Gradient and curvature from the photometric-stereo method, including local confidence estimation

Robert J. Woodham

Laboratory for Computational Intelligence, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

The photometric-stereo method is one technique for three-dimensional shape determination that has been implemented in a variety of experimental settings and that has produced consistently good results. The idea is to use intensity values recorded from multiple images obtained from the same viewpoint but under different conditions of illumination. The resulting radiometric constraint makes it possible to obtain local estimates of both surface orientation and surface curvature without requiring either global smoothness assumptions or prior image segmentation. Photometric stereo is moved one step closer to practical possibility by a description of an experimental setting in which surface gradient estimation is achieved on full-frame video data at near-video-frame rates (i.e., 15 Hz). The implementation uses commercially available hardware. Reflectance is modeled empirically with measurements obtained from a calibration sphere. Estimation of the gradient (p, q) requires only simple table lookup. Curvature estimation additionally uses the reflectance map $R(p,q)$. The required lookup table and reflectance maps are derived during calibration. Because reflectance is modeled empirically, no prior physical model of the reflectance characteristics of the objects to be analyzed is assumed. At the same time, if a good physical model is available, it can be retrofitted to the method for implementation purposes. Photometric stereo is subject to error in the presence of cast shadows and interreflection. No purely local technique can succeed because these phenomena are inherently nonlocal. Nevertheless, it is demonstrated that one can exploit the redundancy in three-light-source photometric stereo to detect locally, in most cases, the presence of cast shadows and interreflection. Detection is facilitated by the explicit inclusion of a local confidence estimate in the lookup table used for gradient estimation.

1. INTRODUCTION

The purpose of computational vision is to produce descriptions of a three-dimensional (3D) world from two-dimensional (2D) images of that world, sufficient to carry out a specified task. Robustness in many tasks is improved when use is made of all the information available in an image, not just that obtained from a sparse set of features. The idea of using multiple images obtained from the same viewpoint but under different conditions of illumination has emerged as a principled way to obtain additional local radiometric constraint in the computation of shape-from-shading. In many cases multiple images overdetermine the solution locally, further improving robustness by permitting local validation of the radiometric model used.

The photometric-stereo method is one technique for shape-from-shading that has been implemented in a variety of experimental settings and that has produced consistently good results. The underlying theory follows from principles of optics. An image irradiance equation is developed to determine image irradiance as a function of surface orientation. This equation cannot be inverted locally because image brightness provides only one measurement, whereas surface orientation has 2 degrees of freedom. Brightness values obtained from the same viewpoint but under different conditions of illumination do make it possible to obtain dense, local estimates of both surface orientation and surface curvature without requiring either global smoothness assumptions or prior image segmentation. In three-light-source photometric stereo the three local intensity measurements overdetermine the 2 degrees of freedom of surface orientation. Similarly,

the six local spatial derivatives of intensity overdetermine the 3 degrees of freedom of surface curvature. Photometric stereo was originally described in Refs. 1 and 2. The first implementation was by Silver.[3] Photometric stereo has since been used for a variety of recognition and localization tasks.[4-9] The relation to surface curvature also has been explored.[10-15]

Optics determines that an image irradiance equation necessarily exists but says little about the particular form that this image irradiance equation must take. Of course, one can exploit situations in which the reflectance properties of a material are known to satisfy a particular functional form. Formal analysis of these situations helps to establish the existence, the uniqueness, and the robustness of solution methods under varying degrees of uncertainty and approximation. Implementation also is facilitated because the resulting computations typically involve equations of known form with unknown coefficients that can be determined as a problem of parameter estimation. By now, the literature on reflectance models applicable to computer graphics and computer vision is quite vast. Some examples are explicitly related to photometric stereo.[16-19]

The usual formulation of an image irradiance equation assumes that each surface element receives illumination only directly from the light source(s). This is correct for scenes consisting of a single convex object. In general, a surface element also receives illumination indirectly from light reflected from other surface elements in the scene. Interreflection is important for two reasons. First, as demonstrated by Gilchrist,[20] interreflection near concave junctions is perceptually salient. It is determined by the intrinsic reflectance (i.e., albedo) of

the surface material, independent of the amplitude of the ambient illumination.[21,22] Second, local shape recovery methods are unreliable if interreflection is not taken into account, as has been shown recently.[23,24]

Indeed, photometric stereo is subject to error in the presence of cast shadows and interreflection. No purely local technique can succeed because these phenomena are inherently nonlocal. Nevertheless, in this paper I demonstrate that one can exploit the redundancy in three-light-source photometric stereo to detect locally, in most cases, the presence of cast shadows and interreflection. Detection is facilitated by the explicit inclusion of a local confidence estimate in the lookup table used for gradient estimation.

Here it is assumed that the objects to be analyzed are made of a single material. Reflectance properties are measured with a calibration sphere made of this material. Measurements from the calibration sphere are directly applicable to the analysis of other objects that are of different shape but made of the same material and that are illuminated and viewed under the same conditions. In this way a material with any reflectance characteristic can be handled, provided that the necessary calibration can be done. In some applications it may be necessary to use paint (or another coating) to match reflectance properties between a calibration sphere and the objects to be analyzed.

Three representations for reflectance data are employed. First, the relationship between measured intensity and surface gradient is represented by the familiar reflectance map $R(p,q)$. Second, for a material with constant albedo, triples of measured intensity values, $[E_1, E_2, E_3]$, are shown to define a 2D surface in the 3D coordinate space whose axes are E_1, E_2, and E_3. In the absence of interreflection and cast shadows, all triples of measured intensity fall on this surface. Interreflection causes some intensity values to be larger than expected. Cast shadows cause some intensity values to be smaller than expected. Thus, in practice, not all measured triples will lie exactly on the specified surface. Third, the relationship between triples of intensity values $[E_1, E_2, E_3]$ and surface gradient is represented in a lookup table for implementation purposes.

In the implementation described, each measured triple is projected onto the 2D calibration surface to estimate the gradient. The distance each measured triple must be projected is used to define a local confidence estimate. Projection tends to decrease error that is due to interreflection. Of course, any claim to quantitative correction would be fortuitous because neither interreflection nor cast shadows are a local effect. On the other hand, the local confidence estimate is a reliable way to notice when a measured triple does not lie on the required surface. This supports the local detection of interreflection and cast shadows. Experimental measurements verify that this is indeed the case. Local detection of interreflection can be used to prevent erroneous surface reconstruction in regions so detected. Should accurate surface reconstruction be the goal, this becomes useful input for schemes intended to reason more globally, such as the one described in Ref. 24.

In this paper photometric stereo is moved closer to practical possibility in two ways. First, an experimental set-ting is described in which the multiple images required for photometric stereo are acquired simultaneously at video rates. No light sources must be turned on and off. Instead spectral multiplexing is used. Three spectrally distinct light sources illuminate the workspace from different directions, and a suitable red–green–blue (RGB) color camera acquires three-channel video images that subsequently are treated as three separate black-and-white (B&W) images, one corresponding to each condition of illumination. Second, commercially available image-processing hardware is used to pass the three-channel video data through a lookup table of length 2^{18}. Lookup-table output includes both the gradient and the local confidence estimate. Overall, processing of full-frame video data occurs at near-video-frame rates (i.e., 15 Hz).

Section 2 presents the background and the theory. Section 3 describes the particular implementation and reports on the experiments performed. Section 4 provides a brief discussion and a summary of the conclusions following from the study reported. Finally, Appendix A derives the results cited in Section 2 that are unique to the Lambertian case.

2. BACKGROUND AND THEORY

A given spatial arrangement of objects made of a given set of materials, illuminated in a given way, and viewed from a given vantage point determines an image according to the laws of optics. Geometric equations determine where each point on a visible surface appears in the image, and corresponding radiometric equations determine its brightness and color.

Reflectance modeling is difficult, in general, because complexities arise at several levels. Local reflectance depends not only on the intrinsic optical properties of a material but also, for example, on its surface roughness. Geometric-ray analysis of surface microstructure can be complex. At finer scales one also must take into account the wave nature of light, adding even more complexity. In scenes consisting of multiple objects geometric analysis at the macroscale also becomes complex, making it difficult to deal effectively with interreflections and cast shadows. Despite this, there is considerable research on reflectance models for computer vision. An edited collection[25] provides a good introduction to the relevant literature.

A key observation, first make by Horn,[21] is that image irradiance can be written as a function only of surface orientation, for many imaging situations of practical importance. Horn's research, formulated for the problem of shape-from-shading, introduced the idea of the reflectance map. Reflectance maps can be derived from formal reflectance models, when they are available, or, as is the case here, when they can be measured empirically.

A. Shape-from-Shading and the Reflectance Map
The standard geometry of shape-from-shading is assumed. That is, let the object surface be given explicitly by $z = f(x,y)$ in a left-handed Euclidean coordinate system, in which the viewer is looking in the positive Z direction, image projection is orthographic, and the image XY axes coincide with the object XY axes. The surface

gradient (p, q) is defined by

$$p = \frac{\partial f(x, y)}{\partial x}, \qquad q = \frac{\partial f(x, y)}{\partial y},$$

so that a surface normal vector is $[p, q, -1]$. Thus the gradient (p, q) is one way to represent surface orientation. An image irradiance equation can be written as

$$E(x, y) = R(p, q),   \tag{1}$$

where $E(x, y)$ is the image irradiance and $R(p, q)$ is the reflectance map. A reflectance map combines information about surface material, scene illumination, and viewing geometry into a single representation that determines image brightness as a function of surface orientation.

Given an image $E(x, y)$ and the corresponding reflectance map $R(p, q)$, shape-from-shading typically is defined as the problem of determining a smooth surface, $z = f(x, y)$, that satisfies the image irradiance equation over some domain Ω, including any initial conditions that may be specified on the boundary $\partial\Omega$ or elsewhere. It should not be assumed, however, that reconstruction of the surface height function, $z = f(x, y)$, is always the goal in shape-from-shading. For example, orientation-based representations of surface shape can be used for object recognition and localization tasks.[4–6,9] These representations use surface orientation directly, without computing the explicit representation of the object surface, $z = f(x, y)$.

Since Horn's original study, a substantial but scattered literature on shape-from-shading has developed. Two essential references are Horn's text[26] and a collection of papers edited by Horn and Brooks.[27] With a single image, shape-from-shading problems typically are solved by exploitation of *a priori* constraints on the reflectance map $R(p, q)$, *a priori* constraints on surface curvature, or global smoothness constraints. Photometric stereo, on the other hand, makes use of additional images.

B. Photometric Stereo

Photometric stereo uses multiple images obtained from the identical geometry but under different conditions of illumination. Three image irradiance equations,

$$E_1(x, y) = R_1(p, q),$$

$$E_2(x, y) = R_2(p, q),$$

$$E_3(x, y) = R_3(p, q),   \tag{2}$$

in general overdetermine the solution at each point (x, y) because three intensity measurements, $E_1(x, y)$, $E_2(x, y)$, and $E_3(x, y)$, are used to estimate two unknowns, p and q.

Conceptually, the implementation of photometric stereo is straightforward. Using a calibration object of known shape, one can build a lookup table mapping triples of measured brightness values $[E_1, E_2, E_3]$ to the corresponding gradient (p, q). Suppose that each image is accurate to $2^8 = 256$ gray values. Then a full table would have $2^8 \times 2^8 \times 2^8 = 2^{24}$ entries. Despite advances in commercial hardware, a lookup table of this size is still prohibitive in terms of memory and real-time throughput capacity. The implementation described in Section 3 uses 2^6 gray values from each image as input for a lookup table with 2^{18} entries to achieve near-real-time throughput at 15 Hz.

Albedo Variation

A reflectance map determines measured intensity as a function of the surface gradient for a particular surface material, a particular scene illumination, and a particular viewing geometry. If the material's bidirectional reflectance factor (i.e., albedo) also varies spatially, independent of the gradient, one would obtain

$$E_1(x, y) = \rho(x, y) R_1(p, q),$$

$$E_2(x, y) = \rho(x, y) R_2(p, q),$$

$$E_3(x, y) = \rho(x, y) R_3(p, q),   \tag{3}$$

where $\rho(x, y)$ is the albedo, a function of (x, y). Thus Eqs. (2) are a special case of Eqs. (3), in which the albedo is constant (and normalized to 1). Equations (3) are three equations in the three unknowns, ρ, p, and q. The equations generally are nonlinear, so a unique solution cannot be guaranteed. For the Lambertian case, Eqs. (3) become linear when unit surface normals are used instead of the gradient to represent surface orientation, as is shown in Ref. 2. In this case every triple of measured brightness values $[E_1, E_2, E_3]$ determines a unique albedo ρ and gradient (p, q).

It is useful to compare the formulation of photometric stereo given in Eqs. (2) with that used by Nayar et al.[24] Both use three light sources. Equations (2) are not restricted to the Lambertian case. Constant albedo is assumed, and as a consequence the problem is overdetermined locally. Nayar's research is restricted to the Lambertian case. Albedo is allowed to vary because it is determined locally, in addition to the gradient. But, as Nayar et al. point out, this determination is erroneous in the presence of interreflection. Errors cannot be detected locally. Instead the pseudoshape is defined to be the (possibly erroneous) shape and albedo that is determined by pointwise solution of the linearized version of Eqs. (3). Nayar et al. further show that this pseudoshape is unique for a given actual shape, independent of the illumination and viewpoint. Finally, they demonstrate an iterative algorithm to reconstruct a possible actual shape and albedo from the measured pseudoshape. This iterative algorithm is nonlocal, requiring explicit reconstruction of the surface height function $z = f(x, y)$, at each iteration step.

The comparison now is clear. Both local solutions are subject to error in the presence of interreflection and cast shadows. When the problem is not overdetermined, as with Nayar et al., there is little one can do locally to detect errors. Their approach necessarily requires global analysis. On the other hand, when the problem is overdetermined, as is the case with Eqs. (2), it becomes possible also to detect inconsistencies locally.

A key benefit of using multiple light sources is the ability to overdetermine the problem locally. If one must account for spatially varying albedo, as in Eqs. (3), it is possible to use a fourth light source, as suggested in Ref. 28. The idea is to define three ratio images $\hat{E}_i = E_i/E_4$, with associated reflectance maps $\hat{R}_i = R_i/R_4$, by using the fourth image E_4 as a reference. If ratios are used, the four-light-source case becomes equivalent to the constant albedo three-light-source case, given by Eqs. (2), because the spatially varying albedo terms cancel when

the ratios are formed. Once the gradient is determined locally, the albedo also can be determined. [In practice, it is preferable to define $\hat{E}_i = (E_i - E_4)/(E_i + E_4)$ so that $\hat{R}_i = (R_i - R_4)/(R_i + R_4)$, but the idea is the same.]

In this paper constant albedo is assumed, so that three-light-source photometric stereo is characterized by Eqs. (2). When albedo varies, as in Eqs. (3), we can reduce the problem to the form of Eqs. (2) by using a fourth light source and ratioed images, as indicated. If the albedo also depends on imaging geometry, then the image irradiance equations become $E_i(x, y) = R_i(p, q, \rho)$, where $i = 1, 2, 3, 4$. Ratioing no longer suffices. Nevertheless, measured intensities are constrained to lie on a 3D manifold in the 4D space of possibilities, demonstrating that the solution is still overdetermined locally.

C. Two-Dimensional Surfaces in the Three-Dimensional Space of Measured Intensities

Equations (2) define the parametric equations, in parameters p and q, of a 2D surface in the 3D coordinate space whose axes are E_1, E_2, and E_3. Thus, with constant albedo and no interreflection or cast shadows, triples of measured intensity values $[E_1, E_2, E_3]$ are constrained to lie on a 2D surface in any three-light-source photometric-stereo situation. This 2D surface exists independent of the particular parameterization. When parameterized in terms of the gradient (p, q), the equations take on the form given by Eqs. (2).

Example: Lambertian Reflectance

When an ideal Lambertian material is illuminated by a single distant light source, image irradiance is proportional to $\cos(i)$, where i is the incident angle (i.e., the angle between the surface normal and a vector in the direction of the light source).

Figure 1 shows the scatter plot obtained for Lambertian reflectance corresponding to three different directions of equal-strength distant-point-source illumination. The gradients corresponding to the light-source directions for $E_1(x, y)$, $E_2(x, y)$, and $E_3(x, y)$ are $(0.7, 0.3)$, $(-0.7, 0.3)$, and $(0, 0)$, respectively.

In Fig. 1 and in all the scatter plots that follow, the 2D surface is shown as an empirical point plot of measured intensity triples $[E_1, E_2, E_3]$. E_1, E_2, and E_3 define the axes of a right-handed Euclidean coordinate system. One can think of the E_1 axis as pointing east, the E_2 axis as pointing north, and the E_3 axis as pointing up. Scatter plots are displayed as orthographic projections from a given viewing direction. The viewing direction is specified by an elevation and an azimuth. Elevation is the angle above the ground (i.e., $E_1 E_2$) plane, and azimuth is measured clockwise from the north (i.e., clockwise from the E_2 axis). Throughout, two particular viewing directions are used: (a) elevation = 10.0, azimuth = 225.0 and (b) elevation = 0.0, azimuth = 0.0. (All angles are in degrees.) View (a) corresponds to a low-altitude view from the southwest, and view (b) corresponds to a projection onto the $E_1 E_3$ plane. (The E_2 axis points directly at the viewer.)

The 2D surface depicted in Fig. 1 is a 6-degree-of-freedom ellipsoid. The particular ellipsoid is determined by the relative strength and configuration of the three light sources. (This result is derived in Appendix A.)

The ellipsoid does not depend on the shape of the object in view or on the relative orientation between object and viewer.

Example: Phong Reflectance

The reflectance of many materials is modeled as a combination of a diffuse component and a specular component. In recent years considerable progress has been made to formulate models based on principles of optics.[25] One early model, well known in both computer vision and graphics, is Phong shading. Phong shading is a purely phenomenological model that now is considered inadequate because, for example, it fails to satisfy Helmholtz's law of reciprocity.[29] A variant of Phong shading, described in Ref. 30, Sec. 39, does satisfy reciprocity. Image irradiance is proportional to

$$\frac{\cos(i)[(1 - \alpha) + \alpha \cos^n(s/2)]}{\cos(g/2)},$$

where i, as above, is the incident angle, g is the phase angle (i.e., the angle between the vector pointing to the light source and the vector pointing to the viewer), s is the off-specular angle (i.e., the angle between the vector pointing to the viewer and the vector that defines, relative to the light-source direction and the surface normal, the direction of perfect specular reflection), α is a fraction, $0 \le \alpha \le 1$, that models how much of the incident light is reflected specularly, and n is a number that models
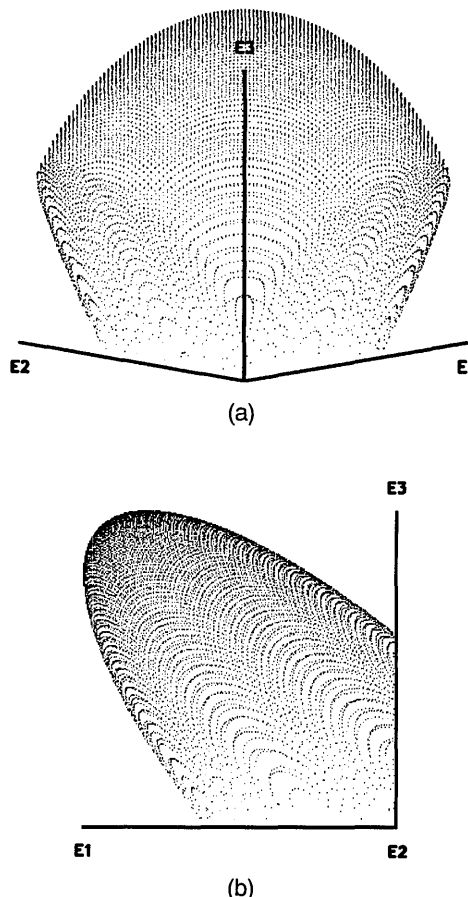

(a)

(b)

Fig. 1. Lambertian reflectance: scatter plot of measured intensity triples. (a) Elevation 10.0, azimuth 225.0; (b) elevation 0.0, azimuth 0.0.
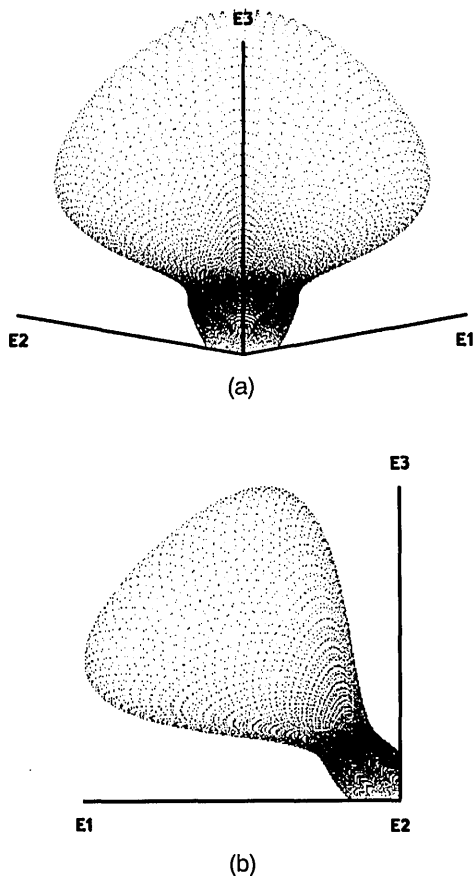
(a)

(b)

Fig. 2.   Phong reflectance: scatter plot of measured intensity triples.   (a), (b) Same as in Fig. 1.

how compact the specular patch is about the direction of perfect specular reflection.   Parameters α and n vary according to the properties of the material modeled.   Phong shading is used here only for illustrative purposes.

Figure 2 shows the scatter plot obtained for Phong reflectance ($\alpha = 0.75, n = 20$) in the same three-light-source configuration used for the Lambertian case (Fig. 1).   The 2D surface defined is still smooth, but clearly it no longer is an ellipsoid.   Similar results would be noted for other non-Lambertian reflectance models.[25]

D.   Surface Curvature

The principal curvatures, and measures derived from principal curvature, are viewpoint invariant and therefore play a potentially valuable role in shape representation for tasks including surface segmentation, object recognition, attitude determination, and surface reconstruction. In differential geometry there are a variety of representations from which principal curvatures can be determined. Many are derived from the explicit surface height representation, $z = f(x, y)$.   In this paper I develop representations for surface curvature based on the gradient (p, q), the reflectance map $R(p, q)$, and the image intensity $E(x, y)$.   In particular, the goal is to determine what independent information about surface curvature can be extracted from the image irradiance equation.

There are 3 degrees of freedom to the curvature at a point on a smooth surface.   Consequently three parameters are required to specify curvature.   One representation is in terms of the 2×2 matrix of second partial

derivatives of the surface $z = f(x, y)$.   For notational convenience let

$$p_x = \frac{\partial^2 f(x, y)}{\partial x^2}, \qquad p_y = \frac{\partial^2 f(x, y)}{\partial x \partial y},$$

$$q_x = \frac{\partial^2 f(x, y)}{\partial y \partial x}, \qquad q_y = \frac{\partial^2 f(x, y)}{\partial y^2}.$$

Now let $\mathbf{H}$ be the matrix

$$\mathbf{H} = \begin{bmatrix} p_x & p_y \\ q_x & q_y \end{bmatrix}.$$

$\mathbf{H}$ is called the Hessian matrix of $z = f(x, y)$.   It may appear that four parameters are required for specification of $\mathbf{H}$.   But for smooth surfaces $\mathbf{H}$ is symmetric.   That is, $p_y = q_x$.   Therefore only three parameters are required after all. $\mathbf{H}$ is a viewer-centered representation of surface curvature because its definition depends on the explicit form of the surface function, $z = f(x, y)$, and on the fact that the viewer is looking in the positive Z direction.

From the Hessian $\mathbf{H}$ and the gradient (p, q) one can determine a viewpoint-invariant representation of surface curvature.   Let $\mathbf{C}$ be the matrix

$$\mathbf{C} = (1 + p^2 + q^2)^{-3/2} \begin{bmatrix} q^2 + 1 & -pq \\ -pq & p^2 + 1 \end{bmatrix} \mathbf{H}. \qquad (4)$$

Now let k_1 and k_2 be the two eigenvalues of $\mathbf{C}$, with associated eigenvectors ω_1 and ω_2.   Then k_1 and k_2 are the principal curvatures, with directions ω_1 and ω_2, at $z = f(x, y)$.   The principal curvatures k_1 and k_2 are viewpoint-invariant surface properties because they do not depend on the viewer-centered XYZ coordinate system.   Equation (4) determines principal curvatures from the Hessian matrix $\mathbf{H}$ and the gradient (p, q).   The terms in Eq. (4) involving the gradient (p, q) can be interpreted as the corrections necessary to account for the geometric foreshortening associated with viewing a surface element obliquely.

The directions ω_1 and ω_2 are viewpoint dependent.   Although the directions of principal curvature are orthogonal in the object-centered coordinate system defined by the local surface normal and tangent plane, they are not, in general, orthogonal when projected onto the image plane.   Thus k_1, k_2, ω_1, and ω_2 together constitute four independent parameters that can be exploited.   (Because they are viewpoint dependent, the directions ω_1 and ω_2 are not typically used in surface representations proposed for object recognition.   Note, however, that Brady et al.[31] argue that, in many cases, the lines of curvature form a natural parameterization of a surface.)

Besl and Jain[32,33] classify sections of a smooth surface into one of eight basic types based on the sign and the zeros of Gaussian and mean curvature.   The Gaussian curvature K, also called the total curvature, is the product, $K = k_1 k_2$, of the principal curvatures.   The mean curvature H is the average, $H = (k_1 + k_2)/2$, of the principal curvatures.   It follows from elementary matrix theory that

$$K = \det(\mathbf{C}), \qquad H = \frac{1}{2} \operatorname{tr}(\mathbf{C}). \qquad (5)$$

The expression for K further simplifies to

$$K = \frac{1}{(1 + p^2 + q^2)^2} \det(\mathbf{H}). \quad (6)$$

Thus the sign of $\det(\mathbf{H})$ is the sign of the Gaussian curvature.

Other local curvature measures can be defined as well. If k is a principal curvature, then $r = 1/k$ is the associated radius of principal curvature. For a smooth surface, the first and the second curvature functions are defined as the sum of the principal radii of curvature and the product of the principal radii of curvature, respectively. For smooth surfaces, the second curvature function is equivalent to what has been called the extended Gaussian image in computer vision.[34] These curvature functions possess desirable mathematical properties that can be exploited for object recognition and attitude determination.[9,35] Also, Koenderink proposes two new curvature measures called curvedness and shape index.[36]

Clearly, if one could locally determine the Hessian $\mathbf{H}$, then one could locally compute the curvature matrix $\mathbf{C}$ by using the gradient (p, q) obtained from photometric stereo and Eq. (4). Given $\mathbf{C}$, one could examine its eigenvalue/eigenvector structure to determine any local curvature representation involving the principal curvatures k_1 and k_2 and their associated directions ω_1 and ω_2, including Gaussian curvature K, mean curvature H, curvedness, and shape index.

E. Determining the Hessian

It would seem that determining the Hessian $\mathbf{H}$ in photometric stereo requires nothing more than numerical differentiation of the gradient estimate (p, q). Although this may seem adequate, differentiating the gradient without prior global smoothing is unstable, especially if the gradient itself is inaccurate or if it has been quantized into too small a set of discrete values. Therefore it is useful to determine what independent information about surface curvature can be extracted from the image irradiance equation.

By taking partial derivatives of the image irradiance equation (1) with respect to x and y, one obtains two equations that can be written as the single matrix equation[10]

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \mathbf{H} \begin{bmatrix} R_p \\ R_q \end{bmatrix}. \quad (7)$$

Subscripts $x, y, p,$ and q denote partial differentiation, and the dependence of E on (x, y) and of R on (p, q) has been omitted for clarity. The vector $[E_x, E_y]$ is normal to the contour of constant brightness in the image at the given point (x, y). The vector $[R_p, R_q]$ is normal to the contour of constant brightness in the reflectance map at the given gradient (p, q). Equation (7) alone is not enough to determine the Hessian $\mathbf{H}$. But with photometric stereo one such equation is obtained for each image. In a two-light-source case, one obtains

$$\mathbf{H} = \begin{bmatrix} E_{1x} & E_{2x} \\ E_{1y} & E_{2y} \end{bmatrix} \begin{bmatrix} R_{1p} & R_{2p} \\ R_{1q} & R_{2q} \end{bmatrix}^{-1}, \quad (8)$$

provided that the required matrix inverse exists. In the three-light-source case, the problem once again is overdetermined. One can write

$$\mathbf{H} = \begin{bmatrix} E_{1x} & E_{2x} & E_{3x} \\ E_{1y} & E_{2y} & E_{3y} \end{bmatrix} \mathbf{R}, \quad (9)$$

where

$$\mathbf{R} = \mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1},$$

with the T denoting the matrix transpose, and

$$\mathbf{M} = \begin{bmatrix} R_{1p} & R_{1q} \\ R_{2p} & R_{2q} \\ R_{3p} & R_{3q} \end{bmatrix},$$

provided that the required matrix inverse exists. Equation (9) is the standard least-squares estimate of the solution to an overdetermined set of linear equations. It can be extended, in the obvious way, to situations in which more than three light sources are used.

The matrices $\mathbf{M}$ and $\mathbf{R}$ are matrix functions of the gradient (p, q). They depend only on the three reflectance maps $R_i(p, q)$, where $i = 1, 2, 3$. The matrix function $\mathbf{R}$ can be determined at the time of gradient lookup-table calibration. It too can be thought of as a (large) lookup table, indexed by the gradient (p, q). The matrix $\mathbf{M}^T\mathbf{M}$, whose inverse is required, is independent of the three images $E_i(x, y)$, where $i = 1, 2, 3$, and hence independent of the particular surface in view. Thus, for a particular surface material, the principal factors that determine the existence and the robustness of the computation are the nature and the distribution of the light sources. No useful local information is obtained when $[R_p, R_q]$ is zero. This occurs at local extrema of $R(p, q)$ and at gradients (p, q) shadowed from the light source. There also may be gradients (p, q) where two of the three $[R_p, R_q]$ vectors are nearly parallel. Local degeneracies in the two-light-source configuration can be eliminated, and the effects of shadows minimized, when three- rather than two-light-source photometric stereo is used.

In Eq. (9) the magnitude of $[R_p, R_q]$ plays the role of a weight that pulls the three-source solution toward an image irradiance equation for which the magnitude of $[R_p, R_q]$ is large (and consequently away from an image irradiance equation for which the magnitude of $[R_p, R_q]$ is small). This factor has a desirable effect because locations in an image at which the magnitude of $[R_p, R_q]$ is small will contribute minimal information, and it is good that they are discounted. Consequently, points that are shadowed with respect to one of the light sources need not be considered as a special case. Indeed, when one of the $[R_p, R_q]$ vectors is zero, the three-light-source solution, given by Eq. (9), reduces to the two-light-source solution, given by Eq. (8).

The lookup table for gradient estimation, the reflectance maps $R_i(p, q)$, where $i = 1, 2, 3$, and the matrix $\mathbf{R}$ are determined during calibration. Subsequently, on a pixel-by-pixel basis, the three local measurements of intensity, $[E_1, E_2, E_3]$, are used to estimate the gradient (p, q). The six partial spatial derivatives of intensity, E_{ix}, E_{iy}, where $i = 1, 2, 3$, together with the gradient (p, q), are used to estimate the three parameters of the Hessian, $p_x, q_y,$ and $p_y = q_x$. Thus a total of nine independent local measurements are used to estimate a total of five local parameters. The estimate of the Hessian $\hat{\mathbf{H}}$ is not strictly independent of the estimate of the gradient
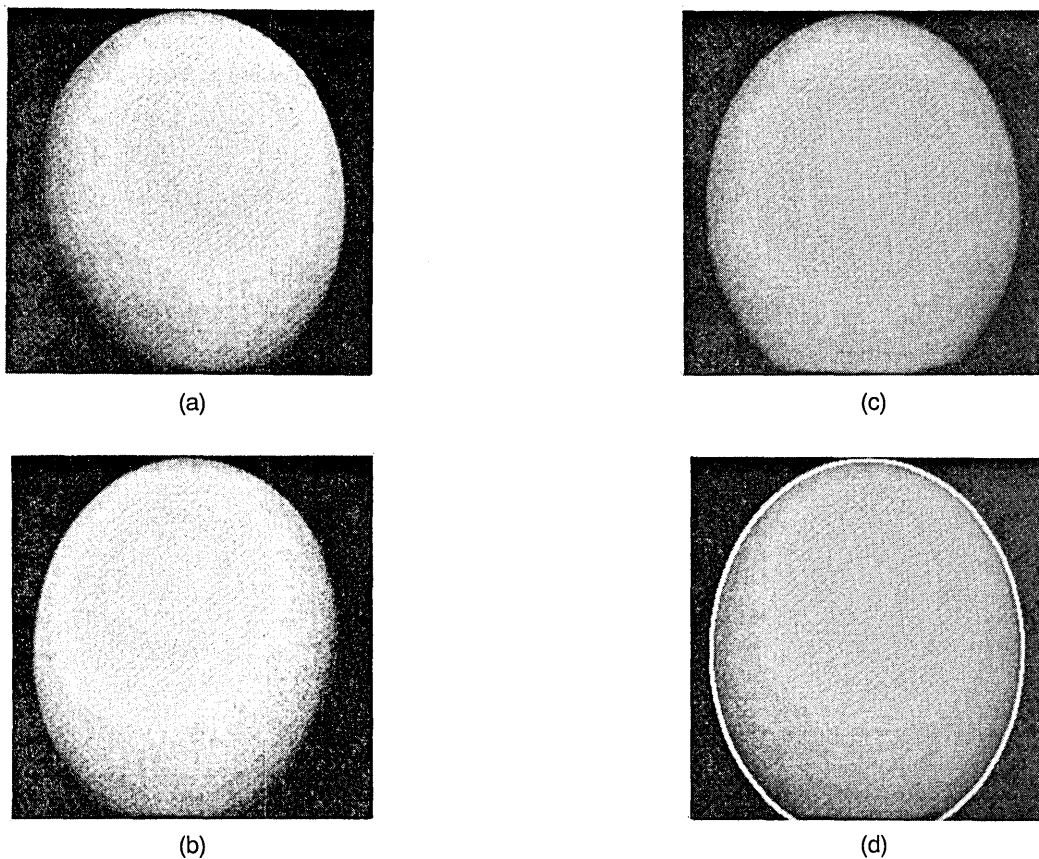
(a)

(c)

(b)

(d)

Fig. 3.   Images of calibration sphere.   (a) Light source 1, (b) light source 2, (c) light source 3, (d) boundary contour overlay.

(p, q) because the gradient is required for determination of the appropriate value of the matrix $\mathbf{R}$ for Eq. (9). Nevertheless, the matrix $\mathbf{R}$ tends to be robust with respect to errors in the gradient (p, q) because, except in regions of highlight or specularity, the error in $[R_p, R_q]$ tends to be small for a given error in (p, q).

Given that estimation of the Hessian is overdetermined, it also becomes possible to detect locations where curvature estimation is unreliable.   In previous research[15] two approaches were described.   First, recall that symmetry of the Hessian $\mathbf{H}$ corresponds to the smoothness (i.e., integrability) of the underlying surface, $z = f(x, y)$.   When Eq. (9) is applied, it is unlikely that the resulting estimate of the Hessian $\hat{\mathbf{H}}$ is exactly symmetric.   One forces symmetry by projecting $\hat{\mathbf{H}}$ onto the so-called symmetric part of $\hat{\mathbf{H}}$, given by

$$\frac{\hat{\mathbf{H}} + \hat{\mathbf{H}}^T}{2}, \qquad (10)$$

prior to estimating the curvature matrix $\mathbf{C}$ with Eq. (4). One can test the assumption that the surface is locally integrable by comparing the norm of the symmetric part of $\hat{\mathbf{H}}$ with that of the antisymmetric part, given by

$$\frac{\hat{\mathbf{H}} - \hat{\mathbf{H}}^T}{2}.$$

Second, one can determine how well the estimated Hessian $\hat{\mathbf{H}}$ accounts for the measured intensity gradients $[E_{ix}, E_{iy}]$, where $i = 1, 2, 3$, based on the error matrix, $\mathbf{E} = [e_{ij}]$, defined by

$$\mathbf{E} = \begin{bmatrix} E_{1x} & E_{2x} & E_{3x} \\ E_{1y} & E_{2y} & E_{3y} \end{bmatrix} - \hat{\mathbf{H}} \begin{bmatrix} R_{1p} & R_{2p} & R_{3p} \\ R_{1q} & R_{2q} & R_{3q} \end{bmatrix}.$$

3.   IMPLEMENTATION AND EXPERIMENTAL RESULTS

A.   Experimental Setting

A calibrated imaging facility was built to control both scene parameters and conditions of imaging.   It is based on a 4 ft × 8 ft (1.22 m × 2.44 m) optical bench with mounting hardware for controlled positioning and motion of cameras, light sources, and test objects.   Specialized equipment includes a Sony DXC-755 3 CCD 24-bit RGB camera with Fujicon 10–120-mm (manual) zoom lens; three Newport MP-1000 Moire (white-light) projectors with associated Nikon lenses and spectral filters; two Daedal rail tables [one 36 in. (91.44 cm) and the other 6 in. (15.24 cm)] and one Daedal rotary table (6 in.); and associated controllers, motors, mounting hardware, and power supplies.   The facility is well integrated with the other vision and robotics equipment in the University of British Columbia Laboratory for Computational Intelligence, including a Datacube image-processing system consisting of DigiColor and MaxVideo-200 subsystems.

Research on photometric stereo and related research on multiple-light-source optical flow[37] requires multiple images of a scene acquired simultaneously under different conditions of illumination.   One achieves this requirement by multiplexing the spectral dimension.   With appropriate filtration the three projectors become spectrally distinct lights sources, one red, one green, and one blue.

The three color-separation filters used are the Newport FS-225 set. The filters are manufactured by Corion (Holliston, Mass.) and are Corion parts CA500 (blue), CA550 (green), and CA600 (red). The projectors illuminate the workspace from different directions. The Sony 3 CCD RGB camera is used to acquire three separate B&W images simultaneously, with each image corresponding to a different condition of illumination.

Care was taken to ensure that the equipment achieves its intended functionality. The light sources and the associated lenses are rated to produce an illumination field uniform to within ±10% over half the spot diameter and to within ±15% over the full spot diameter. The light sources also are dc powered to eliminate the effects of 60-Hz ac line flicker. When experiments are in progress the calibrated imaging facility is enclosed by floor-to-ceiling black curtains, thus isolating it from other light sources in the laboratory.

The precise spectral response of each of the filters was measured by the manufacturer. There is negligible overlap in the visible spectrum between the red-light source and either the green-light source or the blue-light source. There is a small overlap between the green- and the blue-light sources for wavelengths in the 500–520-nm range. Clearly, if either the camera's green channel or its blue channel is sensitive to this common region, there will be some overlap between the images acquired. Indeed, if the camera's spectral responses are quite broad, overlap is possible even if the light sources themselves do not overlap spectrally. Unfortunately the precise spectral response of camera is not provided by the manufacturer, nor was it measured. Instead a simple test was performed to estimate the response of the RGB-camera channels in the given experimental setting. Three RGB video frames were acquired for a test scene consisting of a simple white object. In frame 1 only the red-light source was on, in frame 2 only the green-light source was on, and in frame 3 only the blue-light source was on. The correlation coefficient between the illuminated channel and the other two channels was determined for each frame. The correlation coefficient is a measure of the linear relationship between the two channels. One useful interpretation is that the square of the correlation coefficient is the proportion of the variance accounted for by linear regression. Let $X \rightarrow Y$ denote the influence of light source X on camera channel Y. In four cases ($R \rightarrow G$, $R \rightarrow B$, $B \rightarrow R$, and $G \rightarrow B$) the variance accounted for was less than 1%, indicating excellent spectral separation. In the remaining two cases the variance accounted for was higher, indicating some spectral overlap. The result was 3.5% for $G \rightarrow R$ and 6.5% for $B \rightarrow G$. The slight effect of this overlap is noted in the examples given below.

Two objects are used in the experiments reported. One is a pottery sphere, used for calibration purposes, and the other is a pottery doll face. In this case both objects are made of the same material with the same reflectance properties. Pottery, in bisque form, is a reasonably diffuse reflector, although no particular assumption is made (or required) concerning the underlying surface reflectance function. Other objects made of other materials were used in experiments too. For each material a different calibration sphere is required. In some cases paint was used to achieve the effect of having a calibra-

tion sphere and an object both made of the same material. Although many different objects were tested and many experiments were run, the examples used here, following calibration, all come from an eleven-frame sequence during which the doll face was rotated exactly 3° (about the vertical axis) between successive frames.

B. Calibration

Calibration measures reflectance data by use of an object of known shape. These measurements support the analysis of other objects, provided that the other objects are made of the same material and are illuminated and are viewed under the same imaging conditions. Empirical calibration has the added benefit of automatically compensating for the transfer characteristics of the sensor. Ideally the calibration object is convex, to eliminate interreflection, and has visible surface points spanning the full range of gradients (p, q). A sphere is a good choice and is the calibration shape used here. For a sphere, it is straightforward to determine the gradient (p, q) and the Hessian matrix $\mathbf{H}$ at each visible surface point by geometric analysis of the object's boundary contour. Figures 3(a)–3(c) show the three images of the calibration sphere obtained from the three different light-source directions. The sphere appears as an ellipse because camera calibration, and in particular aspect ratio correction, has not been applied. Nevertheless, the object's boundary contour is easily determined (and becomes part of camera geometric calibration). Let λ denote the relative scaling (i.e., the aspect ratio) of y compared with
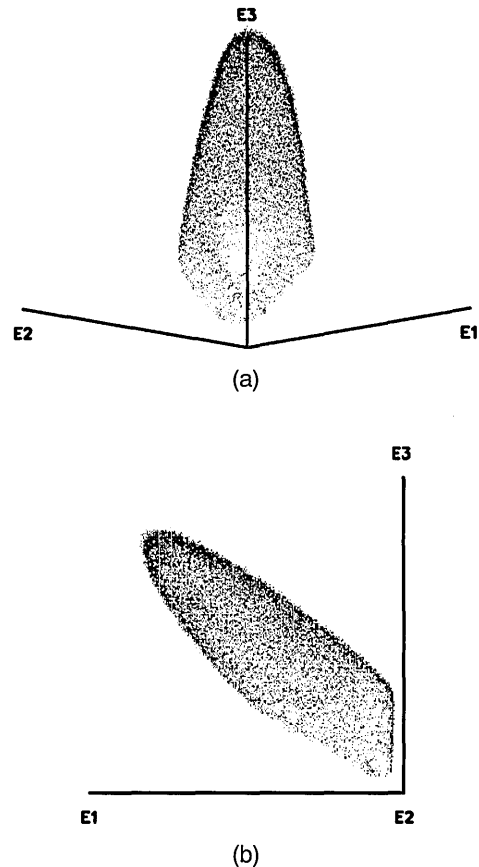

(a)

(b)

Fig. 4.   Calibration sphere: scatter plot of measured intensity triples. (a), (b) Same as in Fig. 1.
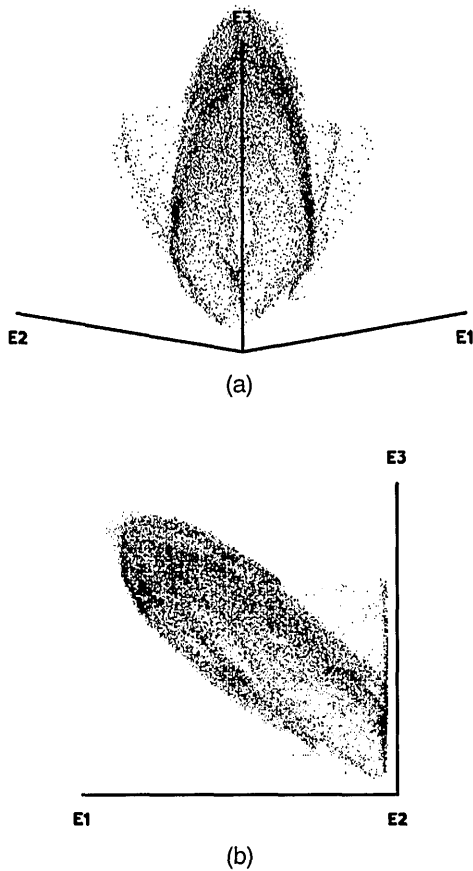
Fig. 5. Doll face (frame 5): scatter plot of measured intensity triples. (a), (b) Same as in Fig. 1.

x. Then the equation of the sphere centered at the origin with radius *r* is

$$x^2 + (\lambda y)^2 + z^2 = r^2 .$$

Again, the standard geometry of shape-from-shading is assumed. That is, the object is defined in a left-handed Euclidean coordinate system in which the viewer is looking in the positive *Z* direction, the image projection is orthographic, and the image *XY* axes coincide with the object *XY* axes. Then the gradient and the Hessian at point (x, y) are, respectively,

$$p = -\frac{x}{z}, \qquad q = -\frac{\lambda^2 y}{z},$$

$$H = -\frac{1}{z^3} \begin{bmatrix} r^2 - \lambda^2 y^2 & \lambda^2 xy \\ \lambda^2 xy & \lambda^2(r^2 - x^2) \end{bmatrix}.$$

Processing the calibration images involves three steps. First, the three images are summed. This is done to ensure that no part of the object's boundary is missed because it lies in shadow. Second, an intensity histogram of the sum image is computed, and a threshold is selected to separate object from background. Simple thresholding is sufficient because, by design, the object is distinct from the black background (i.e., the histogram is clearly bimodal). Third, a simple least-squares method is used to estimate the equation of the ellipse that best fits the object boundary. The equation of the ellipse establishes

the relation between (x, y) and (p, q) for all points on the calibration object (and, as a side effect, determines λ).

Estimation of the ellipse is robust because all boundary points contribute. Figure 3(d) shows the fitted ellipse overlaid on the sum image. (A small portion near the bottom of the ellipse does not figure in the calculation because it corresponds to a region of the sphere obscured by the fixture used to hold the object.) Figure 3 represents a 256 × 256 subwindow of the full 512 × 480 video frame. A total of 566 boundary points contribute to estimation of the ellipse. As a measure of the goodness of fit, the perpendicular distance of each boundary point from the ellipse was calculated. The mean perpendicular distance was 0.429 pixel, and the standard deviation was 0.367.

Figure 4 shows the scatter plot obtained for the calibration sphere. The intensity triples $[E_1, E_2, E_3]$ lie on a single 2D surface. (For comparison purposes see Fig. 5, the scatter plot for frame 5 of the doll-face sequence. The comparison is discussed in Subsection 3.F.) If one chose to force a constant albedo Lambertian model onto this example, one could fit an ellipsoid to the data shown in Fig. 4. This was done as an exercise, and, although the details are not reported here, it is fair to conclude that the result is inferior to the nonparametric approach, which makes no Lambertian assumption and requires no radiometric correction for the transfer characteristics of the sensor.

C. Lookup Tables for Real-Time Photometric Stereo

Lookup-table construction consists of three parts. First, the initial mapping between triples of measured intensity values $[E_1, E_2, E_3]$ and surface orientation is established. Second, table interpolation is used to ensure that the 2D surface defined in the 3D space of measured intensity triples is simply connected. Third, morphological dilation (i.e., expansion) of the table is used to generate surface orientation estimates for intensity triples $[E_1, E_2, E_3]$ that do not fall directly on the 2D calibration surface. As part of lookup table expansion, a distance measure is recorded that determines how far each new point is from a direct table hit. This distance measure is, in turn, used to define the local confidence estimate.

For real-time implementation the lookup table is of dimension $2^6 \times 2^6 \times 2^6 = 2^{18}$. During calibration, surface
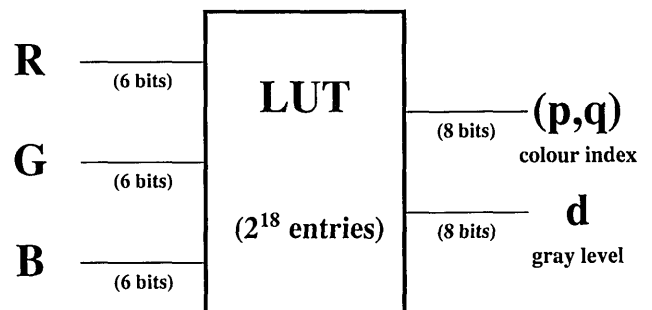

Fig. 6. Lookup tables for photometric stereo. The input combines three video streams, 6 bits each for R, G, and B. The output is two independent 8-bit video streams, one the color-map index for the encoding of the gradient (p, q) and the other the gray-level encoding of distance *d* from a direct table hit. The available hardware does not have a lookup table with 2^{18} entries. Instead, the above procedure is implemented with four hardware lookup tables, each with 2^{16} entries.
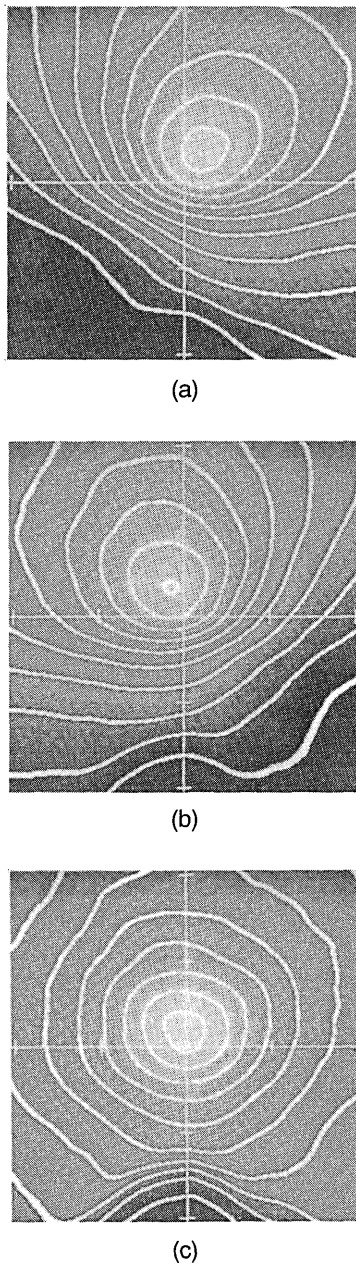
(a)

(b)

(c)

Fig. 7. Three empirically determined reflectance maps with isobrightness contours superimposed. (a) Light source 1, (b) light source 2, (c) light source 3.

orientation is represented by a unit surface normal rather than by the gradient. Each pixel location on the calibration object is sampled. One constructs the intensity triple $[E_1, E_2, E_3]$ by taking the high-order 6 bits from each of the 8-bit values for $E_1(x, y)$, $E_2(x, y)$, and $E_3(x, y)$. The result defines an 18-bit lookup-table index. The unit surface normal corresponding to $[E_1, E_2, E_3]$ is calculated with the equation of the fitted ellipse. The unit surface normal is arithmetically added to the table. Summing unit surface normals averages the triples $[E_1, E_2, E_3]$ that occur more than once on the calibration object. (As a postprocessing step, the summed surface normal vectors are renormalized to unit vectors.) This completes part 1 of lookup-table construction.

It also is likely that there are gaps in the lookup table. This happens, for example, when intensity varies rapidly

enough that the distance between neighboring triples $[E_1, E_2, E_3]$ is greater than 1 in either the E_1, the E_2, or the E_3 dimension. In part 2 of lookup-table construction these gaps are detected, and intermediate table entries are interpolated by subpixel resampling of the calibration images. At the end of part 2 the 2D calibration surface is simply connected, and all table entries are deemed distance $d = 0$ (i.e., direct hit) entries.

In part 3 the lookup table is iteratively expanded n times to extrapolate entries at distances $d = 1, 2, \ldots, n$ from the 2D calibration surface. Let $[i, j, k]$ be a table index at the dth iteration that has no current entry (i.e., no surface normal assigned). The six neighboring points, $[i \pm 1, j \pm 1, k \pm 1]$, are examined. If exactly one of the six neighbors is a table entry, then $[i, j, k]$ is added to the table as a new distance d entry with surface orientation equal to that of the neighbor. (If more than one of the neighbors are table entries, then $[i, j, k]$ is added to the table as a new distance d entry, with surface orientation being a weighted average of the surface normals of those neighbors in which the weights are inversely proportional to the distances assigned to those neighbors.)

If iterative expansion were allowed to run to completion, then every table entry would record a unit surface normal and a distance d corresponding to the iteration at
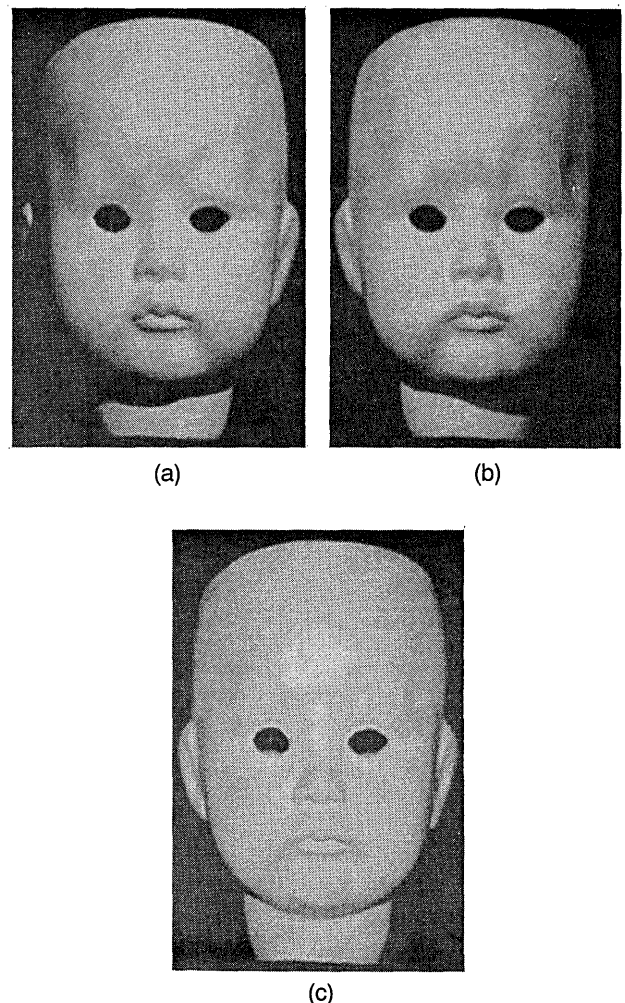

(a)   (b)

(c)

Fig. 8. Three images of doll face (frame 5). (a)–(c) Same as in Fig. 7.
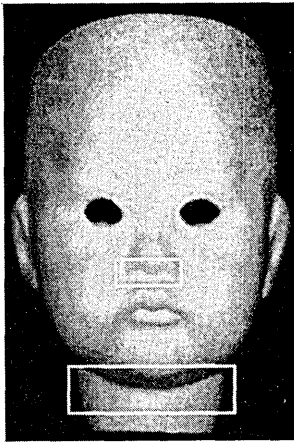
Fig. 9. Selected regions from doll-face (frame 5) light source 3 image. The neck region has cast shadows in the light source 1 and the light source 2 images [see Figs. 8(a) and 8(b)]. The nostril region has significant interreflection.
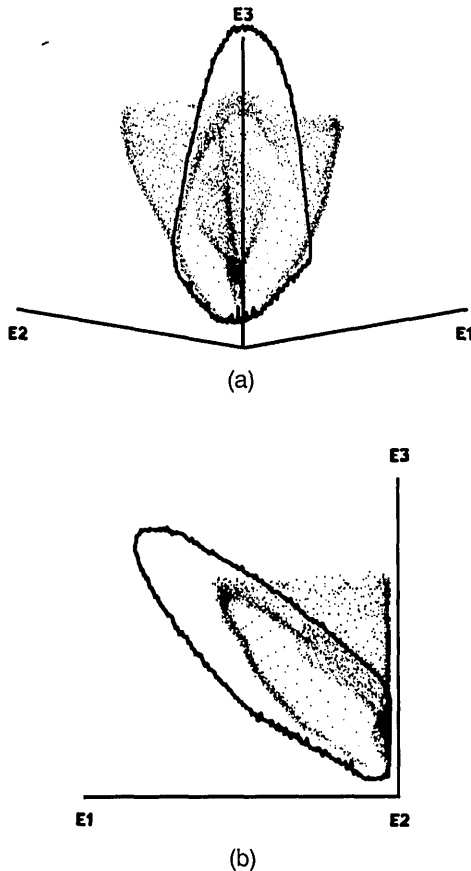

(a)

(b)

Fig. 10. Doll face (frame 5): scatter plot of measured intensity triples taken from the neck region (see Fig. 9) in which cast shadows occur. The outline of the calibration sphere scatter plot is overlaid for comparison purposes. (a) Elevation 10.0, azimuth 225.0; (b) elevation 0.0, azimuth 0.0.

which the surface normal was assigned. In the experiments reported below, $n = 10$ iterations were performed. Intensity triples that have no table entry become points at which photometric stereo assigns no gradient. A bitmap file can be produced to mark those points at which no gradient (p, q) was assigned.

Figure 6 is a conceptual diagram of how the near-real-time implementation is achieved on the Datacube system.

One converts the 24-bit RGB camera output to a $2^6 \times 2^6 \times 2^6 = 2^{18}$ table index by combining the high-order 6 bits of R, G, and B. The lookup table supports two independent 8-bit output streams. In the mode of operation depicted in Fig. 6, one 8-bit stream is an index to a color map used to render the gradient (p, q), and the other 8-bit stream is the gray value representing the distance d. If things were indeed this simple, the implementation would be straightforward. Unfortunately the Datacube MaxVideo-200 does not have a lookup table of length 2^{18}, as depicted. It does, however, have four lookup tables of length 2^{16}. In the actual implementation the high-order 2 bits of the R image are used as a selector to determine which of the four lookup tables of length 2^{16} to use. The MaxVideo-200 operates internally at field rates (i.e., 60 Hz). It is possible to split the incoming video stream, to pass it through the four lookup tables, and to recombine it in a total of four field times leading to an overall throughput (on a 512×480 video frame) of 15 Hz.

D. Determining the Reflectance Maps

Gradient estimation by means of table lookup does not determine the reflectance maps explicitly. But curvature estimation requires their partial derivatives, with respect to p and q, for determination of the matrix **R** of Eq. (9). It is a simple matter to invert the calibra-
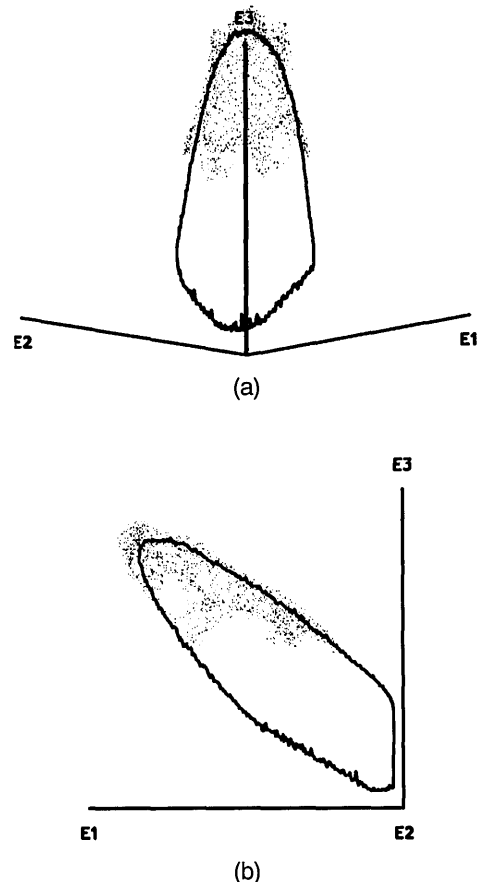

(a)

(b)

Fig. 11. Doll face (frame 5): scatter plot of measured intensity triples taken from the nose region (see Fig. 9) in which significant interreflection occurs. The outline of the calibration sphere scatter plot is overlaid for comparison purposes. (a), (b) Same as in Fig. 10.

tion object gradient equations to determine the image point (x, y) at which to obtain intensity measurements for any particular (p, q). In this way the three reflectance maps $R_i(p, q)$, where $i = 1, 2, 3$, are interpolated from the calibration images. In the implementation the reflectance maps are stored explicitly as arrays for a given domain of p and q and for a given grid spacing. The partial derivatives are computed from the interpolated reflectance maps, on demand. Figure 7 shows the three reflectance maps (with isobrightness contours superimposed for illustration purposes). Tick marks on the (horizontal) p axis and on the (vertical) q axis are plotted one unit apart. Thus the domain covered is approximately $-2.0 \le p \le 2.0$ and $-2.0 \le q \le 2.0$. Because the distance from the origin in gradient space, $\sqrt{p^2 + q^2}$, is the tangent of the slope angle, the domain covered includes all visible object points with slope relative to the image plane less than $\tan^{-1}(2.0) = 63.4°$. (Again, a small region that corresponds to the region of the sphere obscured by the fixture used to mount it for viewing is missing.)

E. Determining Surface Curvature

To determine surface curvature at a point (x, y), we must measure the six partial spatial derivatives, E_{ix}, E_{iy}, where $i = 1, 2, 3$, and we must estimate the gradient (p, q). The reference data are the six partial derivatives, R_{ip}, R_{iq}, where $i = 1, 2, 3$. The gradient (p, q), is obtained by means of lookup table, as described in Subsection 3.A. The reflectance maps $R_i(p, q)$, where $i = 1, 2, 3$, are obtained as described in Subsection 3.D. In the current implementation each image and reflectance map is smoothed with a 2D Gaussian, and the required partial derivatives are estimated with simple local differencing. This computation has not yet been implemented in real time.

At each object point at which the gradient (p, q) is estimated and at which $R_i(p, q)$, where $i = 1, 2, 3$, is defined, Eq. (9) is used to estimate the Hessian matrix $\mathbf{H}$. The resulting estimate, $\hat{\mathbf{H}}$, is made symmetric by means of Eq. (10). The curvature matrix $\mathbf{C}$ is determined with Eq. (4). From the matrix $\mathbf{C}$ the principal curvatures k_1 and k_2, their associated directions ω_1 and ω_2, and other curvature measures are derived, as described in Subsection 2.D. Again, none of these curvature computations has yet been implemented in real time.

F. Experimental Results

For experimental and demonstration purposes, a particular color encoding for the gradient was adopted. Figure 8 shows the three input images for frame 5 of the doll-face sequence. The figure shows a 384×256 subwindow extracted from the full video frame. Light sources 1, 2, and 3 correspond to the red, the green, and the blue illuminants, respectively. Evidence that the blue-light source has some influence on the green channel, but negligible influence on the red, is noted in that the neck shadow region in Fig. 8(b) is not so dark as the corresponding shadow region in Fig. 8(a). Plate 21 shows the corresponding color-encoded gradient as produced in the near-real-time (15-Hz) implementation of photometric stereo. The inset (lower right) shows the color rosette used to encode the gradient. The white lines in the center row and the center column of the rosette represent the p and the q axes, respectively. Angular position about

the origin is encoded as color, and distance from the origin is encoded as brightness. The domain of (p, q) covered is $-1.5 \le p \le 1.5$ and $-1.5 \le q \le 1.5$, so that points with slope less than or equal to $\tan^{-1}(1.5) = 56.3°$ are encoded by distinct colors and brightnesses. The color encoding demonstrates qualitatively the local effectiveness of photometric stereo, as can be seen, for example, around the hollow eye sockets. At the circumference of each eye socket it is evident that the local surface slope and aspect have been recovered.

For integration with other vision modules, it is useful to use the two lookup-table output streams to carry 8 bits each of p and q. This is a trivial modification to what has been described. The necessary information is available from calibration, and the change simply means that a different lookup table is loaded into the MaxVideo-200.

Figure 5 shows the scatter plot obtained for the three images of the doll face (frame 5) shown in Fig. 8. Clearly, it can no longer be said that all intensity triples $[E_1, E_2, E_3]$ lie on a single 2D surface. Even though the doll face is made of the same material and is illuminated in the same way, it is a more complex shape than the calibration sphere. It is nonconvex with regions of cast shadow and regions in which interreflection is significant.

Figure 9 marks two regions for further scrutiny. The neck region has cast shadows in the light source 1 and 2 images [see Figs. 7(a) and 7(b)]. The nostril region has significant local interreflection. Figure 10 shows the
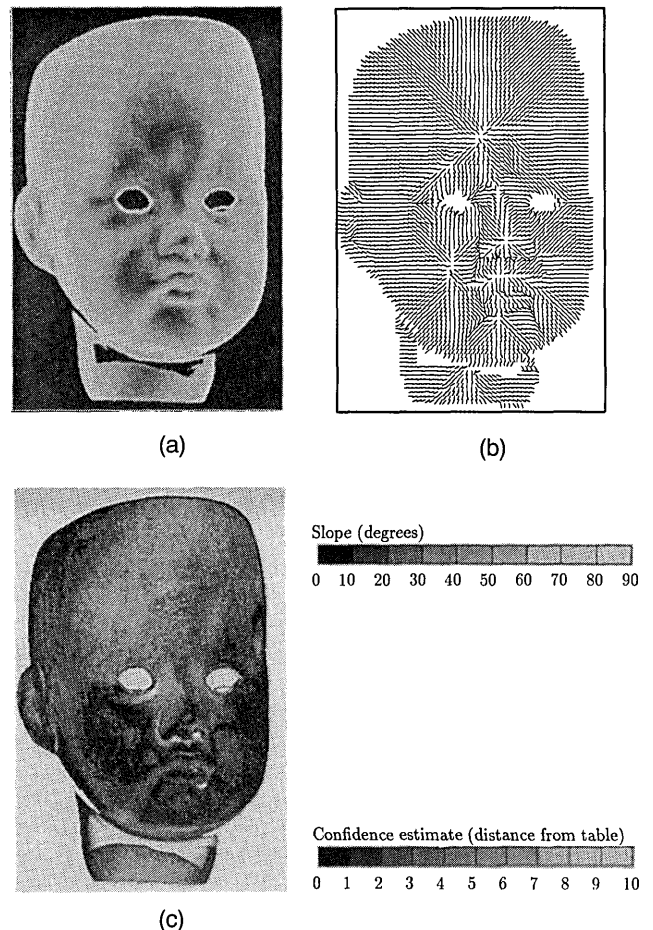

(a)   (b)

Slope (degrees)

0 10 20 30 40 50 60 70 80 90

Confidence estimate (distance from table)

0 1 2 3 4 5 6 7 8 9 10

(c)

Fig. 12. Gradient estimation for doll face (frame 0). (a) Slope, (b) aspect, (c) confidence estimate.

(a)   (b)

Slope (degrees)

0 10 20 30 40 50 60 70 80 90

Confidence estimate (distance from table)

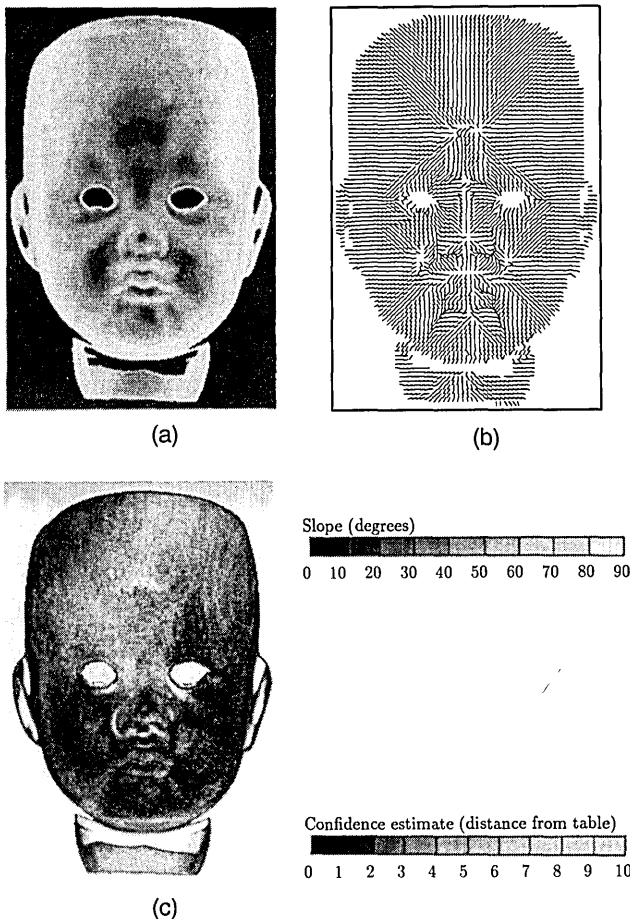0   1   2   3   4   5   6   7   8   9   10

(c)

Fig. 13. Gradient estimation for doll face (frame 5). (a)–(c) Same as in Fig. 12.

scatter plot for the neck region. The outline of the calibration sphere scatter plot from Fig. 4 is overlaid in boldface for comparison purposes. It is clear that many of the points in the scatter plot are outliers with respect to the 2D surface defined by the calibration sphere. In particular, some points are darker in E_1 or E_2 than is consistent with their E_3 value. Similarly, Fig. 11 shows the scatter plot for the nose region. Again, the outline of the calibration sphere scatter plot is overlaid in boldface for comparison purposes. Many of the points in the scatter plot are outliers here too. In particular, many points are brighter than is consistent with the 2D calibration surface.

Figures 12–14 are examples of gradient estimation from frames 0, 5, and 10 of the doll-face sequence. In frame 5 the doll face is oriented directly toward the viewer. In frame 0 it is rotated $3 \times 5 = 15°$ to the left, and in frame 10 it is rotated $3 \times 5 = 15°$ to the right. In these figures the gradient is rendered in B&W. In each figure, (a) encodes the slope angle [i.e., $\tan^{-1}(\sqrt{p^2 + q^2})$] as a gray value, and (b) plots the aspect angle [i.e., $\tan^{-1}(q/p)$] as a short line segment. (To avoid clutter the aspect angle is plotted for every fourth point in x and y.) Slope and aspect are viewpoint-dependent measures. Therefore the values for a given point on the doll face do not remain constant as the object rotates. In each figure, (c) encodes the distance measure d as a gray value. In particular, one can note that, in Fig. 13(c), the d value is

large at many points in the neck and in the nostril region, confirming that these outliers have been detected.

Figures 15–17 are examples of principal curvature estimation from frames 0, 5, and 10 of the doll-face sequence. In each figure, (a) encodes the first principal curvature k_1, the curvature whose magnitude is maximum, and (b) encodes the second principal curvature k_2, the curvature whose magnitude is minimum. Principal curvatures are viewpoint-independent measures. Thus the values for a given point on the doll face should remain constant as the object rotates. In each figure, (c) plots the corresponding principal directions ω_1 (in boldface) and ω_2. (To avoid clutter the principal directions are plotted for every fifth point in x and y.) The principal directions are viewpoint dependent.

4. DISCUSSION AND CONCLUSIONS

Multiple images acquired with the identical viewing geometry but under different conditions of illumination are a principled way to obtain additional local constraint in shape-from-shading. Three-light-source photometric stereo is fast and robust. In the implementation described, surface gradient estimation is achieved on full-frame video data at 15 Hz by use of commercially available hardware. Surface curvature estimation is also demonstrated. Although not yet implemented in real time, curvature estimation also is simple and di-
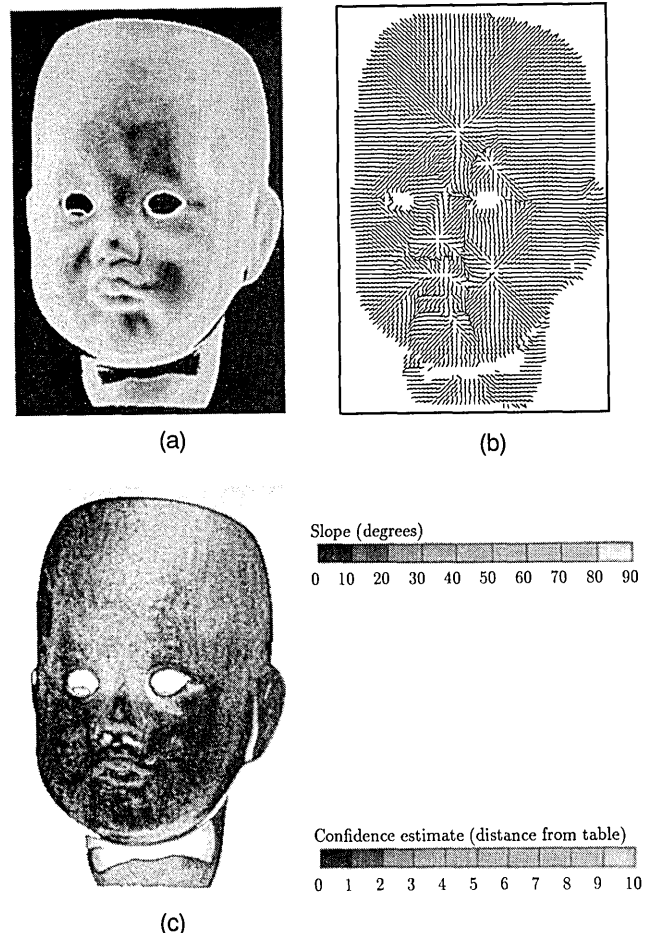

(a)   (b)

Slope (degrees)

0 10 20 30 40 50 60 70 80 90

Confidence estimate (distance from table)

0   1   2   3   4   5   6   7   8   9   10

(c)

Fig. 14. Gradient estimation for doll face (frame 10). (a)–(c) Same as in Fig. 12.

(a)

(b)

Curvature, k, per pixel ($\times 10^{-2}$)

-10 -8 -6 -4 -2 0 2 4 6 8 10

-10 -12.5 -16.7 -25 -50 ∞ 50 25 16.7 12.5 10
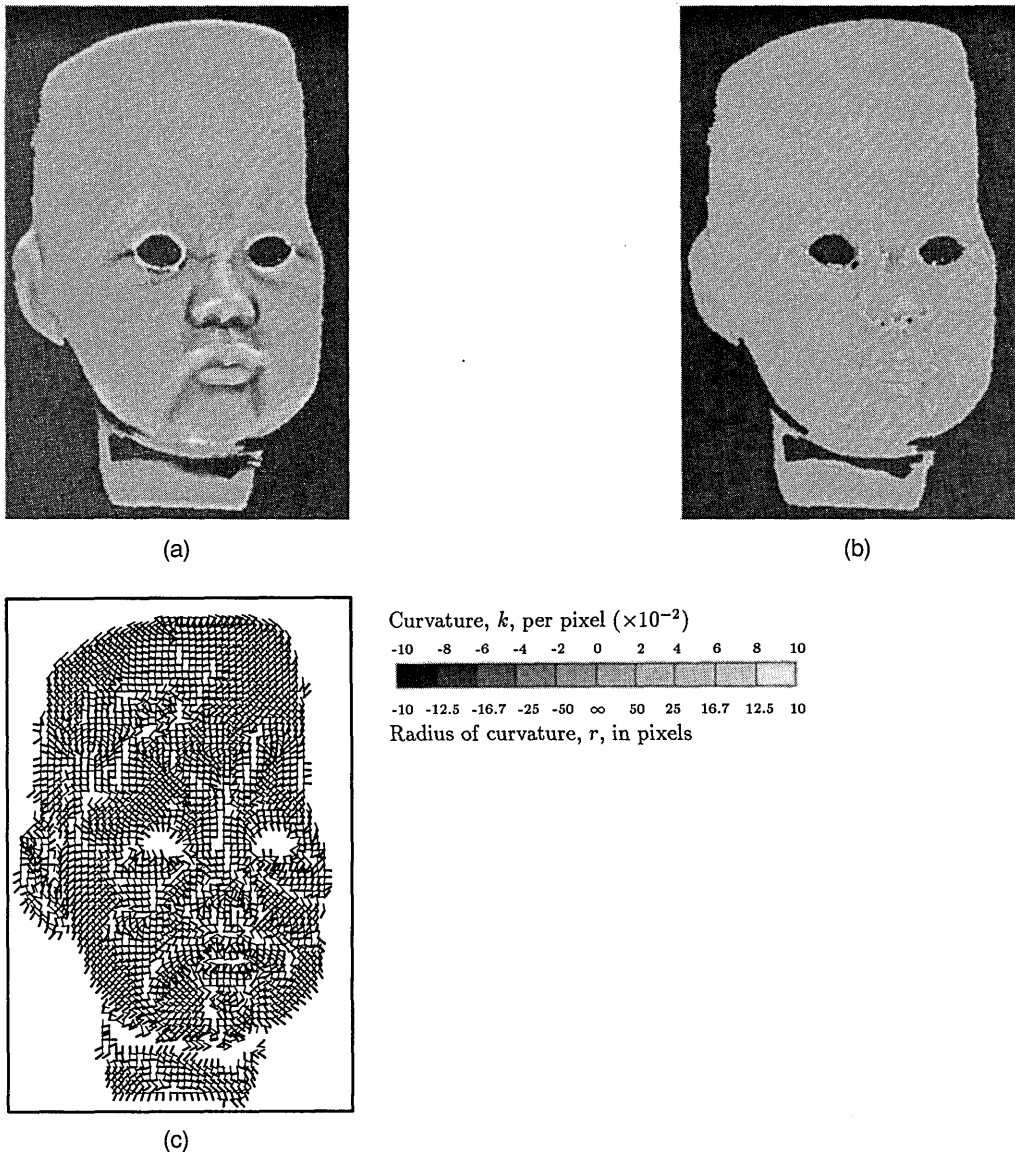Radius of curvature, r, in pixels

(c)

Fig. 15. Principal curvature estimation for doll face (frame 0). (a) Principal curvature k_1, (b) principal curvature k_2, (c) principal directions.

rect, requiring no iteration steps. Overall, the computational requirements for photometric stereo are minimal compared with the iterative schemes typically required for shape-from-shading from a single image. The main practical limits on system performance are data storage, required by calibration, and data throughput, required for processing multiple images, including spatial derivatives, simultaneously.

One practical benefit of real-time implementation is the ability to integrate photometric stereo and motion (based on related research on multiple-light-source optical flow[37]). A high-speed camera-shutter setting ensures that all surface points are imaged at essentially the same instant in time. Real-time implementation of both photometric stereo and optical flow can then determine the 3D structure and the 3D motion of deforming objects. Another benefit relates to tasks in which it may be neither possible nor appropriate to use alternatives, such as laser range sensors. For example, photometric stereo is being evaluated for a medical application involving the acquisi-

tion of 3D models of children's faces (used to plan subsequent medical procedures). Here data acquisition must not harm the patient (so the use of laser ranging is problematic). Because children typically do not remain stationary, data acquisition also must be rapid.

Photometric stereo achieves robustness in several ways. The key is to overdetermine the problem locally. Overdetermination provides noise resilience and protection against local degeneracies. To specify the local properties of a surface up to curvature requires six parameters because there is 1 degree of freedom for range, 2 for surface orientation, and 3 for curvature. If only a single measurement, say, range, is available locally, then the problem is locally underconstrained. The only solution then is to reconstruct a smooth surface globally by combining measurements obtained over extended regions. In three-light-source photometric stereo each image provides three independent pieces of local information, one for intensity and two for the partial spatial derivatives of intensity. (For the information to be truly independent,

(a)

(b)

(c)

Curvature, k, per pixel $(\times 10^{-2})$

-10 -8 -6 -4 -2 0 2 4 6 8 10

-10 -12.5 -16.7 -25 -50 ∞ 50 25 16.7 12.5 10
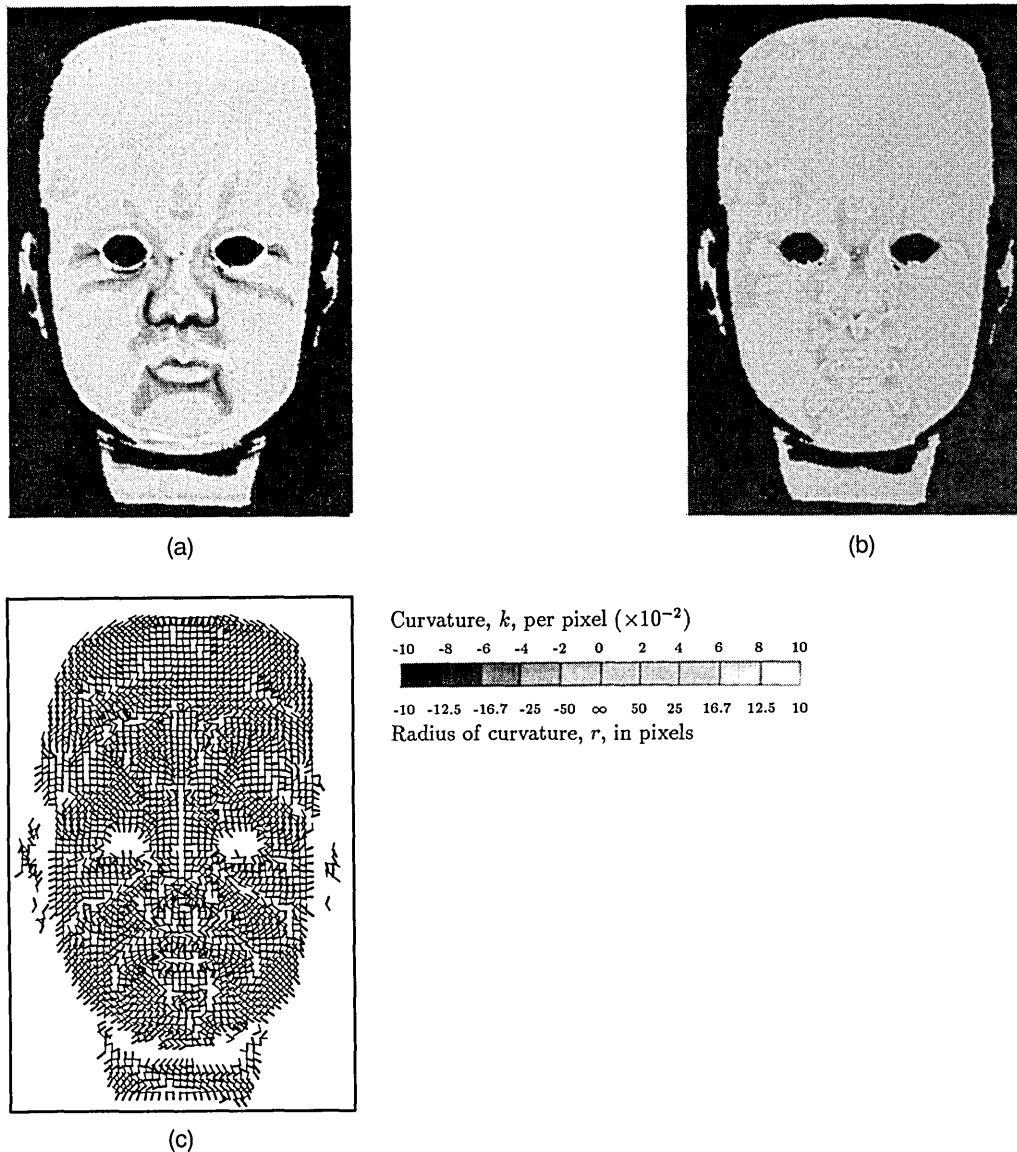Radius of curvature, r, in pixels

Fig. 16.   Principal curvature estimation for doll face (frame 5).   (a)–(c) Same as in Fig. 15.

one would require an image sensor that measured partial derivatives directly.) Thus, with three images, one obtains nine local measurements to overdetermine the five unknowns associated with orientation and curvature. At the implementation level gradient estimation is robust because 18 bits of RGB input are used to estimate 16 bits (8 bits each for p and q) of output.

Overdetermination also supports local detection of modeling errors and other inconsistencies. A (nonparametric) empirical approach to reflectance modeling eliminates errors that arise when the experimental situation does not satisfy assumptions implicit in parametric models. It also eliminates the need to estimate the unknown parameters. For example, in the research described, one need never estimate the directions to the light sources or their relative amplitudes. The empirical approach has the added benefit of automatically compensating for the transfer characteristics of the sensor. It also means that the system is robust to possible spectral overlap in the three color channels used. Indeed, complete spectral

separation is not essential. At a higher level robustness is achieved because an attempt is made to use all the information available in an image, not just that obtained from a sparse set of features.

The claim that photometric stereo is accurate has not been dealt with quantitatively. A careful assessment of accuracy, including comparison with laser range sensors, is an essential next step. Proper error analysis, however, is nontrivial. Issues involved include camera calibration (geometric and radiometric), method of integration (if comparison is made between depth values), range sensor calibration (if range data serve as ground truth), and method of differentiation (if comparison is made between gradients and differentiated range data).

Photometric stereo appears to be a competitive technique for a variety of tasks. One class of task is 3D model acquisition as is required, for example, in computer graphics, CAD/CAM analysis, and rapid prototyping. For this class of task the accuracy of the reconstruction of the surface height function $z = f(x, y)$ is central. But surface

reconstruction itself begs important questions. Standard schemes combine a data fit term with a global smoothness term. Smoothers typically used are viewpoint dependent (Stevenson and Delp[38] is a notable exception). Reconstruction also requires specification of initial boundary conditions. In practice the results of reconstruction tend to be dominated by the choice of smoother and by errors in the initial conditions. Given this factor, it is not clear what is the right approach to surface reconstruction when dense, accurate local orientation and curvature data are available.

Another class of task includes 3D object recognition, localization, and inspection, as is required, for example, in industrial automation. Photometric stereo has been used for object recognition and object localization in ways that do not reconstruct surface height, $z = f(x, y)$. In particular, Li[9] developed a system to determine the 3D attitude of known objects based on dense orientation and curvature data determined by photometric stereo. Her

test objects were precision machined so that accuracy of attitude determination could be assessed. In the end, surface reconstruction may not be the *sine qua non* of shape-from-shading methods.

As shapes treated by machine vision and robotics systems become more complex, segmentation based on surface orientation and curvature becomes more important. Segmentation has always been a chicken-and-egg problem in computer vision. Photometric stereo with three or more light sources allows local surface orientation and curvature to be reliably estimated prior to segmentation. Also, the redundancy in three-light-source photometric stereo makes it possible to detect local inconsistencies that arise, for example, because of cast shadows and interreflection. Detection is facilitated by explicit inclusion of a local confidence estimate in the lookup table used for gradient estimation. The effective interaction between local estimation of surface properties, including local error detection, and global surface reconstruction remains

(a)

(b)

(c)

Curvature, k, per pixel ($\times 10^{-2}$)

| -10 | -8 | -6 | -4 | -2 | 0 | 2 | 4 | 6 | 8 | 10 |

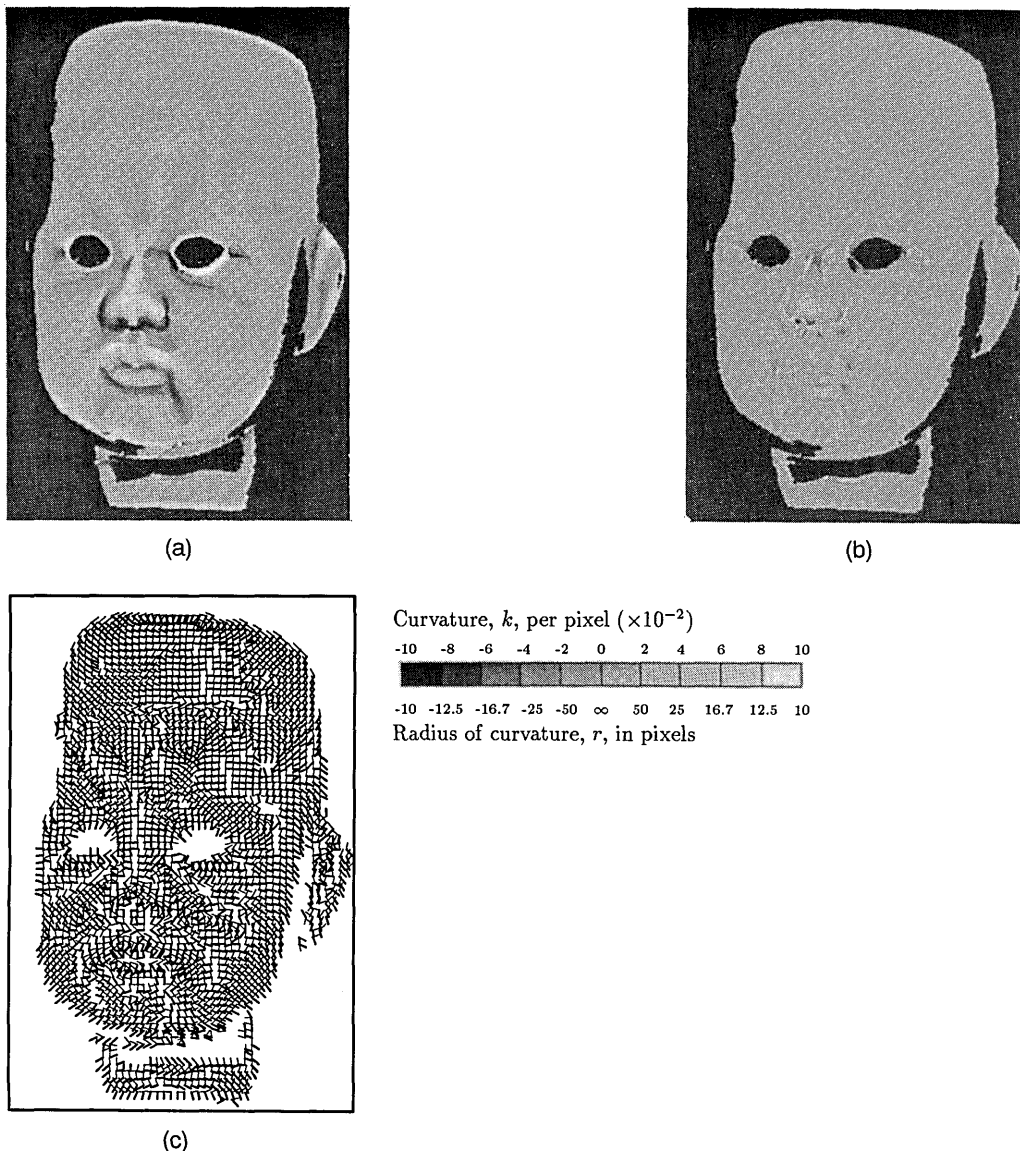| -10 | -12.5 | -16.7 | -25 | -50 | ∞ | 50 | 25 | 16.7 | 12.5 | 10 |

Radius of curvature, r, in pixels

Fig. 17. Principal curvature estimation for doll face (frame 10). (a)–(c) Same as in Fig. 15.

to be explored. The hope is that the present study will allow future segmentation and integration schemes to be more robust.

APPENDIX A: LAMBERTIAN CASE (CONSTANT ALBEDO)

This appendix reconsiders the case of three-light-source photometric stereo under the assumption of Lambertian reflectance and constant albedo. It is shown that, in the absence of interreflection and cast shadows, triples of measured intensity values determine a 6-degree-of-freedom ellipsoid. The ellipsoid characterizes the strengths and the relative positions of the light sources.

The equation characterizing image irradiance for Lambertian reflectance, distant-point-source illumination, orthographic projection, and transmittance through an intervening scatterless medium is

$$E(x,y) = \frac{E_0}{\pi} \rho(x,y)\cos(\theta_i), \qquad (A1)$$

where $E(x,y)$ is the measured irradiance at image point (x,y), $\rho(x,y)$ is the associated bidirectional reflectance factor (i.e., albedo), E_0 is the irradiance of the light source, and θ_i is the incident angle.

The original paper on photometric stereo[2] included a formulation to recover both surface gradient (p,q) and surface reflectance $\rho(x,y)$ under the assumptions of orthographic projection, three distant point sources of illumination, and Lambertian reflectance. This is the basis for pseudoshape recovery, as defined in Ref. 24. To apply this result, however, it is necessary that the three light sources be in a known configuration and of known strength.

Here it is assumed that the directions to, and the relative strengths of, the light sources are not known. Estimation of these parameters becomes part of the problem formulation. Instead, it is assumed that $\rho(x,y)$ is constant at all object points of interest, so the dependence of ρ on (x,y) can be ignored. It follows, without loss of generality, that the scale factor in Eq. (A1) can be taken to be equal to 1, so the image irradiance equation becomes the more familiar

$$E(x,y) = \cos(\theta_i).$$

But we want the relative strengths of the light sources to be distinct. Therefore we write

$$E(x,y) = E \cos(\theta_i), \qquad (A2)$$

where scalar parameter E characterizes the relative strength of the light source.

Directions can be represented by unit vectors, so the cosine of the angle between any two directions is the dot product of the corresponding two unit vectors. This allows the Lambertian case to be formulated as a linear problem. For three-light-source photometric stereo, let $\mathbf{a}_i = [a_{i1}, a_{i2}, a_{i3}]$, where $i = 1,2,3$, be the 1×3 (row) vectors that point in the direction of light source i with magnitude equal to the relative strength E_i of light source i. Let $\mathbf{A}$ be the 3×3 matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Assume that the three light-source directions, given by $\mathbf{a}_i$, where $i = 1,2,3$, are not coplanar, so that the matrix $\mathbf{A}$ is nonsingular.

Let $\mathbf{x} = [x_1, x_2, x_3]^T$ be the unit (column) surface normal vector at some object point of interest. Let $\mathbf{y} = [y_1, y_2, y_3]^T$ be the associated triple of intensity values given by Eq. (A2), applied once for each light-source direction. Then we obtain

$$\mathbf{y} = \mathbf{Ax}. \qquad (A3)$$

Equation (A3) establishes a linear relation between surface shape, given by the unit surface normal vector $\mathbf{x}$, and measured intensity values $\mathbf{y}$.

Of course, if we knew the value of $\mathbf{A}$ then we could determine $\mathbf{x}$ as

$$\mathbf{x} = \mathbf{By}, \qquad (A4)$$

where $\mathbf{B} = \mathbf{A}^{-1}$. Here, however, we do not assume that $\mathbf{A}$ is known. Fortunately, there is more that we can say simply based on the observation that Eq. (A3) is linear.

Consider each unit vector $\mathbf{x}$ to be positioned at the origin. We can then associate all vectors $\mathbf{x}$ with points on the unit sphere centered at the origin. In this way we can think of Eq. (A3) as specifying a linear transformation of the sphere $\|\mathbf{x}\|_2 = (\mathbf{x}^T\mathbf{x})^{1/2} = 1$. It is reasonable to ask, What is the corresponding shape defined by the vectors $\mathbf{y} = \mathbf{Ax}$?

Substitution with Eq. (A4) shows that $\mathbf{x}^T\mathbf{x} = 1$ implies that

$$(\mathbf{By})^T\mathbf{By} = \mathbf{y}^T\mathbf{B}^T\mathbf{By} = \mathbf{y}^T\mathbf{Cy} = 1,$$

where $\mathbf{C} = \mathbf{B}^T\mathbf{B}$ is the 3×3 symmetric positive-definite matrix

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1^T\mathbf{b}_1 & \mathbf{b}_1^T\mathbf{b}_2 & \mathbf{b}_1^T\mathbf{b}_3 \\ \mathbf{b}_2^T\mathbf{b}_1 & \mathbf{b}_2^T\mathbf{b}_2 & \mathbf{b}_2^T\mathbf{b}_3 \\ \mathbf{b}_3^T\mathbf{b}_1 & \mathbf{b}_3^T\mathbf{b}_2 & \mathbf{b}_3^T\mathbf{b}_3 \end{bmatrix}$$

and where $\mathbf{b}_i = [b_{1i}, b_{2i}, b_{3i}]^T$, with $i = 1,2,3$, are the three 3×1 column vectors of $\mathbf{B}$.

Suppose that we now measure intensity triples $\mathbf{y}$ from points on an object of unknown shape. Then these intensity triples are constrained to lie on the quadratic surface $\mathbf{y}^T\mathbf{Cy} = 1$. That is, the intensity triples $\mathbf{y}$ satisfy the equation

$$c_{11}y_1{}^2 + c_{22}y_2{}^2 + c_{33}y_3{}^2 + 2c_{12}y_1y_2 + 2c_{13}y_1y_3$$
$$+ 2c_{23}y_2y_3 - 1 = 0. \qquad (A5)$$

This equation has six unknown coefficients. This follows, of course, from the fact that the matrix $\mathbf{C}$, being symmetric, has only 6 degrees of freedom. Equation (A5) necessarily defines an ellipsoid because the matrix $\mathbf{C}$ is positive definite. In particular, $c_{ii} > 0$, where $i = 1,2,3$.

In Ref. 39 a simple least-squares method is used to estimate the six unknown coefficients of matrix $\mathbf{C}$ from scatter plots of measured intensity triples, even when the

matrix $\mathbf{A}$ is unknown and even when the object shape also is unknown. The constraint that $\mathbf{C}$, in turn, imposes on $\mathbf{A}$ is easiest to interpret when expressed in terms of $\mathbf{C}^{-1}$. Let $\mathbf{D} = \mathbf{C}^{-1}$, so that

$$\mathbf{D} = \mathbf{C}^{-1} = (\mathbf{B}^T \mathbf{B})^{-1} = \mathbf{B}^{-1}(\mathbf{B}^T)^{-1} = \mathbf{B}^{-1}(\mathbf{B}^{-1})^T = \mathbf{A}\mathbf{A}^T .$$

Therefore

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1\mathbf{a}_1^T & \mathbf{a}_1\mathbf{a}_2^T & \mathbf{a}_1\mathbf{a}_3^T \\ \mathbf{a}_2\mathbf{a}_1^T & \mathbf{a}_2\mathbf{a}_2^T & \mathbf{a}_2\mathbf{a}_3^T \\ \mathbf{a}_3\mathbf{a}_1^T & \mathbf{a}_3\mathbf{a}_2^T & \mathbf{a}_3\mathbf{a}_3^T \end{bmatrix} .$$

The matrix $\mathbf{D}$, like the matrix $\mathbf{C}$, is a 3×3 symmetric positive-definite matrix. From $\mathbf{D}$ one can determine the relative strengths of the light sources i, where $i = 1, 2, 3$, and the angle between the vectors to light sources i and j, where $i \neq j$, $i = 1, 2, 3$, and $j = 1, 2, 3$. Specifically, the relative strength of light source i, E_i, is given by

$$E_i = \sqrt{\mathbf{a}_i\mathbf{a}_i^T} = \sqrt{d_{ii}}, \qquad (A6)$$

and the cosine of the angle α_{ij}, where $i \neq j$, between $\mathbf{a}_i$ and $\mathbf{a}_j$ is given by

$$\cos(\alpha_{ij}) = \frac{\mathbf{a}_i\mathbf{a}_j^T}{\sqrt{\mathbf{a}_i\mathbf{a}_i^T}\sqrt{\mathbf{a}_j\mathbf{a}_j^T}} = \frac{d_{ij}}{\sqrt{d_{ii}}\sqrt{d_{jj}}} . \qquad (A7)$$

Equations (A6) and (A7) together represent six constraints on the matrix $\mathbf{A}$. These six constraints can be interpreted geometrically. Let the vectors $\mathbf{a}_i$, where $i = 1, 2, 3$, share a common origin. The vectors $\mathbf{a}_i$ form a triad whose shape, specified by the lengths of the vectors and the angles between them, is known. Any rotation of this triad will not change the shape of the triad and therefore will not violate any of the six constraints. A 3D rotation has 3 degrees of freedom. To fix the triad in a given 3D coordinate system absolutely, three additional constraints would be required.

A simple mathematical argument demonstrates that the ellipsoid $\mathbf{y}^T\mathbf{C}\mathbf{y} = 1$ is indeed invariant under a rotation of the coordinate system used to represent $\mathbf{x}$. Let $\mathbf{R}$ be an arbitrary 3×3 rotation matrix. Consider rotating the unit surface normals by $\mathbf{R}$. That is, let $\hat{\mathbf{x}} = \mathbf{R}\mathbf{x}$. Clearly, the constraint $\hat{\mathbf{x}}^T\hat{\mathbf{x}} = 1$ is preserved because

$$\hat{\mathbf{x}}^T\hat{\mathbf{x}} = (\mathbf{R}\mathbf{x})^T\mathbf{R}\mathbf{x} = \mathbf{x}^T(\mathbf{R}^T\mathbf{R})\mathbf{x} = \mathbf{x}^T\mathbf{x} = 1 .$$

Therefore the corresponding constraint $\mathbf{y}^T\mathbf{C}\mathbf{y} = 1$ also is preserved. It should not be surprising that the ellipsoid $\mathbf{y}^T\mathbf{C}\mathbf{y} = 1$ is invariant under a rotation of the object being viewed because the brightness of a Lambertian reflector is independent of viewpoint.

Finally, there is a generalization to the derivation given here that merits attention. The matrix $\mathbf{A}$ characterizes the directions to, and the relative strengths of, three distant point light sources. It is natural, therefore, to assume that the derivation is valid only when there literally are three distant point light sources. In fact, the result holds more generally as a consequence of another property of Lambertian reflectance. For any Lambertian surface and any spatial distribution of distant illumination, there exists a single distant point source that produces the same

reflectance map for that region of the gradient space not self-shadowed with respect to any part of the illuminant. Silver (Ref. 3, pp. 104–105) provides a formal derivation of this property. The derivation is not repeated here.

Lambertian shading from surface points not shadowed with respect to any part of a spatially distributed distant illuminant is equivalent to that obtained from a single distant-point-source illuminant. Silver's derivation is constructive. Given any spatially distributed distant illuminant, one can determine the equivalent point-source direction and strength. Thus, for Lambertian reflectance, triples of measured intensity values determine a 6-degree-of-freedom ellipsoid, even if one or more of the three images arises from a spatially distributed illuminant. The ellipsoid then characterizes the strengths and the relative positions of the three equivalent distant-point-light sources. Recently, Drew[40] has used this idea to demonstrate that it is possible to recover surface shape from color images of Lambertian surfaces given that a spatially distributed distant illuminant also varies spectrally. Drew also argues that the underlying ellipsoid arising from Lambertian reflectance can be recovered in the presence of specularities, provided that specular points can reliably be detected as outliers. But it is not clear, given the example of Phong reflectance presented in Fig. 2, that specularities always can be effectively treated as outliers to an underlying ellipsoid shape.

ACKNOWLEDGMENTS

REFERENCES

1. R. J. Woodham, "Reflectance map techniques for analyzing surface defects in metal castings," Tech. Rep. AI-TR-457 (Artificial Intelligence Laboratory, MIT, Cambridge, Mass., 1978).
2. R. J. Woodham, "Photometric method for determining surface orientation from multiple images," Opt. Eng. 19, 139–144 (1980).
3. W. M. Silver, "Determining shape and reflectance using multiple images," M. S. thesis (MIT, Cambridge, Mass., 1980).
4. K. Ikeuchi, "Recognition of 3-D objects using the extended Gaussian image," presented at the 7th International Joint Conference on Artificial Intelligence, Vancouver, B.C., Canada, 1981.

5. P. Brou, "Using the Gaussian image to find the orientation of objects," Int. J. Robotics Res. **3**, 89–125 (1984).

6. B. K. P. Horn and K. Ikeuchi, "The mechanical manipulation of randomly oriented parts," Sci. Am. **251**(8), 100–111 (1984).

7. K. Ikeuchi, B. K. P. Horn, S. Nagata, T. Callahan, and O. Feingold, "Picking up an object from a pile of objects," in *Robotics Research: The First International Symposium,* M. Brady and R. Paul, eds. (MIT Press, Cambridge, Mass., 1984), pp. 139–162.

8. K. Ikeuchi, H. K. Nishihara, B. K. P. Horn, P. Sobalvarro, and S. Nagata, "Determining grasp configurations using photometric stereo and the prism binocular stereo system," Int. J. Robotics Res. **5**, 46–65 (1986).

9. Y. Li, "Orientation-based representations of shape and attitude determination," Tech. Rep. TR-93–12 (Dept. of Computer Science, University of British Columbia, Vancouver, BC, 1993).

10. R. J. Woodham, "Analysing images of curved surfaces," Artif. Intell. **17**, 117–140 (1981).

11. M. A. Penna and S. S. Chen, "Shape-from-shading using multiple light sources," Int. J. Intell. Syst. **1**, 263–291 (1986).

12. L. B. Wolff, "Surface curvature and contour from photometric stereo," in *Proceedings of the DARPA Image Understanding Workshop, 1987* (Morgan Kaufman, Los Altos, Calif., 1987), pp. 821–824.

13. L. B. Wolff, "Accurate measurement of second order variations of a smooth object surface using reflectance maps," in *Sensor Fusion: Spatial Reasoning and Scene Interpretation,* P. S. Schenker, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1003**, 59–62 (1988).

14. R. J. Woodham, "Determining surface curvature with photometric stereo," in *Proceedings of IEEE Conference on Robotics and Automation* (Institute of Electrical and Electronics Engineers, New York, 1989), Vol. I, pp. 36–42.

15. R. J. Woodham, "Surface curvature from photometric stereo," in *Physics-Based Vision: Principles and Practice,* L. Wolff, S. Shafer, and G. Healey, eds. (Jones and Bartlett, Boston, Mass., 1992), Vol. 3, pp. 121–155.

16. K. Ikeuchi, "Determining surface orientation of specular surfaces by using the photometric stereo method," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-3,** 661–669 (1981).

17. E. N. Coleman, Jr., and R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," Comput. Graphics Image Process. **18,** 309–328 (1982).

18. S. K. Nayar, K. Ikeuchi, and T. Kanade, "Determining shape and reflectance of hybrid surfaces by photometric sampling," IEEE Trans. Robotics Automation **6,** 418–431 (1990).

19. H. D. Tagare and R. J. P. de Figueredo, "A theory of photometric stereo for a class of diffuse non-Lambertian surfaces," IEEE Trans. Pattern Anal. Mach. Intell. **13,** 133–152 (1991).

20. A. L. Gilchrist, "The perception of surface blacks and whites," Sci. Am. **240**(3), 112–124 (1979).

21. B. K. P. Horn, "Understanding image intensities," Artif. Intell. **8,** 201–231 (1977).

22. J. J. Koenderink and A. J. van Doorn, "Geometrical modes as a general method to treat diffuse interreflections in radiometry," J. Opt. Soc. Am. **73,** 843–850 (1983).

23. D. Forsyth and A. Zisserman, "Reflections on shading," IEEE Trans. Pattern Anal. Mach. Intell. **13,** 671–679 (1991).

24. S. K. Nayar, K. Ikeuchi, and T. Kanade, "Shape from interreflections," Int. J. Comput. Vision **6,** 173–195 (1991).

25. L. Wolff, S. Shafer, and G. Healey, eds., *Physics-Based Vision: Principles and Practice* (Jones and Bartlett, Boston, Mass., 1992), Vol. 1.

26. B. K. P. Horn, *Robot Vision* (MIT Press, Cambridge, Mass., 1986).

27. B. K. P. Horn and M. J. Brooks, eds., *Shape from Shading* (MIT Press, Cambridge, Mass., 1989).

28. B. K. P. Horn, R. J. Woodham, and W. M. Silver, "Determining shape and reflectance using multiple images," Tech. Rep. AI-Memo-490 (AI Laboratory, MIT, Cambridge, Mass., 1978).

29. M. Minnaert, "The reciprocity principle in lunar photometry," Astrophys. J. **93,** 403–410 (1941).

30. B. K. P. Horn, "Hill-shading and the reflectance map," Proc. IEEE **69,** 14–47 (1981).

31. M. Brady, J. Ponce, A. Yuille, and H. Asada, "Describing surfaces," Comput. Vision Graphics Image Process. **32,** 1–28 (1985).

32. P. J. Besl and R. C. Jain, "Three-dimensional object recognition," ACM Comput. Surv. **17,** 75–145 (1985).

33. P. J. Besl and R. C. Jain, "Invariant surface characteristics for 3D object recognition in range images," Comput. Vision Graphics Image Process. **33,** 33–80 (1986).

34. B. K. P. Horn, "Extended Gaussian images," Proc. IEEE **72,** 1671–1686 (1984).

35. Y. Li and R. J. Woodham, "The support function, curvature functions and 3-D attitude determination," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1993), pp. 676–677.

36. J. J. Koenderink, *Solid Shape* (MIT Press, Cambridge, Mass., 1990).

37. R. J. Woodham, "Multiple light source optical flow," in *Proceedings of the 3rd International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, New York, 1990), pp. 42–46.

38. R. L. Stevenson and E. J. Delp, "Viewpoint invariant recovery of visual surfaces from sparse data," IEEE Trans. Pattern Anal. Mach. Intell. **14,** 897–909 (1992).

39. R. J. Woodham, Y. Iwahori, and R. A. Barman, "Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength," Tech Rep. TR-91-18, (Dept. of Computer Science, University of British Columbia, Vancouver, BC, 1991).

40. M. S. Drew, "Shape and specularity from color," Tech. Rep. CSS/LCCR TR-93-01 (Simon Fraser University, Burnaby, B.C., Canada, 1993).