Incorporating the Torrance and Sparrow Model of Reflectance in Uncalibrated Photometric Stereo

Athinodoros S. Georghiades*

Department of Electrical Engineering
Yale University
New Haven, CT 06520-8285, U.S.A.

Abstract

Under the Lambertian reflectance model, uncalibrated photometric stereo with unknown light sources is inherently ambiguous. In this paper, we consider the use of a more general reflectance model, namely the Torrance and Sparrow model, in uncalibrated photometric stereo. We demonstrate that this can not only resolve the ambiguity when the light sources are unknown, but can also result in more accurate surface reconstructions and can capture the reflectance properties of a large number of non-Lambertian surfaces. Our method uses single light source images with unknown lighting and no knowledge about the parameters of the reflectance model. It can recover the 3-D shape of surfaces (up to the binary convex/concave ambiguity) together with their reflectance properties. We have successfully tested our algorithm on a variety of non-Lambertian surfaces demonstrating the effectiveness of our approach. In the case of human faces, the estimated skin reflectance has been shown to closely resemble the measured skin reflectance reported in the literature. We also demonstrate improved recognition results on 4050 images of 10 faces with variable lighting and viewpoint when the synthetic image-based representations of the faces are generated using the surface reconstructions and reflectance properties recovered while assuming the extended reflectance model.

1 Introduction

Assuming Lambertian reflectance [13] in uncalibrated photometric stereo, where the albedo spatially varies and the light sources are unknown, simplifies the problem but has two major drawbacks. First, non-Lambertian objects, such as shiny metals, cannot be reconstructed faithfully, especially when the deviation from the Lambertian assumption is large. Secondly, there is an inherent 9-parameter ambiguity in the surface normals and light source directions [10]. In the case when the integrability constraint is applied to the surface normals, the ambiguity is reduced to the 3-parameter generalized bas-relief (GBR) ambiguity [2, 27, 5]. This ambiguity cannot be resolved any further with photometric data when assuming Lambertian reflectance. Additional information is required.

As a consequence of these two above-mentioned problems, synthetic images of reconstructed non-Lambertian objects viewed under variable lighting will show a lack of the specular component; the shininess of a lacquered object or that of a sweaty face will be missing. Furthermore, synthetic images of these reconstructed objects with variable viewpoint will differ from real images by an unknown image-plane affine transformation [26, 9].

In this paper, we extend uncalibrated photometric stereo to include a non-Lambertian reflectance model, namely the Torrance and Sparrow (T-S) model [23]. With this model, we can capture the reflectance properties of a large number of non-Lambertian surfaces, we can more accurately reconstruct the surface geometry, and we can resolve the 3-parameter GBR ambiguity up to the binary convex/concave ambiguity. Our method assumes no knowledge of the reflectance map, i.e., the light sources and the parameters of the reflectance model are unknown.

When classical photometric stereo was introduced [25, 20, 12], the reflectance map was constrained to be Lambertian and the light sources were assumed to be known. The latter assumption introduces the need for cumbersome calibration which can be difficult to do accurately in practice. In previous work [8, 7], a variant of Lambertian photometric stereo was implemented where no knowledge about the light sources was assumed and where integrability was enforced on the surface normals. The geometry and albedo of a wide variety of surfaces was estimated up to the GBR ambiguity. For faces, their symmetries and similarities were exploited to resolve the three parameters that specify the GBR ambiguity.

In this paper, we show that extending the reflectance model beyond Lambertian can resolve the GBR ambiguity using only photometric data and without resorting to additional extraneous information, such as knowledge of the light source strengths or of some sparse set of albedos. Over the years, many have applied non-Lambertian reflectance models in photometric stereo, but they assumed a known reflectance map and hence knowledge of the light source directions and strengths [22, 3, 14, 16, 21].

Quite recently, an uncalibrated photometric stereo method was presented where the object reflectance is assumed to be the sum of Lambertian and mirror-like reflectance [4]. In that paper, the "consistent viewpoint constraint" is employed on a minimum of four normals of specular points to reduce the ambiguity; in the case when the integrability constraint is also used, the ambiguity is then reduced to the binary convex/concave one. Nevertheless, this method is in essence a Lambertian photometric stereo where the sparse specularities are treated as outliers to the

*athinodoros.georghiades@yale.edu

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

Lambertian model. Hence, this method will, in general, not be able to accurately recover the shape of non-Lambertian objects (which may exhibit specular *lobes*), nor will it be able to recover their non-Lambertian reflectance. Furthermore, the sparse set of specular points required by this algorithm were selected by hand from a set of segmented non-Lambertian regions in the images.

In our method, we incorporate the T-S model in uncalibrated photometric stereo because it has both a Lambertian part and, more importantly, a specular lobe of varying sharpness and strength. This model was derived from a physical model of the roughness of surfaces and can capture the reflectance properties of a wide variety of non-Lambertian surfaces. In our algorithm, the GBR ambiguity can be resolved because the specular lobes in the acquired images with respect to the light sources and camera have to be aligned with the lobes derived from the T-S model. This is, in a sense, a generalization of the consistent viewpoint constraint presented in [4]. Note that the specular lobes of most real-life surfaces can exhibit off-specular behavior, i.e., the maximum intensity for most incidence angles will not be aligned with the perfectly (mirror-like) specular direction utilized by the consistent viewpoint constraint. This off-specular behavior can nevertheless be captured by the T-S model and we have successfully applied it to real, non-Lambertian surfaces, including human faces. We have recovered not only their 3-D shape (up to the binary convex/concave ambiguity) but also their reflectance properties. For faces, the estimated skin reflectance has been shown to closely resemble the measured reflectance function reported by Marschner, et al. [15].

In [6], we have demonstrated the increased photorealism of synthetic images of human faces when using the recovered reflectance properties based on the T-S model. In Section 6, we demonstrate improved recognition rates on 4050 images of 10 faces with variable lighting and viewpoint when synthetic image-based representations of the faces are generated using the surface reconstructions and reflectance properties recovered while assuming the T-S model.

2 Non-Lambertian Reflectance Functions

The surface reflectance of objects is usually represented by the four parameter Bidirectional Reflectance Distribution Function (BRDF). Those parameters are the two incoming light source direction angles, θ_i, ϕ_i, and the two viewing direction angles, θ_r, ϕ_r. The BRDF is defined as follows:

$$r(\theta_i, \phi_i; \theta_r, \phi_r) = \frac{dL_r(\theta_r, \phi_r; \theta_i, \phi_i)}{dE_i(\theta_i, \phi_i)}, \tag{1}$$

where $dL_r(\theta_r, \phi_r; \theta_i, \phi_i)$ is the outgoing irradiance from an infinitesimal patch on the surface, and dE_i is the incident radiance from an infinitesimal source. Due to Helmholtz reciprocity [11], $r(\theta_i, \phi_i; \theta_r, \phi_r) = r(\theta_r, \phi_r; \theta_i, \phi_i)$. This means the BRDF is the same if the light source and the camera (sensor) are interchanged.

The above representation assumes that the light is monochromatic, and that it arrives at and bounces away from the same surface point. This precludes translucency and phosphorescence. Under the isotropy assumption, a common simplification and not an unreasonable one, $r(\theta_i, \phi_i; \theta_r, \phi_r) = r(\theta_i, \theta_r, \phi_r - \phi_i) = r(\theta_i, \theta_r, \Delta\phi)$.

As it can be surmised by looking around in a room, the BRDF function of surfaces can be quite complicated. To simplify things and to make them more tractable in different applications in Computer Vision such as photometric stereo, binocular stereo, tracking, and so on, the surface reflectance has usually been assumed to be Lambertian (i.e., perfectly diffuse). Under the Lambertian assumption, the appearance of an object surface is the same as the viewing direction changes, and is only proportional to the cosine of the angle between the local surface normal and the light source direction. Invoking the Helmholtz reciprocity, it can be shown that the BRDF of a Lambertian surface is constant, i.e.,

$$r_{\text{L}} = a_d \tag{2}$$

where a_d is the Lambertian (diffuse) albedo. Although it is never true in reality, the Lambertian BRDF can nonetheless be a satisfactory approximation to the reflectance of many real surfaces.

On the other extreme is the perfectly specular reflectance exhibited by a perfect mirror. In this case, the BRDF of the surface can be shown to be a Dirac delta function with an infinite magnitude only when the incoming light direction is the reflection of the viewing direction about the surface normal and within the incidence plane. Like before, this is also an idealization of real mirror surfaces.

Over the years various non-Lambertian parametric models have been proposed for the reflectance of real-world surfaces. These parametric models can be divided into physically-based and empirically-based models. One issue that has bedeviled the Computer Graphics community is the complexity of these different models. The best known empirical model is the Phong model [18], and it has been popular because of its simplicity and fairly decent rendering results. Nevertheless, the Phong model has no physical basis because it cannot capture important reflectance effects. These include the significant increase in the BRDF values and the off-specular forward scattering as the incidence angle becomes large. It has been noted that many materials, including metals, oxides [23], and human skin [15], exhibit these properties.

The Torrance and Sparrow model [23] is a physically-based model that can capture those two effects. It assumes that reflectance consists of two components. One is associated with bulk material effects which are assumed to lead to a Lambertian lobe colored by the (diffuse) albedo at a particular position on the surface. The other component is assumed to be purely related to surface scatter.

With the T-S model, the surface is modeled as a large collection of perfectly specular micro-facets whose surface normal deviation from the average surface normal is assumed to be a zero-mean Gaussian—the higher the variance of deviation the rougher the surface, and hence the duller its appearance due to higher scatter. This surface scatter leads to a specular lobe in the forward direction. Combining these two terms gives the following BRDF for the T-S model:

$$r_{\text{TS}} = a_d + a_s Q F(\theta', \eta) \frac{\exp(-\nu^2 \theta_a^2)}{\cos\theta_i \cos\theta_r}, \tag{3}$$

where a_d is the Lambertian (diffuse) albedo, a_s is the specular albedo, ν is the surface roughness (the lower its value the higher the roughness), and θ_a is the angle between the surface normal and the bisector of the (incoming) light source and (outgoing) viewing direction. Q is the bistatic shadowing (also known as the geometric attenuation) factor, and $F(\theta', \eta)$ is the Fresnel reflectivity [19], where θ' is the phase angle, the bisecting angle between the incoming and outgoing directions, and η is the index of refraction. Note that the T-S model BRDF is isotropic and satisfies the Helmholtz reciprocity [11].

The image intensity derived from the T-S model when a single point light source $\mathbf{s}$ at infinity illuminates the object is given by

$$I_{\text{TS}} = a_d||\mathbf{s}||\cos\theta_i + a_s||\mathbf{s}||QF(\theta', \eta)\frac{\exp(-\nu^2\theta_a^2)}{\cos\theta_r}, \quad (4)$$

where $||\mathbf{s}||$ is the light source intensity.

The T-S model deviates from the Phong model on two counts. First, there is the exponential function that comes from the Gaussian assumption mentioned above. It should be noted that the cosine term in the Phong model has no physical basis. The second difference is the existence of the $1/\cos\theta_r$ and $1/\cos\theta_i$ terms in the expression for the T-S model in Equation 3. These two cosine terms can affect the BRDF considerably. They lead to a significant increase in the BRDF values along with (increasingly) off-specular maxima as the incidence angle becomes larger [23]. As noted above, many materials including metals, dielectrics (e.g. oxides) [23], and human skin [15] exhibit this behavior making the T-S model a good candidate for modeling their reflectance.

One aspect of reflectance not captured by the T-S model is the backscatter lobe. This can be observed in metals and in painted surfaces. In this case, the model by Oren and Nayar [17], which is based on the same surface roughness model and assumptions as the T-S model, is more suitable. Nevertheless, the Oren and Nayar model does not capture the forward scatter, a more common effect. Notably, human skin as well as many dielectrics (e.g. oxides) do not exhibit any significant back-scattering properties [15, 23] and the T-S model can capture their reflectance properties quite well.

In the following sections, the bistatic shadowing factor, Q, and the Fresnel reflectivity, F, in Equation 4 are ignored due to their very insignificant effect when the phase angle, θ', is quite small (i.e., when the light source and viewing directions are quite close together [22]). This was the case in the training images in the example reconstructions shown in Section 5.

3 Resolving the GBR ambiguity

As noted before, in the case of Lambertian reflectance, the surface of an object seen under variable lighting, but with fixed viewpoint, can only be reconstructed up to the GBR ambiguity. This means we can only recover $\bar{z}(x,y) = g_3 z(x,y) + g_1 x + g_2 y$ (where the g_i are the three GBR parameters), instead of $z(x,y)$, the true surface height function [2]. As far as the surface normals are concerned, we can only recover

$$\bar{\mathbf{n}}(x,y) = \frac{G^{-T}\mathbf{n}(x,y)}{||G^{-T}\mathbf{n}(x,y)||}, \quad (5)$$

where

$$G^{-1} = \frac{1}{g_3}\begin{bmatrix} g_3 & 0 & 0 \\ 0 & g_3 & 0 \\ -g_1 & -g_2 & 1 \end{bmatrix}, \quad (6)$$

and

$$\mathbf{n}(x,y) = \frac{(z_x(x,y), z_y(x,y), -1)^T}{\sqrt{z_x^2(x,y) + z_y^2(x,y) + 1}}. \quad (7)$$

Note that $z_x(x,y)$ and $z_y(x,y)$ are the $x-$ and $y-$partial derivatives of the surface, and $G^{-T} = (G^{-1})^T = (G^T)^{-1}$.

Under the Lambertian model, the irradiance emanating from the surface of a body is invariant under the GBR transformation. If $\mathbf{b}(x,y) = a_d(x,y)\mathbf{n}(x,y)$, where $a_d(x,y)$ is the (Lambertian) albedo, then it is easy to see that the image intensity $I_{\text{L}}(x,y) = \bar{\mathbf{b}}^T\bar{\mathbf{s}}_i = \mathbf{b}^T G^{-1} G\mathbf{s}_j = \mathbf{b}^T\mathbf{s}_i$, formed by illuminating the surface by a single light source, $\mathbf{s}_i$, is invariant under the GBR transformation.

With the T-S model, on the other hand, the image intensity is not invariant under the GBR transformation. This is evident in the following:

$$\begin{aligned} & I_{\text{TS}}(\bar{\mathbf{b}}(x,y), \bar{\mathbf{s}}_i) \\ &= \bar{\mathbf{b}}^T\bar{\mathbf{s}}_i + \frac{a_s||\bar{\mathbf{s}}_i||\exp(-\nu^2[\cos^{-1}(\frac{\bar{\mathbf{b}}^T(\frac{\bar{\mathbf{s}}_i}{||\bar{\mathbf{s}}_i||}+\hat{\mathbf{v}})}{||\bar{\mathbf{b}}||\,||\frac{\bar{\mathbf{s}}_i}{||\bar{\mathbf{s}}_i||}+\hat{\mathbf{v}}||})]^2)}{\frac{\hat{\mathbf{v}}^T\bar{\mathbf{b}}}{||\bar{\mathbf{b}}||}} \\ &= \mathbf{b}^T G^{-1} G\mathbf{s}_i \\ &\quad + \frac{a_s||G\mathbf{s}_i||\exp(-\nu^2[\cos^{-1}(\frac{\mathbf{b}^T G^{-1}(\frac{G\mathbf{s}_i}{||G\mathbf{s}_i||}+\hat{\mathbf{v}})}{||G^{-T}\mathbf{b}||\,||\frac{G\mathbf{s}_i}{||G\mathbf{s}_i||}+\hat{\mathbf{v}}||})]^2)}{\frac{\hat{\mathbf{v}}^T(G^{-T}\mathbf{b})}{||G^{-T}\mathbf{b}||}}, \end{aligned} \quad (8)$$

where $\hat{\mathbf{v}}$ is the unit-length viewing direction. In general, $I_{\text{TS}}(\bar{\mathbf{b}}(x,y), \bar{\mathbf{s}}_i) \neq I_{\text{TS}}(\mathbf{b}(x,y), \mathbf{s}_i)$ as one traverses an image. Equality (across the whole image) holds only when G is equal to the identity matrix.

The above observation leads us to the natural question of what is the minimum number of $(\mathbf{b}, \mathbf{s}_i)$ pairs required to pin down the solution of the GBR ambiguity. It should be noted that G has three free parameters (or degrees of freedom), namely g_1, g_2, and g_3. The right-hand side of Equation 8 is therefore a non-linear function of those three parameters. With only one $(\mathbf{b}, \mathbf{s}_i)$ pair, the solutions of $I_{\text{TS}}(\bar{\mathbf{b}}(x,y), \bar{\mathbf{s}}_i) = I_{\text{TS}}(\mathbf{b}(x,y), \mathbf{s}_i)$ lie on a 2-D surface in the 3-D space of g_1, g_2, and g_3 (including the point that corresponds to G being equal to the identity matrix). This surface can, in general, be shown to be regular, but not necessarily connected. Note that there are in general two parts, one for $g_3 > 0$ and one for $g_3 \leq 0$. For the purposes of this discussion we will restrict ourselves to the case when $g_3 > 0$.

Each equation $I_{\text{TS}}(\bar{\mathbf{b}}(x,y), \bar{\mathbf{s}}_i) = I_{\text{TS}}(\mathbf{b}(x,y), \mathbf{s}_i)$ associated with a $(\mathbf{b}, \mathbf{s}_i)$ pair provides a single constraint on the

values of g_1, g_2, and g_3. The task is to show that one cannot have more than one solution, other than $g_1 = 0$, $g_2 = 0$, and $g_3 = 1$ (i.e., G is equal to the identity matrix), when there are four or more $(\mathbf{b}, \mathbf{s}_i)$ pairs.

As noted above, with a single pair there is an infinite number of solutions that lie on a 2-D surface in the 3-D space of g_1, g_2, and g_3. We follow here the worst case scenario—even with a single pair the solutions can sometimes lie on a 1-D surface. One usually does not encounter this case except when $I(\mathbf{b}(x,y), \mathbf{s}_i)$ is a critical value (such as a maximum), and hence the 1-D surface is a critical surface. In this discussion, the solution surface will be assumed to be 2-D.

With two light sources (i.e., two $(\mathbf{b}, \mathbf{s}_i)$ pairs), the solutions lie on an 1-D subspace. This is because two 2-D regular surfaces in 3-D in general intersect at an 1-D subspace. Now, this 1-D space, under certain conditions, can intersect a 2-D solution surface defined by a third $(\mathbf{b}, \mathbf{s}_i)$ pair only at a finite (although unknown) number of points. Those conditions stipulate that the $\mathbf{s}_i$, $\forall i$, are distinct in direction, and that the angles between $\mathbf{b}$ and the $\mathbf{s}_i$ are always less than 90^o, i.e., the light sources cast some light on the infinitesimal surface patch associated with $\mathbf{b}$. Furthermore, we assume that this surface patch is strictly visible, i.e., the angle between the viewing direction and $\mathbf{b}$ is less than 90 degrees. Under these conditions, with three $(\mathbf{b}, \mathbf{s}_i)$ pairs, the solutions are in general point solutions, i.e., they are non-degenerative.

This can be seen by looking at the Hessian of the objective function, $O(\mathbf{g}) = \sum_i f_i(\mathbf{g})^2 = \sum_i |I_{\text{TS}}(\bar{\mathbf{b}}(x,y), \bar{\mathbf{s}}_i; \mathbf{g}) - I_{\text{TS}}(\mathbf{b}(x,y), \mathbf{s}_i)|^2$, where $\mathbf{g} = [g_1, g_2, g_3]^T$. Under the conditions mentioned above and in the case of three $(\mathbf{b}, \mathbf{s}_i)$ pairs, the Hessian is full-rank and positive-definite at a solution. This can be demonstrated by taking the second derivative of $O(\mathbf{g})$ w.r.t. $\mathbf{g}$ at a point in the 3-D space where there is a solution, i.e., where $I_{\text{TS}}(\bar{\mathbf{b}}(x,y), \bar{\mathbf{s}}_i) = I_{\text{TS}}(\mathbf{b}(x,y), \mathbf{s}_i)$ for all three pairs simultaneously:

$$\begin{aligned} H(\mathbf{g}) &= \nabla^2 O(\mathbf{g}) \\ &= 2\sum_i (\nabla^2 f_i(\mathbf{g})) f_i(\mathbf{g}) + \nabla f_i(\mathbf{g}) (\nabla f_i(\mathbf{g}))^T \\ &= 2\sum_i \nabla f_i(\mathbf{g}) (\nabla f_i(\mathbf{g}))^T. \end{aligned} \quad (9)$$

If the $\nabla f_i(\mathbf{g})$ are distinct, which in general they are under the conditions given above, then the 3×3 Hessian, $H(\mathbf{g})$, is full rank and positive-definite. This implies that the solution is non-degenerative.

When we have 4 or more distinct, randomly chosen pairs, $(\mathbf{b}, \mathbf{s}_i)$, then there is a zero probability that there will be multiple solutions, i.e., there is a zero probability that a 2-D solution surface defined by the fourth, randomly chosen $(\mathbf{b}, \mathbf{s}_i)$ pair will intersect with any of the finite number of point solutions defined by the first three pairs (except at $g_1 = 0, g_2 = 0$, and $g_3 = 1$).

Alternatively, consider the constraint that the point solutions, in the g_1, g_2, and g_3 space (with three $(\mathbf{b}, \mathbf{s}_i)$ pairs), place on the fourth randomly chosen light source so that the fourth $(\mathbf{b}, \mathbf{s}_i)$ pair is also a solution. A point solution in the g_1, g_2, and g_3 space provides a *single* constraint on the fourth light source direction, which has two degrees of freedom, namely azimuth and elevation. Hence, in the 2-D space of possible azimuths and elevations for the fourth light source direction, there is only an 1-D space where this constraint is satisfied. Nevertheless, when we randomly choose the fourth light source direction out of the possible 2-D space, there is a zero probability that the light source will come from that 1-D solution space. Hence, four $(\mathbf{b}, \mathbf{s}_i)$ pairs are sufficient to resolve the GBR ambiguity.

Granted, the existence of a unique global minimum does not guarantee that $O(\mathbf{g})$ has no local minima in the g_1, g_2, and g_3 space even with 4 or more $(\mathbf{b}, \mathbf{s}_i)$ pairs. Furthermore, showing uniqueness is more challenging when one also optimizes over light sources, surface normals, albedos and the parameters of the reflectance model. In the next section, we outline our surface reconstruction and reflectance recovery algorithm and describe how to avoid local minima in almost all practical cases.

4 Surface Reconstruction and Recovery of Reflectance Function

The algorithm described here is an extension of the uncalibrated photometric stereo technique presented in [7] in which no knowledge about the light sources was assumed and integrability was enforced on the surface normals. We have incorporated in that algorithm the T-S model of reflectance, requiring to estimate, along with the other parameters, a_s and ν of the T-S model shown in Equation 3. All the parameters—surface normals, Lambertian albedo, light sources and the parameters of the T-S model—are iteratively updated using least squares techniques.

While $a_d(x,y)$, the diffuse albedo in Equation 3, varies spatially, the parameters of the surface scatter term in the T-S model are assumed not to be a function of (x,y), which implies that a_s and ν are constant over (x,y). (Note that θ_a in Equation 3 is still a function of (x,y) because it is a function of the surface normal field, $\mathbf{n}(x,y)$.) The spatial invariance of a_s and ν is, of course, not entirely true in real surfaces, e.g. the eyes have a quite different BRDF than the skin part of a face. It is, nonetheless, not an unreasonable assumption and makes the problem more tractable. Allowing a_s and ν to vary across the surface is an interesting problem, but it can be very difficult to solve reliably.

In this exposition, we let the surface of an object $z(x,y)$, whose reflectance is given by the T-S model shown in Equation 3, be viewed by an orthographic camera, and let the camera acquire an n pixel image of this surface. Let the image be denoted by the vector $\mathbf{I} \in \mathbb{R}^n$. Note that the image is written as a vector with each pixel corresponding to an element I_j of the vector $\mathbf{I}$. A pixel I_j samples the image irradiance at some point in the image plane as specified by its Cartesian coordinates (x,y).

We acquire k images of the object by fixing the object's position relative to the camera and moving the light source to some unknown position, $\mathbf{s}_i$, before each acquisition. Let the acquired set of images be denoted by matrix $X = [\mathbf{I}_1, \ldots, \mathbf{I}_k]$ with I_{ij} being the value of the j-th pixel in the i-th image. Note that the images can have shadows

(both cast and attached), and possibly saturations—these do not satisfy the T-S model of reflectance. We therefore determine which pixels are observing either a saturation or a shadow and mark them as being invalid. Any invalid data (both shadows and saturations) are treated as missing measurements by our estimation method.

To determine the 3-D shape and reflectance properties of the object, we solve the following minimization problem:

$$\min_{\mathbf{n}_j,\, a_d(x,y),\, a_s,\, \nu,\, \mathbf{s}_i} O(\mathbf{n}_j,\, a_d(x,y),\, a_s,\, \nu,\, \mathbf{s}_i; \mathbf{g}) = \sum_{ij} m_{ij} |I_{ij} - I_{\text{TS}}(\mathbf{n}_j, a_d(x,y), a_s, \nu, \mathbf{s}_i; \mathbf{g})|^2, \quad (10)$$

where

$$m_{ij} = \begin{cases} 1 & I_{ij} \text{ valid pixel measurement,} \\ 0 & \text{otherwise.} \end{cases}$$

I_{TS} is given in Equations 4 and 8, and the Cartesian position (x, y) corresponds to the j-th pixel. We solve this minimization using steepest descent in an iterative scheme. The algorithm is initialized as in [7] and then let run for 20-30 iterations with only the Lambertian model. After this, the full T-S model is introduced. This almost always prevents the solution from falling into a local minimum during the initial iterations and when the surface normals, albedos, and light sources are very far from their respective solutions. More details about the algorithm are presented in [9] and [6].

5 Reconstruction Results

In this section, we show reconstruction results for part of a lacquered globe and for a face. The results demonstrate that incorporating the T-S model in uncalibrated photometric stereo leads to better surface reconstructions for objects that do not obey the Lambertian reflectance assumption, resolves the GBR ambiguity inherent in uncalibrated photometric stereo, and recovers the parameters of the extended reflectance model. In the case of faces, the recovered skin reflectance function closely resembles the previously measured one [15].

Figure 1 shows the reconstruction of part of the globe. Twelve training images (of size 400×400 pixels) were used which were selected so that the light sources were not centered around the optical axis. This helps remove any accidental symmetry that could influence the resolution of the GBR ambiguity. Note also that no attempt was made to threshold out the specular regions in the training images (Figure 1.a); only saturated pixels were removed.

The results associated with the T-S model are clearly better than those associated with the Lambertian model. For example, note that the (diffuse) albedo of the T-S model reconstruction exhibits much less GBR-induced spatial variation than the albedo of the Lambertian reconstruction [1]. This implies that for the T-S model reconstruction, the GBR ambiguity has, to a large extent, been resolved. Moreover, the $x-$ and $y-$derivatives of the surface for the Lambertian case, in the top row of Figure 1.b, show distortion in the central part due to the presence of specularities in the training images. This is not visible in the $x-$ and $y-$derivatives of the T-S reconstruction in the bottom row because in this case the specularities have been discounted. Last, but more importantly, notice how the Lambertian reconstruction, in

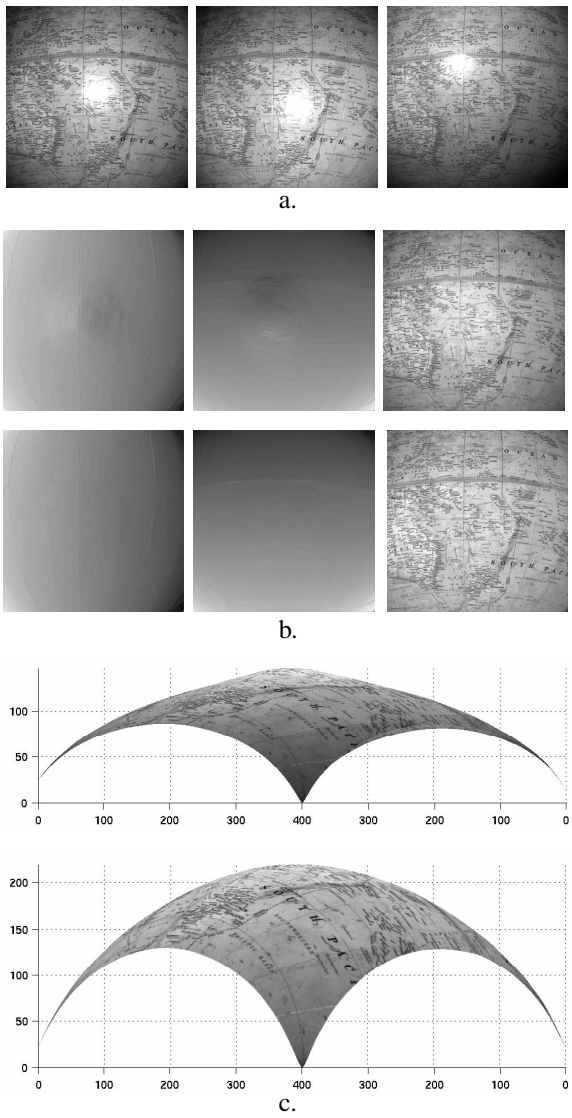


Figure 1: The surface reconstruction of part of a globe: a. Three out of the 12 images used in the reconstructions, where the light source directions can vary up to 24^o from the optical axis. b. The top row shows the $x-$ and $y-$derivatives of the surface and the albedo, $a_d(x, y)$, for the Lambertian model reconstruction. The bottom row shows those corresponding to the T-S model reconstruction. c. The renderings of the Lambertian reconstruction (top) and the T-S reconstruction (bottom) viewed along the $xy-$diagonal. The corresponding albedos $a_d(x, y)$ have been texture-mapped on the surfaces, while the units of the vertical axes are in pixels. The parameters of the T-S model recovered during the reconstruction were: $a_d = 0.0541$ (the average albedo over the whole image), $a_s = 0.0415$, and $\nu = 8.1255$.

[1] Note that the albedo varies as a function of the GBR parameters, as shown in [2].

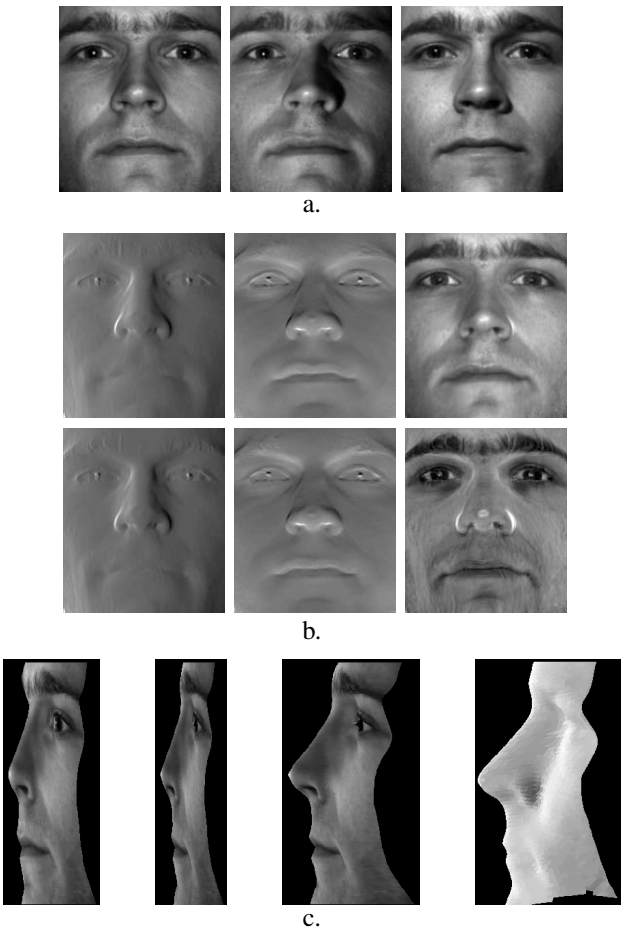


Figure 2: The surface reconstruction of a face: a. Three out of the 12 images used in the reconstructions where the light source direction can vary up to 24^o from the optical axis. b. The top row shows the $x-$ and $y-$derivatives of the surface and the albedo, $a_d(x, y)$, for the Lambertian model reconstruction. The bottom row shows those corresponding to the T-S model reconstruction. c. From left: side-view of the Lambertian model surface reconstruction; the same surface viewed from behind the face so as to demonstrate its flatness; side-view of the T-S model reconstruction; side-view of the structured-light multi-view reconstruction of the face (courtesy of Eyetronics). In the Lambertian case, notice the flattened surface as well as the presence of an additive plane. In the T-S model case, on the other hand, the GBR ambiguity has to a large extent been resolved. (The differences noted around the nose tip region between the T-S model reconstruction and the structured-light reconstruction are mainly due to the presence of occlusions on the left and right and on the edges of the nostrils.)

the top row of Figure 1.c, is not only flattened but also reveals some distortion in the center. The T-S model reconstruction, on the other hand, is very close to a sphere, the expected shape.

Figure 2 shows the reconstruction of a human face. As was the case with the globe reconstructions in Figure 1, the 12 training images (three of which are shown in Figure 2.a) were selected so that they were not centered around the optical axis. The results associated with the T-S model are better. Notice in Figure 2.c how in the T-S case the GBR ambiguity has been resolved to a large extent, while in the Lambertian case the surface appears to be severely flattened and have a residual additive plane.

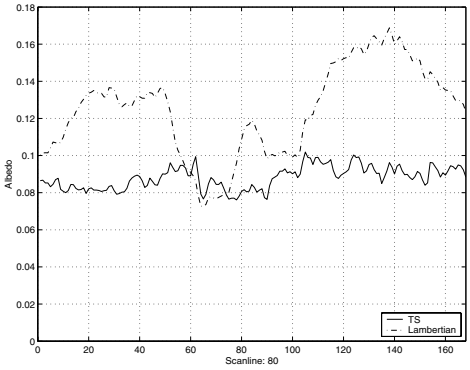


Figure 3: The albedo profiles, $a_d(x, 80)$, of the Lambertian and T-S model reconstructions of the face shown in Figure 2 along a horizontal scanline half-way down the nose ridge where the albedo is expected to be approximately constant. In the Lambertian case, the albedo varies significantly, implying significant flattening, and is higher on the right-hand side, implying an additive plane. The small albedo variation of the T-S reconstruction, on the other hand, implies that the GBR ambiguity has been resolved in that case.

Comparing the albedo $a_d(x, y)$ for the two cases can also reveal the same. Figure 3 shows the albedo profiles for the Lambertian and T-S model reconstructions of the face shown in Figure 2, along a horizontal scanline, traversing the images approximately half-way down the nose ridge and over a part of the face where the albedo is expected to be approximately constant. The albedo in the Lambertian model reconstruction varies significantly, implying significant flattening, and is higher on the right-hand side, implying an additive plane. The small variation of the T-S albedo, on the other hand, signifies that the GBR ambiguity has been resolved in that case.

Significantly, incorporating the T-S model in uncalibrated photometric stereo can also recover the non-Lambertian nature of skin reflectance. The left of Figure 4 shows the skin BRDF recovered using our algorithm. Notice its similarity (up to a global scale factor) with the previously *measured* BRDF of human skin shown on the right. (The plot on the right is courtesy of Stephen R. Marschner [15].)

The skin BRDF is close to Lambertian at small incidence angles, but exhibits strongly increasingly off-specular behavior as the incidence angle becomes larger. Notice also how the scale increases by almost 40 times from the top plot to the bottom. The almost Lambertian behavior when the incidence angle is small is what justified the use of the Lambertian model in reconstructing human faces when the light sources are close to the camera. Of course, in that case, the GBR ambiguity remained unresolved in the surface reconstructions. In our algorithm, the deviations from the Lambertian assumption provide enough information to re-

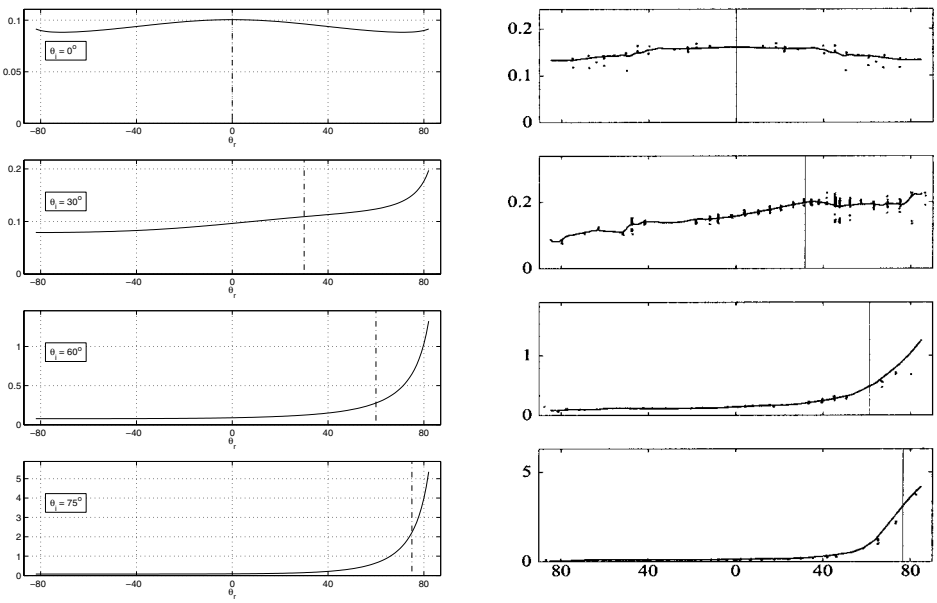


Figure 4: On the left, the estimated BRDF of the face in Figure 2 at different incidence angles, θ_i, estimated using our method. On the right, the *measured* BRDF of human skin. The plots show the BRDF values in the incidence plane as θ_r varies. The parameters of the T-S model recovered during the reconstruction were: $a_d = 0.0776$ (the average albedo over the whole face), $a_s = 0.0229$, and $\nu = 2.2483$. Note that the BRDF is almost Lambertian at small incidence angles (i.e., it is almost constant), but exhibits increasingly off-specular behavior as the incidence angle becomes larger. Observe that the scale increases by almost 40 times from the top plot to the bottom. Furthermore, the specular lobe seems significantly removed from the perfectly specular direction exhibiting a maximum at almost glancing θ_r for most of the range of θ_i. Notice the similarities (up to a global scale factor) of the estimated BRDF with the measured BRDF shown on the right. (The plot on the right is courtesy of Stephen R. Marschner and was presented in [15].)

cover the parameters of the T-S model of reflectance. With these, the skin reflectance can be accurately predicted for a wide range of incidence and viewing angles significantly extrapolating from those in the training images.

Note that although the Fresnel reflectivity was ignored during the reconstruction process due to its insignificant effect at low phase angles (less than 30^o), it was re-introduced in the BRDF plots shown on the left of Figure 4, because the phase angle can be much larger than 30^o. The index of refraction, η, was assumed to be 1.5—the index of refraction of human skin reported in the literature is usually 1.37-1.5 [24].

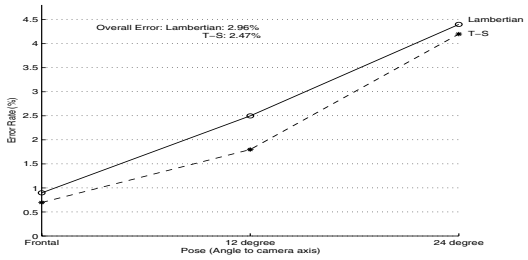
6 Face Recognition

A significant motivation for this research was the use of the reconstructed surface shapes and (non-Lambertian) reflectance properties of faces to generate synthetic image-based representations to be used in face recognition under variable lighting and viewpoint. Incorporating the T-S model of reflectance in uncalibrated photometric stereo obviates the need to correct for the GBR parameters before synthesizing images under variable lighting and viewpoint. Furthermore, the use of the recovered (non-Lambertian) reflectance of faces leads to more photorealistic synthetic images as demonstrated in [6].

We have therefore used this new algorithm to reconstruct the shape and recover the (non-Lambertian) reflectance properties of the 10 faces in the Yale Face Database B[2]. Seven training images with close to frontal, unknown, single light source illumination and with fixed viewpoint were used for each face. The reconstructed shape and surface reflectance were then used to generate synthetic image-based representations for the 10 faces, as described in [7], modeling the appearance of each face under variable lighting and viewpoint. Each representation comprises a collection of (linear subspace approximations) of illumination cones in the space of images—one cone for each sampled viewpoint. (It was shown in [1] that the set of all images of an object (such as a face) in fixed pose, but under all possible illumination conditions, is a convex cone in the image space.)

We have performed face recognition experiments with the face representations generated using the T-S reconstructions comparing their performance with that of face representations created under the Lambertian reflectance assumption. We have performed the tests on 4050 images from the Yale Face Database B. These images contain 405 viewing conditions (9 poses $\times$ 45 illumination conditions) for each of the 10 individuals in the database. The images from each pose are divided into four illumination subsets according to the angle the light source direction makes with the camera's axis. Subset 1 (respectively 2, 3, 4) con-

[2] http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html



Reflectance Model	Error Rate (%) vs. Pose		
	Frontal (Pose 1)	12° (Poses 2 3 4 5 6)	24° (Poses 7 8 9)
Lambertian	0.9	2.5	4.4
T-S	0.7	1.8	4.2

Figure 5: Face recognition results under variable lighting and viewpoint demonstrating the decrease in error rates when synthetic image-based representations of faces are generated using the T-S surface and reflectance reconstructions than the reconstructions recovered under the Lambertian assumption.

tains 70 (respectively 120, 120, 140) images per pose. The database is also divided into three pose subsets: the "Frontal Pose"(Pose 1) which includes 450 test images, the "12 degree" subset (Poses 2, 3, 4, 5, 6) which has 2250 test images, and the "24 degree" (Poses 7, 8, 9) with 1350 test images. (See [7] for more information about the database.)

Figure 5 shows face recognition results using face representations created under the T-S and Lambertian reflectance models. Note that for all three pose subsets the error rate is lower with the T-S representations. The overall error has gone from 2.96% (120 errors in 4050 test images) with the Lambertian representations down to 2.47% (100 errors in 4050 test images) with the T-S representations, a decrease of 16.7%. This can be attributed to the more accurate modeling of appearance under variable lighting and viewpoint when incorporating the T-S model of reflectance in uncalibrated photometric stereo to recover the face surface shape and reflectance used to generate synthetic images.

Acknowledgements

Many thanks to Melissa Koudelka for providing the structured-light reconstruction of the face in Figure 2. Useful discussions with Jonas August, Peter Belhumeur, Patrick Huggins, and Todd Zickler are also appreciated.

References

[1] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible illumination conditions. *Int. J. Computer Vision*, 28(3):245–260, July 1998.

[2] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 1040–1046, 1997.

[3] E. Coleman, Jr. and R. Jain. Obtaining 3-Dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, April 1982.

[4] O. Drbohlav and R. Sara. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. European Conf. on Computer Vision*, page II: 46 ff., 2002.

[5] J. Fan and L. Wolff. Surface curvature and shape reconstruction from unknown multiple illumination and integrability. *Computer Vision and Image Understanding*, 65(2):347–359, February 1997.

[6] A. Georghiades. Recovering 3-D shape and reflectance from a small number of photographs. In *Eurographics Symposium on Rendering*, pages 230–240 and 315, June 2003.

[7] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(6):643–660, June 2001.

[8] A. Georghiades, D. Kriegman, and P. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 52–59, 1998.

[9] A. S. Georghiades. *From Few to Many: Generative Appearance-Based Models for Face Recognition*. PhD thesis, Yale University, 2003.

[10] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *J. Opt. Soc. Am. A*, 11(11):3079–3089, Nov. 1994.

[11] H. v. Helmholtz. *Treatise on Physiological Optics*. Dover, New York, 1925.

[12] B. Horn. *Computer Vision*. MIT Press, Cambridge, Mass., 1986.

[13] J. Lambert. *Photometria Sive de Mensura et Gradibus Luminus, Colorum et Umbrae*. Eberhard Klett, 1760.

[14] S. Lin and S. Lee. Estimation of diffuse and specular appearance. In *Int. Conf. on Comp. Vis.*, pages 855–860, 1999.

[15] S. Marschner, S. Westin, E. Lafortune, K. Torrance, and D. Greenberg. Image-based BRDF measurement including human skin. In *10th Eurographics Workshop on Rendering*, pages 139–152, Grenada, Spain, June 1999.

[16] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Journal of Robotics and Automation*, 6(4):418–431, August 1990.

[17] M. Oren and S. Nayar. Generalization of the Lambertian model and implications for machine vision. *Int. J. Computer Vision*, 14:227–251, 1996.

[18] B. Phong. Illumination for computer generated images. *Communications of the ACM*, 18(6):311–317, June 1975.

[19] R. Siegel and J. Howell. *Thermal Radiation Heat Transfer*. McGraw-Hill, New York, 1972.

[20] W. Silver. *Determining Shape and Reflectance Using Multiple Images*. PhD thesis, MIT, Cambridge, MA, 1980.

[21] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 18(4):449–454, April 1996.

[22] H. Tagare and R. de Figueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 13(2):133–152, February 1991.

[23] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.*, 57:1105–1114, 1967.

[24] M. van Gemert, S. Jacques, H. Sterenborg, and W. Star. Skin optics. *IEEE Trans. on Biomedical Engineering*, 36(12):1146–1154, December 1989.

[25] R. Woodham. Analysing images of curved surfaces. *Artificial Intelligence*, 17:117–140, 1981.

[26] A. Yuille, J. Coughlan, and S. Konishi. The KGBR Viewpoint-Lighting Ambiguity. *J. Opt. Soc. Am. A*, 20(1):24–31, January 2003.

[27] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 158–164, 1997.