Vol. 10, No. 5/May 1993/J. Opt. Soc. Am. A 855

Surface reconstruction from photometric stereo images

Kyoung Mu Lee and C.-C. Jay Kuo

Signal and Image Processing Institute and Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564

Received February 11, 1992; revised manuscript received July 24, 1992; accepted October 21, 1992

In previous research (Tech. Rep. 172, University of Southern California, Los Angeles, Calif., 1991) we developed an iterative shape-from-shading (SFS) algorithm that uses a single image, by combining a triangular-element surface model with a linearized reflectance map. In the current research we generalize the single-image SFS algorithm to the photometric stereo SFS algorithm, which uses multiple images taken under different lighting conditions for more-accurate surface reconstruction. An explicit surface model based on nodal basis-function representation is used so that the integrability problem that arises in conventional photometric stereo SFS algorithms can be solved easily. It is also shown that single-image SFS algorithms have an inherent problem; i.e., the accuracy of the reconstructed surface height is related to the slope of the reflectance-map function defined on the gradient space. The accuracy can be greatly improved by combining two photometric images properly, and the optimal illumination condition that leads to the best shape reconstruction is examined. Simulation results for several test images are given to demonstrate the performance of our new algorithm.

1. INTRODUCTION

Extracting the surface information of an object from its single or multiple shaded image, known as the shape-from-shading (SFS) problem, is one of the fundamental problems in computer vision. The image-formation process consists of three factors: the object shape, the property of the object surface, and the illumination condition (light-source information). If the surface property and the lighting condition are known *a priori,* the shading information provides important cues for three-dimensional (3D) surface reconstruction, since the variation in brightness arises primarily from the change in surface orientation. Research on the SFS problem has been performed extensively in the past two decades.

Many SFS algorithms that use a single image have been developed, such as the characteristic strip method,[1,2] the variational method,[3–11] the Fourier method,[12] and the optimal-control approach.[13] However, the solutions are not reliable because of the lack of either a proper constraint or a proper solution method. Although it has been shown analytically in some recent research[13,14] that the single-SFS problem is not ill posed if the image contains singular points, there are no satisfactory numerical results with real test images so far. Thus researchers have considered the use of multiple images to provide additional information for robust surface reconstruction, including the photometric stereo method[15–23] and the geometric stereo method.[24–27]

The photometric stereo method was first proposed by Woodham[21–23] and has been studied and extended for practical implementation by several researchers.[15,16,18–20] With this method, one uses images taken from the same viewing direction under different lighting conditions. The surface orientation of a local point is determined by its irradiances in these images, which are determined by using the fact that the orientation corresponds to the intersection of constant brightness contours of different reflectance maps defined on the gradient space. The conventional scheme usually requires an additional integration step for constructing a surface with the orientations obtained.

Methods of determining the depth map with the use of geometric stereo and shading information have also been studied. Grimson[24] proposed a method of combining binocular shading information with stereo data to determine the orientation of the surface normal along feature-point contours as well as the reflectance-map parameters. The method uses a general reflectance-map model consisting of both diffuse and specular components. The method is, however, numerically unstable and practically unreliable. Ikeuchi[26] considered another method for constructing the depth map from dual photometric stereo images by combining obtained surface-orientation maps by means of camera geometry transformation and region matching. The result depends heavily on the accuracy of the surface orientations obtained by the conventional photometric stereo algorithm and is therefore sensitive to noise.

In this research we provide a unifying approach to solve both single-image and photometric stereo SFS problems. Our new approach is based on the single-image SFS algorithm proposed in Ref. 28, in which an explicit surface model called the triangular-element surface model and the linear reflectance-map approximation were used. The basic idea can be simply stated as follows. We express the surface height and orientation by using a nodal basis-function representation so that the image brightness on each triangular patch is related directly to the nodal height by means of the reflectance map. The proposed photometric stereo SFS algorithm combines the information of all images simultaneously to recover the underlying surface height. It is formulated as a quadratic functional minimization problem parameterized by surface heights, in which the cost functional is the squares of the brightness error. The surface heights can be obtained by solving the equivalent large, sparse linear

0740-3232/93/050855-14$05.00

system of equations with efficient linear system solvers, such as the multigrid and the preconditioned conjugate gradient methods.

Compared with the conventional photometric stereo method, our new method has two major advantages. First, the new method determines surface heights directly, while the conventional photometric stereo method determines only the surface orientations. Thus the integrability problem arises in the conventional method but not in ours. Second, the new method is a global method that minimizes the squared intensity errors over all points so that it is insensitive to noise. In contrast, the conventional method is a local one and is more sensitive to noise. By examining the characteristics of the reflectance map, we also show that the single-image SFS algorithm may not yield an accurate result, and we explain how to use the photometric stereo information to improve the accuracy of the reconstructed surface. We also discuss the optimal illumination condition that leads to the best shape reconstruction.

2. CONVENTIONAL PHOTOMETRIC STEREO

Under the assumption of orthographic projection, Lambertian surface, and a distant single-point light source, the reflectance map is basically a function that characterizes the relationship between the image irradiance and the orientation of the object surface. It can be derived that

$$R(p,q) = \begin{cases} \eta \mathbf{l}^T\mathbf{n} & \mathbf{l}^T\mathbf{n} \geq 0 \\ 0 & \mathbf{l}^T\mathbf{n} < 0 \end{cases}, \tag{2.1}$$

where η is the albedo of the surface,

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

is the surface normal, p and q represent the derivatives of surface height z, w, r, t, x, and y,

$$\mathbf{l} = (\cos\tau \sin\sigma, \sin\tau \sin\sigma, \cos\sigma)^T$$

is the unit vector of the illumination direction pointing toward the light source, and τ and σ are the tilt and the slant angles that the illumination direction makes with the x and the z axes, respectively. This can also be represented as

$$\mathbf{l} = \frac{(-p_s, -q_s, 1)^T}{(1 + p_s^2 + q_s^2)^{1/2}},$$

where p_s and q_s denote the slope of a surface element perpendicular to the illumination direction. Substituting the vectors $\mathbf{n}$ and $\mathbf{l}$ into Eq. (2.1), we obtain

$$R(p,q) = \begin{cases} \eta K/(1 + p^2 + q^2)^{1/2} & K \geq 0 \\ 0 & K < 0 \end{cases}, \tag{2.2a}$$

where

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

or, equivalently,

$$R(p,q) = \begin{cases} \eta(1 + p_s p + q_s q)/[(1 + p^2 + q^2)(1 + p_s^2 + q_s^2)]^{1/2} & 1 + p_s p + q_s q \geq 0 \\ 0 & 1 + p_s p + q_s q < 0 \end{cases}. \tag{2.2b}$$

The reflectance map $R(p,q)$ is a nonlinear function that can be depicted as nested contours in the gradient space (p,q). The basic equation for the image-formation process can therefore be expressed as

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

which is known as the image-irradiance equation.

Conventional photometric stereo is an algebraic method for solving the image-irradiance equation. With given albedo η, illumination direction $\mathbf{l}$, and image irradiance E, there are two unknown variables, p and q, in Eq. (2.3) so that one needs at least two equations to determine the values of p and q. However, since Eq. (2.3) is nonlinear, it may have more than one solution, and a third image is often needed for formation of an overdetermined system. Based on the reflectance map [Eq. (2.1)], we can provide a simple viewpoint for understanding the solution procedure. Suppose that we have three images obtained with illumination directions $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$. At a given point (x,y) the observed image irradiances in these three images are E_1, E_2, and E_3, respectively. Thus we have three image-irradiance equations,

$$E_i = \eta \mathbf{l}_i^T\mathbf{n}, \qquad i = 1, 2, 3,$$

where $\mathbf{n}$ is the surface normal at the point (x,y). These equations can be written in matrix form as

$$\mathbf{E} = \eta\mathbf{L}\mathbf{n}, \tag{2.4}$$

where

$$\mathbf{E} = \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} \mathbf{l}_1^T \\ \mathbf{l}_2^T \\ \mathbf{l}_3^T \end{bmatrix}.$$

If $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$ are linearly independent, $\mathbf{L}^{-1}$ exists, and Eq. (2.4) can be solved for $\mathbf{n}$. However, note that the surface normal $\mathbf{n}$ is a unit vector with two free variables. Thus Eq. (2.4) is an overdetermined system, and we may obtain its least-squares solution.

One feature of the above formulation is that it applies nicely if the albedo, η, is not known *a priori* and is varying at different image points. That is, we know from Eq. (2.4) that

$$\eta\mathbf{n} = \mathbf{L}^{-1}\mathbf{E}. \tag{2.5}$$

The magnitude of the right-hand side of Eq. (2.5) gives the value of the albedo at (x,y),

$$\eta = \|\mathbf{L}^{-1}\mathbf{E}\|,$$

and the corresponding unit surface normal $\mathbf{n}$ is

$$\mathbf{n} = (1/\eta)\mathbf{L}^{-1}\mathbf{E}.$$

Since the surface orientation instead of the surface height is determined locally by this method, one has to integrate the orientations for surface-height construction. Because the surface orientations are usually not consistent, there may be no surface appropriate for surface-height construction. This is known as the integrability problem.[3,4]

3. DIRECT SURFACE-HEIGHT RECOVERY FROM PHOTOMETRIC STEREO WITH TRIANGULAR-ELEMENT SURFACE MODEL

A. Nodal Basis-Function Representation of Surface Height

Methods of determining the surface height directly from shading information have been reported recently in the literature.[4,28,29] One can remove the integrability problem naturally by introducing a nodal basis-function representation of the surface height. Consider an arbitrary nodal basis function, $\psi_i(x, y)$, whose value is equal to 1 at grid point i and is 0 at other grid points $j \neq i$. Let z_i be the nodal height at point i. Then one surface that interpolates all grid points can be written as

$$z(x, y) = \sum_{i=1}^{M_n} z_i \psi_i(x, y),$$

where M_n is the number of nodal basis functions, and the surface orientation (or normal) can be computed as

$$p(x, y) = \frac{\partial z(x, y)}{\partial x} = \sum_{i=1}^{M_n} z_i \frac{\partial \psi_i(x, y)}{\partial x}, \tag{3.1}$$

$$q(x, y) = \frac{\partial z(x, y)}{\partial y} = \sum_{i=1}^{M_n} z_i \frac{\partial \psi_i(x, y)}{\partial y}. \tag{3.2}$$

Leclerc and Bobick[29] used the first-order numerical derivatives to represent surface orientation in terms of the following discrete forms:

$$p_{i,j} = \tfrac{1}{2}(z_{i+1,j} - z_{i-1,j}),$$

$$q_{i,j} = \tfrac{1}{2}(z_{i,j+1} - z_{i,j-1}).$$

On the one hand, it can be viewed as a special case of the above general approach with an appropriately chosen basis function, $\psi_i(x, y)$. On the other hand, one can derive the relationship by simply considering one-dimensional central differencing of the height variables. The second viewpoint is mainly a numerical technique and is little related to the underlying physical surface model.

A more physically plausible approach[28] is as follows. Consider the approximation of a smooth surface with a union of triangular surface patches called triangular elements over a uniform grid domain, as illustrated in Fig. 1.[30,31] Here $\phi_i(x, y)$ denotes this particular choice of the nodal basis function $\psi_i(x, y)$. Note that $\phi_i(x, y)$ has local compact support, and its value at an arbitrary point (x, y) is obtained by linearly interpolating its three neighboring nodal heights, as shown in Fig. 2.

B. Image Formation on the Modeled Surface

To remove the nonlinearity of the reflectance-map function given by Eq. (2.2a), we take the Taylor series expansion of $R(p, q)$ about a certain reference point, (p_0, q_0), through the first-order term; i.e.,

$$R(p, q) \approx R(p_0, q_0) + (p - p_0) \left.\frac{\partial R(p, q)}{\partial p}\right|_{(p_0, q_0)} + (q - q_0) \left.\frac{\partial R(p, q)}{\partial q}\right|_{(p_0, q_0)}. \tag{3.3}$$

The reference point (p_0, q_0) can be either fixed or varying for different values of (p, q).[28] Combining the triangular surface model with the linearized reflectance-map model given by expression (3.3), we can express the image irradiance on a triangular surface patch directly in terms of nodal heights of triangular elements. By substituting Eqs. (3.1) and (3.2) into expression (3.3), we have

$$E = R(p, q) \approx \alpha p + \beta p + \gamma = \sum_{i=1}^{M_n} \Phi_i z_i + \gamma, \tag{3.4}$$

where

$$\Phi_i(x, y) = \alpha \frac{\partial \phi_i(x, y)}{\partial x} + \beta \frac{\partial \phi_i(x, y)}{\partial y},$$

$$\gamma = R(p_0, q_0) - \alpha p_0 - \beta q_0, \tag{3.4a}$$

$$\alpha = \left.\frac{\partial R(p, q)}{\partial p}\right|_{(p_0, q_0)}, \qquad \beta = \left.\frac{\partial R(p, q)}{\partial q}\right|_{(p_0, q_0)}. \tag{3.4b}$$

Therefore a linear relationship between image irradiance E and nodal height z_i is established.

C. Photometric Stereo SFS Algorithm

For estimating the nodal height z_i based on given J different photometric stereo images E_{oj} with their corresponding

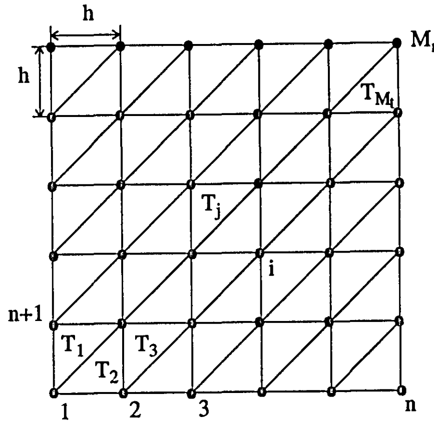


Fig. 1. Uniform triangulation of a square domain, Ω.

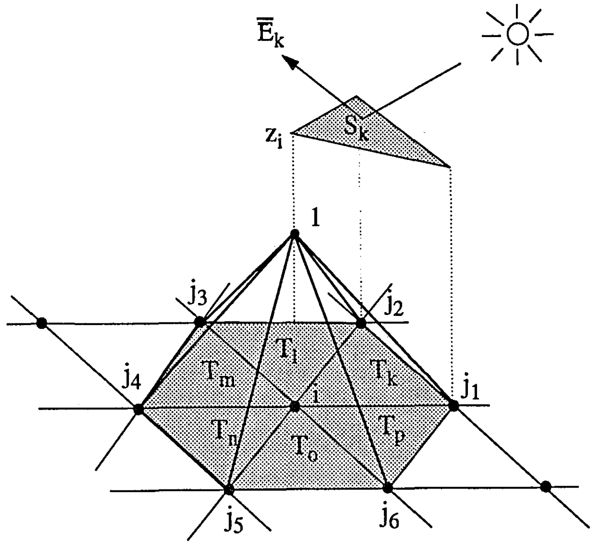


Fig. 2. Nodal basis function, ϕ_i.

reflectance maps $R_j(p, q), j = 1, \ldots, J$, a natural scheme is to minimize the following cost functional over an image domain, Ω, which is divided into a set of nonoverlapping triangles as shown in Fig. 1:

$$\mathscr{E}_b = \iint_\Omega \sum_{j=1}^{J} (E_{oj} - E_j)^2 \mathrm{d}x\mathrm{d}y, \tag{3.5}$$

where E_{oj} and E_j are the jth observed and parameterized images, respectively. By substituting Eq. (3.4) into Eq. (3.5), we obtain

$$\mathscr{E}_b = \sum_{j=1}^{J} \iint_\Omega \left[E_{oj} - \left(\sum_{i=1}^{M_n} \Phi_{ji} z_i + \gamma_j \right) \right]^2 \mathrm{d}x\mathrm{d}y,$$

where Φ_{ji} denote the function Φ_i in Eq. (3.4a) for the jth image. After some manipulation[28] we can rewrite the cost functional as

$$\mathscr{E}_b = \tfrac{1}{2} \mathbf{z}^T \tilde{\mathbf{A}} \mathbf{z} - \tilde{\mathbf{b}}^T \mathbf{z} + c, \tag{3.6}$$

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

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

and where the individual stiffness matrix and load vector can be computed as

$$[\mathbf{A}_j]_{m,n} = 2 \iint_\Omega \Phi_{jm} \Phi_{jn} \mathrm{d}x\mathrm{d}y,$$

$$[\mathbf{b}_j]_m = 2 \iint_\Omega (E_{oj} - \gamma_j) \Phi_{jm} \mathrm{d}x\mathrm{d}y,$$

$$1 \le m, n \le M_n. \tag{3.7}$$

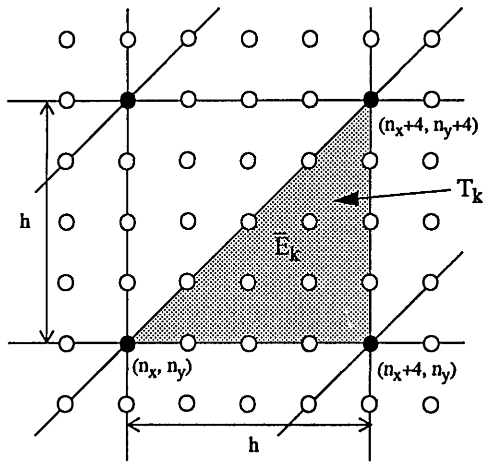
The minimization problem [Eq. (3.6)] is equivalent to the solution of the linear system of equations,

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

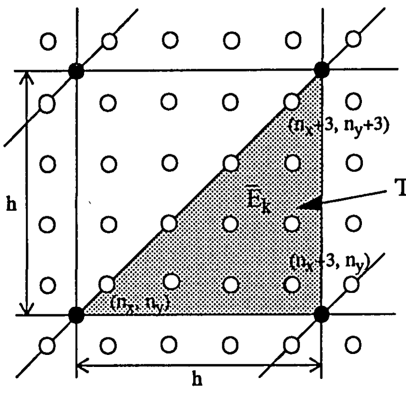
An efficient iterative linear-system solver, such as the multigrid method and the preconditioned-conjugate-gradient method, can be applied for its solution. To obtain a more-accurate reconstructed surface, we apply a successive linearization scheme to the reflectance map.[28] That is, we linearize the reflectance map with respect to the local gradient point of the triangular patch obtained from the previous iteration and perform the above solution procedure repeatedly. Note also that the smoothness constraint is not imposed in the above formulation. For the single-image SFS algorithm, a smoothness term is often incorporated to ensure the nonsingularity of the stiffness matrix (sufficient condition for the unique minimum), and the weighting of the smoothness term can be gradually reduced as iteration continues.[28] However, for the photometric stereo case, the overall stiffness matrix, $\tilde{\mathbf{A}}$, is the sum of individual stiffness matrices, where the singularity is removed by proper combination of photometric stereo images so that the smoothness constraint is not required.

D. Triangulation of an Image

Note that the construction of each load vector $\mathbf{b}_j$ in Eqs. (3.7) requires the average intensity $\overline{E}_{jk}$ over each triangular domain T_k, $k = 1, \ldots, M_t$ for the jth image.[28] To determine the average intensities, we first have to partition discrete image pixels so that they are contained by the triangular domains. For the convenience of notation, let us drop the subscript and consider one image, E. Two possible partitioning schemes for the $h = 4$ spacing between adjacent nodal points are depicted in Fig. 3, where the two-dimensional subscript notation is used for convenience. Figure 3(a) shows the first scheme, in which nodal points belong to a subset of image points; Fig. 3(b) shows the second scheme, in which the nodal points are located between the image points. The average intensity over a triangular domain can be obtained by summing all the intensity values inside the triangle and some fractions of intensity values for nodes lying on the boundary of the triangle and then dividing the sum by the area of the tri-



(a)



(b)

Fig. 3. Triangulation schemes of a discrete image. Open circles, image pixels; filled circles, nodal points.

angle. The average intensities, $\overline{E}_k$, for schemes given in Figs. 3(a) and 3(b), respectively, can be computed as

$$\begin{aligned}\overline{E}_k = {} & [(E_{n_x+2,n_y+1} + E_{n_x+3,n_y+1} + E_{n_x+3,n_y+2}) \\ & + \tfrac{1}{2}(E_{n_x+1,n_y+1} + E_{n_x+2,n_y+2} + E_{n_x+3,n_y+3} + E_{n_x+1,n_y} \\ & + E_{n_x+2,n_y} + E_{n_x+3,n_y} + E_{n_x+4,n_y+1} + E_{n_x+4,n_y+2} \\ & + E_{n_x+4,n_y+3}) + \tfrac{1}{8}(E_{n_x,n_y} + 2E_{n_x+4,n_y} \\ & + E_{n_x+4,n_y+4})]/(h^2/2),\end{aligned}$$

$$\begin{aligned}\overline{E}_k = {} & [(E_{n_x+1,n_y} + E_{n_x+2,n_y} + E_{n_x+3,n_y} + E_{n_x+2,n_y+1} \\ & + E_{n_x+3,n_y+1} + E_{n_x+3,n_y+2}) + \tfrac{1}{2}(E_{n_x,n_y} + E_{n_x+1,n_y+1} \\ & + E_{n_x+2,n_y+2} + E_{n_x+3,n_y+3})]/(h^2/2),\end{aligned}$$

where $h^2/2 = 8$ is the area of every triangle. It turns out that these two schemes give similar results. In all experiments reported in Section 5 below, the first scheme with $h = 1$ is used to triangulate the input image.

It is worthwhile to point out that two different triangulation schemes can also be made by choosing two different directions of the oblique lines, i.e., 45° and 135°. When the image domain is triangulated with the 45° oblique lines, as shown in Fig. 1, there exists a directional preference along the 45° direction, since only six nodal points among the eight nearest neighbors are associated with the central nodal point. The triangulation scheme sometimes may affect the results of the algorithm under various illumination directions. For example, for a spherically symmetric object illuminated by light sources with tilt angles 45° and 135°, the reconstructed surfaces may be slightly different as a result of different triangulations. To avoid the directional preference, we may use two triangulation schemes simultaneously so that individual stiffness matrix $\mathbf{A}_j$ and load vector $\mathbf{b}_j$ can be computed by means of

$$\mathbf{A}_j = \tfrac{1}{2}(\mathbf{A}_{j,r} + \mathbf{A}_{j,l}), \qquad \mathbf{b}_j = \tfrac{1}{2}(\mathbf{b}_{j,r} + \mathbf{b}_{j,l}),$$

where $\mathbf{A}_{j,r}$, $\mathbf{b}_{j,r}$ and $\mathbf{A}_{j,l}$, $\mathbf{b}_{j,l}$ denote the jth stiffness matrices and load vectors with 45° and 135° triangulation, respectively. However, the combination of two triangulation schemes requires extra computational cost. This is especially true for the multigrid method, in which the stiffness matrices have to be computed at every coarse grid level. Even though the single-triangulation scheme is simpler, numerical experiments show that it often gives satisfactory results.

E. Surface-Interpolation Technique

We may use a lower number of nodal basis functions than the number of pixels in the observed image for computational efficiency and convergence of the algorithm and then perform surface interpolation based on computed nodal heights to increase the resolution of the final result. The surface-interpolation technique has been well studied in the context of surface reconstruction from stereo images.[9,32–34] One well-known scheme is the variational spline-fitting algorithm.[32–34] Let $\tilde{z}_i$ be the desired height at point i, and let z_i be the height computed through successive linearization if i happens to be a nodal point, or let z_i be 0 if i is a point to be interpolated. We order first the nodal points with $i = 1, \ldots, M_n$ and then the interpolated points $i = M_n + 1, \ldots, M_i$, where M_i is the total number of nodal and interpolated points. Then the algorithm minimizes the cost functional

$$\tilde{\mathscr{E}}_i = \tilde{\mathscr{E}}_d + \tilde{\lambda}\tilde{\mathscr{E}}_s, \tag{3.8}$$

where

$$\tilde{\mathscr{E}}_d = \tfrac{1}{2}\sum_{i=1}^{M_n}(\tilde{z}_i - z_i)^2,$$

$$\tilde{\mathscr{E}}_s = \tfrac{1}{2}\tilde{\mathbf{z}}^T\tilde{\mathbf{B}}\tilde{\mathbf{z}}.$$

$\tilde{\mathbf{B}}$ is the smoothing matrix characterized by the local stencil:

$$\tilde{\mathbf{B}} : \frac{1}{h_i^2}\left[\begin{array}{ccccc} & & 1 & & \\ & 2 & -8 & 2 & \\ 1 & -8 & -8 & -8 & 1 \\ & 2 & -8 & 2 & \\ & & 1 & & \end{array}\right],$$

where h_i denotes the spacing between adjacent interpolated points. Some special operator stencils for nodal points near the boundary are given in Fig. 4.

The cost functional, Eq. (3.8), can be expressed in matrix form as

$$\tilde{\mathscr{E}}_i = \tfrac{1}{2}\tilde{\mathbf{z}}^T\tilde{\mathbf{C}}\tilde{\mathbf{z}} - \tilde{\mathbf{b}}^T\tilde{\mathbf{z}} + \tilde{c}, \tag{3.9}$$

where

$$\tilde{\mathbf{C}} = \tilde{\mathbf{D}} + \tilde{\lambda}\tilde{\mathbf{B}}$$

and where $\tilde{\mathbf{D}}$ is a diagonal matrix with element 1 for the row corresponding to a known data point and 0 otherwise and $\tilde{\mathbf{b}}$ is the zero-padded vector whose element is z_i for node i. Problem (3.9) can also be solved efficiently by the multigrid method[33,35] or by the preconditioned-conjugate-gradient method with the hierarchical basis preconditioner.[9]

Our numerical experience is that the SFS algorithm with surface interpolation often leads to a surface that is much smoother than the original one. Thus, in order to preserve the quality of the reconstructed surface, one should use $h = 1$ at the finest level resolution. However, the surface-interpolation scheme may be applicable to an image that has large, smooth surfaces and is triangulated with nonuniform patches.

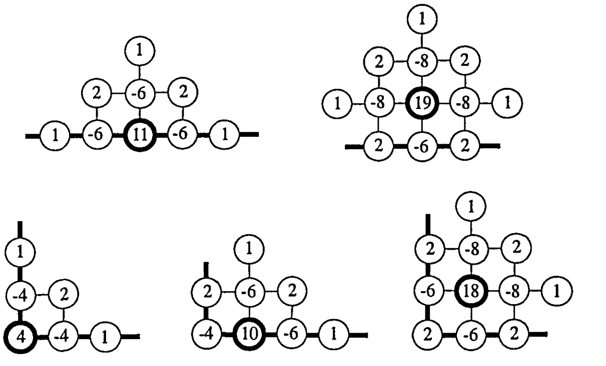


Fig. 4. Stencil forms of the nodal operators of **B** near the boundary.

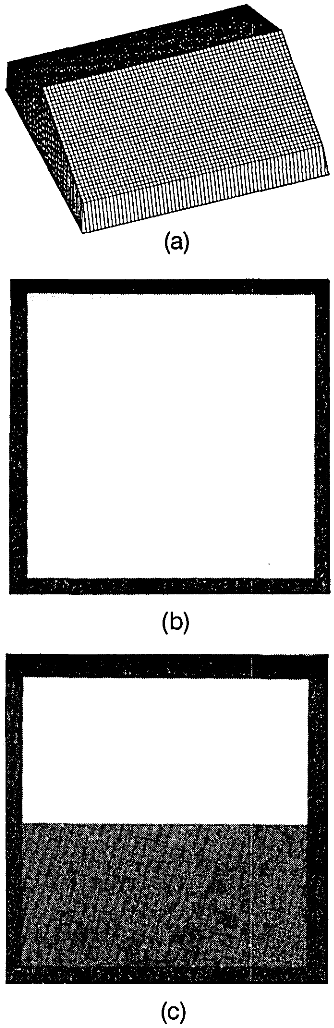


Fig. 5. Example of a roof surface: (a) 3D height plot; (b), (c) synthesized images with (albedo, tilt, slant) = (230, 0°, 45°) and (230, 90°, 45°), respectively.

4. RELATIONSHIP BETWEEN THE PHOTOMETRIC STEREO AND THE SINGLE-IMAGE SFS ALGORITHM

If $J = 1$, the above photometric stereo SFS algorithm reduces to the single-image SFS algorithm given in Ref. 28. In this section we explain why photometric stereo images produce a more accurate surface-reconstruction result than does the single-image case.

A. Limitation of Single-Image SFS Algorithm

The single-image SFS algorithm is limited even if the exact lighting condition and surface reflectivity are known. One extreme case is that the 3D surface information may be totally lost under a certain lighting condition, and thus there is no way to recover the surface orientation. For example, Fig. 5(a) shows a 3D height plot of a roof surface consisting of two planes that have slopes of the same magnitude but different signs. Figures 5(b) and 5(c) are the synthesized images of Fig. 5(a) based on Eqs. (2.1) and (2.3) with illuminating directions (tilt, slant) = (0°, 45°) and (90°, 45°), respectively. Since there is no intensity change in Fig. 5(b), it is natural to conclude that the surface is a plane. In contrast, Fig. 5(c) contains enough information for us to perform the exact 3D shape reconstruction.

To understand the problem better, we may study the characteristics of the reflectance map as follows. For simplicity, let us fix the value $q = q_0$ and view the reflectance map, $R(p, q_0)$, as a function of one variable, p. The corresponding irradiance equation becomes

$$E = R(p, q_o). \tag{4.1}$$

The sensitivity of p with respect to the change in E can be estimated by means of

$$\left|\frac{\Delta p}{\Delta E}\right| = \left|\left[\frac{\partial R(p, q_0)}{\partial p}\right]^{-1}\right|, \tag{4.2}$$

which is inversely proportional to the slope of the reflectance map at point p. Thus, for a fixed value of ΔE, the estimate $\hat{p}$ is more accurate (i.e., the value of Δp is smaller) for the region where $R(p, q_0)$ is steeper. Similar arguments can be made along the q direction; i.e.,

$$\left|\frac{\Delta q}{\Delta E}\right| = \left|\left[\frac{\partial R(p_0, q)}{\partial q}\right]^{-1}\right|. \tag{4.3}$$

The contour plots of two typical reflectance maps are given in Figs. 6(a) and 6(b). They are skewed along the line passing through (0, 0) and (p_s, q_s). If the spacings between the adjacent contour lines are narrower (or wider), the slopes of the reflectance map $R(p, q)$ are steeper (or smoother), or, equivalently, the partial derivatives $R_p(p, q)$ and $R_q(p, q)$ of the reflectance map have larger (or smaller) absolute values. For a given point (p_0, q_0) in the gradient space, the sensitivity defined in Eqs. (4.2) and (4.3) is highest along the gradient direction; i.e.,

$$\nabla R(p_0, q_0) = [R_p(p_0, q_0), R_q(p_0, q_0)],$$

and lowest along the tangential direction, i.e., $[-R_q(p_0, q_0), R_p(p_0, q_0)]$. In practice, real images contain noise, such as the sensor or the quantization noise. Besides, the irradiance values are averaged to obtain a uniform brightness value for each triangular patch for the algorithm given in Section 3. The averaging corresponds to a linear interpolation for the nonlinear reflectance map, which results in a modeling error. Since the image is corrupted by these types of noise, it is relatively difficult to obtain accurate surface orientations or heights in the region where the slope of the reflectance map is smooth.

The same phenomenon can be explained from another viewpoint. The components of the stiffness matrix are determined primarily by the partial derivatives of the reflectance map with respect to p and q, as given by Eqs. (3.4) and (3.7). Small values of α or β cause some elements of the stiffness matrix to be nearly zero and make the problem ill conditioned so that the height or the orientation related to those elements cannot be easily determined. We observe that smaller values of the partial derivatives $R_p(p, q)$ and $R_q(p, q)$ often slow down the convergence rate of the algorithm. The slow convergence behavior is attributed to the large condition number of the original nonlinear minimization problem [Eq. (3.5)].

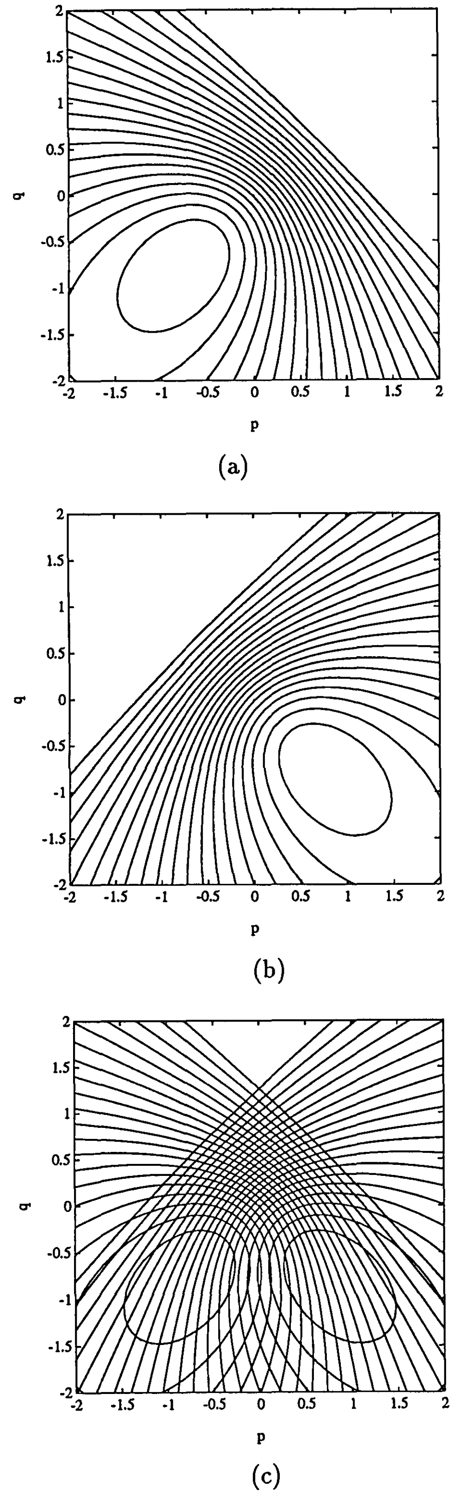


Fig. 6. Contour plots of two reflectance maps with (a) (albedo, tilt, slant) = (250, 45°, 45°) or $(p_s, q_s) = (-0.707, -0.707)$ and (b) (albedo, tilt, slant) = (250, 135°, 45°) or $(p_s, q_s) = (0.707, -0.707)$, (c) combined reflectance map of (a) and (b).

The following example is used to illustrate the above discussion. In Fig. 7 we plot (a) the original heights of a portion of a sphere, (b) the corresponding distribution of the gradients (p, q) of the triangular surface patches, and (c) the synthesized image from the illumination direction (tilt, slant) = (45°, 45°), with albedo = 250. In Fig. 8 we show (a), (b) the reconstructed surface and the corresponding (p, q) plot, respectively, after 1 iteration and (c), (d) after 20 successive iterations. The (p, q) value of all points is set to zero initially. By comparing the (p, q) distributions in Figs. 8(b) and 8(d) with the original distribution in Fig. 7(b), we can easily see that the gradient points spread quickly along the steepest descent direction of the reflectance map and reach quite accurate values in one iteration. In contrast, the gradient points move slowly along the tangential direction of the contour lines and have not yet reached satisfying values even after 20 iterations.

B. Optimal Lighting Condition

The difficulties resulting from a single reflectance map can be overcome by using photometric stereo images, since they provide several reflectance-map functions that enhance the sensitivity of Δp and Δq with respect to ΔE over the gradient domain of interest. It is desirable to

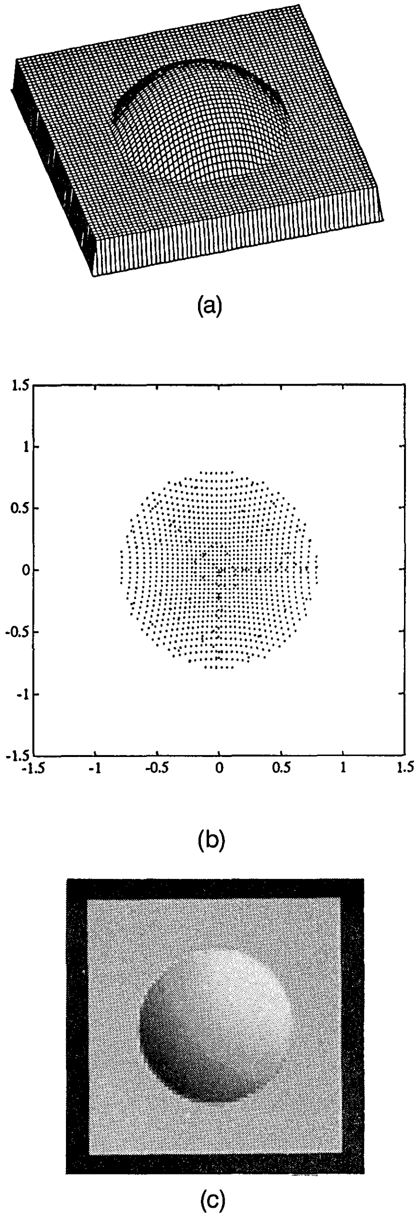
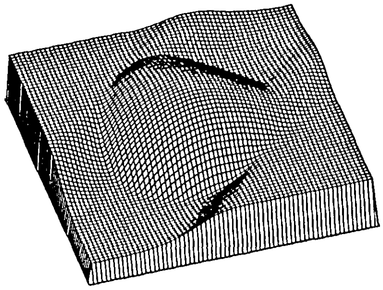
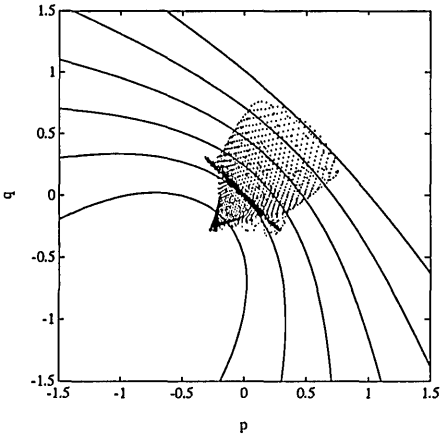


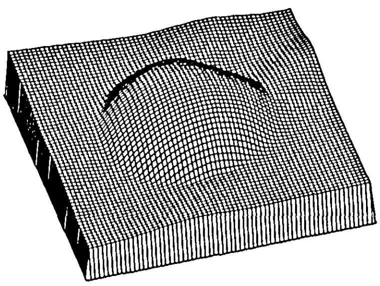
Fig. 7. Example of a sphere surface: (a) ground truth, (b) (p, q) distribution in the gradient space, (c) synthesized image with (albedo, tilt, slant) = (250, 45°, 45°).



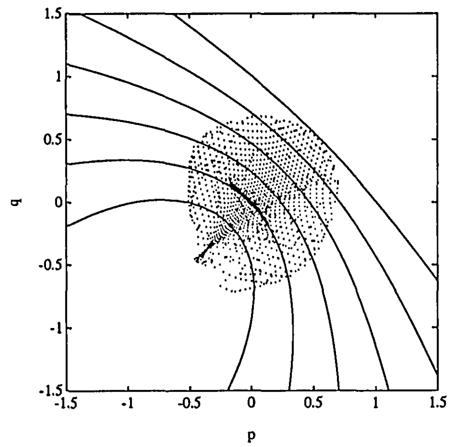
(a)



(b)

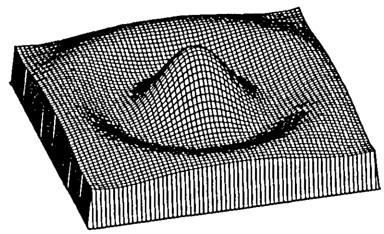


(c)

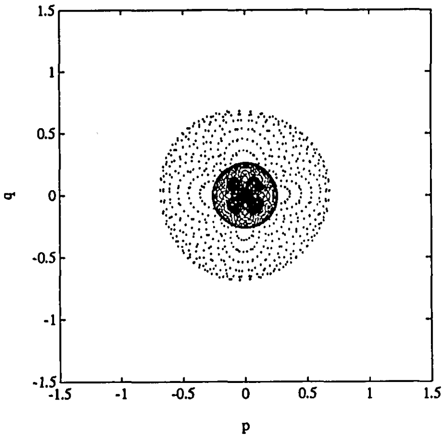


(d)

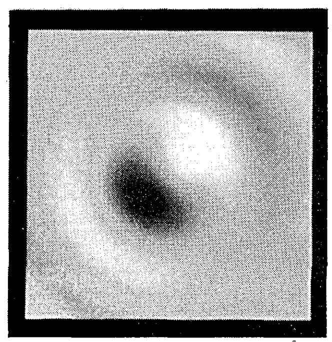
Fig. 8. Results of application of the single-image SFS algorithm to the sphere image of Fig. 7(c): (a), (b) reconstructed height and corresponding (p, q) plot, respectively, after one iteration; (c), (d) reconstructed height and corresponding (p, q) plot, respectively, after 20 successive iterations.



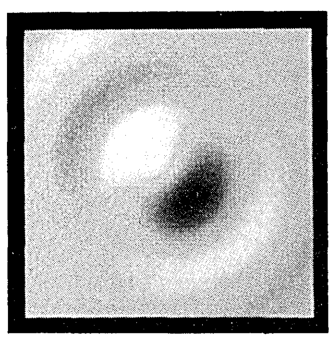
(a)



(b)



(c)



(d)

Fig. 9. Sombrero test problem: (a) ground truth of the sombrero surface, (b) (p, q) distribution in the gradient space, (c) synthesized image with (albedo, tilt, slant) = (250, 45°, 45°), (d) synthesized image with (albedo, tilt, slant) = (250, 135°, 45°).

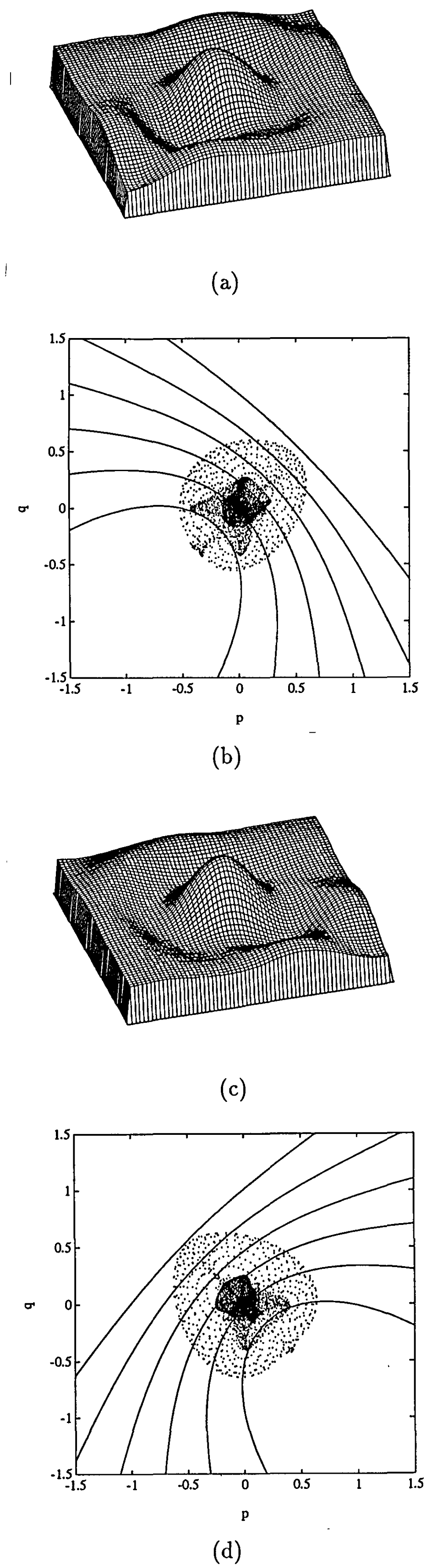


Fig. 10. Results of application of the single-image SFS algorithm to the sombrero images: (a), (b) reconstructed height and corresponding (p, q) plot, respectively, based on the image of Fig. 9(c); (c), (d) reconstructed height and corresponding (p, q) plot, respectively, based on the image of Fig. 9(d).

incorporate reflectance maps that compensate for one another's weaknesses. It is easy to see that the tilt angle determines the orientation of the reflectance map around the origin, whereas the slant angle determines the distance between the origin and (p_s, q_s) in the gradient space as well as the shape of the reflectance map. Therefore the angular distance between the two reflectance maps with tilt angles τ_1 and τ_2 is simply $|\tau_1 - \tau_2|$. If the slant angle is in the range between 30° and 60°, the reflectance map covers the central region of the gradient space, which is our main concern, and has appropriate values in steepness. Thus the optimal lighting condition is dependent primarily on the tilt angles of different light sources and is not sensitive to the slant angles as long as they are between 30° and 60°.

We know from the discussion in Subsection 4.A that the reflectance map provides good sensitivity along the gradient direction but poor sensitivity along the tangential direction. Consider two photometric stereo images illuminated from the same slant angle but from different tilt angles. Ideally the gradient directions of one reflectance map correspond to the tangential directions of the other reflectance map over the region of interest. We can achieve this by choosing the difference of their tilt angles to be 90°. One such example is given by Fig. 6(c), where the contour plots of two reflectance maps are shown together. The tilt angles are 45° and 135°, while the slant angle (= 45°) and the albedo (= 250) are fixed. Note that the gradients of a smooth surface are usually concentrated on the central region of the gradient space, say, $-0.5 < p, q < 0.5$. It is clear from these two figures that

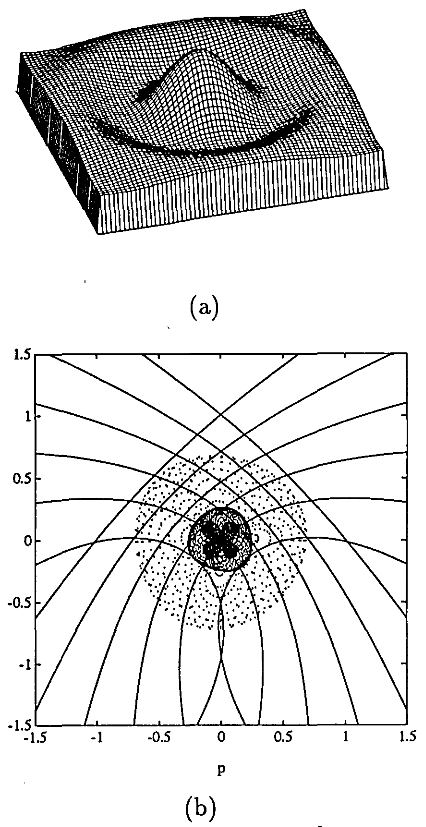


Fig. 11. Results of application of the photometric stereo SFS algorithm to the sombrero images: (a), (b) reconstructed height and corresponding (p, q) plot, respectively.

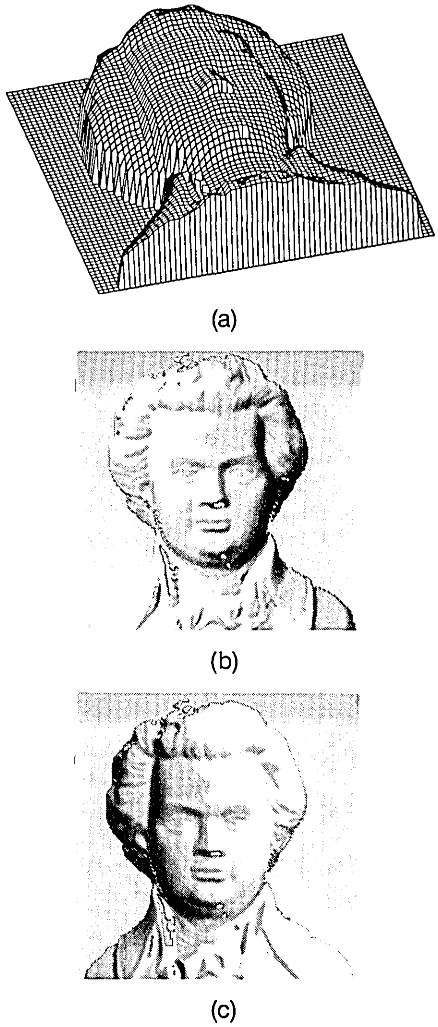


Fig. 12. Mozart test problem: (a) the ground truth of the Mozart statue; two synthetic images illuminated with (b) (albedo, tilt, slant) = (250, 135°, 45°) and (c) (albedo, tilt, slant) = (250, 45°, 45°).

the gradient direction of one reflectance map is the tangential direction of the other, and vice versa, in this region. To summarize, the optimal lighting condition can be written as

$$|\tau_1 - \tau_2| = 90°.$$

This condition has been confirmed experimentally.[17]

5. EXPERIMENTAL RESULTS

Our algorithms have been applied to four sets of photometric stereo images. The first two sets (sombrero and Mozart) are synthetic images, while the last two sets (faces of two statues) are real images. The results are compared with those obtained from the single-image SFS algorithm described in Ref. 28.

Test Problem 1: Sombrero
The tested photometric stereo images are generated from the sombrero surface as shown in Fig. 9(a). The corresponding (p, q) distribution is given in Fig. 9(b). The two photometric stereo images are generated by illuminating from (tilt, slant) = (45°, 45°) and (135°, 45°), with albedo = 250, as shown in Figs. 9(c) and 9(d), respectively. Note that these two images are shaded by light sources with orthogonal tilt angles. Results of the single-image SFS algorithm applied to Figs. 9(c) and 9(d) are given in Fig. 10. The results are not good in some regions. It is easier to see the discrepancies in the (p, q) domain. By comparing the distributions of these (p, q) values with the original distribution, we can see clearly that points move slowly along the tangential directions of the contours. The results of the photometric stereo SFS algorithm are shown in Fig. 11. By comparing the (p, q) distributions of the photometric stereo algorithm and the single-image method with the original distribution, we see that we have achieved a significant improvement by using the photometric stereo SFS method. This improvement is due to the fact that the two reflectance maps help each other and provide good sensitivity over the region of the gradient space of interest, as discussed in Section 4.

Test Problem 2: Mozart Statue
The test images are synthesized from the Mozart statue, and the surface height is obtained from the range data.

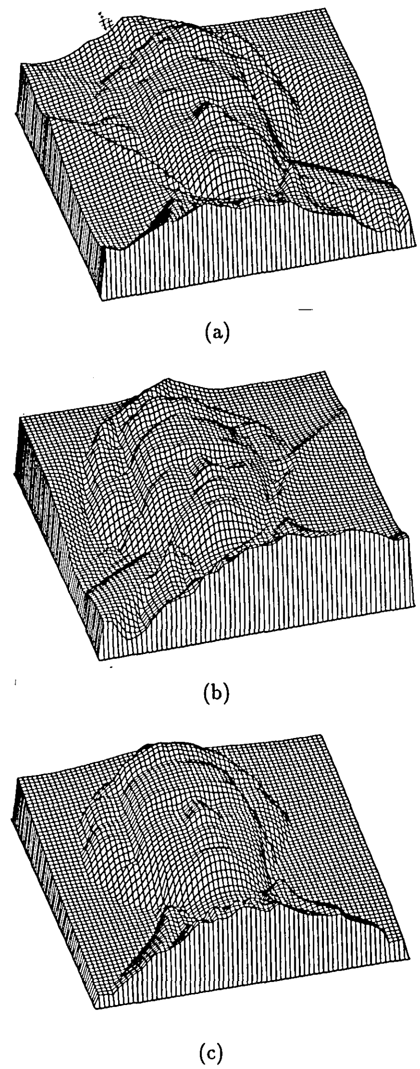
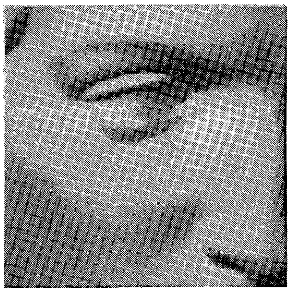


Fig. 13. Results of application of the SFS algorithms to the Mozart images: (a), (b) heights reconstructed from Figs. 12(b) and 12(c), respectively, by the single-image SFS algorithm; (c) heights reconstructed by the photometric stereo SFS algorithm.



(a)

(b)

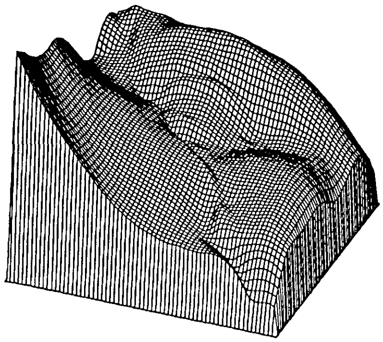
Fig. 14. David test problem: two real images of the David statue illuminated with (a) (tilt, slant) = (135°, 45°), (b) (tilt, slant) = (45°, 45°).

The original 3D surface height is plotted in Fig. 12(a), and two images generated with illuminating directions (tilt, slant) = (135°, 45°) and (45°, 45°) and albedo = 250 are shown in Figs. 12(b) and 12(c), respectively. Note that there are some defects in the original data, such as points along the object boundaries and under the nose. Figures 13(a) and 13(b) show the 3D plots of the surfaces reconstructed by the single-image SFS algorithm applied to images in Figs. 12(b) and 12(c), respectively. The results of the photometric stereo SFS algorithm are shown in Fig. 13(c). The reconstructed surfaces from the single-image SFS algorithm contain errors in regions over the face and the background, depending on the illumination direction or, equivalently, the reflectance map. In contrast, the reconstructed surface with the photometric stereo algorithm shown in Fig. 13(c) is almost the same as the original surface except at the discontinuities along the boundary and at some points that were defective on the original surface.

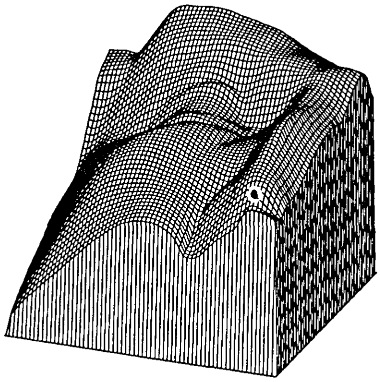
Test Problem 3: Face of the David Statue

The tested images are 128×128 real images of the face of the statue of David illuminated from directions (tilt, slant) = (135°, 45°) and (45°, 45°), as shown in Figs. 14(a) and 14(b), respectively. Even though the plaster statue has a Lambertian surface, we would like to comment on several problems concerning these images. First, it is difficult to get an ideal lighting condition that satisfies the reflectance-map model, and therefore the observed brightness at a given point may have undesirable distortions or variations that are due to the finite distance of the light source as well as to the quantization noise. Second, the shadow problem often appears in real images. Since no shading information is available in shadowed regions, it is difficult to extract the correct shape information in those regions. Additional information, such as shading from different illumination directions and other high-level visual cues, is needed for recovery of the shape in the shadowed region. Note that although there are shadow regions under the nose in both test images, they do not overlap, and hence they provide sufficient information for surface reconstruction.

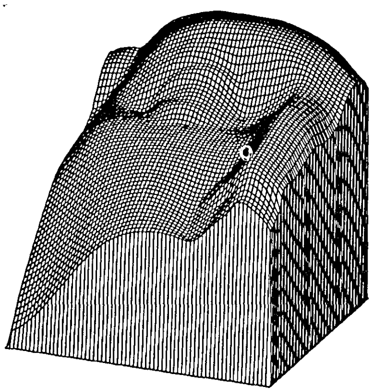
The results of the single-image SFS algorithm based on Figs. 14(a) and 14(b) and the results of the photometric stereo algorithm are shown in Figs. 15(a), 15(b), and 15(c), respectively. By comparing the two reconstructed surfaces of the single-image SFS algorithm in Figs. 15(a) and



(a)



(b)



(c)

Fig. 15. Results of application of the SFS algorithms to the David images: (a), (b) heights reconstructed from Figs. 14(a) and (b), respectively, by the single-image SFS algorithm; (c) heights reconstructed by the photometric stereo SFS algorithm.

(a)

(b)

Fig. 16. Agrippa test problem: two real images of the Agrippa statue illuminated with (a) (tilt, slant) = (135°, 50°), (b) (tilt, slant) = (45°, 45°).

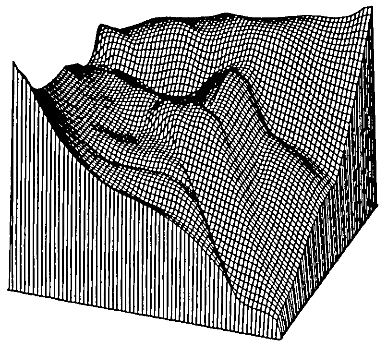
15(b), one can observe that they are not consistent with each other in several regions. Besides, the surface in Fig. 15(a) seems to be erroneous as a result of the ambiguity that occurs in the cheek region, as shown in Fig. 14(a). The reconstructed surface orientations are quite different from their true values, even though the brightness error is small. We can understand this phenomenon by examining the roof-surface example of Fig. 5. A great deal of improvement is achieved with the photometric stereo SFS algorithm. The result shown in Fig. 15(c) looks quite good. Both the brightness distortion effect and the shadow problem are resolved by combining photometric stereo images.

Test Problem 4: Face of the Agrippa Statue

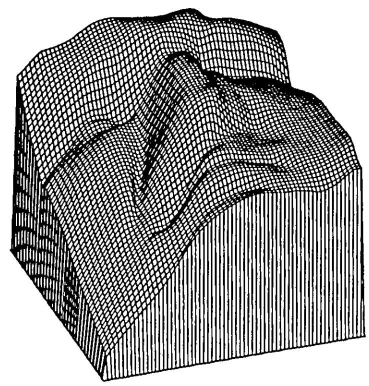
The test images are 128 × 128 real images of the face of the Agrippa statue as shown in Figs. 16(a) and 16(b), where the estimated illumination directions are (tilt, slant) = (135°, 50°) and (45°, 45°), respectively. These images have relatively large regions of shadow. The results of the single-image SFS algorithm and the photometric stereo algorithm are shown in Figs. 17(a)–17(c). We observe from Figs. 17(a) and 17(b) that ambiguities that are similar to those discussed in test problem 3 occur for both surfaces in several regions, including the cheeks. Moreover, the self-shadow problem is quite serious. It is evident from Fig. 17(c) that both ambiguity and shadow problems are resolved by use of the photometric stereo SFS algorithm. Note also that the two images in Figs. 16(a) and 16(b) have common shadowed regions near the eyes. The reconstructed shape of the regions may not be accurate, since sufficient shading information is not provided. The situation often happens in practice and may be handled by incorporating an additional photometric stereo image that gives shading information in these regions.

6. CONCLUSION AND EXTENSION

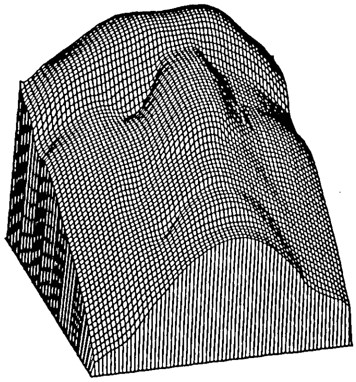
A new photometric stereo algorithm based on a nodal basis-function representation of a surface was proposed. By using the linear approximation of the reflectance map and a triangular-element surface model, we formulated the shape-reconstruction problem as a quadratic functional-minimization problem parameterized by surface heights. The new method does not require any additional integrability constraint or artificial boundary assumption. We also showed that the accuracy of the reconstructed



(a)



(b)



(c)

Fig. 17. Results of application of the SFS algorithms to the Agrippa images: (a), (b) heights reconstructed from Figs. 16(a) and (b), respectively, by the single-image SFS algorithm; (c) heights reconstructed by the photometric stereo SFS algorithm.

surface and the performance of the single-image SFS algorithm are related to the slope of the reflectance-map function in the gradient space and that more-accurate surfaces can be reconstructed by the proper combination of several reflectance maps and the corresponding image information. Our algorithm has been tested for several synthetic and real images. Experimental results show that the proposed photometric SFS algorithm is robust and reliable and produces more-accurate reconstructed surfaces than does the single-image SFS algorithm. Even in the presence of intensity distortions, noise, and shadows in real images, the new photometric SFS algorithm produces robust reconstructed surfaces, while the single-image SFS algorithm does not. The effect of different illumination directions of two light sources was also examined. The best result can be obtained when the two illumination directions are orthogonal to each other, since their reflectance maps complement each other optimally in the central region of the (p, q) domain.

Compared with the conventional photometric stereo method, our new iterative SFS method has two major advantages: it has no integrability problem, and it is insensitive to noise. It would be interesting to see the results of a combination of the conventional and the new method: for example, the application of the conventional method in regions with discontinuities and changing albedo combined with the application of the new method to the remaining region.

ACKNOWLEDGMENT

This research was supported by National Science Foundation Research Initiation Award ASC-9009323.

REFERENCES

1. 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, Mass., 1970).
2. B. K. P. Horn, *Obtaining Shape from Shading Information* (MIT Press, Cambridge, Mass., 1975).
3. R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithm," IEEE Trans. Pattern Anal. Mach. Intell. **10,** 439–451 (1989); in *Shape from Shading,* B. K. P. Horn and M. J. Brooks, eds. (MIT Press, Cambridge, Mass., 1989).
4. B. K. P. Horn, "Height and gradient from shading," Int. J. Comput. Vision **5,** 584–595 (1990).
5. B. K. P. Horn and M. J. Brooks, "The variational approach to shape from shading," Comput. Vision Graphics Image Process. **33,** 174–208 (1986); in *Shape from Shading,* B. K. P. Horn and M. J. Brooks, eds. (MIT Press, Cambridge, Mass., 1989).
6. K. Ikeuchi and B. K. P. Horn, "Numerical shape from shading and occluding boundaries," Artif. Intell. **17,** 141–184 (1981); in *Shape from Shading,* B. K. P. Horn and M. J. Brooks, eds. (MIT Press, Cambridge, Mass., 1989).
7. D. Lee, "A provably convergent algorithm for shape from shading," in *Shape from Shading,* B. K. P. Horn and M. J. Brooks, eds. (MIT Press, Cambridge, Mass., 1989).
8. G. B. Smith, "The relationship between image irradiance and surface orientation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1983), pp. 404–413.
9. R. Szeliski, "Fast surface interpolation using hierarchical basis functions," IEEE Trans. Pattern Anal. Mach. Intell. **12,** 513–528 (1990).
10. R. Szeliski, "Fast shape from shading," Comput. Vision Graphics Image Process.: Image Understanding **53,** 129–153 (1991).
11. Q. Zheng and R. Chellappa, "Estimation of illumination direction, albedo, and shape from shading," IEEE Trans. Pattern Anal. Mach. Intell. **13,** 680–702 (1991).
12. A. P. Pentland, "Shape information from shading: a theory about human perception," in *Proceedings of the IEEE International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, New York, 1988), pp. 404–413.
13. P. Dupuis and J. Oliensis, "Direct method for reconstructing shape from shading," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1992), pp. 453–458.
14. J. Oliensis, "Uniqueness in shape from shading," Int. J. Comput. Vision **6,** 75–104 (1991).
15. 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,** 309–328 (1982).
16. K. Ikeuchi, "Determining surface orientations of specular surfaces by using the photometric stereo method," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-13,** 661–669 (1981).
17. K. M. Lee and C.-C. J. Kuo, "Shape reconstruction from photometric stereo," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1992), pp. 479–484.
18. R. Ray, J. Birk, and R. Kelly, "Error analysis of surface normals determined by radiometry," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-5,** 631–645 (1983).
19. W. Silver, "Determining shape and reflectance using multiple images," master's thesis (Massachusetts Institute of Technology, Cambridge, Mass., 1980).
20. H. D. Tagare and R. J. P. DeFigueiredo, "A theory of photometric stereo for a class of diffuse nonLambertian surface," IEEE Trans. Pattern Anal. Mach. Intell. **13,** 133–152 (1991).
21. R. J. Woodham, "A cooperative algorithm for determining surface orientation from a single view," in *Proceedings of the International Joint Conference on Artificial Intelligence* (Kaufmann, Los Altos, Calif., 1977), pp. 635–641.
22. R. J. Woodham, "Photometric method for determining surface orientation from multiple images," Opt. Eng. **19,** 139–144 (1980).
23. R. J. Woodham, "Analyzing images of curved surfaces," Artif. Intell. **17,** 117–140 (1981).
24. W. E. L. Grimson, "Binocular shading and visual surface reconstruction," Comput. Vision Graphics Image Process. **28,** 19–43 (1984).
25. K. Ikeuchi, "Reconstructing a depth map from intensity maps," in *Proceedings of the IEEE International Conference on Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1984), pp. 736–738.
26. K. Ikeuchi, "Determining a depth map using a dual photometric stereo," Int. J. Robot. Res. **6,** 15–31 (1987).
27. 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,** 219–226 (1991).
28. K. M. Lee and C.-C. J. Kuo, "Shape from shading with a linear triangular element surface model," Tech. Rep. 172 (Signal and Image Processing Institute, University of Southern California, Los Angeles, Calif., 1991); IEEE Trans. Pattern Anal. Mach. Intell. (to be published).
29. Y. G. Leclerc and A. F. Bobick, "The direct computation of height from shading," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1991), pp. 552–558.
30. C. Johnson, *Numerical Solutions of Partial Differential Equations by the Finite Element Method* (Cambridge U. Press, Cambridge, 1987).
31. H. R. Schwartz, *Finite Element Methods* (Academic, New York, 1988).

32. W. E. L. Grimson, "An implementation of a computational theory of visual surface interpolation," Comput. Vision Graphics Image Process. **22,** 39–64 (1983).
33. D. Terzopoulos, "Multilevel computational processes for visual surface reconstruction," Comput. Vision Graphics Image Process. **24,** 52–96 (1983).
34. D. Terzopoulos, "The computation of visual surface representation," IEEE Trans. Pattern Anal. Mach. Intell. **10,** 417–438 (1988).
35. D. Terzopoulos, "Image analysis using multigrid relaxation methods," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-8,** 129–139 (1986).