A Shape-From-Shading Algorithm Using Photometric Stereo

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

Abstract

We present a simple shape-from-shading algorithm that uses two shading images. This enables us to estimate more accurate shapes of objects without any special constraints or conditions than the existing photometric stereo algorithms. The Jacobi's iterative method is applied to the difference between the image and the reflectance function represented as a function of three depth parameters for each of the two images, and the resulting two iterative relations are combined to get a single iterative relation. The algorithm is analytically shown to be numerically stable. Two supplementary methods are also presented to enhance its applicability. Computer experiments clearly show its usefulness and robustness to noise. Typically it takes seconds using a 1.2GHz Pentium III PC to get shapes from two 50x50 shading images.

1 Introduction

Since Horn initiated the research on shape-from-shading more than two decades ago [1], significant developments have been made [2], [3]. A majority of papers focus on estimating shapes from single shading images. Approaches presented so far may be classified in a broader view to local [4], [5], [6] and global [1], [7], [8], [9] ones, and global approaches may be classified to the categories of minimization [10], [11], [12], linear [9], [13] and propagation [1], [14], [15], [16]. A few papers, on the other hand, focus on estimating shapes from multiple images [17], [18].

As for the global approaches, minimization ones are based upon minimizing a given energy criterion to estimate the shape. For example, Zheng and Chellappa [11] introduced image gradient and integrability constraints to obtain fine details. Worthington and Hancock [12] introduced curvature consistency and image gradient ones as constraints. Linear approaches linearlize the reflectance map in tilts or depth. Pentland applies Fourier transform after linearlizing the function in tilts, but its applicability is limited due to the approximations. Tsai and Shah [9] approximate the relation linearlized in the depth and use the Newton-Raphson method to obtain an iterative relation. Their algorithm is fast in time but lacks reconstruction quality of shape. We also presented a nonlinear approach by applying the Jacobi iterative technique to the image irradiance equation represented by the three depth parameters, where we carry out bi-directional estimation to minimize shape distortions [19]. Propagation approaches are to obtain a shape starting from some initial curve, which uses such special points as the brightest or the darkest [1]. For example, Kimmel et al. show that good shape reconstruction is possible with boundary conditions [16], which, however, may be too demanding and may make the processing complicated as the object is more complex.

As for the estimation using multiple images, Woodham proposes to use three shading images to uniquely decide the gradient map of the surface [17]. Using two images, on the other hand, may lead to a partially incorrect map [18]. Rocchini et al. take six shading images and choose good ones free from effects of specular reflection or shadow [20]. The conversion from the gradient map to the depth or height one, however, is not straightforward. Frankot et al., for example, present a method using Fourier transform [21]. But the method tends to give rise to noticeable undulant shapes especially to flat or near flat parts.

In this paper we present a simple shape-from-shading algorithm by developing our previous method to the case of using two images, aiming at estimating accurate and reliable shapes for a variety of objects and illuminating conditions, without any special conditions or constraints. We also present supplemental methods to enhance the applicability of the algorithm. The principle, an analysis on numerical stability, and comprehensive computer experiments are given.

2 Principle

The object is illuminated sequentially from each of two different directions to obtain two shading images. Given a reflectance function $R(p,q)$, there holds the relation between the function and the image $I(x,y)$ for each illumination:

$$R_i(p,q) = I_i(x,y)\ ,\ i = 1, 2 \tag{1}$$

where $x, y = 1, \ldots, N$, and the images are assumed to be normalized to unity. Let **P** and **S** be the surface normal of the object and illuminant vector, respectively:

$$\mathbf{P} = \frac{(p,q,1)}{\sqrt{p^2+q^2+1}} \tag{2}$$

$$\mathbf{S}_i = \frac{(s_i,t_i,1)}{\sqrt{s_i^2+t_i^2+1}}\ ,\ i = 1, 2 \tag{3}$$

where p and q are given by $-\partial z/\partial x$ and $-\partial z/\partial y$, respectively. Then, for the *Lambertian* surface, the reflectance function, normalized by *albedo*, is given by their scalar product:

$$R_i(p,q) = \frac{1+ps_i+qt_i}{\sqrt{1+p^2+q^2}\sqrt{1+s_i^2+t_i^2}}\ ,\ i = 1, 2 \tag{4}$$

Letting $z(x,y)$ be the depth or shape, the discrete forms of p and q may be given by

$$\begin{aligned} p &= z(x-1,y) - z(x,y) \\ q &= z(x,y-1) - z(x,y) \end{aligned} \tag{5}$$

This simplest approximation appears to suffice in our case, but using more sophisticated approximations [22] could possibly improve the accuracy in shape when the object shape is very complex.

Then, R_i can be regarded to be a function of three variables $z(x,y)$, $z(x\text{-}1,y)$ and $z(x,y\text{-}1)$. Applying the Jacobi's iterative method to the function $f_i(x,y)$ defined by

$$f_i(x,y) \equiv I(x,y) - R_i(p,q)\ ,\ i = 1, 2 \tag{6}$$

we obtain the following iterative relations

$$-f_i(x,y)^{(n-1)} = \left(\frac{\partial f_i(x,y)}{\partial z(x,y)}\right)^{(n-1)} \left(z(x,y)^{(n)} - z(x,y)^{(n-1)}\right)$$
$$+\left(\frac{\partial f_i(x,y)}{\partial z(x-1,y)}\right)^{(n-1)} \left(z(x-1,y)^{(n)} - z(x-1,y)^{(n-1)}\right) \qquad i=1,2 \quad (7)$$
$$+\left(\frac{\partial f_i(x,y)}{\partial z(x,y-1)}\right)^{(n-1)} \left(z(x,y-1)^{(n)} - z(x,y-1)^{(n-1)}\right)$$

where n is the number of iterations. These can be rewritten in matrix form as

$$-\mathbf{f}_i^{(n-1)} = \mathbf{g}_i^{(n-1)}(\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}),\ n = 1, 2, \ldots,\ i = 1, 2 \qquad (8)$$

where $\mathbf{f}_i$, i = 1, 2, are vectors of N^2 elements of $f_i(x,y)$, $\mathbf{z}$ is a vector of N^2 elements of $z(x, y)$, and $\mathbf{g}_i$, i = 1, 2, are matrices of N^2xN^2 elements that are made of $\partial f_i(x,y)/\partial z(x,y)$, $\partial f_i(x,y)/\partial z(x\text{-}1,y)$ and $\partial f_i(x,y)/\partial z(x,y\text{-}1)$. The two matrix relations can be combined to give the following single relation:

$$-\mathbf{F}^{(n-1)} = \mathbf{G}^{(n-1)}\left(\mathbf{z}^{(n)} - \mathbf{z}^{(n-1)}\right) \qquad (9)$$

$$\mathbf{F} = \begin{bmatrix}\mathbf{f}_1 & \mathbf{f}_2\end{bmatrix}^T \qquad (10)$$

$$\mathbf{G} = \begin{bmatrix}\mathbf{g}_1 & \mathbf{g}_2\end{bmatrix}^T \qquad (11)$$

Eq. (9) can be solved for $\mathbf{z}^{(n)}$ as

$$\mathbf{z}^{(n)} = \mathbf{z}^{(n-1)} - \left(\mathbf{G}^{(n-1)^T}\mathbf{G}^{(n-1)}\right)^{-1}\left(\mathbf{G}^{(n-1)^T}\mathbf{F}^{(n-1)}\right), \qquad (12)$$
$$n = 1,2,\ldots$$

Thus, the shape is estimatable iteratively using the relation in Eq. (12), typically beginning with null values $\mathbf{z}^{(0)} = \mathbf{0}$.

3 Numerical Stability

In order for Eq. (12) to be carried out, the determinant of $\mathbf{G}^T\mathbf{G}$ must be significant for any number of iterations. Since the determinant is given by the product of all the diagonal elements, the condition can be reduced to that all the diagonal elements, that is, all the eigenvalues must be significant. We obtain those eigenvalues from Eqs. (7)-(12) as

$$\lambda(x,y) = \begin{cases} \sum_{i=1}^{2}\left[\left(\frac{\partial f_i(x,y)}{\partial z(x,y)}\right)^2 + \left(\frac{\partial f_i(x+1,y)}{\partial z(x,y)}\right)^2 + \left(\frac{\partial f_i(x,y+1)}{\partial z(x,y)}\right)^2\right] & \text{for } 1 \le x \le N-1, 1 \le y \le N-1 \\ \sum_{i=1}^{2}\left[\left(\frac{\partial f_i(x,y)}{\partial z(x,y)}\right)^2 + \left(\frac{\partial f_i(x,y+1)}{\partial z(x,y)}\right)^2\right] & \text{for } x = N, 1 \le y \le N-1 \\ \sum_{i=1}^{2}\left[\left(\frac{\partial f_i(x,y)}{\partial z(x,y)}\right)^2 + \left(\frac{\partial f_i(x+1,y)}{\partial z(x,y)}\right)^2\right] & \text{for } 1 \le x \le N-1,\ y = N \\ \sum_{i=1}^{2}\left(\frac{\partial f_i(x,y)}{\partial z(x,y)}\right)^2 & \text{for } x = N,\ y = N \end{cases} \qquad (13)$$

It is seen that all of them are non-negative, and that $(N\text{-}1)^2$ in N^2 eigenvalues consist of six terms, $2(N\text{-}1)$ in N^2 eigenvalues consist of four terms, and one in N^2 eigenvalues consists of two terms.

In the case where only a single shading image is available, on the other hand, Eq. (8) can be used for the shape estimation rather than Eq. (9). In this case the determinant of $\mathbf{g}$ must be significant throughout the iteration. This is given by the product of all the diagonal terms, which are given from Eqs. (7) and (8) as

$$\lambda(x,y) = \frac{\partial f(x,y)}{\partial z(x,y)} \qquad (14)$$

It is seen that they are not non-negative. Actually they can be null for the image parts where the surface normal $\mathbf{P}$ of the shape being reconstructed is parallel to the illuminant vector $\mathbf{S}$.

In the system being presented, the terms in Eq. (13) may not in general be null simultaneously because the two illuminating directions are different and because three kinds of derivatives are different of each other in their characteristics. Thus we are able to execute the iteration in Eq. (12) for most objects and illuminating conditions to obtain converged solutions for the shape.

4 Two Methods to Enhance Applicability

4.1 Rotation of Images

It is noted that in case the eigenvalues for the two images are null simultaneously at $x = N$ and $y = N$ for $\mathbf{z}^{(0)} = \mathbf{0}$, rotating the coordinates, or the images together with their illuminant vectors instead, may enable us to carry out the estimation. For example, let $\mathbf{S}_1$ and $\mathbf{S}_2$ be equal to (-1,1,1) and (1,-1,1), or (0,0,1) and (-1,1,1), respectively, hereafter the normalizing coefficients are omitted. These make the eigenvalue null at $(x,y) = (N,N)$ for the initial value of $\mathbf{z}^{(0)} = \mathbf{0}$, making impossible to carry out the iteration. If we rotate the coordinates by 90 degrees, then $\mathbf{S}_1$ and $\mathbf{S}_2$ are (1,1,1) and (-1,-1,1), or (0,0,1) and (1,1,1), respectively, enabling to carry out the iteration using the initial value.

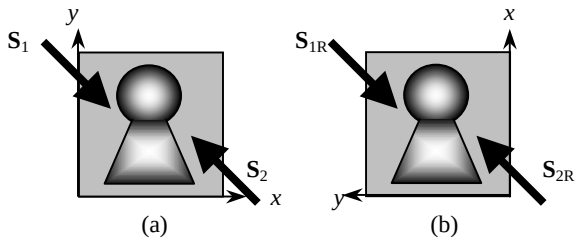
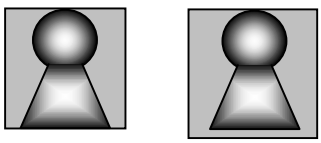


Fig. 1 The illuminant vectors $\mathbf{S}_1$=(-1,1,1) and $\mathbf{S}_2$=(1,-1,1) can equivalently be changed to $\mathbf{S}_{1R}$=(1,1,1) and $\mathbf{S}_{2R}$=(-1,-1,1) by rotating the coordinates by 90 degrees.

4.2 Boundary Conditions

We can make an assumption in Eq. (4) that $p = 0$ along the line of $x = 1$ and $1 \le y \le N$, and $q = 0$ along the line of $y = 1$ and $1 \le x \le N$. It is needless to make a similar assumption along the other two boundaries, when we use Eq. (5). We can also make another assumption that $p(1,y) = p(2,y)$, $1 \le y \le N$ and $q(1,y) = q(2,y)$, $1 \le x \le N$. If these assumptions do not fully agree with the shape, significant shape distortions may appear, as will be shown.

We adopt two methods to avert such distortions. One is to set the coordinates axes along the flat intensity boundaries, if they are available, so that the first assumption for Eq. (4) holds. If not, such flat intensity boundaries are added to the image, as shown in Fig. 2, so that the first assumption for Eq. (4) holds. The addition works for most cases of the illuminant vectors, as will be shown.



(a) (b)

Fig. 2 If the image does not have two neighboring flat intensity boundaries as in (a), they are added to the image as in (b).

5 Computer Experiments

Two shapes that are mostly used are shown in Fig. 3. One is a semi-sphere and the other is the measured data of a Mozart sculpture [23]. Assuming the surfaces to be *Lambertian*, shading images were computationally generated which have the size of 50x50 in pixel.

First, for comparison, shapes were reconstructed from single shading images for a variety of illuminant vectors using the basic algorithm [19]:

(1) The coordinates are rotated so that the illuminant vector has the tilt angle of 45 degrees. If this accompanies the increase in image area, the image intensity value corresponding to the flat surface is filled in the vacant areas.
(2) The shape is estimated iteratively using Eq. (8), where the iteration is stopped when the minimal eigenvalue decreases to a value less than, for example, 0.1. This estimation is carried out bi-directionally.
(3) The two shapes are averaged with appropriate weights that is determined by evaluating average magnitudes of (p,q) of the shapes.
(4) The shape within the region of the original images is extracted.

Shading images for eight illuminant vectors, which are different in tilt angle, and the obtained shapes are shown in Figs. 4 and 5, respectively. Also, shading images for five illuminant vectors, which are different in slant angle, and obtained shapes are shown in Figs. 6 and 7, respectively. It is seen from the results in Fig. 5 that shapes for tilt angles of 0, 45, 180, 225 degrees are relatively good but those for tilt angles of 90, 135, 270, 315 degrees are not. It is seen from the results in Fig. 7 that shapes for slant angles of 45 and 60 degrees are relatively good but those for other angles are not. So the shape obtainable from a single image is limited in quality.

In the case of using two images, shapes estimated from typical pairs of shading images are shown in Fig. 8, first, for the semi-sphere. In this case all the images have flat intensity boundaries, so the shape can be estimated in either forward or backward direction. The results show that the reconstructed shapes are very close to the original one regardless of the differences in tilt or slant angle of the illuminant vectors. An error distribution in depth and a profile of the convergence are shown in Fig. 9 for the pair of **S**=(5,5,7) and (-5,5,7). The distribution is very similar to those of the other pairs in Fig. 8, and the profile is also very similar to them except for the pair of (1,0,1) and (-1,0,1), which once is on the verge of divergence before convergence. The reconstructed shapes have the maximal depth differences of, from top to bottom in Fig. 8, 87, 87, 82, 83, 92, 94% of the original one. This may come from that the surface gradients obtained are smaller than the actual ones along the circle where the hemi-sphere and the flat part meet, as the distribution in Fig. 9 indicates.

Shapes estimated for typical pairs of the illuminant vectors for the Mozart are shown in Fig. 10, where backward estimation was carried out because of the boundary condition. All the shapes in Fig. 10 appear very similar to the original one except for minor distortions, such as non-flat part for the (0,0,1) and (5,5,7) pair. The convergence is fast similarly to the semi-sphere case, as shown in Fig. 11. Figure 12 shows the robustness of the algorithm to such illuminant vectors that make reconstructing a shape from a single image difficult, although distortions tend to increase. It is not possible to estimate the shape for the images with the illuminant vectors, (0,0,1) and (5,-5,7), but rotating the images together with the vectors by 90 degrees makes it possible as shown in Fig. 13. Figure 14 shows that when the estimation is carried out in the forward direction in the absence of the flat intensity boundaries, the obtainable shape changes partially or wholly depending on the boundary condition assumed, but that if we add flat intensity boundaries to the images we can obtain a good shape. It is noted that the top-right shape in Fig. 14 is not a converged one but deteriorates further with the iteration.

Our method was applied also to the images of a computer mouse to reconstruct its shape. The shape data was obtained with a laser range scanner, and two 80x80 pixel shading images computer-generated for **S** = (-5,-5,7) and (5,-5,7) are shown in Fig. 15. Significant structural noise is observable over the mouse and letters on the mouse appear to be local noise or depth variations. The estimation was carried out in the forward direction. In Fig. 15 the estimated shape is represented both by wire-frame and with the superposition of the texture in (a), and it is viewed from the front and back. It is seen that the shape is finely reconstructed and that it is clearly robust to noise.

6 Conclusions

We presented a shape-from-shading method using two images that can reconstruct accurate shapes for most cases. The algorithm is very simple, and its execution is not in real time but does not take time either. We are studying the algorithm that enables us to get more accurate shapes for any illuminating conditions.

References

[1] B.K.P. Horn and M.J. Brooks, *Shape from Shading* (Cambridge, MA: MIT Press, 1989).

[2] B.K.P. Horn, Obtaining Shape from Shading Information, *The psychology of computer vision*, P. H. Winston (Ed.) (New York: McGraw Hill, 1975), 115-155.

[3] R. Zhang, P. Tsai, J. E. Cryer, and M. Shah, Shape from Shading: A Survey, *IEEE Trans. PAMI*, 21(8), 1999, 690-705.

[4] P. Pentland, Local shading analysis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6(2), 1984, 170-187.

[5] C.H. Lee and A. Rosenfeld, Improved Methods of estimating Shape from Shading Using the Light Source Coordinate System, *Artificial Intelligence*, 26, 1985, 125-143.

[6] F. P. Ferrie and J. Lagarde, Curvature Consistency Improves Local Shading Analysis, *Proc. IEEE Int'l Conf. Pattern Recognition, I*, 1990, 70-76.

[7] K. Ikeuchi and B. P. Horn, Numerical Shape from Shading and Occluding Boundaries, *Artificial Intelligence*, 17(3), 1981, 141-184.

[8] B.K.P. Horn and M.J. Brooks, The Variational approach to Shape from Shading, *Computer Vision, Graphics, and Image Processing*, 33(2), 1986, 174-208.

[9] P. S. Tsai and M. Shah, Shape from Shading Using Linear Approximation, *J. Imaging and Vision Computing*, 12(8), 1994, 487-498.

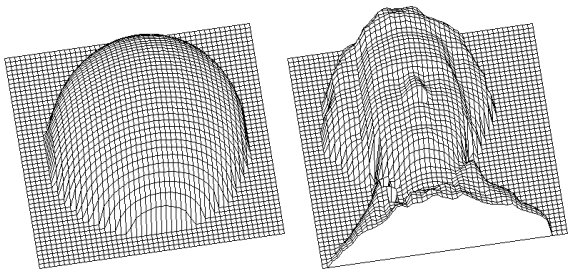
[10] M. J. Brooks and B. K. P. Horn, Shape and Source from Shading, *Proc. Int'l Joint Conf. Artificial Intelligence*, 1985, 932-936.

[11] Q. Zheng and R. Chellappa, Estimation of Illuminant Direction, Albedo, and Shape from Shading, *IEEE Trans. PAMI, 13*(7), 1991, 680-702.

[12] P. L. Worthington and E. R. Hancock, New Constraints on Data-Closeness and Needle Map Consistency for Shape-from-Shading, *IEEE Trans. PAMI*, 21(12), 1999, 1250-1267.

[13] A. Pentland, Shape Information from Shading: A Theory about Human Perception, *Proc. Int'l Conf. Computer Vision*, 1988, 404-413.

[14] J. Oliensis, Shape from Shading as a Partially Well-Constrained Problem, *Computer Vision, Graphics, and Image Processing*, 54, 1991, 163-183.
[15] M. Bichsel and A. Pentland, A Simple Algorithm for Shape from Shading, *IEEE Proc. Computer Vision and Pattern Recognition*, 1992, 459-465.
[16] R. Kimmel and A.M. Bruckstein, Tracking Level Sets by Level Sets: A Method for Solving Shape from Shading Problem, *Computer Vision and Image Understanding*, 62(1), 1995, 47-58.
[17] R.J. Woodham, Photometric Method for Determining Surface Orientation from Multiple Images, Chap. 17 in *Shape from Shading*(Cambridge, MA: MIT Press, 1989).
[18] R. Klette, A. Koschan and K.Schluns, *Three-Dimensional Data from Images* (Singapore, Springer, 1998), Chap. 8.
[19] O. Ikeda, A Novel Shape-From-Shading Algorithm Using Jacobi Iterative Method and Bi-Directional Estimation, *Proc. 5th IASTED International Conference on Computer Graphics and Imaging*, 2002, 56-61.
[20] C. Rocchini, P. Cignoni, C. Montani and R. Scopigno, Acquiring, stitching and blending appearance attributes on 3D Models, *The Visual Computer*, 18, 186-204, 2002.
[21] R.T. Frankot and R. Chellappa, A Method for Estimating Integrability in Shape from Shading Algorithms, *IEEE Trans. PAMI*, 10(4), 1988, 439-451.
[22] S.Collins, R.Kozera and L.Noakes, A Piecewise Quadratic Approach to Single Image Shape from Shading, *Proc. IEEE ICIP*, 2002, 126-129.
[23] eustis.cs.ucf.edu (132.170.108.42)



(a) Shape of a semi-sphere (b) Shape of a Mozart sculpture

Fig. 3 Two shapes used in the computer experiments.

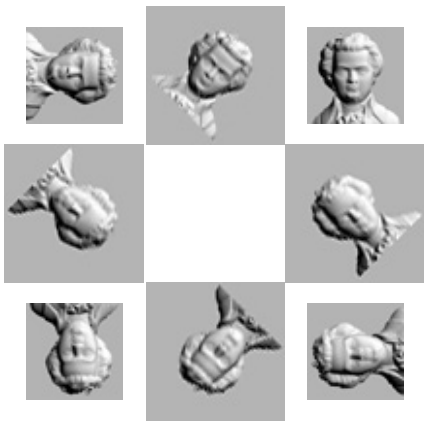


Fig. 4 Eight shading images with original illuminant vectors (1,0,1), (5,5,7), (0,1,1), (-5,5,7), (-1,0,1), (-5,-5,7), (0,-1,1), (5,-5,7) unclockwise from the middle right. They were rotated so that the tilt angles are equal to 45 degrees.

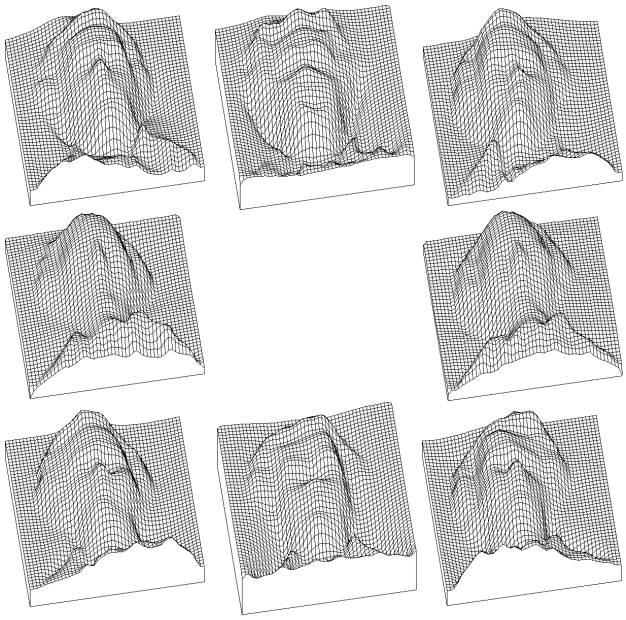


Fig. 5 Estimated shapes from the single shading images in Fig. 4.

Fig. 6 Five shading images with illuminant vectors (1,1,6), (2,2,5), (5,5,7), (6,6,5) and (5,5,2), from let to right. They correspond to the slant angles of 13, 30, 45, 60, and 74 degrees, respectively.

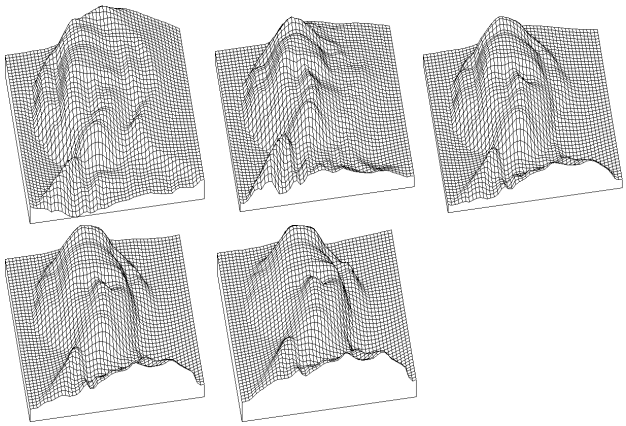


Fig.7 Reconstructed shapes for the images in Fig. 6, where the top left is for **S** = (1,1,6) and the bottom center is for (5,5,2).

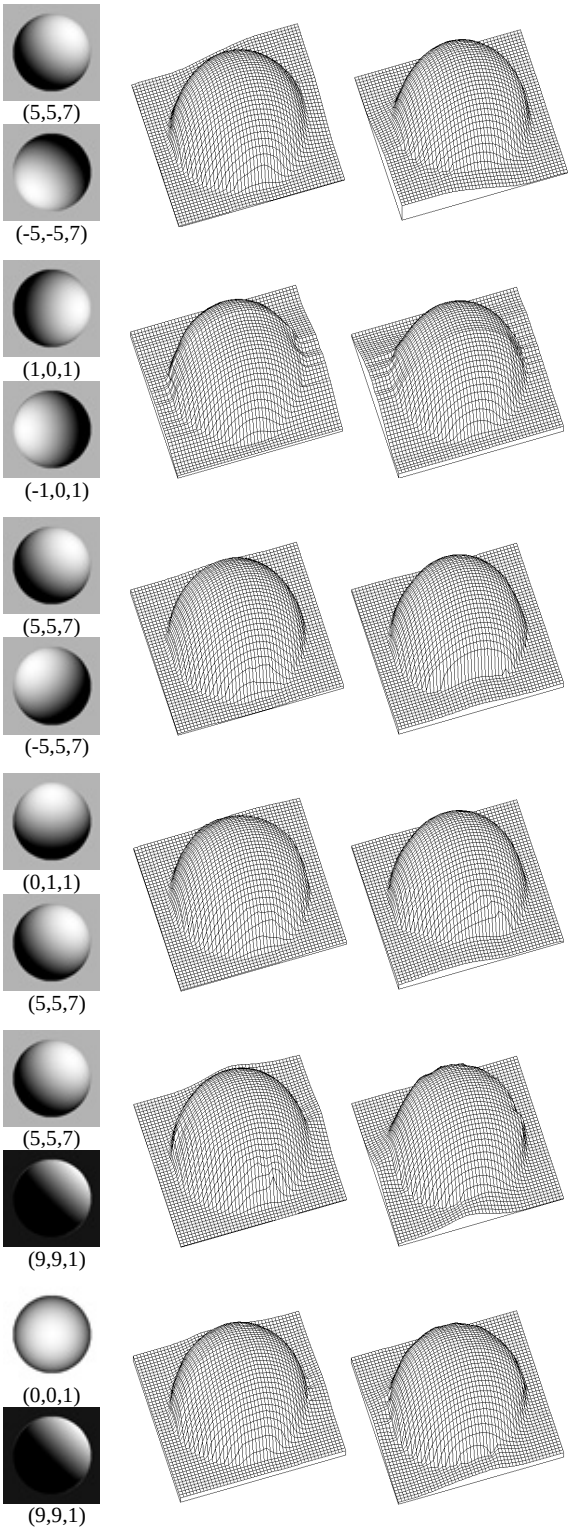


Fig. 8 Reconstructed shapes for six pairs of two shading images. The differences in slant and tilt angles of the two illuminant vectors are, from top to bottom, (0,180), (0,90), (0,90), (0,45), (40,0), (85,0) degrees. On the center are frontviews of the shapes and on the right are their backviews.

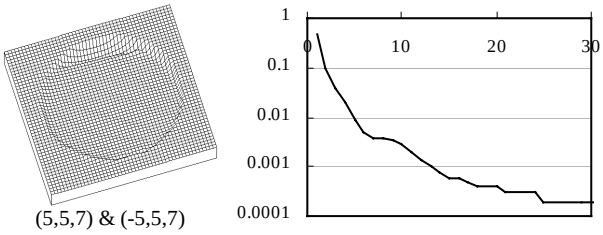


Fig. 9 An error distribution in depth value and a profile of the difference between the neighboring shapes which are normalized to [0,1]. One iteration takes about 2 seconds.

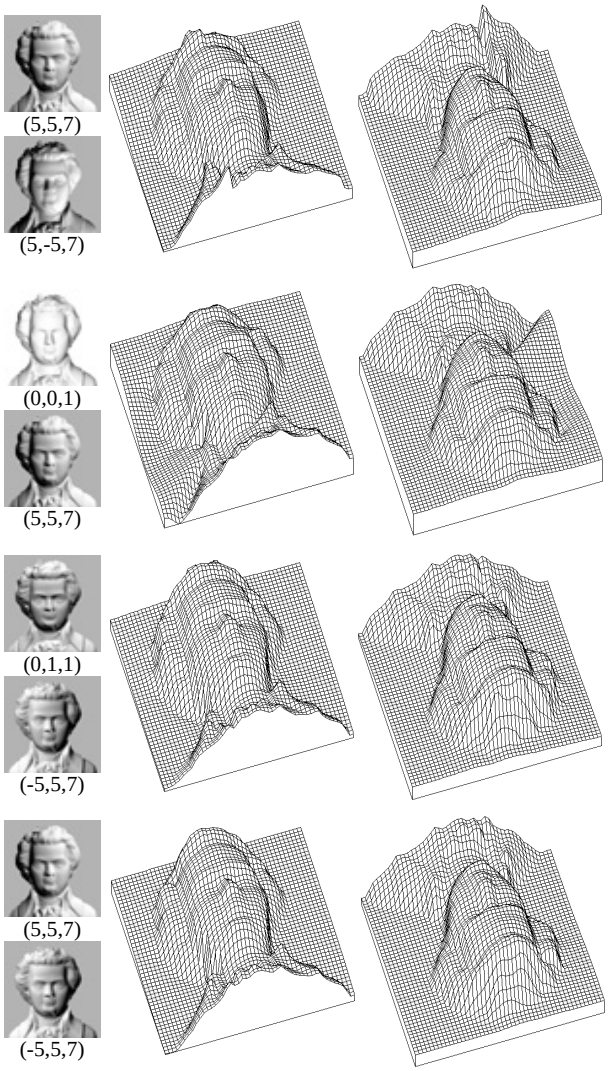


Fig. 10 Front and back views of reconstructed shapes for four typical pairs of two shading images, which are on the left with their illuminant vectors.

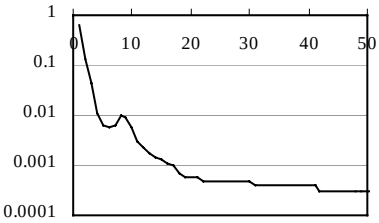


Fig. 11 Profile in the difference of the neighboring shapes in the iteration for the Mozart images with the vectors of (5,5,7) and (5,-5,7), where the shapes are normalized to [0,1].

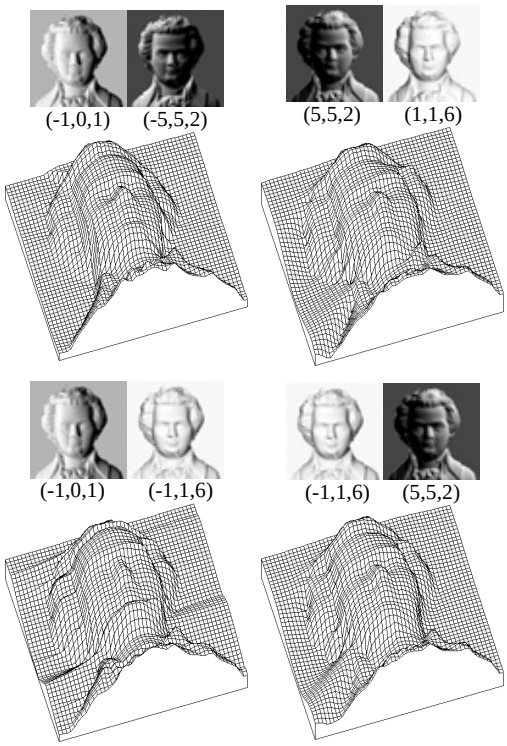


Fig. 12 Reconstructed shapes for such illuminant vectors that make it hard to obtain good shapes from single shading images.

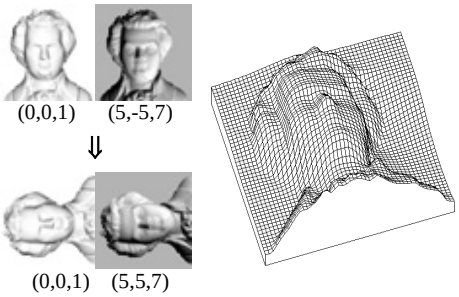


Fig. 13 Rotating images and their illuminant vectors enables shape reconstruction, where the estimation is carried out in the forward direction after the rotation.

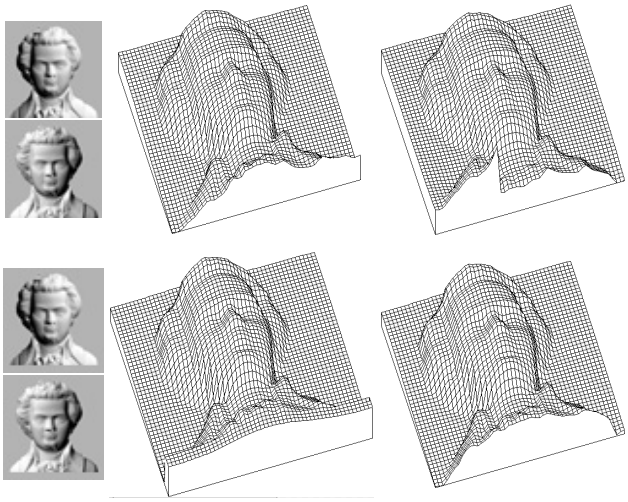


Fig. 14 Effects of adding the flat intensity boundaries to the original images, where the estimation is carried out in the forward direction. Top-left, shape obtained using the assumption of $p(x=1,y)=0$ and $q(x,y=1)=0$ for the original two shading images with $\mathbf{S}$ = (5,5,7) and (-5,5,7), top-right, shape obtained using the other assumption of $p(x=1,y) = p(x=2,y)$ and $q(x,y=1) = q(x,y=2)$ for the same images, bottom-left, shape obtained by adding the flat intensity boundaries combined with the first assumption, and bottom-right, shape within the original shading images.

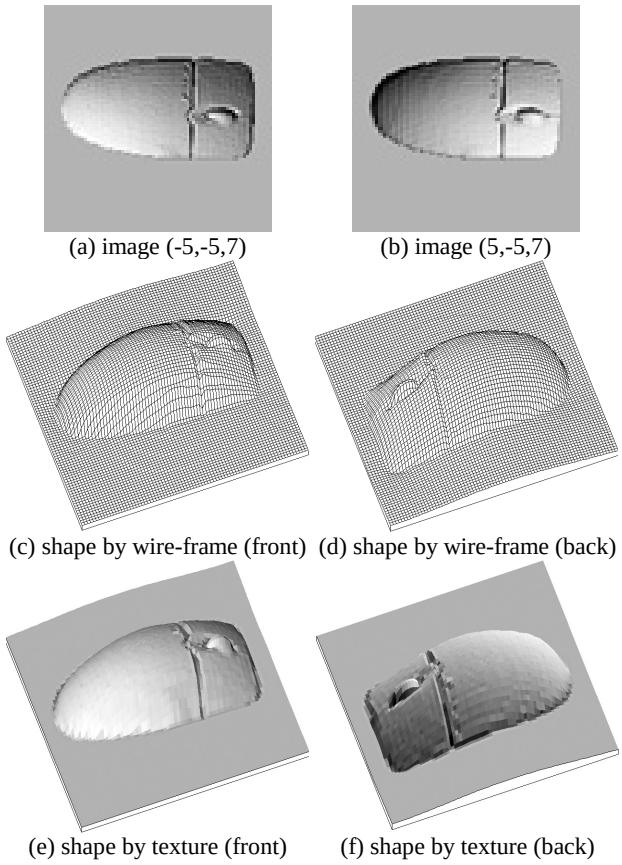


Fig. 15 Shape reconstruction from two synthetic images of a computer mouse. (a), (b), shading images, (c), (d), estimated shape represented by wire-frame, (e) and (f), that by the texture in (a), and front and back views are given for the estimate.