A ROBUST SHAPE-FROM-SHADING USING MULTIPLE SURFACE NORMAL APPROXIMATIONS

Osamu Ikeda
Faculty of Engineering, Takushoku University
815-1 Tate, Hachioji, Tokyo, 193-0985 Japan

ABSTRACT

We present a robust shape reconstruction method from a single shading image, which is able to give good estimate especially for the light direction having a relatively large slant angle. We derive four iterative shape-estimating relations by using the Jacobi iterative method and the four approximations of the surface normal, and combine them with weights to get a single relation. We show that it is stable enough to attain convergence. We also optimize the light direction through an eigenvalue analysis for the matrix of the combined iterative relation, to improve the shape by avoiding the small eigenvalues that cause serious distortions. Numerical examples using both synthetic and real images show the usefulness of the new method.

1. INTRODUCTION

Shape estimation from a single shading image has been studied, producing a variety of approaches, called minimization [1], [2], linear [3], propagation [4], [5], and deformable model [6]. They, however, still fail to give good shapes partly because iterative operations employed tend to lack numerical stability. For example, the minimization approach presented by Zheng et al. extrapolates the surface normals to estimate them on the boundaries [1]. This local estimation may affect the surface normal estimation on the whole region through iterations, which may cause the numerical instability. They stop the iteration to avoid the instability. The approach given by Tsai et al. linearizes the reflectance map in a single depth parameter to iteratively estimate the shape [3]. The estimation is subject to the consistency between the image and the map and it uses the Lambertian assumption, so that instability always occurs on the brightest parts where the surface normal and the illuminant vectors are parallel. To avoid it, they restrict the iteration in number strong enough to be able to regard their method to be just another transformation of the image to a depth distribution. As a result, the light direction does not play a great role in reconstructing the shape, giving inaccurate shapes for many images. Propagation approaches estimate shapes starting from some initial curves at special points as the brightest or the darkest [5]. If there exist only a few such special points, the methods may work and give good shapes, provided that the illumination is right from the frontal. But if there are many such points, which may be usual, a large fraction of the image may have to be normalized to a value less than unity, as they did, in order to avoid the complex processing; otherwise instability may occur. The normalization, however, may make the resulting shapes inaccurate. The method using the deformable model has a stabilization factor in the estimation, which was introduced in [7] to give the damping effect. The method is applicable to a relatively wide range of illuminating direction, but as far as the numerical results reported are concerned, it appears that the shape accuracy is sacrificed to a certain degree in return for the stability.

In this paper we present a robust shape-from-shading approach, aimed both at realizing stable estimation of the shape and at improving the shape by getting rid of possible distortions. We derive four iterative relations for shape estimation using the four approximations of the surface normal, and combine them with weights to get a single relation. This makes it possible to enhance robustness to a great degree, which may lead to better estimates. Also we optimize the light direction in slant angle through the eigenvalue analysis for the matrix of the combined iterative relation, to improve the shape by avoiding the small eigenvalues that cause serious distortions. We use a combination of the rank of the matrix and the minimal eigenvalue to construct a criterion for the optimal slant angle.

Our method works well basically for light directions with relatively large slant angles, just as humans can best guess shapes of objects by shedding light at them not from the front but rather from the side.

2. EXTENDED ITERATIVE RELATION

The consistency between a shading image, $I(x,y)$, and the reflectance map, $R(p,q)$, is used:

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

0-7803-8554-3/04/$20.00 ©2004 IEEE.

where x,y=1,…,N, and the image is assumed to be normalized to unity. Let **P** and **S** be the surface normal vector of the shape, $z(x,y)$, and the light direction vector, respectively:

$$\mathbf{P} = (p, q, 1)^T / P_0 \tag{2}$$

$$\mathbf{S} = (S_x, S_y, S_z)^T / S_0 \tag{3}$$

$$p = -\partial z/\partial x,\ q = -\partial z/\partial y \tag{4}$$

where P_0 and S_0 are normalizing factors. Then, assuming the *Lambertian* surface, $R(p,q)$, normalized by the *albedo*, is given by

$$R(p, q) = \mathbf{P} \bullet \mathbf{S} \tag{5}$$

Here we use the following four simplest approximations for p and q:

$$(p, q) = \begin{cases} (z(x-1, y) - z(x, y), z(x, y-1) - z(x, y)) & \text{(6a)} \\ (z(x, y) - z(x+1, y), z(x, y) - z(x, y+1)) & \text{(6b)} \\ (z(x-1, y) - z(x, y), z(x, y) - z(x, y+1)) & \text{(6c)} \\ (z(x, y) - z(x+1, y), z(x, y-1) - z(x, y)) & \text{(6d)} \end{cases}$$

Let the function $f(x,y)$ be defined by

$$f(x, y) \equiv I(x, y) - R(p, q) \tag{7}$$

Applying the Jacobi's iterative method to $f(x,y)$, we obtain the following four relations in correspondence to the four approximations in Eq. (6), respectively:

$$-\mathbf{f}_m^{(n-1)} = \mathbf{g}_m^{(n-1)}(\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}),\ m=1,\ldots,4,\ n=1,2,\ldots \tag{8}$$

where $\mathbf{f}_m$ and $\mathbf{z}$ are N^2-elements column vectors of $f_m(x,y)$ and $z(x,y)$, respectively, and $\mathbf{g}_m$, m=1,2,3,4, are N^2xN^2 elements matrices made of $\partial f(x,y)/\partial z(x,y)$, $\partial f(x,y)/\partial z(x\text{-}1,y)$, $\partial f(x,y)/\partial z(x\text{+}1,y)$, $\partial f(x,y)/\partial z(x,y\text{-}1)$ and/or $\partial f(x,y)/\partial z(x,y\text{+}1)$.

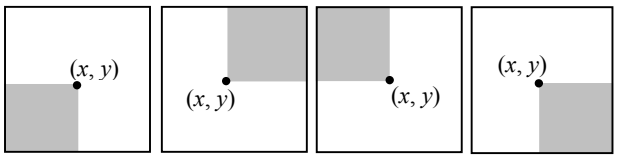


Fig. 1 Integrations of the form $\mathbf{g}_m^{-1}\mathbf{f}_m$ are carried out in the shaded regions to give values at (x, y) for the four approximations in Eq. (6).

The inverses of the four $\mathbf{g}_m$ matrices take values in completely different regions from each other on the (x,y) coordinates, as shown in Fig. 1. The elements of $\mathbf{f}_m$ are multiplied by those of $\mathbf{g}_m^{-1}$ and are integrated, to give values of $\mathbf{g}_m^{-1}\mathbf{f}_m$ at (x,y). So we can obtain four shapes correspondingly to the four different integration regions. Considering that shape distortions may tend to accumulate as the integration region is larger in area, it may be beneficial to weigh the elements of $\mathbf{g}_m$ and $\mathbf{f}_m$ to indirectly average the four results as:

$$-\begin{pmatrix} \mathbf{f}_{w1}^{(n-1)} \\ \mathbf{f}_{w2}^{(n-1)} \\ \mathbf{f}_{w3}^{(n-1)} \\ \mathbf{f}_{w4}^{(n-1)} \end{pmatrix} = \begin{pmatrix} \mathbf{g}_{w1}^{(n-1)} \\ \mathbf{g}_{w2}^{(n-1)} \\ \mathbf{g}_{w3}^{(n-1)} \\ \mathbf{g}_{w4}^{(n-1)} \end{pmatrix} (\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}),\ n=1,2,\ldots \tag{9}$$

where their explicit forms are given later. Using **F** and **G** given by

$$\mathbf{F}^{(n)} = \left((\mathbf{f}_{w1}^{(n)})^T, (\mathbf{f}_{w2}^{(n)})^T, (\mathbf{f}_{w3}^{(n)})^T, (\mathbf{f}_{w4}^{(n)})^T \right)^T \tag{10}$$

$$\mathbf{G}^{(n)} = \left((\mathbf{g}_{w1}^{(n)})^T, (\mathbf{g}_{w2}^{(n)})^T, (\mathbf{g}_{w3}^{(n)})^T, (\mathbf{g}_{w4}^{(n)})^T \right)^T \tag{11}$$

Eq. (9) is rewritten as

$$-\mathbf{F}^{(n-1)} = \mathbf{G}^{(n-1)}(\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}),\ n=1,2,\ldots \tag{12}$$

Then, following the least square error procedure, the shape is estimated following the iterative relation:

$$\mathbf{z}^{(n)} = \mathbf{z}^{(n-1)} - \left\{ (\mathbf{G}^{(n-1)})^T (\mathbf{G}^{(n-1)}) \right\}^{-1} \left\{ (\mathbf{G}^{(n-1)})^T \mathbf{F}^{(n-1)} \right\} \tag{13}$$

with $\mathbf{z}^{(0)}$=**0** as typical initial values, where n =1,2,….

In Eq. (9), the two terms, $\mathbf{g}_{wm}$ and $\mathbf{f}_{wm}$, are given, respectively, by

$$\mathbf{g}_{wm}^{(n)} = [w_{m_i} g_{m_{i,j}}^{(n)}] \tag{14}$$

$$\mathbf{f}_{wm}^{(n)} = [w_{m_j} f_{m_j}^{(n)}] \tag{15}$$

In the same way **z** is expressed as

$$\mathbf{z}^{(\mathbf{n})} = [z_i^{(n)}] \tag{16}$$

where i or j =x+Ny; that is, the weights vary depending on the coordinates (x,y). Then, the terms, $\mathbf{G}^T\mathbf{G}$ and $\mathbf{G}^T\mathbf{F}$ in Eq. (13), are given, respectively, by

$$\mathbf{G}^{(\mathbf{n})^T} \mathbf{G}^{(\mathbf{n})} = \left[\sum_{m=1}^{4} \sum_{k=1}^{N^2} w_{m_k}^2 g_{m_{k,i}}^{(n)} g_{m_{k,j}}^{(n)} \right] \tag{17}$$

$$\mathbf{G}^{(\mathbf{n})^T} \mathbf{F}^{(\mathbf{n})} = \left[\sum_{m=1}^{4} \sum_{k=1}^{N^2} w_{m_k}^2 g_{m_{k,i}}^{(n)} f_{m_k}^{(n)} \right] \tag{18}$$

It may be seen if we check the distributions of $\mathbf{g}_m^{-1}$ that the effective averaging region may roughly be elliptical around the reconstruction point with the long axis in the direction of the tilt angle of the light direction and that the ellipse varies in shape depending on the tilt angle; for example, it is line-like when the angle is null and it is most circular when the angle is 45 degrees. It may also be seen in terms of the magnitude of the eigenvalue that the two approximations of Eqs. (6a) and (6b) play major roles for the tilt angle in the range of 0 to 90 or 180 to 270 degrees, while those of Eqs. (6c) and (6d) do so for the other ranges.

The matrix, $\mathbf{G}^T\mathbf{G}$, is a sparse one, and we can see that its eigenvalues are given by the diagonal elements:

$$\lambda(x,y) = \left(w_1^2 + w_2^2 + w_3^2 + w_4^2\right)\left(\frac{\partial f(x,y)}{\partial z(x,y)}\right)^2 + \left(w_2^2 + w_4^2\right)\left(\frac{\partial f(x-1,y)}{\partial z(x,y)}\right)^2 + \left(w_1^2 + w_3^2\right)\left(\frac{\partial f(x+1,y)}{\partial z(x,y)}\right)^2 + \left(w_2^2 + w_3^2\right)\left(\frac{\partial f(x,y-1)}{\partial z(x,y)}\right)^2 + \left(w_1^2 + w_4^2\right)\left(\frac{\partial f(x,y+1)}{\partial z(x,y)}\right)^2$$
$$\text{for } 1 \le x \le N \text{ and } 1 \le y \le N \qquad (19)$$

where the terms outside of the region are to be neglected; that is, the eigenvalue consists of four terms along the four edge lines and three terms at the four corners. It is seen that they are non-negative and symmetric, away from the weights. It also may be seen that the five terms have different intensity distributions from each other for the light directions of our interest. So, we may be able to expect that the determinant of $\mathbf{G}^T\mathbf{G}$ has a much more significant value by using the four approximations. This may be crucial for the iteration to converge.

The light direction, given or estimated, is optimized in the reconstruction in the slant angle through the eigenvalue analysis to avoid small eigenvalues that may cause serious shape distortions. Let the minimal eigenvalue, λ_{min}, and the rank, *rank*, of the matrix $\mathbf{G}^T\mathbf{G}$ in the converged state of the iteration be maximal for the values of S_z in Eq. (3), $S_{z,min}$ and $S_{z,rank}$, respectively. Then, we regard their average value as the optimal one, $S_{z,opt}$; in other words, we adopt the shape obtained for this value.

To enhance numerical stability, we enclose the image with a uniform shading part, assuming that the part is flat, and impose z=0 as the boundary condition. In this case we add shaded lines between the image and the added part based on the assumption that the object rests on a stand.

3. COMPUTER EXPERIMENTS

Several images were used to examine the new method. We did not observe instability as far as they are concerned. In this paper we pick up the four images for which the optimization makes a significant difference. Among the four objects, the computer mouse was measured using a laser range scanner. Shading images were synthesized assuming the *Lambertian* surface. The image for the mouse has lateral and longitudinal noise that resulted from errors occurred in converting the three-dimensional measured data to that on the two-dimensional grid. In the reconstruction, we continued the iteration until the average change in depth reaches 10^{-3} to 10^{-4}.

For the semi-sphere object, it is seen in Fig. 2 that using the optimized slant angles significantly reduces the distortions and improves the shape. The profiles of λ_{min} and *rank* in Fig. 3 gives ($S_{z,min}$, $S_{z,rank}$) =(34,52), so that we obtain the average of $S_{z,opt}$=43 for the given S_z=50. For the mouse, a significant improvement is seen with the optimization, as shown in Fig. 4. In this case ($S_{z,min}$,$S_{z,rank}$) are (26,30) for the given S_z=30. For the Mozart sculpture, the optimization also makes a difference, as shown in Fig. 5. In this case ($S_{z,min}$, $S_{z,rank}$) are (32,34) for the given S_z=30. For the real image "Lena", first, its slant and tilt angles were estimated using Zheng and Chellappa's algorithm [1] as σ=58 and τ=3 degrees. As mentioned in Sec. 2, the distribution of $\mathbf{g}_m^{-1}$ for this light direction is effectively line-like, giving rise some degree of ruggedness varying in y direction to the resulting shape. So the image was rotated counter-clockwise by 42 degrees after adding flat image parts and four shaded lines, as shown in Fig. 6. We obtained ($S_{z,rank}$,$S_{z,min}$)= (35,36), and the shape obtained for the optimal condition appears to be good, as the comparison of the texture-mapped shapes pixel by pixel for three different values of S_z demonstrates.

4. CONCLUSIONS

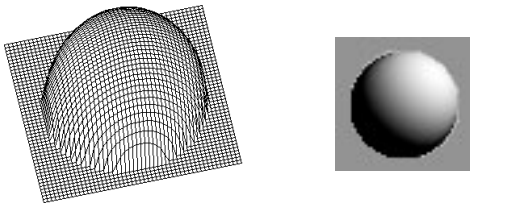
We constructed a robust shape-from-shading algorithm by using the four approximations of the surface normal and by combining the resulting four iterative relations to get a single relation. We also optimized the light direction through the eigenvalue analysis for the matrix of the combined iterative relation, to improve the shape while avoiding the small eigenvalues that causeed serious distortions. The method is able to reconstruct stable and good shapes from images whose light directions have relatively large slant angles.

REFERENCES

[1] Q. Zheng and R. Chellappa, "Estimation of Illuminant Direction, Albedo, and Shape from Shading," *IEEE Trans. PAMI*, vol. 13, pp. 680-702, 1991.

[2] P. L. Worthington and E. R. Hancock, "New Constraints on Data-Closeness and Needle Map Consistency for Shape-from-Shading," *IEEE Trans. PAMI*, vol.21, pp. 1250-1267, 1999.

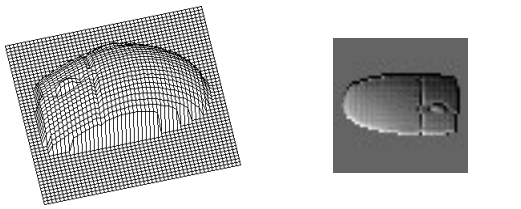
[3] P. S. Tsai and M. Shah, "Shape from Shading Using Linear Approximation," *J. Imaging and Vision Computing*, vol. 12, pp. 487-498, 1994.
[4] M. Bichsel and A. Pentland, "A Simple Algorithm for Shape from Shading," *Proc. CVPR*, pp. 459-465, 1992.
[5] R. Kimmel and A.M. Bruckstein, "Tracking Level Sets by Level Sets: A Method for Solving Shape from Shading Problem," *CVIU*, vol. 62, pp. 47-58, 1995.
[6] D.Samaras and D.Metaxas, Incorporating Illumination Constraints in Deformable Models, *Proc.CVPR*, pp.322-329, 1998.
[7] D.Metaxas and D.Terzopoulos, "Shape and Nonrigid Motion Estimation through Physics-Based Synthesis," *IEEE Trans. PAMI*, vol.15, pp.580-591, 1993.



(a) given shape (b) image for **S**=(50,50,50)

(c) with **S**=(50,50,50) (d) with **S**=(50,50,43)

Fig. 2 Shapes reconstructed from the semi-sphere image without and with the optimized light direction.



(a) given shape (b) image for (-50,-50,30)

(c) with **S**=(-50,-50,30) (d) with **S**=(-50,-50,28)

Fig. 4 Shapes reconstructed from the mouse image without and with the optimized light direction.

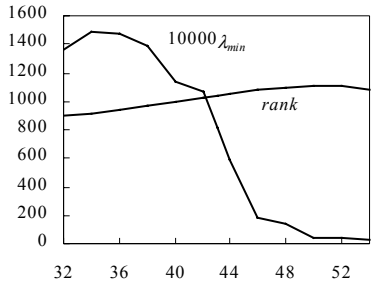
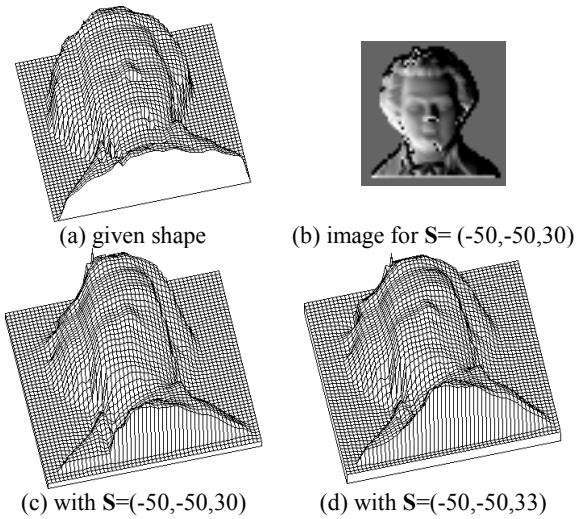


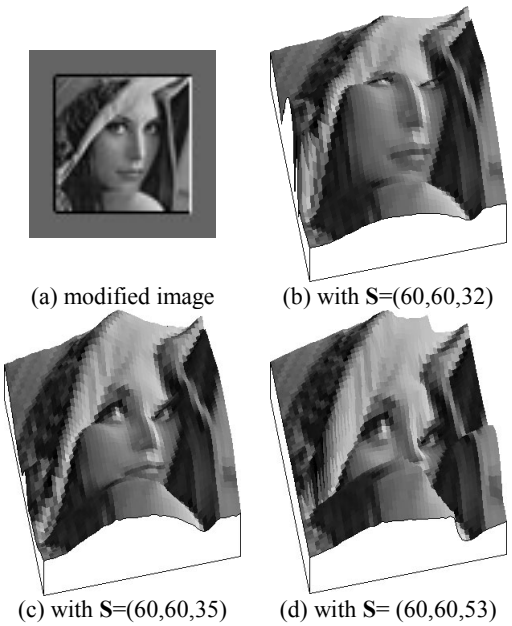
Fig. 3 Profiles of λ_{min} and the rank for the semi-sphere image.



(a) given shape (b) image for **S**= (-50,-50,30)

(c) with **S**=(-50,-50,30) (d) with **S**=(-50,-50,33)

Fig. 5 Shapes reconstructed from the Mozart image without and with the optimized light direction.



(a) modified image (b) with **S**=(60,60,32)

(c) with **S**=(60,60,35) (d) with **S**= (60,60,53)

Fig. 6 Shapes reconstructed from the "Lena" image without and with the optimized light direction.