Shape estimation of concave specular object from multiview polarization
Daisuke Miyazaki, Ryosuke Furuhashi, Shinsaku Hiura,
Journal of Electronic Imaging,
vol. 29, no. 4, 041006, 2020.1

Shape estimation of concave specular object from multiview polarization

Daisuke Miyazaki[a,*], **Ryosuke Furuhashi**[a,b], **Shinsaku Hiura**[a,c]
[a]Hiroshima City University, Japan
[b]Present address: DISEC Corporation, Japan
[c]Present address: University of Hyogo, Japan

Abstract. This paper proposes a method to estimate the surface normal of concave objects. The target object of our method has 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. In order 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 different views. It 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. In order to get out of a tough spot, we assume that the target object is almost planar. Under this assumption, the surface normal of the object is uniquely determined. This paper shows that the surface normal of the non-planar part can be also estimated using the proposed method.

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

***Daisuke Miyazaki**, miyazaki@hiroshima-cu.ac.jp

1 Introduction

Factories in industrial field have a high demand to estimate the shape of crack since it is quite important for quality control of the products. Although there are many methods which detect cracks,[1] little method 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 which estimates the surface normal of black specular object with concave shape, by analyzing the polariztion state of the reflected light, where the target object is observed from multiple views.

3D modeling techniques have been intensively investigated in the field of computer vision. The techniques used can be categorized into two types, the geometric approach and the photometric approach. Geometric approach uses the geometrical structure of the scene, such as time-of-flight laser range sensor, multinocular stereo, or structured light projection. 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 clue 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 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 viewpoints in order 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 3D geometry obtained by a time-of-flight (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 structre-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 3D 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. In order to determine the corresponding point among multiple views, we assume the target object as planar. However, this assumption solely can simply produce a planar shape, thus we additionally use polarization information in order to estimate the non-planar 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

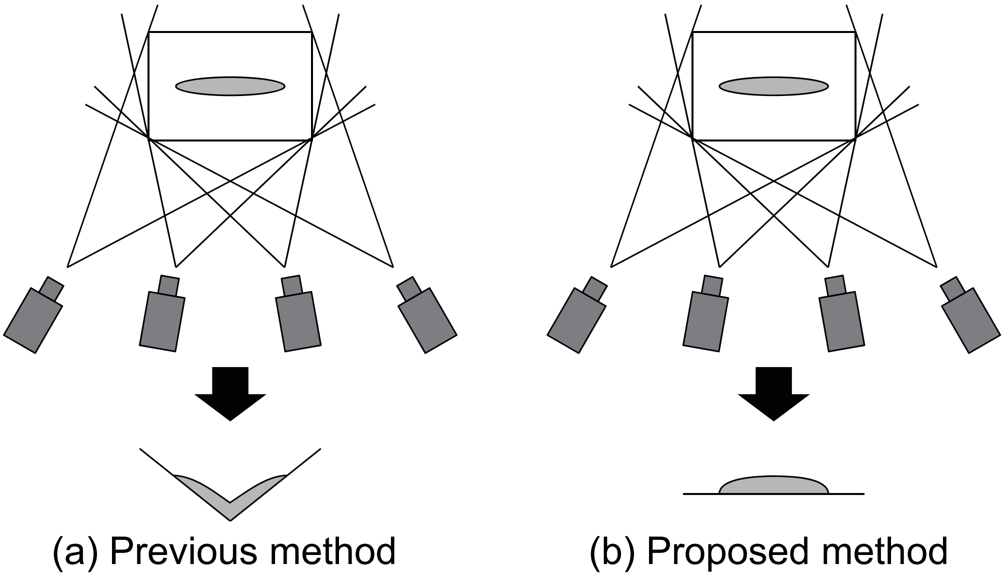

Fig 1 Our contribution: (a) Previous method which is based on visual hull which is not suited to estimate planar shapes and (b) proposed method which is suited to estimate concave shapes which is almost planar.

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 small size.

Miyazaki's method[27] relies on the visual hull. It is difficult to estimate a planar shape using visual hull, and 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 than their method[27] if the target object is almost planar.

We describe our method in Section 2 and present our results in Section 3. The theory shown in Section 2 assumes that the target object must be completely planar. However, Section 3 emprically 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 Section 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).

Since we observe the target object from multiple viewpoints, we calibrate each viewpoint in order 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 square for clarity. The area which is surrounded by these markers is the target area. Using these markers, we estimate the homography $\mathbf{H}$ (Section 2.6) and rotation $\mathbf{R}$ (Section 2.5). Fig. 3 shows the homographic projection from each view to canonical square. Canonical square can be any square defined by the engineer.

Polarization camera captures the azimuth angle ϕ of the target object (Section 2.2). We denote the 90° rotation of ϕ as vector $\mathbf{a}$, which would be orthogonal to surface normal (Section 2.3). Using the vector $\mathbf{a}$ and rotation matrix $\mathbf{R}$ of camera parameter, surface normal $\mathbf{n}$ is calculated using SVD (singular value decomposition) (Section 2.4).

Finally, 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 electromagnetic wave oscillating in only one direction is said to have perfectly linear polarization, while electromagnetic wave oscillating isotropically in all directions is called unpolarized light. The intermediate state of such light is called partially polarized light.

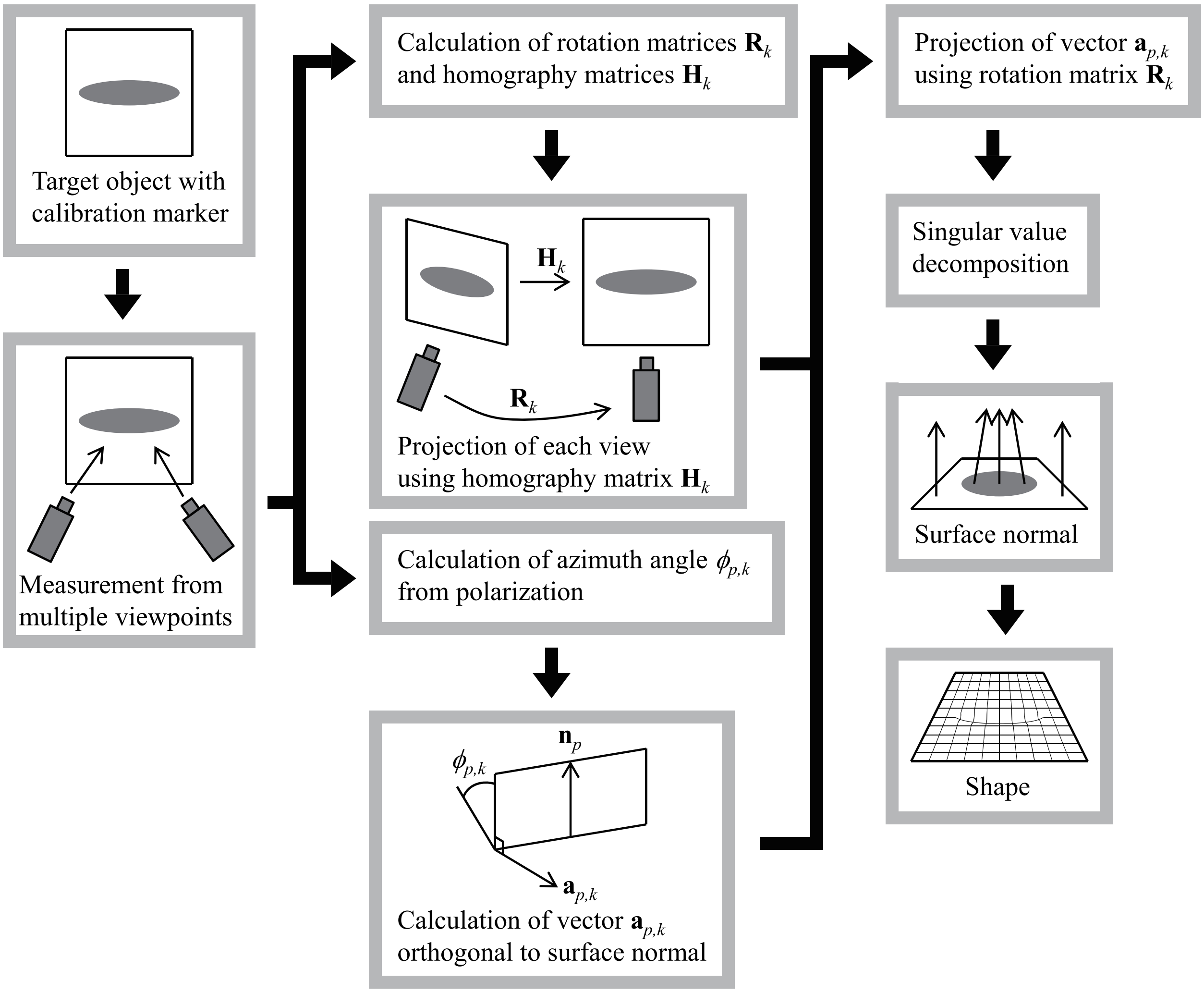

Fig 2 Algorithm flow.

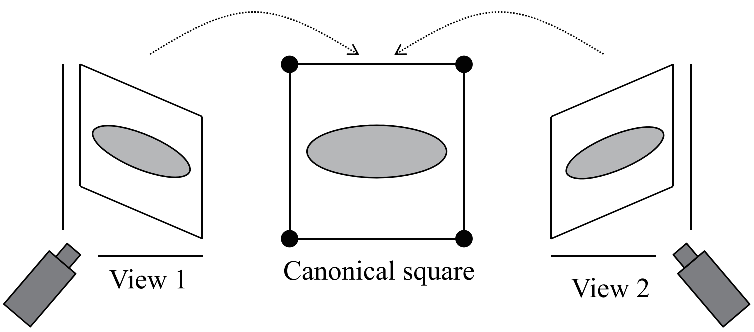

Fig 3 Transformation to canonical square.

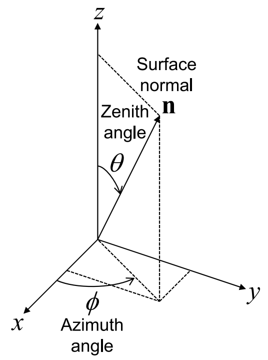

Fig 4 Polar coordinates of surface normal.

DOP (degree of polarization)[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 $180°$-ambiguity since linear polarizer has $180°$ cycle. Thus, the azimuth angle of the surface normal will be either ϕ or $\phi + 180°$. 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 $\phi + 180°$, which is defined on a certain xy-plane and is defined as an angle between x-axis and the reflection plane projected on xy-plane. Since we capture images with a camera, the x-axis and the y-axis of the image coordinates is used.

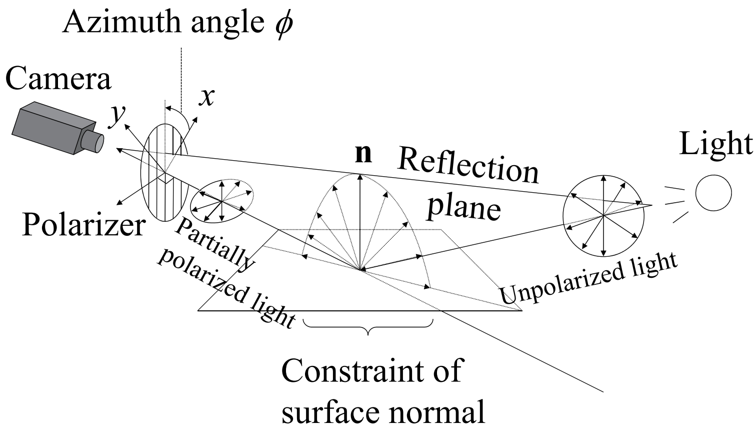

Fig 5 Relationship between the surface normal and the reflection plane when observed from a single viewpoint.

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.

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

We analyze the two reflection plane angles at the same surface point, corresponding to the known 3D geometry. Our method assumes that the 3D 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

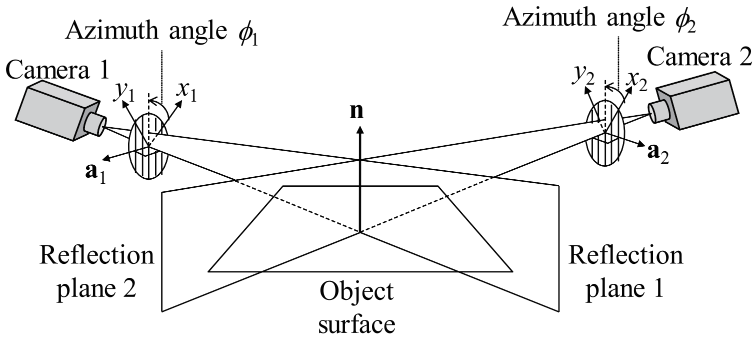

Fig 6 Relationship between the surface normal and the reflection plane when observed from two viewpoints.

ϕ_2, and the normal vector of the reflection plane, represented as $\mathbf{a}_1$ and $\mathbf{a}_2$, is shown in Eq. (1).

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

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 can 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

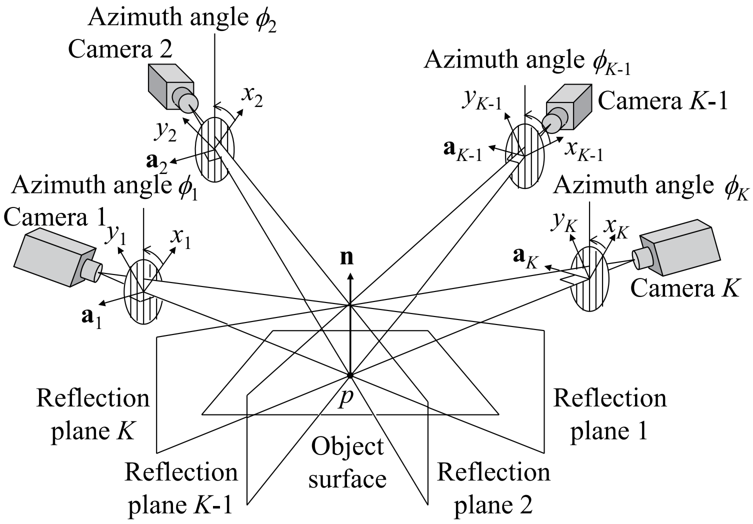

Fig 7 Relationship between the surface normal and the azimuth angle observed from multiple viewpoints.

be expressed, in orther 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 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), Eq. (3) or Eq. (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 other form,

$$\mathbf{A}\mathbf{n} = \mathbf{0}\,. \tag{4}$$

The surface normal $\mathbf{n}$, which satisfies Eq. (4) in the least-squares sense, can be estimated using SVD (singular value decomposition).[28] The $K\times3$ matrix $\mathbf{A}$ can be decomposed by SVD as follows.

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

Here, $\mathbf{U}$ is a $K\times3$ 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 the matrix $\mathbf{W}$ is the singular value of the 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 the matrix $\mathbf{A}$ is at most 2; thus, one of the three singular values becomes 0. Please see Miyazaki[27] for the proof. The surface normal $\mathbf{n}$ can be represented as Eq. (6),[28] which can be calculated from the singular vector that has the smallest singular value, namely, the third row of $\mathbf{V}^\top$ in Eq. (5).

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

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 be 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 the matrix $\mathbf{A}$ is 2.

2.5 Camera parameters

Eq. (3) or Eq. (4) calculates the surface normal from the azimuth angle under multiple viewpoints. In order 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 Section 2.6. Eq. (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 3D vertex (X, Y, Z) to 2D 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 skip to describe the camera center parameter (C_x, C_y) for clarity, since we assume pinhole camera model. We skip to 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 nor lack of information. Therefore, it is natural to use homography in our work since the target object is almost planar.

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') can be 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

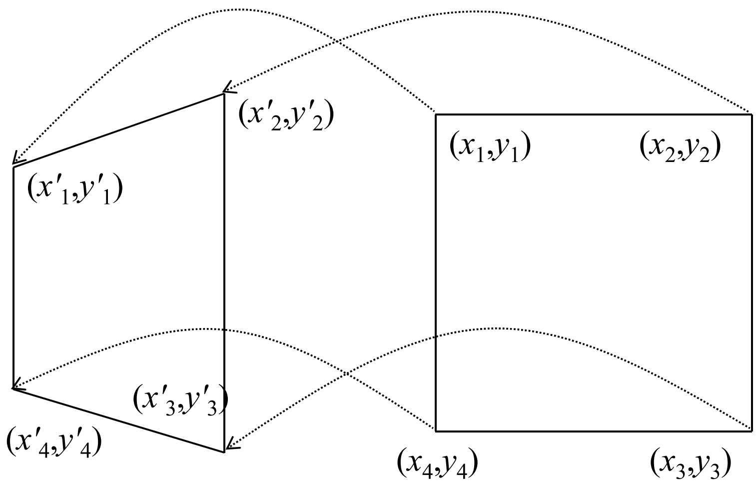

Fig 8 Homographic projection from a certain quadrangle to another certain quadrangle.

(x, y) is projected to the point $(x', y') = (\xi_1/\xi_3, \xi_2/\xi_3)$ by this homography matrix.

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

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

Fig. 8 is an example where 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 same transformation, thus, we fix one element as follows in order to uniquely determine the homography matrix.

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

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

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

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

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

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1{x'}_1 & -y_1{x'}_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1{y'}_1 & -y_1{y'}_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2{x'}_2 & -y_2{x'}_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2{y'}_2 & -y_2{y'}_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3{x'}_3 & -y_3{x'}_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3{y'}_3 & -y_3{y'}_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4{x'}_4 & -y_4{x'}_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4{y'}_4 & -y_4{y'}_4 \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{pmatrix} = \begin{pmatrix} {x'}_1 \\ {y'}_1 \\ {x'}_2 \\ {y'}_2 \\ {x'}_3 \\ {y'}_3 \\ {x'}_4 \\ {y'}_4 \end{pmatrix} . \tag{15}$$

Since we have 8 unknowns $(h_{11}, h_{12}, \ldots, h_{32})$ and 8 equations (8 rows of the leftmost matrix in

Eq. (15)), closed-form solution exists. Solving this results in 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) can be expressed by Eq. (9). Suppose that the homography of camera 1 is $\mathbf{H}_1$ and that of camera 2 is $\mathbf{H}_2$. Fig. 3 shows the homographic projection from each view to canonical square. Canonical square can be 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) can be 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) are corresponded while the point (x, y) acted as a mediator.

3 Experiment

3.1 Experimental setup

As is 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 photometric stereo, which needs to move the light

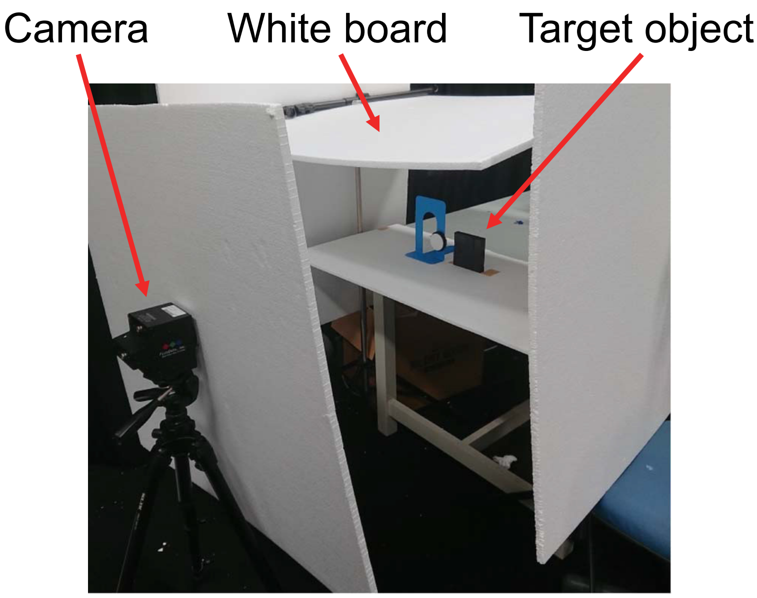

Fig 9 Experimental environment.

Table 1 Specification of the camera.

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

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

The camera we used is shown in Fig. 10 and Table 1. Since we have only one camera (because polarization camera is expensive), we rotated 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 same.

Fig 10 Polarization camera.

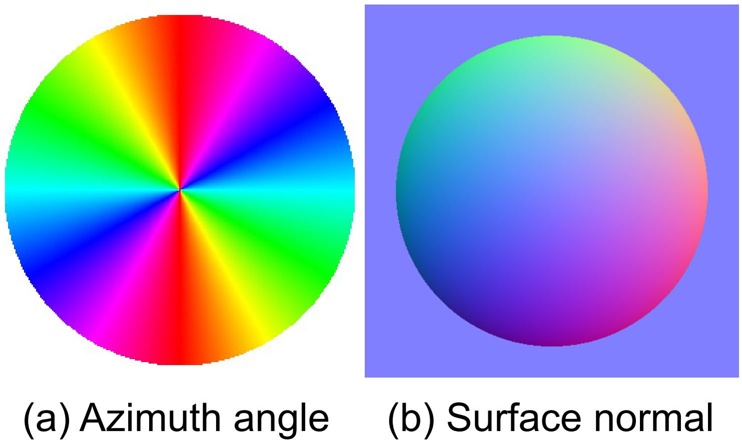

Fig 11 Pseudo color representation of an ideal sphere: (a) Azimuth angle and (b) surface normal.

3.2 Pseudo color representation of the result

Following sections show some results of our method. For visualization, the azimuth angle and the surface normal are represented by pseudo-color. Fig. 11 (a) and Fig. 11 (b) show the pseudo color representation of the azimuth angle and the surface normal of ideal hemishpere, respectively.

3.3 *Result of ellipsoid*

The target object is shown in Fig. 12. We generated the object using 3D printer, so that 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 [cm] $\times$ 10 [cm], the diameter of long axis of the ellipse is 7.5 [cm], the diameter of 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 a shape of cracks. First of all, we evaluate the performance of the proposed method. In order to guarantee the statistical reliability, we need to estimate the surface normal with wide variety and wide area. That is why we first measure the concave ellipsoid.

We took one image each from 15 different direction (Fig. 13). Pseudo-color representation of surface normal of our method is shown in Fig. 14, and that of ground truth is shown in Fig. 15. Note that our method successfully estimated the shape which is almost the same as true shape. The estimated shape is shown in Fig. 16 and Fig. 18 (c), while ground truth is shown in Fig. 17 and Fig. 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), where the average error was 4.49 [deg].

3.4 *Comparison to photometric stereo*

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

Photometric stereo from 15 lights is applied to the object shown in Fig. 12, and the input images are shown in Fig. 20. Surface normal of the photometric stereo is shown in Fig. 21, and the estimated shape is shown in Fig. 22 and Fig. 18 (b). Photometric stereo assumes Lambertian reflection though the actual reflection is specular reflection, thus, the shape is distorted. The error

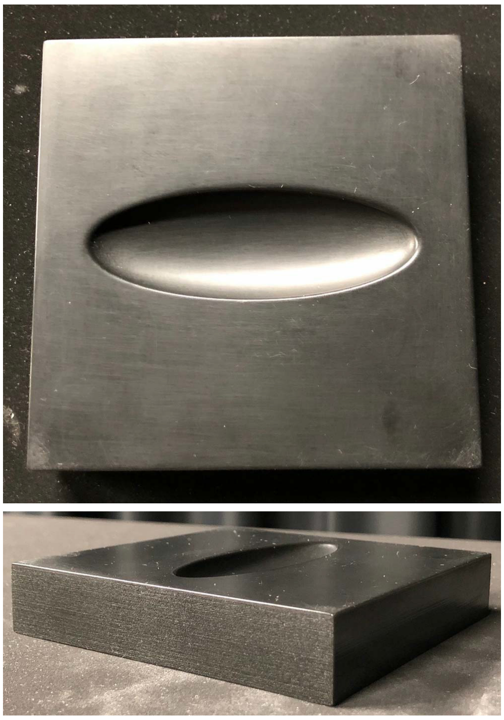

Fig 12 Target object [ellipsoid].

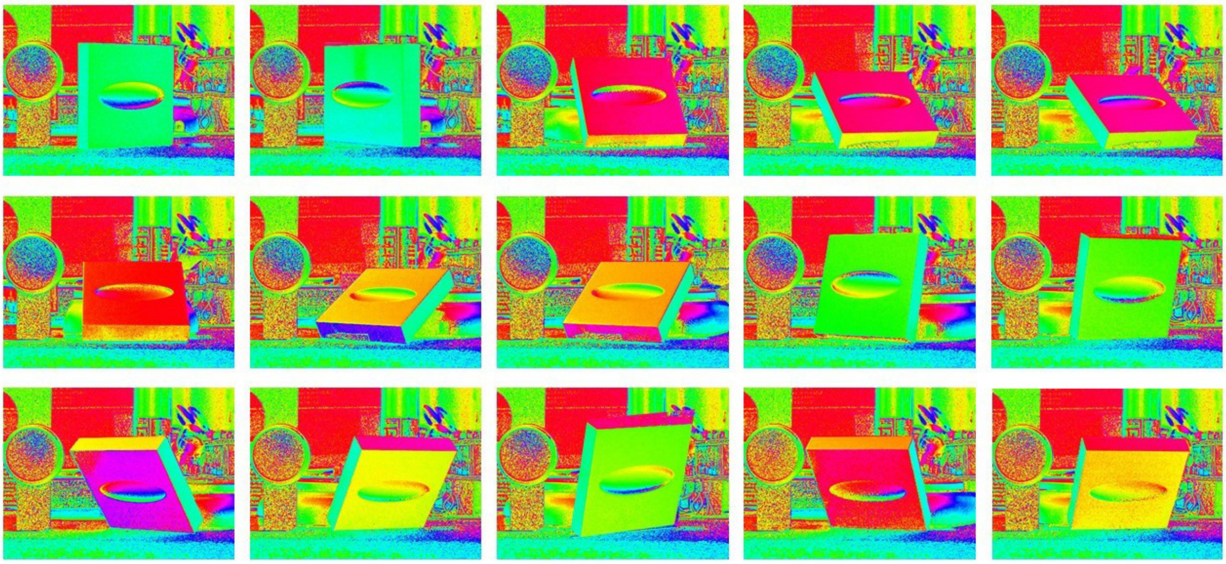

Fig 13 Input image [ellipsoid].

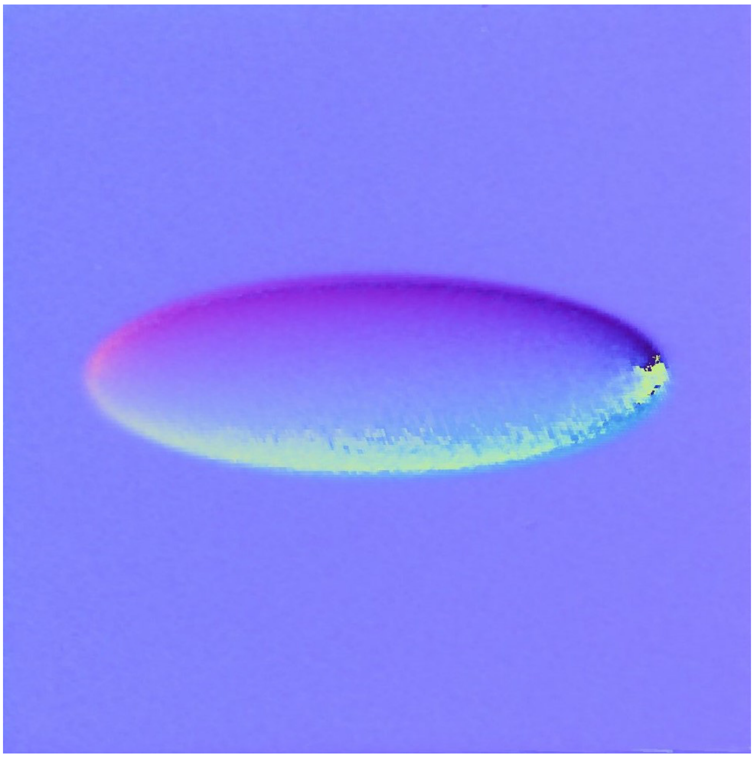

Fig 14 Estimated surface normal [ellipsoid].

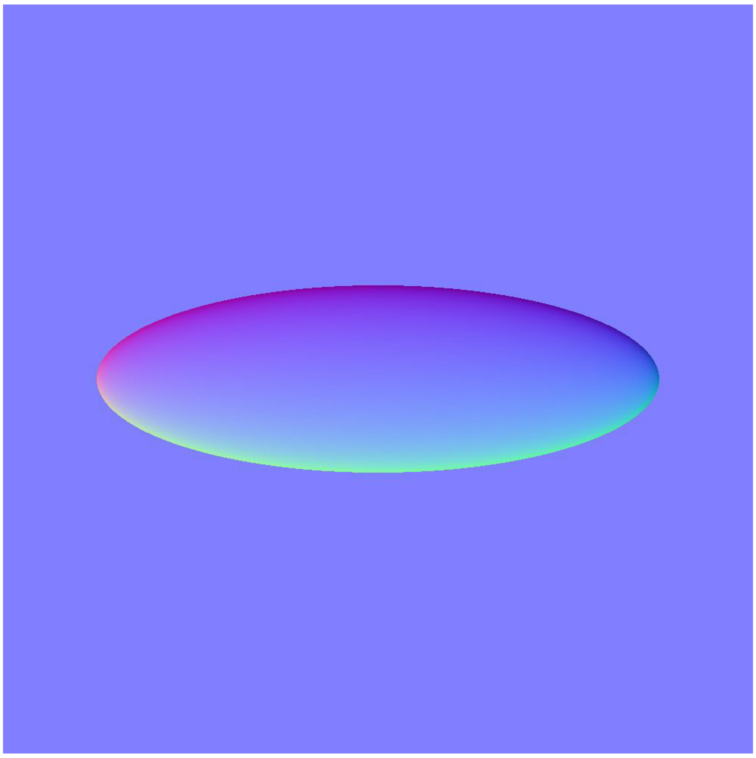

Fig 15 Ground truth of surface normal [ellipsoid].

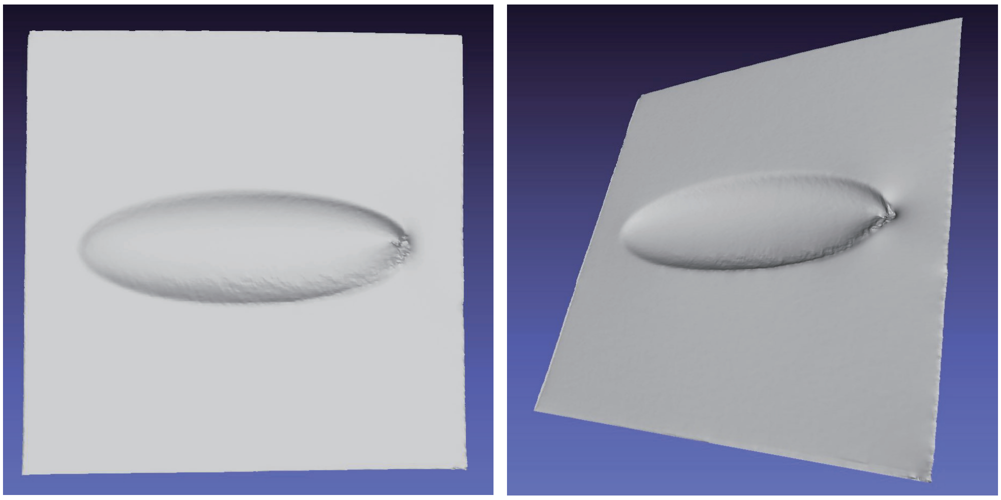

Fig 16 Estimated shape [ellipsoid].

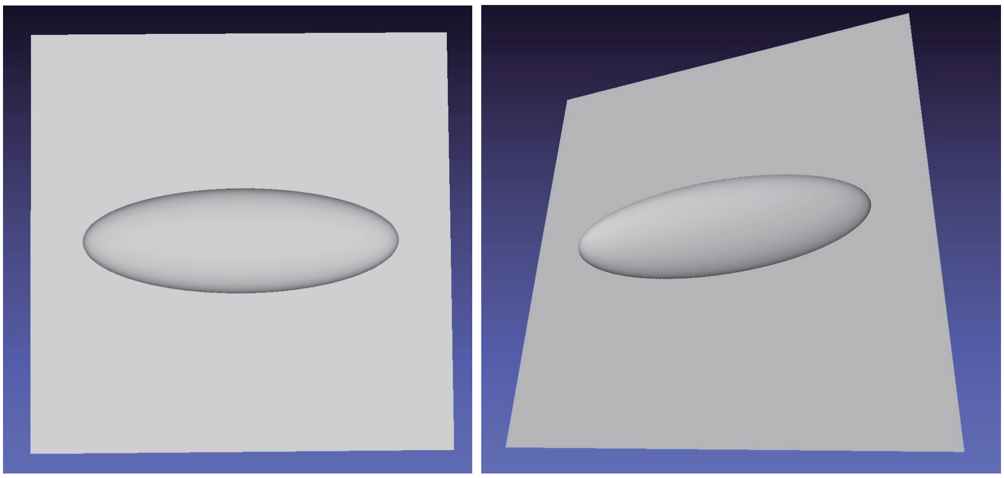

Fig 17 Ground truth of shape [ellipsoid].

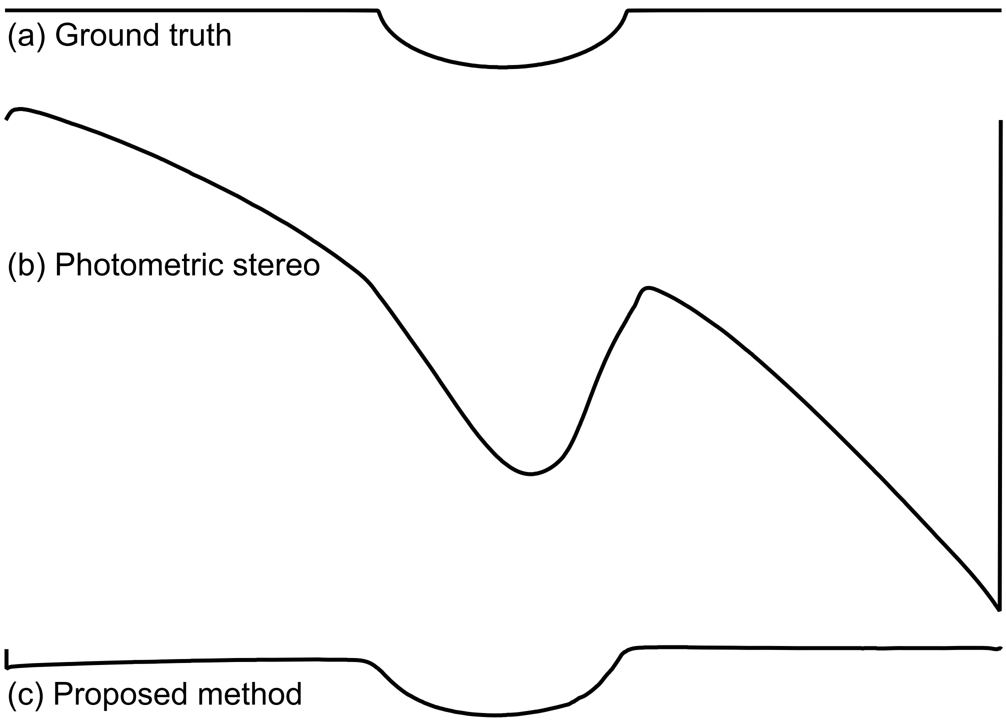

Fig 18 Intersection shape [ellipsoid]: (a) Ground truth, (b) photometric stereo, and (c) proposed method.

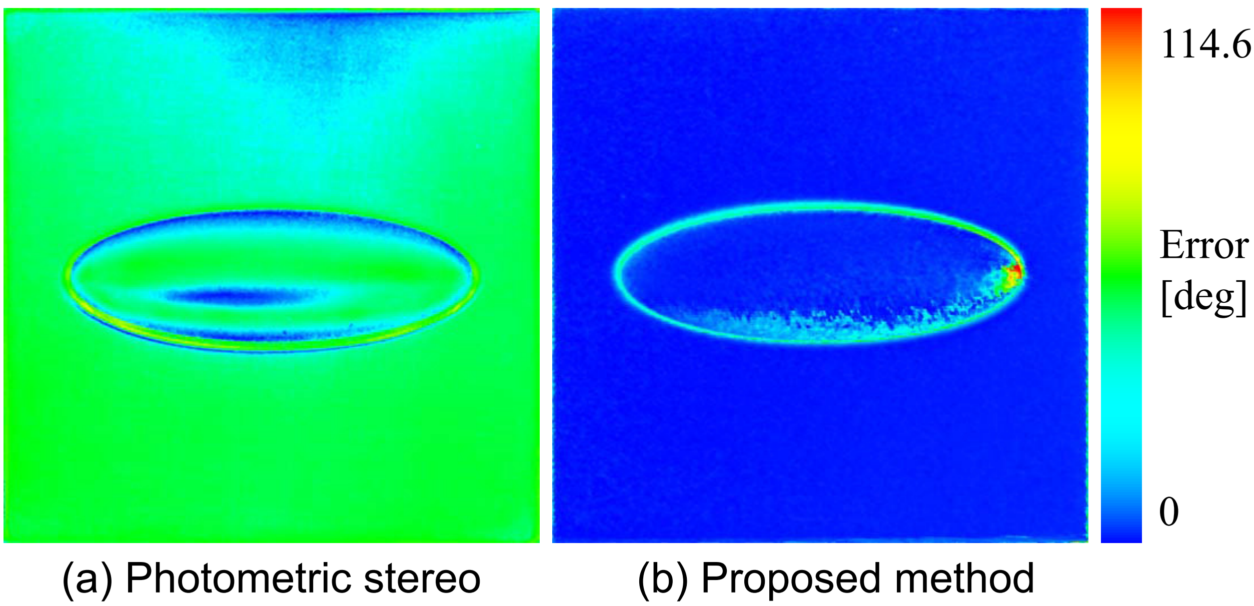

Fig 19 Estimation error [ellipsoid]: (a) Photometric stereo and (b) proposed method.

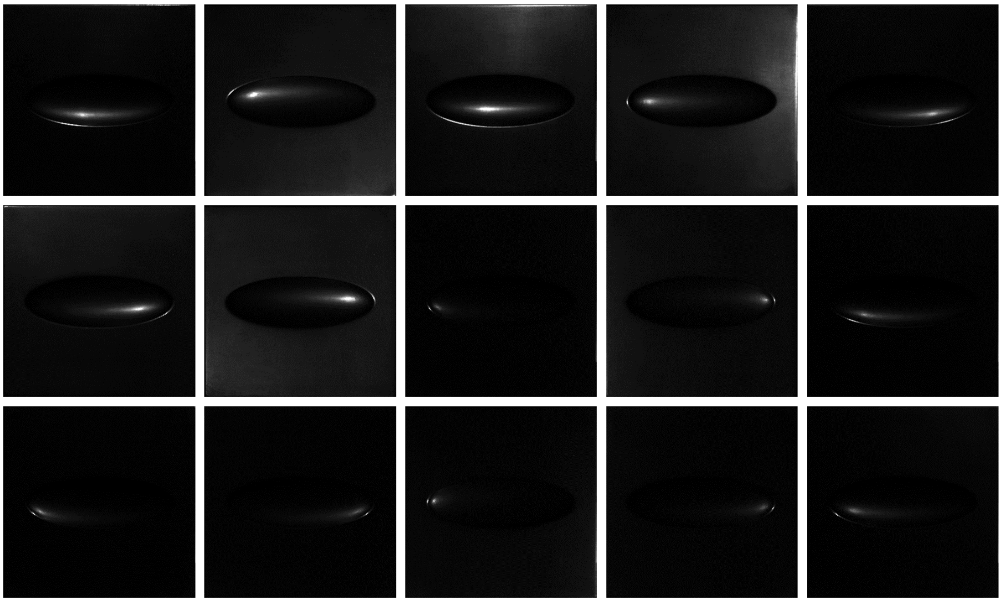

Fig 20 Input image of photometric stereo [ellipsoid].

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.

3.5 Result of convex object

Our method can not only be applied to concave objects but also be applied to convex objects. In order 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 is shown in Fig. 24, while those of photometric stereo is shown in Fig. 25. The surface normal of the ground truth, the photometric stereo, and the proposed method is shown in Fig. 26, Fig. 27, and Fig. 28. The shape of the ground truth, the photometric stereo, and the proposed method is shown in Fig. 29, Fig. 30, and Fig. 31. The cross section of the shape of the ground truth, the photometric stereo, and the proposed method is shown in Fig. 32 (a), Fig. 32 (b), and Fig. 32 (c). The error of the photometric stereo and the proposed method is shown in Fig. 33 (a) and Fig. 33 (b). The average error of the photometric stereo was 47.0 [deg] while that of the proposed method was 12.9 [deg].

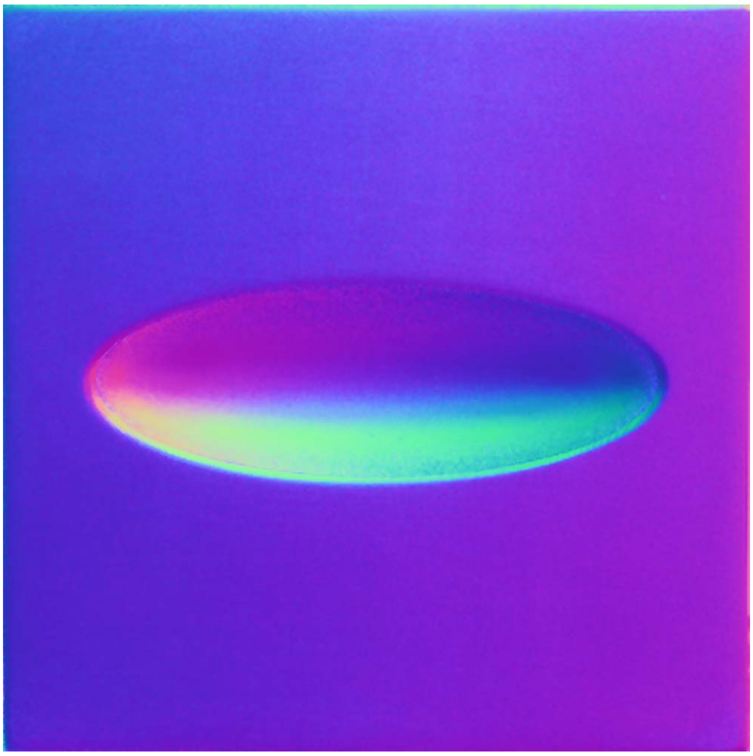

Fig 21 Surface normal of photometric stereo [ellipsoid].

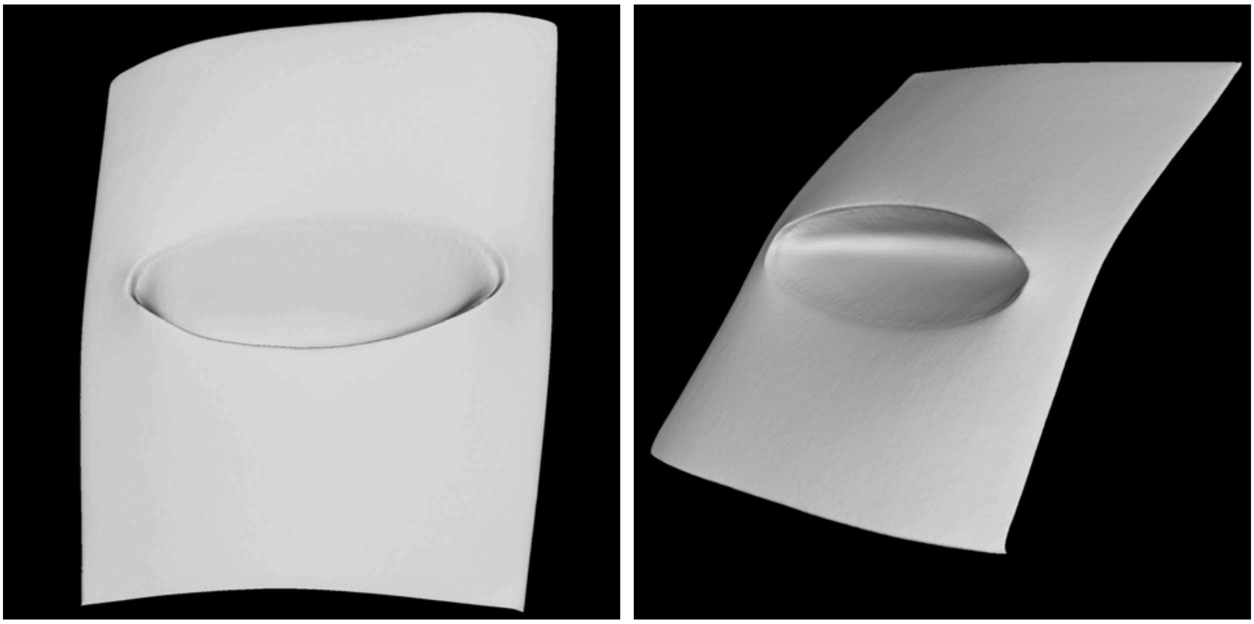

Fig 22 Shape of photometric stereo [ellipsoid].

Fig 23 Photograph of target object [convex].

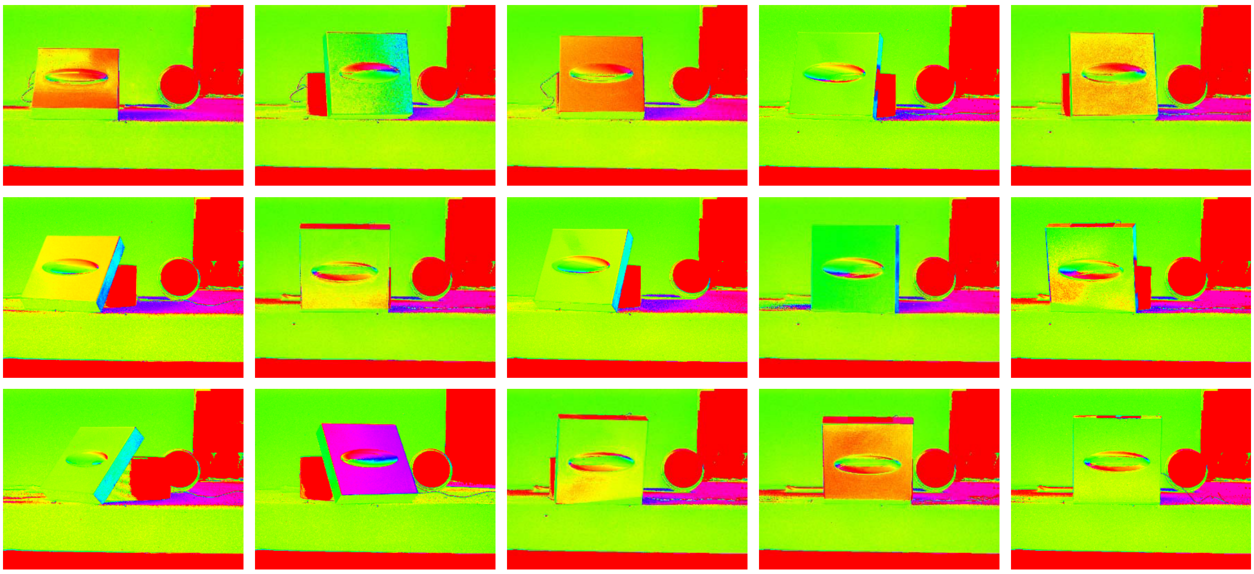

Fig 24 Input data of our method [convex].

Fig 25 Input images of photometric stereo [convex].

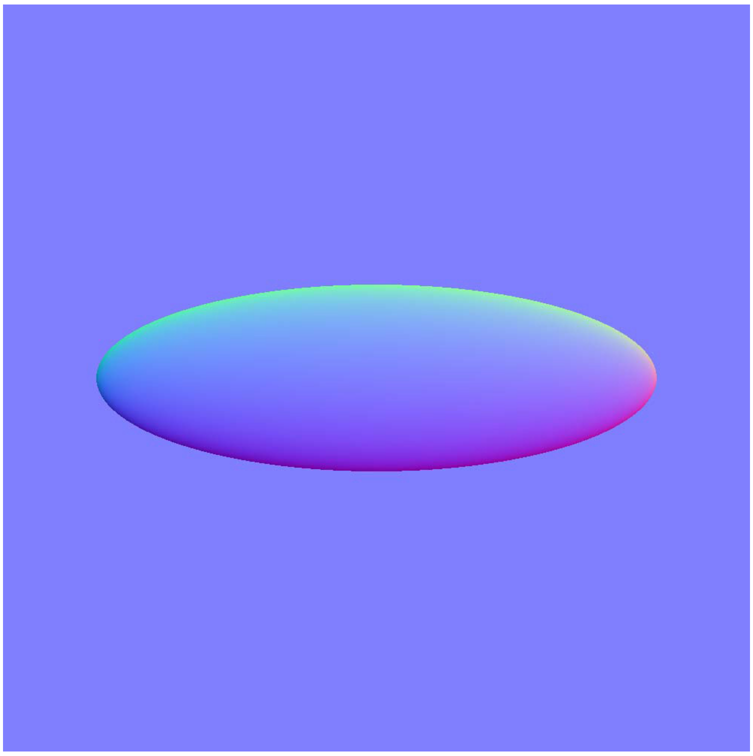

Fig 26 Surface normal of ground truth [convex].

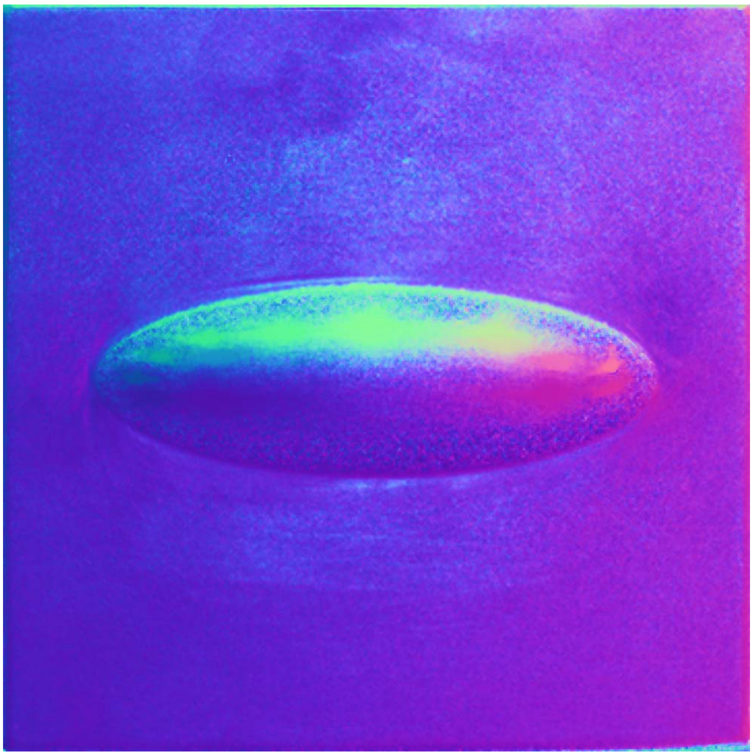

Fig 27 Surface normal of photometric stereo [convex].

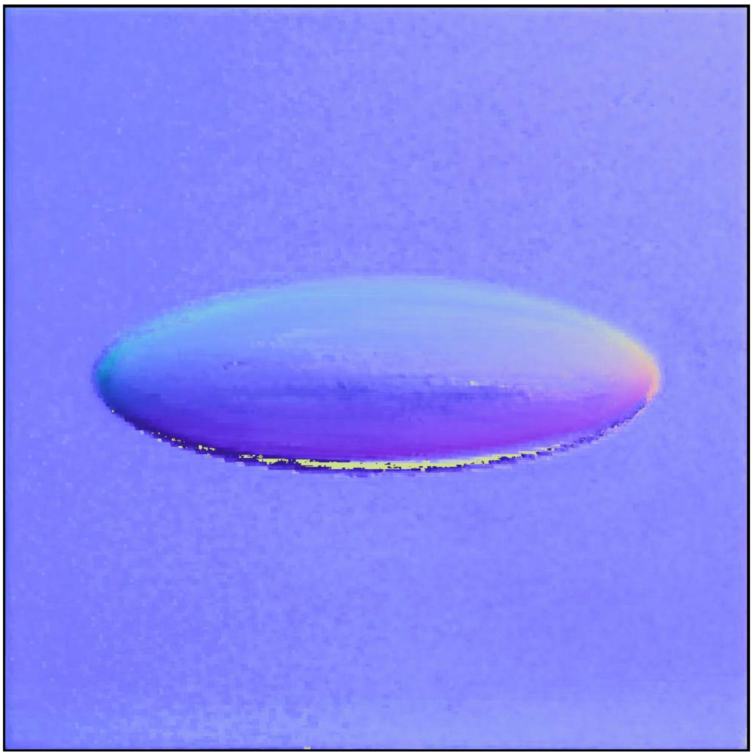

Fig 28 Surface normal of proposed method [convex].

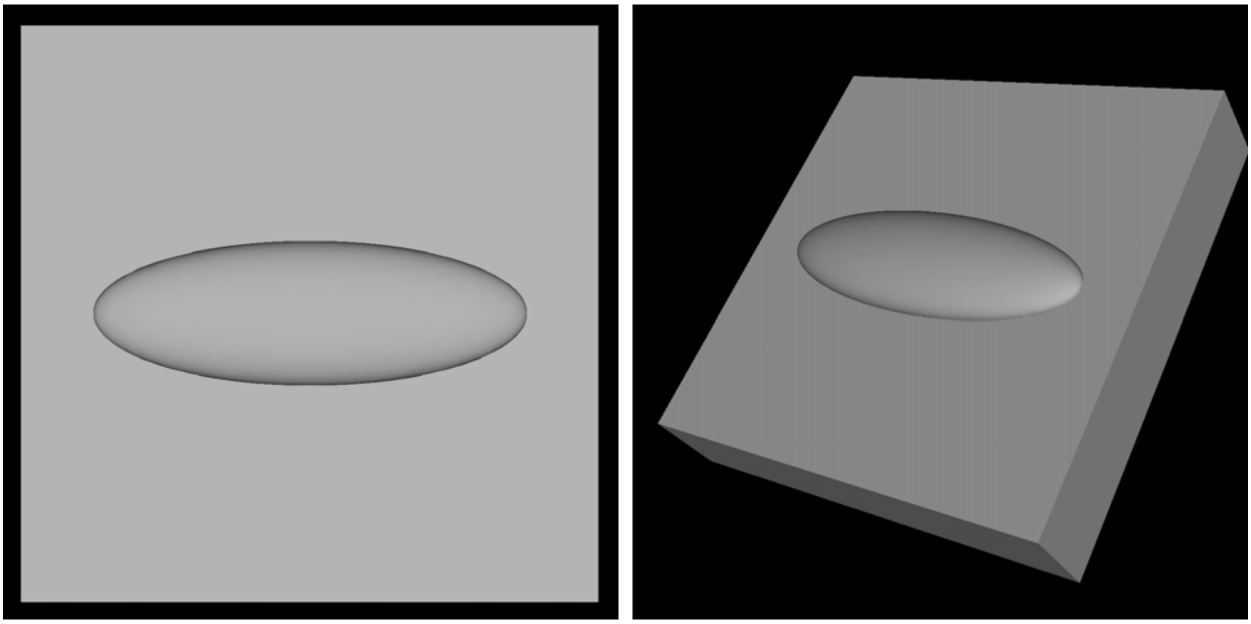

Fig 29 Shape of ground truth [convex].

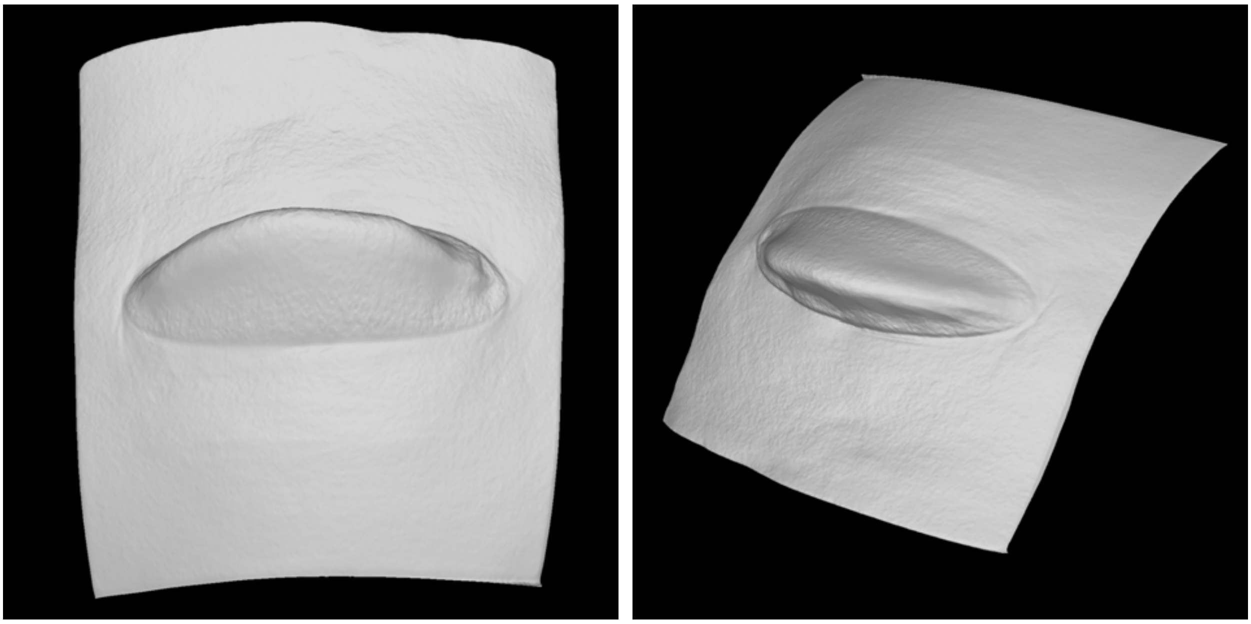

Fig 30 Shape of photometric stereo [convex].

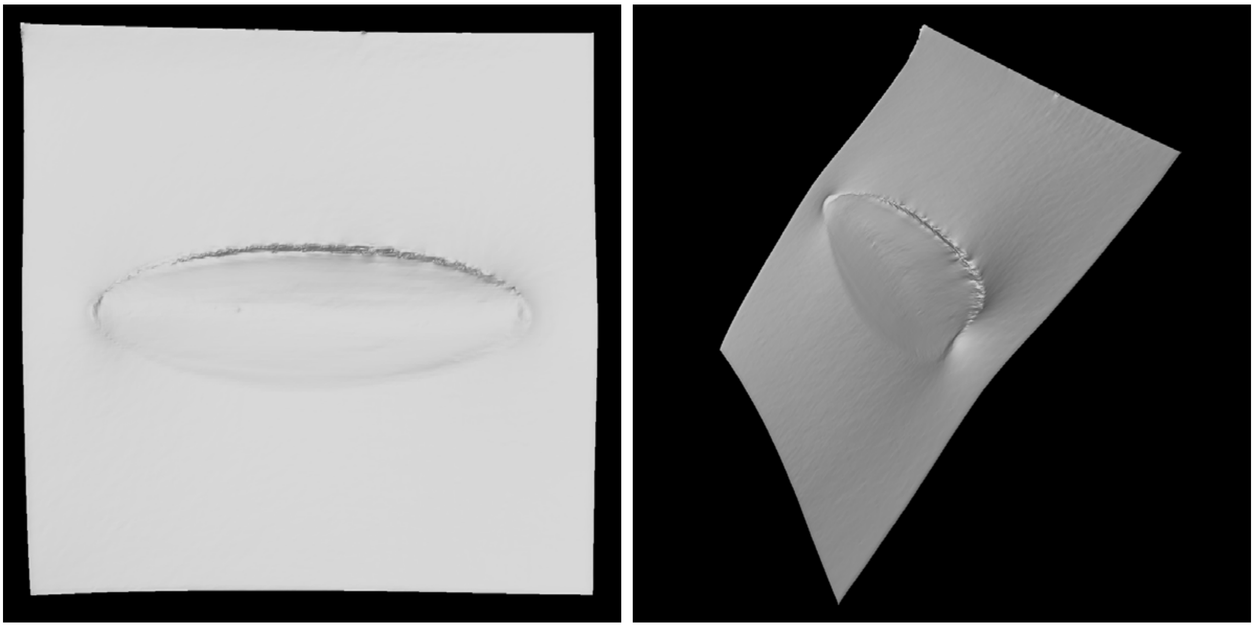

Fig 31 Shape of proposed method [convex].

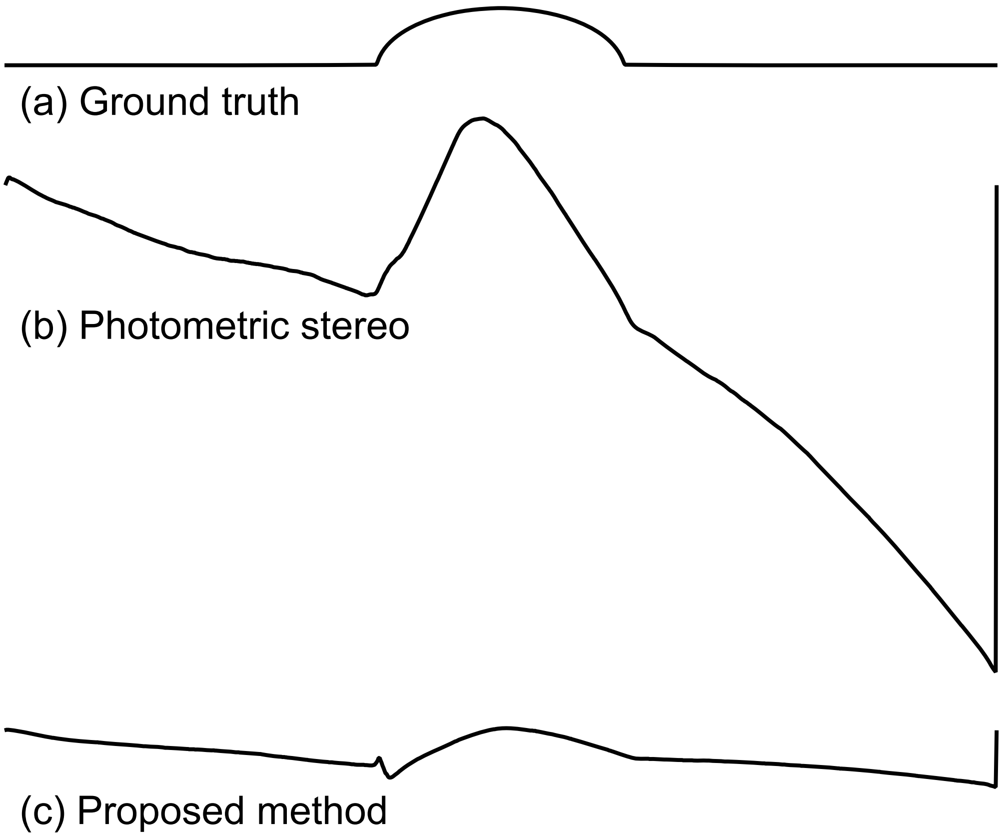

Fig 32 Intersection shape [convex]: (a) Ground truth, (b) photometric stereo, and (c) proposed method.

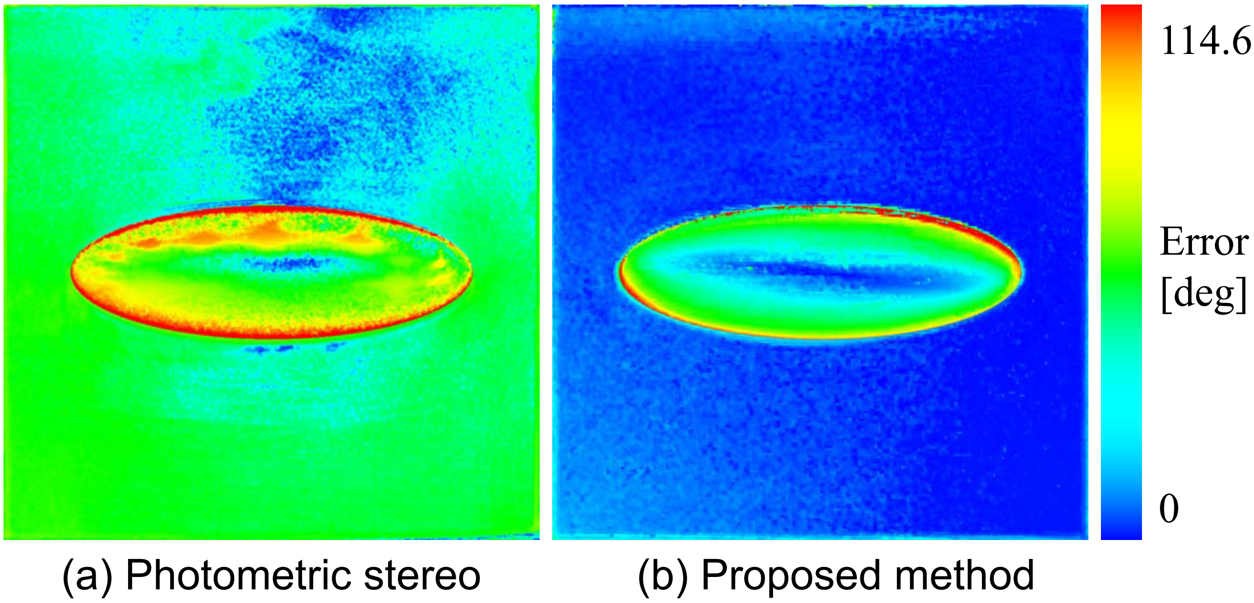

Fig 33 Estimation error [convex]: (a) Photometric stereo and (b) proposed method.

3.6 Result of stripes

In order to evaluate the performance of our method depending on the width of cracks, three different concave shapes with different width are measured. Fig. 34 shows the target object and Fig. 35 shows the input images. Also, 15 images are taken, one for each direction. 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 Fig. 38 and Fig. 40 (b), while ground truth is shown in Fig. 39 and Fig. 40 (a). The error map is shown in Fig. 41, and the average error was 7.18 [deg].

3.7 Result of worm

In order to simulate an acutal situation, we applied our method to a cracks which is not shaped in a straight line. Fig. 42 shows the target object, Fig. 43 shows the input images, Fig. 44 shows the estimated surface normal, and Fig. 45 shows the estimated shape.

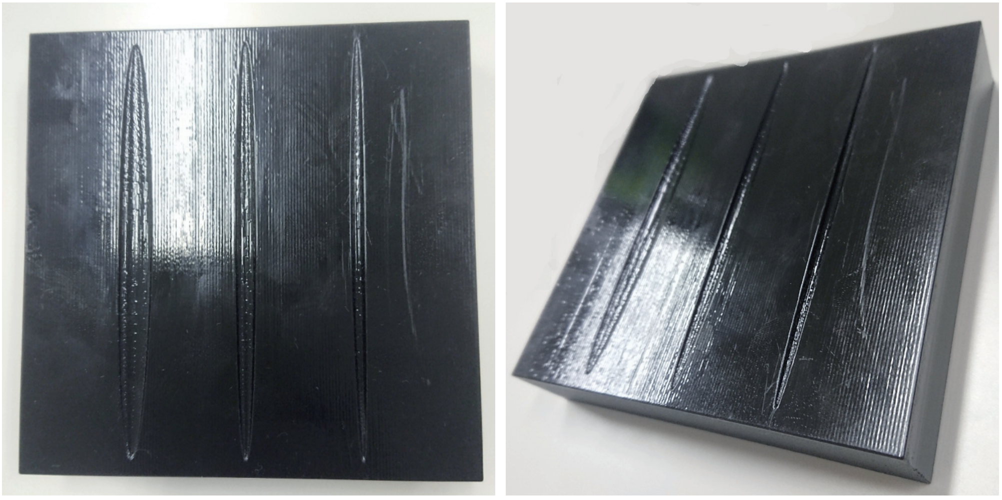

Fig 34 Target object [stripe].

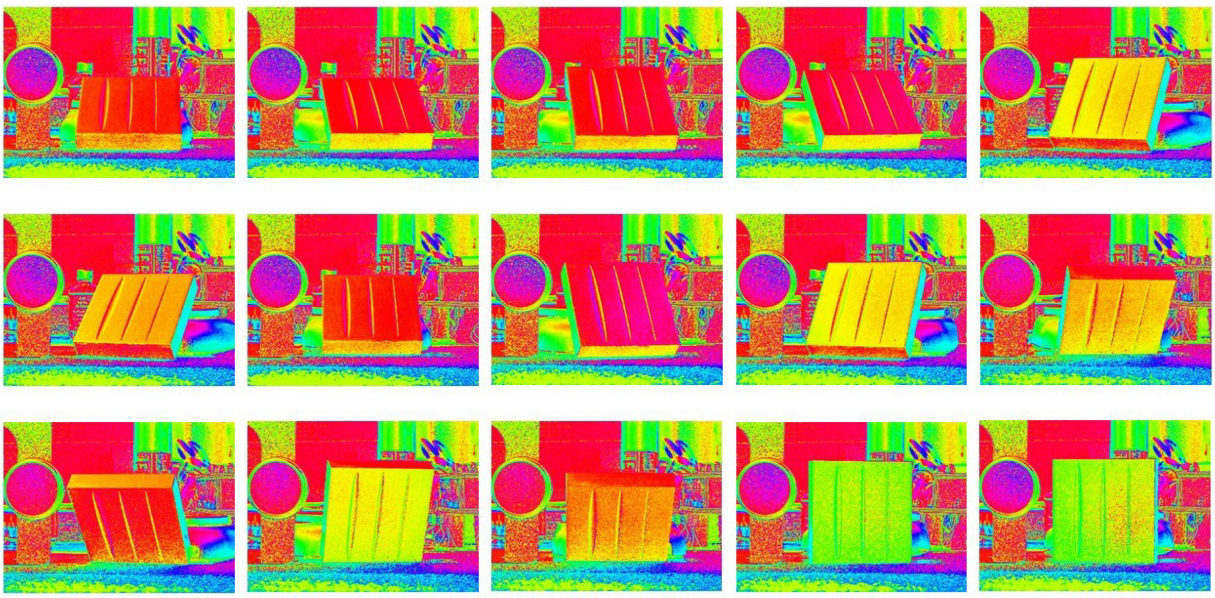

Fig 35 Input images [stripe].

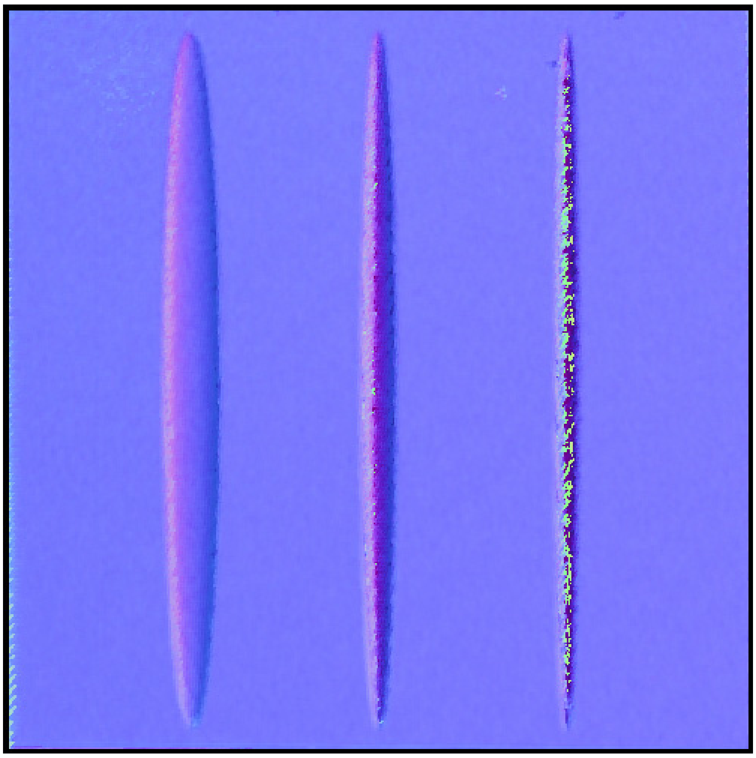

Fig 36 Estimated surface normal [stripe].

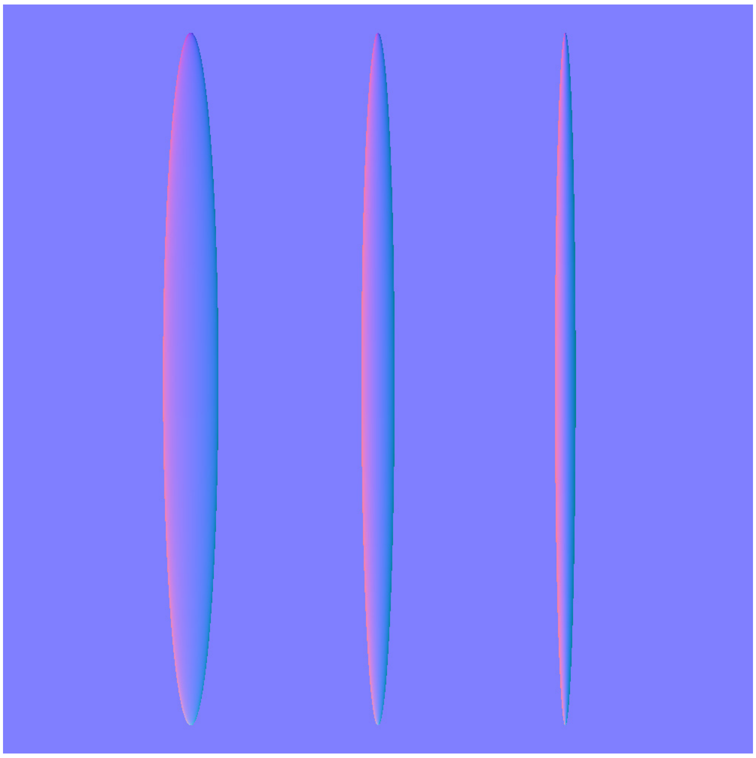

Fig 37 Ground truth of surface normal [stripe].

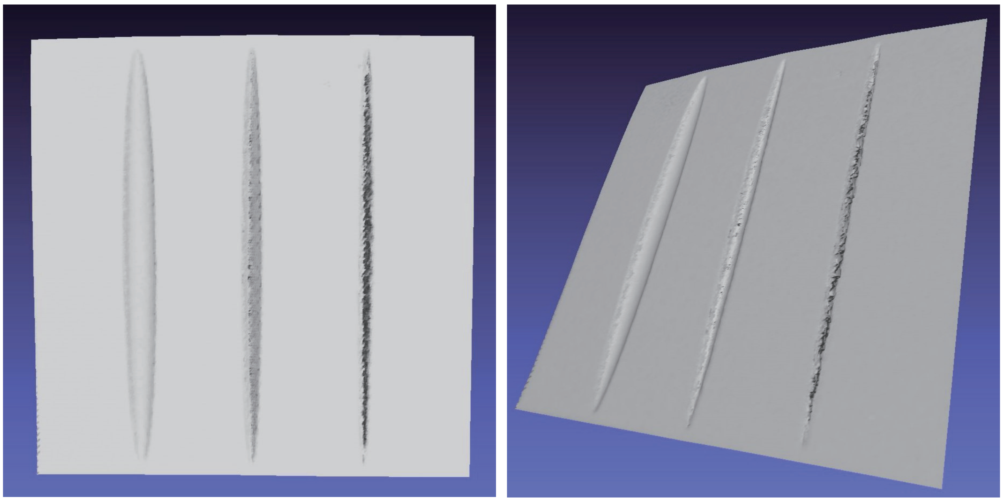

Fig 38 Estimated shape [stripe].

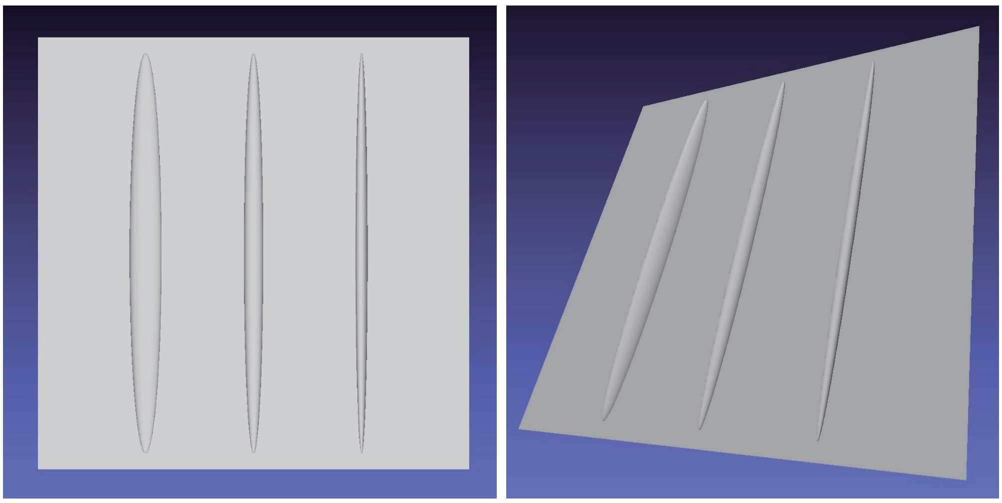

Fig 39 Ground truth of shape [stripe].

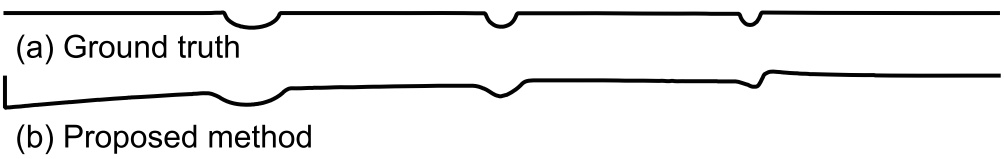

Fig 40 Intersection shape [stripe]: (a) Ground truth and (b) estimated shape.

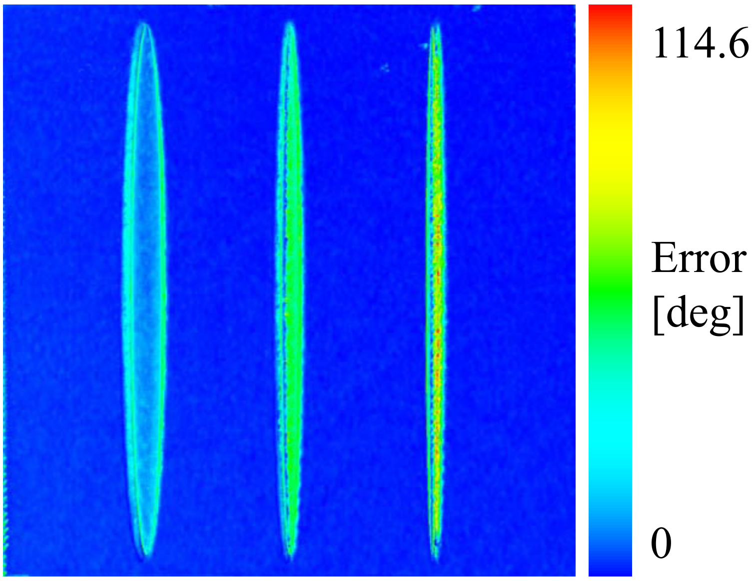

Fig 41 Error of our method [stripe].

Fig 42 Target object [worm].

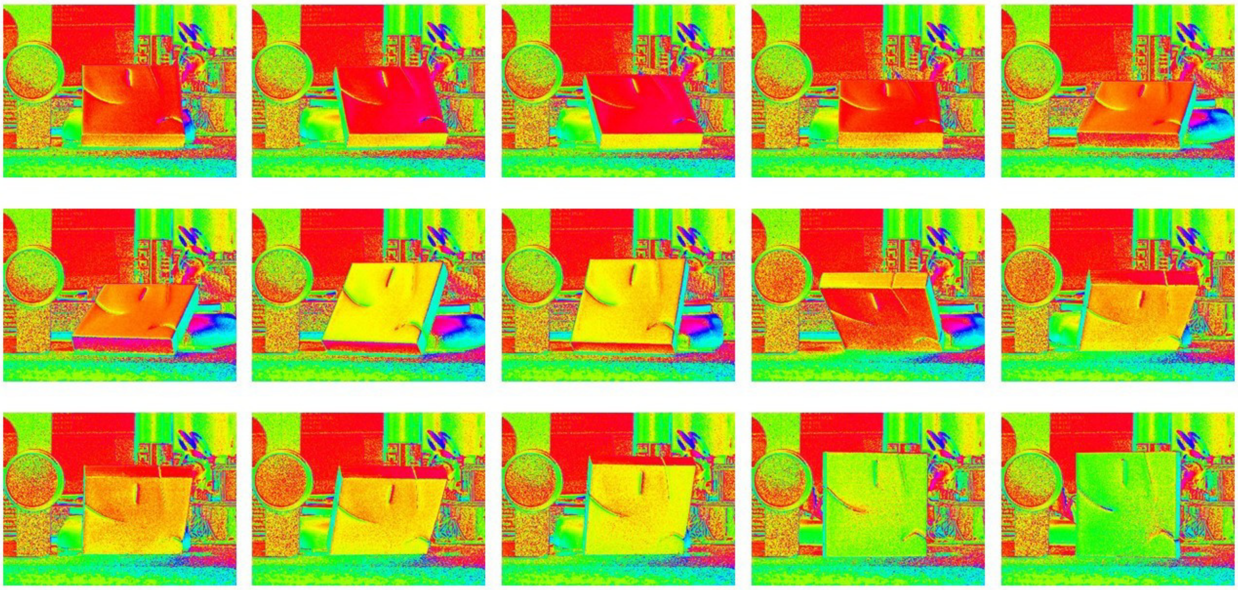

Fig 43 Input images [worm].

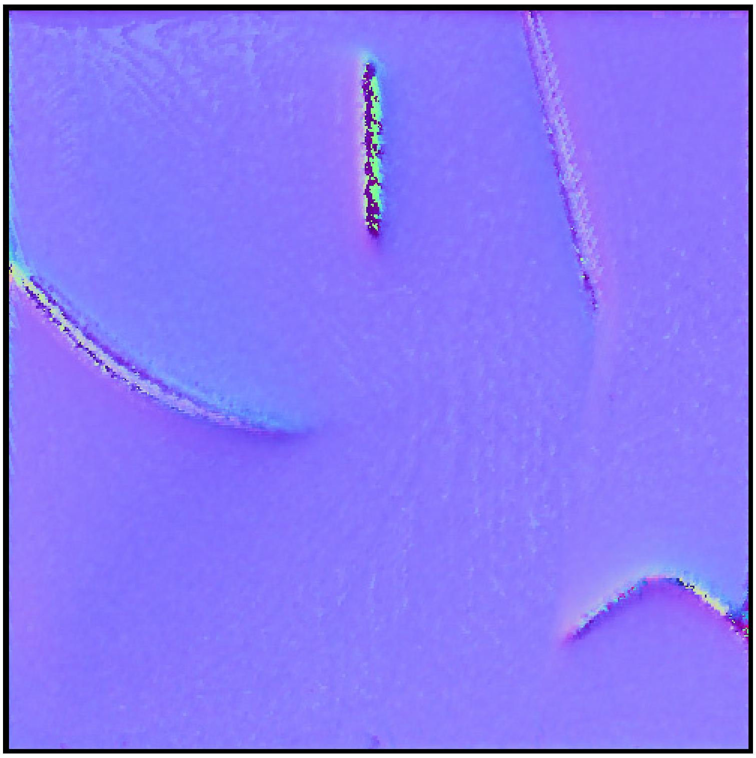

Fig 44 Estimated surface normal [worm].

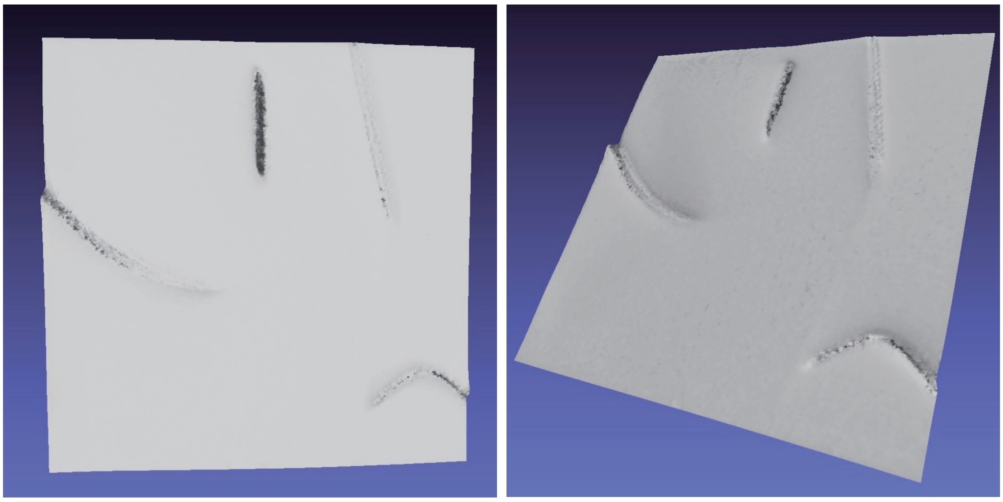

Fig 45 Estimated shape [worm].

3.8 Discussion

As is shown in Fig. 41, the narrower the concave part is, the worse the result is. This is because the light is not illuminated satisfactorilly to the narrow concave part. In addition, interreflection becomes strong at narrow concave part.

4 Conclusion

We propose a shape estimation method from polarization images obtained from multiple view-points. The proposed method computes the surface normal using SVD to minimize the least-squared error. It can estimate the shapes of concave part of planar objects which is black and has high specularity. It is usually difficult to estimate the shape of planar object with small details, however, our algorithm fully utilizes the property that the target object is almost planar. 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 investigation of product inspection in factory, damage inspection in architecture, age estimation from skin wrinkle, and so on. For example, factories want to know the reason of the defect of the product since they want to fix the problem and decrease the defects. In order to analyze the reason, the shape of the defects is necessary, and our method is useful for this purpose.

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

Acknowledgments

This research was supported in part by MEXT KAKENHI Grant Number 15H05925, in part by MEXT KAKENHI Grant Number 15H02727, and in part by JSPS KAKENHI Grant Number 24700176.

References

1 A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, **57**, 787–798 (2018). [doi:10.1016/j.aej.2017.01.020].

2 I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich, "Transparent and specular object reconstruction," *Computer Graphics Forum*, **29**(8), 2400–2426 (2010). [doi:10.1111/j.1467-8659.2010.01753.x].

3 M. Born and E. Wolf, *Principles of optics*, Pergamon Press, London (1959).

4 E. Hecht, *Optics*, Addison-Wesley, Reading, Mass. (2002).

5 W. A. Shurcliff, *Polarized light: production and use*, Harvard University Press, Cambridge, Mass (1962).

6 K. Koshikawa and Y. Shirai, "A model-based recognition of glossy objects using their polarimetrical properties," *Advanced Robotics*, **2**(2), 137–147 (1987). [doi:10.1163/156855387X00129].

7 G. C. Guarnera, P. Peers, P. Debevec, and A. Ghosh, "Estimating surface normals from spherical Stokes reflectance fields," in *Computer Vision - ECCV2012. Workshops and Demonstrations*, 340–349 (2012). [doi:10.1007/978-3-642-33868-7_34].

8 O. Morel, C. Stolz, F. Meriaudeau, and P. Gorria, "Active lighting applied to three-dimensional reconstruction of specular metallic surface by polarization imaging," *Applied Optics*, **45**(17), 4062–4068 (2006). [doi:10.1364/AO.45.004062].

9 M. Saito, Y. Sato, K. Ikeuchi, and H. Kashiwagi, "Measurement of surface orientations of transparent objects by use of polarization in highlight," *Journal of Optical Society of America A*, **16**(9), 2286–2293 (1999). [doi:10.1364/JOSAA.16.002286].

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, T. Morimoto, I. Sato, Y. Mukaigawa, T. Tomono, and K. Ikeuchi, "Reconstructing shapes and appearances of thin film objects using RGB images," in *IEEE Conference on Computer Vision and Pattern Recognition*, 3774–3782 (2016). [doi:10.1109/CVPR.2016.410].

12 D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi, “Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths,” *Journal of Optical Society of America A*, **19**(4), 687–694 (2002). [doi:10.1364/JOSAA.19.000687].

13 D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi, “Polarization-based inverse rendering from a single view,” in *Proceedings of IEEE International Conference on Computer Vision*, 982–987 (2003). [doi:10.1109/ICCV.2003.1238455].

14 D. Miyazaki, M. Kagesawa, and K. Ikeuchi, “Transparent surface modeling from a pair of polarization images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**(1), 73–82 (2004). [doi:10.1109/TPAMI.2004.1261080].

15 D. Miyazaki and K. Ikeuchi, “Shape estimation of transparent objects by using inverse polarization raytracing,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(11), 2018–2030 (2007). [doi:10.1109/TPAMI.2007.1117].

16 Y. Yu, D. Zhu, and W. A. Smith, “Shape-from-polarisation: a nonlinear least squares approach,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2969–2976 (2017). [doi:10.1109/ICCVW.2017.350].

17 L. B. Wolff and T. E. Boult, “Constraining object features using a polarization reflectance model,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**(7), 635–657 (1991). [doi:10.1109/34.85655].

18 S. Rahmann and N. Canterakis, “Reconstruction of specular surfaces using polarization imaging,” in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 149–155 (2001). [doi:10.1109/CVPR.2001.990468].

19 S. Rahmann, “Reconstruction of quadrics from two polarization views,” in *Proceed-*

ings of Iberian Conference on Pattern Recognition and Image Analysis, 810–820 (2003). [doi:10.1007/978-3-540-44871-6_94].

20 G. A. Atkinson and E. R. Hancock, "Shape estimation using polarization and shading from two views," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(11), 2001–2017 (2007). [doi:10.1109/TPAMI.2007.1099].

21 G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," *IEEE Transactions on Image Processing*, **15**(6), 1653–1664 (2006). [doi:10.1109/TIP.2006.871114].

22 C. P. Huynh, A. Robles-Kelly, and E. R. Hancock, "Shape and refractive index from single-view spectro-polarimetric images," *International Journal of Computer Vision*, **101**, 64–94 (2013). [doi:10.1007/s11263-012-0546-3].

23 A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar, "Polarized 3D: High-quality depth sensing with polarization cues," in *Proceedings of IEEE International Conference on Computer Vision*, 3370–3378 (2015). [doi:10.1109/ICCV.2015.385].

24 W. Smith, R. Ramamoorthi, and S. Tozza, "Height-from-polarisation with unknown lighting or albedo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**(12), 2875–2888 (2018). [doi:10.1109/TPAMI.2018.2868065].

25 Z. Cui, J. Gu, B. Shi, P. Tan, and J. Kautz, "Polarimetric multi-view stereo," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1558–1567 (2017). [doi:10.1109/CVPR.2017.47].

26 L. Yang, F. Tan, A. Li, Z. Cui, Y. Furukawa, and P. Tan, "Polarimetric dense monocular

slam," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3857–3866 (2018). [doi:10.1109/CVPR.2018.00406].

27 D. Miyazaki, T. Shigetomi, M. Baba, R. Furukawa, S. Hiura, N. Asada, "Surface normal estimation of black specular objects from multiview polarization images," *SPIE Optical Engineering*, **56**(4), 041303-1–041303-17 (2016). [doi:10.1117/1.OE.56.4.041303].

28 W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C: the art of scientific computing*, Cambridge University Press, Cambridge UK (1997).

29 R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, Cambridge UK (2004).

30 R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, **19**(1), 139–144 (1980). [doi:10.1117/12.7972479].

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 degree in information science and technology in 2005. His research interests include physics-based vision. He is a member of ACM and IEEE.

Biographies and photographs of the other authors are not available.

List of Figures

List of Tables