CVGIP: IMAGE UNDERSTANDING
Vol. 54, No. 3, November, pp. 416–427, 1991

Depth and Shape from Shading Using the Photometric Stereo Method

BYUNGIL KIM AND PETER BURGER

Computer Vision and Graphics Group, Department of Computing, Imperial College of Science, Technology and Medicine, University of London, 180 Queen's Gate, London SW7 2BZ, England

Received September 20, 1989; accepted May 9, 1990

The independent calculation of local position and orientation of the Lambertian surface of an opaque object is proposed using the photometric stereo method. A number of shaded video images are taken using different positions of an ideal point light source which is placed close to the object. Normally, three images are required for a uniform and four for a textured Lambertian surface. By restricting three light sources to lie in a straight line, the depth calculations for an arbitrary surface with textured Lambertian reflection characteristics can be also determined; however, in this case the orientation of the surface cannot be calculated independently. It is shown that for both uniform and textured Lambertian surfaces the equations which are functions of three independent variables, namely, depth (D) and surface normal direction vector ($\mathbf{n} = [p, q, -1]$), can be reduced to a single nonlinear equation of depth, i.e., the distance between the camera and the point on the surface. Both convergence and a unique solution are ensured because of the simple behavior of the nonlinear equation within a practical range of depth and gradient values. The robustness of the algorithm is demonstrated by synthetic as well as experimental data. The calculation of the approximate positions and orientations of discontinuous surfaces is demonstrated when random noise is added to the synthetically calculated image intensities. Two parallel planes with a gap, two sloped planes, and a spherical surface are used to demonstrate that the algorithms work well. An important feature of calculating both depth and orientation independently is that for smooth surfaces they must obey the partial differential expressions $p = \delta D/\delta x$ and $q = \delta D/\delta y$. If we are certain that the experimental errors are within a known limit then the numerical approximation to these partial derivative expressions can be used to determine discontinuities within the image. On the other hand, if we know that the surfaces are smooth then errors in the numerical evaluation of these differential expressions allow the estimation of experimental errors. © 1991 Academic Press, Inc.

1. INTRODUCTION

According to our common everyday experience, the intensity of light reaching our retinas when we view a small area of an illuminated object depends on the absolute position of the object, its surface orientation, the illumination sources, the reflectivity properties of the viewed portion of the surface, and the transmission media between the object and our eyes. In other words, shape and depth of the object influence the intensity we perceive. Depth cuing in computer graphics is one application of this common dependency between perceived intensity (shade) and depth when successively darker portions of a computer generated scene seem to recede away from the viewer.

The inverse procedure of recovering shape (or depth) from shading is by now a classical topic of computer vision research which is still waiting for a large scale practical application. Since the pioneering work of Horn on *shape from shading* [1] a whole research area has grown and over the last 15 or so years a large number of researchers have investigated the feasibility of determining the shape of objects from their video image [see references 1–10]. We will refer to the shape (or depth) from shading problem as the calculation of the gradient vector components $p = \partial D/\partial x$ and $q = \partial D/\partial y$ (and depth D) at a given three dimensional position in space $[x, y, D]$ from the measured intensity function $I(x^*, y^*)$, where x^* and y^* specify positions in the image plane which are related to the three-dimensional coordinates $[x, y, D]$ through an appropriate geometrical projection function. Our earlier reports [11–13] already demonstrated the feasibility of determining depth as well as shape from image intensities when point light sources were used.

The shape from shading problem applied to real images faces many complications. In general, the distribution of incident illumination is nonuniform and is often unknown. The reflection properties of real surfaces are equally difficult to determine precisely. Furthermore, the resulting equations are highly nonlinear. In order to be able to arrive at a solvable problem, several strong assumptions have to be made. We have also shown that surface shape and depth could be determined from a single image; however, in this case known illumination and known reflection coefficient (albedo) must be assumed since any variation in reflected intensity can be due to the surface gradient, the incident illumination, and the reflection coefficient. If the latter two are not known, the surface gradient cannot be determined. (In other words, a

1049-9660/91 $3.00

darker spot in the image may be equally due to less incident illumination, to a darker colored spot on the surface, or to change in the direction of the gradient vector).

The strongest assumptions for the general shape from shading problem provide the simplest problem. These are parallel and uniform incident light illumination (such as from a very distant point light source) and uniform Lambertian surface (perceived intensity is independent of the viewer's position and is proportional to the cosine of the angle between the incident light direction and the surface normal). A further simplification is to use orthographic projection ($x^* = x, y^* = y$) which indicates a very distant camera position. Under these assumptions the measured light intensity can be expressed by

$$I(x, y) = I_0 \rho_0 (\mathbf{s} \cdot \mathbf{n})/|\mathbf{n}|, \tag{1.1}$$

where I_0 and ρ_0 are constants, the unit direction vector $\mathbf{s} = [s_x, s_y, s_z]$ is parallel to the incident light direction and points towards the light source, and the outside normal vector to the surface is $\mathbf{n}$ which can be also expressed by the gradient components as $[p, q, -1]$. Therefore

$$I(x, y) = I_0 \rho_0 (p s_x + q s_y - s_z)/[p^2 + q^2 + 1)^{1/2}. \tag{1.2}$$

Under these assumptions the perceived illumination is invariant to depth (since both the light source and the camera are far away).

It is not possible to use one image and Eq. 1.1 for the determination of shape locally since there are two unknowns (p and q) and only one equation. In order to arrive at a solution from a single image, smoothness conditions of the surface and the consideration of a continuous uniform region must be considered. Hence, we may call methods of solution for a single image *global*. Various research workers have shown that such global techniques are feasible when strong smoothness constraints are assumed [see Refs. 1–5]. Since we are interested in the photometric stereo technique which uses multiple images, we will not discuss these global methods in detail here.

2. PHOTOMETRIC STEREO METHOD

Woodham was the first to formulate the photometric stereo method [3]. In order to solve the illumination equation at any given surface point $[x, y, z]$ we must have at least as many equations as unknowns. We may generate multiple intensity maps by using different illumination conditions. Let us first relax one of the strong assumptions and allow the albedo to be a function of position $\rho(x, y)$ (variable albedo means textured surface), although the surface is still assumed to be purely Lambertian. Then, using far away light sources producing incoming light from different directions, we get

$$I_i(x, y) = I_{0i} \rho(x, y)(\mathbf{s}_i \cdot \mathbf{n})/|\mathbf{n}|, \tag{2.1}$$

where the index "i" indicates different light source directions and I_{0i} is a brightness constant for each source.

If three light sources are used (say, $i = 0, 1, 2$) then we can form two ratios

$$\begin{aligned} R_{10}(x, y) &= I_1/I_0 \\ &= (K_{10})(p s_{x1} + q s_{y1} - s_{z1})/(p s_{x0} + q s_{y0} - s_{z0}), \\ R_{20}(x, y) &= I_2/I_0 \\ &= (K_{20})(p s_{x2} + q s_{y2} - s_{z2})/(p s_{x0} + q s_{y0} - s_{z0}) \end{aligned} \tag{2.2}$$

which form a linear system of equations in two unknowns (p and q). The constants K_{10} and K_{20} express the relative strengths of the two light sources I_1 and I_2 with respect to the light source I_0. The two equations can be solved for the two unknowns p and q as long as the determinant is not zero. Depending on the directions of the incoming light rays, the determinant may become zero at isolated points but will have a finite value for the other points on the surface. It has been assumed so far that the illuminations are from faraway point sources and the ratios of the incident light intensity constants are known. These ratios may be considered a calibration problem in a practical system where the constants K_{10} and K_{20} can be measured for a known surface orientation (say $p = q = 0$) and then kept constant for all other measurements.

Even the uniformity of the light sources may be relaxed if the assumption on the parallel nature of the incoming light rays is retained. In this case, with a known calibration surface (say, a plane with $p = q = 0$) the light intensity variations as a function of position can be experimentally determined and look-up tables constructed [8, 9]. The practical applicability of this technique depends only on the following strong assumptions: parallel light rays, measured light intensity not a function of depth, purely Lambertian surface, and constant light sources.

3. NEAR POINT LIGHT SOURCES

Our main interest is to find solutions to the shape (and depth) from shading problem when the light sources are brought near to the viewed surface. Since the actual three-dimensional positioning of the light sources can be significant and the possibilities are limitless, we restrict ourselves for this study to the physical arrangement shown in Fig. 1. The camera is placed at the origin of our coordinate system and its axis is aligned with the z axis. Thus, depth (D) is equal to the z ordinate value. Since we

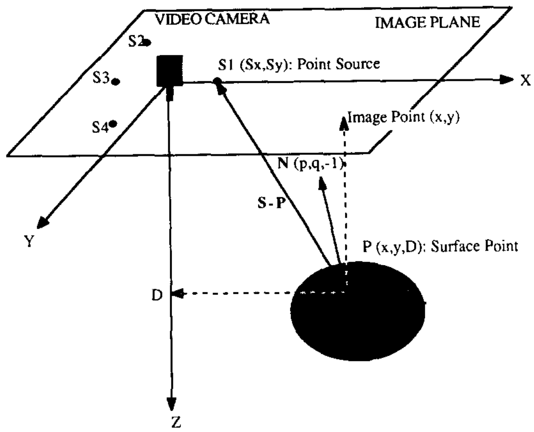

FIG. 1. Physical arrangement of object, light sources, and video camera.

use orthographic projection, the actual z position of the camera is unimportant as long as we know its position with respect to the plane of the light sources. To find the correct depth, the z ordinate of the camera has to be added to the calculated depth values.

As shown, all the light sources are assumed to be point sources and are placed in the $z = D = 0$ plane. Keeping the assumption of orthographic projection, the intensity map equation similar to Eq. 2.1 can now be derived as

$$I_i(x, y) = I_{0i}\rho(x, y)(\mathbf{s}_i \cdot \mathbf{n})/\{|\mathbf{n}|\ |\mathbf{s}_i|^3\}, \tag{3.1}$$

where $\mathbf{s}_i$ is now *not* a unit vector but is the difference vector between the source position and the surface point with components

$$\mathbf{s}_i = [S_{xi} - x,\ S_{yi} - y,\ - D]$$

The terms in the denominator of Eq. 3.1 arise because the incident intensity at the surface drops off as the square of the distance between the point light source and the object. Since orthographic projection is used, the surface coordinates (x, y) are equal to the image coordinates (x^*, y^*). For perspective projection (x^*, y^*) becomes a function of surface coordinates (x, y, z) and camera focal length f. In the case of a simple pinhole camera, (x, y) can be replaced by $[(Df)x^*, (Df)y^*]$.

Substituting the terms $\Delta x_i = S_{xi} - x$ and $\Delta y_i = S_{yi} - y$ and expanding Eq. 3.1 we get

$$I_i(x, y) = I_{0i}\rho(x, y)(p\ \Delta x_i + q\ \Delta y_i + D)/ \\ \{(1 + p^2 + q^2)^{1/2}(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2}\}. \tag{3.2}$$

We have now four unknowns (D, p, q, and ρ), hence four different illuminations must be chosen. If we assign indices 0 to 3 for the different measurements and take ratios as before, we get

$$R_{i0}(x, y) = I_i(x, y)/I_0(x, y) \\ = \frac{I_{0i}}{I_{00}} \frac{(p\ \Delta x_i + q\ \Delta y_i + D)(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2}}{(p\ \Delta x_0 + q\ \Delta y_0 + D)(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2}}, \tag{3.3}$$

where the index i has values 1, 2, and 3. From the three equations we can derive one nonlinear equation for D which is of the form

$$(A_3 B_2 - A_2 B_3) C_1 + (A_1 B_3 - A_3 B_1) C_2 \\ + (A_2 B_1 - A_1 B_2) C_3 = 0, \tag{3.4}$$

where A_i, B_i and C_i are all functions of D and are given by

$$A_i = [\Delta x_i - \Delta_{x0} F_i(D)]$$
$$B_i = [\Delta y_i - \delta y_0 F_i(D)]$$
$$C_i = D[F_i(D) - 1]$$

for $i = 1, 2, 3$, and

$$F_i(D) = K_{i0}(x, y)[(\Delta x_i^2 + \Delta y_i^2 + D^2)/ \\ (\Delta x_0^2 + \Delta y_0^2 + D^2)]^{3/2}, \tag{3.5}$$

where $K_{i0}(x, y) = R_{i0}(x, y)(I_{00}/I_{i0})$. Here values for $R_{i0}(x, y)$ are derived from the measured image intensities and constants I_{00} and I_{i0} are determined by calibration. Assuming that the system of equations is not singular, Eq. 3.4 will yield one or more solutions for D after which the gradient components p and q can be determined.

Thus, with the assumptions of constant point light sources, orthographic projection, and Lambertian surfaces, we have reduced the local solution to the shape and depth from shading problem to the solution of one nonlinear equation in D. It is worthwhile to compare our method to that used for faraway light sources and camera positions. Since we have one more variable (depth), we need one more light source. Also, our equations become nonlinear. These are the disadvantages. The advantage is that since we calculate D, p, and q independently, in the case of a smooth surface we have redundancy. For a smooth surface $p = \partial D/\partial x$ and $q = \partial D/\partial y$ and since these equations may be checked numerically, we can form a measure of the errors inherent in our solution. Furthermore, if the absolute value of the errors are within some limit for most parts of the image but are much larger along selected contours, then discontinuities such as edges may be detected at these positions. As far as the numerical (iterative) solution of nonlinear equations is concerned, it is particularly convenient that a system of nonlinear equations in three unknowns can be reduced to one nonlinear equation of a single variable.

The other advantage is practical. Obviously, the photometric stereo method relies on the illumination of objects by controlled and, for simplicity, uniform light sources. In practice one may have ambient lighting conditions which would be impossible to control. However, by adding a bright light illumination to the scene it may be possible to use two images, one with and one without the light source, in which case the illumination due to the point light source can be recovered. It seems that this method could succeed more easily if the additional light source was a point light source brought close to the object.

Obviously, the success of this method depends on whether the nonlinear equation in D is solvable and whether it has only one solution. This will be influenced by the positioning of the light sources as well as other experimental conditions. First we investigate the positioning of the light sources; then we show that the equations are solvable using standard numerical techniques.

4. POSITIONING OF LIGHT SOURCES AND UNIQUENESS OF SOLUTIONS

We now investigate different arrangements of the point light sources and the feasibility and uniqueness of the solutions. The uniqueness of the solutions will be determined within a range of parameter values. In order to extend these results for practical cases, a normalization will be carried out which results in dimensionless quantities. The average distance between the camera and the light sources will be used as the normalizing length and for other distances (depth, for example) the practical range of 0.1 to 10 will be used.

4.1. *Constant Albedo*

First, we make the additional assumption of constant albedo, in which case only three light sources are needed. The three equations are now

$$I_i(x, y) = K_i(p\,\Delta x_i + q\,\Delta y_i + D)/\{(1 + p^2 + q^2)^{1/2}\,(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2}\}, \tag{4.1}$$

where the constant K_i contains both the strength of the light source and the constant albedo and is determined by calibrating the experimental system. Assigning one of the equations index 0 and the other two indices 1 and 2, two ratios can be formed:

$$R_{i0}(x, y) = I_i(x, y)/I_0(x, y) = \frac{K_i}{K_0}\,\frac{(p\,\Delta x_i + q\,\Delta y_i + D)(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2}}{(p\,\Delta x_0 + q\,\Delta y_0 + D)(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2}}, \tag{4.2}$$

From Eq. 4.2 both p and q can be expressed as nonlinear equations of D, i.e.,

$$p\,A_i + q\,B_i = C_i, \tag{4.3}$$

where A_i, B_i, and C_i are all functions of D and for $i = 1, 2$ they are given by

$$\begin{aligned} A_i &= R_{i0}(x, y)\,\Delta x_0\,(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2} \\ &\quad - K_{i0}\Delta x_i\,(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ B_i &= R_{i0}(x, y)\,\Delta y_0\,(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2} \\ &\quad - K_{i0}\Delta y_i\,(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ C_i &= D\{K_{i0}(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ &\quad - R_{i0}(x, y)\Delta x_i\,(\Delta x_i^2 + \Delta y_i^2 + D^2)^{3/2}\} \end{aligned} \tag{4.4}$$

The ratio $R_{i0}(x, y) = I_i(x, y)/I_0(x, y)$ which is measured experimentally for each position (x, y) and K_{i0} is the calibration constant (K_i/K_0). The terms Δx_i and Δy_i are the differences between the source positions $\mathbf{S}_i$ and the coordinates $[x, y]$. From Eq. 4.3 we get

$$\begin{aligned} p &= (C_1\,B_2 - C_2B_1)/(A_1\,B_2 - A_2\,B_1) \quad \text{and} \\ q &= (C_2\,A_1 - C_1\,A_2)/(A_1\,B_2 - A_2\,B_1), \end{aligned} \tag{4.5}$$

which can be substituted into Equation 4.1 with $i = 0$ to get a single nonlinear equation for distance D:

$$\begin{aligned} &I_0(x, y)(1 + p^2 + q^2)^{1/2}(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ &\quad - K_0(p\,\Delta x_0 + q\,\Delta y_0 + D) = 0. \end{aligned} \tag{4.6}$$

Before the nonlinear equation for D is solved, a normalization by the average distance between the light sources and the camera is useful. Without loss of generality, we can consider all distances normalized to this average length S. For example, if we select three symmetrically placed light sources around a circle with radius S, the source coordinates become

$$S_{x0} = 1;\quad S_{y0} = 0;\quad S_{x1} = -1/2;\quad S_{y1} = \sqrt{3}/2;$$
$$S_{x2} = -1/2;\quad S_{y2} = -\sqrt{3}/2.$$

For this nonlinear equation it would be difficult if not impossible to analytically prove the uniqueness of the solution. By using normalized quantities we can search for solutions within a reasonable practical range, say

$$0.1 < D < 10 \quad \text{and} \quad (p^2 + q^2)^{1/2} < 10.$$

The search starts by assigning values to p, q, and D. Then, for a selected number of positions $[x, y]$, the values for I_0, I_1, and I_2 are calculated and the left hand side

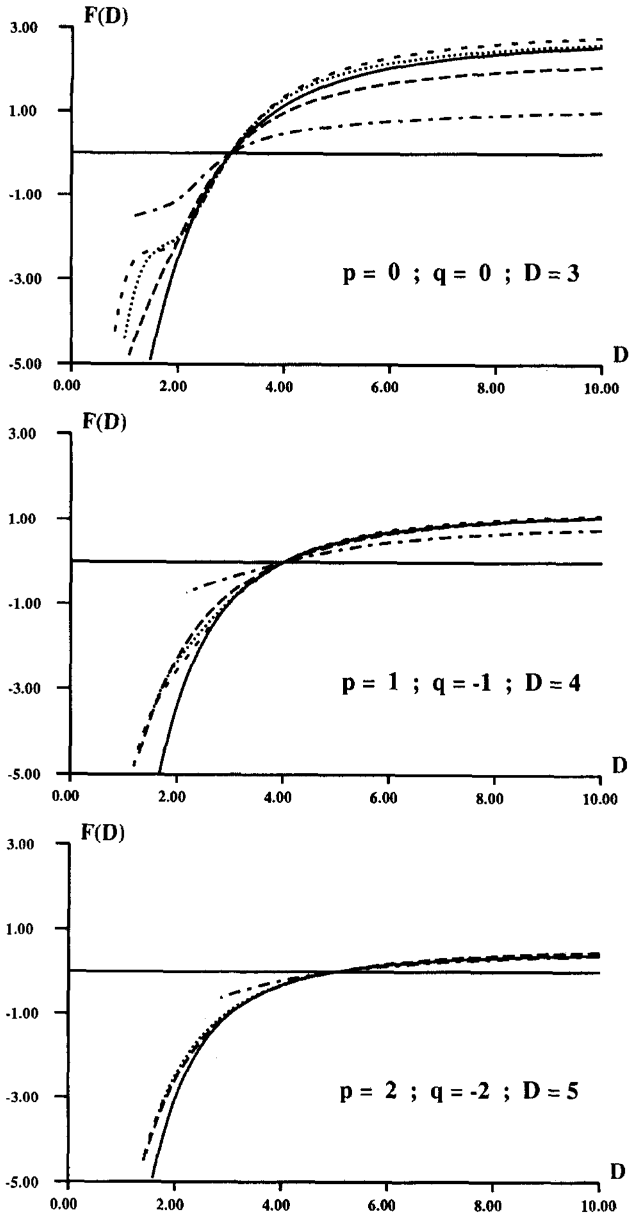

FIG. 2.

of Eq. 4.6 is plotted for a range of D values between 0 and 10. Three such plots are shown in Fig. 2 with different values for p and q. At a point near to the origin ($x = 0$, $y = 0$) we have found a unique solution for all values of p and q which we tested. For larger p and q values and points outside the circle with radius 1 a second solution seems to emerge for very small values of D. However, we have also found that for these conditions the determinant connecting p and q to D (Eq. 4.3) is near singular. Since we can detect the near singularity of the determinant, we can discard this solution and end up with a unique solution for the practical ranges we have considered. The third diagram in Fig. 2 demonstrates the increasing ambiguity of the solution as the derivative of the $F(D)$ curve approaches zero near the solution. Hence, we

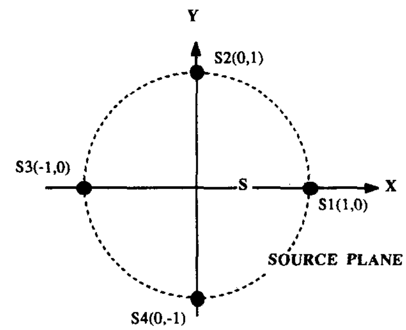

FIG. 3. Symmetric placement of the four light sources in the source plane.

should expect numerical difficulties at large values of D, p, and q. Otherwise, the simple behavior of the $F(D)$ function indicates that it will be straightforward to solve numerically the $F(D) = 0$ equation.

Another aspect of these curves which is not immediately apparent is that only a limited range of parameter values can be tested since, depending on the values of p, q, and D, one or more surfaces may be in shadow.

4.2. *Textured Surfaces*

We have shown earlier that for textured surfaces four light sources are required. Our first choice for the position of the sources was in a symmetrical arrangement (see Fig. 3):

$$S_{x0} = 1;\quad S_{y0} = 0;\quad S_{x1} = 0;\quad S_{y1} = 1;\quad S_{x2} = -1;$$
$$S_{y2} = 0;\quad S_{x3} = 0;\quad S_{y3} = -1.$$

Unfortunately, we found that this symmetrical arrangement provides a singular point in the center ($x = 0$, $y = 0$). Since the center is where our solution should be the most robust, obviously, the symmetrical arrangement of the four light sources is unsatisfactory. Thus, we have returned to the original symmetrical arrangement for three light sources and moved the fourth light source 2 normalized distances away from the origin. The positions of the sources are shown in Fig. 4 and are given by

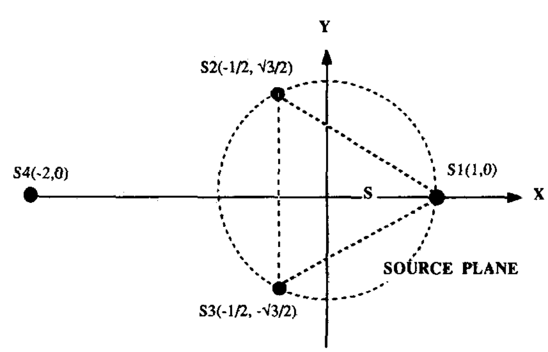

FIG. 4. Nonsymmetric placement of the four light sources in the source plane.

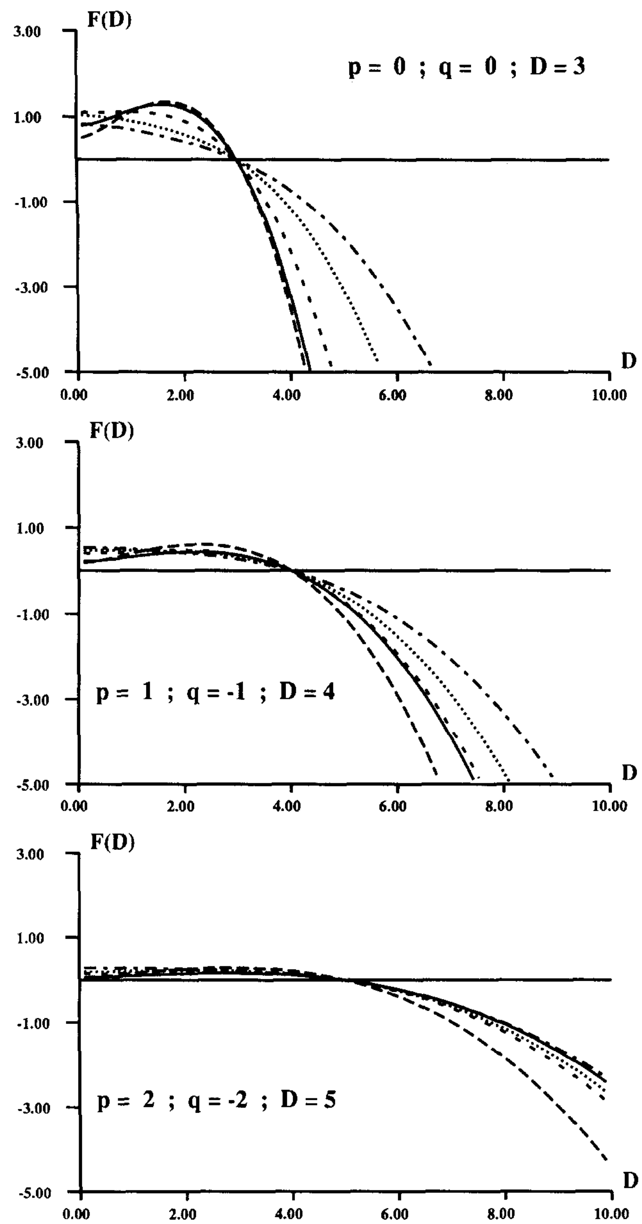

FIG. 5.

$$S_{x0} = 1; \quad S_{y0} = 0; \quad S_{x1} = -1/2; \quad S_{y1} = \sqrt{3}/2;$$
$$S_{x2} = -1/2; \quad S_{y2} = -\sqrt{3}/2; \quad S_{x3} = -2;$$
$$S_{y3} = 0.$$

Three representative curves for $F(D)$ (Eq. 3.5) are shown in Fig. 5. For all of the values we have tested, when none of the four illuminations produced self-shadow, only one solution could be found.

4.3. *Three Light Sources on a Straight Line*

An interesting anomaly of these nonlinear equations was discovered when different light arrangements were tried. When only three light sources are used for the textured (variable albedo) case and are placed in a straight line, the depth may be determined. In this case we have

$$S_{x1} = S_{x0} + \delta x; \quad S_{y1} = S_{y0} + \delta y; \quad S_{x2} = S_{x0} - \delta x;$$
$$S_{y2} = S_{y2} - \delta y.$$

Using Eq. 3.2 and taking differences we get

$$\begin{aligned} &(I_1(x, y)/I_{01})(\Delta x_1^2 + \Delta y_1^2 + D^2)^{3/2} \\ &\quad - (I_0(x, y)/I_{00})\,(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ &= \rho(x, y)(-p\,\delta x - q\,\delta y)/(1 + p^2 + q^2)^{1/2} \\ &= F_1(p, q, D) \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} &(I_2(x, y)/I_{02})(\Delta x_2^2 + \Delta y_2^2 + D^2)^{3/2} \\ &\quad - (I_0(x, y)/I_{00})\,(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ &= \rho(x, y)(p\,\delta x + q\,\delta y)/(1 + p^2 + q^2)^{1/2} \\ &= -\,F_1(p, q, D). \end{aligned} \tag{4.8}$$

Hence, the left hand sides of Eqs. 4.7 and 4.8 can be used to arrive at a single nonlinear equation for D,

$$\begin{aligned} F(D) &= (I_1(x, y)/I_{01})(\Delta x_1^2 + \Delta y_1^2 + D^2)^{3/2} \\ &\quad - 2\,(I_0(x, y)/I_{00})(\Delta x_0^2 + \Delta y_0^2 + D^2)^{3/2} \\ &\quad + (I_2(x, y)/I_{02})(\Delta x_2^2 + \Delta y_2^2 + D^2)^{3/2} = 0, \end{aligned} \tag{4.9}$$

which can be solved numerically. We can again investigate the behavior of $F(D)$ as a function of D for various values of a self-consistent solution for p, q, and D. Three sets are shown in Fig. 6 which indicate again unique solutions for moderate values of p and q but they do warn us about numerical difficulties for large p and q.

It is odd that in general four images are required for a textured case but for these special lighting conditions three suffices. We find the anomaly when we try to determine values for p and q. Once the depth (D) is known, the three equations are reduced to two linear equations in p and q. However, for this case the determinant of the system of two equations is singular. Thus, by eliminating one light source we must give up the independent determination of values for p and q. This fact suggests an effective way for arranging four light sources. If three light sources are placed in a straight line then there are two ways to calculate the depth at each point—an additional redundancy which could help in the estimation of

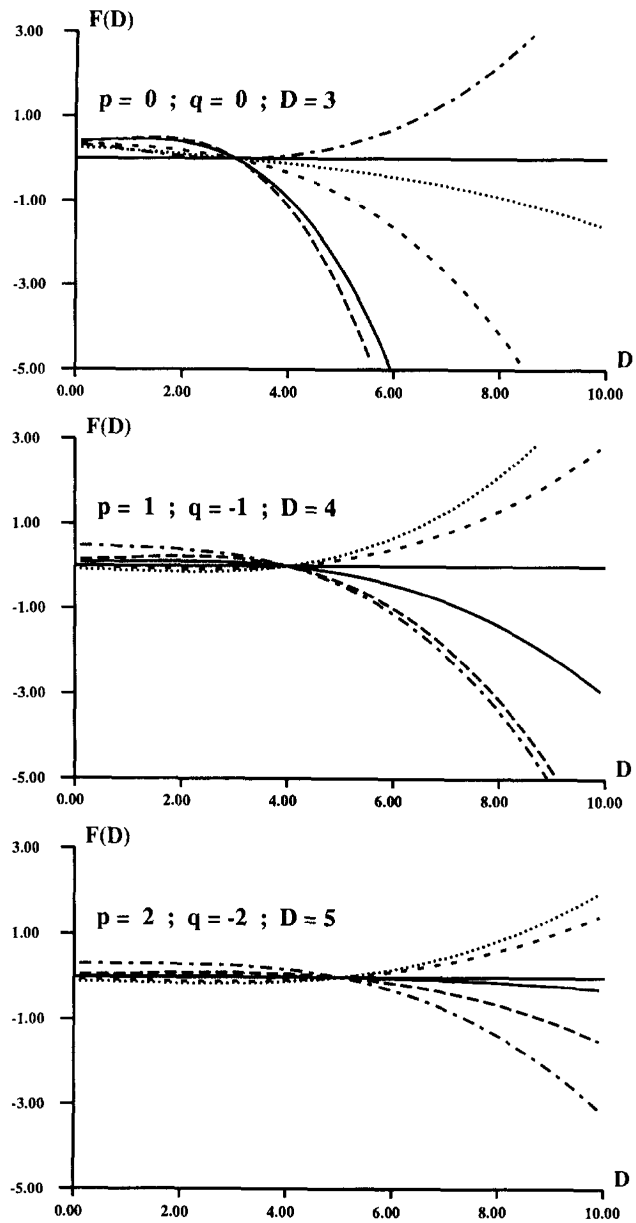

FIG. 6.

the experimental errors. We hope to explore this feature in the future.

5. INITIAL CONDITIONS

We have reduced now the determination of local position and orientation from a number of intensity maps to the solution of a single nonlinear equation for D for both constant and variable albedo. The use of an iterative nonlinear equation requires initial conditions, i.e., an initial guess for D in order to start the iteration. The plots for $F(D)$ in all cases we have investigated seem to be well behaved; however, a good initial guess for D speeds up convergence. It seems reasonable to assume that in a small neighborhood the surface may be approximated by a plane. Let us choose small increments in the x and the y directions as δx and δy around an arbitrary point on the surface $[x, y, D]$, then using the equation of a plane we get

$$x' = x + \delta x; \quad y' = y + \delta y; \quad D' = D + p\,\delta x + q\,\delta y;$$
$$p' = p; \quad q' = q; \quad \rho' = \rho(x', y'), \quad \text{and} \quad I' = I(x', y'), \tag{5.1}$$

where the prime (′) indicates the neighborhood point. The ratios of the measured intensities for these two neighboring points and one light source produce a greatly simplified expression,

$$\{I'/I\}^{2/3} = \frac{(\rho')^{2/3}(\Delta x^2 + \Delta y^2 + D^2)}{\rho^{2/3}\{(\Delta x - \delta x)^2 + (\Delta y - \delta y^2) + (D + p\delta x + q\delta y)^2\}}, \tag{5.2}$$

where I and I' are measured data, ρ and ρ' are the values of albedo at (x, y, D) and at (x', y', D'), $\Delta x = S_x - x$, $\Delta y = S_y - y$, and the light source has Cartesian coordinates $[S_x, S_y, 0]$. Equation 5.2 holds for any light source. For the i'th light source we can write

$$R_i(x, y)^{2/3}(\Delta x_i^2 + \Delta y_i^2 + D^2)\{\rho'/\rho\}^{2/3} - (\Delta x_i - \delta x)^2 - (\Delta y_i - \delta y)^2 = (D + p\delta x + q\delta y)^2, \tag{5.3}$$

where, as before, the point source positions are at coordinates $[S_{xi}, S_{yi}, 0]$ and the ratio

$$R_i(x, y) = \{I_i/I_i'\}$$

is determined from the measured data. The right hand side of Eq. 5.3 does not depend on the position of the light source; thus, subtracting the left hand sides of Equation 5.3 for two different light sources, we get

$$\begin{aligned} &R_2^{2/3}(\Delta x_2^2 + \Delta y_2^2 + D^2) - R_0^{2/3}(\Delta x_0^2 + \Delta y_0^2 + D^2) \\ &\quad = \{\rho/\rho'\}^{2/3}\{(\Delta x_2 - \delta x)^2 + (\Delta y_2 - \delta y)^2 - (\Delta x_0 - \delta x)^2 - (\Delta y_0 - \delta y)^2\} \\ &R_1^{2/3}(\Delta x_1^2 + \Delta y_1^2 + D^2) - R_0^{2/3}(\Delta x_0^2 + \Delta y_0^2 + D^2) \\ &\quad = \{\rho/\rho'\}^{2/3}\{(\Delta x_1 - \delta x)^2 + (\Delta y_1 - \delta y)^2 - (\Delta x_0 - \delta x)^2 - (\Delta y_0 - \delta y)^2\}. \end{aligned} \tag{5.4}$$

From Eq. 5.4 we get an expression for D^2:

$$\{R_2^{2/3}(\Delta x_2^2 + \Delta y_2^2 + D^2) - R_0^{2/3}(\Delta x_0^2 + \Delta y_0^2 + D^2)\}$$
$$\{(\Delta x_1 - \delta x)^2 + (\Delta y_2 - \delta y)^2 - (\Delta x_0 - \delta x)^2 - (\Delta y_0 - \delta y)^2\}$$
$$= R_1^{2/3}(\Delta x_1^2 + \Delta y_1^2 + D^2) - R_0^{2/3}(\Delta x_0^2 + \Delta y_0^2 + D^2)\}$$
$$\{(\Delta x_2 - \delta x)^2 + (\Delta y_2 - \delta y)^2 - (\Delta x_0 - \delta x)^2 - (\Delta y_0 - \delta y)^2\}. \tag{5.5}$$

Equation 5.5 is left in terms of arbitrary distance parameters δx and δy. Incidentally, these equations show that only three light sources are needed for the determination of the three parameters p, q, and D for a plane. This is a result that may be used for practical shape from shading determination in the future.

Returning to our original aim of finding initial conditions for D, only three of the four light sources and two points are needed for an initial estimate. If we take a small three-by-three area and keep the center point constant, we may use eight different neighboring points and a number of combinations of light sources. For four light sources we have 3×8 or 24 ways of estimating the initial D value. For a general scene we also have the choice of the center point. Thus, we can search for a point where the 8 or 24 calculations produce similar results. This is an indication that the measured intensities are consistent with the assumed reflected light intensity function. The average of the 8 or 24 estimates then should produce an unbiased initial estimate for D and the iteration can start.

6. RESULTS

First we show results from synthetic data for our proposed algorithms. The synthetic data are derived from planar and spherical surfaces with and without pseudo random noise. In order to demonstrate that our method is a truly local one, we have chosen two discontinuous surfaces; one has a jump in value, the other a discontinuity in its derivative. Figure 7 is for exact data (no noise) where the results show perfect reconstruction of all the surfaces. In addition to the perspective view of the reconstructed surfaces, the spread of surface points with respect to one of the normalized distances (x/S) is also shown. The points on the spherical surface lie on a small band of circular arcs; the planes show up as straight lines. These should be compared to Fig. 8 where the same algorithm is used but 10% pseudo random noise is added to the "recorded" intensity values.

As described in Section 5, the initial D value is estimated at the center of the 32×32 grid and we use the hybrid method [14] for solving our nonlinear equation for D. Once the actual D value is known at one point, instead of recalculating the initial estimate, it is used as an initial estimate for the eight neighbouring points. In this way, solutions start in the middle and continue out to the boundaries. Of course, at every point a self-consistent solution is found. It takes approximately one second to find the solution for one pixel using an IBM AT with an 80386 processor running at 16 MHz.

In addition to these synthetic surfaces which demonstrate that reconstruction from a shaded image is feasible, we have conducted experiments with prepared surfaces that approximated Lambertian reflectors well. The illuminations we used were close to ideal point sources and the experimental results were similar to the synthetic ones. These experiments are described elsewhere [13]. A set of results of two sloped planes and uniform Lambertian surfaces is shown in Fig. 9. Experimentally we have demonstrated that the two other algorithms for textured surfaces, one using four light sources and the other using three sources in a straight line, also produce similar results. Overall, even for our experimental measurements, relative errors were less than 15%.

7. CONCLUSIONS

It has been demonstrated that illumination by nearby point light sources produces shaded video images from which the three-dimensional position and orientation of the reflecting surface can be reconstructed. In order to recover the three independent quantities, namely depth (D) and the two components of the surface normal vector: $\mathbf{n} = [p, q, -1]$ when the surface is uniform Lambertian, three images resulting from three different illumination conditions are required. If the positions of the light sources are chosen symmetrically in a plane perpendicular to the camera optical axis, uniqueness of the solutions within a practical range of depth and gradient values is ensured. The three-dimensional position and orientation of the surface are independently evaluated for each image point. If the surface is textured, in general, four images are required. But if only the depth value is required then three light sources in a straight line can be used.

The robustness of the algorithms is demonstrated by both synthetic and experimental data. It is shown that since this is a local technique, discontinuous surfaces can be handled and the errors in the calculated depth values are proportional to the noise added to the recorded image intensity values. This is a new way of obtaining dense range data since no special equipment is necessary. However, the accuracy of the method does rely on controlled sources which approximate the characteristics of an ideal point light source well, on a controlled calibration technique, and on near-Lambertian surfaces.

One immediate extension of this technique is to include perspective projection which, looking at the equations, should not be very difficult. Another way of extending

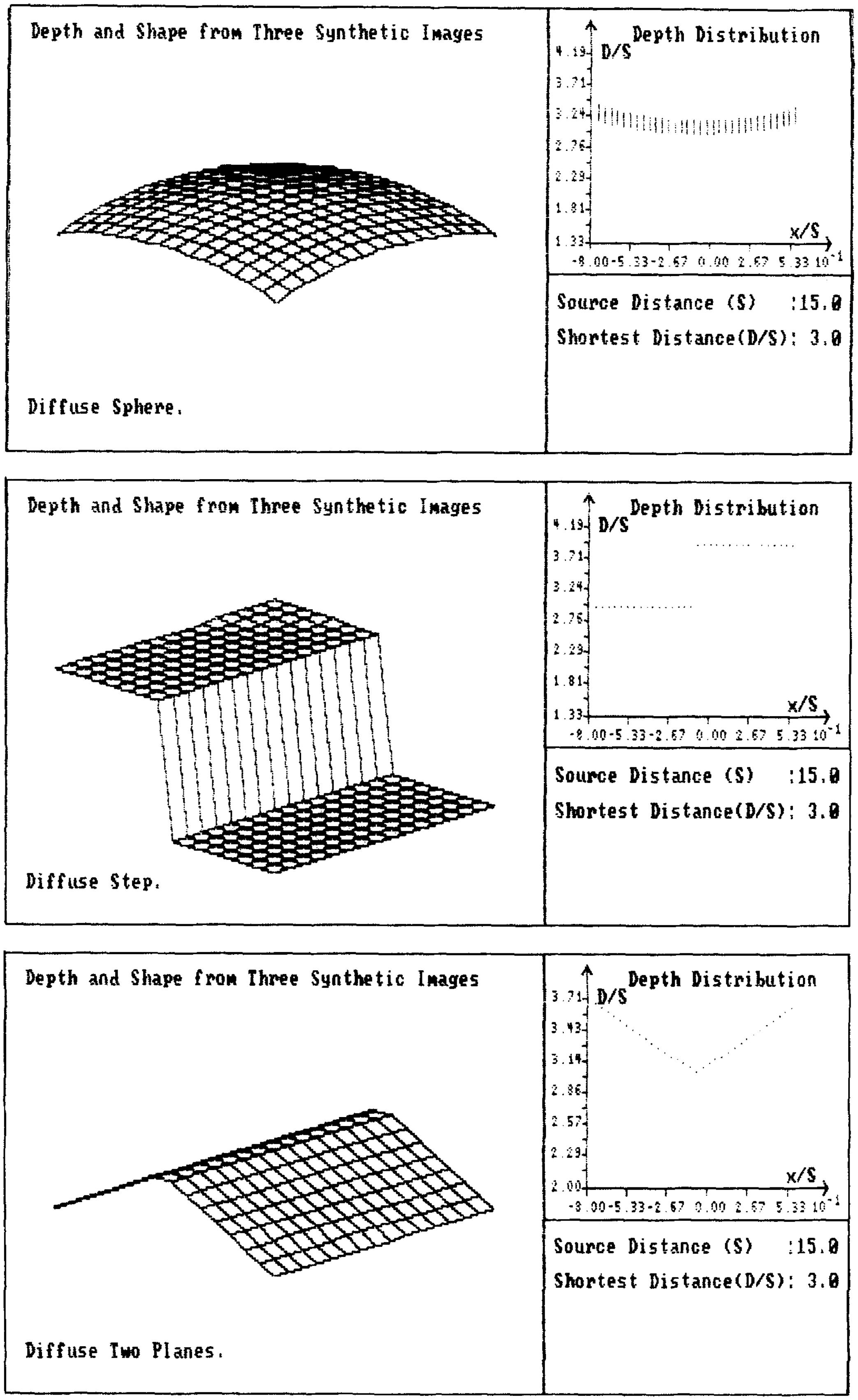

FIG. 7. Shapes generated by the algorithm from three synthetic images.

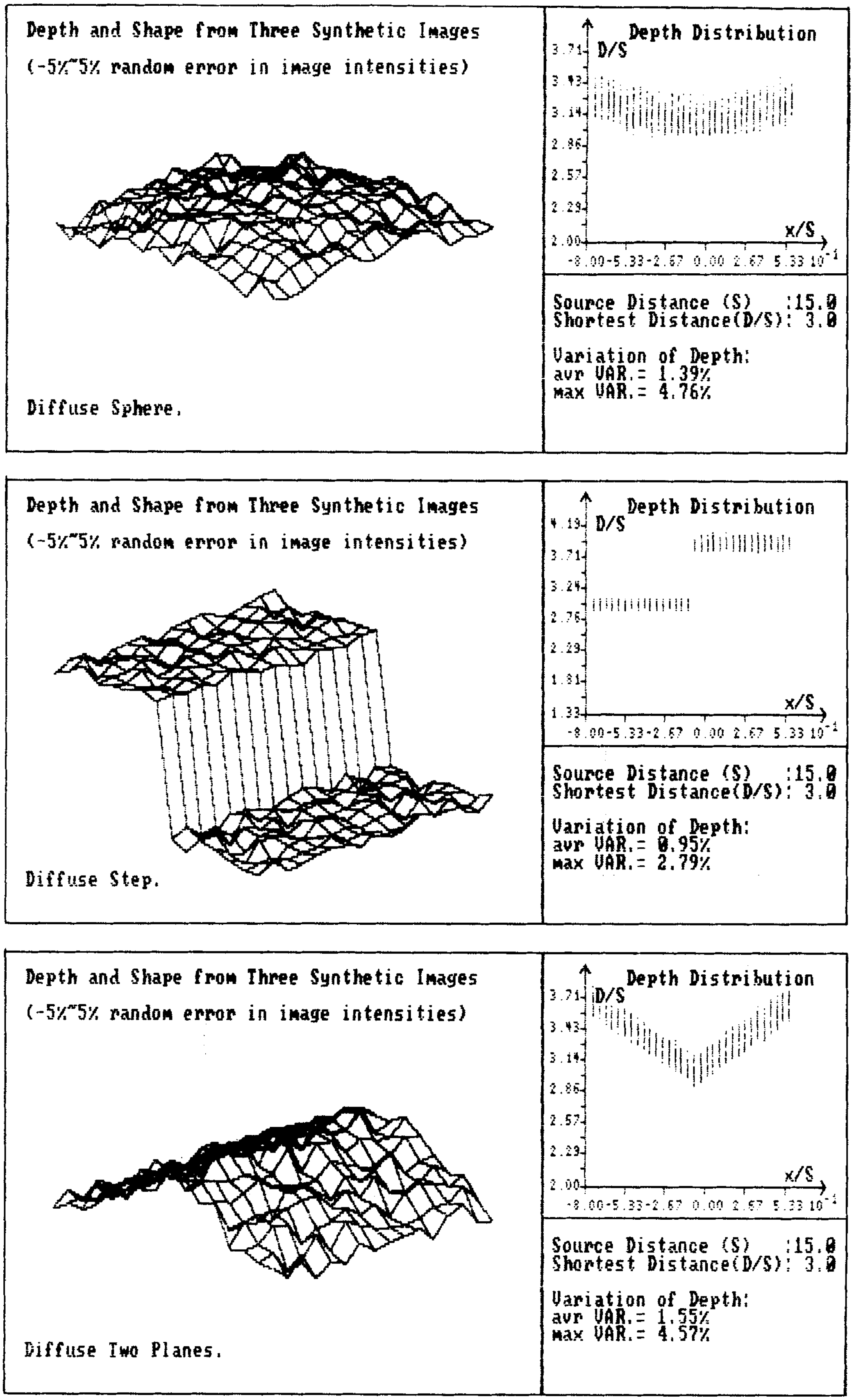

FIG. 8. Shapes generated by the algorithm from three synthetic images with 10% random errors.

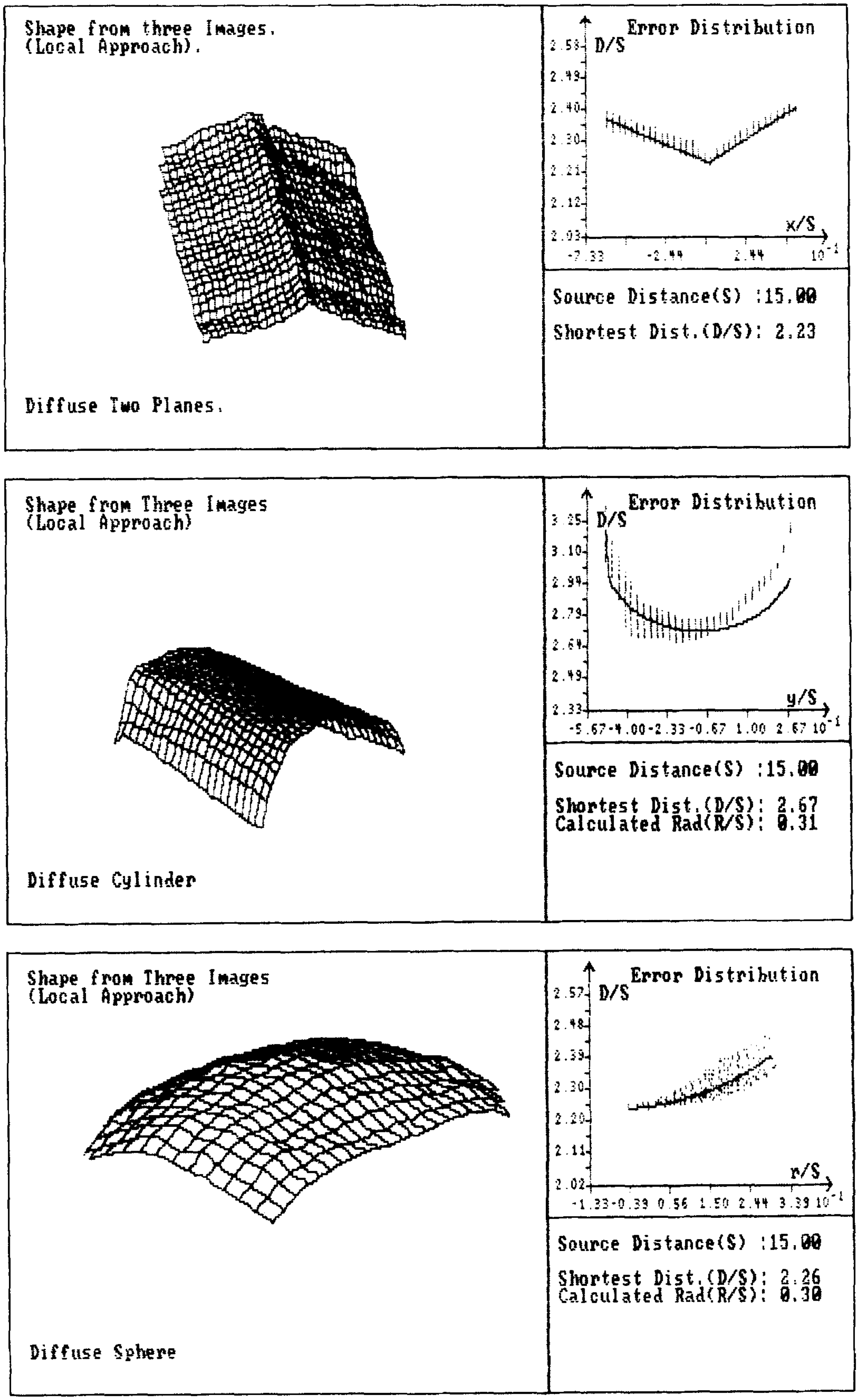

FIG. 9. Shapes generated by the algorithm from three real images placed at 45 cm.

this technique to more practical situations is to allow light sources to have more general illumination characteristics. This will result in the numerical mapping (calibration) of the source characteristics. Then, when the nonlinear equation for the depth value is solved, the actual illumination received by the point on the surface may be taken into account.

The extension of this local method to specular surfaces seems much more difficult than the mentioned two extensions. However, Kim has found [13] that a global method for the determination of position and orientation for smooth surfaces can be combined with the photometric stereo method, and specular highlights can be handled correctly. Thus, if local and global techniques are combined and described extensions for the better approximation of practical situations are introduced, it seems reasonable to expect that the determination of dense range data from shaded video images will become practical in the future.

REFERENCES

1. B. K. P. Horn, Obtaining shape from shading information, in *The Psychology of Computer Vision* (P. H. Winston, Ed.), pp. 115–155, McGraw-Hill, New York, 1975.
2. B. K. P. Horn, Image intensity understanding, *Artificial Intelligence* **8,** 1977, 201–231.
3. R. J. Woodham, A cooperative algorithm for determining surface orientation from a single view, in *Proceedings of the International Joint Conference on Artificial Intelligence, Cambridge, Massachusetts, August 22–25, 1977,* pp. 635–641.
4. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding boundaries, *Artificial Intelligence* **17,** 1981, 141–184.
5. B. K. P. Horn and M. J. Brooks, *The Variational Approach to Shape from Shading,* MIT AI Lab. Memo 813, March, 1985.
6. B. K. P. Horn, R. J. Woodham, and W. Silver, *Determining Shape and Reflectance Using Multiple Images,* MIT AI Lab. Memo 490, August, 1978.
7. R. J. Woodham, Photometric stereo: A reflectance map technique for determining surface orientation from a single view," in *Proceedings, SPIE 22nd Annual Technical Symposium, Image Understanding Systems and Industrial Applications, San Diego, California, August 28–31, 1978,* Vol. 155, pp. 136–143.
8. W. M. Silver, *Determining Shape and Reflectance Using Multiple Images, S.M. Thesis,* Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, June, 1980.
9. K. Ikeuchi, Determining surface orientations of specular surfaces by using the photometric stereo methods," *IEEE Trans. Pattern Anal. Mach. Intelligence* **3,** 1981, 661–669.
10. E. N. Coleman, Jr., and R. Jain, Obtaining 3-dimensional shape of textured and specular surface using four-source photometry, *Comput. Graphics Image Process.* **18,** 1982, 309–328.
11. B. Kim and P. Burger, Calculation of surface position and orientation using the photometric stereo method, in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition (IEEE CVPR '88), Ann Arbor, Michigan, June, 1988,* pp. 492–497.
12. B. Kim and P. Burger, Calculation of surface position and orientation from a single shaded image, in *Proceeding of the International Congress on Optical Science and Engineering, Hamburg, W. Germany, September, 1988,* pp. 122–129.
13. B. Kim, *Depth and Shape from Shading,* Ph.D Thesis, Imperial College of Science, Technology and Medicine, Univ. of London, London, 1989.
14. D. W. Marquardt, An algorithm for least squares estimation of nonlinear parameters" *J. Soc. Industrial. Appl. Math.* **11,** 1963, 431–441.