# Shape from Shading with a Generalized Reflectance Map Model* 

Kyoung Mu Lee<br>Department of Radio Engineering, Hong-Ik University, 72-1 Sangsu-Dong, Mapo-Gu, Seoul, Korea 121-791<br>and<br>C.-C. Jay Kuo $\dagger$<br>Signal and Image Processing Institute and the Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564<br>Received June 22, 1994; accepted March 14, 1996

Most conventional SFS (shape from shading) algorithms have been developed under three basic assumptions on surface properties and imaging geometry to simplify the problem, namely, a Lambertian surface, a distant point light source and orthographic projection. In this research, we derive a physics-based generalized reflectance map model which includes diffuse and specular reflection effects, a nearby point light source and perspective projection, and then we develop a new direct shape recovery algorithm from shaded images. The basic idea of our solution method is to discretize the image irradiance equation with a finite triangular element surface model, to express the resulting nonlinear system of equations in terms of depth variables only and to recover the object shape by linearizing the nonlinear equations and minimizing a quadratic cost functional. We perform numerical experiments with one or multiple photometric stereo images to demonstrate the performance of the derived physics-based reflectance map model and the proposed SFS algorithm. © 1997 Academic Press

## 1. INTRODUCTION

There has been a considerable amount of interest and effort on shape extraction from image intensities in computer vision research for the last several decades. The brightness of a pixel in an image is generated through an image formation process governed by the optical, physical, and geometrical factors including the object shape, the surface reflectance property, and the illumination and sensor characteristics. Thus, to get an accurate reconstructed surface, it is crucial to understand and model the whole complicated imaging process. Most conventional SFS algo-

[^0]rithms have been developed under three simple but restrictive assumptions on the image formation model to simplify the problem, namely, the ideal Lambertian surface, a distant point light source and orthographic projection [2, 8, $11,14,15,17,18,24,30,32,39,41]$, etc. We refer to [40] for a detailed performance analysis and comparison of several important algorithms. By means of orthographic projection and distant light source, the image intensity can be simply characterized by a reflectance map which gives scene radiance as a function of only two variables $(p, q)$, the gradient of the corresponding surface point, in a viewercentered coordinated system [7]. With the Lambertian assumption, the reflectance map can be further simplified to be independent on viewer's direction. However, these assumptions are not valid in practical environments for the following reasons. First, surface reflection in general contains both the diffuse and the specular components. Second, the light source is often located in a finite distance from the object so that the image brightness depends on the distance between every object point and the light source. Third, images are formed through a pin-hole camera which should be modeled by perspective projection. Since conventional SFS algorithms do not use accurate physical and optical models, apparent distortions of reconstructed surfaces often occur in many real applications.

To make SFS algorithms practically useful, we have to consider a more realistic model of surface reflection and the imaging process. Several attempts to relax the above restrictive assumptions have been made so far $[1,5,6,10$, $13,19,20,28,33]$. We summarize important developments below. First, the use of more sophisticated non-Lambertian reflectance maps in the SFS problem has been studied by quite a few researchers. Ikeuchi [10] used a double-delta specular model to determine orientations of a specular surface with photometric stereo images generated by distributed light sources. Coleman and Jain [1] used four light
source photometric stereo images to extract the shape of textured and specular surfaces. Healy and Binford [5] and Tsai and Shah [36] employed the Torrance-Sparraw model in their work to determine the shape of a specular surface with a single image. Recently, Tagare and deFigueiredo [33] presented and analyzed a class of $m$-lobed non-Lambertian reflectance maps and developed a photometric stereo SFS method to determine the local surface normal and the reflectance map parameters. Second, the nearby point light source for the SFS problem was examined by Kim and Burger [13]. They derived a reflectance map with a nearby point light source under the orthographic projection and Lambertian surface assumptions, and determined local positions and orientations with photometric stereo images. Finally, the perspective projection model has been applied to the SFS problem by Penna [28,29] and the authors [18]. Penna presented a local SFS analysis of a single perspective image of a Lambertian polyhedronshaped surface and proposed an algorithm which recovers the local shape of the polyhedron by solving a nonlinear system of algebraic equations. The potential extension of his method to a non-Lambertian object was also mentioned. However, this algorithm is practically unreliable due to its sensitivity to noise and numerical finite precision. Lee and Kuo [18] developed a more robust SFS algorithm based on a triangular element surface model and a linear approximation of a Lambertian reflectance map. It recovers the shape of a Lambertian surface directly with single or multiple photometric stereo images taken under perspective projection. An effort trying to solve the SFS problem in a more general setting can be found in the original work of Horn [6], in which he used a perspective projection model, a nearby light source, and an arbitrary reflectivity function. A first-order partial differential equation known as the image brightness equation was derived in terms of five dependent variables $(x, y, z, p, q)$, and then a set of equivalent five ordinary partial equations were derived and solved by the characteristic strip method. However, this method suffers from several practical problems including noise sensitivity and error accumulation in numerical integration of the differential equations.

In this research, we derive a physics-based generalized reflectance map model which includes perspective projection, a nearby light source, and diffuse and specular reflection effects, and then we develop a new unified algorithm for recovering depth variables from shaded images with such a generalized reflectance map. We employ the Torrance and Sparrow BRDF (bidirectional reflectance distribution function) to model diffuse and specular reflections of the reflectance map based on geometrical optics. Furthermore, we incorporate the nearby point light source and perspective projection in the model. The complete model is quite complicated where the brightness at a surface point in an image depends on many factors including
the surface property parameters, the surface normal at that point, the light source direction, specular direction, the viewer direction, and the distance between the surface point and the light source. The SFS problem would be very difficult if we attempted to determine the surface orientation as well as the position of a surface point by solving the nonlinear image irradiance equation with this generalized reflectance map. However, the solution procedure can be greatly simplified by introducing a triangular element surface model for the discretization and parameterization of the image irradiance equation. Through such a procedure, we can express the resulting discrete nonlinear system of equations in terms of depth variables only, and we recover the object shape by linearizing the equations and minimizing a quadratic cost functional. Since our method recovers depth variables directly, no integrability constraint is required.

The paper is organized as follows. In Section 2, we derive a generalized reflectance map model which takes the following factors into account: the specular and diffuse components of light reflection, a nearby point light source, and perspective projection. We consider the discretization and parameterization of the nonlinear reflectance map based on a triangular element surface model in Section 3. We are able to simplify the discrete reflectance map for several special cases, and these simplifications are examined also in this section. The formulation and solution of the SFS problem is studied in Section 4. We perform numerical experiments with one or multiple photometric stereo images to demonstrate the performance of the derived phys-ics-based reflectance map model and the proposed SFS algorithm in Section 5. Some concluding remarks are given in Section 6.

## 2. GENERALIZED REFLECTANCE MAP MODEL

The reflectance map was first defined by Horn [7] as a function that relates reflected radiance to local surface orientation in a viewer-centered coordinate. It is independent of all other variables such as position in the image. This notion of reflectance map is not general enough for describing the situation where there is perspective projection and a finite distant light source. In this section, by incorporating more accurate surface reflectance phenomena and image formation effects, we propose a generalized reflectance map which depends on not only gradient variables but also variables of the position and the distance of light source. The radiance of the reflected light from a surface patch is derived in terms of BRDF and the incident irradiance due to a nearby point light source in Section 2.1. By combining the imaging geometry with perspective projection and the result in Sections 2.1, we obtain a phys-ics-based generalized reflectance map model and the corresponding image irradiance equation in Section 2.2.


FIG. 1. Reflection geometry model.

### 2.1. Reflection Due to a Nearby Point Light Source

We illustrate the reflection geometry used in this paper in Fig. 1, where $r$ is the distance between a surface point $\mathbf{P}$ and a point light source $\mathbf{S}, \mathbf{n}$ is the unit surface normal at $\mathbf{P}, \mathbf{i}$ is the unit vector toward the light source, $\mathbf{v}$ is the unit vector toward the camera, and $\mathbf{h}$ is the unit vector along the specular direction. By definition [34], the bisector of $\mathbf{i}$ and $\mathbf{v}$ specifies the specular direction so that we have

$$
\mathbf{h}=\frac{\mathbf{i}+\mathbf{v}}{\|\mathbf{i}+\mathbf{v}\|}
$$

The $\vartheta_{i}, \vartheta_{v}$, and $\alpha$ denote the angles between $\mathbf{n}$ and $\mathbf{i}, \mathbf{n}$ and $\mathbf{v}$, and $\mathbf{n}$ and $\mathbf{h}$, respectively. The zenith and azimuth angles $(\theta, \phi)$ with a proper subscript represent a unit vector in a given polar coordinate system.

The light reflection depends on the radiance of the light source and the bidirectional reflectance distribution function (BRDF). The BRDF proposed by Nicodemous et al. [23] is a useful tool for characterizing light reflection from solid surfaces and it provides the information of the brightness of a surface patch with given viewing and illumination directions. By definition, it is written as

$$
\begin{equation*}
f=L / E \tag{2.1}
\end{equation*}
$$

where $L$ and $E$ denote the reflected radiance in the emitting direction and the irradiance in the direction of incident light, respectively.

The model general light reflection from surfaces, the BRDF $f$ consists of both the diffuse component $f_{d}$ and the specular component $f_{s}$ [21, 33]; i.e.,

$$
\begin{equation*}
f=\rho_{d} f_{d}+\rho_{s} f_{s} \tag{2.2}
\end{equation*}
$$

where $\rho_{d}$ and $\rho_{s}$ are the weighting factors of diffuse and
specular components. The Lambertian model and the Torrance-Sparrow model have been widely used for the diffuse and specular reflection in the community of computer vision and computer graphics.

Several comprehensive models for diffuse reflection have been recently examined. Oren and Nayar [25] proposed a generalized Lambertion model for rough diffuse surfaces by taking into account complex geometrical effects. The Oren-Naylar model has been used not only for vision but also graphics [26] and visual psychophysics [22, 27]. The accuracy and generality of the model was well illustrated by the experimental results in their work. Wolff [37, 38] also presented an accurate diffuse reflectance model for the smooth dielectric surface. Even though being highly nonlinear involving functions of the viewing direction, interreflection, and Fresnel coefficient, the above two models are able to capture the reflectance of real-world surfaces.

Furthermore, Wolff [37] showed that the proposed model can be well approximated by the Lambertian model when both the angles of incident and emittance are simultaneously less than $50^{\circ}$ and the angle between the viewer direction and the illumination direction is less than $60^{\circ}$. Under these assumptions, we employ the Lambertian model for the diffuse component, and the TorranceSparrow model for the specular components, respectively, in this paper. That is,

$$
f_{d}(\mathbf{i}, \mathbf{n}, \mathbf{v})=1 / \pi
$$

and

$$
\begin{aligned}
f_{s}(\mathbf{i}, \mathbf{n}, \mathbf{v}) & =\frac{1}{\cos \vartheta_{i} \cos \vartheta_{v}} \exp \left\{-k \alpha^{2}\right\} \\
& =\frac{1}{\left(\mathbf{i}^{\mathrm{T}} \mathbf{n}\right)\left(\mathbf{v}^{\mathrm{T}} \mathbf{n}\right)} \exp \left\{-k\left[\cos ^{-1}\left(\mathbf{h}^{\mathrm{T}} \mathbf{n}\right)\right]^{2}\right\}
\end{aligned}
$$

where $k$ is the surface roughness parameter.
As indicated in (2.1), the reflected radiance $L$ depends on the incident irradiance $E$ as well as the BRDF $f$. Since the amount of light energy falling on a surface patch is proportional to the foreshortened area, the irradiance $E$ at that patch is proportional to a cosine function of the angle between the surface normal direction and the light source direction. Moreover, since the distance between the surface patch and the light source is finite, the irradiance is inversely proportional to square of the distance. Assume that the point light source has isotropic radiance $I_{0}$. Then, it is easy to derive the irradiance $E$ at a surface point from a point light source in the direction $\mathbf{i}$ or $\left(\left(\theta_{i}, \phi_{i}\right)\right.$.

$$
\begin{equation*}
E(\mathbf{x}, \mathbf{n}, r)=\frac{I_{0}}{r^{2}} \max \left[0, \mathbf{i}^{\mathrm{T}} \mathbf{n}\right] \delta_{\omega}(\mathbf{x}-\mathbf{i}) \tag{2.3}
\end{equation*}
$$

where $\delta_{\omega}$ is a solid angle delta function [33] defined by

$$
\delta_{\omega}(\mathbf{x}-\mathbf{i})=\delta_{\omega}\left(\theta_{x}-\theta_{i}, \phi_{x}-\phi_{i}\right)=\frac{\delta\left(\theta_{x}-\theta_{i}\right) \delta\left(\phi_{x}-\phi_{i}\right)}{\sin \theta_{i}}
$$

where $\delta$ is the Dirac delta function.
Given a BRDF $f$ of a surface and a point light source, the radiance of reflected light along the viewer direction $\mathbf{v}$ can be computed as

$$
\begin{align*}
L(\mathbf{i}, \mathbf{n}, \mathbf{v} r) & =\int_{\omega_{x}} f(\mathbf{x}, \mathbf{n}, \mathbf{v}) E(\mathbf{x}, \mathbf{n}, r) d \omega_{x}  \tag{2.4}\\
& =f(\mathbf{i}, \mathbf{n}, \mathbf{v}) \frac{I_{0}}{r^{2}} \max \left[0, \mathbf{i}^{\mathrm{T}} \mathbf{n}\right] .
\end{align*}
$$

By substituting (2.2) into (2.4), we can represent the reflected hybrid radiance as

$$
\begin{align*}
L(\mathbf{i}, \mathbf{n}, \mathbf{v}, r) & =\left[\rho_{d} f_{d}(\mathbf{i}, \mathbf{n}, \mathbf{v})+\rho_{s} f_{s}(\mathbf{i}, \mathbf{n}, \mathbf{v})\right] \frac{I_{0}}{r^{2}} \max \left[0, \mathbf{i}^{\mathrm{T}} \mathbf{n}\right] \\
& =\rho_{d} L_{d}(\mathbf{i}, \mathbf{n}, \mathbf{v}, r)+\rho_{s} L_{s}(\mathbf{i}, \mathbf{n}, \mathbf{v}, r) \tag{2.5}
\end{align*}
$$

where

$$
\begin{align*}
& L_{d}=\frac{I_{0}}{r^{2} \pi} \max \left[0, \mathbf{i}^{\mathrm{T}} \mathbf{n}\right], \\
& L_{s}=\frac{I_{0}}{r^{2}} \frac{\exp \left\{-k\left[\cos ^{-1}\left(\mathbf{h}^{\mathrm{T}} \mathbf{n}\right)\right]^{2}\right\}}{\left(\mathbf{i}^{\mathrm{T}} \mathbf{n}\right)\left(\mathbf{v}^{\mathrm{T}} \mathbf{n}\right)} \max \left[0, \mathbf{i}^{\mathrm{T}} \mathbf{n}\right] \tag{2.6}
\end{align*}
$$

are the diffuse and specular radiance components, respectively.

### 2.2. Derived Reflectance Map Model and Image Irradiance Equation

The general perspective projection, which models the ideal pinhole camera, is employed in this work. As depicted in Fig. 2, we consider a camera-centered Cartesian coordinate system with the lens at the origin called the center of projection and the optical axis aligned with the $-Z$-axis. The actual film is located at $Z=f$, where $f$ is the focal length behind the lens. However, to avoid the sign reversal of the coordinates, we assume without loss of generality that the image plane $x-y$ is located at $Z=-f$ in front of the lens. With this model, the mapping between a surface point $\mathbf{P}=(X, Y, Z)^{\mathrm{T}}$ and the projected image point $\mathbf{p}=$ $(x, y,-f)^{\mathrm{T}}$ obeys the relationship


FIG. 2. Imaging geometry model.

$$
\begin{equation*}
x=-f \frac{X}{Z}, \quad y=-f \frac{Y}{Z} \tag{2.7}
\end{equation*}
$$

The irradiance $E$ at an image point of the film is obtained through the lens system via the radiance $L$ of the corresponding surface point. It was shown in [9] that not all, but only a portion of reflected light comes through the lens system which is known as the lens collection [4]. It is described by

$$
\begin{equation*}
E=\frac{\pi}{4}\left(\frac{d}{f}\right)^{2} \cos ^{4} \gamma L \tag{2.8}
\end{equation*}
$$

where $d$ is the diameter of the lens and $\gamma$ is the angle between the ray from the object point to the center of projection and the optical axis.

In general, the converted image intensity or gray value through an electronic sensor device can be written as

$$
\begin{equation*}
I=g E+b \tag{2.9}
\end{equation*}
$$

where $g$ and $b$ are the sensor gain and the bias, respectively, which can be determined by proper camera sensor calibration. As a direct consequence of (2.9), we can define the generalized reflectance map function

$$
\begin{equation*}
R(\mathbf{i}, \mathbf{n}, \mathbf{v}, r)=g \frac{\pi}{4}\left(\frac{d}{f}\right)^{2} \cos ^{4} \gamma L(\mathbf{i}, \mathbf{n}, \mathbf{v}, r)+b \tag{2.10}
\end{equation*}
$$

so that the image intensity $I$ at a point $\mathbf{p}$ in the image plane can be characterized by the image irradiance equation

$$
\begin{equation*}
I(\mathbf{p})=R(\mathbf{i}(\mathbf{S}, \mathbf{P}), \mathbf{n}(\mathbf{P}), \mathbf{v}(\mathbf{P}), r(\mathbf{S}, \mathbf{P})) \tag{2.11}
\end{equation*}
$$

where the relationship between the image point $\mathbf{p}$ and the corresponding surface point $\mathbf{P}$ is defined by the perspective law in (2.7). The cosine function of off-axis angle $\gamma$ can be written as

$$
\cos \gamma=-\mathbf{v}^{\mathrm{T}} \mathbf{n}_{z}
$$

where $\mathbf{n}_{s}=(0,0,1)$ is the unit normal of $X-Y$ plane. By substituting (2.5) and (2.6) into (2.10), the generalized reflectance map can be written as

$$
\begin{aligned}
& R(\mathbf{i}, \mathbf{n}, \mathbf{v}, r) \\
&= g \frac{I_{0}}{r^{2}}\left(\mathbf{v}^{\mathrm{T}} \mathbf{n}_{z}\right)^{4} \frac{\pi}{4}\left(\frac{d}{f}\right)^{2} \\
& \cdot\left(\frac{\rho_{d}}{\pi}+\rho_{s} \frac{\exp \left\{-k\left[\cos ^{-1}\left(\mathbf{h}^{\mathrm{T}} \mathbf{n}\right)\right]^{2}\right\}}{\left(\mathbf{i}^{\mathrm{T}} \mathbf{n}\right)\left(\mathbf{v}^{\mathrm{T}} \mathbf{n}\right)}\right) \max \left[0, \mathbf{i}^{\mathrm{T}} \mathbf{n}\right]+b \\
&= \begin{cases}g \frac{I_{0}}{r^{2}}\left(\mathbf{v}^{\mathrm{T}} \mathbf{n}_{z}\right)^{4} \\
\cdot\left(\beta_{d} \mathbf{i}^{\mathrm{T}} \mathbf{n}+\beta_{s} \frac{\exp \left\{-k\left[\cos ^{-1}\left(\mathbf{h}^{\mathrm{T}} \mathbf{n}\right)\right]^{2}\right\}}{\mathbf{v}^{\mathrm{T}} \mathbf{n}}\right)+b, & \mathbf{i}^{\mathrm{T}} \mathbf{n} \geq 0 \\
0, & \text { otherwise }\end{cases}
\end{aligned}
$$

where the diffuse and specular albedo, $\beta_{d}$ and $\beta_{s}$, are constants of proper dimension that makes $R$ a valid reflectance map. We note that in order to make our algorithm work properly and reconstruct accurate surfaces, the generalized reflectance map parameters, $\beta_{d}, \beta_{s}$, and $k$ of an object should be known a priori. Several algorithms for estimating these parameters of hybrid reflection have been proposed in the literatures, and we refer to $[12,33]$ for more detailed discussion. Methods for estimating the albedo and illumination direction for Lambertian surface can be found in [16, 41].

## 3. DISCRETE GENERALIZED REFLECTANCE MAP PARAMETERIZED BY DEPTH VARIABLES

The SFS problem can be viewed as a problem of solving the image irradiance equation (2.11) with given (observed) image intensity at $\mathbf{p}$. The generalized reflectance map $R$ in (2.10) is a function of the light source position $\mathbf{S}$, the position of a surface point $\mathbf{P}$ and the surface normal at that point which can be expressed as

$$
\mathbf{n}=\frac{(-p,-q, 1)}{\sqrt{p^{2}+q^{2}+1}}, \quad \text { where } p=\frac{\partial Z(X, Y)}{\partial X}, \quad q=\frac{\partial Z(X, Y)}{\partial Y}
$$

Once the light source position and the projected point $\mathbf{p}$
in the image plane are specified, $R$ reduces to a function of three variables, i.e., depth $Z$ and the gradients $p$ and $q$, since the components of $\mathbf{P}$ can be represented in terms of $Z$ by the perspective law with known $\mathbf{p}$. Thus, we are led to a problem of solving three dependent variables, $Z, p$, and $q$, characterizing the underlying surface with one given equation. To solve them uniquely, conventional photometric stereo methods use additional image irradiance equations due to other light sources as constraints [13, 33]. This approach has one difficulty, namely, since the variables $Z$, $p$, and $q$ are viewed as independent variables, the recovered depth variables and surface normals are often inconsistent. This is known as the integrability problem. Note also that the idea of using discrete approximations of $p$ and $q$ by a straightforward finite difference method, which has been applied to the orthographic SFS problem [15, 35], is no longer applicable in this context since the position components $(X, Y)$ are in fact functions of the depth $Z$.

By introducing a triangular element surface model, we represent the surface normal as functions of the nodal depth. Consequently, the reflectance map $R$ can be discretized and parameterized with only the nodal depth variables. Such a procedure greatly simplify the SFS problem formulation and its solution.

### 3.1. Discretization and Parameterization with Triangular Element Surface Model

Our basic idea of discretization and parameterization is to approximate a smooth object surface by the union of triangular surface patches called triangular elements such that the approximating surface can be written as a linear combination of a set of nodal basis functions of compact support. Let us triangulate a square image domain $\Omega$ by dividing it into a set of $M_{t}$ nonoverlapping triangles $T_{i}$, $i=1, \ldots, M_{t}$, with $M_{n}$ nodal points $\mathbf{p}_{i}, i=1, \ldots, M_{n}$, so that the intensity within each triangle is almost homogeneous. Then, we approximate a smooth object surface by a piecewise linear surface consisting of triangular surface patches $S_{i}$ with nodal points $\mathbf{P}_{i}$ in such a way that $S_{i}$ and $\mathbf{P}_{i}$ are perspectively projected to $T_{i}$ and $\mathbf{p}_{i}$, respectively, in the image plane. The approximating surface can be uniquely specified by $\mathbf{P}_{i}$, or equivalently by the surface nodal depth variables $Z_{i}$ associated with $\mathbf{p}_{i}, i=$ $1, \ldots, M_{n}$.

Let us now focus on a triangular surface patch $S_{k}$ and the corresponding projected triangle $T_{k}$ on the image plane as shown in Fig. 3. We denote the nodal vectors (control points) of three vertices of $S_{k}$ as

$$
\mathbf{P}_{i}=\left(X_{i}, Y_{i}, Z_{i}\right), \quad \mathbf{P}_{j}=\left(X_{j}, Y_{j}, Z_{j}\right), \quad \mathbf{P}_{l}=\left(X_{l}, Y_{l}, Z_{l}\right),
$$

and the corresponding projected nodal points in the image plane as


FIG. 3. Image formation on a triangular surface patch.

$$
\mathbf{p}_{i}=\left(x_{i}, y_{i},-f\right), \quad \mathbf{p}_{j}=\left(x_{j}, y_{j},-f\right), \quad \mathbf{p}_{l}=\left(x_{l}, y_{l},-f\right) .
$$

The center of $S_{k}$ is

$$
\begin{aligned}
\mathbf{P}_{k c} & =\left(X_{k c}, Y_{k c}, Z_{k c}\right) \\
& =\left(\frac{X_{i}+X_{j}+X_{l}}{3}, \frac{Y_{i}+Y_{j}+Y_{l}}{3}, \frac{Z_{i}+Z_{j}+Z_{l}}{3}\right) .
\end{aligned}
$$

By using the perspective relationship in (2.7), we can rewrite the components of $\mathbf{P}_{k c}$ in matrix form as

$$
\left(\begin{array}{c}
X_{k c} \\
Y_{k c} \\
Z_{k c}
\end{array}\right)=\frac{-1}{3 f}\left(\begin{array}{ccc}
x_{i} & x_{i} & x_{l} \\
y_{i} & y_{j} & y_{l} \\
-f & -f & -f
\end{array}\right)\left(\begin{array}{c}
Z_{i} \\
Z_{j} \\
Z_{l}
\end{array}\right) .
$$

We assume that $f$ is known and the points $\left(x_{i}, y_{i}\right),\left(x_{j}, y_{j}\right)$, and $\left(x_{i}, y_{l}\right)$ are observed in the image plane. Then, $\mathbf{P}_{k c}$
depends only on the depth values of three nodal points $Z_{i}$, $Z_{j}$, and $Z_{l}$. The unit vector $\mathbf{i}_{k}$ of the light source direction at the center of the surface patch $S_{k}$ can be written as

$$
\mathbf{i}_{k}=\frac{1}{r_{k}}\left(\mathbf{S}-\mathbf{P}_{k c}\right)=\frac{1}{r_{k}}\left(S_{x}-X_{k c}, S_{y}-Y_{k c}, S_{z}-Z_{k c}\right),
$$

where

$$
\begin{align*}
r_{k} & =\left\|\mathbf{S}-\mathbf{P}_{k c}\right\| \\
& =\left[\left(S_{x}-X_{k c}\right)^{2}+\left(S_{y}-Y_{k c}\right)^{2}+\left(S_{z}-Z_{k c}\right)^{2}\right]^{1 / 2} \tag{3.1}
\end{align*}
$$

is the distance between the light source $\mathbf{S}$ and $\mathbf{P}_{k c}$. Similarly, the unit vector $\mathbf{v}_{k}$ along the viewing direction from the surface patch $S_{k}$ can be represented by

$$
\begin{align*}
\mathbf{v}_{k} & =\frac{\mathbf{O}-\mathbf{P}_{k c}}{\left\|\mathbf{O}-\mathbf{P}_{k c}\right\|}  \tag{3.2}\\
& =\frac{1}{\left(X_{k c}^{2}+Y_{k c}^{2}+Z_{k c}^{2}\right)^{1 / 2}}\left(-X_{k c},-Y_{k c},-Z_{k c}\right)
\end{align*}
$$

The unit vector $\mathbf{h}_{k}$ along the specular direction on $S_{k}$ can also be determined by

$$
\mathbf{h}_{k}=\frac{\mathbf{i}_{k}+\mathbf{v}_{k}}{\left\|\mathbf{i}_{k}+\mathbf{v}_{k}\right\|}
$$

Besides, the surface normal $\mathbf{n}_{k}$ of the triangular surface patch $S_{k}$ can be uniquely determined by its three nodal vectors $\mathbf{P}_{i}, \mathbf{P}_{j}$, and $\mathbf{P}_{l}$ via

$$
\begin{align*}
\mathbf{n}_{k} & =\frac{\left(\mathbf{P}_{j}-\mathbf{P}_{i}\right) \times\left(\mathbf{P}_{l}-\mathbf{P}_{i}\right)}{\left|\left(\mathbf{P}_{j}-\mathbf{P}_{i}\right) \times\left(\mathbf{P}_{l}-\mathbf{P}_{i}\right)\right|} \\
& =\frac{\left(X_{j}-X_{i}, Y_{j}-Y_{i}, Z_{j}-Z_{i}\right) \times\left(X_{l}-X_{i}, Y_{l}-Y_{i}, Z_{l}-Z_{i}\right)}{\left|\left(X_{j}-X_{i}, Y_{j}-Y_{i}, Z_{j}-Z_{i}\right) \times\left(X_{l}-X_{i}, Y_{l}-Y_{i}, Z_{l}-Z_{i}\right)\right|} . \tag{3.3}
\end{align*}
$$

By using the perspective relationship in (2.7), we can rewrite $\mathbf{n}_{k}$ in terms of the location of the image points as

$$
\begin{aligned}
\mathbf{n}_{k} & =\frac{\left(\frac{1}{f}\left(x_{i} Z_{i}-x_{j} Z_{j}\right), \frac{1}{f}\left(y_{i} Z_{i}-y_{j} Z_{j}\right), Z_{j}-Z_{i}\right) \times\left(\frac{1}{f}\left(x_{i} Z_{i}-x_{l} Z_{l}\right), \frac{1}{f}\left(y_{i} Z_{i}-y_{l} Z_{l}\right), Z_{l}-Z_{i}\right)}{\left|\left(\frac{1}{f}\left(x_{i} Z_{i}-x_{j} Z_{j}\right), \frac{1}{f}\left(y_{i} Z_{i}-y_{j} Z_{j}\right), Z_{j}-Z_{i}\right) \times\left(\frac{1}{f}\left(x_{i} Z_{i}-x_{l} Z_{l}\right), \frac{1}{f}\left(y_{i} Z_{i}-y_{l} Z_{l}\right), Z_{l}-Z_{i}\right)\right|} \\
& =\frac{\left(-\varphi_{k},-\nu_{k}, \mu_{k}\right)^{\mathrm{T}}}{\left(\varphi_{k}^{2}+\nu_{k}^{2}+\mu_{k}^{2}\right)^{1 / 2}}
\end{aligned}
$$

where

$$
\left(\begin{array}{c}
\varphi_{k} \\
\nu_{k} \\
\mu_{k}
\end{array}\right)=\left(\begin{array}{ccc}
f\left(y_{i}-y_{j}\right) & f\left(y_{l}-y_{i}\right) & f\left(y_{j}-y_{l}\right) \\
f\left(x_{j}-x_{i}\right) & f\left(x_{i}-x_{l}\right) & f\left(x_{l}-x_{j}\right) \\
\left(x_{i} y_{j}-x_{j} y_{i}\right) & \left(x_{l} y_{i}-x_{i} y_{l}\right) & \left(x_{j} y_{l}-x_{l} y_{j}\right)
\end{array}\right)\left(\begin{array}{c}
Z_{i} Z_{l} \\
Z_{i} Z_{j} \\
Z_{l} Z_{j}
\end{array}\right) .
$$

Note that since the vectors $\mathbf{i}_{k}, \mathbf{v}_{k}, \mathbf{n}_{k}, \mathbf{h}_{k}$, and $r_{k}$ are all expressed in terms of $Z_{i}, Z_{j}$, and $Z_{l}$ in the above discussion, the reflectance map $R_{k}$ is only a function of $Z_{i}, Z_{j}$, and $Z_{l}$, i.e., the depth variables of three vertices of $S_{k}$.

Finally, by using the image irradiance equation, we can relate the image intensity $I_{k}$ of a triangle $T_{k}$ directly to the nodal depth values of the corresponding surface patch $S_{k}$ :

$$
\begin{align*}
I_{k} & =R_{k}\left(\mathbf{i}_{k}, \mathbf{n}_{k}, \mathbf{v}_{k}, r_{k}\right) \\
& = \begin{cases}\frac{I_{0}}{r_{k}^{2}}\left(\mathbf{v}_{\mathbf{k}}^{\mathrm{T}} \mathbf{n}_{z}\right)^{4} & \\
\cdot\left(\beta_{d} i_{k}^{\mathrm{T}} \mathbf{n}_{k}+\beta_{s} \frac{\exp \left\{-k\left[\cos ^{-1}\left(\mathbf{h}_{k}^{\mathrm{T}} \mathbf{n}_{k}\right)\right]^{2}\right\}}{\mathbf{v}_{k}^{\mathrm{T}} \mathbf{n}_{k}}\right), & \mathbf{i}_{k}^{\mathrm{T}} \mathbf{n}_{k} \geq 0, \\
0, & \text { otherwise },\end{cases} \\
& =R_{k}\left(Z_{i}, Z_{j}, Z_{l}\right) \tag{3.4}
\end{align*}
$$

### 3.2. Special Cases

We discuss several simplified reflectance map models for some special cases below.

### 3.2.1. Distant Light Source

If the light source is located far away from the object and the camera, the relative depth difference between surface points is negligible compared to the average distance from the object to the light source. For this case, the radiance flux arriving at a surface patch can be approximated as

$$
\begin{equation*}
d \Phi_{i}=\frac{I_{0}}{r^{2}} d A \max \left[0, \cos \vartheta_{i}\right] \approx I_{l} d A \max \left[0, \cos \vartheta_{i}\right] \tag{3.5}
\end{equation*}
$$

where

$$
\frac{I_{0}}{r^{2}} \approx I_{l} \equiv I_{0} / \mathbf{r}^{2}
$$

and where $\mathbf{r}$ is the average distance between the object surface and the light source. Besides, since the light source is far away, we can assume that incident rays from the light source are in parallel with each other. Then, the unit vector $\mathbf{i}_{k}$ of the light source direction in (3.1) is independent of the position of the surface patch so that we can drop the subscription and approximate it as

$$
\begin{equation*}
\mathbf{i} \approx \frac{\mathbf{S}-\mathbf{O}}{\|\mathbf{S}-\mathbf{O}\|}=\frac{\left(S_{x}, S_{y}, S_{z}\right)}{\left(S_{x}^{2}+S_{y}^{2}+S_{z}^{2}\right)^{1 / 2}} \tag{3.6}
\end{equation*}
$$

### 3.2.2. Distant Object

When an object is far away from the camera while the light source is near the object, the perspective projection model can be approximated by the simpler orthographic projection model. With orthogonal projection, we have the following relationship:

$$
\begin{equation*}
x=X, \quad y=Y \tag{3.7}
\end{equation*}
$$

For this case, since all rays from the surface points are in parallel with each other and orthogonal to the image plane, the unit vector $\mathbf{v}_{k}$ along the viewer direction from the surface patch $S_{k}$ in (3.2) becomes independent on the position of the patch, i.e.,

$$
\begin{equation*}
\mathbf{v}_{k}=\mathbf{v}=\mathbf{n}_{z}=(0,0,1)^{\mathrm{T}} \tag{3.8}
\end{equation*}
$$

Moreover, since

$$
\begin{equation*}
\cos \gamma=-\mathbf{n}_{z}^{\mathrm{T}} \mathbf{n}_{z}=-1 \tag{3.9}
\end{equation*}
$$

the lens collection is independent on the surface position and (2.8) reduces to

$$
E=C L
$$

where $C$ is a constant collection factor $\pi d^{2} /\left(4 f^{2}\right)$. Besides, by using the orthographic projection relationship in (3.7), the surface normal in (3.3) can be approximated by

$$
\begin{align*}
\hat{\mathbf{n}}_{k} & =\frac{\left(x_{j}-x_{i}, y_{j}-y_{i}, Z_{j}-Z_{i}\right) \times\left(x_{l}-x_{i}, y_{l}-y_{i}, Z_{l}-Z_{i}\right)}{\left|\left(x_{j}-x_{i}, y_{j}-y_{i}, Z_{j}-Z_{i}\right) \times\left(x_{l}-x_{i}, y_{l}-y_{i}, Z_{l}-Z_{i}\right)\right|} \\
& =\frac{\left(-\hat{\varphi}_{k},-\hat{\nu}_{k}, \hat{\mu}_{k}\right)^{\mathrm{T}}}{\left(\hat{\varphi}_{k}^{2}+\hat{\nu}_{k}^{2}+\hat{\mu}_{k}^{2}\right)^{1 / 2}}, \tag{3.10}
\end{align*}
$$

where

$$
\begin{aligned}
\hat{\varphi}_{k} & =\left(y_{j}-y_{l}\right) Z_{i}+\left(y_{i}-y_{j}\right) Z_{l}+\left(y_{l}-y_{i}\right) Z_{j} \\
\hat{\nu}_{k} & =\left(x_{l}-x_{j}\right) Z_{i}+\left(x_{j}-x_{i}\right) Z_{l}+\left(x_{i}-x_{l}\right) Z_{j} \\
\hat{\mu}_{k} & =\left(x_{j} y_{l}-x_{l} y_{j}\right)+\left(x_{i} y_{j}-x_{j} y_{i}\right)+\left(x_{l} y_{i}-x_{i} y_{l}\right)
\end{aligned}
$$

### 3.2.3. Distant Light Source and Object

If the light source, the object and the camera are far away from one another, both the distant light point source and orthographic projection assumptions hold. For this


FIG. 4. Test Problem 1: (a) ground truth, and reconstructed results with (b) a single image, (c) two photometric stereo images, (d) three photometric stereo images; (e) three noisy photometric stereo images corrupted by $10 \%$ randim noise; (f) the 1D sliced plot of several reconstructed results.
case, the vectors $\mathbf{i}_{k}, \mathbf{v}_{k}$, and $\mathbf{h}_{k}$ are no longer dependent on the surface position. Once the light source direction is specified, the brightness of a pixel in an image plan is determined only by the surface normal $\mathbf{n}_{k}$. By using (3.5), (3.6), (3.8), (3.9), (3.10) and dropping off the unnecessary subscripts, we obtain the simplified reflectance map

$$
\begin{aligned}
& R_{k}\left(\mathbf{i}, \hat{\mathbf{n}}_{k}, \mathbf{v}\right) \\
& \quad= \begin{cases}I_{1}\left(\beta_{d} \mathbf{i}^{\mathrm{T}} \hat{\mathbf{n}}_{k}+\beta_{s} \frac{\exp \left\{-k\left[\cos ^{-1}\left(\mathbf{h}^{\mathrm{T}} \hat{\mathbf{n}}_{k}\right)\right]^{2}\right\}}{\mathbf{v}^{\mathrm{T}} \hat{\mathbf{n}}_{k}}\right), & \mathbf{i}^{\mathrm{T}} \hat{\mathbf{n}}_{k} \geq 0, \\
0, & \text { otherwise. }\end{cases}
\end{aligned}
$$

## 4. FORMULATION OF THE SFS PROBLEM

The shape of an object, which is approximated with a union of linear triangular surface patches, can be characterized by the coordinates of the nodal points $\mathbf{P}_{m}=\left(X_{m}, Y_{m}\right.$, $\left.Z_{m}\right), m=1, \ldots, M_{n}$. Since the positions of the points $\mathbf{p}_{m}=\left(x_{m}, y_{m},-f\right), m=1, \ldots, M_{n}$, in the image plane are given, one can determine the positions of $\mathbf{P}_{m}$ by calculating the nodal depths $Z_{m}, m=1, \ldots, M_{n}$, and applying the perspective law (2.7) to find out the corresponding $X_{m}$ and $Y_{m}$. The reflectance map $R_{k}$ given by (3.4) is a nonlinear function of depth variables $Z_{i}, Z_{j}$, and $Z_{l}$ which makes

(e) $\mathrm{H}-\mathrm{H} 3$ with noise

(f)

FIGURE 4—Continued
the SFS problem difficult to solve. However, we can simplify the solution process by linearizing the reflectance map, solving the linearized problem and then by applying a successive linearization scheme to improve the accuracy of the computed solution.
By successive linearization, we mean that the nodal values obtained from the previous iteration are used as the reference points for linearization for the current iteration.

In terms of mathematics, we can derive a linear approximation of $R_{k}$ at the $n$th iteration by taking the Taylor series expansion about the reference depth values obtained at the $(n-1)$ th iteration as

$$
\begin{aligned}
& R_{k}\left(Z_{i}, Z_{j}, Z_{l}\right) \\
& \approx R_{k}\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right) \\
&+\left.\sum_{m=i, j, l}\left(Z_{m}-Z_{m}^{n-1}\right) \frac{\partial R_{k}\left(Z_{i}, Z_{j}, Z_{l}\right)}{\partial Z_{m}}\right|_{\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right)} \\
&=\left.\sum_{m=i, j, l} \frac{\partial R_{k}\left(Z_{i}, Z_{j}, Z_{l}\right)}{\partial Z_{m}}\right|_{\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right)} Z_{m} \\
&+\left\{R_{k}\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right)\right. \\
&\left.-\left.\sum_{m=i, j, l} \frac{\partial R_{k}\left(Z_{i}, Z_{j}, Z_{l}\right)}{\partial Z_{m}}\right|_{\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right)} Z_{m}^{n-1}\right\}
\end{aligned}
$$

Since the second term of the above equation is equal to a constant, the reflectance map $R_{k}$ over $S_{k}$ is a linear function of depth values $Z_{i}, Z_{j}$, and $Z_{l}$ of the three vertices of $S_{k}$. We may rewrite $R_{k}$ in terms of all nodal depth variables $Z_{m}, m=1, \ldots, M_{n}$. That is,

$$
\begin{equation*}
R_{k} \approx \sum_{m=1}^{M_{n}} \omega_{k m} Z_{m}+\xi_{k}, \tag{4.1}
\end{equation*}
$$

where

$$
\omega_{k m}= \begin{cases}\left.\frac{\partial R_{k}\left(Z_{i}, Z_{j}, Z_{l}\right)}{\partial Z_{m}}\right|_{\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right)}, & \text { if } m \in V_{k}=\{i, j, l\} \text { of } T_{k},  \tag{4.2}\\ 0, & \text { otherwise },\end{cases}
$$

and

TABLE 1
Depth and Orientation Error of the Reconstructed Surfaces in Test Problem 1

| Error | H-H1 | H-H2 | H-H3 | H-H3 <br> (with noise) |
| :--- | :---: | :---: | :---: | :---: |
| Mean-absolute Z | 1.3578 | 0.9729 | 0.0092 | 0.2778 |
| Std-absolute Z | 0.5279 | 0.3275 | 0.0023 | 0.0648 |
| Mean-relative Z (\%) | 2.4243 | 1.7879 | 0.0163 | 0.4753 |
| Mean-pq | 0.1760 | 0.0976 | 0.0584 | 0.0660 |



FIG. 5. Test Problem 2: (a) ground truth of a spherical surface; (b)-(c) photometric stereo images of a Lambertian surface; (d)-(g) photometric stereo images of a hybrid surface.

$$
\begin{equation*}
\xi_{k}=R_{k}\left(Z_{i}^{n-1}, Z_{j}^{n-1}, Z_{l}^{n-1}\right)-\sum_{m=1}^{M_{n}} \omega_{k m} Z_{m}^{n-1} \tag{4.3}
\end{equation*}
$$

and where $V_{k}$ denotes the index set of vertices of $T_{k}$.
Our objective is to determine the nodal depth $Z_{m}$ with one or multiple images. To achieve this goal, we employ a cost functional minimization approach with $J$ different photometric stereo images taken by various illumination directions while keeping the camera position fixed. The scheme reduces to the single image SFS algorithm for $J=1$. The cost functional is chosen to be

$$
\begin{equation*}
\mathscr{E}=\sum_{k=1}^{M_{t}} \sum_{j=1}^{J} \mathscr{E}_{k}{ }_{k}=\sum_{k=1}^{M_{t}} \sum_{j=1}^{J}\left(I_{k}^{j}-R_{k}^{j}\right)^{2}, \tag{4.4}
\end{equation*}
$$

where $\mathscr{E}_{k}{ }_{k}$ denotes the cost term corresponding to the $k$ th triangular domain of the $j$ th image, and $I_{k}^{j}$ and $R_{k}^{j}$ are the observed image irradiance and the reflectance map over the $k$ th triangular domain of the $j$ th image, respectively. It is worthwhile to emphasize that no regularization term is used in (4.4).

By substituting (4.1) into (4.4) and simplifying the expression, we obtain

$$
\begin{align*}
\mathscr{E} & =\sum_{k=1}^{M_{t}} \sum_{j=1}^{J}\left[I_{k}^{j}-\left(\sum_{m=1}^{M_{n}} \omega_{k m}^{j} Z_{m}+\xi_{k}^{j}\right)\right]^{2}  \tag{4.5}\\
& =\frac{1}{2} \mathbf{z}^{\mathrm{T}} \mathbf{A} \mathbf{z}-\mathbf{b}^{\mathrm{T}} \mathbf{z}+c, \quad \mathbf{z}=\left[Z_{1}, Z_{2}, \ldots, Z_{M_{n}} \mathrm{~T}^{\mathrm{T}},\right.
\end{align*}
$$

where the stiffness matrix $\mathbf{A}$ and the load vector $\mathbf{b}$ are the sum of each individual stiffness matrix $\mathbf{A}_{j}$ and the load vector $\mathbf{b}_{j}$ of $j$ th image, respectively. In terms of mathematics, we have

$$
\mathbf{A}=\sum_{j=1}^{J} \mathbf{A}_{j}, \quad \mathbf{b}=\sum_{j=1}^{J} \mathbf{b}_{j},
$$

where the individual stiffness matrix $\mathbf{A}_{j}$ and the load vector $\mathbf{b}_{j}$ can be determined by

$$
\begin{aligned}
{\left[\mathbf{A}_{j}\right]_{m, n} } & =2 \sum_{k=1}^{M_{t}} \omega_{k m}^{j} \omega_{k n}^{j} \\
{\left[\mathbf{b}_{j}\right]_{m} } & =2 \sum_{k=1}^{M_{t}}\left(I_{k}^{j}-\xi_{k}^{j}\right) \omega_{k m}^{j}, \quad 1 \leq m, n \leq M_{n}
\end{aligned}
$$

and $\omega^{j}$ and $\xi^{j}$ are the coefficients in (4.2) and (4.3) for the $j$ th image. The minimization of the quadratic functional in (4.5) with respect to the nodal variables $\mathbf{z}$ is equivalent to finding the solution of a linear system of equations

$$
\mathbf{A z}=\mathbf{b}
$$

Since the stiffness matrix $\mathbf{A}$ is sparse and symmetric, the system can be efficiently solved by iterative methods such as the multigrid method and the preconditioned conjugate gradient method [3, 31].

## 5. EXPERIMENTAL RESULTS

We present some experimental results to demonstrate the performance of the proposed SFS algorithm in this section. In the experiment, we put light sources on the $Z=0$ plane centered around the origin. When two light sources are used, they are chosen to be orthogonal to each other in the azimuth angle, and when three are used, they are placed $120^{\circ}$ apart in the azimuth angle. The angle between the light ray to the object center and camera optical axis is maintained at less than $45^{\circ}$ for each light source. Unless specified otherwise, the initial depth estimates $Z_{i}^{0}, i=1, \ldots, M_{n}$, are set at an arbitrary constant and no a priori knowledge about the true depth is assumed. Since the nodal points whose depth values can be determined by the SFS algorithm are irregular and sparse on the object domain, we perform interpolation to increase the resolution and visibility of the reconstructed surface. To show the performance of the algorithm in a quantitive way, the absolute and relative depth error, as well as the orientation error, of the reconstructed surfaces are calculated and illustrated for each of the synthetic test problems where the ground truth are known.

Test Problem 1: Spherical polyhedron. We examine the performance of the proposed algorithm applied to a spherical polyhedron whose surface is composed of piecewise triangular patches so that it fits the model described in Section 3.1. The $17 \times 17$ surface depth values associated with image nodal points are shown in Fig. 4a. The ideal image intensity $I_{k}$ of each triangle $T_{k}$ with respect to each triangular surface patch $S_{k}$ can be exactly determined by (3.8) so that $I_{k}$ contains no noise. We set the focal length $f=30$, the diffuse and specular weighting factors $\beta_{d}=0.6, \beta_{s}=0.4$, and the surface roughness parameter $k=10$, respectively.
The recovered surface depth values with a single image generated by a light source located at $\left(S_{x}, S_{y}, S_{z}\right)=$ $(52,30,0)$ are shown in Fig. 4b. The result with two photometric stereo images generated by two light sources $\left(S_{x}, S_{y}, S_{z}\right)=(52,30,0)$ and $(-30,52,0)$ is shown in Fig. 4c. We also show in Fig. 4d the reconstructed surface with three photometric stereo images with light sources at $\left(S_{x}, S_{y}, S_{z}\right)=(60,0,0),(-30,52,0)$, and $(-30,-52,0)$. To examine the robustness with respect to noise, we apply our algorithm to three noisy photometric stereo images generated by adding $10 \%$ pseudo random intensity noise to the original images and show the result in Fig. 4 e.


FIG. 6. Reconstruction results of the spherical surface test problem: (a) applying the Lambertian model to two photometric stereo images of a Lambertian surface; (b) applying the Lambertian model to two photometric stereo images of a hybrid surface; (c) applying the Lambertian model to three photometric stereo images of a hybrid surface; (d) applying the general reflectance model to three photometric stereo images of hybrid surfaces; (e) 1D sliced plot of reconstructed Lambertian surfaces; (f) 1D sliced plot of reconstructed hybrid surfaces.

To see the accuracy of the recovery results more clearly, we present the 1 D sliced view of three reconstructed surfaces along with the original one in Fig. 4f, where the solid, dotted, dashdot, "+" and "o" marked lines are used to represent the ground truth and the results depicted in Figs. 4b, c, d, and e, respectively. The initial surface was to be a plane $Z=-100$ in the experiment. We also calculate the mean and standard deviation of the absolute depth error, the mean of the relative percentage depth error, and the average orientation error of each reconstructed surface in Figs. 4b-e and summarize them in Table 1.

We observe from these results that a single image does not provide accurate depth as well as orientation information. Moreover, it is interesting to see that, unlike the Lambertian case where two images are sufficient to recover accurate results $[19,18]$, the non-Lambertian surface can hardly be recovered correctly with two photometric stereo images. The result in Fig. 4 and Table 1 shows that with three photometric stereo images, we can obtain robust and very accurate reconstructions of the non-Lambertian hybrid test surface. Even with images corrupted by $10 \%$ random noise, the algorithm recovers the surface robustly with the relative depth error of less than $0.5 \%$ and the


FIGURE 6-Continued
average orientation error of 0.066 . However, we observe that when the specular factor $\beta_{s}$ or the surface roughness $k$ becomes larger, the accuracy of the reconstructed depth even with three images becomes degraded.

Test Problem 2: Sphere. We sythesize test images by illuminating a $129 \times 129$ spherical surface as shown in Fig. 5a via a pointwise mapping of (2.11) with the surface normal approximated by

$$
\mathbf{n}=\frac{(-p,-q, 1)^{\mathrm{T}}}{\sqrt{p^{2}+q^{2}+1}}
$$

where $p=Z(X+1, Y)-Z(X, Y), q=Z(X, Y+1)-$ $Z(X, Y)$, and we estimate the average intensity $I_{k}$ of each image triangle $T_{k}$ on the tessellated image domain as discussed in [18]. By setting $f=200$, we obtain a set of perspective test images of size $64 \times 64$ by varying both surface reflection parameters and the source position as shown in Figs. 5b-g. Figures 5b and c are two photometric stereo images of a Lambertian surface ( $\beta_{d}=1, \beta_{s}=0$ ) with light sources at $\left(S_{x}, S_{y}, S_{z}\right)=(-150,260,0),(260$, 150,0 ), while Figs. $5 \mathrm{~d}-\mathrm{g}$ are those of a hybrid surface with $\left(\beta_{d}, \beta_{s}, k\right)=(0.5,0.5,7)$ and light sources at $\left(s_{x}, S_{y}, S_{z}\right)$ $=(-150,260,0),(-260,150,0),(260,150,0)$, and $(0$, $-300,0)$, respectively.

Figures 6a and b show the reconstructed Lambertian and non-Lambertian surfaces by using a Lambertian reflectance map model with two sets of photometric stereo images in Figs. 5b and c and Figs. 5d and f, respectively. The results of applying the Lambertian reflectance map and the generalized reflectance map to three photometric images of the hybrid surface in Figs. 5e, f, and $g$ are shown in Figs. 6c and d, respectively. The mean and standard deviations of the absolute depth error, the mean of relative depth error, and the average orientation error of the reconstructed surfaces in Figs. 6a-d are shown in Table 2. By comparing these results, we see clearly that applying the Lambertian model to a non-Lambertian hybrid surface using two or even three photometric stereo images produces a substantial amount of reconstruction error and is thus not appropriate in practice.

TABLE 2
Depth and Orientation Error of the Reconstructed Surfaces in Test Problem 2

| Error | L-L2 | H-L2 | H-L3 | H-H3 | L-L2 (w. depth constr.) | H-H3 <br> (w. depth constr.) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Mean-absolute Z | 4.1282 | 83.2214 | 70.5369 | 6.4409 | 0.5869 | 0.2458 |
| Std-absolute Z | 0.4084 | 11.4276 | 11.3204 | 0.3720 | 0.5008 | 0.2676 |
| Mean-relative Z (\%) | 1.1501 | 12.0248 | 17.7108 | 1.6058 | 0.2920 | 0.1993 |
| Mean-pq | 0.0618 | 0.1285 | 0.2573 | 0.0599 | 0.0646 | 0.0635 |

TABLE 3
The Effect of the Distance of the Object from the Camera in Test Problem 3

| Error-distance | 300 | 400 | 500 |
| :--- | :---: | :---: | ---: |
| Mean-absolute Z | 2.7531 | 6.4409 | 10.6165 |
| Mean-relative Z (\%) | 0.9335 | 1.6058 | 2.1435 |
| Mean-pq | 0.0664 | 0.5999 | 0.0614 |

In Figs. 6e and f, we show the 1D sliced view of reconstructed surfaces, where the solid lines denote the ground truth while the dotted lines represent results by using correct reflectance maps (i.e., results in Figs. 6a and d). Although the overall shapes are close to the original one, compared to the absolute depth of the ground truth, the reconstructed surfaces are shifted globally by a certain amount, which is also represented in the error statistics in
the first and fourth columns of Table 2, that the mean values are relatively large, while the standard deviations and the orientation error are small. By referring the results of the ideal case in Fig. 4c, this discrepancy is primarily due to the quantization noise and noise caused by estimating the image intensity $I_{k}$ via an averaging process. However, this depth discrepancy can be reduced by incorporating the depth information of a single surface point. We used $Z_{\text {ref }}=-380$, the depth of the center points, as a hard constraint. We show the results in Figs. 6e and $f$ with dashed lines, and also we illustrate the statistics of depth error in the last two columns of Table 2. We note that the mean absolute depth error, as well as the mean relative error, are greatly reduced. These data show that more accurate reconstructed surfaces can be obtained by imposing a depth constraint when the noise level in the image is relatively small. We also tested the effect of the distance of an object from the camera. The test results of placing the same spher-


FIG. 7. Test Problem 3: (a) ground truth of a penny surface; (b)-(d) photometric stereo images.


FIG. 8. Results of the penny surface test problem: (a) reconstructed surface using a general reflectance map model with a single depth constraint; (b) 1D sliced view of several reconstructed surfaces.
ical object 300,400 , and 500 units from the camera, are reported in Table 3. These results show that the depth error tends to increase as the object becomes far from the camera. This can be easily expected since the projected image resolution and thus the depth information becomes degraded when the object gets far away. We note, however, due to the smooth varing curvature property of the spherical object, the orientation error does not change so much.

Test Problem 3: Penny. The penny surface is shown in Fig. 7a, and test images synthesized with parameters
$f=250, \beta_{d}=0.4, \beta_{s}=0.6, k=10$ and light sources at $\left(S_{x}, S_{y}, S_{z}\right)=(285,285,0),(-386,104,0)$ and $(104,-386$, 0 ), are given in Figs. 7b-d, respectively. Figure 8a is the reconstructed depth by applying the general reflectance map with a depth constraint $Z_{\text {ref }}=-491$. In this case, the reconstruction error at the depth discontinuities is quite visible. For comparison, we show in Fig. 8b the 1D slice plot of several reconstructed surfaces, where the solid, dashdot, dotted, and dashed lines denote the ground truth and the results obtained by applying a Lambertian model, a generalized reflectance model without and with a depth constraint, respectively. Table 4 compares the depth and orientation error of those reconstructed surfaces.

Compared to previous test problems, both the depth and orientation error becomes substantial even in the case when a depth constraint is imposed. This is partially due to the error occurring around the four corner regions of the object, where the original surface has relatively large depth discontinuities. Besides, since the test surface is more complicated than the smooth spherical and cylindrical surfaces, with considerable local depth variation, the averaging effect in estimating $I_{k}$ is more serious. This explains why the reconstructed surface with the general reflectance model, but without using the depth constraint, becomes worse for this problem than for that of Test Problem 2.

Test Problem 4: Golf ball. In this test, real images taken by a CCD camera are used for the experiment. The three phometric stereo images of a golf ball are shown in Figs. 9a-c. The ball is located at a distance of 45 cm in front of the camera, and the three-point light sources are placed on the $Z=0$ plane $120^{\circ}$ apart from each other in a circular fashion to make the illumination angle to the center of the object approximately $30^{\circ}$. We performed the test with the estimated reflection parameters $\beta_{d}=0.8$, $\beta_{s}=0.2$, and $k=8$ without any depth constraint. The reconstructed surface is shown in Fig. 9d. We also show the $1 D$ depth profile along a diagonal direction in Fig. 9 e . Note that the object surface is inhomogeneous and laminated with transparent paint. Even though the reconstructed surface does not produce very accurate detailed

TABLE 4
Depth Error of the Reconstructed Surfaces in Test Problem 3

| Error | H-L3 | H-H3 | H-H3 <br> (w. depth constr.) |
| :--- | ---: | ---: | :---: |
| Mean-absolute Z | 184.28 | 71.5691 | 13.4443 |
| Std-absolute Z | 11.2425 | 4.7793 | 5.2935 |
| Mean-relative Z (\%) | 38.4671 | 14.8325 | 1.9752 |
| Mean-pq | 0.7712 | 0.4499 | 0.4617 |



FIG. 9. Test Problem 4: (a)-(c) photometric stereo images of a golf ball; (d) reconstructed surface; (e) 1D slice plot.
information which may be caused by the inaccurarcy of estimated reflection parameters, one can still see from the result that the proposed algorithm does recover the complicated surface well qualitatively.

Test Problem 5: Mannequin. The test object is a head of a mannequin as shown in Fig. 10a. Three real photometric stereo images are obtained by arranging light sources $120^{\circ}$ apart from each other in a circular fashion on the source plane $Z=0$. The focal length of the camera is 8 mm and the object is placed approximately 80 cm form the camera. In this test, we make a guess on the reflection parameters $\beta_{d}=0.7, \beta_{s}=0.3$, and $k=7$, and do not impose any depth constraint. The reconstructed surface and the corresponding level contour plot of the marked region of the face in Fig. 10a are illustrated in Figs. 10b and c , respectively. Although there might be substantial error in the reconstructed depth due to inaccuracy in estimating the light source, camera and reflectance parameters, as well as the measurement noise, the overall shape of the reconstructed surface looks good.

## 6. CONCLUSION

A new general physics-based reflectance map model which includes diffuse and specular reflection effects, a nearby point source and perspective projection was derived in this research. We discussed the discretization and parameterization of the derived reflectance map in terms of nodal depth variables by using a triangular element surface representation, and we proposed a direct shape recovery algorithm from shaded images by successively linearizing the reflectance map and minimizing a quadratic cost functional. The proposed method is practically attractive, since it recovers a broad range of object surfaces with different reflective properties and under various geometric and lighting environments. We performed some experiments with synthetic and real images to demonstrate the excellent performance of the new method. The accurate and robust estimation of the surface reflectivity parameters and the light source geometry information is crucial for the accurate surface reconstruction, which is the research topic we are currently investigating.

(a)

(b)

(c)

FIG. 10. Test Problem 5: (a) ground truth of a mannequin; (b) reconstructed surface; and (c) level contour plot.

## REFERENCES

1. E. N. Coleman Jr. and R. C. Jain, Obtaining 3-dimensional shape of textured and specular surface using four-source photometry, Comput. Graphics Image Process. 18, 1982, 309-328.
2. P. Dupuis and J. Oliensis, Direct method for reconstructing shape from shading, in IEEE Conference on Computer Vision and Pattern Recognition, Champaign, IL, June 1992. pp. 453-458.
3. W. Hackbusch, Multi-Grid Methods and Applications, SpringerVerlag, Berlin, 1985.
4. R. M. Haralick and L. G. Shapiro, Computer and Robot Vision, Addison-Wesley, Reading, MA, 1992.
5. G. Healey and T. O. Binford, Local shape from specularity, Computer Vision, Graphics, and Image Processing 42, 1988, 62-86.
6. B. K. P. Horn, Obtaining Shape from Shading Information, MIT Press, Cambridge, MA, 1975.
7. B. K. P. Horn, Understanding image intensities, Artif. Intell. 8, No. 2, 1977, 201-231.
8. B. K. P. Horn, "Height and gradient from shading," International Journal of Computer Vision 5, 1990. 584-595.
9. B. K. P. Horn and R. W. Sjoberg, Calculating the reflectance map, App. Opt. 18, 1979, 1770-1779; in Shape from Shading (B. K. P. Horn and M. J. Brooks, Eds.), MIT Press, Cambridge, MA, 1989.
10. K. Ikeuchi, Determining surface orientations of specular surfaces by
using the photometric stereo method, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-3, No. 6, 1981, 661-669.
11. K. Ikeuchi and B. K. P. Horn, "Numerical shape from shading and occluding boundaries," Artificial Intelligence 17, 1981, 141-184; in Shape from Shading, B. K. P. Horn and M. J. Brooks (eds.) 1989, MIT Press, Cambridge, MA.
12. K. Ikeuchi and K. Sato, Determining reflectance properties of an object using range and brightness images, IEEE Trans. Pattern Analysis and Machine Intelligence 13, No. 11, 1991, 1139-1153.
13. B. Kim and P. Burger, Depth and shape from shading using the photometric stereo method, Comput. Vision Graphics Image Process. Image Understanding 54, No. 3, 1991, 416-427.
14. R. Kozera, Existence and uniqueness in photometric stereo, I, Appl. Math. Comput. 44, 1991, 1-103.
15. Y. G. Leclerc and A. F. Bobick, The direct computation of height from shading, in IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, May 1991, 552-558.
16. C. H. Lee and A. Rosenfeld, Improved methods of estimating shape from shading using light source coordinate system, Artificial Intelligence 26, 1985; 125-143. Also in Shape from Shading, B. K. P. Horn and M. J. Brooks (eds.) 1989, MIT Press, Cambridge, MA.
17. K. M. Lee and C.-C. J. Kuo, Shape from shading with a linear triangular element surface model, IEEE Trans. Pattern Analysis and Machine Intelligence 15, 1993, 815-822.
18. K. M. Lee and C.-C. J. Kuo, Surface reconstruction from photometric stereo images, Journal of Optical Society of America: A 10, 1993, 855-868.
19. K. M. Lee and C.-C. J. Kuo, Shape from shading with perspective projection, Computer Vision, Graphics, Image Processing: Image Understanding, 59, 1994.
20. S. K. Nayar, K. Ikeuchi, and T. Kanade, Determining shape and reflectance of hybrid surfaces by photometric sampling, IEEE Trans. Robotics and Automation 6, 1990, 418-431.
21. S. K. Nayar, K. Ikeuchi, and T. Kanada, Surface reflection: physical and geometricl perspectives, IEEE Trans. Pattern Analysis and Machine Intelligence 13, 1991, 611-634.
22. S. K. Nayar and M. Oren, Visual Appearance of Matte Surfaces, Science 267, No. 5201, 1995, 1153-1156.
23. F. E. Nicodemus, J. C. Richmond, J. J. Hsia, and I. W. Ginsberg, Geometrical Considerations and Nomenclature for Reflectance, Boulder CO: National Bureau of Standards, 1977.
24. J. Oliensis, Shape from Shading as a Partially Well-Constrained Problem, Computer Vision, Graphics, and Image Processing: Image Understanding 54, No. 2, 1991, 163-183.
25. M. Oren and S. K. Nayar, Diffuse Reflectance from Rough Surfaces,
in IEEE Conference on Computer Vision and Pattern Recognition, 1993, pp. 290-295.
26. M. Oren and S. K. Nayar, Generalization of Lambert's reflection model, in Proc. of ACM SIGGRAPH, Orlando, Florida, July 1994.
27. M. Oren and S. K. Nayar, Generalization of the Lambertian model and implication for machine vision, International Journal of Computer Vision 14, No. 3, 1995, 227-251.
28. M. A. Penna, Local and semi-local shape from shading for a single perspective image of a smooth object, Computer Vision, Graphics, and Image Processing 46, 1989, 346-366.
29. M. A. Penna, A shape from shading analysis for a single perspective image of a polyhedron, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-11, 1989, 545-554.
30. A. P. Pentland, Shape information from shading: a theory about human perception, in Proc. of International Conf. on Computer Vision 1988, 404-413.
31. R. Szeliski, Fast surface interpolation using hierarchical basis functions, IEEE Trans. Pattern Analysis and Machine Intelligence 12, No. 6, 1990, 513-528.
32. R. Szeliski, Fast shape from shading Computer Vision, Graphics, Image Processing: Image Understanding 53, No. 2, 1991, 129-153.
33. H. D. Tagare and R. J. P. DeFigueiredo, A theory of photometric stereo for a class of diffuse non-Lambertian surface, IEEE Trans. Pattern Analysis and Machine Intelligence 13, No. 2, 1991, 133-152.
34. K. E. Torrance and E. M. Sparrow, Theory for off-specular reflection from roughened surfaces, Journal of Optical Society of America 57, No. 9, 1967, 1105-1114.
35. P. S. Tsai and M. Shah, A fast linear shape from shading, in IEEE Conference on Computer Vision and Pattern Recognition, Champaign, Illinois, June 1992, pp. 734-736.
36. P. S. Tsai and M. Shah, A Simple Shape from Shading Algorithm, Tech. Rep. CS-TR-92-24, CS Dept., Univ. of Central Florida, 1992.
37. L. B. Wolff, Diffuse-reflectance model for smooth dielectric surface, Journal of Optical Society of America: A 11, 1994, 2956-2968.
38. L. B. Wolff, On the relative brightness of specular and diffuse reflection, in IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 369-376.
39. R. J. Woodham, Analyzing images of curved surfaces, Artificial Intelligence 17, No. 1-3, 1981, 117-140.
40. R. Zhang, P. S. Tsai, J. E. Cryer, and M. Shad, Analysis of Shape from Shading Techniques, in IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 377-384.
41. Q. Zheng and R. Chellappa, Estimation of illumination direction, albedo, and shape from shading, IEEE Trans. Pattern Analysis and Machine Intelligence 13, No. 7, 1991, 680-702.

[^0]:    * This work was supported by the National Science Foundation Presidential Faculty Fellow Award ASC-9350309.
    $\dagger$ E-mail: cckuo@sipi.usc.edu.

