suggestion is simple. If the width of light slab is small, the holes are just filled by the maximum color value surround the hole after thresholding. It still produces reasonable results. Alternatively, we can sweep another slab of rays in a direction *orthogonal* to the first sweeping slab. To recover the missing $I_{v,i}$, the following relation is used: $A \cup B = A + B - A \cap B$, where A is the contribution to I_{scat} from the first sweeping slab, and B is the contribution to I_{scat} from the second sweeping slab. The holes produced in the images from the first sweep can be filled in by the corresponding pixels in the images obtained after the second sweep. We call the resulting image an PSF image (Fig 2(c)).

Finally, we produce the $I_{scat}(x)^{1/n_d}$ by summing up the n_d-th root of the PSF images, and then scale the image sum by the C we found earlier.

By using our proposed image acquisition method, not only $I_{scat,k}(x)$ can be robustly recovered for $k = 1, 2, 3$, but also n_d. These parameters are fed into the algorithm described in section 5, thus reducing the number of local and global parameters to four respectively: four global parameters $\{a_1, a_2, a_3, n_s\}$, and four local paramters $\{B(x), D_1(x), D_2(x), D_3(x)\}$. This makes our parameter estimation even more stable. Besides, the image acquisition step 1 and 2 mentioned above can be used to capture all the images $I_k(x)$.

7 Implementation and Results

We experiment our reflectance separation on real images, and use the separated components to perform novel view synthesis and highlight removal. Fig. 3 shows the image

(a) (b) (c)

Fig. 2. (a) The response of a mango pudding upon illumination by a light slab. This is $I_{seq,i}$. (b) Thresholded $I_{seq,i}$ to remove local diffuse and specular response. (c) The PSF image.

acquisition system we built. The video camera we used is Sony PC-100 Handycam. A thin light slab is produced by using a cardboard with a vertical slit set in the middle. Constant speed is maintained by moving the cardboard on a track, using a stepper motor with adjustable speed. During the capturing of $I_{seq,i}$, since the slab of rays is thin, the global radiance from the object is very small. Therefore, if we use normal exposure setting for regular image capture, the resulting PSF images will be very dark, and noise may dominate the image. Therefore, we use a different exposure setting for capturing $I_{seq,i}$.

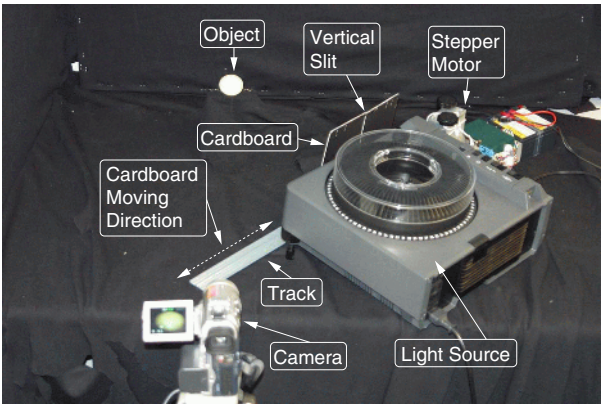


Fig. 3. Experimental setup.

When all the global and local parameters have been estimated using our method in section 5 and 6, we can reconstruct a novel image illuminated by any lighting vector, by setting the corresponding values of $\alpha_1, \alpha_2, \alpha_3$ and using Eqn. (21). We have conducted two experiments on our derived model with two different real objects: mango pudding and wax.

The mango pudding is located about 3 feet from the camera. The ray box can be placed very close to the object since it is set to approximate directional light source. The mango pudding is inside a thin and transparent plastic container which is hexagonal in

shape. The outer surface of the container is very rough so the effect of specular reflection is very small. There is a white paper label with the "ECCV04" logo on the front face. This label is cut up from a piece of white paper, which is translucent under illumination. The label has a different level of translucency compared with that of the pudding. The separated reflectance components for $k = 1$ are shown in Fig. 4(a-d)

The local diffuse component is very dark because the pudding is a highly scattering medium, almost all the energy has contributed to the non-directional subsurface scattering component. Since the paper label is a less scattering media, it is darker than the pudding in the separated I_{scat}. Although the container surface exhibits very little off-specular effect, this experiment shows that our method can still separate it out robustly.

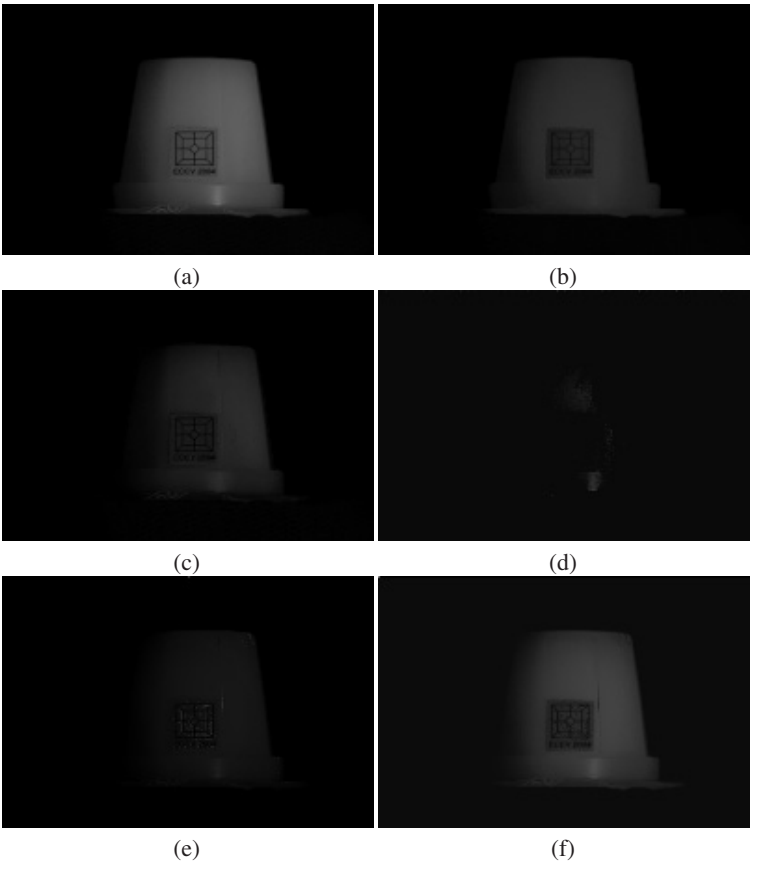
Compared with the result produced by using the model of Lin and Lee [8], as shown in Fig. 4(e) and (f), where only BRDF is assumed, our result is much more reasonable. The off-specular component produced by their model covers the whole pudding because subsurface scattering is not modelled there.

Fig. 5 shows the same pudding re-lit by a novel lighting direction and intensity. Compared with the actual image, the synthesized image looks reasonable. Since the actual image is captured as regular images while our images I_k, $k = 1, 2, 3$, are captured by our new image acquisition method, they exhibit different intensity ranges. Therefore, some artifact can be noticed. It is suggested that all six images should be captured by our new image acquisition method, instead of the essential three images only, to improve the result.

Another object is a piece of wax. The surface of the wax is smooth, therefore the perfect mirror specular effect is very strong, which is not modelled in our reflectance representation. However, we want to test the robustness of our model. The separated reflectance components for $k = 1$ are shown in Fig. 6(a-d). The results produced using our model are similar to that of mango pudding. The local diffuse component D is dark and the subsurface scattering component I_{scat} is bright. Although our reflectance model does not model mirror specular component explicitly, the off-specular component approximates mirror specular effect very well.

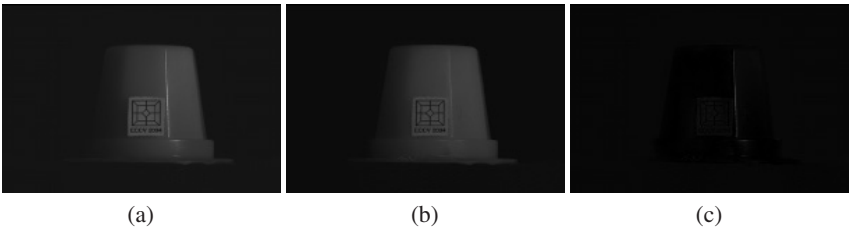
In addition, Fig. 7 shows a novel image constructed by our method. Compared with the actual, ground truth image, the difference of the diffuse components (local and subsurface) is undistinguisable. The image difference is also shown. However, the specular components look quite different when compared with the specular components of the ground truth, which looks much brighter. It is due to mirror specular reflection. Besides, the edge of the wax shows some artifacts. This is because during image acquisition, we have to change the exposure setting of the camera by hand. The remote control of the camera does not provide this function. Some camera movement is unavoidable.

Although we use grey level images to separate reflectance, our separation method can be readily extended to color images. The color results are shown in the supplementary material.



(a) (b) (c) (d) (e) (f)

Fig. 4. Reflectance separation of a scattering medium mango pudding : (a) original image I_k, (b) non-directional subsurface scattering component $I_{scat,k}$, (c) local diffuse component $I_{local,k}$, (d) off-specular component $I_{s,k}$, (e) diffuse component produced by using [8], (f) off-specular component produced by using [8]. We only show the case $k = 1$. See supplementary material of larger images, and for $k = 2, 3$.



(a) (b) (c)

Fig. 5. Novel image synthesis: (a) Real image, (b) synthetic image, (c) image difference between the real and the synthetic image. See supplementary material for larger images.

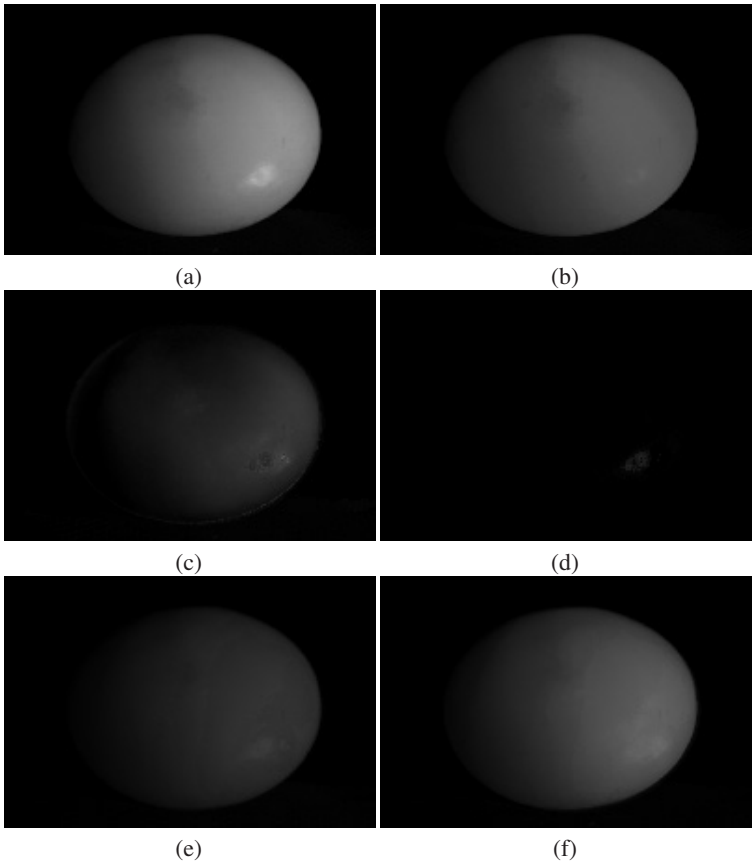


Fig. 6. Reflectance separation of a scattering medium wax : (a) original image I_k, (b) non-directional subsurface scattering component $I_{scat,k}$, (c) local diffuse component $I_{local,k}$, (d) off-specular component $I_{s,k}$, (e) diffuse component produced by using [8], (f) off-specular component produced by using [8]. We only show the case $k = 1$. See supplementary material of larger images, and for $k = 2, 3$.

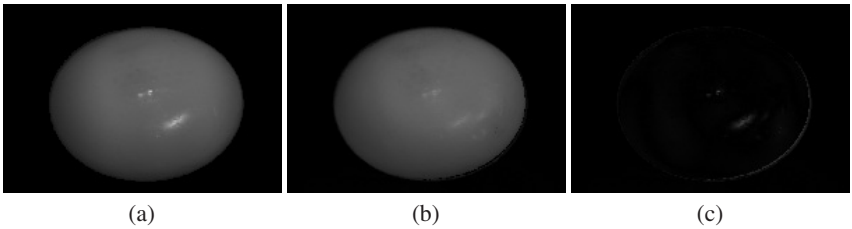


Fig. 7. Novel image: (a) Real image, (b) synthetic image, (c) image difference between the real and the synthetic image. See supplementary material for larger images.

8 Conclusion and Future Work

In this paper, we present an appearance-based model that allows for the separation of off-specular, non-Lambertian local diffuse and non-directional of subsurface scattering reflectance for real world objects. The BSSRDF model is necessary to capture subsurface scattering. A point spread function is estimated. Based on our mathematical derivation, a new and practical image acquisition method is proposed to capture non-directional subsurface scattering component, which complements and improves our parameter estimation process. Some successful reflectance separation experiments were conducted. Faithful novel images under different lighting condition can be generated by using our appearance model. We have also demonstrated improved result on hightlight removel in highly scattering medium, which is traditionally difficult for approaches based on BRDF. Since we do not model the effect of single scattering, our future work focuses on incorporating this component in order to support a wider range of translucent materials.

References

1. R.L. Cook and K.E. Torrance. A reflectance model for computer graphics. *Computer Graphics*, 15:307–316, 1981.
2. P. Hanrahan and W. Krueger. Reflection from layered surfaces due to subsurface scattering. In *SIGGRAPH93*, pages 165–174, 1993.
3. A. Hertzmann and S.M. Seitz. Shape and materials by example: a photometric stereo approach. In *CVPR03*, pages I: 533–540, 2003.
4. H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *SIGGRAPH01*, pages 511–518, 2001.
5. E. P. F. Lafortune, S. Foo, K. E. Torrance, and D. P. Greenberg. Non-linear approximation of reflectance function. In *SIGGRAPH97*, pages 117–126, 1997.
6. S. Lin and S. Lee. Estimation of diffuse and specular appearance. In *ICCV99*, pages 855–860, 1999.
7. S. Lin and S. Lee. A representation of specular appearance. In *ICCV99*, pages 849–854, 1999.
8. S. Lin and S.W. Lee. An appearance representation for multiple reflection components. In *CVPR00*, pages I: 105–110, 2000.
9. M. Oren and S.K. Nayar. Generalization of the lambertian model and implications for machine vision. *IJCV*, 14(3):227–251, April 1995.
10. Pierre Poulin and Alain Fournier. A model for anisotropic reflection. *Computer Graphics*, 24(4):273–282, August 1990.
11. H. Ragheb and E.R. Hancock. Highlight removal using shape-from-shading. In *ECCV02*, page II: 626 ff., 2002.
12. A. Shashua. Geometry and photometry in 3d visual recognition. In *MIT AI-TR*, 1992.