CVGIP: IMAGE UNDERSTANDING
Vol. 59, No. 2, March, pp. 202–212, 1994

Shape from Shading with Perspective Projection*

KYOUNG MU LEE AND C.-C. JAY KUO

Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564

Received September 11, 1992; accepted September 24, 1993

Most conventional SFS (shape from shading) algorithms have been developed under the assumption of orthographic projection. However, the assumption is not valid when an object is not far away from the camera and, therefore, it causes severe reconstruction error in many real applications. In this research, we develop a new iterative algorithm for recovering surface heights from shaded images obtained with perspective projection. By dividing an image into a set of nonoverlapping triangular domains and approximating a smooth surface by the union of triangular surface patches, we can relate image brightness in the image plane directly to surface nodal heights in the world space via a linearized reflectance map based on the perspective projection model. To determine the surface height, we consider the minimization of a cost functional defined to be the sum of squares of the brightness error by solving a system of equations parameterized by nodal heights. Furthermore, we apply a successive linearization scheme in which the linearization of the reflectance map is performed with respect to surface nodal heights obtained from the previous iteration so that the approximation error of the reflectance map is reduced and accuracy of the reconstructed surface is improved iteratively. The proposed method reconstructs surface heights directly and does not require any additional integrability constraint. Simulation results for synthetic and real images are demonstrated to show the performance and efficiency of our new method. © 1994 Academic Press, Inc.

1. INTRODUCTION

Recovering explicit geometric information of 3D objects from 2D projected images has been important for many computer vision problems. It supplies a great deal of useful information for both basic and higher level vision tasks. Among various available depth cues, shading often provides reliable dense depth information of an underlying smooth object. Since initiated by Horn [7, 8], the shape from shading (SFS) problem has received a lot of attention. Among many significant developments made in this area, we want to point out the following: characteristic strip expansion [7, 8], variational approach [5, 10, 11, 15, 18, 30–32, 39], Fourier transform approach [26], direct height recovery [10, 17, 19, 35], optimal control approach [4], photometric stereo [3, 13, 16, 21, 23, 27, 29, 33, 36–38], etc. For a thorough review for work before 1990, we refer to [12, 10]. More recent work includes [1, 2, 4, 16, 17, 19, 21, 22].

The modeling of the image formation process, which defines a transformation between surface shape and projected image intensity, is crucial for the SFS problem. The image formation process depends mainly on the projection model, the surface reflectance property, surface orientations, and the lighting condition. A projection model defines the geometric relationship between a surface point and the projected point on image plane. Two commonly used projection models are *perspective* and *orthographic* projections [9]. By perspective projection, we mean an ideal image forming model which works like a pin hole camera that forces the light rays to go through one single point, and thus the position of a projected point on the image plane is dependent on both the depth and position of a surface point in the world space as well as the focal length of the camera. If the size of the objects in view is small compared to the viewing distance, perspective projection can be approximated as orthographic projection in which all rays are parallel to the optical axis, and the position of a projected point is independent of the depth of a surface point. Most conventional SFS algorithms assume orthographic projection to simplify the analysis. However, objects are often not located very far away from the camera in many real applications so that orthographic projection is not appropriate and distortions of reconstructed surfaces occur. To make the SFS algorithm practically useful, one may have to adopt the more general perspective projection model.

Few attempts have so far been made to solve the SFS problem under the perspective projection assumption. Penna [24, 25] proposed a local SFS algorithm applicable to a single image of a smooth object or an opague polyhedron with perspective projection. He derived a system of algebraic equations from the image irradiance equation,

* This work was supported by a National Science Foundation Young Investigator Award ASC-9258396.

1049-9660/94 $6.00

which describes the local surface geometry of an object, and solved it by using the Newton method. However, no experimental result with real images was reported in [24, 25]. According to our experience, local SFS algorithms in general suffer from noise and are therefore not very useful when applied to real images. In this paper, we present a new iterative SFS algorithm by using images under perspective projection. Our algorithm is robust, fast, and insensitive to noise, since it is based on the global minimization approach. Given one reference nodal height, it determines the absolute surface height directly by discretizing and solving the minimization problem. No integrability problem arises in our formulation. More recently, methods of combining shading and geometric stereo information for the SFS problem were proposed [14, 17, 20, 28]. However, since orthographic and perspective projections are assumed, respectively, for SFS and stereopsis, there exists an obvious inconsistency in their work. In contrast, imposing stereo data information in our new perspective SFS algorithm is natural and consistent, since the same perspective projection imaging model is used for both SFS and stereopsis.

The basic idea of our approach is based on a triangular element surface model and a linearized reflectance map. The projected image is divided into a set of nonoverlapping triangular domains with homogeneous brightness in each triangle. The surface normal of each triangular patch can be represented by its three nodal vectors in the world space. By using the perspective relationship between the projected and true surface nodal points, one can establish a linear relationship between image intensity and the surface nodal heights via the image irradiance equation with a linearized reflectance map. To determine the nodal heights, we minimize a global cost functional defined to be the sum of squared brightness error over each projected triangular image domain. The minimization problem is equivalent to solving a linear system of algebraic equations parameterized by the nodal heights, and thus can be efficiently solved by MG (multi-grid) or PCG (preconditioned conjugate gradient) algorithms. To minimize the approximation error of the linearized reflectance map, we employ a successive linearization scheme in which nodal heights obtained from the previous iteration are used as reference points for linearization at the current iteration.

This paper is organized as follows. In Section 2, we discuss models of projection, reflectance map and triangular element surface approximation. The perspective SFS problem is formulated as a finite dimensional optimization problem which minimizes a quadratic cost functional consisting of squares of brightness error, and the construction of stiffness matrix and load vector is examined in Section 3. The formulation is generalized to incorporate multiple photometric stereo images in Section 4. Experimental results on several synthetic and real images are given in Section 5.

2. DISCUSSION ON THE MODEL

Similar to most other SFS work, the mate object surface and a distant single light source are assumed in this work. Distinctive features of our model such as the reflectance map with perspective projection and the triangular element surface approximation will be examined in detail in this section.

2.1. *Reflectance Map with Perspective Projection*

A projection model defines a relationship between points in the 3D space and their corresponding pixels in a 2D picture. The perspective projection imaging model provides a natural first-order approximation to image formation process, in which all rays pass through a fixed point, called the *center of projection* or the *view point*. Consider a viewer-centered Cartesian coordinate system with the origin at the center of projection and the $-Z$-axis aligned with the optical axis as depicted in Fig. 1. Let the x- and y-axes of the image plane be parallel to the X- and Y-axes and the *focal length*, the distance between the origin and the image plane, be f. Then, a surface point $\mathbf{P} = (X, Y, Z)^T$ is transformed to a projected image point $\mathbf{p} = (x, y, -f)^T$ via the perspective relationship

$$x = -f\frac{X}{Z}, \quad y = -f\frac{Y}{Z}, \tag{2.1}$$

which relates the image coordinates (x, y) to the world coordinates (X, Y, Z) of a point with a given focal length f. Orthogonal projection can be viewed as a special case of the perspective projection model when an object is far away from the viewer. That is, if the depth range of an object is quite small compared with the average distance

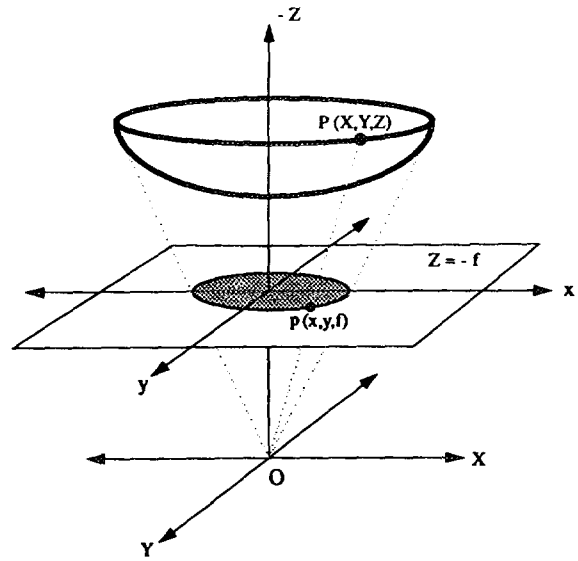


FIG. 1. Perspective projection.

from the viewer to the object, we have the approximations,

$$x \approx mX, \quad y \approx mY, \tag{2.2}$$

where $m = -f/Z_0$ and Z_0 is the average depth value of Z. Since m is positive constant, one may rescale the image coordinate in a way such that $x = X$ and $y = Y$ for convenience. Thus, by orthographic projection, one assumes all rays are perpendicular to the image plane and parallel to optical axis as shown in Fig. 2.

Note that under orthographic projection, the spatial relationship between image points (x, y) is preserved in the world plane but not under a perspective projection since the relationship is distorted by Z at each point by (2.1). Since one has to determine three dependent variables (X, Y, Z) for given image point (x, y) with perspective projection, but only the variable Z with orthographic projection, the perspective SFS problem is more difficult than the orthographic SFS problem.

With perspective projection, an image intensity $E(x, y)$ at point (x, y) on the image plane is formed through the *image irradiance* equation,

$$E(x, y) = E(X, Y, Z) = R(p(X, Y), q(X, Y)), \tag{2.3}$$

where R is the reflectance map function, $p(X, Y) = \partial Z(X, Y)/\partial X$ and $q(X, Y) = \partial Z(X, Y)/\partial Y$. If we assume ideal Lambertian surface illuminated by a single distant point light source, the reflectance map can be written as

$$R(p, q) = \begin{cases} \eta \dfrac{K}{\sqrt{1 + p^2 + q^2}}, & K \geq 0, \\ 0, & K < 0, \end{cases} \tag{2.4}$$

where

$$K = -p \cos \tau \sin \sigma - q \sin \tau \sin \sigma + \cos \sigma,$$

and where η is the composite albedo of the surface, τ and σ are the *tilt* and *slant* angles of the illumination direction pointing toward the light source. This is typically a nonlinear function depending on the surface gradient $p(X, Y)$ and $q(X, Y)$.

2.2. *Triangular Element Surface Approximation*

A triangular element surface model has been recently proposed and applied successfully to the SFS problem with orthographic projection [19, 21]. By the model, we approximate a smooth surface with a union of finite triangular surface patches so that the orientation (surface normal) of each triangular patch can be uniquely determined by its three vertices vectors. Projection of triangular surface patches with nodal points through the perspective law

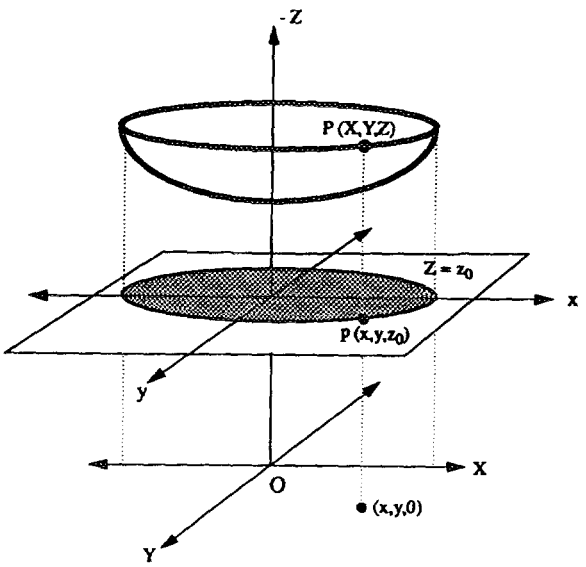


FIG. 2. Orthogonal projection.

(2.1) produces corresponding triangles and nodal points in the image plane. We assume that the image intensity within each triangular image domain is homogeneous and equal to that generated by the corresponding triangular surface patch so that a direct relationship between image intensity and surface nodal heights can be established via the image irradiance equation (2.3).

3. FORMULATION OF PERSPECTIVE SFS

3.1. *Image Formation on a Triangular Surface Patch*

For a smooth and non-occluded object surface, there exists a one-to-one correspondence between triangular patches on a surface and triangles in the image plane. They are denoted respectively by S_i and T_i, $i = 1, 2, \ldots, M_t$, where M_t is the total number of triangles. Similarly, nodal points on a surface and the corresponding projected points in the image plane are denoted by $\mathbf{P}_i$ and $\mathbf{p}_i$, $i = 1, 2, \ldots, M_n$, respectively, where M_n is the total number of nodal points. Consider a small triangular surface patch S_k and projected triangle T_k on image plane as shown in Fig. 3.

Let V_k be the index set of vertices points of T_k or S_k. Then, from Fig. 3, we have

$$V_k = \{i, j, l\}.$$

The image intensity E_k over T_k is determined by

$$E_k = R_k(p_k, q_k) = \eta \frac{\cos \sigma - p_k \cos \tau \sin \sigma - q_k \sin \tau \sin \sigma}{\sqrt{(1 + p_k^2 + q_k^2)}}, \tag{3.1}$$

where (p_k, q_k) is the gradient of S_k. The surface normal $\mathbf{n}_k$ of the triangular surface patch S_k can be determined with its three nodal vectors $\mathbf{P}_i$, $\mathbf{P}_j$, and $\mathbf{P}_l$ by

$$\mathbf{n}_k = \frac{(\mathbf{P}_j - \mathbf{P}_i) \times (\mathbf{P}_l - \mathbf{P}_i)}{|(\mathbf{P}_j - \mathbf{P}_i) \times (\mathbf{P}_l - \mathbf{P}_i)|} = \frac{(X_j - X_i, Y_j - Y_i, Z_j - Z_i) \times (X_l - X_i, Y_l - Y_i, Z_l - Z_i)}{|(X_j - X_i, Y_j - Y_i, Z_j - Z_i) \times (X_l - X_i, Y_l - Y_i, Z_l - Z_i)|}. \tag{3.2}$$

By using (2.1), we can rewrite (3.2) as

$$\mathbf{n}_k = \frac{((1/f)(x_iZ_i - x_jZ_j), (1/f)(y_iZ_i - y_jZ_j), Z_j - Z_i) \times ((1/f)(x_iZ_i - x_lZ_l), (1/f)(y_iZ_i - y_lZ_l), Z_l - Z_i)}{|((1/f)(x_iZ_i - x_jZ_j), (1/f)(y_iZ_i - y_jZ_j), Z_j - Z_i) \times ((1/f)(x_iZ_i - x_lZ_l), (1/f)(y_iZ_i - y_lZ_l), Z_l - Z_i)|}.$$

Besides, the surface normal $\mathbf{n}_k$ is related to the gradient (p_k, q_k) via

$$\mathbf{n}_k = \frac{(-p_k, -q_k, 1)^T}{\sqrt{1 + p_k^2 + q_k^2}}.$$

Then, it is straightforward to verify that

$$p_k = \frac{\alpha_k}{\gamma_k}, \quad q_k = \frac{\beta_k}{\gamma_k}, \tag{3.3a}$$

where

$$\begin{aligned} \alpha_k &= f\{(y_i - y_j)Z_iZ_j + (y_l - y_i)Z_iZ_l + (y_j - y_l)Z_jZ_l\}, \\ \beta_k &= f\{(x_j - x_i)Z_iZ_j + (x_i - x_l)Z_iZ_l + (x_l - x_j)Z_jZ_l\}, \\ \gamma_k &= (x_iy_j - x_jy_i)Z_iZ_j + (x_ly_i - x_iy_l)Z_iZ_l + (x_jy_l - x_ly_j)Z_jZ_l. \end{aligned} \tag{3.3b}$$

Since the image points (x_i, y_i), (x_j, y_j), and (x_l, y_l) are known, α_k, β_k, and γ_k are functions of only three nodal height variables Z_i, Z_j, and Z_l. By substituting (3.3a) and (3.3b) into (3.1), we obtain a nonlinear reflectance map function with variables Z_i, Z_j, and Z_l,

$$R_k(Z_i, Z_j, Z_l) = \eta \frac{\gamma_k \cos\sigma - \alpha_k \cos\tau \sin\sigma - \beta_k \sin\tau \sin\sigma}{(\alpha_k^2 + \beta_k^2 + \gamma_k^2)^{1/2}}. \tag{3.4}$$

Local linear approximation of the reflectance map can be achieved by taking Taylor series expansion of $R_k(Z_i, Z_j, Z_l)$ about a reference nodal height (Z_i^0, Z_j^0, Z_l^0) through the first-order term,

$$\begin{aligned} R_k(Z_i, Z_j, Z_l) \approx R_k(Z_i^0, Z_j^0, Z_l^0) &+ (Z_i - Z_i^0)\frac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_i}\bigg|_{(Z_i^0, Z_j^0, Z_l^0)} \\ &+ (Z_j - Z_j^0)\frac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_j}\bigg|_{(Z_i^0, Z_j^0, Z_l^0)} \\ &+ (Z_l - Z_l^0)\frac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_l}\bigg|_{(Z_i^0, Z_j^0, Z_l^0)}. \end{aligned} \tag{3.5}$$

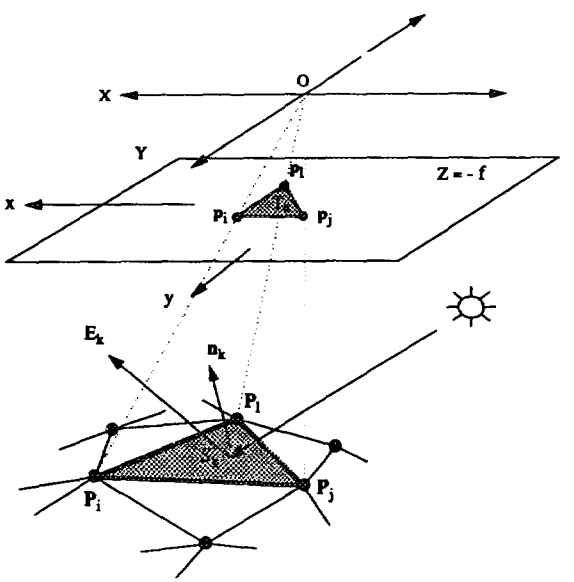


FIG. 3. Image formation on a triangular surface patch.

Thus, we can express the brightness E_k over T_k as a linear combination of the surface nodal heights Z_i, $i = 1, \ldots, M_n$; i.e.,

$$E_k \approx \sum_{m=1}^{Mn} \omega_{km} Z_m + \xi_k, \tag{3.6}$$

where

$$\omega_{km} = \begin{cases} \left.\dfrac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_m}\right|_{(Z_i^0, Z_j^0, Z_l^0)}, & \text{if } m \in V_k = \{i, j, l\} \text{ of } T_k, \\ 0, & \text{otherwise,} \end{cases} \tag{3.7}$$

and

$$\xi_k = R_k(Z_i^0, Z_j^0, Z_l^0) - \sum_{m=1}^{Mn} \omega_{km} Z_m^0. \tag{3.8}$$

3.2. *Perspective SFS with a Single Image*

Our objective is to determine the surface height Z_i, $i = 1 \ldots, M_n$. This can be achieved by minimizing a cost functional defined as

$$E = \sum_{k=1}^{Mt} E_k = \sum_{k=1}^{Mt} (E_k - \hat{E}_k)^2, \tag{3.9}$$

where E_k denotes the cost term corresponding to the kth triangular domain, and E_k and $\hat{E}_k$ are the observed and reconstructed image intensities over the kth triangular domain, respectively. By substituting (3.6) into (3.9), we have

$$\begin{aligned} E &= \sum_{k=1}^{Mt} \left[E_k - \left(\sum_{m=1}^{Mn} \omega_{km} Z_m + \xi_k \right) \right]^2 \\ &= \sum_{k=1}^{Mt} \left[\sum_{i=1}^{Mn} \sum_{j=1}^{Mn} \omega_{ki} \omega_{kj} Z_i Z_j - 2 \right. \\ &\quad \left. (E_k - \xi_k) \sum_{i=1}^{Mn} \omega_{ki} Z_i + (E_k - \xi_k)^2 \right] \\ &= \tfrac{1}{2} \sum_{i=1}^{Mn} \sum_{j=1}^{Mn} \left(2 \sum_{k=1}^{Mt} \omega_{ki} \omega_{kj} \right) Z_i Z_j - \sum_{i=1}^{Mn} \left[2 \sum_{k=1}^{Mt} (E_k - \xi_k) \omega_{ki} \right] Z_i \\ &\quad + \sum_{k=1}^{Mt} (E_k - \xi_k)^2 \\ &= \tfrac{1}{2} \mathbf{z}^{\mathrm{T}} \mathbf{A} \mathbf{z} - \mathbf{b}^{\mathrm{T}} \mathbf{z} + c, \quad \mathbf{z} = [Z_1, Z_2, \ldots, Z_{Mn}]^{\mathrm{T}}. \end{aligned} \tag{3.10}$$

The elements a_{ij} and b_i of the stiffness matrix $\mathbf{A}$ and the load vector $\mathbf{b}$ can be determined, respectively, as

$$a_{ij} = 2 \sum_{k=1}^{Mt} \omega_{ki} \omega_{kj}, \quad b_i = 2 \sum_{k=1}^{Mt} (E_k - \xi_k) \omega_{ki}, \quad 1 \le i, j \le M_n. \tag{3.11}$$

Thus, by minimizing the above quadratic functional, one can determine the nodal height $\mathbf{z}$ and this can be done by solving the equivalent linear system of equations,

$$\mathbf{A}\mathbf{z} = \mathbf{b}. \tag{3.12}$$

Note that the element of the $\mathbf{A}$, $a_{ij} = 0$, if i and j are not the neighboring nodal points, since in that case no triangle has both nodal points i and j as its vertices and thus either ω_{ki} and ω_{kj} is zero. Suppose that node i has six neighboring nodal points $j_1, \ldots, j_6$ and the six triangular domains surrounding it, $T_k, \ldots, T_p$ as shown in Fig. 4(a). Then, we can represent the stiffness matrix using a seven-point stencil operator form as illustrated in Fig. 4(b). The nonzero elements of stiffness matrix $\mathbf{A}$ and load vector $\mathbf{b}$ associated with node i are determined in terms of the coefficients ω's in (3.7) and brightness $E_k, \ldots, E_p$, and they are summarized in Table 1.

3.3. *Special Case: Orthographic Projection*

It is clear that the orthographic SFS problem is a special case of the perspective SFS problem so that the framework developed above applies to both orthographic and perspective projected images. Under the orthographic projection assumption and by using (2.2) with $m = 1$, one can approximate the surface normal $\mathbf{n}_k$ by

$$\mathbf{n}_k = \frac{(x_j - x_i, y_j - y_i, Z_j - Z_i) \times (x_l - x_i, y_l - y_i, Z_l - Z_i)}{|(x_j - x_i, y_j - y_i, Z_j - Z_i) \times (x_l - x_i, y_l - y_i, Z_l - Z_i)|}, \tag{3.13}$$

Then, the gradient (p_k, q_k) become

$$p_k = \frac{\hat{\alpha}_k}{\hat{\gamma}_k}, \quad q_k = \frac{\hat{\beta}_k}{\hat{\gamma}_k}, \tag{3.14a}$$

TABLE 1
Elements of Stiffness Matrix A and Load Vector b

$a_{i,i}$	$2(\omega_{ki}^2 + \omega_{li}^2 + \omega_{mi}^2 + \omega_{ni}^2 + \omega_{oi}^2 + \omega_{pi}^2)$
a_{ij_1}	$2(\omega_{pi}\omega_{pj_1} + \omega_{ki}\omega_{kj_1})$
a_{ij_2}	$2(\omega_{ki}\omega_{kj_2} + \omega_{li}\omega_{lj_2})$
a_{ij_3}	$2(\omega_{li}\omega_{lj_3} + \omega_{mi}\omega_{mj_3})$
a_{ij_4}	$2(\omega_{mi}\omega_{mj_4} + \omega_{ni}\omega_{nj_4})$
a_{ij_5}	$2(\omega_{ni}\omega_{nj_5} + \omega_{oi}\omega_{oj_5})$
a_{ij_6}	$2(\omega_{oi}\omega_{oj_6} + \omega_{pi}\omega_{pj_6})$
b_i	$2((E_k - \xi_k)\omega_{ki} + (E_l - \xi_l)\omega_{li} + (E_m - \xi_m)\omega_{mi} + (E_n - \xi_n)\omega_{ni} + (E_o - \xi_o)\omega_{oi} + (E_p - \xi_p)\omega_{pi})$

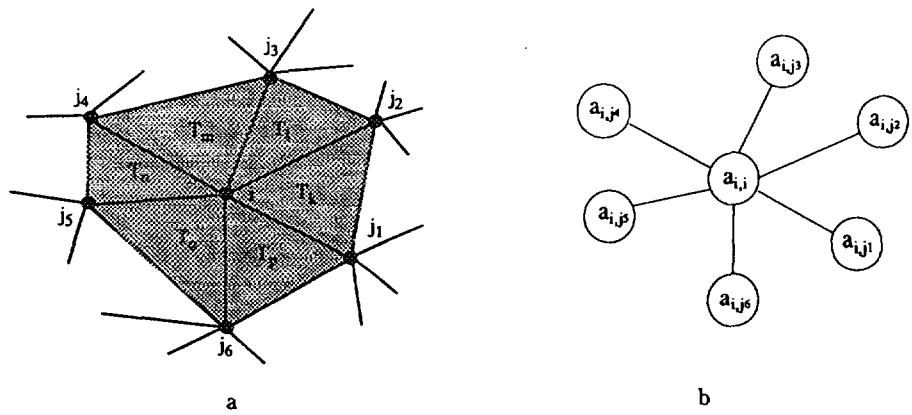


FIG. 4. (a) An example of a node configuration associated with node i; (b) corresponding seven-point stencil operator.

where

$$\hat{\alpha}_k = (y_j - y_l)Z_i + (y_i - y_j)Z_l + (y_l - y_i)Z_j,$$
$$\hat{\beta}_k = (x_l - x_j)Z_i + (x_j - x_i)Z_l + (x_i - x_l)Z_j,$$
$$\hat{\gamma}_k = (x_j y_l - x_l y_j) + (x_i y_j - x_j y_i) + (x_l y_i - x_i y_l). \quad (3.14b)$$

Note here that the surface gradients p_k and q_k derived with orthographic projection are linear functions of Z_i, Z_j, and Z_l, whereas they are nonlinear functions with perspective projection as shown in (3.3). Similarly, we can express the reflectance map as a function of variables Z_i, Z_j, and Z_l by substituting (3.14a) and (3.14b) into (3.1), i.e.,

$$R_k(Z_i, Z_j, Z_l) = \eta \frac{\hat{\gamma}_k \cos\sigma - \hat{\alpha}_k \cos\tau \sin\sigma - \hat{\beta}_k \sin\tau \sin\sigma}{(\hat{\alpha}_k^2 + \hat{\beta}_k^2 + \hat{\gamma}_k^2)^{1/2}}. \quad (3.15)$$

By taking linear approximation of this reflectance map about a certain reference nodal heights, the same shape reconstruction procedure as described in Section 3.2 follows.

4. PERSPECTIVE SFS WITH PHOTOMETRIC STEREO IMAGES

We use the same framework to solve a related problem which is more robust and practically important. That is, surface reconstruction from multiple perspective photometric stereo images taken under different lighting conditions with the camera position fixed. Let us consider a new cost functional which is the sum of squared brightness error associated with each image. Suppose we have J different photometric images, then the cost functional to be minimized is

$$E = \sum_{k=1}^{Mt} \sum_{j=1}^{J} (E_k^j - \hat{E}_k^j)^2,$$
$$= \sum_{k=1}^{Mt} \sum_{j=1}^{J} \left[E_k^j - \left(\sum_{m=1}^{Mn} \omega_{km}^j Z_m + \xi_k^j \right) \right]^2 \quad (4.1)$$
$$= \tfrac{1}{2}\mathbf{z}^T \tilde{\mathbf{A}} \mathbf{z} - \tilde{\mathbf{b}}^T \mathbf{z} + \tilde{c},$$

where the overall stiffness matrix $\tilde{\mathbf{A}}$, the load vector $\tilde{\mathbf{b}}$, and $\tilde{c}$ are the sum of each individual stiffness matrix $\mathbf{A}_j$, the load vector $\mathbf{b}_j$, and c_j, respectively, i.e.,

$$\tilde{\mathbf{A}} = \sum_{j=1}^{J} \mathbf{A}_j, \quad \tilde{\mathbf{b}} = \sum_{j=1}^{J} \mathbf{b}_j, \quad \tilde{c} = \sum_{j=1}^{J} c_j,$$

and the individual stiffness matrix and load vector can be computed by (3.11). The minimization of the quadratic functional (4.1) leads to the solution of the system of equations

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

Simple analysis and observation shows that two photometric stereo images taken under illumination directions for which the tilt angles are orthogonal to each other give the best results. They also make the algorithm stable and its convergence fast [21].

Note that since an approximation error may result from the linear approximation of a reflectance map, the reconstructed surface height may not be accurate. The approximation error of a reflectance map depends primarily on the choice of the reference surface nodal height values Z_i^0, $i = 1, \ldots, M_n$. A successive linearization scheme has been developed and successfully applied in our previous work for the orthogonal SFS problem [19, 21] to reduce the linear approximation error of the reflectance map. In this scheme, the surface gradients for each trian-

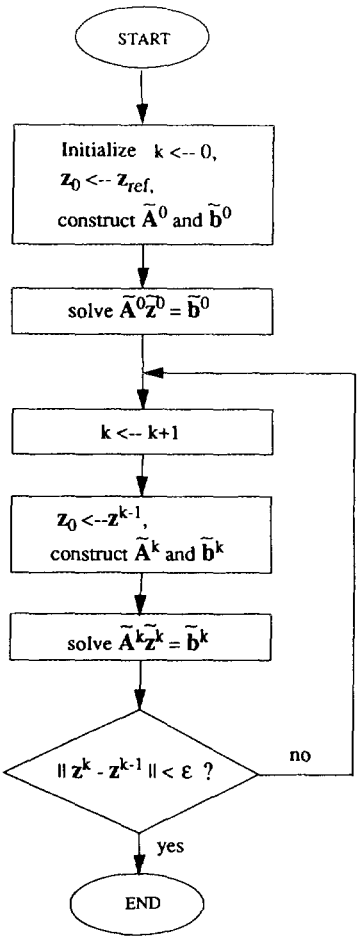


FIG. 5. Proposed perspective SFS algorithm.

gular patch determined from the previous iteration were used as the reference gradient points for the linearization of the reflectance map at the current iteration. The same scheme can be used for our perspective SFS algorithm. The only difference is that, since in the perspective SFS algorithm, the linearization is performed with respect to nodal heights Z_i directly while it is performed with respect to surface gradients (p, q) in an orthogonal SFS algorithm, explicit surface gradient (p, q) information is not needed in this case. Figure 5 shows the flow chart of a perspective SFS algorithm with the successive linearization scheme.

Under perspective projection, even for finely and regularly sampled nodal points on an image, the corresponding surface nodal points in the world plane may become sparse and irregular due to the depth effect (see (2.1)). As a result, we often obtain a smaller number of irregular surface height data points with respect to the actual size of the object. One way to increase the resolution of the final result is to perform surface interpolation using the sparse data points, which has been well studied in the context of surface reconstruction from sampled data [6, 34, 31].

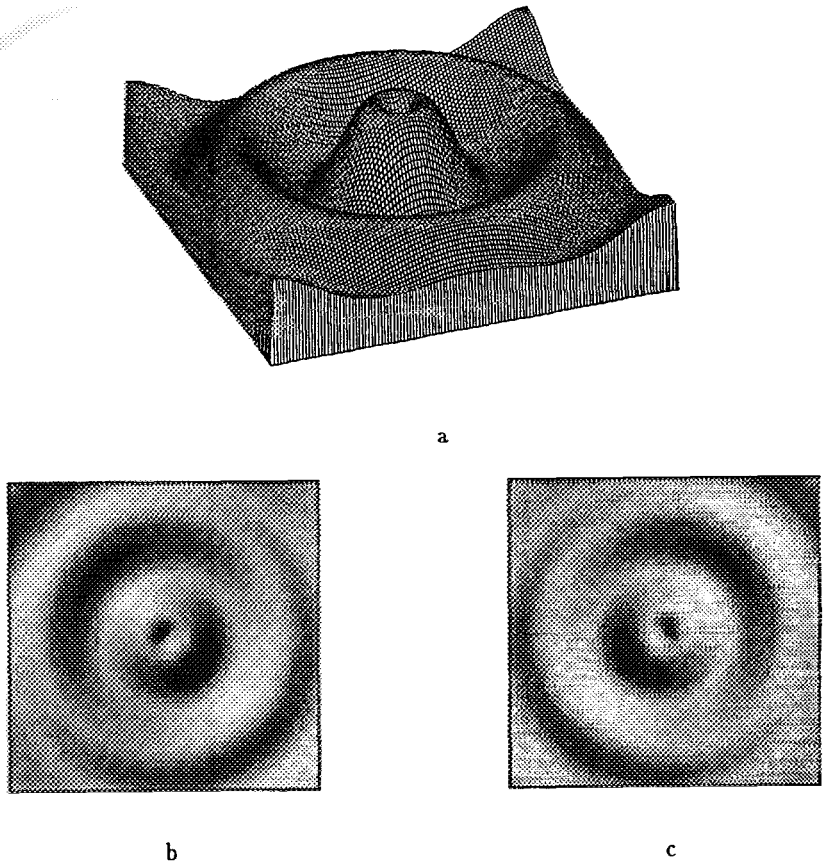
5. EXPERIMENTAL RESULTS

We have tested the proposed SFS algorithm on several synthetic and real perspective photometric stereo images. For all test images, we perform the triangulation using the finest grid points so that all nodal points coincide with image points. We assume that the height of one surface nodal point (the ground point, say $Z_1 = Z_{ref}$) is known and fixed all the time. We set $Z_i^0 = Z_{ref}$, $i = 1, \ldots, M_n$, at the first iteration. For the following iterations, the linearization of the reflectance map is performed with respect to the surface nodal height obtained from the previous iteration. We present reconstructed results based on a single image and a couple of photometric stereo images generated by light sources with orthogonal tilt angles for all the test problems. Similar to the results given in [21] with orthogonal projection, the single image case gives a much less accurate results. To display a finer and smoother reconstructed surface, interpolation based on the determined finite sparse surface nodal height data has been performed.

TEST PROBLEM 1 (Modified sombrero). The test object is a modified sombrero surface. A portion of its ground truth with 128 × 128 surface points is shown in Fig. 6(a). We set the focal length of the camera f and the reference height Z_{ref} to be 150 and −300, respectively. Figures 6(b) and (c) illustrate two 64 × 64 photometric stereo images generated with perspective projection and illuminated by a distant light source from $(\tau, \sigma) = (45°, 45°)$ and $(135°, 45°)$, respectively. Note that the two light sources have tilt angles which are orthogonal to each other. Figures 7(a) and (b) show the reconstructed surfaces from the single perspective image shown in Fig. 7(b) and the combined photometric stereo images, respectively.

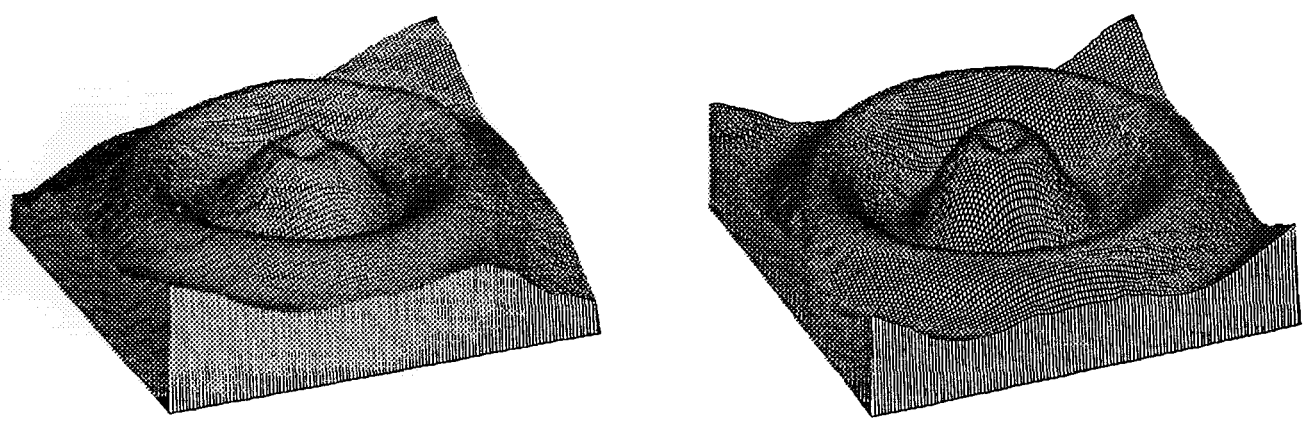
TEST PROBLEM 2 (Mozart). The test images are synthesized from the 256 × 256 Mozart statue height data obtained from the range data. We set the focal length of the camera f and the reference height Z_{ref} to be 300 and −600, respectively. The original ground truth of the surface and the two 128 × 128 shaded images with illumination directions $(\tau, \sigma) = (45°, 45°)$ and $(135°, 45°)$ are shown in Figs. 8(a)–(c). Note that since the original surface has discontinuities along the object boundaries and on several points inside the object, the resulting images are not ideal and have noises at the corresponding points. The reconstructed surfaces from a single perspective image given in Fig. 8(c) and from combined images are shown, respectively, in Figs. 9 (a) and (b). We observe some reconstruction error along the discontinuous boundary even with photometric stereo images.

TEST PROBLEM 3 (Agrippa). Two real perspective images are used for this test problem. The 128 × 128 test images of Agrippa statue obtained by illuminating from



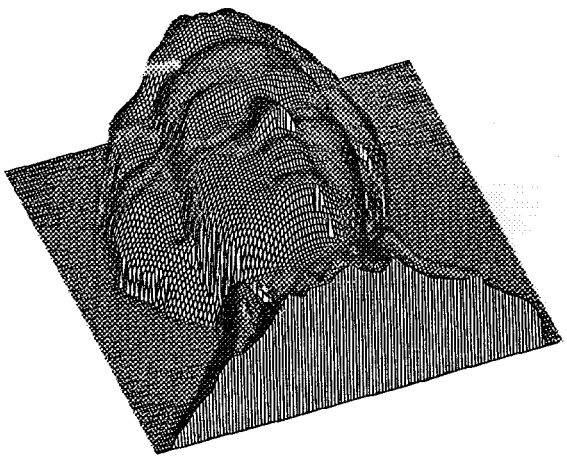
a b c

FIG. 6. The modified sombrero test problem: (a) the ground truth of the modified sombrero surface; and two synthetic images illuminated with (b) (albedo, tilt, slant) = (250, 135°, 45°) and (c) (albedo, tilt, slant) = (250, 45°, 45°).



a b

FIG. 7. Results of the perspective SFS algorithms applied to the modified sombrero images: (a) the reconstructed surface from single image in Fig. 6(c); (b) the reconstructed surface using the photometric stereo images.

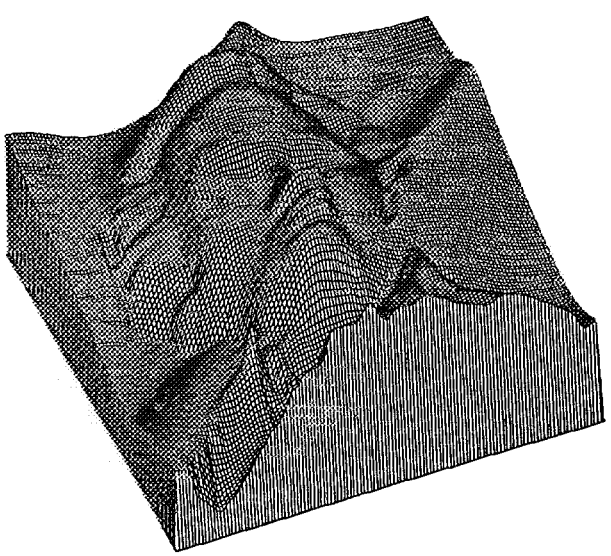


a

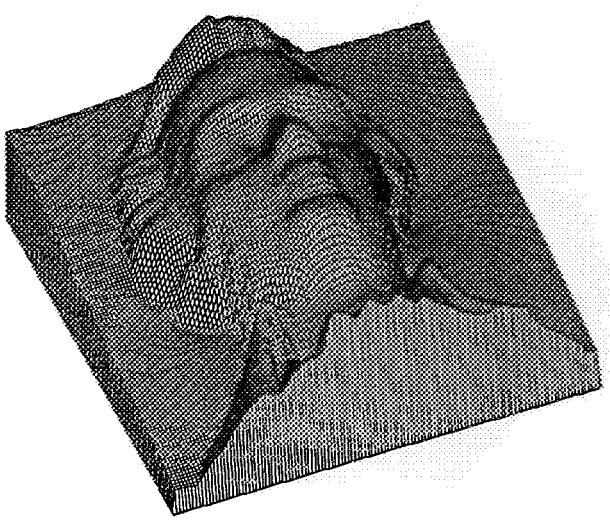
b

c

FIG. 8. The Mozart test problem: (a) the ground truth of the Mozart statue; and two synthetic images illuminated with (b) (albedo, tilt, slant) = (250, 135°, 45°) and (c) (albedo, tilt, slant) = (250, 45°, 45°).



a



b

FIG. 9. Results of the perspective SFS algorithms applied to the Mozart images: (a) the reconstructed surface from single image in Fig. 8(c); (b) the reconstructed surface using the photometric stereo images.

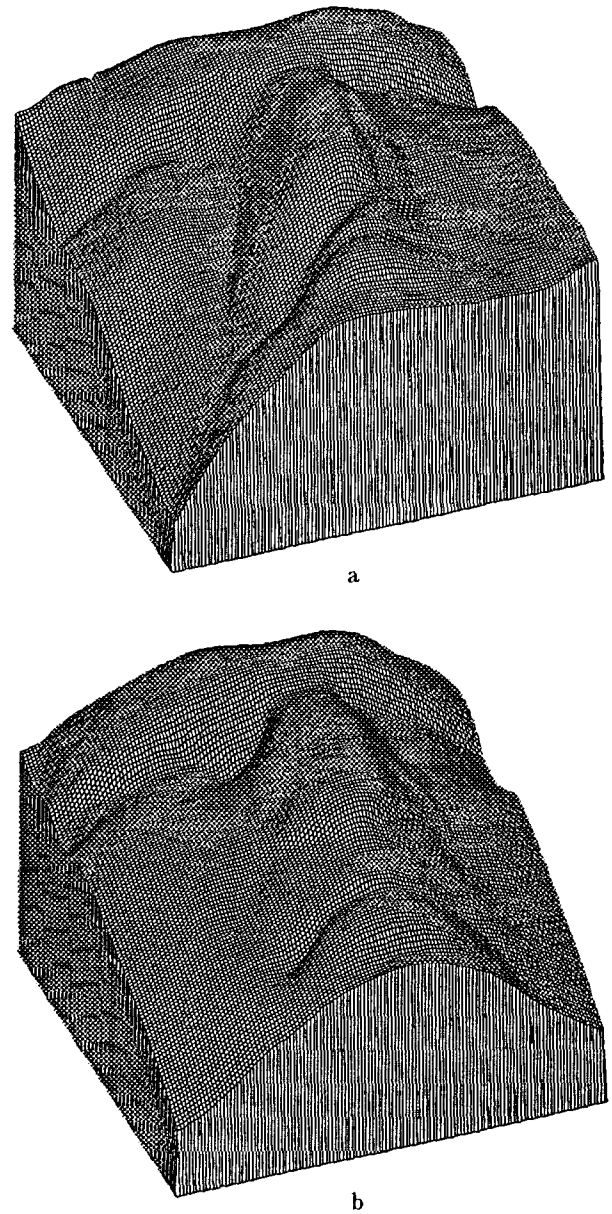
a b

FIG. 10. The Agrippa test problem: two real images of the Agrippa statue illuminated with (a) (albedo, tilt, slant) = (190, 135°, 50°); (b) (albedo, tilt, slant) = (190, 45°, 45°).

directions $(\tau, \sigma) = (45°, 45°)$ and $(135°, 50°)$ are shown in Figs. 10(a) and (b), respectively. We assumed the focal length of the camera f and the reference height Z_{ref} to be 300 and -600. Figures 11(a) and (b) show the reconstructed surface from the single image given in Fig. 10(b) and from photometric stereo images, respectively.

6. CONCLUSION

We have proposed a new algorithm for recovering shape from shaded images under perspective projection. Our algorithm is based on a triangular element surface approximation and a linearized reflectance map around a certain nodal height value. We relate image intensity of a triangle to the surface nodal heights via an image irradiance equation, and we determine surface nodal heights by minimizing the cost functional consisting of the sum of brightness error associated with each triangle. It recovers surface heights directly instead of surface orientations. No additional integrability constraint is required, and no boundary conditions and smoothness constraints are employed. The recovery algorithm refines the reconstructed surface iteratively by successive linearization scheme, in which the linearization is performed based on updated surface nodal heights obtained from the previous iteration. We have included several experimental results including both synthetic and real perspective images to demonstrate the performance of the proposed algorithm.



a b

FIG. 11. Results of the perspective SFS algorithms applied to the Agrippa images: (a) the reconstructed surface from single image in Fig. 10(b); (b) the reconstructed surface by the photometric stereo images.

REFERENCES

1. M. Bichsel and A. P. Pentland, A simple algorithm for shape from shading, in *IEEE Conference on Computer Vision and Pattern Recognition, Champaign, Illinois, June 1992*, pp. 459–465.
2. M. J. Brooks, W. Chojnacki, and R. Kozera, Impossible and ambiguous shading patterns, *Int. J. Comput. Vision* **7**(2), 1992, 119–126.
3. 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.
4. P. Dupuis and J. Oliensis, Direct method for reconstructing shape from shading, in *IEEE Conference on Computer Vision and Pattern Recognition, Champaign, Illinois, June 1992*, pp. 453–458.
5. R. T. Frankot and R. Chellappa, A method for enforcing integrability in shape from shading algorithm, *IEEE Trans. Pattern Anal. Machine Intell.* **10**(4), 1989, 439–451; in *Shape from Shading* (B. K. P. Horn and M. J. Brooks, Eds.), MIT Press, Cambridge, MA, 1989.

6. W. E. L. Grimson, An implementation of a computational theory of visual surface interpolation, *Comput. Vision Graphics Image Process.* **22,** 1983, 39–64.
7. B. K. P. Horn, *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View,* Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1970.
8. B. K. P. Horn, *Obtaining Shape from Shading Information,* MIT Press, Cambridge, MA, 1975.
9. B. K. P. Horn, *Robot Vision,* The MIT Press, Cambridge, MA, 1986.
10. B. K. P. Horn, Height and gradient from shading, *Int. J. Comput. Vision* 5, 1990, 584–595.
11. B. K. P. Horn and M. J. Brooks, The variational approach to shape from shading, *Comput. Vision Graphics Image Process.* **33,** 1986, 174–208; in *Shape from Shading* (B. K. P. Horn and M. J. Brooks, Eds.), MIT Press, Cambridge, MA, 1989.
12. B. K. P. Horn and M. J. Brooks, *Shape from Shading,* MIT Press, Cambridge, MA, 1989.
13. K. Ikeuchi, Determining surface orientations of specular surfaces by using the photometric stereo method, *IEEE Trans. Pattern Anal. Machine Intel.* **PAMI-13**(6), 1981, 661–669.
14. K. Ikeuchi, Determining a depth map using a dual photometric stereo, *Int. J. Rob. Res.* **6**(1), 1987, 15–31.
15. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding boundaries, *Artif. Intell.* **17,** 1981, 141–184; in *Shape from Shading* (B. K. P. Horn and M. J. Brooks, Eds.), MIT Press, Cambridge, MA, 1989.
16. R. Kozera, Existence and uniqueness in photometric stereo, *App. Math. Comput.* **44,** 1991, Part I, 1–103.
17. 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,* pp. 552–558.
18. D. Lee, A provably convergent algorithm for shape from shading, in *Shape from Shading* (B. Horn and M. Brooks, Eds.), pp. 349–373, MIT Press, Cambridge, MA, 1989.
19. K. M. Lee and C.-C. J. Kuo, *Shape from Shading with a Linear Triangular Element Surface Model, IEEE Trans. on Pattern Analysis and Machine Intel.* **PAMI-15**(8), 1993, 815–822.
20. K. M. Lee and C. C. J. Kuo, Robust shape reconstruction from combined shading and stereo information, in *Proceedings of SPIE Conference (Vol. 1771): Applications of Digital Image Processing XV, San Diego, CA, July 1992.*
21. K. M. Lee and C. C. J. Kuo, Surface reconstruction from photometric stereo, *J. Opt. Soc. Amer. A.* **10**(5), 1993, 855–868.
22. J. Oliensis, Uniqueness in shape from shading, *Int. J. Comput. Vision* **6**(2), 1991, 75–104.
23. R. Onn and A. Bruckstein, Integrability disambiguates surface recovery in two-image photometric stereo, *Int. J. Comput. Vision* 5(1), 1990, 105–113.
24. M. A. Penna, Local and semi-local shape from shading for a single perspective image of a smooth object, *Comput. Vision Graphics Image Process.* **46,** 1989, 346–366.
25. M. A. Penna, A shape from shading analysis for a single perspective image of a polyhedron, *IEEE Trans. Pattern Anal. Machine Intel.* **PAMI-11,** 1989, 545–554.
26. A. P. Pentland, Shape information from shading: A theory about human perception, in *Proceedings International Conf. on Computer Vision, 1988.* pp. 404–413.
27. R. Ray, J. Birk, and R. Kelly, Error analysis of surface normals determined by radiometry, *IEEE Trans. Pattern Anal. Machine Intel.* **PAMI-5**(6), 1983, 631–645.
28. M. Shao, R. Chellappa, and T. Simchony, Reconstructing a 3-D depth map from one or more images, *Comput. Vision Graphics Image Process.: Image Understanding,* **53,** 1991, 219–226.
29. W. Silver, Determining shape and reflectance using multiple images, Master's thesis, MIT, Cambridge, MA, 1980.
30. G. B. Smith, The relationship between image irradiance and surface orientation, in *IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983,* pp. 404–413.
31. R. Szeliski, Fast surface interpolation using hierarchical basis functions, *IEEE Trans. Pattern Anal. Machine Intel.* **PAMI-12**(6), 1990, 513–528.
32. R. Szeliski, Fast shape from shading, *Computer Vision Graphics Image Process.: Image Understanding* **53**(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 Anal. Machine Intel.* **PAMI-13**(2), 1991, 133–152.
34. D. Terzopoulos, Multilevel computational processes for visual surface reconstruction, *Comput. Vision Graphics Image Process.* **24,** 1983, 52–96.
35. P. S. Tsai and M. Shah, A fast linear shape from shading, in *IEEE Conference on Computer Vision and Pattern Recognition, Champaign, IL, June 1992,* pp. 734–736.
36. R. J. Woodham, A cooperative algorithm for determining surface orientation from a single view, in *International Joint Conference on Artificial Intelligence, Cambridge, MA, August 1977,* pp. 635–641.
37. R. J. Woodham, Photometric method for determining surface orientation from multiple images, *Opt. Eng.* **19**(1), 1980, 139–144.
38. R. J. Woodham, Analysing images of curved surfaces, *Artif. Intel.* **17,**(1–3), 1981, 117–140.
39. Q. Zheng and R. Chellappa, Estimation of illumination direction, albedo, and shape from shading, *IEEE Trans. Pattern Anal. Machine Intel.* **PAMI-13**(7), 1991, 680–702.