Median Photometric Stereo as Applied to the Segonko Tumulus and Museum Objects
Daisuke Miyazaki, Kenji Hara, Katsushi Ikeuchi,
International Journal of Computer Vision,
vol. 86, no. 2-3, pp. 229-242, 2010.01

http://www.cg.info.hiroshima-cu.ac.jp/~miyazaki/

Noname manuscript No.
(will be inserted by the editor)

Median Photometric Stereo as Applied to the Segonko Tumulus and Museum Objects

Daisuke Miyazaki · Kenji Hara · Katsushi Ikeuchi

Received: 2 February 2008 / Accepted: 5 May 2009

Abstract One of the necessary techniques for constructing a virtual museum is to estimate the surface normal and the albedo of the artwork which has high specularity. In this paper, we propose a novel photometric stereo method which is robust to the specular reflection of the object surface. Our method can also digitize the artwork arranged inside a glass or acrylic display case without bringing the artwork out of the display case. Our method treats the specular reflection at the object surface or at the display case as an outlier, and finds a good surface normal evading the influence of the outliers. We judiciously design the cost function so that the outlier will be automatically removed under the assumption that the object's shape and color are smooth. At the end of this paper, we also show some archived 3D data of Segonko Tumulus and objects in the University Museum at The University of Tokyo that were generated by applying the proposed method.

Keywords Photometric stereo · M-estimation

1 Introduction

It is a benefit to society to allow people to become familiar with precious artworks in museums by broadcasting these objects through the Internet or mobile phones. We can do this by digitally archiving these works. However, these precious objects are exhibited inside glass display cases and are not allowed to be removed from these cases. We propose a method to estimate the surface normal and the albedo of these objects without removing them from the display case.

D. Miyazaki · K. Ikeuchi
Institute of Industrial Science, The University of Tokyo,
4-6-1 Komaba, Meguro-ku, 153-8505 Tokyo, Japan
E-mail: miyazaki@hiroshima-cu.ac.jp
Present address: of D. Miyazaki
Graduate School of Information Sciences, Hiroshima City University,
3-4-1 Ozukahigashi, Asaminami-ku, 731-3194 Hiroshima, Japan

K. Hara
Faculty of Design, Kyushu University,
4-9-1 Shiobaru, Minato-ku, 815-8540 Fukuoka, Japan

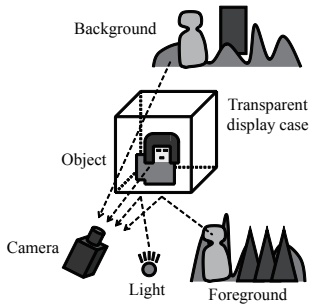


Fig. 1 Foreground and light reflected, and background and object transmitted through a transparent display case.

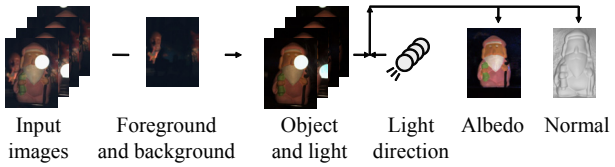


Fig. 2 Flow diagram. We first subtract the image under the light when it is off from the images under the light when it is on. We apply our median photometric stereo method to the subtracted images.

If we take a photograph of such objects in a well-lighted museum, we also observe the scene reflected by the display case. Many researchers have proposed methods to separate such reflections [14, 52, 59, 51, 47, 27, 28, 48, 41]. To add to the complication of separating the object from its reflection, if we illuminate the object by a lamp, the light is also reflected at the surface of the glass. In this paper, we partially remove the reflection at the glass surface, and estimate the surface normal and the albedo of the object by using photometric stereo. Our method assumes that the target object has both diffuse reflection and specular reflection.

The problem we are addressing in this paper is the situation when the object is set inside a transparent display case, as is shown in Fig. 1. The foreground, the scene in front of the object, is reflected at the transparent surface. The background, the scene behind the object, is transmitted through the transparent surface. The light illuminating the object also reflects at the transparent surface. The goal of this research is to estimate the surface normal of the object covered by a planar transparent surface.

The flow of the proposed algorithm is shown in Fig. 2. First, we take multiple images of the object under a different single light source. The image includes not only the object and the background but also the light and the foreground (Fig. 1). In this paper, we remove the foreground and the background by using a simpler method than previous methods. However, after this process, the output image still includes the reflection of the light as well as the object; thus, we cannot apply conventional photometric stereo. Therefore, we use so-called "four-light photometric stereo [10, 56, 4, 9]." Four-light photometric stereo can be applied to specular objects; however, it cannot be applied to specular objects kept inside a display case. Therefore, we use five or more lights. If we use four lights, we obtain four candidates for surface normal. Four-light photometric stereo, originally proposed by Coleman and Jain [10], investigated the rule for estimating the surface normal using these four candidates. Our

algorithm, called “median photometric stereo,” provides a framework for determining the surface normal when we have more than four candidates.

Some photometric stereo methods use multiple images to enhance the quality of the output. Hayakawa [20] proposed an uncalibrated photometric stereo using singular value decomposition (SVD) based on the method proposed by Woodham *et al.* [68]. Yuille *et al.* [73] analyzed Hayakawa’s method, and proposed a method to solve the bas-relief ambiguity. While Hayakawa’s method assumes a directional light source, Kim and Burger [24] estimated the position of the point light source. Basri *et al.* [5] used spherical harmonics to apply uncalibrated photometric stereo to arbitrary illumination. Sato *et al.* [50] used ISOMAP to estimate the surface normal of an object with specular reflection that follows the Torrance-Sparrow reflection model. Tan *et al.* [60] enhanced the resolution of the surface normal by using photometric stereo. Chandraker *et al.* [8] removed the shadow by using the graph cut method, and estimated the surface normal robustly from only four images. Wu *et al.* [71,62,69] used the graph cut method and the belief propagation method in order to estimate the smooth surface. Wu and Tang [70] estimated the surface normal and the albedo robustly using the Expectation-Maximization algorithm. Sato and Ikeuchi [49] applied the photometric stereo to a huge outdoor object. Some photometric stereo methods can be applied to non-Lambertian objects [44,61]. Hertzmann and Seitz [21] used a spherical object colored with the same paintings as the target object, and they robustly estimated the surface normal of the object. Goldman *et al.* [17] estimated the BRDF (bidirectional reflectance distribution function) of the object’s surface using photometric stereo, under the condition that the BRDF can be categorized into two types or a combination of those types. Lu and Little [32] and Shen *et al.* [54] also estimated both the surface normal and the BRDF. Helmholtz stereo [33,75,65,76] applied stereo and photometric stereo to a non-Lambertian model. Alldrin *et al.* [3] applied photometric stereo to non-Lambertian objects with a different approach from Helmholtz stereo. Photometric Sampler can also estimate the surface normal of specular objects [36]. Narasimhan *et al.* [35] proposed that the surface normal of an object that is placed inside a scattering medium can be estimated from five light sources. In this paper, we propose a robust photometric stereo method for a Lambertian object under the condition that only a small number of images is supplied. This sparse photometric stereo method is useful in a wide field of applications such as applying it to a robot’s eye for object recognition, obtaining the surface normal in a small space by a gastrofiberscope, using it to create a handheld 3D scanner for entertainment, and so on.

The basic idea of our approach is the same as the successful algorithm called ShadowCuts [8]. We believe that this idea employed by us and Chandraker *et al.* [8] is the most beneficial idea for solving the sparse photometric stereo problem. Each pixel is illuminated by several light sources, and there is a combination of light sources that produces the best result. A notable difference is that ShadowCuts chose the surface normal using an image irradiance equation, while we choose it using a geometric property, namely, the smoothness constraint. If there is an error in the light source direction, ShadowCuts estimates the wrong surface normal due to bas-relief ambiguity [7]. If the object surface can be assumed as smooth, our method is more useful than ShadowCuts.

We describe the separation method of the reflection in Section 2. We propose the median photometric stereo method in Section 3. We discuss how many lights are needed to remove the specularities in Section 4. We present some experimental results in Section 5, and conclude the paper in Section 6.

2 Removing Ambient Light Reflected at a Transparent Surface

As is shown in Fig. 1, the observed image includes foreground, background, light, and object. The scene in front of the display case is reflected, and is observed as a foreground. The light is also reflected at the display case. We observe the scene behind the display case, defined as a background. The object is also affected by ambient light. In addition, some additional value is detected by the camera due to its dark current. Moreover, some amount of light reaches to the background. By considering these effects, the observed image with the light on (I_{on}) can be formulated as follows:

$$\begin{aligned} I_{\mathrm{on}} = {} & I_{\mathrm{foreground,ambient}} + I_{\mathrm{background,ambient}} + I_{\mathrm{background,light}} \\ & + I_{\mathrm{light}} + I_{\mathrm{object,light}} + I_{\mathrm{object,ambient}} + I_{\mathrm{dark}} \,, \end{aligned} \tag{1}$$

where $I_{\mathrm{foreground,ambient}}$ represents the scene in front of the display case reflected at the transparent surface, I_{light} represents the light coming from the lamp and reflecting at the transparent surface, and I_{dark} represents the dark current of the camera. The object and the scene behind the object are illuminated by both the lamp and the surrounding environment. The brightness of the object illuminated by the lamp is expressed as $I_{\mathrm{object,light}}$ and that illuminated by the surrounding environment is expressed as $I_{\mathrm{object,ambient}}$. The background, or the scene behind the object, is transmitted through the transparent surface. The brightness of the background illuminated by the lamp is expressed as $I_{\mathrm{background,light}}$ and that illuminated by the surrounding environment is expressed as $I_{\mathrm{background,ambient}}$. The surrounding environment contains both the foreground and the background and is called the ambient light. We assume that the dark current in a camera is contaminated by the camera noise which can be represented as a white noise (Gaussian noise). We take multiple images and calculate their average; thus, the noise (*i.e.*, dark current) can be represented as a constant value.

Next, we take an image with the lamp switched off. In this case, the observed light will be

$$\begin{aligned} I_{\mathrm{off}} = {} & I_{\mathrm{foreground,ambient}} + I_{\mathrm{background,ambient}} \\ & + I_{\mathrm{object,ambient}} + I_{\mathrm{dark}} \,. \end{aligned} \tag{2}$$

The difference between Eq. (1) and Eq. (2) can be expressed as follows if both values are not saturated:

$$I_{\mathrm{on}} - I_{\mathrm{off}} = I_{\mathrm{light}} + I_{\mathrm{object,light}} + I_{\mathrm{background,light}} \,. \tag{3}$$

These two images are taken from the same camera with fixed position and fixed camera settings. The background does not affect the appearance of the object; thus, only the object and the light are obtained from the pixels in the object region. We find that the image represented by Eq. (3) does not have the effect of ambient light, foreground, or background. In addition, we take multiple images and calculate their average also for I_{off}. Thus, the image represented by Eq. (3) has no dark current. The influence of the room light disappears in Eq. (3), and we can measure the object whether the room light is on or off. However, in practice, the specular reflection of the object caused by the room light will often be saturated, so it is better to switch off the room light in some situations. Through this procedure, we can separate the effect of the light source and the effect of ambient light (Fig. 3). The separation in our case is different from the previously mentioned separation methods [14, 52, 59, 51, 47, 27, 28, 48, 41], which separate the foreground and the background, since the light reflecting at the transparent surface I_{off} can be regarded as the foreground.

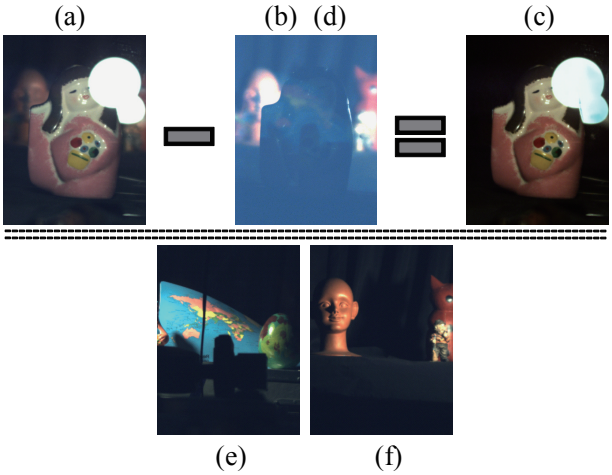


Fig. 3 Separation result: (a) Input image with the light on, I_{on}, (b)(d) input image with the light off, I_{off}, (brightness is multiplied by 4 for visibility), and (c) separated image with the light and the object, $I_{\mathrm{on}} - I_{\mathrm{off}}$. (e) is the foreground image, $I_{\mathrm{foreground,ambient}}$, and (f) is the background image, $I_{\mathrm{background,ambient}}$. These two images are only shown for reference for the readers, and are not used in this method.

In the field of shape-from-shading research, calculating the difference between the light-on image and the light-off image is used as a preprocessing step for removing the dark current and the ambient light. Effectively using the light-on image and the light-off image is simple but beneficial, and it is also used aggressively in recent research [45, 57, 42, 13, 31]. Therefore, we use this method for separation.

The image represented by Eq. (3) contains not only the object $I_{\mathrm{object,light}}$ but also the light I_{light}; thus, we cannot apply conventional photometric stereo. The direct reflection of the light on the display surface I_{light} is too bright, and ordinary polarizers cannot completely remove the reflection. In our experiment, we avoid saturation of the diffuse reflection component I_{diffuse} (where $I_{\mathrm{object,light}} = I_{\mathrm{specular}} + I_{\mathrm{diffuse}}$); however, the direct reflection of the light I_{light} often causes saturation. Section 3 presents a method to estimate the surface normal and the albedo even if the pixels that contain the reflection of light I_{light} or I_{specular} are saturated. If we need to analyze the specular reflection of the object, we carefully avoid its saturation.

3 Median Photometric Stereo

3.1 M-Estimation Using Laplace Distribution

We obtain L number of brightnesses per pixel from L number of input images. L number of brightnesses sometimes includes the specular reflection or the shadow. Therefore, we adopt a strategy to calculate the optimal surface normal and albedo through iterative computation.

First, we explain the smoothness constraint of the albedo ρ. The albedo is expressed by three values, namely RGB channel, and the proposed algorithm is applied per channel. By minimizing

$$\varepsilon_{\rho,\mathrm{avg}} = (\rho_x(x,y))^2 + (\rho_y(x,y))^2 \ , \tag{4}$$

we obtain the following albedo ρ

$$\rho(x,y) = \mathrm{average}\big(\rho(x+1,y), \rho(x-1,y), \rho(x,y+1), \rho(x,y-1)\big)\,. \tag{5}$$

Eq. (5) is derived by the discussion presented in Appendix A.

Eq. (4) uses L2-norm, which means that it uses Gaussian distribution as the M-estimator [43]. This M-estimator is not robust to outliers. We use Laplace distribution (double exponential distribution) as the M-estimator, which is more robust than Gaussian distribution [43]. In this case, the smoothness constraint of the albedo is represented as the following L1-norm.

$$\varepsilon_{\rho,\mathrm{med}} = |\rho_x(x,y)| + |\rho_y(x,y)|\,. \tag{6}$$

The albedo ρ that minimizes this equation would be

$$\rho(x,y) = \mathrm{median}\big(\rho(x+1,y), \rho(x-1,y), \rho(x,y+1), \rho(x,y-1)\big)\,. \tag{7}$$

Eq. (7) is derived by the discussion presented in Appendix B.

Next, we explain the optimization of the albedo. We represent the L number of input images with the subscript $i = \{1,\ldots,L\}$. Since there are some outliers, such as shadows, we use Laplace distribution as the M-estimator.

$$\varepsilon_{\rho,\mathrm{opt}} = \sum_{i=1}^{L} \left| \rho - \frac{I_i}{S_i \cdot N} \right|\,, \tag{8}$$

where S_i represents the direction of the light source, which is given a priori, and N represents the surface normal. By fixing the surface normal N, the albedo ρ which minimizes Eq. (8) will be given as follows:

$$\rho = \mathrm{median}\big(\frac{I_i}{S_i \cdot N} \;\big|\; i = 1,\ldots,L\big)\,. \tag{9}$$

The overall cost for estimating the albedo is as follows:

$$\varepsilon_\rho = \iint \varepsilon_{\rho,\mathrm{opt}} + \lambda_{\rho,\mathrm{med}}\varepsilon_{\rho,\mathrm{med}} + \lambda_{\rho,\mathrm{avg}}\varepsilon_{\rho,\mathrm{avg}} dxdy\,, \tag{10}$$

where $\lambda_{\rho,\mathrm{med}}$ and $\lambda_{\rho,\mathrm{avg}}$ are the user-defined constant values representing the smoothness of the object's surface. By increasing $\lambda_{\rho,\mathrm{med}}$, the method will be robust to outliers, and by increasing $\lambda_{\rho,\mathrm{avg}}$, the method will be robust to local minima.

Finally, we explain the optimization of the surface normal. In order to calculate the surface normal by photometric stereo, we need at least three data sets. The number of all combinations for selecting the three data sets from L data is ${}_L\mathrm{C}_3 = \binom{L}{3}$. We represent each surface normal estimated from each combination as $N_m (m = 1,\ldots,{}_L\mathrm{C}_3)$. Each surface normal is calculated by conventional photometric stereo [67]. We estimate the surface normal by minimizing the following formula:

$$\varepsilon_{\mathrm{N,opt}} = \sum_{m=1}^{{}_L\mathrm{C}_3} |N - N_m|\,. \tag{11}$$

The surface normal is expressed by three values, namely, the XYZ axis, and the proposed algorithm is applied per axis. We calculate each element of normal vector one by one; thus, minimizing Eq. (11) results in the following formula (Fig. 4):

$$N = \mathrm{median}\big(N_m \;\big|\; m = 1,\ldots,{}_L\mathrm{C}_3\big)\,. \tag{12}$$

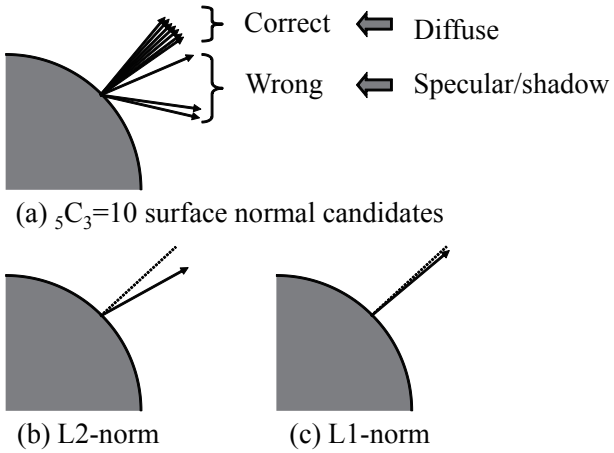


Fig. 4 Estimating the surface normal by taking the median of candidate surface normals.

Therefore, the overall cost for estimating the surface normal is as follows:

$$\varepsilon_{\mathrm{N}} = \iint \varepsilon_{\mathrm{N,opt}} + \lambda_{\mathrm{N,med}}\varepsilon_{\mathrm{N,med}} + \lambda_{\mathrm{N,avg}}\varepsilon_{\mathrm{N,avg}} dxdy\,, \tag{13}$$

where $\lambda_{\mathrm{N,med}}$ and $\lambda_{\mathrm{N,avg}}$ are the user-defined constant values representing the smoothness of the object's surface. By increasing $\lambda_{\mathrm{N,med}}$, the method will be robust to outliers, and by increasing $\lambda_{\mathrm{N,avg}}$, the method will be robust to local minima.

We first minimize Eq. (13) until convergence in order to estimate the surface normal. Eq. (13) only depends on the surface normal and the input data; thus, its convergence is guaranteed. After convergence, we minimize Eq. (10) until convergence in order to estimate the albedo. Although Eq. (10) also depends on the surface normal, we use the estimated surface normal, which is treated as a constant value during this step; thus, its convergence is guaranteed.

3.2 Proposed Algorithm

Let us organize the algorithm. First, we set the appropriate initial values for the surface normal and the albedo. We utilize conventional photometric stereo [67] to calculate the initial values from the whole input images although the images contain specular reflections and shadows.

Next, we calculate all $_L\mathrm{C}_3$ number of candidates of the surface normal for all pixels. Here, we represent a set of three light sources as follows:

$$\{i, j, k\} \in \mathcal{M}\,, \tag{14}$$

where

$$|\mathcal{M}| = {}_L\mathrm{C}_3 \tag{15}$$
$$i \in \{1, \ldots, L\} \tag{16}$$
$$j \in \{1, \ldots, L\} \tag{17}$$
$$k \in \{1, \ldots, L\}\,. \tag{18}$$

The candidate surface normal is calculated as follows:

$$N_m = \begin{bmatrix} S_i^T \\ S_j^T \\ S_k^T \end{bmatrix}^{-1} \begin{bmatrix} I_i \\ I_j \\ I_k \end{bmatrix} . \tag{19}$$

If the invert matrix does not exist, we set a large number for N_m so that it can be judged as an outlier.

In order to estimate the surface normal, we iterate the following process until convergence:

1. We initialize a set of candidates.
   $\mathcal{A} \leftarrow \emptyset$.
2. We add the surface normal calculated from input images to the set of candidates.
   $\mathcal{A} \leftarrow \mathcal{A} \cup \{N_m \mid m = 1, \ldots, {}_L\mathrm{C}_3\}$.
3. We also add the neighboring surface normals weighting with an integer value $\lfloor \lambda_{\mathrm{N,med}} \rfloor$.
   $\mathcal{A} \leftarrow \mathcal{A} \cup \bigcup_{\lfloor \lambda_{\mathrm{N,med}} \rfloor} \{N(x+1,y), N(x-1,y), N(x,y+1), N(x,y-1)\}$.
4. We calculate the median of the set.
   $N_{\mathrm{opt}} = \mathrm{median}(\mathcal{A})$.
5. We also calculate the average of the neighboring surface normal.
   $N_{\mathrm{avg}} = \frac{1}{4}\left(N(x+1,y) + N(x-1,y) + N(x,y+1) + N(x,y-1)\right)$.
6. We calculate the weighted sum for the final surface normal and normalize it.
   $N = \mathrm{normalize}\left(\frac{N_{\mathrm{opt}} + \lambda_{\mathrm{N,avg}} N_{\mathrm{avg}}}{1 + \lambda_{\mathrm{N,avg}}}\right)$.

The above calculation is done for each pixel. After calculating all pixels, we calculate this step again. Iterations stop when the average change of the surface normal between the current step and the previous step is less than the convergence criterion.

After that, we also estimate the albedo by iterating the following process until convergence:

1. We initialize a set of candidates.
   $\mathcal{A} \leftarrow \emptyset$.
2. We add the albedo calculated from input images to the set of candidates.
   $\mathcal{A} \leftarrow \mathcal{A} \cup \{\frac{I_i}{S_i \cdot N} \mid i = 1, \ldots, L\}$.
3. We also add the neighboring albedos weighting with an integer value $\lfloor \lambda_{\rho,\mathrm{med}} \rfloor$.
   $\mathcal{A} \leftarrow \mathcal{A} \cup \bigcup_{\lfloor \lambda_{\rho,\mathrm{med}} \rfloor} \{\rho(x+1,y), \rho(x-1,y), \rho(x,y+1), \rho(x,y-1)\}$.
4. We calculate the median of the set.
   $\rho_{\mathrm{opt}} = \mathrm{median}(\mathcal{A})$.
5. We also calculate the average of the neighboring albedo.
   $\rho_{\mathrm{avg}} = \frac{1}{4}\left(\rho(x+1,y) + \rho(x-1,y) + \rho(x,y+1) + \rho(x,y-1)\right)$.
6. We calculate the weighted sum for the final albedo.
   $\rho = \frac{\rho_{\mathrm{opt}} + \lambda_{\rho,\mathrm{avg}} \rho_{\mathrm{avg}}}{1 + \lambda_{\rho,\mathrm{avg}}}$.

The above calculation is done for each pixel. After calculating all pixels, we calculate this step again. Iterations stop when the average change of the albedo between the current step and the previous step is less than the convergence criterion.

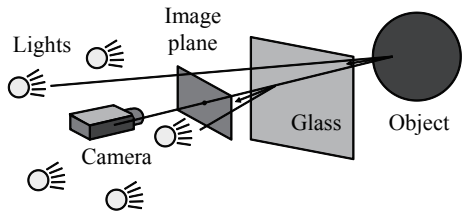


Fig. 5 The principle of measuring the surface normal by photometric stereo using five light sources.

Finally, the surface normal is integrated to the height map by the relaxation method (*i.e.*, Poisson solver) while enforcing the natural boundary condition [23, 22]. The relaxation method is one of the best methods for surface reconstruction if the surface is continuous and the noise function of the estimated surface normal is uniform through each pixel. However, if the surface is discontinuous in a small area or the estimated surface normal has an outlier, another robust method such as the method proposed by Agrawal *et al.* [2] is more useful. We assume the object is continuous and the noise function of each pixel is uniform, thanks to our robust method of estimating the surface normal; thus, we use the relaxation method.

In our algorithm, the surface normal and the albedo are updated using four-neighbor information in order to avoid falling into a local minimum. RANSAC [15] is a famous technique to remove the outliers; however, if the input images are small (*i.e.*, if L is small), we can calculate all ${}_L\mathrm{C}_3$ candidates and remove the outliers more robustly. If the number of the images is large, calculating all candidates takes a long time (${}_L\mathrm{C}_3 = L(L-1)(L-2)/6 \simeq O(L^3)$); thus, RANSAC is more effective than our approach in such a case. Mukaigawa *et al.* [34] removed the outliers from many input images using RANSAC, and generated the images to which the photometric stereo can be applied.

Barsky and Petrou [4] used color information [26] in order to detect the specular pixels. This approach has a problem in that it cannot be applied if the object is white or the specular pixel is saturated, so we do not adopt this approach. Barsky and Petrou also used another type of algorithm to detect the specular pixel, which reinforces the idea that their color information approach is not always applicable.

Conventional binocular stereo cannot be applied because the specular reflection interferes with finding the corresponding points. Our method treats the specular reflection and the shadow as outliers. Therefore, our method can estimate the surface normal and the albedo even if the specular pixels are saturated. There are many methods to separate diffuse reflection and specular reflection; however, most of them cannot be applied if the pixel is saturated, and most of them cannot be applied if the object is white. We assume that the direction of the light sources is known; thus, we can estimate the orientation of the glass case and detect the reflection of the display case. However, we do not estimate it. Even if we detect the reflection of the display case, we cannot detect the specular reflection of the object and the shadow; thus, we propose the outlier removal algorithm.

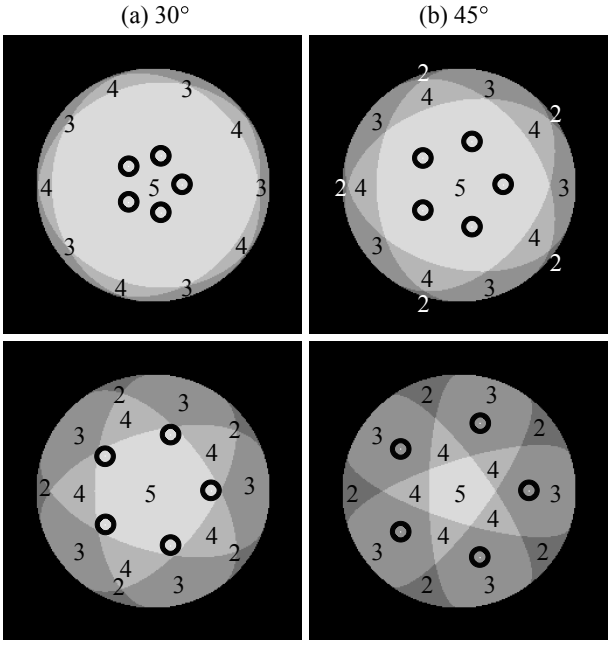


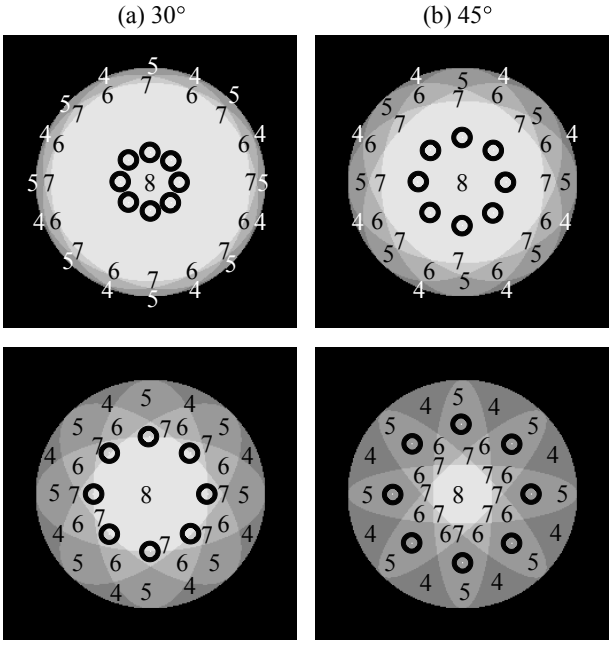
Fig. 6 Gaussian sphere of the object surface when illuminated by five lights. Cases when the angle between the light and the camera is (a) 30°, (b) 45°, (c) 60°, and (d) 75°.

4 Analysis of Gaussian Sphere

4.1 Five-Light Photometric Stereo

Considering the reflection at the glass case, five lights give a good result in practice (Fig. 5). We will explain the detail below by using a Gaussian sphere viewed from above (Fig. 6). Five lights are placed as an equilateral pentagon. Fig. 6 (a)–(d) shows the Gaussian spheres when the angle between the light source and the viewing direction is 30°, 45°, 60°, and 75°, respectively. The marked circle represents the specular reflection of the object's surface. The numbers represent the numbers of the lights that illuminate the surface. In the case of "30°," all regions are illuminated by three or more lights; thus, the surface normal can be calculated in all regions. The specular reflection of the object's surface appears in region "5"; thus, we have four input data sets to calculate the surface normal. Even if the specular reflection of the glass display appears in region "5," we have three input data sets to calculate the surface normal. Region "2," which is illuminated by only two lights, appears in case "45°." However, the observed area is small, and we can interpolate from neighboring pixels. The area of region "2" becomes larger in case "60°." Consequently, if the angle between the light source and the viewing direction is less than 45°, we can estimate the surface normal and the albedo from five lights.

The above discussion considers the attached shadow, while the cast shadow is not considered. It also assumes the specular spike but not the specular lobe [38]. In addition, we do



(a) 30° (b) 45° (c) 60° (d) 75°

Fig. 7 Gaussian sphere of the object surface when illuminated by eight lights. Cases when the angle between the light and the camera is (a) 30°, (b) 45°, (c) 60°, and (d) 75°.

not know which region each pixel belongs to when the surface normal is unknown. Therefore, we have proposed an iterative algorithm in Section 3.

4.2 Eight-Light Photometric Stereo

Practically, five lights are enough to estimate a specular object kept inside a glass case. However, some part of the object is only illuminated by two light sources. In order to overcome this problem, we have to increase the number of the light sources. If the specular object is not kept inside a glass case, six lights are enough for the measurement [58]. In this section, we show that eight lights are enough if the specular object is kept inside a glass case.

The Gaussian sphere in this case is shown in Fig. 7. Eight lights are placed as an equilateral octagon. According to the figure, any part of the whole object surface is illuminated by at least four light sources. The specular reflection of the object appears in region "5" (and "7," "8"), and it does not appear in region "4." Therefore, even if the glass reflection appears in region "4" or "5," at least three diffuse pixels can be obtained at such regions; thus, we can estimate the surface normal for every region.

The orientation of the light source can be represented as follows:

$$\begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}. \tag{20}$$

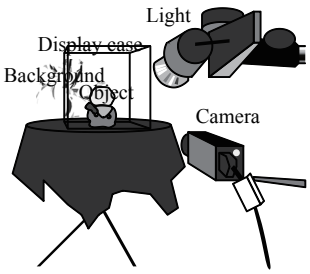


Fig. 8 Experimental setup.

If the light source can be represented by the above formula, the position of the specular reflection on the Gaussian sphere can be represented by the following:

$$\begin{pmatrix} \sin(\theta/2)\cos\phi \\ \sin(\theta/2)\sin\phi \\ \cos(\theta/2) \end{pmatrix}. \tag{21}$$

Suppose that we lit the object by the light sources where $\phi = 0^\circ, \pm 45^\circ, \pm 90^\circ, \pm 135^\circ, 180^\circ$. Let us consider the specular reflection caused at $\phi = 0^\circ$. The specular reflection occurs in region "5," and is lit by the light sources that are placed at $\phi = 0^\circ, \pm 45^\circ, \pm 90^\circ$. Now, we derive the inclination angle of the light source θ when this region becomes region "3." This region "3" is only illuminated by the light sources that are placed in $\phi = 0^\circ, \pm 45^\circ$. Therefore, θ is derived as follows:

$$\begin{pmatrix} \sin\theta\cos\pm 90^\circ \\ \sin\theta\sin\pm 90^\circ \\ \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \sin(\theta/2)\cos 0^\circ \\ \sin(\theta/2)\sin 0^\circ \\ \cos(\theta/2) \end{pmatrix} = 0. \tag{22}$$

The above equation is simplified as follows:

$$\cos\theta\cos(\theta/2) = 0. \tag{23}$$

The solution of this equation is $\theta = \pm 90^\circ, \pm 180^\circ$; however, in this case, the solution is $\theta = 90^\circ$. Therefore, if the light source is set on the same side as the camera, there exists no region "3." To conclude, eight lights are enough in order to estimate the surface normal of specular objects kept inside a transparent display case, if the light sources are not set behind the objects.

The above discussion assumes that the specular reflection is a perfect mirror reflection. Usually, the object has widely spread specular reflection. In order to prevent the overlap of specular reflection as much as possible, the eight lights should be placed as an equilateral octagon. In addition, the angle between the light and the camera should be large.

5 Experimental Result

5.1 Experimental Setup

The experimental setup is shown in Fig. 8 and Fig. 9. The target object is covered with a glass or acrylic display case, and is observed with one camera and five or more lamps. LED lamps

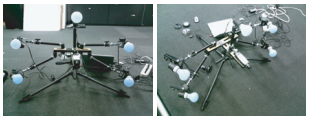


Fig. 9 Measurement system “Photometric Star.”

Fig. 10 Input images: (a) The images taken with the light on, (b) one of the images taken with the light off.

or fluorescent lamps are more appropriate than incandescent lamps since precious artworks would be sometimes damaged by far-infrared light or deep-ultraviolet light. However, the experiments in this paper use incandescent lamps. The camera is fixed, and we take the image with three-band RGB. The light source direction is obtained a priori from the mirror sphere. The light source brightness is obtained a priori by illuminating a white diffuse object. We take four or more images for a diffuse-only object, and take five or more images for an object that includes specularities. For each shot, we take both the image with the light switched on and the image with the light switched off. Each face of the display case should be transparent, and its thickness and orientation are unknown. However, we assume that the display case is a box-type object structured by transparent planes with uniform thickness.

5.2 Evaluation

Fig. 10 is the input image and Fig. 11 is the separated result. Fig. 12 (a)–(c) are the estimated albedo, height, and shape, respectively. Fig. 13 is an example image rendered by using estimated parameters with a different viewpoint and different light direction from the acquisition stage. In Fig. 12 (c), we did not remove the background; however, it can be easily removed by simply thresholding the dark brightness of the background, or using region segmentation techniques such as Lazy Snapping [29], GrabCut [46], or Adobe Photoshop [1]. All results in Fig. 12 are obtained from five input images.

Fig. 14 shows the result of applying conventional photometric stereo using only three images under three different lights. The conventional photometric stereo cannot handle the reflection at the display case or the specular reflection of the object surface; thus, it is impossible to estimate the true surface normal and albedo. The conventional photometric stereo is not robust; thus, it is affected by the error in the brightness of the light source. Calibrated photometric stereo assumes the light source direction and the brightness are known. Although we compensated for these factors before the experiment, some amount of the error of the compensation still remains as a noise.

Fig. 15 shows the result of light-stripe range scanning (active stereo). We scan from exactly the front of the acrylic box with a Konica Minolta VIVID 910 [25]. Owing to the

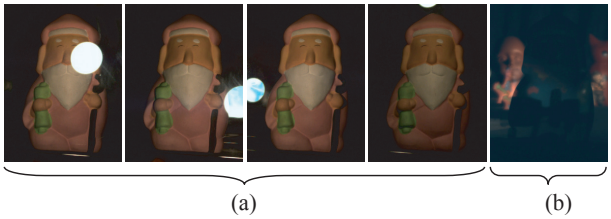


Fig. 11 Separation result: (a) The images that only contain the object and the light, (b) one of the images that contain the foreground and the background.

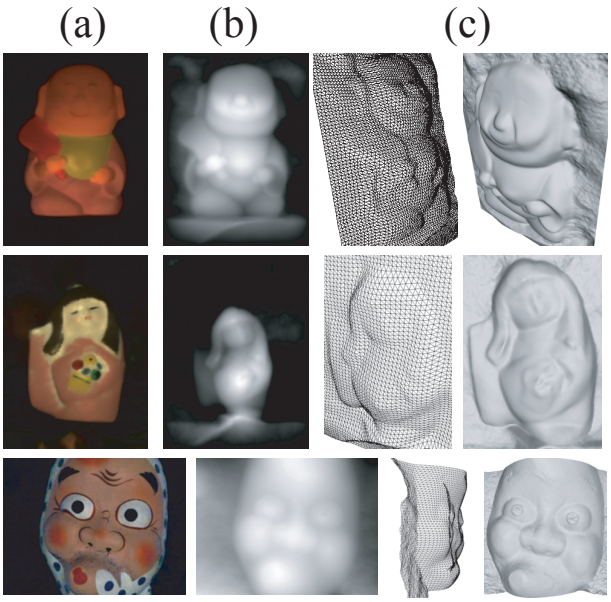


Fig. 12 Estimated result: (a) Albedo, (b) height, (c) shape.

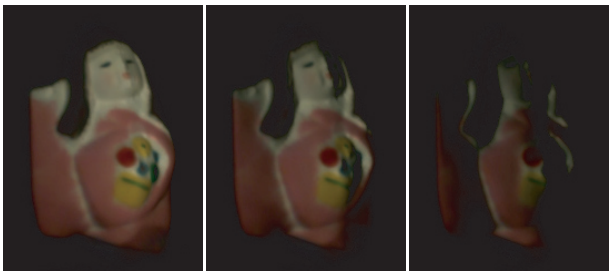
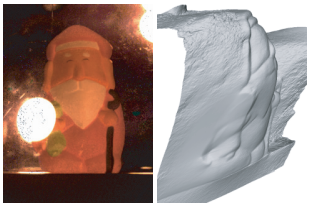
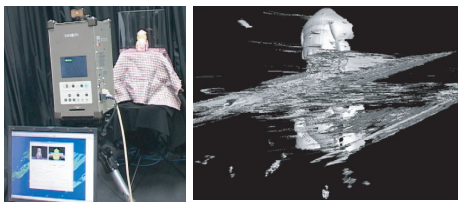


Fig. 13 Rendered image with a different viewpoint and light direction from the acquisition stage.



(a) (b)

Fig. 14 Result of conventional photometric stereo: (a) Albedo, (b) shape.



(a) (b)

Fig. 15 The result of light-stripe range scanning: (a) Setup, (b) estimated shape.

Table 1 Evaluation.

	Height error (RMSE)	Normal error (RMSE)
Light-stripe range scanning	2.43cm	36.0°
Conventional photometric stereo	1.32cm	28.7°
Proposed method	0.86cm	10.1°

reflection from the acrylic surface, the sensor cannot obtain the true shape. However, the sensor can possibly determine the shape if it is observed on a slant. The proposed method can estimate the shape by observing the object not only from the front of the case but also from a slanting direction.

Fig. 16 (b) shows the result of our method applied to an object 15 cm tall, from six input images. Fig. 16 (a) is a ground truth obtained by scanning the object brought out from the acrylic case by VIVID910. Fig. 16 (c) is the result of conventional photometric stereo from three certain input images. Fig. 16 (d) is the same as Fig. 15. The comparison in Table 1 shows that the proposed method is more effective than the other methods. Fig. 16 (e) (3) represents the height error, and we can find out that the estimated shape is bent. The distortion is caused when the height is calculated from the surface normal. A possible solution would be to combine the proposed method with a stereo method. Fig. 16 (e) (4) represents the surface normal error, and we find out that the error is caused at the concave part. This error is caused by the cast shadow, the interreflection, and the smoothness constraint. A possible solution would be to remove the cast shadow from many light sources or to apply interreflection-removing algorithms.

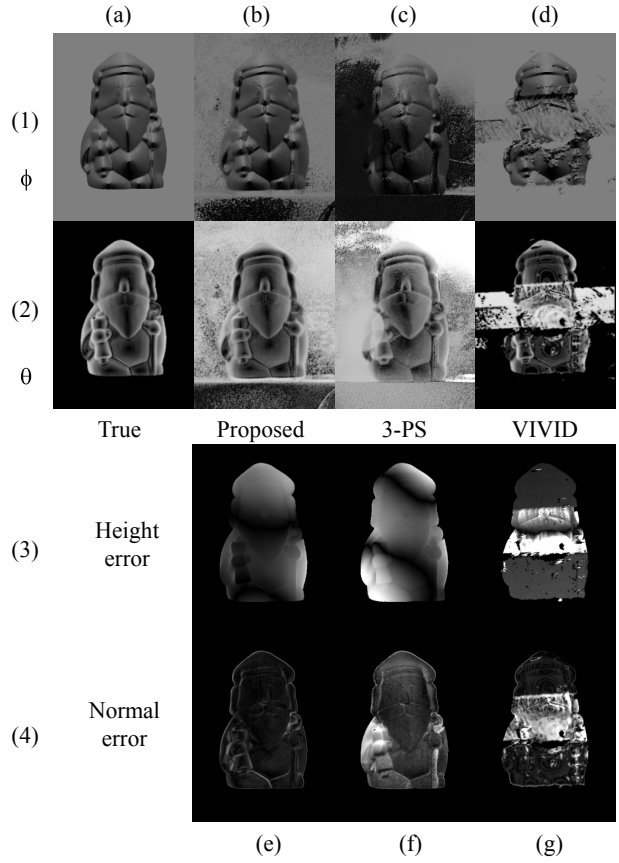


Fig. 16 Comparison: (a) True value, (b) the result of the proposed method, (c) the result of conventional photometric stereo, (d) the result of light-stripe range scanning, (e) the error of the proposed method, (f) the error of conventional photometric stereo, (g) the error of light-stripe range scanning; (1) Azimuth angle (white: upper direction, black: lower direction) (2) zenith angle (black: 0°, white: 90°) (3) the difference of the height (the brighter the noisier), (4) the difference of the surface normal (the brighter the noisier).

5.3 Application to Museum

Photometric stereo is effectively used for digital archiving of artworks [63, 19]. For the starting point of the virtual museum project, we went to The University Museum located at The University of Tokyo, in order to verify our method. We digitized a shell collected in Australia called “chlamys australis,” which is kept in an acrylic display case (Fig. 17). This colorful shell is a kind of scallop, and tastes better than commonly consumed scallops. We took six images for reconstructing the surface. The result is shown in Fig. 18. The wavy surface of the shell is clearly modeled by the proposed method.

(a) (b)

Fig. 17 The University Museum, The University of Tokyo: (a) Scanning setup, and (b) the chlamys australis shell.

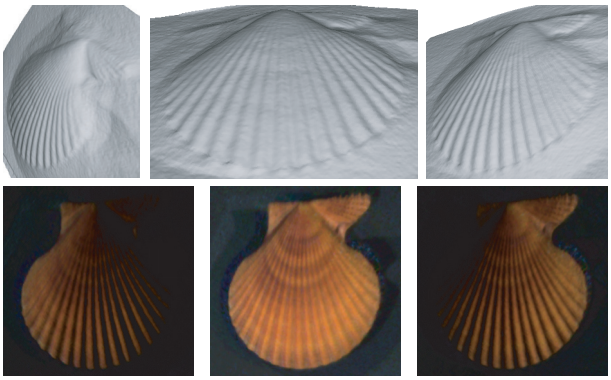


Fig. 18 The result of applying our method to the shell.

5.4 Application to Segonko Tumulus

We also applied our method to a cultural asset called the Segonko Tumulus, which is located in the Kumamoto prefecture in Japan (Fig. 19). The tumulus was built around A.D. 500. Its wall is not only carved but also painted with red, green, and yellow. To preserve the current state of the painting, the tumulus is not open to the public; thus, providing it digitally for common view is important.

We developed two types of systems: The Photometric Diamond, shown in Fig. 20 (a) and the Photometric Wing, shown in Fig. 20 (b), where both have eight lights. The Photometric Wing proved to be better since it has a wide baseline and the weight is light. The light sources are not placed completely evenly in this system. The target object is not kept in a display case, and it is almost flat; thus, there is no problem in using a system whose lights are not placed in an equilateral octagon.

The tumulus is a very small chamber, and only one or two people can go inside (Fig. 21). It is difficult to illuminate the wall with a large number of light sources. A system that has only eight lights is useful for these kinds of small places.

The input images are shown in Fig. 22 (a). The images rendered from the estimated albedo and the surface normal are shown in Fig. 22 (b). A typhoon attacked us on the first day of our scanning mission. We captured the images another day; however, the wall was

(a) (b)

Fig. 19 Segonko Tumulus: (a) Replica, and (b) inside the tumulus.

(a) (b)

Fig. 20 Measurement system used for Segonko Tumulus: (a) Photometric Diamond, and (b) Photometric Wing.

Fig. 21 The measurement system capturing images inside the small cave.

wet. The specularities occurred on the surface of the wall, but our method was not affected by such specularities. Note that the specularities are removed in Fig. 22 (b).

The result of applying our method to the carvings in the Segonko Tumulus is shown in Fig. 23. We can clearly detect the hole in the center of the concentric circles, and we can give evidence to the opinion of the archaeologist that the circle was carved by using compasses.

Fig. 22 (a) Input images, and (b) rendered images of Segonko Tumulus.

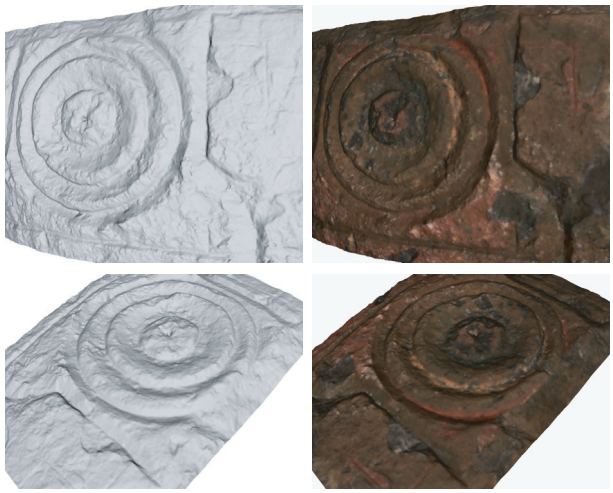


Fig. 23 Shape estimation results of Segonko Tumulus.

6 Conclusion

In this paper, we propose a method for digitizing artworks to be included in a virtual museum. The proposed method can digitize artwork that is kept inside a glass display case as well as outdoor cultural assets. In this paper, we describe how to improve photometric stereo so that it will be less affected by specularities and shadows.

Binocular stereo cannot be applied if there is a reflection at the display case or if there is a specular reflection at the object's surface. However, after generating the diffuse-only images by our method, we can apply binocular stereo. A light-stripe range sensor cannot estimate the object's shape if the object is observed from the front of the display case, but our method can do this. We are now planning to estimate a more precise shape by observing

the object from multiple viewpoints with a method combining photometric stereo and multi-view stereo [30, 74, 18].

We propose an algorithm that can be applied to the input image taken with a machine like those shown in Fig. 9 and Fig. 20. While the current method is adequate when the number of lights is small, if we could increase the number of lights, we could obtain much more information [20, 73, 55, 12, 6, 16]. Our method assumes that the diffuse reflection can be modeled as a Lambertian model; extending our method to other reflection models would also be important. The quality of the shape would increase by considering the interreflection; thus, it might be improved by applying the solutions suggested by [37, 66, 53, 39, 72].

The proposed method can be applied to an object that is not covered by a display case, and it can be applied also to specular objects. Our optimization algorithm is robust to outliers caused by specular reflection or shadow. We can also remove the effect of dark current and ambient light. The method can be applied in a variety of applications, and this warrants further investigation.

Acknowledgements This research was supported in part by the Ministry of Education, Culture, Sports, Science and Technology under the Leading Project, "Development of High Fidelity Digitization Software for Large-Scale and Intangible Cultural Assets." SONY XCD-X710CR, FUJINON lens, TECHNO SCOPE CFW-46, Chori Imaging VFS-42, SHARP laptop PC, LPL lamp holder, SLIK and ETSUMI camera plates, Edmund Optics articulated arms, Velbon tripods, Labsphere white reference, and TOKYU HANDS specular sphere were used in the experiment. KONICA MINOLTA VIVID 910, the alignment software developed by Oishi *et al.* [40], and the camera calibration [64] software developed by Ryo Kurazume, Hiroki Unten, Ryo Ohkubo, Fujitsu Ltd., and Atsuhiko Banno, were used in the evaluation. Fast algorithm for calculating the median introduced by Press *et al.* [43] is used. The measurement of "chlamys australis" was supported by Hirdy Miyamoto and The University Museum, The University of Tokyo. The measurement of "Segonko Tumulus" was supported by Nobuaki Kuchitsu, Kumamoto Prefectural Board of Education, and Toppan Printing Co., Ltd. The authors thank Joan Knapp for proofreading and editing this manuscript. They also thank anonymous reviewers for their careful reviews of the paper.

References

1. Adobe Photoshop, `http://www.adobe.com/products/photoshop/` (ADOBE)
2. A. K. Agrawal, R. Raskar, and R. Chellappa, "What is the range of surface reconstructions from a gradient field?," in *Proceedings of European Conference on Computer Vision*, pp. 578–591 (2006)
3. N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (2008)
4. S. Barsky and M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1239–1252 (2003)
5. R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *International Journal of Computer Vision*, vol. 72, no. 3, pp. 239–257 (2007)
6. N. Birkbeck, D. Cobzas, P. Sturm, and M. Jagersand, "Variational shape and reflectance estimation under changing light and viewpoints," in *Proceedings of European Conference on Computer Vision*, pp. 536–549 (2006)
7. P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," *International Journal of Computer Vision*, vol. 35, no. 1, pp. 33–44 (1999)
8. M. Chandraker, S. Agarwal, and D. Kriegman, "ShadowCuts: Photometric stereo with shadows," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (2007)
9. C. P. Chen and C. S. Chen, "The 4-source photometric stereo under general unknown lighting," in *Proceedings of European Conference on Computer Vision*, pp. 72–83 (2006)
10. E. N. Coleman Jr. and R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," *Computer Graphics and Image Processing*, vol. 18, no. 4, pp. 309–328 (1982)
11. R. Courant and D. Hilbert, *Methods of mathematical physics*, p. 560. Wiley, New York (1953)
12. M. S. Drew, "Reduction of rank-reduced orientation-from-color problem with many unknown lights to two-image known-illuminant photometric stereo," in *Proceedings of International Symposium on Computer Vision*, pp. 419–424 (1995)

13. E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 673–678 (2004)
14. H. Farid and E. H. Adelson, "Separating reflections from images by use of independent component analysis," *Journal of the Optical Society of America A*, vol. 16, no. 9, pp. 2136–2145 (1999)
15. M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395 (1981)
16. A. S. Georghiades, "Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 816–825 (2003)
17. D. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 341–348 (2005)
18. P. Fua and Y. G. Leclerc, "Object-centered surface reconstruction: combining multi-image stereo and shading," *International Journal of Computer Vision*, vol. 16, no. 1, pp. 35–56 (1995)
19. D. V. Hahn, D. D. Duncan, K. C. Baldwin, J. D. Cohen, and B. Purnomo, "Digital Hammurabi: design and development of a 3D scanner for cuneiform tablets," in *Proceedings of SPIE*, vol. 6056, pp. 130–141 (2006)
20. H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *Journal of the Optical Society of America A*, vol. 11, no. 11, pp. 3079–3089 (1994)
21. A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: Shape reconstruction with general, varying BRDFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1254–1264 (2005)
22. B. K. P. Horn, *Robot vision*, p. 509. MIT Press, Cambridge Mass. (1986)
23. K. Ikeuchi, "Reconstructing a depth map from intensity maps," in *Proceedings of International Conference on Pattern Recognition*, pp. 736–738 (1984)
24. B. Kim and P. Burger, "Depth and shape from shading using the photometric stereo method," *CVGIP: Image Understanding*, vol. 54, no. 3, pp. 416–427 (1991)
25. Konica Minolta VIVID 910, `http://www.minolta3d.com/products/vi910-en.asp` (MINOLTA3D)
26. G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," *International Journal of Computer Vision*, vol. 2, no. 1, pp. 7–32 (1988)
27. A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," in *Proceedings of European Conference on Computer Vision*, pp. 602–613 (2004)
28. A. Levin, A. Zomet, and Y. Weiss, "Separating reflections from a single image using local features," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 306–313 (2004)
29. Y. Li, J. Sun, C. K. Tang, and H. Y. Shum, "Lazy snapping," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 303–308 (2004)
30. J. Lim, J. Ho, M. H. Yang, and D. Kriegman, "Passive photometric stereo from motion," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1635–1642 (2005)
31. C. Lu and M. S. Drew, "Practical scene illuminant estimation via flash/no-flash pairs," in *Proceedings of Color Imaging Conference*, (2006)
32. J. Lu and L. Little, "Reflectance function estimation and shape recovery from image sequence of a rotating object," in *Proc. International Conference on Computer Vision*, pp. 80–86 (1995)
33. S. Magda, D. J. Kriegman, T. Zickler, and P. N. Belhumeur, "Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 391–398 (2001)
34. Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Analysis of photometric factors based on photometric linearization," *Journal of the Optical Society of America A*, vol. 24, no. 10, pp. 3326–3334 (2007)
35. S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal, "Structured light in scattering media," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 420–427 (2005)
36. S. K. Nayar, K. Ikeuchi, and T. Kanade, "Determining shape and reflectance of hybrid surface by photometric sampling," *IEEE Transactions on Robotics and Automation*, vol. 6, no. 4, pp. 418–431 (1990)
37. S. K. Nayar, K. Ikeuchi, and T. Kanade, "Shape from interreflections," *International Journal of Computer Vision*, vol. 6, no. 3, pp. 173–195 (1991)
38. S. K. Nayar, K. Ikeuchi, and T. Kanade, "Surface reflection: Physical and geometrical perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 6119–634 (1991)
39. S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 935–944 (2006)

40. T. Oishi, A. Nakazawa, R. Kurazume, and K. Ikeuchi, “Fast simultaneous alignment of multiple range images using index images,” in *Proceedings of International Conference on 3-D Digital Imaging and Modeling*, pp. 476–483 (2005)
41. T. Oo, H. Kawasaki, Y. Ohsawa, and K. Ikeuchi, “The separation of reflected and transparent layers from real-world image sequence,” *Machine Vision and Applications*, vol. 18, no. 1, pp. 17–24 (2007)
42. G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, and K. Toyama, “Digital photography with flash and no-flash image pairs,” *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 664–672 (2004)
43. W. H. Press *et al.*, *Numerical recipes in C: the art of scientific computing*, p. 994. Cambridge University Press, Cambridge England (1997)
44. H. Ragheb and E. R. Hancock, “Surface normals and height from non-Lambertian image data,” in *Proceedings of International Symposium on 3D Data Processing, Visualization and Transmission*, pp. 18–25 (2004)
45. R. Raskar, K. H. Tan, R. Feris, J. Yu, and M. Turk, “Non-photorealistic camera: Depth edge detection and stylized rendering using multi-flash imaging,” *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 679–688 (2004)
46. C. Rother, V. Kolmogorov, and A. Blake, “GrabCut: Interactive foreground extraction using iterated graph cuts,” *ACM Transactions on Graphics*, vol. 23, no. 3, pp.=309–314 (2004)
47. B. Sarel and M. Irani, “Separating transparent layers through layer information exchange,” in *Proceedings of European Conference on Computer Vision*, pp. 328–341 (2004)
48. B. Sarel and M. Irani, “Separating transparent layers of repetitive dynamic behaviors,” in *Proceedings of IEEE International Conference on Computer Vision*, pp. 26–32 (2005)
49. Y. Sato and K. Ikeuchi, “Reflectance analysis under solar illumination,” in *Proceedings of IEEE Workshop on Physics-Based Modeling and Computer Vision*, pp. 180–187 (1995)
50. I. Sato, T. Okabe, Q. Yu, and Y. Sato, “Shape reconstruction based on similarity in radiance changes under varying illumination,” in *Proc. International Conference on Computer Vision* (2007)
51. Y. Y. Schechner, N. Kiryati, and R. Basri, “Separation of transparent layers using focus,” *International Journal of Computer Vision*, vol. 39, no. 1, pp. 25–39 (2000)
52. Y. Y. Schechner and J. Shamir, “Polarization and statistical analysis of scenes containing a semireflector,” *Journal of the Optical Society of America A*, vol. 17, no. 2, pp. 276–284 (2000)
53. S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, “A theory of inverse light transport,” in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1440–1447 (2005)
54. L. Shen, T. Machida, and H. Takemura, “Efficient photometric stereo technique for three-dimensional surfaces with unknown BRDF,” in *Proc. Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pp. 326–333 (2005)
55. D. Simakov, D. Frolova, and R. Basri, “Dense shape reconstruction of a moving object under arbitrary, unknown lighting,” in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1202–1209 (2003)
56. F. Solomon and K. Ikeuchi, “Extracting the shape and roughness of specular lobe objects using four light photometric stereo,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 4, pp. 449–454 (1996)
57. J. Sun, Y. Li, S. B. Kang, and H. Y. Shum, “Flash matting,” *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 772–778 (2006)
58. J. Sun, M. Smith, L. Smith, S. Midha, and J. Bamber, “Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities,” *Image and Vision Computing*, vol. 25, no. 7, pp. 1050–1057 (2007)
59. R. Szeliski, S. Avidan, and P. Anandan, “Layer extraction from multiple images containing reflections and transparency,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 246–253 (2000)
60. P. Tan, S. Lin, and L. Quan, “Resolution-enhanced photometric stereo,” in *Proceedings of European Conference on Computer Vision*, pp. 58–71 (2006)
61. P. Tan, S. P. Mallick, L. Quan, D. J. Kriegman, and T. Zickler, “Isotropy, reciprocity and the generalized bas-relief ambiguity,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (2007)
62. K. Tang, C. Tang, and T. Wong, “Dense photometric stereo using tensorial belief propagation,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 132–139 (2005)
63. S. Tominaga, M. Nakagawa, and N. Tanaka, “Image rendering of art paintings -total archives considering surface properties and chromatic adaptation-,” in *Proceedings of Color Imaging Conference*, pp. 70–75 (2004)
64. R. Y. Tsai, “An efficient and accurate camera calibration technique for 3D machine vision,” in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 364–374 (1986)

65. P. Tu and P. R. S. Mendonca, "Surface reconstruction via Helmholtz reciprocity with a single image pair," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2003)
66. T. Wada, H. Ukida, and T. Matsuyama, "Shape from shading with interreflections under a proximal light source: Distortion-free copying of an unfolded book," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 125–135 (1997)
67. R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 19, no. 1, pp. 139–144 (1980)
68. R. J. Woodham, Y. Iwahori, and R. A. Barman, "Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength," in *Technical Report* (1991)
69. T. P. Wu and C. K. Tang, "Dense photometric stereo using a mirror sphere and graph cut," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 140–147 (2005)
70. T. P. Wu and C. K. Tang, "Dense photometric stereo by expectation maximization," in *Proceedings of European Conference on Computer Vision*, pp. 159–172 (2006)
71. T. P. Wu, K. L. Tang, C. K. Tang, and T. T. Wong, "Dense photometric stereo: a Markov random field approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1830–1846 (2006)
72. J. Yang, D. Zhang, N. Ohnishi, and N. Sugie, "Determining a polyhedral shape using interreflections," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 110–115 (1997)
73. A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur, "Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability," *International Journal of Computer Vision*, vol. 35, no. 3, pp. 203–222 (1999)
74. L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 618–625 (2003)
75. T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *International Journal of Computer Vision*, pp. 215–227 (2002)
76. T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Toward a stratification of Helmholtz stereopsis," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2003)

A Minimization of L2-Norm

Here, we solve the following minimization problem.

$$\iint \left(\rho_x(x,y)\right)^2 + \left(\rho_y(x,y)\right)^2 dxdy\,. \tag{24}$$

We obtain the following Euler-Lagrange differential equation [11].

$$\rho_{xx} + \rho_{yy} = 0\,, \tag{25}$$

constrained with Dirichlet boundary condition or Neumann boundary condition. We represent the discretization of second-order derivative as follows.

$$\rho_{xx}(x,y) = \rho(x-1,y) - 2\rho(x,y) + \rho(x+1,y) \tag{26}$$
$$\rho_{yy}(x,y) = \rho(x,y-1) - 2\rho(x,y) + \rho(x,y+1)\,. \tag{27}$$

As a result, Eq. (25) becomes the following well-known formula.

$$\rho(x,y) = \frac{1}{4}\left(\rho(x+1,y) + \rho(x-1,y) + \rho(x,y-1) + \rho(x,y+1)\right)\,. \tag{28}$$

Iteratively computing this formula for each pixel with Gauss-Seidel method under the boundary constraint minimizes Eq. (24).

B Minimization of L1-Norm

Here, we solve the following minimization problem:

$$\iint |\rho_x(x,y)| + |\rho_y(x,y)|\, dxdy\,. \tag{29}$$

We solve this problem heuristically. If we only concentrate on solving the ρ at pixel (x,y) when fixing the ρ of other pixels, we just have to minimize the following for each pixel.

$$|\rho_x(x,y)| + |\rho_y(x,y)|\ . \tag{30}$$

There are two possibilities for discretizing first-order derivative.

$$\rho_x(x,y) = \rho(x+1,y) - \rho(x,y)\,, \tag{31}$$

or

$$\rho_x(x,y) = \rho(x,y) - \rho(x-1,y)\,. \tag{32}$$

We use both of these equations in order to decrease the discretization error as much as possible. Therefore, we minimize the following formula instead of Eq. (30).

$$\begin{aligned}
&|\rho_x(x,y)| + |\rho_x(x,y)| + |\rho_y(x,y)| + |\rho_y(x,y)| \\
&= |\rho(x,y) - \rho(x+1,y)| + |\rho(x,y) - \rho(x-1,y)| + |\rho(x,y) - \rho(x,y+1)| + |\rho(x,y) - \rho(x,y-1)| \\
&= \sum_{i=1}^{4} |\rho(x,y) - \rho_i|\ ,
\end{aligned} \tag{33}$$

where

$$\rho_1 = \rho(x+1,y),\ \rho_2 = \rho(x-1,y),\ \rho_3 = \rho(x,y+1),\ \rho_4 = \rho(x,y-1)\,. \tag{34}$$

By minimizing Eq. (33), we obtain the following formula.

$$\rho(x,y) = \mathrm{median}\left(\rho(x+1,y), \rho(x-1,y), \rho(x,y+1), \rho(x,y-1)\right)\ . \tag{35}$$

Iteratively computing this formula minimizes Eq. (29).