3-D Stereo Using Photometric Ratios

Lawrence B. Wolff and Elli Angelopoulou

Computer Vision Laboratory
Department of Computer Science, The Johns Hopkins University
Baltimore, MD 21218

Abstract. We present a novel robust methodology for corresponding a dense set of points on an object surface from photometric values, for 3-D stereo computation of depth. We use two stereo pairs of images, each pair taken of exactly the same scene but under different illumination. By respectively dividing the left images and the right images of these pairs, a stereo pair of photometric ratio images is produced. We formally show that for diffuse reflection the photometric ratio is invariant to camera characteristics, surface albedo, and viewpoint. Therefore the same photometric ratio in both images of a stereo pair implies the same equivalence class of geometric physical constraints. We derive a shape-from-stereo methodology applicable to perspective views and not requiring precise knowledge of illumination conditions. This method is particularly applicable to smooth featureless surfaces. Experimental results of our technique on smooth objects of known ground truth shape are accurate to within 1% depth accuracy.

1 Introduction

There has been extensive work on computational stereo vision (see [12], [11], [5], [15], [1], and [13]). Most of the methods for computing depth from stereo vision involves the correspondence of image features such as intensity discontinuities or zero crossings determining image edges. A possible disadvantage of these techniques is that such data can be sparse and a number of methods have been developed to interpolate smooth surfaces to sparse depth data from stereo [5], [18]. There are considerable problems with shape determination of smooth featureless objects using feature-point based stereo vision algorithms.

Grimson [6] was the first to consider utilizing the diffuse shading information from two camera views to determine surface orientation at zero crossings. This additional information increased the accuracy of the surface interpolation. Smith [17] considered the correspondence of points in a stereo pair of images of a smooth featureless Lambertian reflecting surface utilizing an elegant mathematical formulation he termed the *Stereo Integral Equation*. There has been work by Blake, Brelstaff,

This research was supported in part by an NSF Research Initiation Award, grant IRI-9111973, DARPA contract F30602-92-C-0191 and an NSF Young Investigator Award IRI-9357757.

Lecture Notes in Computer Science, Vol. 801
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

Zisserman and others [2], [3], [24] that has exploited the geometry of specular reflection viewed from a stereo pair of cameras to derive constraints on surface shape.

The major advantage of being able to accurately correspond photometric values between a stereo pair of images, besides being able to determine the shape of smooth featureless surfaces, is that it provides a very dense depth map. We utilize 2 illumination conditions but do not need to know them precisely. Our methodology corresponds the left and right ratio images of the same scene produced under these illumination conditions. We prove that the photometric ratios arising from diffuse reflection are invariant to almost all characteristics varying between a stereo pair of cameras as well as viewpoint and diffuse surface albedo. These varying characteristics make pixel gray values by themselves unreliable for stereo correspondence. In general, photometric ratios are invariant for reflectance functions that are separable with respect to incident light angular variables, and, viewing angular variables. Corresponding photometric ratios is equivalent to corresponding classes of well-defined physical constraints on object points, which makes this method formally robust. Furthermore, correspondence of photometric ratios can be done to subpixel accuracy using interpolation. We examine the *isoratio image curves* produced from these geometric physical constraints, which are image curves with equal photometric ratio. Due to their invariance to viewpoint and surface albedo, isoratio image curves may be useful for object recognition.

A technique that is somewhat related to our stereo methodology is "dual photometric stereo" pioneered by Ikeuchi [9]. There are however major conceptual and implementation differences. Ikeuchi does have the extra information provided by surface orientation to refine his depth map. However, his method is restricted to nearly orthographic views and known incident orientation of at least 3 distant light sources. The 3-D stereo method using multiple illumination that we are proposing does not require precise knowledge of any of the multiple illumination conditions and is applicable to full perspective views.

The experimentation that we present with ground truth objects shows that we can determine the shape of cylindrical and spherical objects, using stereo correspondence of photometric ratios, with a depth accuracy to well within 1%. Our stereo correspondence algorithm can recover the shape of more complex surfaces (like a face mask) with comparable accuracy.

2 Problem Background

To describe the problematic issues of comparing image intensities between a stereo pair of cameras we need to understand the image formation process. We describe the formation of image intensity values beginning with the familiar relation from Horn and Sjoberg[8]:

$$E = L_r(\pi/4)\ (D/i)^2 \cos^4\alpha \qquad (1)$$

which relates image irradiance, E, to reflected radiance, L_r. Let D be the lens diameter, i the image distance, and α the light angle relative to the optic axis incident on the camera lens. Equation (1) assumes ideal pinhole optics. The effective diameter, D, of a lens can be controlled with an aperture iris the size of which is measured on

an F-stop scale. Image irradiance is therefore very sensitive to F-stop. While a stereo pair of cameras can use identical model lenses at exactly the same F-stop setting, the effective lens diameters can still be slightly different. The focal lengths as well can be slightly different and by the classical thin lens law this will influence the image distance, i, in turn effecting the image irradiance. Even in the ideal case where focal lengths are precisely equal, the image distance, i, can be slightly different for a stereo pair of images even though the images "appear" equivalently in focus. On top of all this is the dependence of image irradiance in perspective images on pixel location relative to the optical center of the image plane. The farther a pixel is radially away from the optical center, the larger is the light incident angle, α, which strongly affects image irradiance. Image irradiances arising from the same object point appear in different parts of a stereo pair of images making them difficult to compare.

Equation (1) only takes into account the optics involved in image formation. Image irradiance is converted into pixel gray value using electronics. In general, the conversion of image irradiance, E, into pixel gray value, I, can be described by the expression

$$I = gE^{1/\gamma} + d \tag{2}$$

where g is termed the *gain*, d is the *dark reference*, and, γ, controls the non-linearity of gray level contrast. It is typically easy to set $\gamma = 1.0$ producing a linear response, and easy to take a dark reference image with the lens cap on, then subtracting d out from captured images. However, we have observed that not only can the gain, g, be variable between identical model cameras but this can change over time especially for relatively small changes in temperature. Unless g is calibrated frequently, comparing pixel gray values for identical image irradiances between a stereo pair of cameras can be difficult.

A widely used assumption about diffuse reflection from materials is that they are Lambertian [10] meaning that light radiance, L, incident through solid angle, $d\omega$, at angle of incidence, ψ, on an object point with diffuse albedo ρ, produces reflected radiance:

$$L \cdot \rho \cdot \cos\psi \cdot d\omega$$

independent of viewing angle. The independence of diffuse reflected radiance with respect to viewing angle makes it theoretically feasible to associate radiance values with object points in a stereo pair of images.

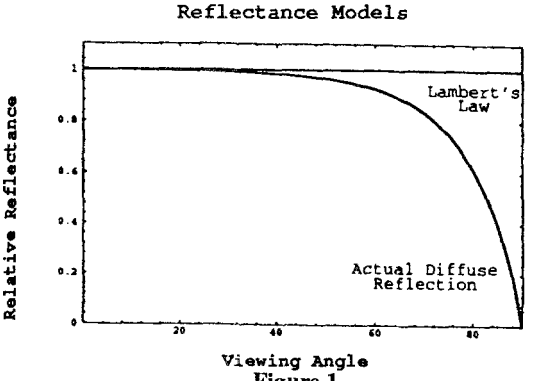


Figure 1

The physical reality of diffuse reflection makes it even more practically difficult to associate diffuse reflected radiance with object points across a stereo pair of images. A recently proposed diffuse reflectance model for smooth dielectric surfaces [19], [20], empirically verified to be more accurate than Lambert's Law, expresses the dependence of diffuse reflected radiance on both angle of incidence, ψ, and viewer angle, ϕ, as

$$L \cdot \rho \cdot [1 - F(\psi, n)] \cdot \cos\psi \cdot \left[1 - F\left(\sin^{-1}\left(\frac{\sin\phi}{n}\right), 1/n\right)\right] \cdot d\omega \qquad (3)$$

where the functions, $F()$, are the Fresnel reflection coefficients [16], and, n, is the index of refraction of the dielectric surface, and, ρ, is the diffuse albedo computed from physical parameters. Figure 1 shows the significant dependence of diffuse reflection upon viewer angle.

Diffuse reflected radiance from an object point as seen from the two different viewpoints of a stereo pair of cameras will almost always not be equal. The dependence of specular reflection upon viewpoint is even more severe due to its highly directional nature and the geometry of angle of incidence equals angle of reflection.

3 Using Photometric Ratios for 3-D Stereo

We show that the ratio image produced by dividing 2 images of diffuse reflection from the same scene with respect to 2 different illumination conditions is invariant to the differences in physical characteristics of cameras discussed in the previous section, as well as viewpoint and diffuse surface albedo. Furthermore, these photometric ratios can be associated with well-defined geometric physical constraints on object points making them suitable for robust correspondence in a stereo pair of images.

3.1 Photometric Ratios as an Invariant

Combining equations (1), (2) and (3) gives us an expression which precisely relates pixel gray value, I, to diffuse reflection as a function of imaging geometry and camera parameters. For incident radiance L through a small solid angle, $d\omega$, at an angle of incidence, ψ, the gray value formed from viewing angle, ϕ (assuming we subtract out the dark reference, d) is:

$$I = g\{(\pi/4)(D/i)^2 \cos^4\alpha (L\rho[1 - F(\psi, n)] \cos\psi \left[1 - F\left(\sin^{-1}\left(\frac{\sin\phi}{n}\right), 1/n\right)\right] d\omega)\}^{1/\gamma}$$

For a general incident radiance distribution, $L(\psi,\theta)$, (θ is azimuth about the surface normal) with respect to an object point, integrating over the incident hemisphere, the gray value will be:

$$I = g\{(\frac{\pi}{4})(\frac{D}{i})^2 \cos^4\alpha \int (L(\psi, \theta)\rho[1 - F(\psi, n)] \cos\psi \left[1 - F\left(\sin^{-1}\left(\frac{\sin\phi}{n}\right), \frac{1}{n}\right)\right] d\omega)\}^{1/\gamma}$$

$$= g\{(\frac{\pi}{4})(\frac{D}{i})^2 \cos^4\alpha \left[1 - F\left(\sin^{-1}\left(\frac{\sin\phi}{n}\right), \frac{1}{n}\right)\right]\rho\}^{1/\gamma} \{\int L(\psi, \theta)[1 - F(\psi, n)] \cos\psi d\omega\}^{1/\gamma} \qquad (4)$$

Consider this object point first illuminated with an incident radiance distribution, $L_1(\psi,\theta)$, and then illuminated with an incident radiance distribution, $L_2(\psi,\theta)$. From

equation (4) the photometric ratio of gray values is:

$$\frac{I_1}{I_2} = \frac{\{\int L_1(\psi, \theta)\,[1 - F(\psi, n)]\cos\psi d\omega\}^{1/\gamma}}{\{\int L_2(\psi, \theta)\,[1 - F(\psi, n)]\cos\psi d\omega\}^{1/\gamma}} \tag{5}$$

where the terms outside the integral signs in the numerator and the denominator cancel out. This photometric ratio expresses a well-defined geometric physical constraint determined by an illumination distribution relative to surface orientation at an object point. It is a function of only incident light geometry relative to local surface orientation, and is invariant to all camera parameters (except γ), viewing angle, ϕ, and diffuse surface albedo, ρ.

It is important to note that the reason why the viewing angle cancels out in the photometric ratio, expression (5), is because expression (3) for diffuse reflection is a separable function with respect to incident and viewing angular variables, ψ, and, ϕ. In general, it is clear that this photometric ratio is invariant to viewing for any reflectance function that is separable with respect to incident light and viewing angular variables, including azimuth dependent reflectance functions. This photometric ratio will not be such an invariant for combined specular and diffuse reflection, since specular reflection at an object point adds the term

$$L \cdot F(\psi, n) \cdot \delta\ (\psi - \phi) \cdot \delta\ (\theta_0 + 180 - \theta)$$

(incidence angle ψ, incident azimuth angle θ_0, emittance angle ϕ and emittance azimuth angle θ) to expression (3), and therefore combined specular and diffuse reflection is not separable with respect to ψ and ϕ.

3.2 Isoratio Image Curves and Physical Constraints

The ratio of equation (5) expresses a physical constraint consisting of the interrelationship between the local surface orientation at an object point and the two illumination distributions, $L_1(\psi,\theta)$, and, $L_2(\psi,\theta)$. Object points producing the same photometric ratios form equivalence classes that project onto the image plane forming what we term in this paper *isoratio image curves*. Different than isophotes which are image curves of equal gray value, an object point belongs to an isoratio image curve based on the geometric relationship of its surface normal with respect to two illuminations, independent of diffuse surface albedo, and independent of viewpoint even as diffuse reflection is viewpoint dependent. Corresponding photometric ratios along epipolar lines between a stereo pair of images is identical to corresponding points that are at the intersection of isoratio image curves and the epipolar lines. For best correspondence of photometric ratios we would like isoratio image curves to intersect epipolar lines as perpendicularly as possible. This tends to maximize the spatial variability (i.e., gradient) of photometric ratios along an epipolar line so that correspondences can be more accurately localized. See [21] for more detailed analysis.

3.3 Corresponding Photometric Ratios

Most cameras have a default setting of linear response (i.e., γ=1.0). If not, the intensity values can be linearized by inverse γ-correction computational processing.

For monotonically varying photometric ratios across an epipolar line, the numerical value I_1/I_2 in the left image can be matched accurately with the corresponding one in the right image and by using interpolation we get subpixel accuracy. We have found the photometric ratio I_1/I_2 in general to be remarkably smoothly varying across epipolar lines, particularly as compared with gray values themselves which typically fluctuate more. Below we present two algorithms for non-monotonic distributions of photometric ratios across epipolar lines. The first one is based on the preservation of left-to-right ordering. The second one is a more sophisticated hierarchical algorithm that can handle complex photometric ratio distributions produced from complex object surfaces.

4 Experimental Results

We tested the accuracy of a dense depth map determined from correspondence of photometric ratios between a stereo pair of images on two objects of known ground truth, a cylinder Figure 3 and a sphere Figure 8, and then experimented with the recovery of the depth of a far more complicated surface, a face mask Figure 12.

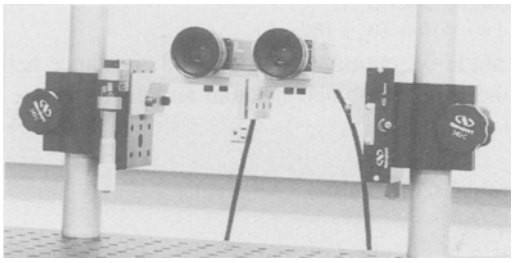


Figure 2

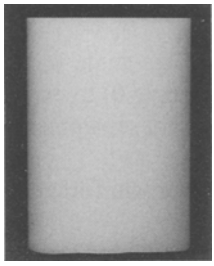


Figure 3

A pair of Sony XC-77 cameras (about 100:1 signal-to-noise ratio) with 25mm lenses were used with a stereo baseline of 3 inches. The cameras were mounted on precision controlled vertical moving platforms, and camera roll was precision controlled so that the corresponding epipolar lines were the scanlines themselves (Figure 2). The radius of the smooth plastic cylinder was precisely machined to 1 3/8 inches, and the radius of the smooth sphere is a precisely 1 3/16 inch radius billiard cue ball. The objects were placed so that their closest point was 20 inches from the stereo baseline which is far relative to the sizes of these objects themselves. At the distance that the cylinder and sphere were placed away from the baseline, the maximum depth variation across each object is about 5% of the total distance from the baseline. This means that for a depth map to have any resemblance to the shape of these objects our depth recovery from stereo correspondence must have an accuracy significantly better than 5%. Each illumination condition was produced from one of 2 point light sources incident at approximately 20° on each side of the horizontal perpendicular bisector of the baseline, and at about 12 inches away from the objects. The light sources were not precision mounted, nor were their incident orientations precisely known. To minimize the occurrence of specularities on the objects, we used cross-polarization.

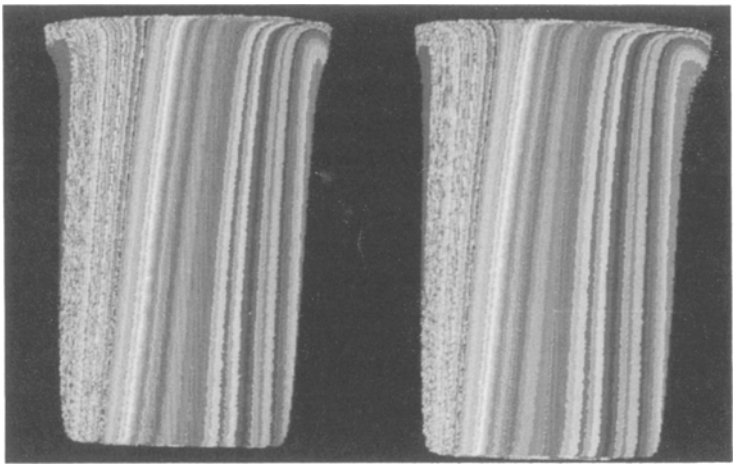


Figure 4

Figures 4 and 6 show the photometric ratio images with respect to the left and right views of a cylinder (shown in Figure 3) and a sphere (shown in Figure 5) respectively . All the photometric ratio images were derived by dividing the image of the object illuminated by the left light source, by the image of the object illuminated by the right light source. Notice that more of the left side of the cylinder is visible from the left view than from the right view, as can be seen by the width of the left-most dark isoratio curve. The inverse is noticeable on the right side of the cylinder. The shape of the isoratio image curves on the various objects shows the well-ordered structure that is induced on objects by photometric ratios. Note the two tiny spots near the center of the sphere on each of the isoratio image curve images. These are due to parts of specularities that we did not completely cancel out by cross polarization.

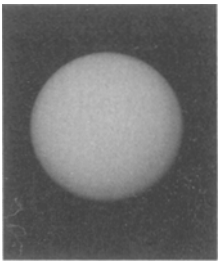


Figure 5

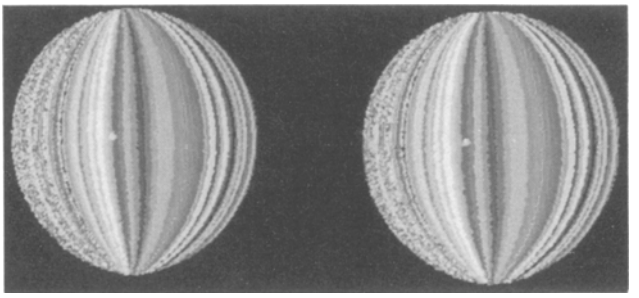


Figure 6

Figures 7, 8 show the actual photometric ratio values going across a pair of corresponding epipolar horizontal scanlines in the left and right images, approximately midway through the cylinder, and the sphere respectively. In Figure 8 the displayed corresponding epipolar scanlines pass through where specular reflections are produced from the left light source (the small peak) and from the right light source (the small valley).

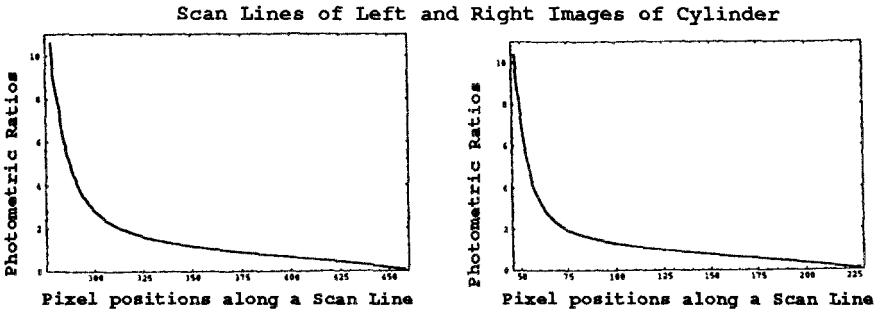


Figure 7

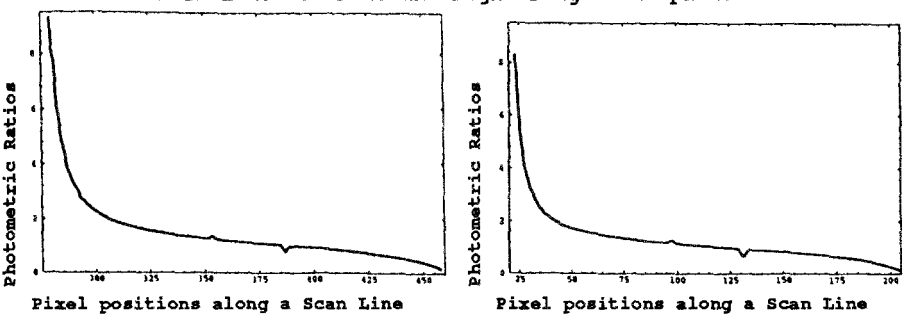


Figure 8

For these simpler objects we had success with the following photometric ratio matching algorithm.

Correspondence_Algorithm_1

For each scan line

Repeat

1. In the **left image**, find the *leftmost* uncorresponded pixel that has some ratio value.
2. In the **right image**, find the corresponding pixel, which has to be to the *right* of the last match. To do this:
   * either find the *leftmost* pixel with the same ratio value;
   * or a) find the *two adjacent leftmost* pixels with values larger and smaller than that ratio value.
     b) interpolate linearly.

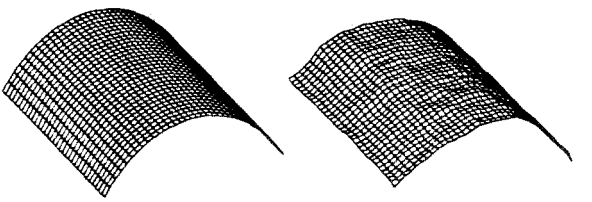


Figure 9

Figure 9 shows the ground truth depth map of the cylinder (left) and the depth

map of that cylinder derived from Correspondence_Algorithm_1 (right). Similarly, Figure 10 shows the ground truth depth map of the sphere (left) and the depth map of the sphere derived from that stereo correspondence algorithm (right). The slight "dimple" near the center of the recovered sphere depth map is caused by the slight presence of specular reflection. The average depth error for the cylinder compared to the ground truth was 0.17 inches (0.85% depth variation at 20"). The average depth error across the sphere, was 0.09 inches (0.45% depth variation at 20")!

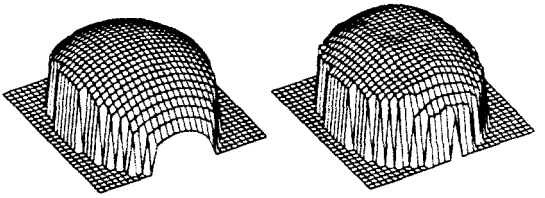


Figure 10

Figure 11 shows the actual image gray values across the same epipolar scanlines as Figure 7. There are noisy fluctuations in these image gray values, which cause great difficulty to the correspondence problem. Notice also that the left-hand portion of the left graph in Figure 11 is generally slightly higher than the same left-hand portion of the right graph in Figure 11. The reverse effect is noticeable at the right-hand portion of the graphs. This is due to the dependence of diffuse reflection on viewing angle, ϕ, as described by equation (3). These differences are compounded by the slightly different camera characteristics in the cameras used in the stereo pair resulting in large errors.

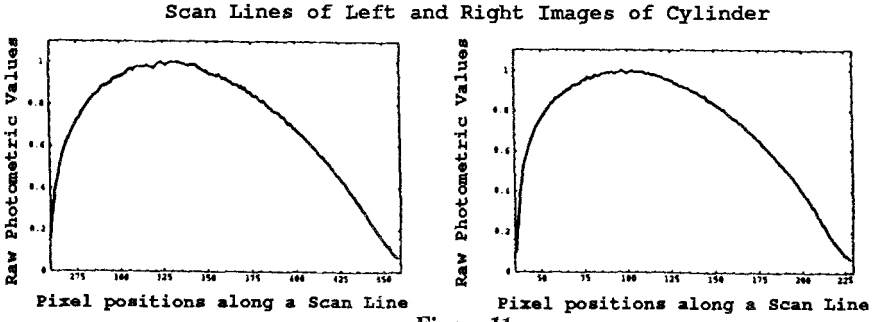


Figure 11

Figure 12 shows a much more complicated surface, a face mask. Figure 13 shows the resulting isoratio image curves for left and right views. Due to the significant complexity of the surface of the face mask some isoratio image curves appear thicker and sometimes more like "regions" rather than curves on some of the flatter portions of the mask.

Figure 14 shows photometric ratio distributions across corresponding epipolar scanlines going through the eyes and the bridge of the nose on the face mask. This is an example of one of the more difficult photometric ratio distributions to correspond. There are significantly wide flat plateaus across which there are many minor fluctuations due to noise that do not consistently appear across corresponding epipolar scanlines. It is straightforward to distinguish peaks and valleys produced by noise, from true peaks and valleys caused by real object characteristics.

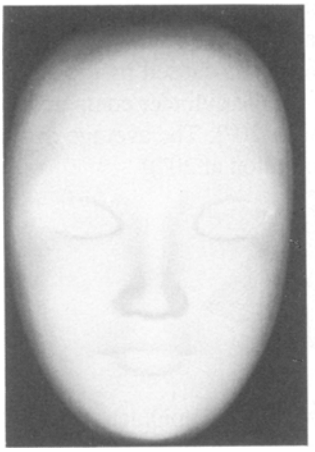
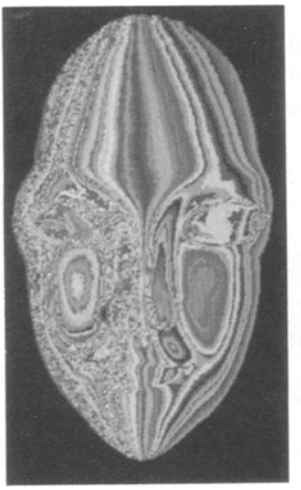
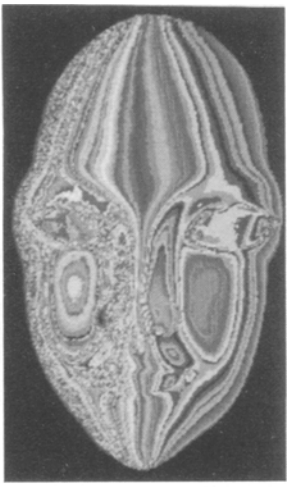


Figure 12



(a)



(b)

Figure 13

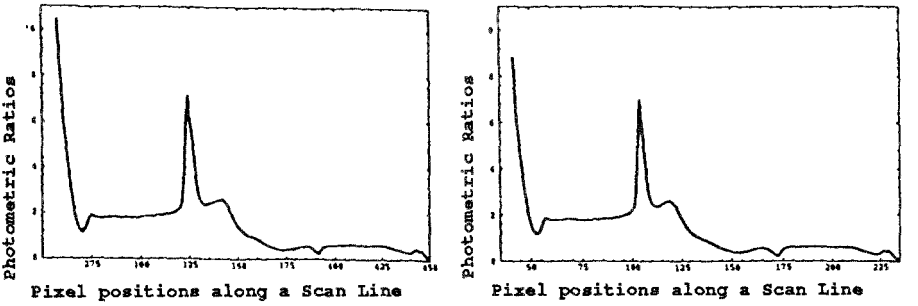


Figure 14

Let the *degree of non-monotonicity* of a photometric ratio distribution across an epipolar line be the total number of occurring maxima and minima. The two properties most salient to peaks and valleys due to noise are that: (i) their degree of non-monotonicity is high over small local pixel neighborhoods, typically of degree at least 3 within a 4 pixel neighborhood; (ii) they typically have a much smaller photometric ratio gradient, ε, than true peaks and valleys. These observations motivated us to develop the following hierarchical stereo matching algorithm.

Correspondence_Algorithm_2

For each scan line

Set the gradient ε to some prescribed lower bound

Repeat for current ε

Repeat

Choose the *uncorresponded* non-noisy pixels in the **left image**, that satisfy ε. Run Correspondence_Algorithm_1 on these pixels, using the previously corresponded pixels to calculate the region where the matching pixel should lie.

Until no more pixels
Decrease ε
Until ε very small
Linearly interpolate uncorresponded pixels.

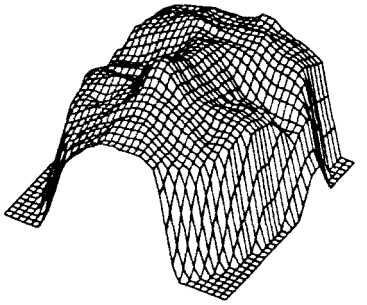


Figure 15

The depth map in Figure 15 was produced by using Correspondence_Algorithm_ 2 starting with a lower bound of photometric ratio gradient of 0.035, successively reducing it by 0.005 for a total of six passes. The depth accuracy attained appears to be on the same order as that for the recovery of the depth maps for the cylinder and the sphere.

5 Conclusions and Future Work

The stereo method presented here using correspondence of photometric ratios can be a very practical tool in machine vision, and could also be very useful for computing the shape of small smooth particles in microscopy. Correspondence of photometric ratios requires minimal calibration, does not require a sophisticated lighting set-up, subpixel accuracy of corresponding photometric ratios can be more precise than corresponding image edge features from projected grids (photometric ratios provide denser depth maps literally providing depth data at every single pixel where there is diffuse reflection), and, is relatively computationaly inexpensive.

In addition to proposing photometric ratios as a reliable way of corresponding a stereo pair of images, we introduced the notion of the isoratio image curve which (unlike isophotes) are invariant to diffuse surface albedo. Therefore isoratio curves are more directly related to the actual geometry of the surface itself and can yield more of information for object recognition. We are currently studying the use of isoratio image curves for improving the performance of object recognition in robotic environments.

References

1. N. Ayache. *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception.* MIT Press, 1989.
2. A. Blake. "Specular Stereo." *Proceedings of IJCAI*, pp. 973-976, 1985.
3. G.J. Brelstaff and A. Blake."Detecting specular reflections using Lambertian

constraints." *Proceedings of the IEEE Second International Conference on Computer Vision (ICCV)*, pp.297-302, Tampa, Florida, December 1988.

4. D. Clarke and J.F. Grainger. *Polarized Light and Optical Measurement*. Pergamon Press, 1971.
5. W.E.L. Grimson. *From Images to Surfaces: A Computational Study of the Human Early Visual System*. MIT Press, 1981.
6. W.E.L. Grimson. "Binocular Shading and Visual Surface Reconstruction." *Computer Vision Graphics and Image Processing*, 28 (1): 19-43, 1984.
7. B.K.P. Horn. "Understanding Image Intensities." *Artificial Intelligence*, pp. 1-31, 1977.
8. B.K.P. Horn and R.W. Sjoberg. "Calculating the Reflectance Map." *Applied Optics*, 18 (11): 1770-1779, June 1979.
9. K. Ikeuchi. "Determining a Depth Map Using a Dual Photometric Stereo." *International Journal of Robotics Research*, 6 (1): 15-31, 1987.
10. J.H. Lambert. "Photometria sive de mensura de gratibus luminis, colorum et umbrae" *Eberhard Klett*. Ausberg, Germany, 1760.
11. D.Marr. Vision. Freeman, San Francisco, 1982.
12. D. Marr and T. Poggio. "A theory of human vision." *Proceedings of the Royal Society of London*. B, 204: 301-328, 1979.
13. J.E.W. Mayhew and J.P. Frisby. *3D Model Recognition from Stereoscopic Cues*. MIT Press, 1991.
14. B.T. Phong. "Illumination for computer generated images." *Communications of the ACM*, 18(6): 311-317, June 1975.
15. S. Pollard and J. Mayhew and J. Frisby. "PMF: a stereo correspondence algorithm using the disparity gradient limit." *Perception*, 14: 449-470, 1985.
16. R. Siegal and J.R. Howell. *Thermal Radiation Heat Transfer*. McGraw-Hill, 1981.
17. G.B. Smith. "Stereo Integral Equation." *Proceedings of the AAAI*, pp. 689-694, 1986.
18. D. Terzopoulos. "The role of constraints and discontinuities in visible-surface reconstruction." *Proceedings of IJCAI*, pp. 1073-1077, 1983.
19. L.B. Wolff. "Diffuse Reflection." *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 472-478, June 1992.
20. L.B. Wolff. "Diffuse and Specular Reflection." *Proceedings of the DARPA Image Understanding Workshop*, April 1993.
21. L.B. Wolff and E. Angelopoulou. "3-D Stereo Using Photometric Ratios." *Johns Hopkins University Technical Report CS-93-10*, October 1993.
22. L.B. Wolff and T.E. Boult. "Constraining Object Features Using a Polarization Reflectance Model." *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 13 (7): 635-657, July 1991.
23. R.J. Woodham. "Reflectance map techniques for analyzing surface defects in metal castings." *Ph.D. thesis, MIT AI Lab Tech Report AI-TR-457*, June 1978.
24. A. Zisserman and P. Giblin and A. Blake. "The information available to a moving observer from specularities". *Image and Vision Computing*, 7 (1): 38-42, 1989.