

Shape-from-Shading for Oblique Lighting with Accuracy Enhancement by Light Direction Optimization

Osamu Ikeda

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

Received 16 December 2004; Revised 12 February 2006; Accepted 18 February 2006

Recommended for Publication by Dimitrios Tzovaras

We present a shape-from-shading approach for oblique lighting with accuracy enhancement by light direction optimization. Based on an application of the Jacobi iterative method to the consistency between the reflectance map and image, four surface normal approximations are introduced and the resulting four iterative relations are combined as constraints to get an iterative relation. The matrix that converts the shading information to the depth is modified so as to be uniform over the whole image region, making the iteration stable and, as a result, the resulting shape more accurate. Then, to enhance the accuracy, the light direction is optimized for slant angle using two criteria based on the initial boundary value and the rank of the converting matrix. The method is examined using synthetic and real images to show that it is superior to the current state-of-the-art methods and that it is effective for oblique light direction whose slant angle ranges from 55 to 75 degrees.

Copyright © 2006 Hindawi Publishing Corporation. All rights reserved.

1. INTRODUCTION

Shape reconstruction from a single shading image has been studied for decades [1], producing a variety of approaches using minimization [2, 3], linearization [4], propagation [5, 6], deformable model [7], viscosity solutions [8–11], an attenuation term of the form $1/r^2$ [12], and the Helmholtz reciprocity [13] that also uses the attenuation term. They, however, still fail to produce acceptable shapes. One of the reasons may be that the iterative operations used in those methods necessitate a tradeoff of accuracy in shape for numerical stability. For example, the minimization approach presented by Zheng and Chellappa extrapolates the surface normals to estimate them on the boundaries [2]. This might appear to be local, but its effects are global through the iteration, causing the numerical instability. They stop the iteration to avoid the instability at the cost of accuracy. The approach given by Tsai and Shah expands the reflectance map in a single linear depth parameter and iteratively estimates the shape [4]. It uses the Lambertian reflection and the consistency between the image and the reflectance map, so that instability occurs at the brightest parts. To avoid it, they limit the number of the iteration, making it difficult to obtain an accurate shape for many images. Propagation approaches estimate shapes starting from some initial curves at such special points as the brightest or the darkest. When many such parts

are present, which may be usual, the image is normalized to a value less than unity [6] to avoid the complex processing of combining many shape patches, making the resulting shapes inaccurate. The method using the deformable model has also a stabilization factor in the estimation, introduced in [14], to give a damping effect. The shape accuracy may be sacrificed in return for stability. The two methods using the attenuation term [12, 13] are stable, but the resulting shapes appear to lack accuracy.

Another reason has to do with the use of parametric constraints and their heuristic optimization. For example, two Lagrange multipliers are used in the minimization approach [2], a single parameter is used in the linear approach in the iterative process of revising the shape [4], normalization of the image to a value less than unity is made in the propagation approach [6], and the initial function is chosen in the method using the deformable model [7], in which case the resulting shape may depend on the function. As an extreme case, methods using viscosity solutions require knowledge of the boundary [11], the heights at the local minimal points [8], or at least part of the shape information on the boundary [15], wherein the Morse functions used to form shapes [16] may lack generality.

In addition, shadows are present in images which have no shading information. The downright light may be best in view of this, but there will be ambiguities between convex

and concave shapes. An oblique light, on the other hand, is most informative from our perceptual viewpoint, but most likely there exist shadows in the image.

In this paper, we present a novel shape-from-shading method, which uses neither adjusting parameters nor a priori or additional information, and which appears more accurate for oblique light cases than the current methods. In this method, based on an application of the Jacobi iterative method to the consistency between the image and the reflectance map, we introduce four surface normal approximations and the resulting four iterative relations are combined as constraints to get an iterative relation. The methods using viscosity solutions also use multiple depths at neighboring grid points. Specifically, they use two gradients in each direction, but this results in spatially blurring shape reconstruction. On the other hand, we use four surface normals, which result in better shape through enhancement of stability. Then, the matrix that converts the shading information to the depth is modified so as to be uniform over the whole image region, making the iteration stable and, as a result, the resulting shape accurate. Then, in order to enhance the accuracy, the light direction is optimized using two criteria based on the initial boundary value in the iteration and the rank of the converting matrix.

2. ITERATIVE RELATION FOR RECONSTRUCTION AND OPTIMIZATION OF LIGHT DIRECTION

We use the consistency between a given image $I(x, y)$ and the reflectance map $R(p, q)$. Let $\mathbf{P} \propto (p, q, 1)^T$ be the surface normal of the object's depth $z(x, y)$, $x, y = 1, \dots, N$, and $\mathbf{S} \propto (S_x, S_y, S_z)^T$ be the light direction, where orthographic projection from a point source is assumed. Then, for the Lambertian surface, the map normalized by the albedo is given by the scalar product of \mathbf{P} and \mathbf{S} :

$$R(p, q) = \frac{pS_x + qS_y + S_z}{\sqrt{p^2 + q^2 + 1}\sqrt{S_x^2 + S_y^2 + S_z^2}}. \quad (1)$$

The Lambertian surface does not represent real surfaces of objects, but is a good approximation if we use polarization filters when taking pictures to eliminate specular reflection components. We do not impose $R(p, q)$ to be smooth. The surface normal components, p and q , are given by $-\partial z/\partial x$ and $-\partial z/\partial y$, respectively, where the negative sign is used for the convenience. Here we consider the four approximations for them as follows:

$$(p, q) = \begin{cases} (z(x-1, y) - z(x, y), z(x, y-1) - z(x, y)), \\ (z(x, y) - z(x+1, y), z(x, y) - z(x, y+1)), \\ (z(x-1, y) - z(x, y), z(x, y) - z(x, y+1)), \\ (z(x, y) - z(x+1, y), z(x, y-1) - z(x, y)), \end{cases} \quad (2)$$

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

$$f_m(x, y) \equiv J_m(x, y) - R_m(p, q), \quad m = 1, \dots, 4, \quad (3)$$

where the image, $J_m(x, y)$, is shifted corresponding to the approximations:

$$J_m(x, y) = \begin{cases} I(x, y) & \text{for } m = 1, \\ I(x+1, y+1) & \text{for } m = 2, \\ I(x, y+1) & \text{for } m = 3, \\ I(x+1, y) & \text{for } m = 4. \end{cases} \quad (4)$$

$I(x, y)$ is normalized to unity, and (p, q) in (2) are used in $R_m(p, q)$, $m = 1, \dots, 4$. The shifts are necessary to avoid the deterioration in shape resolution. Applying the Jacobi iterative method to $f_m(x, y)$, $m = 1, \dots, 4$, we obtain the following four iterative relations, respectively:

$$\begin{aligned} -f_{1,x,y}^{(n-1)} &= \left(\frac{\partial f_{1,x,y}}{\partial z_{x,y}} \right)^{(n-1)} (z_{x,y}^{(n)} - z_{x,y}^{(n-1)}) + \left(\frac{\partial f_{1,x,y}}{\partial z_{x-1,y}} \right)^{(n-1)} \\ &\quad \times (z_{x-1,y}^{(n)} - z_{x-1,y}^{(n-1)}) + \left(\frac{\partial f_{1,x,y}}{\partial z_{x,y-1}} \right)^{(n-1)} (z_{x,y-1}^{(n)} - z_{x,y-1}^{(n-1)}), \\ -f_{2,x,y}^{(n-1)} &= \left(\frac{\partial f_{2,x,y}}{\partial z_{x,y}} \right)^{(n-1)} (z_{x,y}^{(n)} - z_{x,y}^{(n-1)}) + \left(\frac{\partial f_{2,x,y}}{\partial z_{x+1,y}} \right)^{(n-1)} \\ &\quad \times (z_{x+1,y}^{(n)} - z_{x+1,y}^{(n-1)}) + \left(\frac{\partial f_{2,x,y}}{\partial z_{x,y+1}} \right)^{(n-1)} (z_{x,y+1}^{(n)} - z_{x,y+1}^{(n-1)}), \\ -f_{3,x,y}^{(n-1)} &= \left(\frac{\partial f_{3,x,y}}{\partial z_{x,y}} \right)^{(n-1)} (z_{x,y}^{(n)} - z_{x,y}^{(n-1)}) + \left(\frac{\partial f_{3,x,y}}{\partial z_{x-1,y}} \right)^{(n-1)} \\ &\quad \times (z_{x-1,y}^{(n)} - z_{x-1,y}^{(n-1)}) + \left(\frac{\partial f_{3,x,y}}{\partial z_{x,y+1}} \right)^{(n-1)} (z_{x,y+1}^{(n)} - z_{x,y+1}^{(n-1)}), \\ -f_{4,x,y}^{(n-1)} &= \left(\frac{\partial f_{4,x,y}}{\partial z_{x,y}} \right)^{(n-1)} (z_{x,y}^{(n)} - z_{x,y}^{(n-1)}) + \left(\frac{\partial f_{4,x,y}}{\partial z_{x+1,y}} \right)^{(n-1)} \\ &\quad \times (z_{x+1,y}^{(n)} - z_{x+1,y}^{(n-1)}) + \left(\frac{\partial f_{4,x,y}}{\partial z_{x,y-1}} \right)^{(n-1)} (z_{x,y-1}^{(n)} - z_{x,y-1}^{(n-1)}), \end{aligned} \quad (5)$$

where $f_{m,x,y} \equiv f_m(x, y)$ and $z_{x,y} \equiv z(x, y)$. These can be rewritten in matrix form as

$$-\mathbf{f}_m^{(n-1)} = \mathbf{g}_m^{(n-1)} (\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}), \quad m = 1, \dots, 4, \quad n = 1, 2, \dots, \quad (6)$$

where \mathbf{f}_m and \mathbf{z} are N^2 -element column vectors of $f_m(x, y)$ and $z(x, y)$, respectively, and \mathbf{g}_m are $N^2 \times N^2$ -element sparse matrices made of one to three derivatives of $f_m(x, y)$ with respect to $z(x, y)$, $z(x-1, y)$, $z(x+1, y)$, $z(x, y-1)$, or $z(x, y+1)$. The derivatives have positive or negative values.

The inverses of the four \mathbf{g}_m matrices take values in different regions from each other, as shown in Figure 1. The elements of \mathbf{f}_m within these shaded regions are multiplied by those of \mathbf{g}_m^{-1} and are integrated to give values of $\mathbf{g}_m^{-1}\mathbf{f}_m$. For finer details, it is seen from the distribution of the values of \mathbf{g}_m^{-1} that the effective averaging region is roughly elliptical around the reconstruction point with the long axis being in the direction of the tilt angle, τ , of the light direction and that the ellipse is most circular for $\tau = 45 + 90u$, $u = \text{integer}$, while

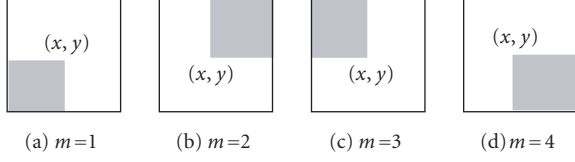


FIGURE 1: Integral operations of the form $\mathbf{g}_m^{-1} \mathbf{f}_m$ are carried out in the different shaded regions to give depth maps for the four different approximations.

it is just a line for $\tau = 90u$. Thus, the method is not specifically sensitive to noise or shadows in the image due to the integration; and in finer details, their effects may be largest for the case of $\tau = 90u$ and smallest for $45 + 90u$.

We combine the four iterative relations as follows:

$$-\begin{pmatrix} \mathbf{f}_1^{(n-1)} \\ \mathbf{f}_2^{(n-1)} \\ \mathbf{f}_3^{(n-1)} \\ \mathbf{f}_4^{(n-1)} \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1^{(n-1)} \\ \mathbf{g}_2^{(n-1)} \\ \mathbf{g}_3^{(n-1)} \\ \mathbf{g}_4^{(n-1)} \end{pmatrix} (\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}), \quad n = 1, 2, \dots \quad (7)$$

Using \mathbf{F} and \mathbf{G} given by

$$\begin{aligned} \mathbf{F}^{(n)} &= \left((\mathbf{f}_1^{(n)})^T, (\mathbf{f}_2^{(n)})^T, (\mathbf{f}_3^{(n)})^T, (\mathbf{f}_4^{(n)})^T \right)^T, \\ \mathbf{G}^{(n)} &= \left((\mathbf{g}_1^{(n)})^T, (\mathbf{g}_2^{(n)})^T, (\mathbf{g}_3^{(n)})^T, (\mathbf{g}_4^{(n)})^T \right)^T. \end{aligned} \quad (8)$$

Equation (7) is rewritten as

$$-\mathbf{F}^{(n-1)} = \mathbf{G}^{(n-1)} (\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}), \quad n = 1, 2, \dots \quad (9)$$

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

$$\mathbf{z}^{(n)} = \mathbf{z}^{(n-1)} - (\mathbf{G}_2^{(n-1)})^{-1} \mathbf{F}_2^{(n-1)}, \quad n = 1, 2, \dots, \quad (10)$$

typically with $\mathbf{z}^{(0)} = \mathbf{0}$ as initial values, where

$$\mathbf{G}_2 = \mathbf{G}^T \mathbf{G}, \quad \mathbf{F}_2 = \mathbf{G}^T \mathbf{F}. \quad (11)$$

Let us express the terms \mathbf{g}_m , \mathbf{f}_m , and \mathbf{z} as

$$\mathbf{g}_m^{(n)} = [\mathbf{g}_{m_{i,j}}^{(n)}], \quad \mathbf{f}_m^{(n)} = [f_{m_j}^{(n)}], \quad \mathbf{z}^{(n)} = [z_i^{(n)}], \quad (12)$$

where i or j is equal to $x + Ny$, then \mathbf{G}_2 and \mathbf{F}_2 are given, respectively, by

$$\mathbf{G}_2^{(n)} = \left[\sum_{m=1}^4 \sum_{k=1}^{N^2} \mathbf{g}_{m_{k,i}}^{(n)} \mathbf{g}_{m_{k,j}}^{(n)} \right], \quad (13)$$

$$\mathbf{F}_2^{(n)} = \left[\sum_{m=1}^4 \sum_{k=1}^{N^2} \mathbf{g}_{m_{k,i}}^{(n)} f_{m_k}^{(n)} \right]. \quad (14)$$

It is seen from (13) that the matrix, \mathbf{G}_2 , is also sparse and its eigenvalues are given by the diagonal elements as

$$\begin{aligned} \lambda(x, y) &= \sum_{m=1}^4 \left(\frac{\partial f_m(x, y)}{\partial z(x, y)} \right)^2 + \sum_{m=2,4} \left(\frac{\partial f_m(x-1, y)}{\partial z(x, y)} \right)^2 \\ &+ \sum_{m=1,3} \left(\frac{\partial f_m(x+1, y)}{\partial z(x, y)} \right)^2 + \sum_{m=2,3} \left(\frac{\partial f_m(x, y-1)}{\partial z(x, y)} \right)^2 \\ &+ \sum_{m=1,4} \left(\frac{\partial f_m(x, y+1)}{\partial z(x, y)} \right)^2, \\ &2 \leq x \leq N-1, \quad 2 \leq y \leq N-1. \end{aligned} \quad (15)$$

The eigenvalues on the four boundary lines are also given by (15) if we retain only those terms within the region of $1 \leq x \leq N$ and $1 \leq y \leq N$. That is, they consist of five kinds of the squared derivatives in the region $2 \leq x \leq N-1$ and $2 \leq y \leq N-1$, four kinds of such terms on the four boundary lines and three kinds of such terms at the four corners.

We can see by inserting (3) and the relevant expressions in (15) that nine depths at (x, y) and at its eight neighboring points contribute to the eigenvalue in the region $2 \leq x \leq N-1$ and $2 \leq y \leq N-1$. Similarly it is seen from (14) that the elements of \mathbf{F}_2 also have similar symmetric expressions. The symmetric property, which is achieved by combining the four approximations, indicates that the reconstruction uses the entire image in correspondence with Figure 1. The symmetry, however, is not complete due to the existence of the nonsymmetric terms on the boundary lines, which possibly seriously damages the shape reconstruction. Those terms could make the determinant of the converting matrix insignificant, making the iteration unstable.

We restrict the reconstruction in the region $2 \leq x \leq N-1$ and $2 \leq y \leq N-1$. In this case all the eigenvalues are given by (15). Letting \mathbf{z}' , \mathbf{G}'_2 , and \mathbf{F}'_2 have the elements of \mathbf{z} , \mathbf{G}_2 , and \mathbf{F}_2 , respectively, in this region, the following iterative relation holds:

$$\mathbf{z}'^{(n)} = \mathbf{z}'^{(n-1)} - (\mathbf{G}'_2)^{-1} \mathbf{F}'_2, \quad n = 1, 2, \dots \quad (16)$$

It is noted that the values of (p, q) in the entire area are still needed, as seen from (15).

We impose the solid boundary condition in order to ensure stability. The depth on the boundary lines is set to the same value as the initial depth value in the iteration, so that we impose the following in each iteration:

$$z(x, y) = 0 \quad \text{except for } 2 \leq x \leq N-2, \quad 2 \leq y \leq N-2. \quad (17)$$

In case the image varies on the boundary lines, where the condition in (17) cannot be applied, we enclose the image with flat-shaped strips whose shading value is determined from the lighting direction, and we shade the boundary between the object and the surrounding flat part on the assumption that the object is positive in depth along the boundary.

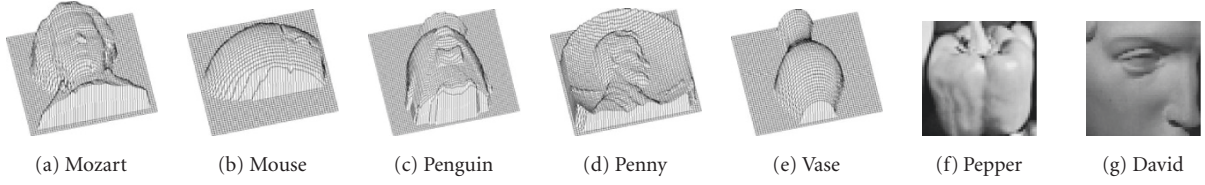


FIGURE 2: Five shapes and two real images used in experiments.

It is possible that this imposition affects the resulting shape around the boundary. In general, as the height of the object on the boundary is larger, the effect may be larger. In most synthetic image examples we have in this paper the height is null on the boundary, so we may have no such effect. In real image examples the height may not be null on the boundary, so the resulting shape around the boundary may be affected by the imposition. The degree and spatial extent of the effect may depend on the reconstruction capability; that is, as the capability is larger, the degree and the affecting area may be smaller. As far as the real images used are concerned, the effect appears to be insignificant. This, in turn, implies that our method is superior in reconstruction capability.

We use two criteria to optimize the light direction. One evaluates how accurately the flat part with the initial null value is reconstructed, and the other evaluates the rank of the converting matrix,

$$\mathbf{S}_{\text{Min } Z} = \arg \text{Max}_{\mathbf{S}} \{ \text{Min}_{(x,y)} \{ z(x,y) \} \}, \quad (18)$$

$$\mathbf{S}_{\text{rank}} = \arg \text{Max}_{\mathbf{S}} \left\{ \sum_{x=2}^{N-1} \sum_{y=2}^{N-1} \frac{\lambda(x,y)}{\text{Max}_{x,y} \{ \lambda(x,y) \}} \right\}. \quad (19)$$

As the light direction is different from the true direction, shape distortions may increase and the minimal depth is observed to usually be smaller than the initial value which is null. And at the same time the number of insignificant eigenvalues may increase. It should be noted that optimization possibly compensates for the insufficient reconstruction characteristic of the method.

3. COMPUTER EXPERIMENTS

Five objects and two real images shown in Figure 2 were used, among which shapes of the mouse and the penguin were measured using a laser range scanner. Some shape errors, generated when converting the three-dimensional data to that on the two-dimensional grid, are noticeable in the synthetic images. Shading images of 50×50 to 96×96 pixels were synthesized from the objects. The number of iterations using (16) was typically 100, resulting in the average change in shape less than 0.1 percent for most cases. Taking into account the orthographic projection, the error of the reconstructed shape was evaluated as

$$e_a = \frac{\text{Min}_c \{ \sum_{(x,y)} |z_{\text{rec}}(x,y) - c - z_{\text{gvn}}(x,y)| \}}{\text{Max}_{(x,y)} \{ |z_{\text{gvn}}(x,y)| \} \sum_{(x,y)}}, \quad (20)$$

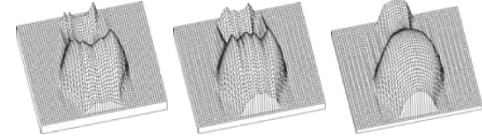


FIGURE 3: Shapes reconstructed from the vase images, where $\mathbf{S} = (0, 1, 1)$ and the depths are, from left, 100, 25, and 12% of the true one. The ratio of the number of zero-valued pixels to that of entire pixels is, from left to right, 1.1, 0.4, and 0%. The reconstructed shapes are normalized in height.

where z_{rec} and z_{gvn} represent the reconstructed and the ground-truth shapes, respectively. As for the real images, the David may have a Lambertian surface to a great degree. The specular reflection components in the pepper image were reduced to create smoother brightness profiles.

The results in Figure 3, which were obtained for vase images for three different magnitudes of the object depth, show that the reconstruction is successful when there is no shadow in the image, while it fails when there are shadows. The effects of shadows are most serious for $\mathbf{S} = (0, 1, 1)$ and $(0, -1, 1)$ due to the symmetry of the shape and the existence of cliffs at the top and bottom. It appears that the variability in object shape does not contribute significantly to the results. The results in Figure 4, which were obtained for the Mozart object with one twentieth of the true height, show that when images have few shadows, the reconstruction is successful for a wide slant angle range of 30 to 87 degrees for the shape error of 5%.

Figure 5 shows shape errors and shadow ratios as a function of slant angle for the five objects. It is seen that shadows increase with increasing slant angle, degrading the accuracy for a large slant angle, and that the shape also tends to be worse for smaller slant angle due to increasing number of singular points. The eigenvalues of the converting matrix are small at singular points, resulting in large depth changes. When the object has a shape as shown in Figure 2, this property may not often give contradictory results for a large slant angle, but it may often give contradictory results for a small slant angle. As a result, the effective slant angle range is restricted to 55 to 75, as shown in Figure 5, where there exists no ambiguity between convex and concave shapes. Examples of reconstructed shapes are shown in Figure 6 for three slant angles of 54.7, 67.0, and 82.0 and for a tilt angle of 0, -45, or 45 degrees. Figure 7 shows shadow ratios and the shape errors as a function of tilt angle, and examples of reconstructed

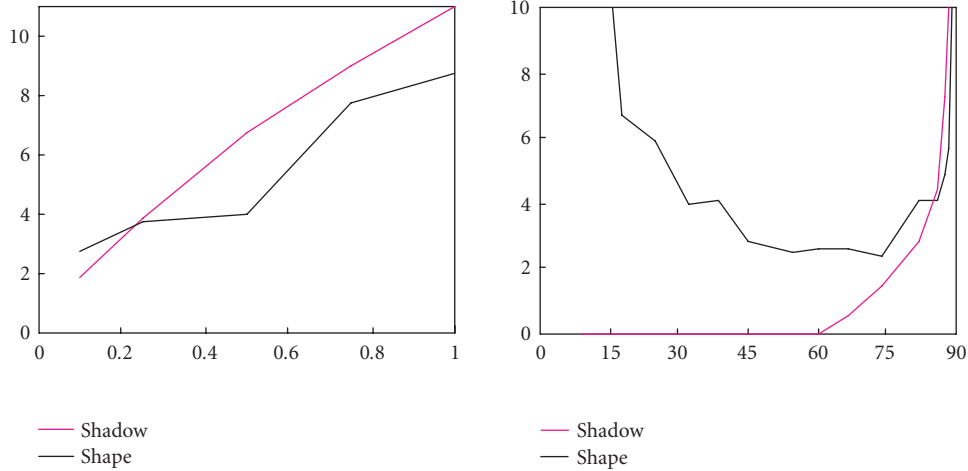


FIGURE 4: Left: shadow ratio and the error of reconstructed shape both in percentage as a function of the ratio of the maximal height to the true height of the Mozart, where $\tau = 45$. Right: the shadow ratio and the reconstructed shape error for $\sigma = 0$ to 90 and $\tau = 45$ for a Mozart object with one twentieth of the true height.

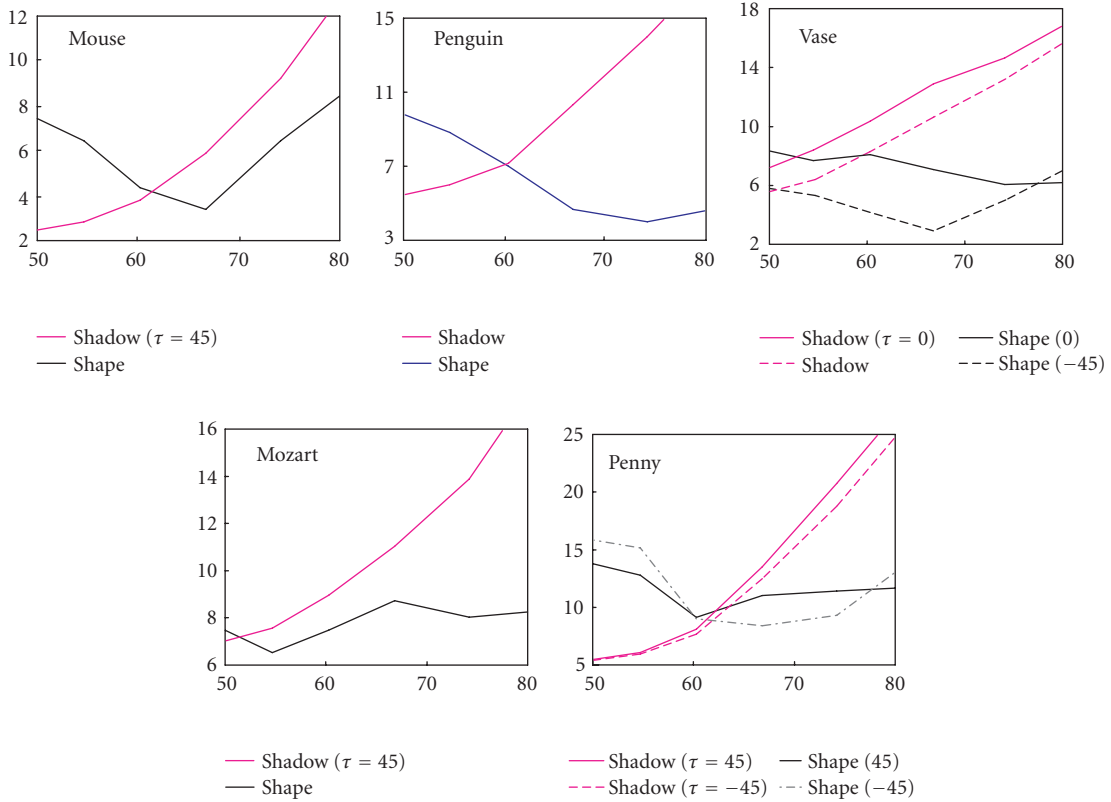


FIGURE 5: Ratios of the number of shadow pixels and shape errors as a function of slant angle for the five objects.

shapes are shown in Figure 8. Figure 7 shows that shapes tend to be worst when the vector normal to the cliff-like part of the object has the same tilt angle as the light. It is in Figure 8 that lighting with a tilt angle of -45 or 45 degrees gives smoother shapes compared to -90 , 0 , or 90 as described previously.

Our method is compared with the current state-of-the-art methods in Table 1, where DM stands for deformable model [7], BEST is a group of six methods [17], and in our method the surface normal components are derived from the shape. Shapes are normalized in height so as to have the same range in order to obtain statistics on shape accuracy

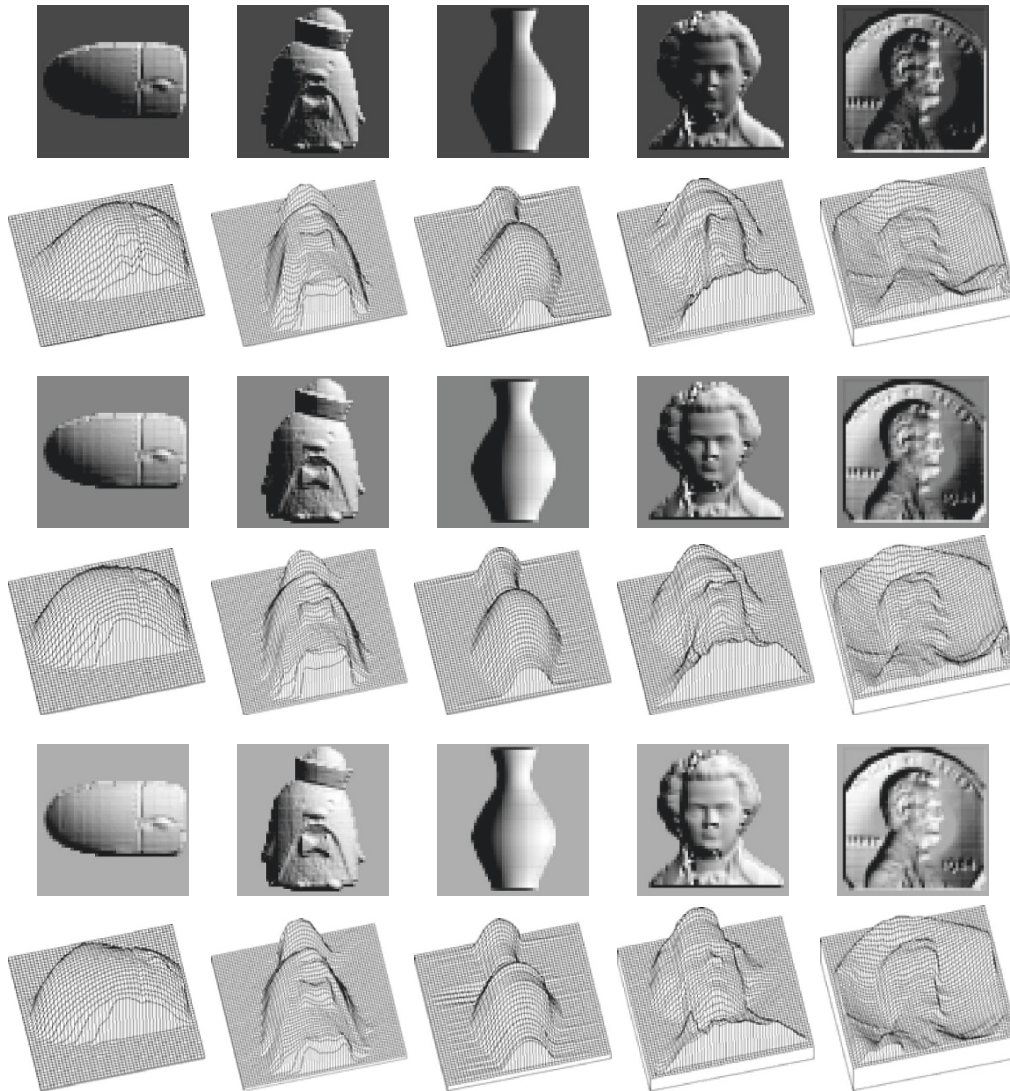


FIGURE 6: Shading images and their reconstructed shapes for three slant angles of, from top, 82.0, 67.0, and 54.7 degrees for the five objects. The tilt angle is 45 degrees for mouse and penguin, 0 for vase, 45 for Mozart, and -45 for penny.

and compare them. The results show that our method is better than those state-of-the-art methods for the three object examples in terms of the absolute depth error and its standard deviation. Especially the small standard deviation values mean that our method can reconstruct similar shapes for different light directions. Our method is also better in terms of the surface normal error, except for the Mozart example, in which case our method is inferior to DM for smoothness of shape.

Figure 9 shows examples of the optimal slant and tilt angles, estimated using the two criteria in (18) and (19), relative to the true ones as a function of the slant angle, where Mozart images for the case of $\mathbf{S} = (5, 5, S_z)$ are used. As shown in Figure 10, it is more advantageous to use S_{rank} to get a better shape than S_{minZ} . It can be observed for those objects that optimization for the slant angle tends to have much more significant effects than that for the tilt angle. The

difference in effects between the two criteria is more clearly seen in the real images of pepper and David, as shown in Figures 11 and 12, respectively. The true light directions given in the references are $(\sigma, \tau) = (45, 40)$ and $(45, 135)$, respectively. It is seen from the results that optimal directions based on (18) are $(59, 40)$ and $(59, 135)$ for pepper and David, respectively, while the directions optimized with respect to the slant angle based on (19) are $(59, 40)$ and $(65.8, 135)$, respectively. It is seen from Figure 11 that optimization improves the shape and from Figure 12 that optimization based on rank improves the shape more than that based on minimal depth. The right cheek of the David is noticeably distorted, but it can be corrected by using a slightly different direction from the optimal light direction, which indicates a need to improve the criterion. Hence, the criterion using rank may be more effective than that using minimal depth.

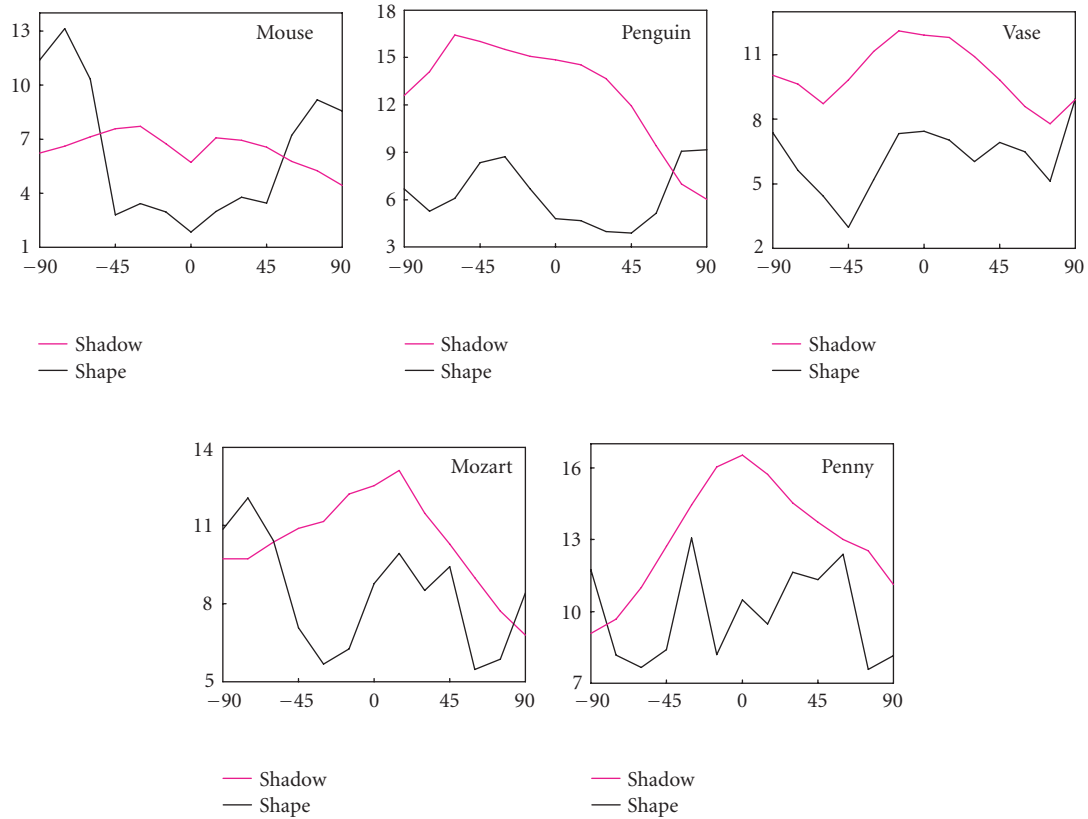


FIGURE 7: Ratios of shadow pixels and the shape errors as a function of tilt angle for the five objects.

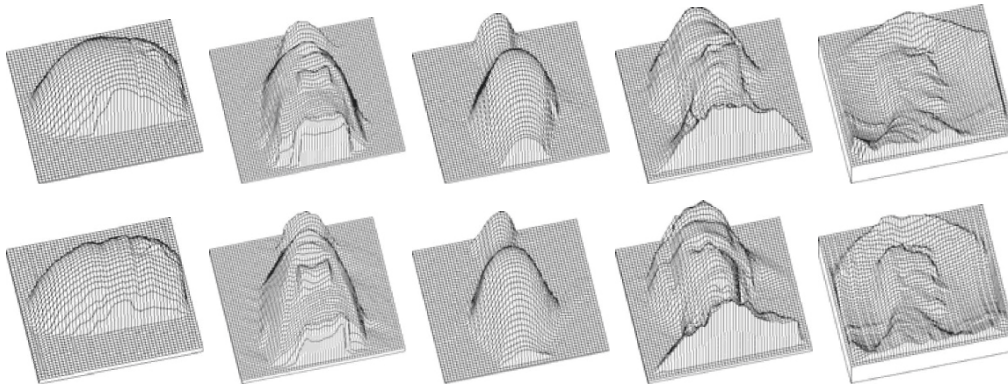


FIGURE 8: Examples of reconstructed shapes, where $S = (71, 71, 42)$ (top) and $(0, 99, 42)$ (bottom) for mouse, $(71, 71, 35)$ and $(71, -71, 35)$ for penguin, $(71, -71, 47)$ and $(71, 71, 47)$ for vase, $(71, 71, 42)$ and $(71, -71, 42)$ for Mozart, and $(71, 71, 42)$ and $(0, -99, 42)$ for penny.

Relatively small-sized images are used in these experiments, but the method can be applied to larger images to reconstruct more details of the shape, as an example is shown in Figure 13.

4. CONCLUSIONS

We presented a shape-from-shading method for oblique lighting with accuracy enhancement by light direction

optimization. Based on an application of the Jacobi iterative method to the consistency between the reflectance map and image, four surface normal approximations were introduced, and the matrix of the resulting relation was made uniform over the image region to obtain a more stable and accurate shape. Then, the light direction was optimized in slant angle based on the rank of the converting matrix to enhance the accuracy. Examination using synthetic and real images showed that the method was superior to the current state-of-the-art

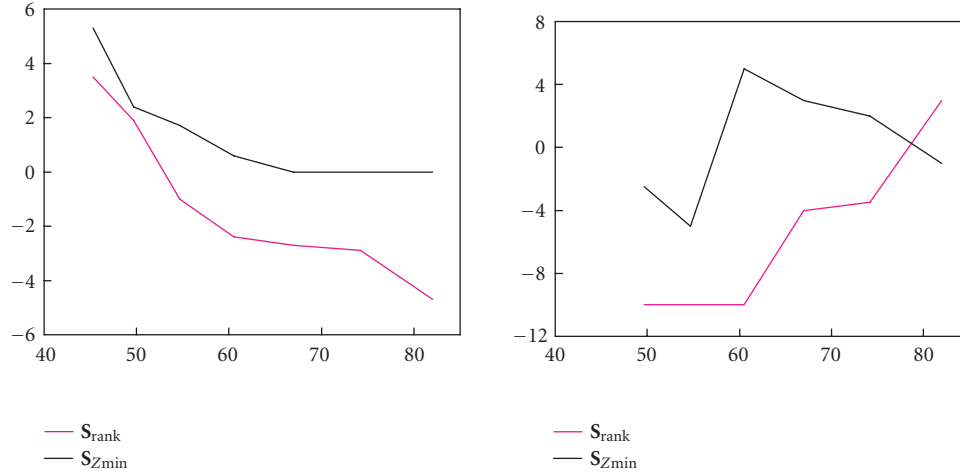


FIGURE 9: Optimal light directions relative to the true ones for slant (left) and tilt (right) angles using the two criteria for the case of Mozart.

TABLE 1: Comparison of our method with the current state-of-the-art methods, BEST and DM, where BEST consists of six methods and its figure means the best value among the six and DM stands for deformable model. The first figure in each cell is for $\mathbf{S} = (5, 5, 7)$ for BEST and DM, while it is an average for $(5, 5, S_z)$, $S_z = 1-5$, for our method, and the second figure is for $\mathbf{S} = (1, 0, 1)$ for BEST and DM, while it is an average for $(7, 0, S_z)$, $S_z = 1-5$, for our method. Z_{avg} means the absolute depth error, std. dev. the standard deviation of the absolute depth error, and (p, q) the surface normal components error.

Methods		Vase	Penny	Mozart
BEST [7, 17]	Z_{avg}	7.5, 7.9	4.7, 4.4	8.8, 7.7
	std. dev.	12.9, 13.9	7.3, 5.5	5.8, 3.5
	(p, q)	0.9, 0.9	1.1, 1.0	0.7, 0.5
DM [7]	Z_{avg}	3.7, 4.4	2.2, 2.9	4.5, 4.2
	std. dev.	3.3, 3.3	1.9, 2.1	5.8, 3.5
	(p, q)	0.3, 0.5	0.4, 0.4	0.3, 0.3
Our method	Z_{avg}	1.8, 1.3	2.1, 2.5	4.2, 4.3
	std. dev.	0.1, 0.1	0.2, 0.5	0.6, 0.6
	(p, q)	0.2, 0.2	0.2, 0.3	0.5, 0.5

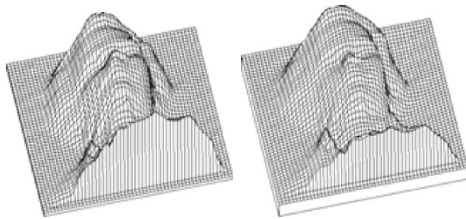


FIGURE 10: The shape of the Mozart reconstructed for $\mathbf{S} = (5, 5, 2)$ on the left has the maximal minimal depth and an error of 8.0%, while that for $\mathbf{S} = (50, 50, 24)$ on the right has the maximal rank and an error of 6.4%.

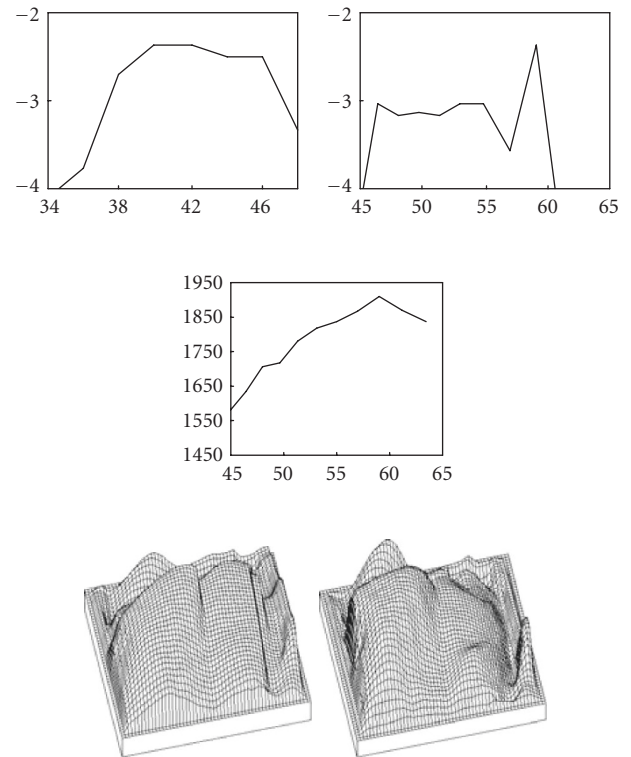


FIGURE 11: Results for the pepper image. Top: minimal depth profile as a function of tilt angle (left), and minimal depth (center) and rank (right) profiles as a function of slant angle. Bottom: shapes reconstructed for $\mathbf{S} = (0.766, 0.642, 1)$ (left) and $(0.766, 0.642, 0.6)$ (right).

methods and that it effectively worked for oblique light direction ranging from 55 to 75 degrees in slant angle without convex/concave ambiguities. A more sophisticated optimizing method is under study.

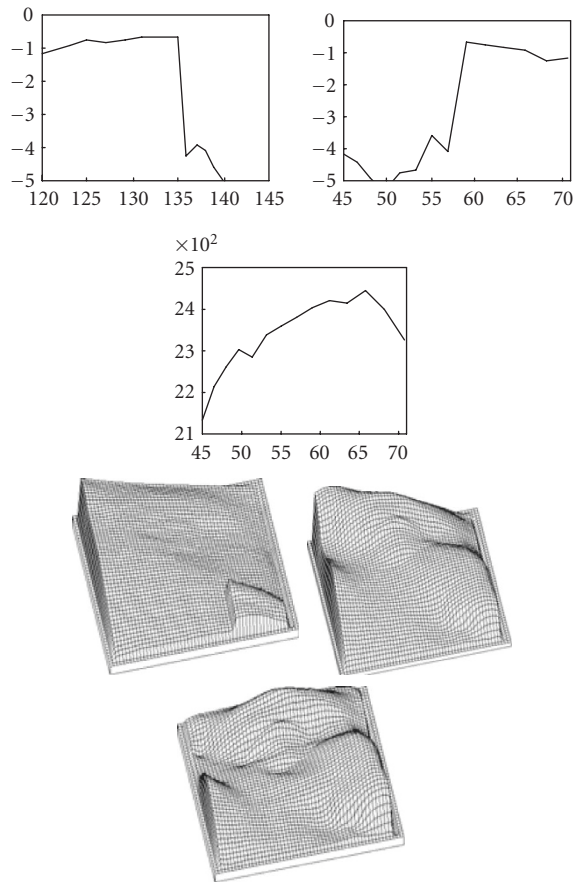


FIGURE 12: Results for the David image. Top: minimal depth profile as a function of tilt angle (left), and minimal depth (center) and rank (right) profiles as a function of slant angle. Bottom: shapes reconstructed for $S = (-0.707, 0.707, 1)$, $(-0.707, 0.707, 0.6)$, and $(-0.707, 0.707, 0.45)$, from left to right.

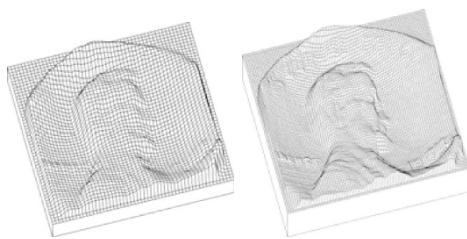


FIGURE 13: Comparison between shapes reconstructed from two different-sized penny images of 54×54 (left) and 96×96 (right) pixels.

ACKNOWLEDGMENT

We would like to thank the reviewers for kind advice and contributions.

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, NY, USA, 1975.

[2] Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo, and shape from shading," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 680–702, 1991.

[3] P. L. Worthington and E. R. Hancock, "New constraints on data-closeness and needle map consistency for shape-from-shading," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1250–1267, 1999.

[4] P.-S. Tsai and M. Shah, "Shape from shading using linear approximation," *Image and Vision Computing*, vol. 12, no. 8, pp. 487–498, 1994.

[5] M. Bichsel and A. Pentland, "A simple algorithm for shape from shading," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR '92)*, pp. 459–465, Champaign, Ill, USA, June 1992.

[6] R. Kimmel and A. M. Bruckstein, "Tracking level sets by level sets: a method for solving the shape from shading problem," *Computer Vision and Image Understanding*, vol. 62, no. 1, pp. 47–58, 1995.

[7] D. Samaras and D. Metaxas, "Incorporating illumination constraints in deformable models for shape from shading and light direction estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 247–264, 2003.

[8] R. Kimmel and J. A. Sethian, "Optimal algorithm for shape from shading and path planning," *Journal of Mathematical Imaging and Vision*, vol. 14, no. 3, pp. 237–244, 2001.

[9] A. Tankus, N. Sochen, and Y. Yeshurun, "Perspective shape-from-shading by fast marching," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 1, pp. I-43–I-49, Washington, DC, USA, June–July 2004.

[10] E. Prados, O. Faugeras, and E. Rouy, "Shape from shading and viscosity solutions," in *Proceedings of 7th European Conference on Computer Vision (ECCV '02)*, vol. 2, pp. 790–804, Copenhagen, Denmark, May 2002.

[11] J.-D. Durou, M. Falcone, and A. Sagona, "A survey of numerical methods for shape from shading," Report of IRT 2004-2-R, 2004.

[12] E. Prados and O. Faugeras, "Shape from shading: a well-posed problem?" in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '2005)*, vol. 2, pp. 870–877, San Diego, Calif, USA, June 2005.

[13] P. Tu and P. R. S. Mendonça, "Surface reconstruction via Helmholtz reciprocity with a single image pair," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '03)*, vol. 1, pp. 541–547, Madison, Wis, USA, June 2003.

[14] D. Metaxas and D. Terzopoulos, "Shape and nonrigid motion estimation through physics-based synthesis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 6, pp. 580–591, 1993.

[15] E. Prados, F. Camilli, and O. Faugeras, "A viscosity method for shape-from-shading without boundary data," INRIA Research Report 5296, 2004.

[16] R. Kimmel and A. M. Bruckstein, "Global shape from shading," *Computer Vision and Image Understanding*, vol. 62, no. 3, pp. 360–369, 1995.

[17] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, "Shape from shading: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 690–706, 1999.

Osamu Ikeda received his Master and Doctor of Engineering degrees in control engineering from Tokyo Institute of Technology in 1972 and 1976, respectively. Since then, he had been a Research Associate at the institution, working in the fields of computed imaging, optical information processing, and signal and multidimensional signal processing. Since 1987, he has been in the Faculty of Engineering at Takushoku University in Tokyo. His current research interests include computer vision, multimedia, pattern recognition, image processing, and image retrieval. He has published 100 peer-reviewed journal and conference papers. He is a Member of the IEEE, ACM, and IEICE of Japan.

