Journal of Electronic Imaging

JElectronicImaging.org

Shape estimation of concave specular object from multiview polarization

Daisuke Miyazaki
Ryosuke Furuhashi
Shinsaku Hiura

SPIE. IS&T imaging.org

Daisuke Miyazaki, Ryosuke Furuhashi, Shinsaku Hiura, "Shape estimation of concave specular object from multiview polarization," *J. Electron. Imaging* **29**(4), 041006 (2020), doi: 10.1117/1.JEI.29.4.041006

Shape estimation of concave specular object from multiview polarization

Daisuke Miyazaki,* Ryosuke Furuhashi,† and Shinsaku Hiura‡
Hiroshima City University, Hiroshima, Japan

Abstract. We propose a method to estimate the surface normal of concave objects. The target object of our method has a specular surface without diffuse reflection. We solve the problem by analyzing the polarization state of the reflected light. The polarization analysis gives a constraint to the surface normal. However, polarization data from a single view has an ambiguity and cannot uniquely determine the surface normal. To solve this problem, the target object should be observed from two or more views. However, the polarization of the light should be analyzed at the same surface point through the different views. This means that both the camera parameters and the surface shape should be known. The camera parameters can be estimated *a priori* using known corresponding points. However, it is a contradiction that the shape should be known in order to estimate the shape. To resolve this problem, we assume that the target object is almost planar. Under this assumption, the surface normal of the object is uniquely determined. We show that the surface normal of the nonplanar part can also be estimated using the proposed method.
 [DOI: 10.1117/1.JEI.29.4.041006]

Keywords: polarization; shape-from-X; surface normal; concavity; specular reflection; crack.

Paper 190891SS received Sep. 27, 2019; accepted for publication Dec. 26, 2019; published online Jan. 17, 2020.

1 Introduction

Factories in industrial fields have a high demand for estimating the shape of cracks since it is quite important for quality control of the products. Although there are many methods for detecting cracks,[1] few methods have been proposed for estimating the shape of cracks. Therefore, there is a great demand for estimating the shape of concave objects of highly specular surfaces since it is a challenging task. This paper proposes a method that estimates the surface normal of a black specular object with a concave shape by analyzing the polarization state of the reflected light in which the target object is observed from multiple views.

Three-dimensional (3-D) modeling techniques have been intensively investigated in the field of computer vision. These techniques can be categorized into two types: the geometric approach and the photometric approach. The geometric approach uses the geometrical structure of the scene, such as time-of-flight (ToF) laser range sensor, multinocular stereo, or structured light projection. The photometric approach uses the light reflected from the scene, such as photometric stereo or shape from polarization. Shape from specularity has been extensively surveyed by Ihrke et al.[2]

A smooth surface normal can be obtained using a photometric approach. Polarization[3–5] is one of the photometric clues that can be used to obtain a smooth surface normal. Koshikawa and Shirai[6] used circular polarization to estimate the surface normal of a specular object. Guarnera et al.[7] extended their method to determine the surface normal uniquely, by changing the lighting conditions in two configurations. Morel et al.[8] also disambiguated it using multiple illumination; however, they did not solve the ambiguity of the degree of polarization (DOP)[3–5] because they did not use circular polarization. Saito et al.[9] proposed the basic theory for estimating the surface normal of a transparent object using polarization. Barbour[10] approximated the relation between

*Address all correspondence to Daisuke Miyazaki, E-mail: miyazaki@hiroshima-cu.ac.jp
†Present Address: DISEC Corporation, Japan
‡Present Address: University of Hyogo, Japan

the surface normal and the DOP and developed a commercial sensor for shape from polarization. Kobayashi et al.[11] estimated the surface normal of transparent thin objects using DOP. They also estimated the thickness by analyzing the light interference. Miyazaki et al.[12] estimated the surface normal of a transparent object by analyzing the polarization state of the thermal radiation from the object. Miyazaki et al.[13] attempted to estimate the surface normal of a diffuse object from a single view. Miyazaki et al.[14] used a geometrical invariant to match the corresponding points from two views to estimate the surface normal of a transparent object. Miyazaki and Ikeuchi[15] solved the inverse problem of polarization ray tracing to estimate the surface normal of a transparent object. These methods first calculate the polarization data from input images, while Yu et al.[16] used the input images themselves to estimate the surface normal without explicitly calculating the DOP.

Wolff and Boult[17] developed the basic theory for showing that polarization analysis can estimate a surface normal from two views if the corresponding points are known. Rahmann and Canterakis[18] estimated the surface normal of a specular object from multiple views by iteratively finding the corresponding points of these views. Rahmann[19] proved that only the quadratic surfaces are estimated if the corresponding points are searched iteratively. Atkinson and Hancock[20] analyzed the local structure of an object to find the corresponding points between two view points to calculate the surface normal from the polarization of two views. Atkinson and Hancock[21] also provided a detailed investigation of surface normal estimation for a diffuse object from a single view. Huynh et al.[22] estimated not only the surface normal but also the refractive index.

Kadambi et al.[23] combined the 3-D geometry obtained by a ToF sensor and the surface normal obtained from the DOP. Smith et al.[24] combined the depth sensor and the shape from polarization.

Cui et al.[25] used structure from motion, while Yang et al.[26] used SLAM in addition to the shape from polarization. Miyazaki et al.[27] combined the visual hull and the shape from polarization.

In this study, we propose a method for creating a 3-D model using both polarization analysis and planarity assumption. The principal target objects are smooth surfaces with high specular reflection and low diffuse reflection, which are annoying targets in conventional techniques. We first calibrate multiple cameras to calculate the geometrical relationships among them. We observe the object from multiple viewpoints using a polarization imaging camera. To determine the corresponding point among multiple views, we assume the target object as planar. However, this assumption simply produces only a planar shape; thus we also use polarization information to estimate the nonplanar part of the object. The shape-from-polarization method can estimate the shape of black objects with high specularity, which cannot be estimated using the photometric stereo method because there are no diffuse reflections. The polarization information of the object is obtained from multiple viewpoints using a polarization imaging camera. The polarization data must be analyzed at identical points on the object surface when observed from multiple viewpoints. Thus, the planarity assumption can be used for estimating the surface normal from polarization data. The target object of our method is almost planar except for a crack with a small size.

Miyazaki et al.'s method[27] relies on the visual hull. It is difficult to estimate a planar shape using a visual hull; in addition, it is impossible to estimate a planar shape with infinite size [Fig. 1(a)]. Our method can also be applied to infinite plane [Fig. 1(b)]. Thus, our method overcomes the disadvantage of their method,[27] which means that the proposed method is fundamentally superior to their method[27] if the target object is almost planar.

We describe our method in Sec. 2 and present our results in Sec. 3. The theory shown in Sec. 2 assumes that the target object must be completely planar. However, Sec. 3 empirically proves that our method can successfully estimate the surface normal even if the object is not completely planar. We discuss the advantages and disadvantages of our method and conclude the paper in Sec. 4.

2 Using Polarization in Estimating the Surface Normal of Concave Objects

2.1 *Algorithm Flow*

First, we explain the flow of our method (Fig. 2).

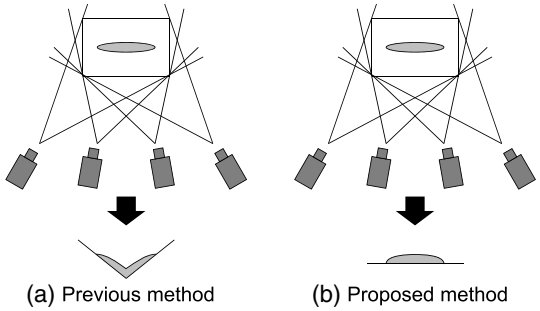


Fig. 1 Our contribution: (a) previous method, which is based on visual hull that is not suited to estimating planar shapes and (b) the proposed method, which is suited to estimating the concave shapes that is almost planar.

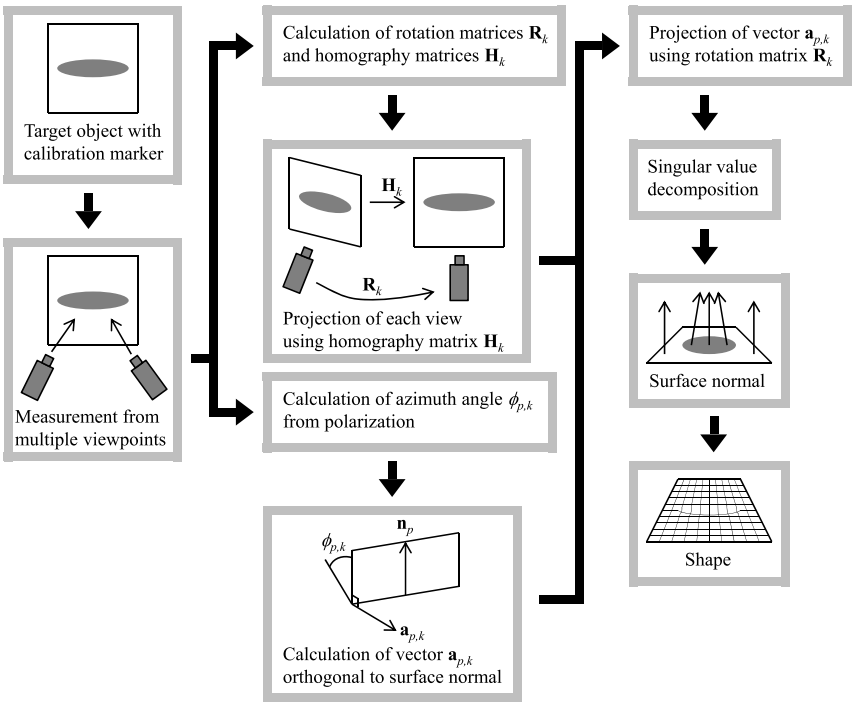


Fig. 2 Algorithm flow.

Since we observe the target object from multiple viewpoints, we calibrate each viewpoint to obtain each camera parameter. Although any calibration pattern works well, this paper assumes that each camera is calibrated using four points marked at the vertices of a square for clarity. The area that is surrounded by these markers is the target area. Using these markers, we estimate the homography **H** (Sec. 2.6) and rotation **R** (Sec. 2.5). Figure 3 shows the homographic projection from each view to the canonical square. The canonical square is any square defined by the engineer.

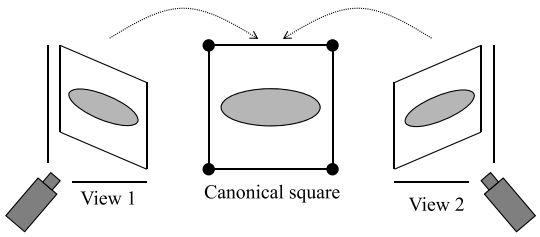


Fig. 3 Transformation to canonical square.

A polarization camera captures the azimuth angle ϕ of the target object (Sec. 2.2). We denote the 90-deg rotation of ϕ as vector **a**, which would be orthogonal to the surface normal (Sec. 2.3). Using the vector **a** and rotation matrix **R** of the camera parameter, the surface normal **n** is calculated using singular value decomposition (SVD) (Sec. 2.4).

Finally, the surface normal is integrated to height field.[15]

2.2 *Polarization*

We explain only linear polarization since circular polarization is not related to our method. Light is an electromagnetic wave, and an electromagnetic wave oscillating in only one direction is said to have perfectly linear polarization, while an electromagnetic wave oscillating isotropically in all directions is called unpolarized light. The intermediate state of such light is called partially polarized light. DOP[3–5] is one of the metrics used to represent the polarization state of light. Its value varies from 0 to 1, with 1 representing perfectly polarized light and 0 representing unpolarized light.

The maximum light observed while rotating the polarizer is denoted as $I_{\max}$, and the minimum light is denoted as $I_{\min}$. In this paper, the polarizer angle at which $I_{\min}$ is observed is called the azimuth angle ϕ. The surface normal is represented in polar coordinates, where the azimuth angle is denoted as ϕ and the zenith angle is denoted as θ (Fig. 4). The azimuth angle calculated from the polarization has 180-deg ambiguity since a linear polarizer has a 180-deg cycle. Thus, the azimuth angle of the surface normal will be either ϕ or $\phi + 180$ deg. The plane consisting of the incident light and surface normal vectors is called the reflection plane. The reflected light vector is also coplanar with the reflection plane since the surface is optically smooth. The orientation of the reflection plane is the same as the azimuth angle ϕ and $\phi + 180$ deg, which is defined on a certain xy plane and is defined as an angle between the x axis and the reflection plane projected on the xy plane. Since we capture images with a camera, the x axis and the y axis of the "image" coordinates is used.

2.3 *Calculating the Surface Normal from Two Viewpoints*

Section 2.2 described the relationship between the surface normal and the azimuth angle obtained from polarization. However, we cannot determine the surface normal uniquely because only the orientation of the reflection plane including the surface normal is obtained. We must observe the object from two viewpoints to solve this problem.

Figure 5 represents the situation of our problem. A camera has its coordinate system x axis, y axis, and z axis. The camera's z axis is along the optical axis. The reflection plane angle ϕ is the angle between the x axis of the camera coordinate system and the line caused by the intersection between the reflection plane and the xy plane.

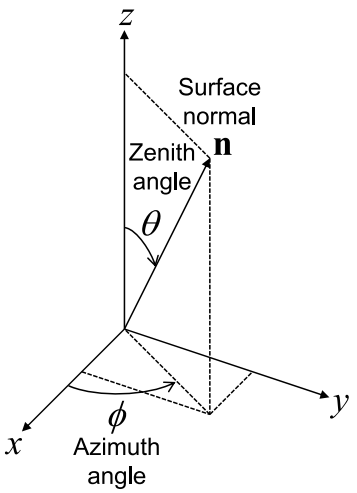


Fig. 4 Polar coordinates of surface normal.

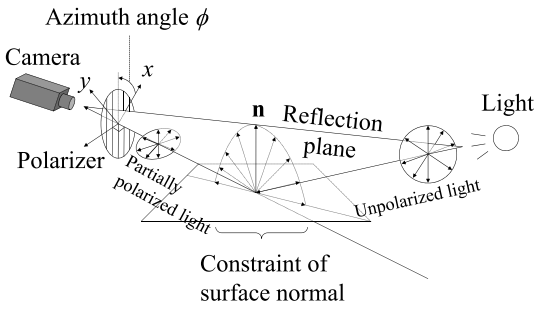


Fig. 5 Relationship between the surface normal and the reflection plane when observed from a single viewpoint.

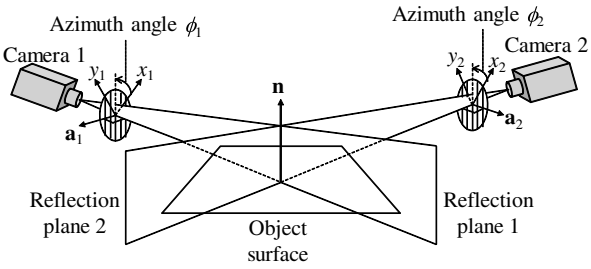


Fig. 6 Relationship between the surface normal and the reflection plane when observed from two viewpoints.

We analyze the two reflection plane angles at the same surface point, corresponding to the known 3-D geometry. Our method assumes that the 3-D geometry of the target object is almost a plane. The relationship between the surface normal vector and the azimuth angle is shown in Fig. 6. The relationship between the azimuth angles for each of the cameras, represented as ϕ_1 and ϕ_2, and the normal vector of the reflection plane, represented as $\mathbf{a}_1$ and $\mathbf{a}_2$, is shown in Eq. (1).

$$\mathbf{a}_1 = \begin{pmatrix} \cos(\phi_1 + 90 \ \text{deg}) \\ \sin(\phi_1 + 90 \ \text{deg}) \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} \cos(\phi_2 + 90 \ \text{deg}) \\ \sin(\phi_2 + 90 \ \text{deg}) \\ 0 \end{pmatrix}. \tag{1}$$

As shown in Fig. 6, the surface normal $\mathbf{n}$ is orthogonal to the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. After projecting the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ to the world coordinate system, we calculate the surface normal $\mathbf{n}$. The rotation matrix projecting the world coordinate system to each camera coordinate system is represented as $\mathbf{R}_1$ and $\mathbf{R}_2$. The inverse of each of these rotation matrices is its transpose, and they project back from the camera coordinate system to the world coordinate system. Thus, this situation is represented as Eq. (2).

$$\begin{pmatrix} \mathbf{a}_1^\top \mathbf{R}_1 \\ \mathbf{a}_2^\top \mathbf{R}_2 \\ \mathbf{0}^\top \end{pmatrix} \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2}$$

Namely, the world coordinate of $\mathbf{a}_1$ and $\mathbf{a}_2$ are $\mathbf{R}_1^\top \mathbf{a}_1$ and $\mathbf{R}_2^\top \mathbf{a}_2$. Since $\mathbf{R}_1^\top \mathbf{a}_1$ and $\mathbf{R}_2^\top \mathbf{a}_2$ are orthogonal to the surface normal $\mathbf{n}$, $(\mathbf{R}_1^\top \mathbf{a}_1) \cdot \mathbf{n} = 0$ and $(\mathbf{R}_2^\top \mathbf{a}_2) \cdot \mathbf{n} = 0$ hold. These formulae can be expressed, in another form, as $\mathbf{a}_1^\top \mathbf{R}_1 \mathbf{n} = 0$ and $\mathbf{a}_2^\top \mathbf{R}_2 \mathbf{n} = 0$ [Eq. (2)].

2.4 *Calculating the Surface Normal from Multiple Viewpoints*

This section explains the estimation process for the surface normal from the azimuth angle obtained from multiple viewpoints.

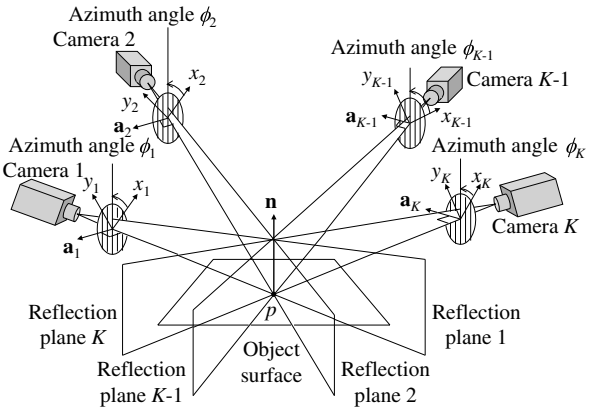


Fig. 7 Relationship between the surface normal and the azimuth angle observed from multiple viewpoints.

Figure 7 shows the relationship between the surface normal $\mathbf{n}$ of the surface point p and the azimuth angle obtained from K viewpoints. In Fig. 7, ϕ_k represents the azimuth angle of the surface point p observed by the camera $k = (1,2,\cdots,K)$, and $\mathbf{a}_k$ represents the vector orthogonal to the reflection plane under the coordinate system of the camera k.

The rotation matrix $\mathbf{R}_k$ represents the transformation from the world coordinate system to the local coordinate system of the camera indicated by k. Similar to Eq. (2), Eqs. (3) or (4) holds.

$$\begin{pmatrix} \mathbf{a}_1^\top \mathbf{R}_1 \\ \mathbf{a}_2^\top \mathbf{R}_2 \\ \vdots \\ \mathbf{a}_K^\top \mathbf{R}_K \end{pmatrix} \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{3}$$

Or, in an other form,

$$\mathbf{An} = \mathbf{0}. \tag{4}$$

The surface normal $\mathbf{n}$, which satisfies Eq. (4) in the least-squares sense, is estimated using SVD.[28] The $K \times 3$ matrix $\mathbf{A}$ is decomposed by SVD as follows:

$$\begin{pmatrix} \mathbf{a}_1^\top \mathbf{R}_1 \\ \mathbf{a}_2^\top \mathbf{R}_2 \\ \vdots \\ \mathbf{a}_K^\top \mathbf{R}_K \end{pmatrix} = \mathbf{UWV}^\mathrm{T} = \mathbf{U} \begin{pmatrix} w_1 & & \\ & w_2 & \\ & & 0 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{pmatrix}. \tag{5}$$

Here, $\mathbf{U}$ is a $K \times 3$ orthogonal matrix, $\mathbf{W}$ is a 3×3 diagonal matrix with non-negative values, and $\mathbf{V}^\top$ is a 3×3 orthogonal matrix. The diagonal element w_i of matrix $\mathbf{W}$ is the singular value of matrix $\mathbf{A}$, and the singular vector corresponding to w_i is $\mathbf{v}_i$. Owing to the relationship between the surface normal and the reflection planes, the rank of matrix $\mathbf{A}$ is at most 2; thus, one of the three singular values becomes 0 (see Miyazaki et al.,[27] for the proof). The surface normal $\mathbf{n}$ is represented as Eq. (6),[28] which is calculated from the singular vector that has the smallest singular value, namely, the third row of $\mathbf{V}^\top$ in Eq. (5).

$$\mathbf{n} = s\mathbf{v}_3^\top. \tag{6}$$

In general, s is an arbitrary scalar coefficient; however, since the surface normal and the singular vectors are normalized vectors, s would be either $+1$ or -1. Whether s is positive or negative is determined so that the surface normal faces toward the camera. The surface normal estimated by Eq. (6) is the optimal value that minimizes the squared error of Eq. (4) formulated by

K equations. The input data must be obtained from two or more viewpoints since the rank of matrix $\mathbf{A}$ is 2.

2.5 *Camera Parameters*

Equation (3) or Eq. (4) calculates the surface normal from the azimuth angle under multiple viewpoints. To solve Eq. (4), the azimuth angle should be analyzed at corresponding points among multiple viewpoints. The corresponding points are determined by homography, as shown in Sec. 2.6. Equation (4) also requires the rotation matrices of each camera, namely, the extrinsic parameter of each camera should be known.

Our paper represents the projection from 3-D vertex (X, Y, Z) to two-dimensional vertex (x, y) as Eq. (7).[29]

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{7}$$

In Eq. (7), we do not describe the camera center parameter (C_x, C_y) for clarity since we assume the pinhole camera model. We do not explain the detailed implementation to estimate these parameters $f, t_1, t_2, t_3, r_{11}, r_{12}, \cdots, r_{33}$.

2.6 *Homography Transform*

Homography is a projection from a certain quadrangle to another certain quadrangle represented under the homographic projection. Homography represents one-to-one correspondence between two planes without redundancy or lack of information. Therefore, it is natural to use homography in our work since the target object is almost planar.

The homogeneous coordinate is defined as follows using (ξ_1, ξ_2, ξ_3) $(\xi_3 \neq 0)$, where one element is added to the coordinates (x', y').

$$x' = \frac{\xi_1}{\xi_3}, \quad y' = \frac{\xi_2}{\xi_3}. \tag{8}$$

Homographic projection from a certain quadrangle (x, y) to another certain quandrangle (x', y') is represented as follows:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \sim \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{9}$$

Namely, homographic projection is represented by homography matrix $h_{11}, h_{12}, \cdots, h_{33}$. Point (x, y) is projected to the point $(x', y') = (\xi_1/\xi_3, \xi_2/\xi_3)$ by this homography matrix.

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \tag{10}$$

$$y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}. \tag{11}$$

Figure 8 is an example in which vertices of quadrangle (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4) correspond to vertices of quadrangle (x_1', y_1'), (x_2', y_2'), (x_3', y_3'), and (x_4', y_4').

Scaling the 3×3 homography matrix $h_{11}, h_{12}, \ldots, h_{33}$ results in the same transformation. Thus, we fix one element as follows to uniquely determine the homography matrix:

$$h_{33} = 1. \tag{12}$$

Substituting the above equation into Eqs. (10) and (11) results in Eqs. (13) and (14).

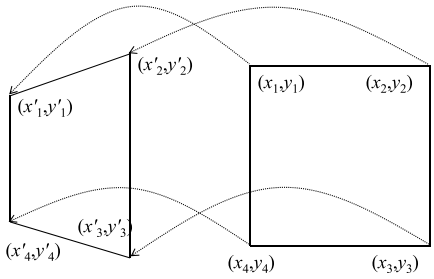


Fig. 8 Homographic projection from a certain quadrangle to another certain quadrangle.

$$xh_{11} + yh_{12} + h_{13} - xx'h_{31} - yx'h_{32} = x', \tag{13}$$

$$xh_{21} + yh_{22} + h_{23} - xy'h_{31} - yy'h_{32} = y'. \tag{14}$$

Concatenating Eqs. (13) and (14) for four vertices results in Eq. (15).

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x'_2 & -y_2x'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y'_2 & -y_2y'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3x'_3 & -y_3x'_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3y'_3 & -y_3y'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x'_4 & -y_4x'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y'_4 & -y_4y'_4 \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{pmatrix} = \begin{pmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \\ x'_4 \\ y'_4 \end{pmatrix}. \tag{15}$$

Since we have eight unknowns $(h_{11}, h_{12}, \ldots, h_{32})$ and eight equations [eight rows of the leftmost matrix in Eq. (15)], a closed-form solution exists. Solving this results in the homography matrix shown below:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix}. \tag{16}$$

Using the homography matrix $\mathbf{H}$ [Eq. (16)], the corresponding points between two quadrangles (Fig. 8) are expressed by Eq. (9). Suppose that the homography of camera 1 is $\mathbf{H}_1$ and that of camera 2 is $\mathbf{H}_2$. Figure 3 shows the homographic projection from each view to the canonical square. The canonical square is any square defined by the engineer. Suppose that the pixel position of the canonical square is (x, y). The corresponding points of camera 1 (x'_1, y'_1) and camera 2 (x'_2, y'_2) are calculated as follows:

$$\begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix} \sim \mathbf{H}_1 \begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x'_2 \\ y'_2 \end{pmatrix} \sim \mathbf{H}_2 \begin{pmatrix} x \\ y \end{pmatrix}. \tag{17}$$

Namely, the two points (x'_1, y'_1) and (x'_2, y'_2) correspond while the point (x, y) acted as a mediator.

3 Experiment

3.1 *Experimental Setup*

As shown in Fig. 9, the target object is surrounded by white material. This white environment acts as a light source and illuminates the target object from every direction. Cheap foaming polystyrene is used in our experiment, and it is located not strictly but roughly. Once we set this white enclosure, we do not need to move it like with photometric stereo, which needs

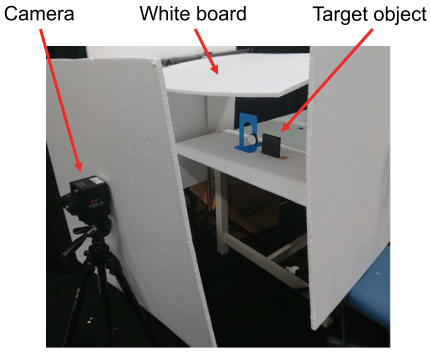


Fig. 9 Experimental environment.

Fig. 10 Polarization camera.

Table 1 Specification of the camera.

Manufacturer	FluxData Inc., New York
Product name	FD-1665P
Sensor	Sony ICX414
Resolution	659×494
Pixel size	$9.9 \times 9.9\ \mu$m
Configuration	0, 45, 90 linear polarizer
Frame rate	74 fps
Interface	IEEE-1394b

to move the light sources. The white board is illuminated by ordinary room light which is set in an ordinary room. Often, the white board is unnecessary since wall, floor, and ceiling act as an illuminator.[18]

The camera used is shown in Fig. 10 and Table 1. Since we have only one camera (because a polarization camera is expensive), we rotate the target object instead of rotating the camera. Note that observing a target object with multiple cameras and observing the target object rotated in multiple angles with a single camera are mathematically the same.

3.2 *Pseudo-Color Representation of the Result*

The following sections show some results of our method. For visualization, the azimuth angle and the surface normal are represented by pseudo-color. Figures 11(a) and 11(b) show the

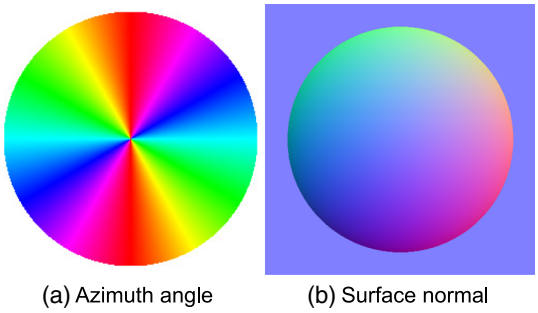


Fig. 11 Pseudo-color representation of an ideal sphere: (a) azimuth angle and (b) surface normal.

pseudo-color representation of the azimuth angle and the surface normal of an ideal hemisphere, respectively.

3.3 *Result of Ellipsoid*

The target object is shown in Fig. 12. We generated the object using a 3D printer, so we can compare the result with the ground truth, which is the digital data input to the 3D printer. The size of the square is 10×10 cm, the diameter of the long axis of the ellipse is 7.5 cm, the diameter of the short axis of the ellipse is 2.5 cm, and the maximum deepness of the concave part is 0.625 cm. The unique characteristic of our method is that we can estimate the shape of the cracks. First, we evaluate the performance of the proposed method. To guarantee the statistical reliability, we need to estimate the surface normal with wide variety and wide area. This is why we first measure the concave ellipsoid.

We took one image from each of 15 different directions (Fig. 13). Pseudo-color representation of the surface normal of our method is shown in Fig. 14 and that of the ground truth is shown in Fig. 15. Note that our method successfully estimated the shape, which is almost the same as the true shape. The estimated shape is shown in Figs. 16 and 18(c), while the ground truth is shown in Figs. 17 and 18(a). The error is calculated as the angle between two surface normals of the estimated and the ground truth. Error is shown in Fig. 19(b); the average error was 4.49 deg.

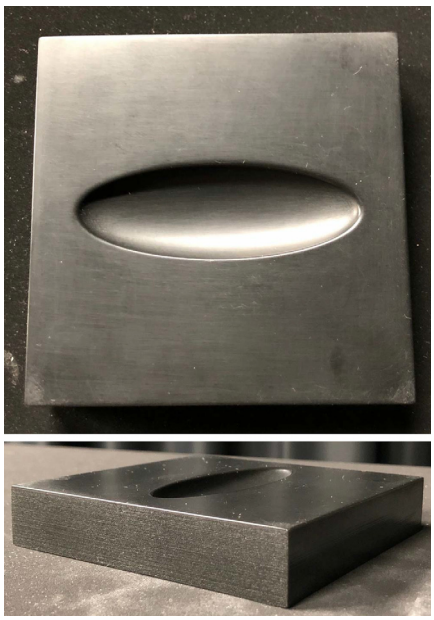


Fig. 12 Target object (ellipsoid).

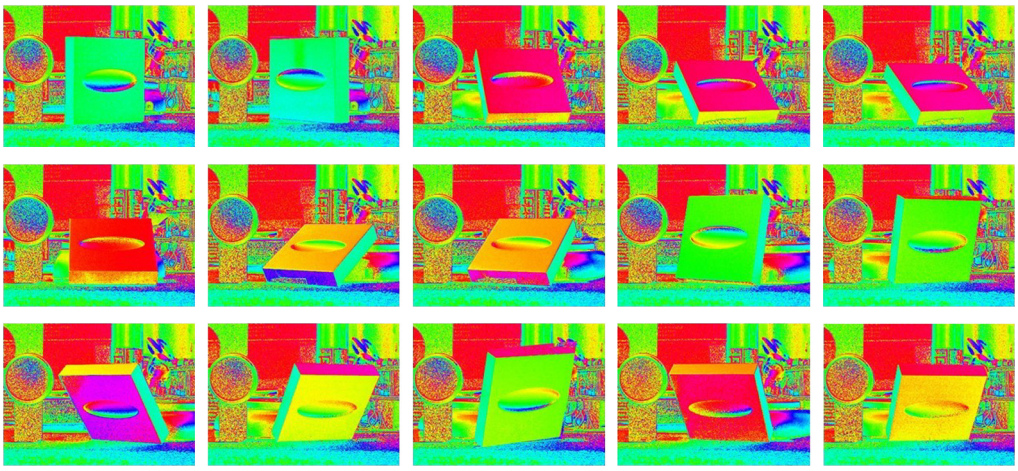


Fig. 13 Input image (ellipsoid).

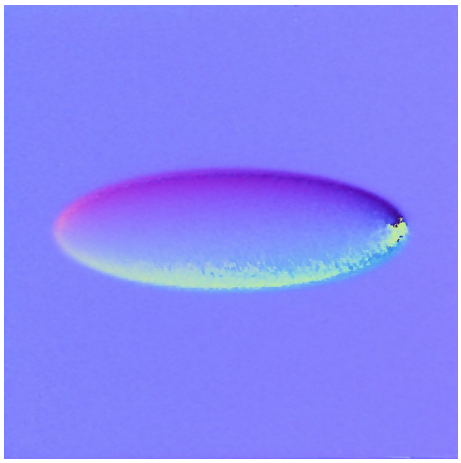


Fig. 14 Estimated surface normal (ellipsoid).

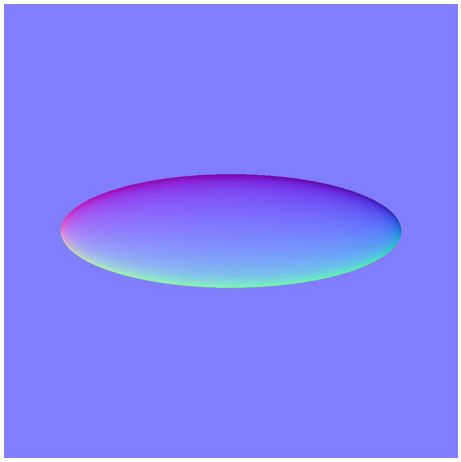


Fig. 15 Ground truth of surface normal (ellipsoid).

3.4 *Comparison to Photometric Stereo*

To prove the effectiveness of our method, we compare our method with the result of the photometric stereo method.[30]

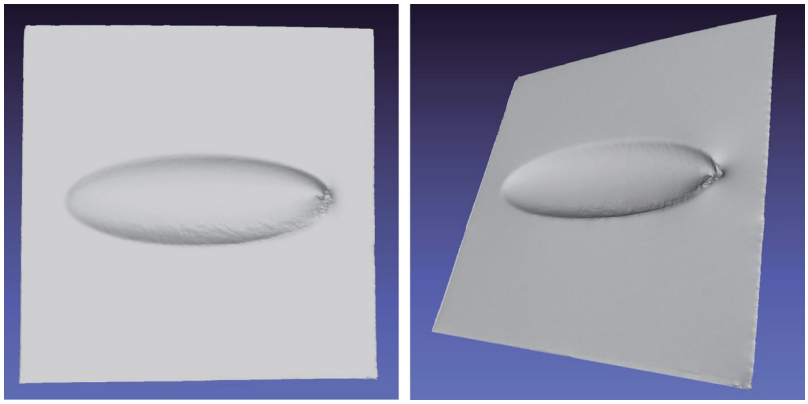


Fig. 16 Estimated shape (ellipsoid).

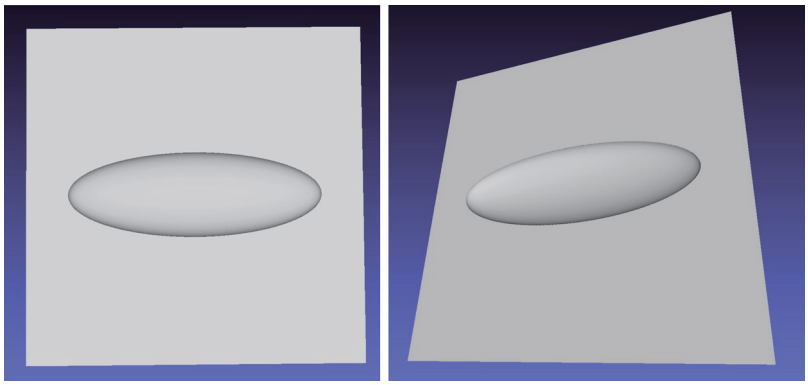


Fig. 17 Ground truth of shape (ellipsoid).

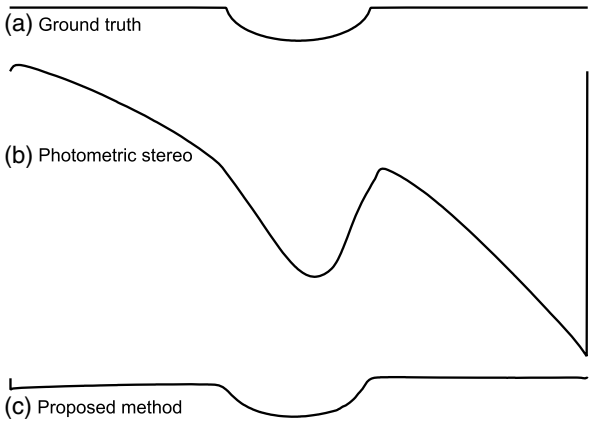
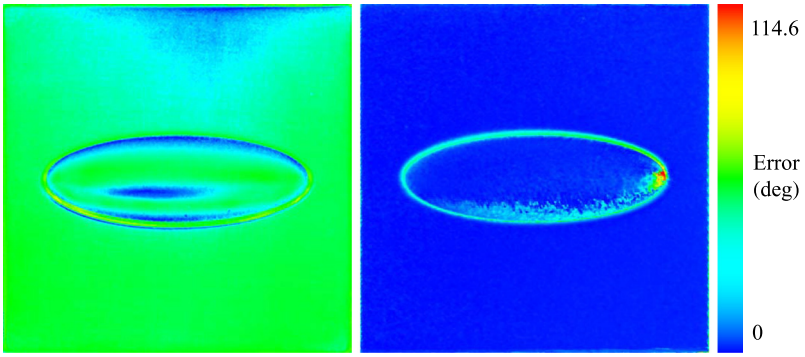


Fig. 18 Intersection shape (ellipsoid): (a) ground truth, (b) photometric stereo, and (c) proposed method.

Photometric stereo from 15 lights is applied to the object shown in Fig. 12, and the input images are shown in Fig. 20. The surface normal of the photometric stereo method is shown in Fig. 21, and the estimated shape is shown in Figs. 18(b) and 22. Photometric stereo assumes Lambertian reflection, though the actual reflection is specular reflection, and thus, the shape is distorted. The error is shown in Fig. 19(a), and the average error was 42.3 deg. Since our error is 4.49 deg, the performance of our method is better.



(a) Photometric stereo (b) Proposed method

Fig. 19 Estimation error (ellipsoid): (a) photometric stereo and (b) proposed method.

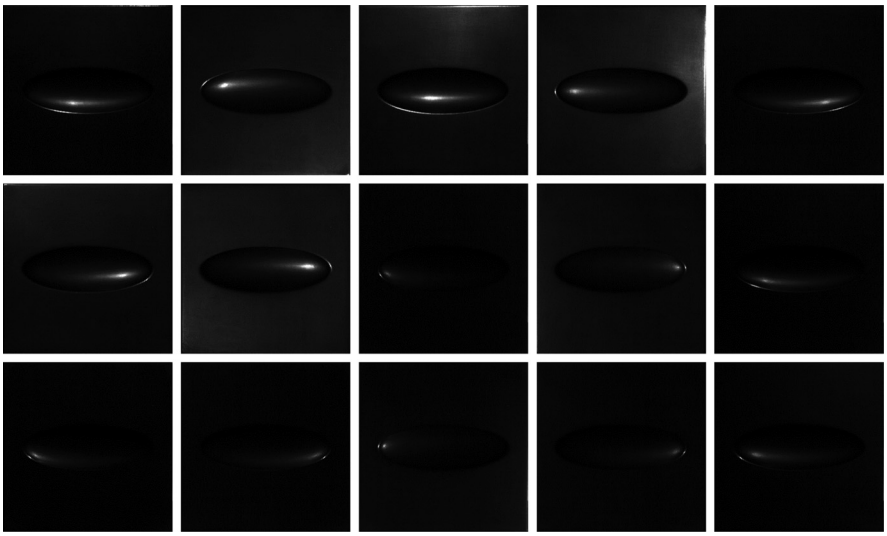


Fig. 20 Input image of photometric stereo (ellipsoid).

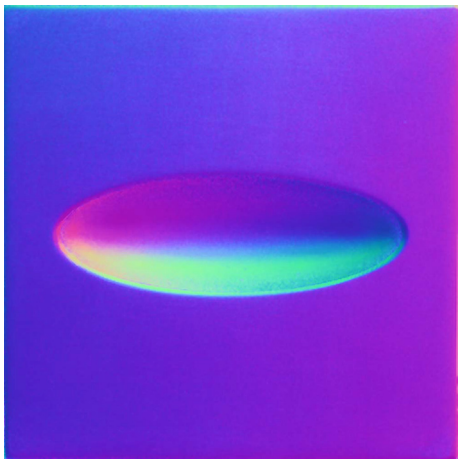


Fig. 21 Surface normal of photometric stereo (ellipsoid).

3.5 *Result of Convex Object*

Our method can be applied not only to concave objects but also to convex objects. To prove the wide applicability of our method, we measure a convex object. The target object is shown in Fig. 23. The input images of our method are shown in Fig. 24, while those of photometric stereo

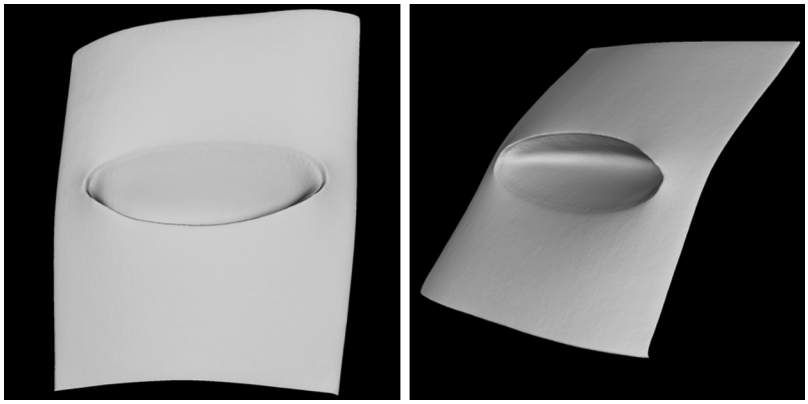


Fig. 22 Shape of photometric stereo (ellipsoid).

Fig. 23 Photograph of target object (convex).

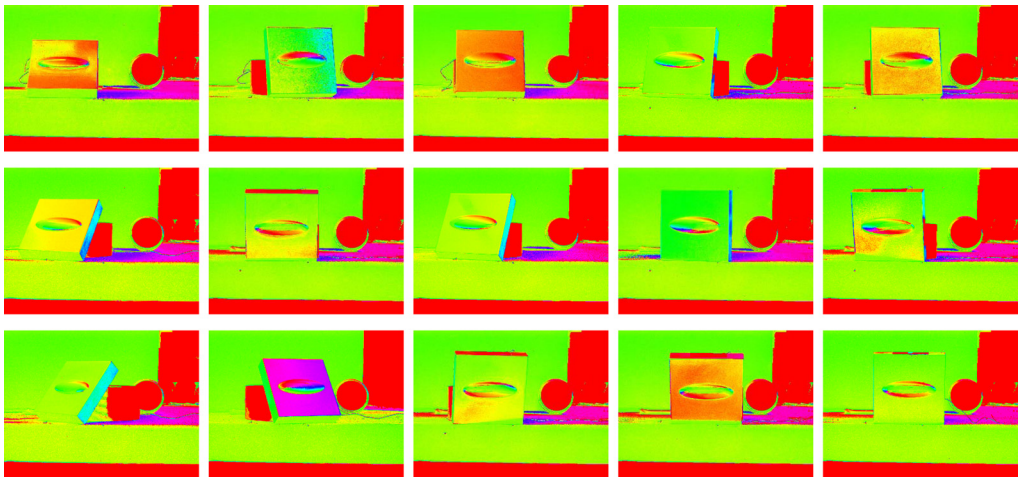


Fig. 24 Input data of our method (convex).

Fig. 25 Input images of photometric stereo (convex).

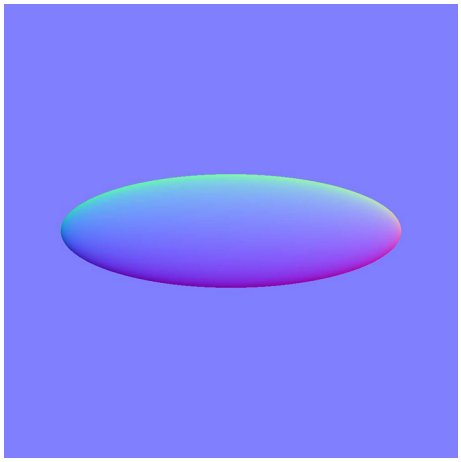


Fig. 26 Surface normal of ground truth (convex).

are shown in Fig. 25. The surface normal of the ground truth, the photometric stereo method, and the proposed method is shown in Figs. 26–28. The shape of the ground truth, the photometric stereo method, and the proposed method is shown in Figs. 29–31. The cross section of the shape of the ground truth, the photometric stereo method, and the proposed method is shown in Figs. 32(a)–32(c). The error of the photometric stereo method and the proposed method is shown in Figs. 33(a) and 33(b). The average error of the photometric stereo method was 47.0 deg, while that of the proposed method was 12.9 deg.

3.6 *Result of Stripes*

To evaluate the performance of our method depending on the width of cracks, three different concave shapes with different widths are measured. Figure 34 shows the target object, and Fig. 35 shows the input images. Also, 15 images are taken, one for each direction. The surface normal of our method is shown in Fig. 36, and that of ground truth is shown in Fig. 37. The estimated shape is shown in Figs. 38 and 40(b), while ground truth is shown in Figs. 39 and 40(a).

The error map is shown in Fig. 41; the average error was 7.18 deg.

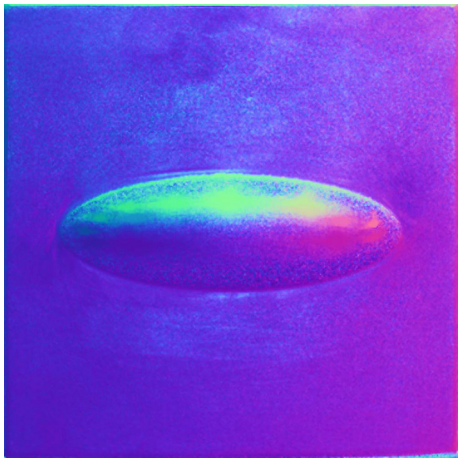


Fig. 27 Surface normal of photometric stereo (convex).

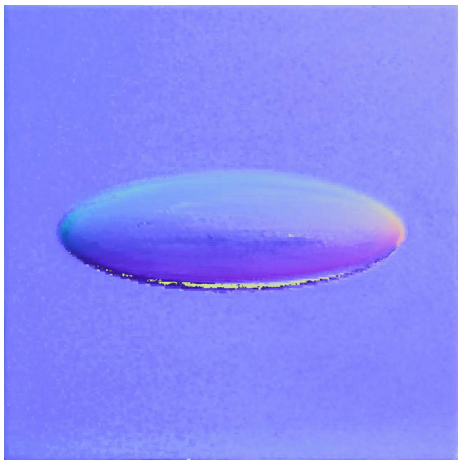


Fig. 28 Surface normal of proposed method (convex).

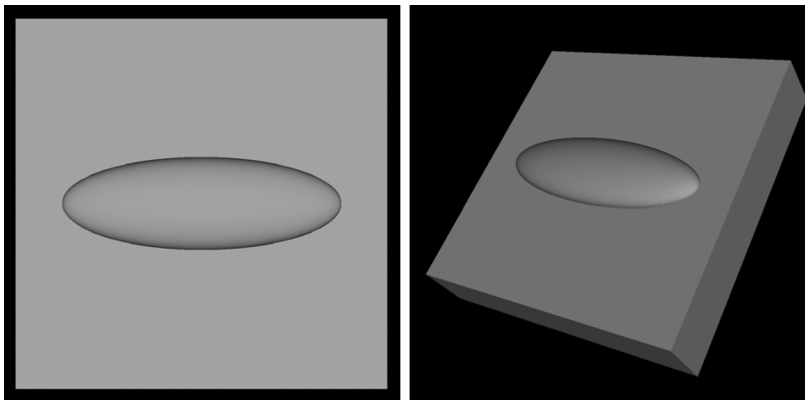


Fig. 29 Shape of ground truth (convex).

3.7 *Result of Worm*

To simulate an acutal situation, we applied our method to a crack that is not shaped in a straight line. Figs. 42 and 43 show the target object and the input images, respectively, and Figs. 44 and 45 show the estimated surface normal and the estimated shape, respectively.

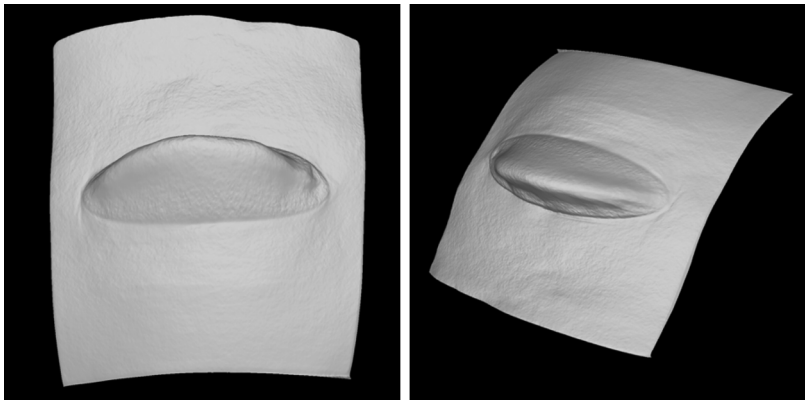


Fig. 30 Shape of photometric stereo (convex).

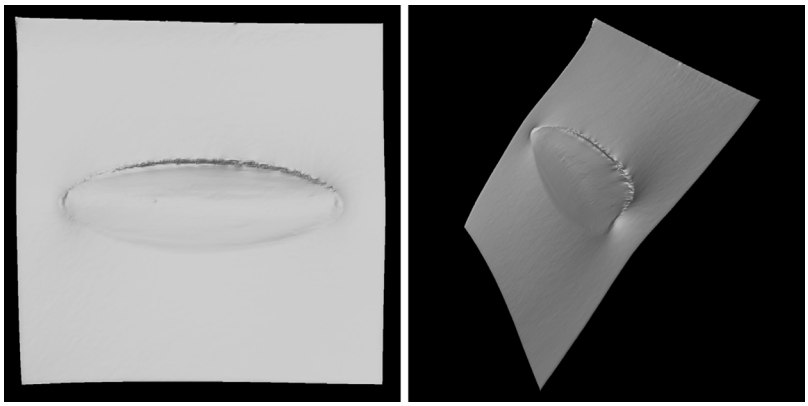


Fig. 31 Shape of proposed method (convex).

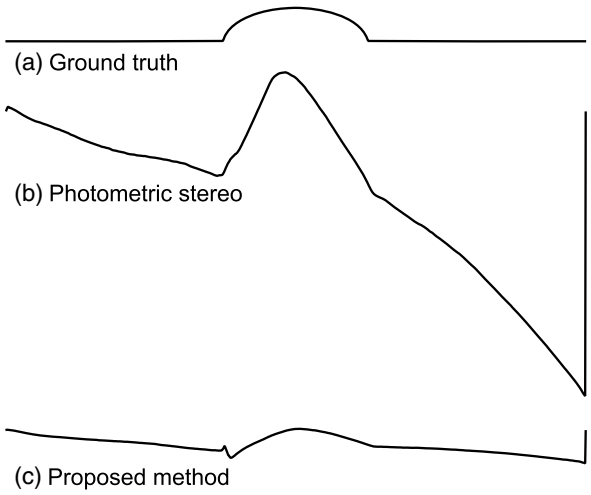
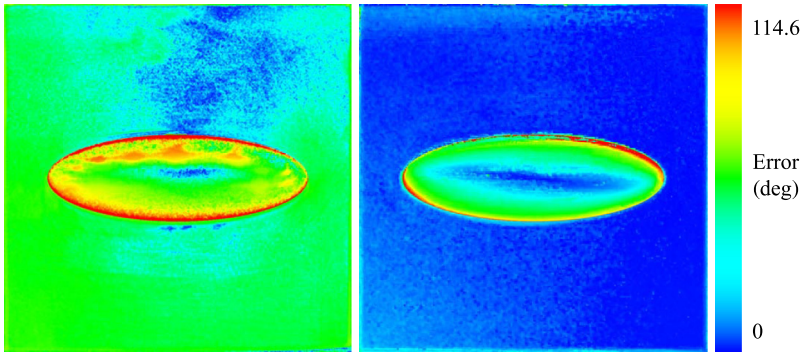


Fig. 32 Intersection shape (convex): (a) ground truth, (b) photometric stereo, and (c) proposed method.

3.8 *Discussion*

As shown in Fig. 41, the narrower the concave part is, the worse the result is. This is because the light is not illuminated satisfactorily in the narrow concave part. In addition, inter-reflection becomes strong at the narrow concave part.



(a) Photometric stereo (b) Proposed method

Fig. 33 Estimation error (convex): (a) photometric stereo and (b) proposed method.

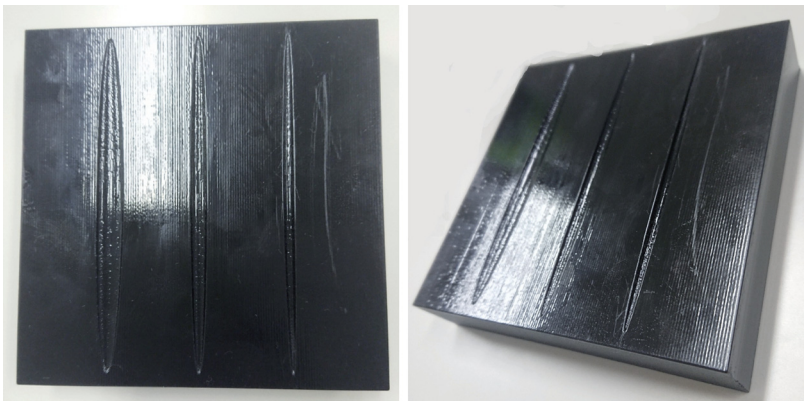


Fig. 34 Target object (stripe).

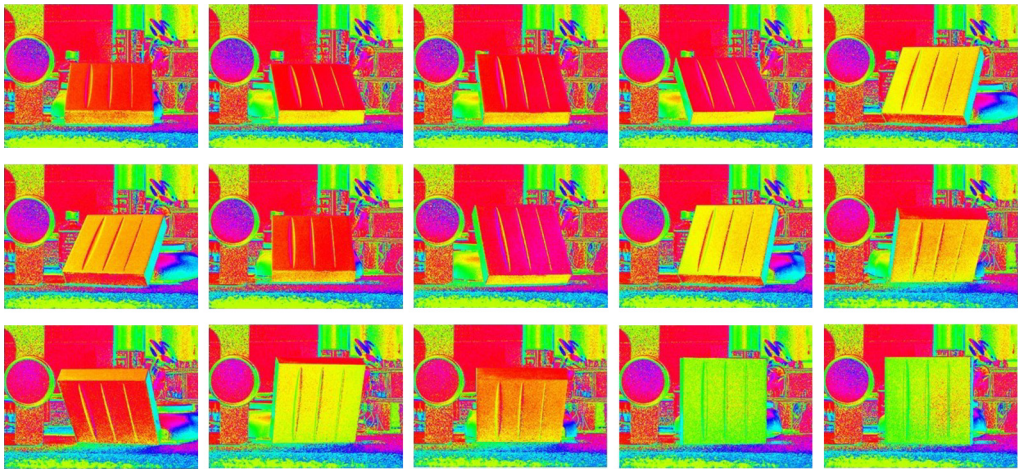


Fig. 35 Input images (stripe).

4 Conclusion

We propose a shape estimation method from polarization images obtained from multiple viewpoints. The proposed method computes the surface normal using SVD to minimize the least-squared error. It can estimate the shapes of the concave part of planar objects that are black and have high specularity. It is usually difficult to estimate the shape of a planar object with small details; however, our algorithm fully utilizes the property that the target object is almost planar.

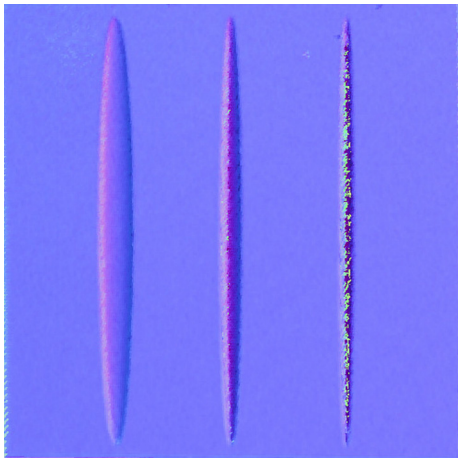


Fig. 36 Estimated surface normal (stripe).

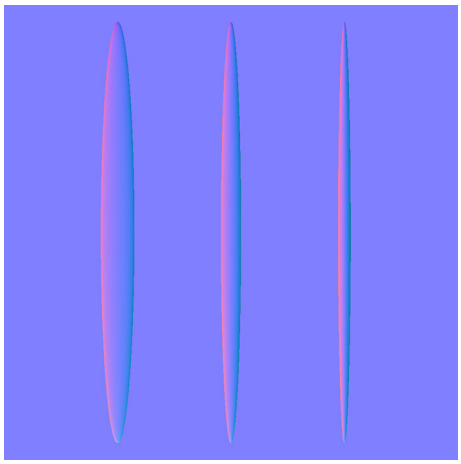


Fig. 37 Ground truth of surface normal (stripe).

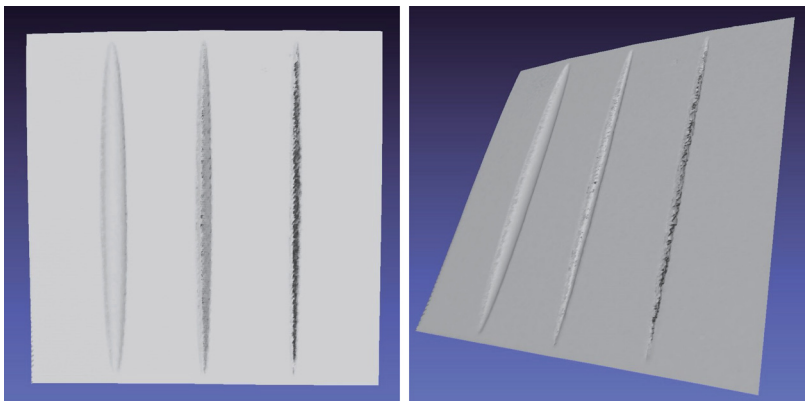


Fig. 38 Estimated shape (stripe).

What is interesting in our method is that, even if we assume that the object is planar, the shape of the concave part is also successfully determined.

The experiments show that our method can estimate the shape of the crack. This property demonstrates that our method is useful for investigating product inspection in factories, damage

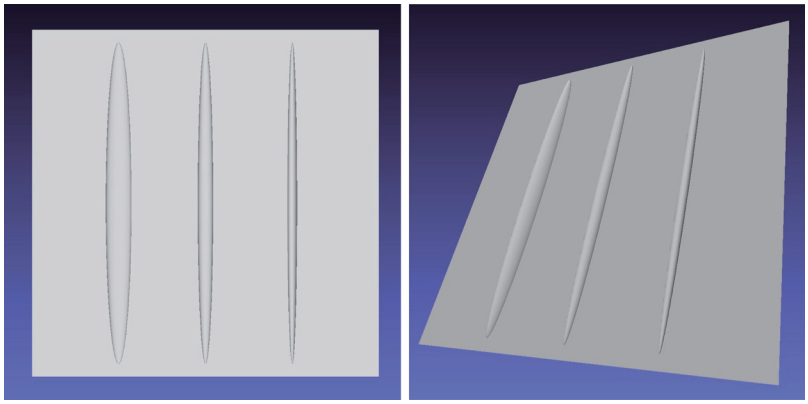


Fig. 39 Ground truth of shape (stripe).

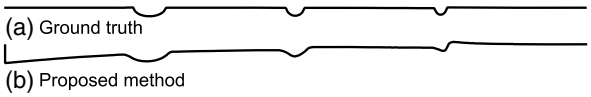


Fig. 40 Intersection shape (stripe): (a) ground truth and (b) estimated shape.

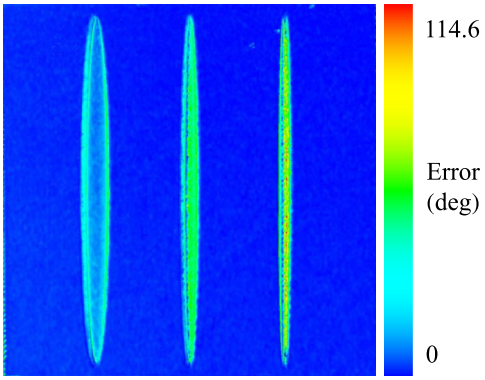


Fig. 41 Error of our method (stripe).

Fig. 42 Target object (worm).

inspection in architecture, age estimation from skin wrinkles, and so on. For example, factories want to know the reason for the defect of the product since they want to fix the problem and decrease the defects. To analyze the reason, the shape of the defects is necessary, and our method is useful for this purpose.

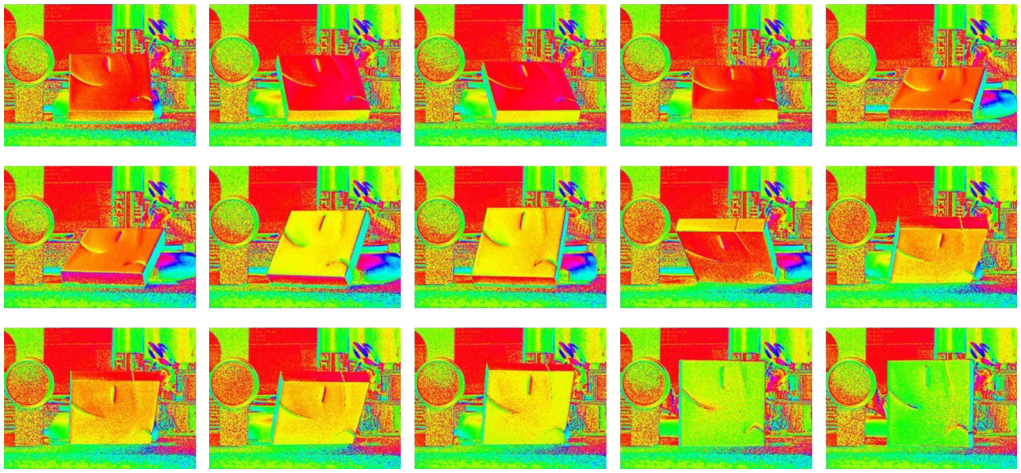


Fig. 43 Input images (worm).

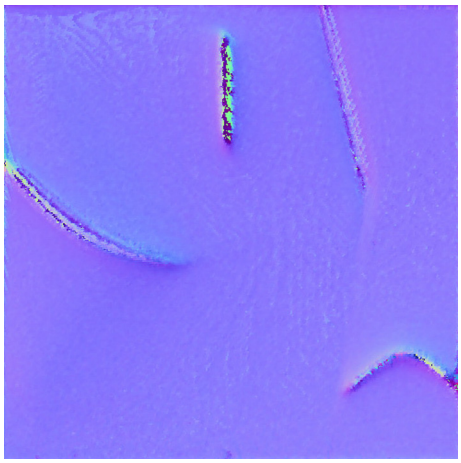


Fig. 44 Estimated surface normal (worm).

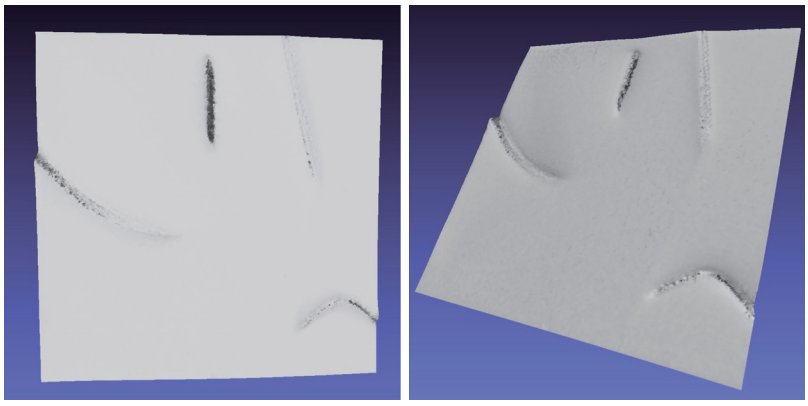


Fig. 45 Estimated shape (worm).

The disadvantage of our method is that the shape where the light has not reached cannot be estimated. However, this disadvantage does not apply only to our method but also to any method in the image processing field since an image cannot be observed if the scene is not illuminated. Our future work will develop a measurement system that illuminates the target object from any directions.

Acknowledgments

This research was supported in part by the Ministry of Education, Culture, Sports, Science, and Technology KAKENHI Grant Nos. 15H05925 and 15H02727 and in part by the Japan Society for the Promotion of Science KAKENHI Grant No. 24700176.

References

1. A. Mohan and S. Poobal, "Crack detection using image processing: a critical review and analysis," *Alexandria Eng. J.* **57**, 787–798 (2018).
2. I. Ihrke et al., "Transparent and specular object reconstruction," *Comput. Graphics Forum* **29**(8), 2400–2426 (2010).
3. M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, London (1959).
4. E. Hecht, *Optics*, Addison-Wesley, Reading, Massachusetts (2002).
5. W. A. Shurcliff, *Polarized Light: Production and Use*, Harvard University Press, Cambridge, Massachusetts (1962).
6. K. Koshikawa and Y. Shirai, "A model-based recognition of glossy objects using their polarimetrical properties," *Adv. Rob.* **2**(2), 137–147 (1987).
7. G. C. Guarnera et al., "Estimating surface normals from spherical Stokes reflectance fields," *Lect. Notes Comput. Sci.* **7584**, 340–349 (2012).
8. O. Morel et al., "Active lighting applied to three-dimensional reconstruction of specular metallic surface by polarization imaging," *Appl. Opt.* **45**(17), 4062–4068 (2006).
9. M. Saito et al., "Measurement of surface orientations of transparent objects by use of polarization in highlight," *J. Opt. Soc. Am. A* **16**(9), 2286–2293 (1999).
10. B. A. Barbour, "Apparatus and method for extracting information from electromagnetic energy including target 3D structure and materials," U S. Patent, 8320661 (2012).
11. Y. Kobayashi et al., "Reconstructing shapes and appearances of thin film objects using RGB images," in *IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 3774–3782 (2016).
12. D. Miyazaki et al., "Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths," *J. Opt. Soc. Am. A* **19**(4), 687–694 (2002).
13. D. Miyazaki et al., "Polarization-based inverse rendering from a single view," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 982–987 (2003).
14. D. Miyazaki, M. Kagesawa, and K. Ikeuchi, "Transparent surface modeling from a pair of polarization images," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(1), 73–82 (2004).
15. D. Miyazaki and K. Ikeuchi, "Shape estimation of transparent objects by using inverse polarization raytracing," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(11), 2018–2030 (2007).
16. Y. Yu, D. Zhu, and W. A. Smith, "Shape-from-polarisation: a nonlinear least squares approach," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 2969–2976 (2017).
17. L. B. Wolff and T. E. Boult, "Constraining object features using a polarization reflectance model," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(7), 635–657 (1991).
18. S. Rahmann and N. Canterakis, "Reconstruction of specular surfaces using polarization imaging," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit.*, pp. 149–155 (2001).
19. S. Rahmann, "Reconstruction of quadrics from two polarization views," in *Proc. Iberian Conf. Pattern Recognit. and Image Anal.*, pp. 810–820 (2003).
20. G. A. Atkinson and E. R. Hancock, "Shape estimation using polarization and shading from two views," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(11), 2001–2017 (2007).
21. G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," *IEEE Trans. Image Process.* **15**(6), 1653–1664 (2006).
22. C. P. Huynh, A. Robles-Kelly, and E. R. Hancock, "Shape and refractive index from single-view spectro-polarimetric images," *Int. J. Comput. Vision* **101**, 64–94 (2013).
23. A. Kadambi et al., "Polarized 3D: high-quality depth sensing with polarization cues," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 3370–3378 (2015).

24. W. Smith, R. Ramamoorthi, and S. Tozza, "Height-from-polarisation with unknown lighting or albedo," *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(12), 2875–2888 (2018).
25. Z. Cui et al., "Polarimetric multi-view stereo," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1558–1567 (2017).
26. L. Yang et al., "Polarimetric dense monocular slam," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 3857–3866 (2018).
27. D. Miyazaki et al., "Surface normal estimation of black specular objects from multiview polarization images," *Proc. SPIE* **56**(4), 041303 (2016).
28. W. H. Press et al., *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge (1997).
29. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge (2004).
30. R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.* **19**(1), 139–144 (1980).

Daisuke Miyazaki is an associate professor at Hiroshima City University. He received his BS and MS degrees in science from the University of Tokyo in 2000 and 2002, respectively, and his PhD in information science and technology in 2005. His research interests include physics-based vision. He is a member of ACM and IEEE.

Biographies of the other authors are not available.