Daisuke Miyazaki, Kenji Hara, Katsushi Ikeuchi,
"Photometric Stereo beyond Glass: Active Separation of Transparent Layer and Five-Light
Photometric Stereo with M-Estimator Using Laplace Distribution for a Virtual Museum,"
Proceedings of International Workshop on Photometric Analysis for Computer Vision,
Rio de Janeiro, Brazil, 2007.10

Photometric Stereo beyond Glass:
Active Separation of Transparent Layer and Five-Light Photometric Stereo with M-Estimator Using Laplace Distribution for a Virtual Museum

Daisuke Miyazaki[†], Kenji Hara[††], and Katsushi Ikeuchi[†]

[†]The University of Tokyo

Institute of Industrial Science, Komaba 4–6–1, Meguro-ku, Tokyo, 153–8505 Japan

[††]Kyushu University

Faculty of Design, Shiobaru 4–9–1, Minato-ku, Fukuoka, 815–8540 Japan

Abstract

One of the necessary techniques for constructing a virtual museum is to digitize the artwork arranged inside a glass or acrylic display case without bringing the artwork out of the display case. By using photometric stereo, we estimate the shape (surface normal) and the reflectance (albedo) of the artwork arranged inside a transparent display case. If we illuminate the display case, the light will reflect at its surface; thus, we cannot apply conventional photometric stereo as is. In this paper, we propose a five-light photometric stereo that estimates the shape and the reflectance of an object that has specularities under the circumstances that the light is reflected at the surface of the display case.

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.

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 [1–9]. 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.
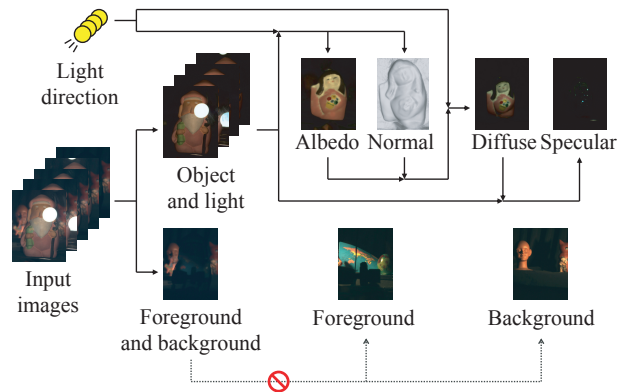

Figure 1. Flow diagram.

Our method assumes that the target object has both diffuse reflection and specular reflection.

The flow of the proposed algorithm will go as shown in Fig. 1. 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. 2). 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 [10–12]. 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 need to extend the four-light photometric stereo so that it can be applied to the objects inside the display case, and we denote the extended algorithm as "five-light photometric stereo." The prototype of our measurement device is illustrated in Fig. 3.

Some photometric stereo methods use multiple images to enhance the quality of the output. Hayakawa [13] proposed
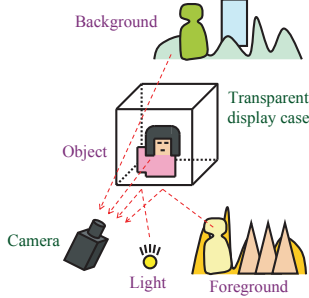
Figure 2. Foreground and light reflected, and background and object transmitted through a transparent display case.
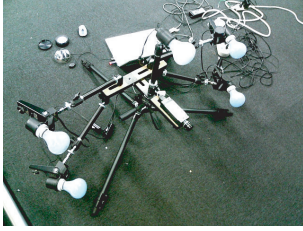

Figure 3. Prototype of the five-light camera.

an uncalibrated photometric stereo using singular value decomposition (SVD). Yuille et al. [14] analyzed Hayakawa's method, and proposed a method to solve the bas-relief ambiguity. Basri et al. [15] used spherical harmonics to apply uncalibrated photometric stereo to arbitrary illumination. Tan et al. [22] enhanced the resolution of the surface normal by using the photometric stereo. Chandraker et al. [23] removed the shadow by using the graph cut method, and estimated the surface normal robustly from only four images. Some photometric stereo methods can be applied to non-Lambertian objects [16, 17]. Goldman et al. [18] estimated the BRDF (bidirectional reflectance distribution function) of the object's surface using photometric stereo, under the condition that the BRDF can be categorized in two types or combination of those types. Helmholtz stereo [19, 20] applied stereo and photometric stereo to a non-Lambertian model. Photometric Sampler can also estimate the surface normal of specular objects [21]. In this paper, we propose a robust photometric stereo method for Lambertian object under the condition that only a small number of images is supplied. Such sparse photometric stereo is useful in a wide field of applications such as a robot's eye for object recognition, obtaining the surface normal in a narrow area, developing a commercial camera for entertainment, and so on.

We describe the separation method of the reflection in Section 2. Also, we discuss the advantage and the disadvantage of the separation. We propose the five-light photometric stereo in Section 3. We show that five lights are sufficient for measuring the object in a display case by using photometric stereo. We present some experimental results in Section 4, and conclude the paper in Section 5.

2. Active Separation of Transparent Layers

As is shown in Fig. 2, the observed image includes foreground, background, light, and object. The scene in front of the display case will be reflected, and is observed as a foreground. The light is also reflected at the display case. We also observe the scene behind the display case, defined as a background. Also, the object is affected by ambient light. In addition, some noise occurs due to the dark current of the camera. Moreover, some amount of the light reaches to the background. By considering these effects, the observed image can be reformulated as follows:

$$I_{\text{on}} = I_{\text{foreground}} + I_{\text{background}} + I_{\text{background,light}}$$
$$+ I_{\text{light}} + I_{\text{object,light}} + I_{\text{object,ambient}} + I_{\text{noise}} . \quad (1)$$

We assume the camera noise can be represented as a white noise (Gaussian noise). Therefore, by taking multiple images and calculating the average, the noise 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

$$I_{\text{off}} = I_{\text{foreground}} + I_{\text{background}}$$
$$+ I_{\text{object,ambient}} + I_{\text{noise}} . \quad (2)$$

Then, if we calculate the difference between Eq. (1) and Eq. (2), we obtain

$$I_{\text{on}} - I_{\text{off}} = I_{\text{light}} + I_{\text{object,light}} + I_{\text{background,light}} . \quad (3)$$

The background does not affect the appearance of the object; thus, only the object and the light are obtained at the pixels in object region. We can find out that the image represented by Eq. (3) does not have noise, effect of ambient light, foreground, or background. The light might reach the foreground by reflecting at the transparent surface of the display case; however, this can be ignored because the light is attenuated after traveling a long distance. The result of the separation is shown in Fig. 4.

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 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 [24–28]. Therefore, we use this method for separation.

The image represented by Eq. (3) contains not only the object but also the light; thus, we cannot apply conventional photometric stereo. The direct reflection of the light on the display surface is too bright, and the ordinary polarizers cannot completely remove the reflection. In our experiment, we avoid saturation of the diffuse reflection component; however, the direct reflection of the light often causes saturation. The separation procedure works successfully even for saturated pixels. Section 3 presents a method to estimate the surface normal and the albedo even if reflection of the light is saturated. If we need to analyze the specular reflection of the object, we carefully avoid its saturation.
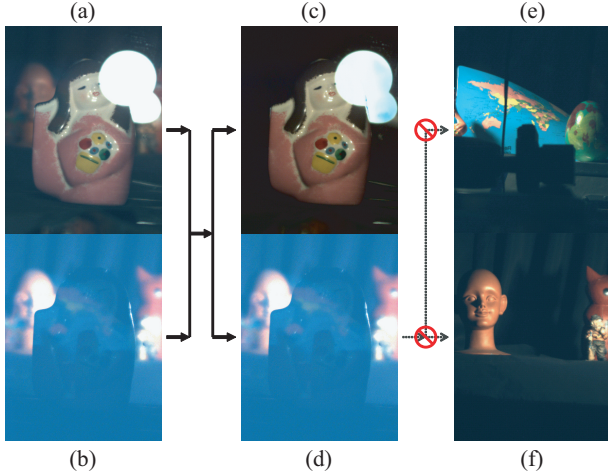
(a)   (c)   (e)

(b)   (d)   (f)

Figure 4. Separation result: (a) Input image with the light on, (b) input image with the light off (intensity is multiplied by 4 for visibility), (c) separated image with the light and the object, (d) separated image with the foreground and the background (intensity is multiplied by 4 for visibility), (e) foreground image (for reference), (f) background image (for reference).

3. Photometric Stereo with Five Lamps

3.1. Five-Light Photometric Stereo

Photometric stereo is effectively used for digital archiving of artworks [29, 30]. Conventional photometric stereo cannot be applied to an object that has specular reflection; moreover, it cannot be applied if there is a reflection at the glass case.

In this paper, we follow the idea of the four-light photometric stereo [10–12]. Conventional photometric stereo solves the equation from three images taken under three different light sources. Let us increase the number of the light sources to four. Four combinations can be calculated if we choose three intensities from four intensities. From the chosen three intensities, the conventional photometric stereo can estimate the surface normal. Therefore, there are four candidates for the surface normal at each pixel. If all four intensities are caused by diffuse reflection, the four surface normals will be close together. If one of the intensities includes specular reflection, the four surface normals will be different. In this case, the surface normal can be calculated from the darkest three intensities. This is the basic idea of the four-light photometric stereo proposed by Coleman and Jain [10].

Considering the reflection at the glass case, five lights will be enough for the measurement (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) are the Gaussian spheres when the angle between the light source and the viewing direction is 30°, 45°, 60°, and 75°, respectively. The red circle rep-
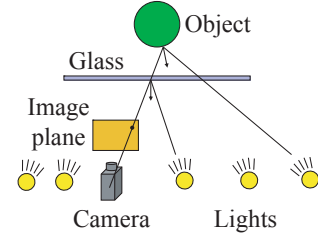
Figure 5. The principle of measuring the surface normal by photometric stereo using five light sources.
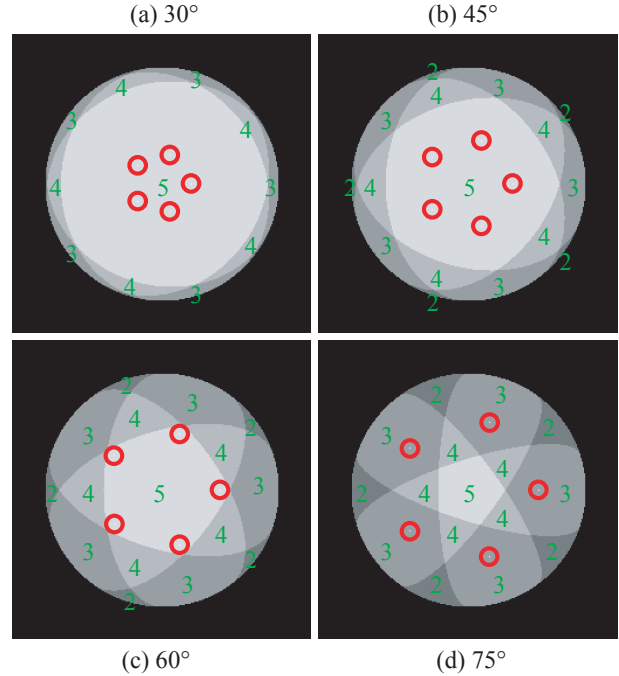

(a) 30°   (b) 45°

(c) 60°   (d) 75°

Figure 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°.

resents the specular reflection of the object's surface. The numbers represent the numbers of the lights that illuminate the surface. In case "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 appeared in region "5," we have three input data sets to calculate the surface normal. Region "2," which is illuminated only by two lights, will appear 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. In addition, we do not know which region each pixel belongs to, when the surface normal is unknown. Therefore, we propose an iterative algorithm in Section 3.2.

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.

3.2. Robust Iterative Computation

3.2.1 M-Estimation Using Laplace Distribution

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

First, we explain the smoothness constraint of the albedo. By minimizing

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

we obtain the following albedo ρ

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

Eq. (5) is derived by differentiating Eq. (4) with ρ at pixel (x,y) fixing other pixels, and setting "=0."

Eq. (4) uses L2-norm, which means that it uses Gaussian distribution as M-estimator [31]. This M-estimator is not robust to outliers. We use Laplace distribution (double exponential distribution) as M-estimator, which is more robust than Gaussian distribution [31]. In this case, the smoothness constraint of the albedo is represented as the following L1-norm.

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

The albedo ρ which minimizes this equation would be

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

Eq. (7) is derived by differentiating Eq. (6) with ρ at pixel (x,y) fixing other pixels, and setting "=0". We present the proof in the Appendix.

Next, we explain the optimization of the albedo. We represent the L input images with the subscript $i = \{1, \ldots, L\}$.

Since there are some outliers like shadows, we use Laplace distribution as the M-estimator.

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

By fixing the surface normal N, the albedo ρ which minimizes Eq. (8) will be given as follows.

$$\rho = \text{median}\Big(\frac{I_i}{\cos L_i \cdot N} \mid i = 1, \ldots, L \Big) . \tag{9}$$

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 $_L C_3 = \binom{L}{3}$. We represent each surface normal estimated from each combination as $N_m (m = 1, \ldots, {_L C_3})$. Each surface normal is calculated by conventional photometric stereo [32]. We estimate the surface normal by minimizing the following formula.

$$\varepsilon_{\text{N,opt}} = \sum_{m=1}^{_L C_3} |N - N_m| . \tag{10}$$

By minimizing Eq. (10), we obtain the following solution.

$$N = \text{median}\big(N_m \mid m = 1, \ldots, {_L C_3} \big) . \tag{11}$$

3.2.2 Proposed Algorithm

Let us organize the algorithm. First, we set the appropriate initial values for the surface normal and the albedo. Next, we calculate all $_L C_3$ candidates of the surface normal for all pixels. Then, we iterate the following process until convergence.

1. We estimate L numbers of albedo from the current surface normal. We also add the albedos of four neighborhoods. We calculate the median ρ_{opt} of these albedos.
2. We calculate the four-neighbor average of the albedo, ρ_{avg}. We calculate the new albedo from the weighted average of ρ_{opt} and ρ_{avg}.
3. We calculate the median N_{opt} of the $_L C_3$ candidates of surface normal and the surface normals of four neighborhoods.
4. We calculate the four-neighbor average of the surface normal, N_{avg}, and calculate the new surface normal from the weighted average of N_{opt} and N_{avg}.

Finally, the surface normal is integrated to the height map using the natural boundary condition [33]. The surface normal and the albedo are updated using the four-neighbor information in order to avoid falling into a local minimum. RANSAC 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 $_L C_3$ candidates and remove the outliers more robustly. If the number of the images are large, calculating all candidates takes a long time
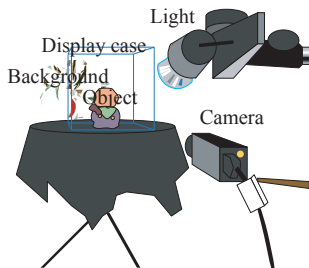
Figure 7. Experimental setup.

$(_LC_3 = L(L-1)(L-2)/6 \simeq O(L^3))$; thus, RANSAC is more effective than our approach in such a case.

4. Experimental Result

4.1. Experimental Setup

The experimental setup is shown in Fig. 7. The target object is covered with a glass or acrylic display case, and is observed with one camera and one lamp. The camera is fixed, and we take the image with three-band RGB. The light source direction and its intensity are assumed to be known. We move the light source, and 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.

4.2. Experimental Result

Fig. 8 is the input image and Fig. 9 is the separated result. Fig. 10(a)–(c) are the estimated albedo, height, and shape, respectively. Fig. 11 is an example image rendered by using estimated parameters with a different viewpoint and different light direction from the acquisition stage. In Fig. 10(c), we did not remove the background; however, it can be easily removed by simply thresholding the dark intensity of the background, or using background cut techniques such as Lazy Snapping [34], GrabCut [35], or Adobe Photoshop CS3. Both results in Fig. 10 are obtained from five input images.

Fig. 12 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 intensity of the light source. Calibrated photometric stereo assumes the light source direction and the intensity
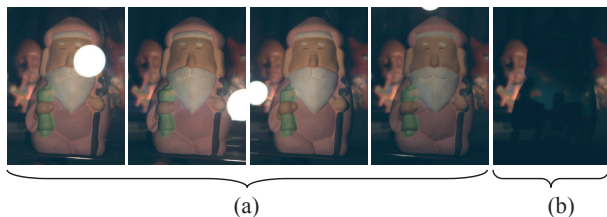

Figure 8. Input images: (a) The images taken with the light on, (b) one of the images taken with the light off.
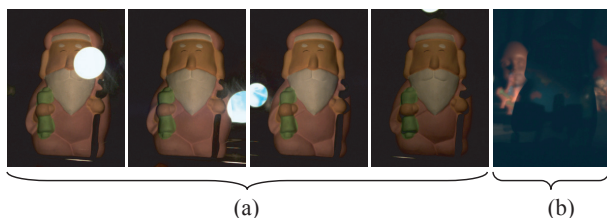

Figure 9. Separation result: (a) The images that only contain the object and the light, (b) one of the images that contain the foreground and the background.
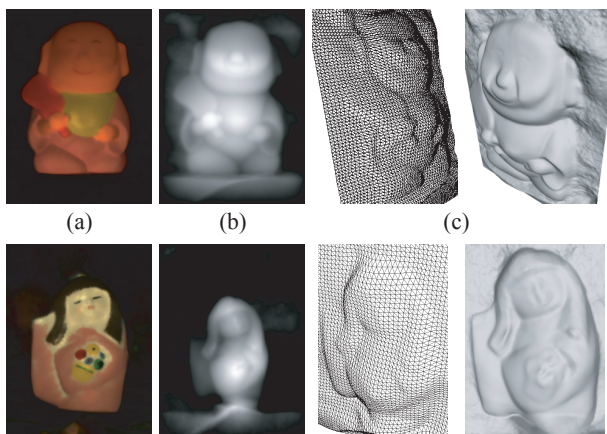

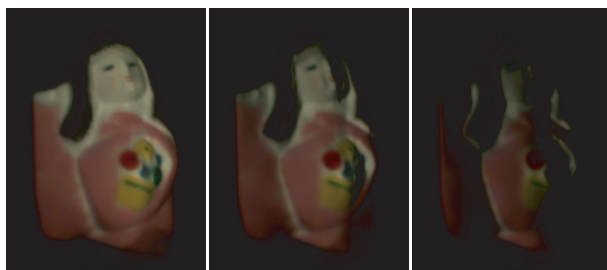Figure 10. Estimated result: (a) Albedo, (b) height, (c) shape.

Figure 11. Rendered image with a different viewpoint and light direction from the acquisition stage.

are known. Although we compensated for these factors before the experiment, some amount of the error of the compensation still remain as a noise.

Fig. 13 shows the result of light-stripe range scanning (active stereo). We scan from exactly the front of the acrylic
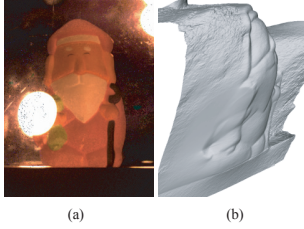
Figure 12. Result of conventional photometric stereo: (a) Albedo, (b) shape.
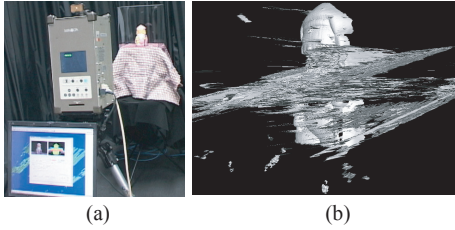

Figure 13. The result of light-stripe range scanning: (a) Setup, (b) estimated shape.

box with a Konica Minolta VIVID 910. Owing to the 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. 14(b) shows the result of our method applied to an object 15cm tall, from six input images. Fig. 14(a) is a ground truth obtained by scanning the object brought out from the acrylic case by VIVID910. Fig. 14(c) is the result of conventional photometric stereo from three certain input images. Fig. 14(d) is the same as Fig. 13. The comparison in Table 1 shows that the proposed method is more effective than the other methods. Fig. 14(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. 14(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.

5. 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 artwork that is taken out of the case. In this paper, we improve photometric stereo so that it will be suited for this application.
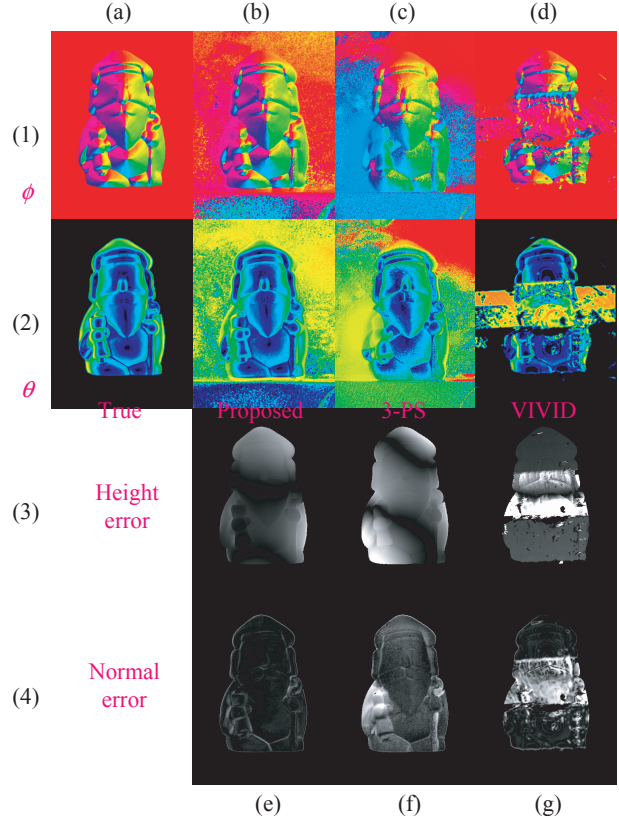

Figure 14. Comparison: (a) True value, (b) the result of the proposed method, (c) the result of conventional photometric stereo, (d) the result of light-stripe range scanning, (e) the error of the proposed method, (f) the error of conventional photometric stereo, (g) the error of light-stripe range scanning; (1) Azimuth angle (red: upper direction, blue: left-bottom direction, green: right-bottom direction) (2) zenith angle (blue: $0°$, red: $90°$) (3) the difference of the height (the brighter the noisier), (4) the difference of the surface normal (the brighter the noisier).

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°

One of our future goals is to estimate specular reflection parameters. We can render the diffuse reflection component as is shown in Fig. 15(b). We can obtain the specular image (Fig. 15(c)) by subtracting the rendered image (Fig. 15(b)) from the input image (Fig. 15(a)). The specular pixels were saturated in this experiment; however, we can estimate the specular reflection parameter if we adjust the exposure adequately.

Binocular stereo cannot be applied if there is a reflec-

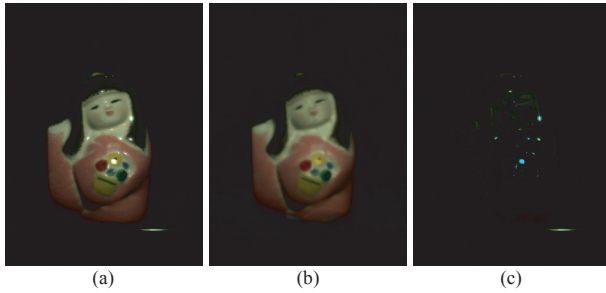<div align="center">(a)   (b)   (c)</div>

Figure 15. Rendering result of diffuse and specular reflection: (a) Input image, (b) rendered image of diffuse reflection, (c) calculated image of specular reflection.

tion 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; however, our method can estimate the object's shape even if the object is observed from the front of the display case. 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 [36, 37].

We propose an algorithm that can be applied to the input image taken with a machine like Fig. 3. Therefore, the method is adequate when the number of lights is small. If we increase the number of lights, we can obtain much more information [13, 14, 42, 43]. Our method assumes that the diffuse reflection can be modeled as a Lambertian model; and extending our method to other reflection models is also important. The quality of the shape will increase by considering the interreflection, thus; it might be improved by applying the solutions suggested by [38–41].

The proposed method can be applied also to an object which is not covered by a display case, and can be applied also to specular objects. Our optimization algorithm is robust to noise and outliers. 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.

References

[1] H. Farid, E.H. Adelson, "Separating reflections from images by use of independent component analysis," *J. Opt. Soc. Am. A*, **16**(9), 2136-2145, 1999.

[2] Y.Y. Schechner, J. Shamir, "Polarization and statistical analysis of scenes containing a semireflector," *J. Opt. Soc. Am. A*, **17**(2), 276-284, 2000.

[3] R. Szeliski, S. Avidan, P. Anandan, "Layer extraction from multiple images containing reflections and transparency," *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, 246-253, 2000.

[4] Y.Y. Schechner, N. Kiryati, R. Basri, "Separation of trans-

parent layers using focus," *Int'l J. Comput. Vis.*, **39**(1), 25-39, 2000.

[5] B. Sarel, M. Irani, "Separating transparent layers through layer information exchange," *Proc. IEEE Eur. Conf. Comput. Vis.*, 328-341, 2004.

[6] A. Levin, Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *Proc. IEEE Eur. Conf. Comput. Vis.*, 602-613, 2004.

[7] A. Levin, A. Zomet, Y. Weiss, "Separating reflections from a single image using local features," *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, 306-313, 2004.

[8] B. Sarel, M. Irani, "Separating transparent layers of repetitive dynamic behaviors," *Proc. IEEE Int'l Conf. Comput. Vis.*, 26-32, 2005.

[9] T. Oo, H. Kawasaki, Y. Ohsawa, K. Ikeuchi, "The separation of reflected and transparent layers from real-world image sequence," *J. Mach. Vis. Appl.*, **18**(1), 17-24, 2007.

[10] E.N. Coleman Jr., R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," *Comput. Graph. Image Process.*, **18**(4), 309-328, 1982.

[11] F. Solomon, K. Ikeuchi, "Extracting the shape and roughness of specular lobe objects using four light photometric stereo," *IEEE Trans. Patt. Anal. Mach. Intell.*, **18**(4), 449-454, 1996.

[12] S. Barsky, M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Trans. Patt. Anal. Mach. Intell.*, **25**(10), 1239-1252, 2003.

[13] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *J. Opt. Soc. Am. A*, **11**(11), 3079-3089, 1994.

[14] A.L. Yuille, D. Snow, R. Epstein, P.N. Belhumeur, "Determining generative models of objects under varying illumination: shape and albedo from multiple images using SVD and integrability," *Int'l J. Comput. Vis.*, **35**(3), 203-222, 1999.

[15] R. Basri, D. Jacobs, I. Kemelmacher, "Photometric stereo with general, unknown lighting," *Int'l J. Comput. Vis.*, **72**(3), 239-257, 2007.

[16] H. Ragheb, E.R. Hancock, "Surface normals and height from non-Lambertian image data," *Proc. 3DPVT*, 18-25, 2004.

[17] P. Tan, S.P. Mallick, L. Quan, D.J. Kriegman, T. Zickler, "Isotropy, reciprocity and the generalized bas-relief ambiguity," *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, 1-8, 2007.

[18] D. Goldman, B. Curless, A. Hertzmann, S.M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," *Proc. IEEE Int'l Conf. Comput. Vis.*, 341-348, 2005.

[19] S. Magda, D.J. Kriegman, T. Zickler, P.N. Belhumeur, "Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs," *Proc. IEEE Int'l Conf. Comput. Vis.*, 391-398, 2001.

[20] T. Zickler, P.N. Belhumeur, D.J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *Int'l J. Comput. Vis.*, 215-227, 2002.

[21] S.K. Nayar, K. Ikeuchi, T. Kanade, "Determining shape and reflectance of hybrid surface by photometric sampling," *IEEE Trans. Rob. Autom.*, **6**(4), 418-431, 1990.

[22] P. Tan, S. Lin, L. Quan, "Resolution-enhanced photometric stereo," *Proc. IEEE Eur. Conf. Comput. Vis.*, 58-71, 2006.

[23] M. Chandraker, S. Agarwal, D. Kriegman, "ShadowCuts: Photometric stereo with shadows," *Proc. IEEE Conf. Comput. Vis. Patt. Recognit.*, 2007.

[24] R. Raskar, K.H. Tan, R. Feris, J. Yu, M. Turk, "Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging," *ACM Trans. Graph.*, **23**(3), 679-688, 2004.

[25] J. Sun, Y. Li, S.B. Kang, H.Y. Shum, "Flash matting," *ACM Trans. Graph.*, **25**(3), 772-778, 2006.

[26] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. Graph.*, **23**(3), 664-672, 2004.

[27] E. Eisemann, F. Durand, "Flash photography enhancement via intrinsic relighting," *ACM Trans. Graph.*, **23**(3), 673-678, 2004.

[28] C. Lu, M.S. Drew, "Practical scene illuminant estimation via flash/no-flash pairs," *Proc. Color Imaging Conf.*, 2006.

[29] S. Tominaga, M. Nakagawa, N. Tanaka, "Image rendering of art paintings -total archives considering surface properties and chromatic adaptation-," *Proc. 12th Color Imaging Conf.*, 70-75, 2004.

[30] D.V. Hahn, D.D. Duncan, K.C. Baldwin, J.D. Cohen, B. Purnomo, "Digital Hammurabi: design and development of a 3D scanner for cuneiform tablets," *Proc. SPIE*, **6056**, 130-141, 2006.

[31] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical recipes in C: the art of scientific computing*, 994, Cambridge University Press, 1992.

[32] R.J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.*, **19**(1), 139-144, 1980.

[33] K. Ikeuchi, "Reconstructing a depth map from intensity maps," *Proc. Int'l Conf. Patt. Recognit.*, 736-738, 1984.

[34] Y. Li, J. Sun, C.K. Tang, H.Y. Shum, "Lazy Snapping," *ACM Trans. Graph.*, **23**(3), 2004.

[35] C. Rother, V. Kolmogorov, A. Blake, "GrabCut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, **23**(3), 2004.

[36] J. Lim, J. Ho, M.H. Yang, D. Kriegman, "Passive photometric stereo from motion," *Proc. IEEE Int'l Conf. Comput. Vis.*, 1635-1642, 2005.

[37] L. Zhang, B. Curless, A. Hertzmann, S.M. Seitz, "Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo," *Proc. IEEE Int'l Conf. Comput. Vis.*, 2003.

[38] S.K. Nayar, K. Ikeuchi, T. Kanade, "Shape from interreflections," *Int'l J. Comput. Vis.*, **6**(3), 173-195, 1991.

[39] T. Wada, H. Ukida, T. Matsuyama, "Shape from shading with interreflections under a proximal light source: distortion-free copying of an unfolded book," *Int'l J. Comput. Vis.*, **24**(2), 125-135, 1997.

[40] S.M. Seitz, Y. Matsushita, K.N. Kutulakos, "A theory of inverse light transport," *Proc. IEEE Int'l Conf. Comput. Vis.*, 1440-1447, 2005.

[41] S.K. Nayar, G. Krishnan, M.D. Grossberg, R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," *ACM Trans. Graph.*, **25**(3), 935-944, 2006.

[42] D. Simakov, D. Frolova, R. Basri, "Dense shape reconstruction of a moving object under arbitrary, unknown lighting," *Proc. IEEE Int'l Conf. Comput. Vis.*, 1202-1209, 2003.

[43] M.S. Drew, "Reduction of rank-reduced orientation-from-color problem with many unknown lights to two-image known-illuminant photometric stereo," *Proc. IEEE Int'l Symp. Comput. Vis.*, 419-424, 1995.

A. Maximum Likelihood Estimation of Laplace Distribution

Eq. (6) can be discretized as follows:

$$\varepsilon_{\rho,\mathrm{med}} = |\rho(x+1, y) - \rho(x,y)| + |\rho(x,y) - \rho(x-1,y)|$$
$$+ |\rho(x, y+1) - \rho(x,y)| + |\rho(x,y) - \rho(x,y-1)| \; . \quad (12)$$

By applying the following theorem in minimizing Eq. (12), we obtain Eq. (7).

Theorem A.1 *Suppose that we want to estimate the parameter p from data d_i by maximizing the following likelihood:*

$$\Pr = \prod_{i=1}^{n} \frac{1}{2b} \exp\left(-\frac{|p - d_i|}{b}\right) \; . \quad (13)$$

This probability density function is called Laplace distribution, and b is called the scale parameter. The parameter p of this solution will be the following:

$$p = \mathrm{median}\big(d_i \mid i = 1, \ldots, n\big) \; . \quad (14)$$

Proof. By calculating the logarithm of Eq. (13), we obtain the following formula:

$$\mathrm{pr} = -n \log 2b - \frac{1}{b} \sum_{i=1}^{n} |p - d_i| \; . \quad (15)$$

Therefore, maximizing Eq. (13) results in minimizing the following formula:

$$\varepsilon = \sum_{i=1}^{n} |p - d_i| \; . \quad (16)$$

Let us express the number of the data d_i which is smaller than p as j, and the number of the data d_i which is greater than p as k, where $j + k = n$. Therefore, Eq. (16) can be expressed as follows:

$$\varepsilon = \sum_{i=1}^{j} (p - d_i) + \sum_{i=j+1}^{j+k} (d_i - p) \; , \quad (17)$$

by representing the subscript properly. By differentiating Eq. (17) by p, we obtain

$$\frac{\partial \varepsilon}{\partial p} = \sum_{i=1}^{j} - \sum_{i=j+1}^{j+k} = j - k \; . \quad (18)$$

By letting Eq. (18) be zero, we obtain $j = k$, namely, the number of the data d_i smaller than p and the number of the data d_i greater than p is the same. Therefore, Eq. (14) is obtained. □