Surface Normal Deconvolution: Photometric Stereo for Optically Thick Translucent Objects

Chika Inoshita[1], Yasuhiro Mukaigawa[2]
Yasuyuki Matsushita[3], and Yasushi Yagi[1]

[1] Osaka University, Japan
[2] Nara Institute of Science and Technology, Japan
[3] Microsoft Research Asia, China

Abstract. This paper presents a photometric stereo method that works for optically thick translucent objects exhibiting subsurface scattering. Our method is built upon the previous studies showing that subsurface scattering is approximated as convolution with a blurring kernel. We extend this observation and show that the original surface normal convolved with the scattering kernel corresponds to the blurred surface normal that can be obtained by a conventional photometric stereo technique. Based on this observation, we cast the photometric stereo problem for optically thick translucent objects as a deconvolution problem, and develop a method to recover accurate surface normals. Experimental results of both synthetic and real-world scenes show the effectiveness of the proposed method.

1 Introduction

Photometric stereo estimates the surface normals of a scene from multiple shading images taken under different lighting conditions [1]. While conventional methods are developed for simple Lambertian diffuse surfaces [2], recent generalizations can handle more complex reflections in real-world scenes [3, 4]. However, surface normal estimation of translucent materials is still a difficult task, where subsurface scattering is significant [5].

In a translucent object, incident light travels randomly and exits from various neighboring locations. This global light transport effect makes it hard to directly associate the shading observations with its surface geometry. As a result, shape-from-intensity techniques that only assume local illumination models naturally suffer from the unmodeled error of subsurface scattering. One of the directions to address this issue is to remove the subsurface scattering component from the observations, therefore conventional shape-from-intensity techniques can be used to the remaining direct lighting component. Recently, Nayar *et al.* [6] have demonstrated an approach to effectively remove subsurface scattering from the scene observations, with an expense of additional measurements under high-frequency illuminations. In general, removing the subsurface scattering component requires

D. Fleet et al. (Eds.): ECCV 2014, Part II, LNCS 8690, pp. 346–359, 2014.
© Springer International Publishing Switzerland 2014

an additional preprocessing stage, and a shape-from-intensity method that can directly account for subsurface scattering is wanted.

While exact modeling of subsurface scattering is still a difficult task that requires complicated models, prior studies in the field of computer graphics show that the image formation model of subsurface scattering can be well approximated as convolution of the scattering kernel and surface radiance on optically thick materials, which distribute light regardless of the incident directions [7]. We use this approximation to develop *surface normal deconvolution*, which recovers original surface normal from the *blurry* surface normal obtained by conventional photometric stereo on translucent objects. This idea is similar to Dong *et al.*'s method [8], which estimates surface normal by deconvolved input images to remove the subsurface scattering effect. While Dong *et al.* assume parametric subsurface scattering, *i.e.*, photon beam diffusion of optically homogeneous media, we represent subsurface scattering by non-parametric convolution kernels for either optically homogeneous or inhomogeneous media. The convolution kernels can be either calibrated or estimated, and various deconvolution techniques in the literature (such as image deblurring methods) can be used for the implementation to recover deblurred surface normal. We show estimation results by both our deconvolution formulation and existing deconvolution in experiments.

2 Related Works

Conventional photometric stereo methods recover surface normals at a pixel-level detail based on local illumination models. While the original work of Woodham [1] uses a simple Lambertian reflection model [2], more recent approaches make various generalizations by explicitly accounting for more flexible reflectance models [9, 10], or by robust estimation framework [3, 4]. These methods are shown effective; however, for translucent objects, global light interactions need to be accounted for to achieve accurate surface normal estimation. A seminal work of Nayar *et al.* [11] explicitly takes interreflections into account, which are global light transports among opaque surfaces. While the problem is similar, subsurface scattering remains as an un-addressed light transport effect in shape-from-intensity methods.

Recently, structured-light methods for measuring the shape of translucent surfaces are proposed. To reduce effect of subsurface scattering, combination of polarizers and phase shifting [12], multiplication of low-frequency and high-frequency projection pattern [13], and high frequency illumination with multiplexed light source [14] have been used. In addition, Gupta *et al.* [15] use several binary projection codes to decrease estimation errors caused by subsurface scattering and interreflections. These techniques are shown effective, with an expense of specialized hardware setups.

Modeling subsurface scattering has been more studied in computer graphics as bidirectional scattering surface reflection distribution function (BSSRDF). Although the general BSSRDFs can represent various translucent appearances, it is difficult to exactly model BSSRDFs because of its high-dimensionality. Hence,

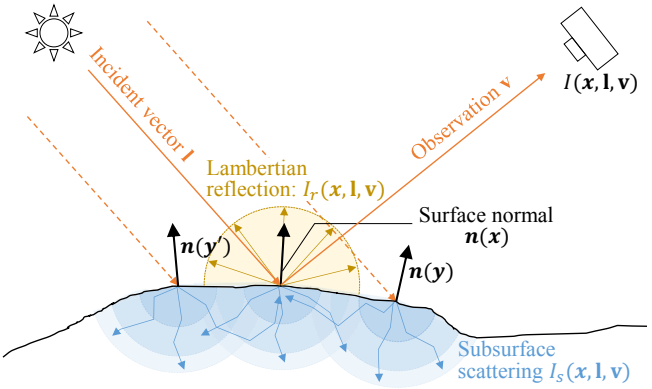


Fig. 1. Light interactions on a translucent surface. Incident light is partially reflected off the surface, while the rest of the light transmits and spreads inside the subsurface.

researchers previously approximate BSSRDFs as a low dimensional function with an assumption of homogeneous media [16], or isotropic scattering based on the diffusion theory [17–19]. In our work, we also approximate subsurface scattering by a simple model with an assumption of optically thick materials. In optically thick materials, incident light repeatedly scatters and loses its directionality, and as a result, the scattering strength becomes invariant to the illumination and observation directions. Based on this characteristics, subsurface scattering is approximated by a convolutional model in [7]. Our method is built upon this observation and extends it to develop a photometric stereo method for optically thick translucent objects.

3 Convolutional Image Formation Model

We begin with the image formation model for a translucent surface. When light illuminates a translucent surface, it is partially reflected, transmitted and also absorbed as depicted in Fig. 1. A portion of the transmitted light comes back to the surface via subsurface scattering; thus, the radiance $I(x, \mathbf{l}, \mathbf{v})$ at a scene point x with incident vector $\mathbf{l}$ and observation vector $\mathbf{v}$ becomes the sum of the reflection $I_r(x, \mathbf{l}, \mathbf{v})$ and subsurface scattering $I_s(x, \mathbf{l}, \mathbf{v})$ components as

$$I(x, \mathbf{l}, \mathbf{v}) = I_r(x, \mathbf{l}, \mathbf{v}) + I_s(x, \mathbf{l}, \mathbf{v}). \tag{1}$$

The subsurface scattering component $I_s(x, \mathbf{l}, \mathbf{v})$ is modeled as [17]

$$I_s(x, \mathbf{l}, \mathbf{v}) = \gamma(x)\ F(\mathbf{v}, \mathbf{n}(x), \eta) \int_{y \in A} R(x, y) F(\mathbf{l}, \mathbf{n}(y), \eta) \mathbf{n}(y)^T \mathbf{l} dy, \tag{2}$$

where $\gamma(x)$ is a scale factor for the subsurface scattering component, F represents Fresnel transmission, and $\mathbf{v}, \mathbf{n}, \mathbf{l} \in \mathbb{R}^3$ are the obervation, surface normal,

and incident vectors, respectively. η is a refractive index, $R(x, y)$ represents an extinction term for light from scene point x to its neighbor y such as dipole model [17], and A defines a neighboring area. Generally, the subsurface scattering component describes a nonlinear relation between the surface normals and observed intensity due to the Fresnel transmission term. To relax this complexity, we approximate the original model in a simpler form by assuming an optically thick material, as in [7]. On the surface of an optically thick material, subsurface scattering does not depend on the direction of the light, because the transmitted light scatters uncountable times and loses its directionality as same as diffusion approximation. Thus, subsurface scattering is invariant to the incident direction and outgoing direction, and the Fresnel terms F can be regarded as constants on optically thick materials. As a result the subsurface scattering component $I_s(x, \mathbf{l}, \mathbf{v})$ is simplified to

$$I_s(x, \mathbf{l}) = \gamma'(x) \int_{y \in A} R(x, y)\mathbf{n}(y)^T \mathbf{l} dy, \tag{3}$$

where $\gamma'(x)$ is a new scale factor of subsurface scattering that includes constant Fresnel transmission terms.

Assuming a Lambertian reflectance model for the reflection component $I_r(x, \mathbf{l}) = \rho(x)\mathbf{n}(x)^T\mathbf{l}$ with a diffuse albedo $\rho(x)$, the intensity observation $I(x, \mathbf{l}, \mathbf{v})$ can be written as

$$I(x, \mathbf{l}) = \left(\rho(x)\mathbf{n}(x) + \gamma'(x) \int_{y \in A} R(x, y)\mathbf{n}(y) dy \right)^T \mathbf{l}. \tag{4}$$

The first factor of Eq. (4) can be regarded as a simple convolution model as

$$I(x, \mathbf{l}) = \left(\int_{y \in A} h(x, y)\mathbf{n}(y) dy \right)^T \mathbf{l} = (h * \mathbf{n}(x))^T \mathbf{l}, \tag{5}$$

where $*$ is the convolution operation, the kernel h represents a scattering effect for the surface normals as

$$h(x, y) = \rho(x)\delta(x - y) + \gamma'(x)R(x, y). \tag{6}$$

A similar convolutional approximation of subsurface scattering is also discussed in the work of Munoz *et al.* [7] for the forward rendering of optically thick materials. This method is also inspired by the works of convolutional approximated subsurface scattering by d'Eon *et al.* [18] for the rendering of human skin and Donner *et al.* [16] for multi-layered materials. Unlike their method where the extinction term $R(x, y)$ is defined as a function parameterized only by the relative position of x and y, our method allows more flexibility for the extinction term $R(x, y)$ so that inhomogeneous translucent materials can also be handled.

4 Solution Method

Based on the convolutional image formation model, we develop a photometric stereo method for estimating the surface normals of an optically thick translucent surface. Our input is the same as traditional photometric stereo: A set of images is taken under varying lighting directions from a fixed view point. To simplify the discussion, we assume that the light directions are calibrated and the observations do not include shadows. In the rest of the paper, we work in the discretized pixel sites u and v that correspond to scene points x and y, respectively; thus Eq. (5) becomes

$$I(u, \mathbf{l}) = (h(u, v) * \mathbf{n}(u))^T \mathbf{l}. \tag{7}$$

The convolution equation Eq. (7) has a simple linear algebraic expression as

$$\mathbf{D} = \mathbf{HNL}, \tag{8}$$

where $\mathbf{D} \in \mathbb{R}^{m \times k}$ is an observation matrix, m and k are the numbers of pixels and light directions, respectively, $\mathbf{H} \in \mathbb{R}^{m \times m}$ is a scattering matrix, $\mathbf{N} \in \mathbb{R}^{m \times 3}$ is a surface normal matrix, and $\mathbf{L} \in \mathbb{R}^{3 \times k}$ is an incident light matrix, which is assumed to be known. This linear expression indeed has a similarity to the Lambertian photometric stereo [1], where the observation $\mathbf{D}$, scaled surface normal $\mathbf{N}_s$, and light matrix $\mathbf{L}$ has the following relationship:

$$\mathbf{D} = \mathbf{N}_s \mathbf{L}. \tag{9}$$

From Eqs. (8) and (9), we can see that the scaled surface normal $\mathbf{N}_s$ corresponds to $\mathbf{HN}$ as

$$\mathbf{N}_s = \mathbf{HN}. \tag{10}$$

Therefore, we could regard the scaled surface normal $\mathbf{N}_s$ as a *blurry* version of the original surface normal $\mathbf{N}$ that we wish to estimate. In the following we call $\mathbf{N}_s$ a smoothed surface normal.

Based on this observation, we estimate the surface normal $\mathbf{N}$ by taking the following two-step approach: (a) Obtain the smoothed surface normal $\mathbf{N}_s$ by Lambertian photometric stereo [1], (b) Estimate surface normal $\mathbf{N}$ in a deconvolution framework using the subsurface scattering matrix $\mathbf{H}$.

(a) Estimation of smoothed surface normal $\mathbf{N}_s$. We use a conventional Lambertian photometric stereo [1] for deriving the smoothed surface normal $\mathbf{N}_s$ as

$$\mathbf{N}_s = \mathbf{DL}^\dagger, \tag{11}$$

where † represents a Moore-Penrose pseudo inverse.

(b) Estimation of original surface normal $\mathbf{N}$. Once the smoothed surface normal $\mathbf{N}_s$ is obtained, we use Eq. (10) for deriving the original surface normal $\mathbf{N}$. If the scattering matrix $\mathbf{H}$ is available and invertible, we can directly obtain the estimate of the original surface normal $\mathbf{N}$ in a linear least-squares fashion as $\mathbf{N} = \mathbf{H}^{-1}\mathbf{N}_s$. Since the estimation result produced by such a simple deconvolution is often degraded by ringing artifacts due to the loss of high-frequency information in the original signal, we use a smoothness constraint to stabilize the estimation. We design the smoothness term s as a weighted second-order difference of $\mathbf{n}(u)$ among u's neighborhood locations t and v as

$$\mathbf{n}''(u) = w(t,u)\left(\mathbf{n}(t) - \mathbf{n}(u)\right) - w(u,v)\left(\mathbf{n}(u) - \mathbf{n}(v)\right). \tag{12}$$

The weight $w(u,v)$ controls the discontinuity of surface normals by taking the difference of intensity observations across varying lightings $\mathbf{l}_i$ as

$$w(u,v) = \exp\left(-\frac{1}{m}\sum_{i}^{k}\left(I(u,\mathbf{l}_i) - I(v,\mathbf{l}_i)\right)^2\right). \tag{13}$$

The matrix expression of the smoothness $\mathbf{N}''$ is given as

$$\mathbf{N}'' = \mathbf{W}\mathbf{N}, \tag{14}$$

where $\mathbf{W} \in \mathbb{R}^{a\times m}$ is a matrix of the second-order derivative filter, a is the number of triplets used for computing the second-order derivatives. In our case, we define the triplets along horizontal and vertical directions in the image coordinates. Finally, our estimation problem becomes a ridge regression problem as

$$\hat{\mathbf{N}} = \underset{\mathbf{N}}{\operatorname{argmin}} \, ||\mathbf{H}\mathbf{N} - \mathbf{N_s}||_F^2 + \lambda||\mathbf{W}\mathbf{N}||_F^2, \tag{15}$$

where λ controls smoothness of the estimates. An explicit solution to this problem is given by setting its first-order derivative to be zero as

$$\mathbf{N} = \left(\mathbf{H}^T\mathbf{H} + \lambda\mathbf{W}^T\mathbf{W}\right)^{-1}\mathbf{H}^T\mathbf{N_s}. \tag{16}$$

In this manner, the estimates for the original surface normal $\mathbf{N}$ can be obtained in a closed-form.

The mathematical expression of the problem is equivalent to the image deblurring problem, where the original sharp image is recovered via deconvolution. The important difference is, however, that our problem deals with the deconvolution of surface normals. Therefore, conventional image priors that are developed for natural images may not be suitable. Other than this aspect, existing deconvolution techniques can be alternatively used for estimating the surface normal $\mathbf{N}$ from the smoothed surface normal $\mathbf{N}_s$. The convolution kernel $\mathbf{H}$ is generally unknown but can be either calibrated (non-blind deconvolution) or estimated (blind deconvolution). While most of image deblurring techniques are limited to spatially-invariant point spread functions (PSFs), which corresponds to handling optically homogeneous materials in our case, the formulation of Eq. (16) can naturally handle optically inhomogeneous materials, corresponding to the case of spatially-varying PSFs.

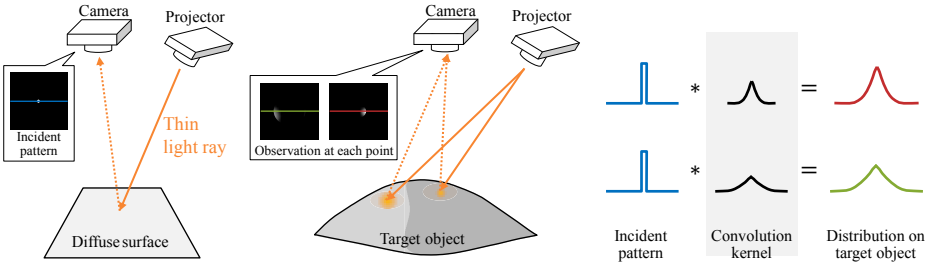


Fig. 2. Setting for measuring the convolution kernel. Projector casts a thin light ray to target object. We estimate the convolution kernel from the incident pattern and light distributions on the target object. In the case of inhomogeneous media, we capture light distributions at optically different regions.

4.1 Calibration of Convolution Kernel

As mentioned above, the surface normal deconvolution can be performed without knowing the convolution kernel by using blind deconvolution techniques; however, the knowledge of the convolution kernel is useful for stabilizing the estimation. Here we describe a simple procedure for measuring the convolution kernel. Fig. 2 shows our setting for measuring the convolution kernel. By illuminating a diffuse surface and the target translucent material individually by a thin ray emitted from a projector, we obtain the measurements of the incident light distribution and scattering response on the surface, respectively. The measured scattering response corresponds to the convolution between the incident light distribution and the convolution kernel. From this relationship, we calibrate the convolution kernel h. When target media is optically inhomogeneous, we need to calibrate convolution kernels at each optically different region.

5 Experiments

Now we evaluate our method using both synthetic and real-world data for the purposes of quantitative and qualitative evaluations.

5.1 Synthetic Data

Homogeneous Media. For the synthetic data, we use two types of scenes, scene A and B, as shown in Fig. 3. The image sizes of scene A and B are 150×113 and 160×160 pixels, respectively. For synthesizing the input images under varying lightings, we use Eq. (1) with the subsurface scattering model of Eq. (2). For the extinction term $R(x, y)$ in Eq. (2), we use the Dipole model [17] with the same parameters that are described in their paper. The camera model is orthographic and a pixel area corresponds to $(4/15)^2[\text{mm}^2]$ in the metric system.

Figure 3 (b) shows the result of Lambertian photometric stereo based on Eq. (11) and its angular error in the pseudo color. Although the estimated surface normals

are smoothed out due to subsurface scattering, especially around edges, the low-frequency signal of the overall surface normal directions are largely obtained.

To apply our surface normal deconvolution of Eq. (16), we use the extinction term $R(x, y)$ as the convolution kernel. The distance between scene points x and y is approximated to the distance between pixel sites u and v in the image coordinates. Figures 3 (c) and (d) show the result of our method with varying smoothness factors, $\lambda = 0.01$ and $\lambda = 0.1$, respectively. While the results with a small smoothness factor $\lambda = 0.01$ yield sharper reconstructions, they suffer from ringing artifacts around surface normal edges. Although the choice of the proper value for λ is scene-dependent and thus difficult as is the case with any regularization techniques, with a proper value of λ, our method significantly improves the reconstruction accuracy over the Lambertian photometric stereo that only considers the local illumination model, even though we also assume the same Lambertian model as the reflectance component. Table 1 summarizes the maximum and mean angular errors of the surface normal estimates using various material parameters. In general, we have observed that the smaller magnitude of subsurface scattering yields better accuracy, because stronger subsurface scattering cuts off the high-frequency signals more significantly. It shows that, by properly accounting for subsurface scattering, the accuracy improves by roughly $2 \sim 5$ times in comparison with the baseline technique that only considers the local illumination model.

For optically homogeneous materials, we can also use conventional deconvolution methods in place of solving Eq. (16). Figures 4 and 5 show results of conventional non-blind deconvolution and blind deconvolution for scene B, respectively. For the non-blind deconvolution methods, we use the same convolution kernel with the one that is used for producing the result of Fig. 3. The results show consistent improvement over Lambertian photometric stereo, although these original methods are not particularly designed for deblurring surface normal fields. In addition, the results of blind deconvolution methods in Fig. 5, where the convolution kernel is not given but simultaneously estimated, also show improvement. While the blind deconvolution is a harder problem than non-blind deconvolution and the results are generally worse, when the knowledge of the convolution kernel is unavailable, it is a viable option for our method.

Inhomogeneous Media. Our solution method is naturally applicable to the case of inhomogeneous materials as long as the convolution kernel $\mathbf{H}$ in Eq. (16) is defined. To evaluate the performance of our method for inhomogeneous materials, we produce synthetic images that contain different optical thicknesses using masks that indicate the material regions as shown in Fig. 6 (a) and (b). Due to the difference of the magnitudes of subsurface scattering in the material regions, the surface normal estimates obtained by Lambertian photometric stereo, shown in Fig. 6 (d) and (e), exhibit varying smoothnesses; smoother in the gray mask region, and sharper in the white mask region.

By applying our method, the surface normal field is consistently improved regardless of the material regions as shown in the figure. This recovery shows higher accuracy than that of Fig. 3, because of the fact that this inhomogeneous example contains a region where scattering is less significant.

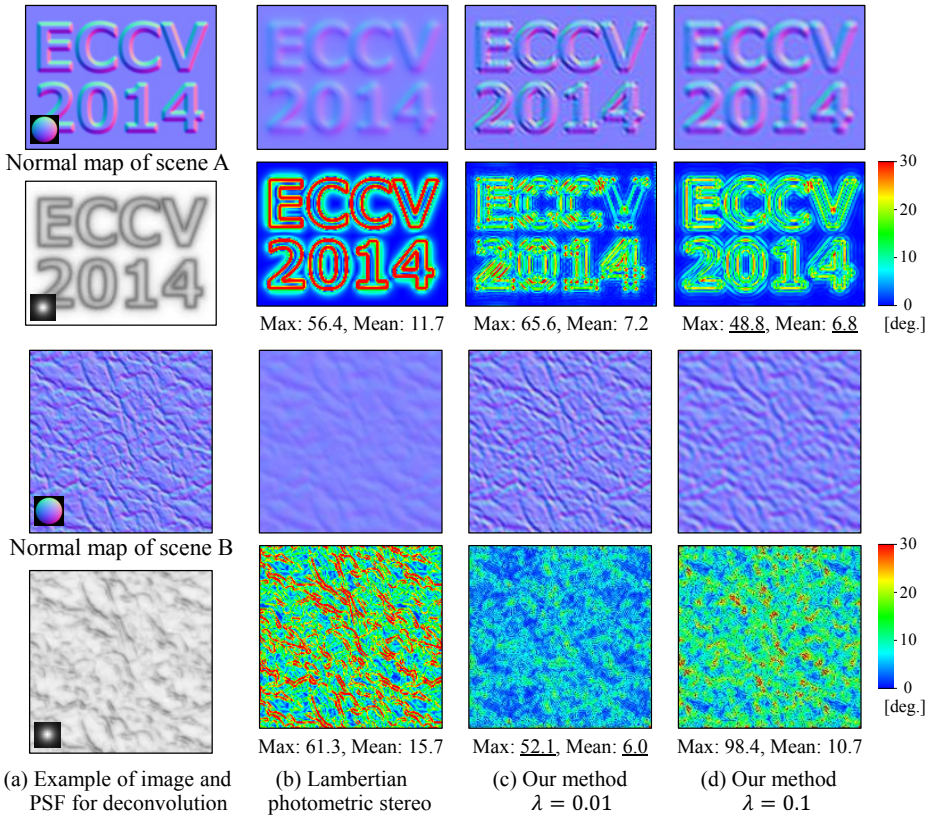
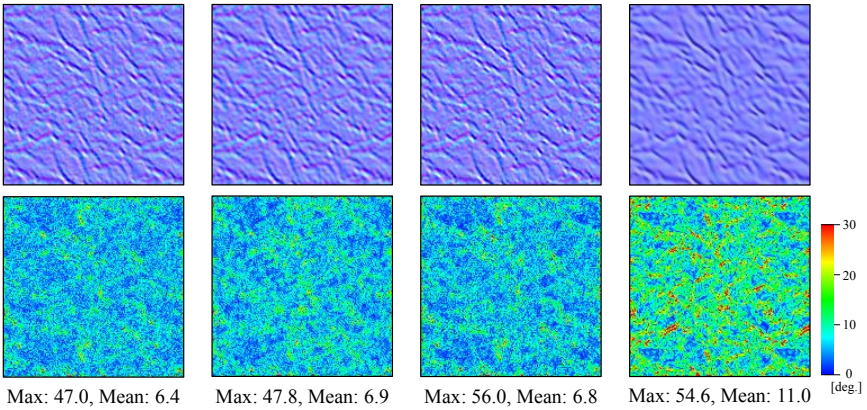


Fig. 3. Result of synthetic scenes A and B. (a) shows an example of synthetic images using the Dipole model with the skim milk parameters in [17]. (b) is the surface normal and error maps of the Lambertian photometric stereo. More faithful surface normals are obtained with our method in (c) and (d) with the varying smoothness factor λ.

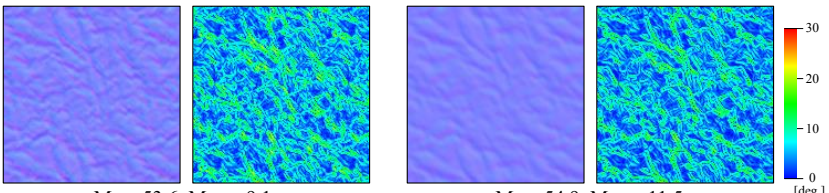
Table 1. Max and mean angular errors [deg.] of scene A and B with various materials. Parameters of each materials are described in [17].

Plot of kernels	Scene A						Scene B					
	Lambertian PS		Our method $\lambda = 0.01$		Our method $\lambda = 0.1$		Lambertian PS		Our method $\lambda = 0.01$		Our method $\lambda = 0.1$	
	max	mean	max	mean	max	mean	max	mean	max	mean	max	mean
1. Marble	41.5	7.7	40.1	3.0	30.5	2.6	56.2	11.9	29.1	1.9	36.2	5.6
2. Skim milk	56.4	11.7	65.6	7.2	48.8	6.8	61.3	15.7	52.1	6.0	98.4	10.7
3. Whole milk	37.7	6.6	30.7	2.5	25.4	1.9	52.4	10.7	22.1	1.5	28.5	4.4
4. Skin1	54.4	10.7	61.4	7.6	50.6	6.3	63.4	15.3	43.1	6.5	105.1	10.6
5. Skin2	50.3	9.8	59.5	5.1	44.1	4.7	61.5	14.3	47.9	4.2	86.2	8.7



(a) Wiener filter [20] (b) Richardson Lucy [21] (c) Cho *et al.* [22] (d) Dong *et al.* [23]

Fig. 4. Surface normal estimates of scene B using non-blind deconvolution methods

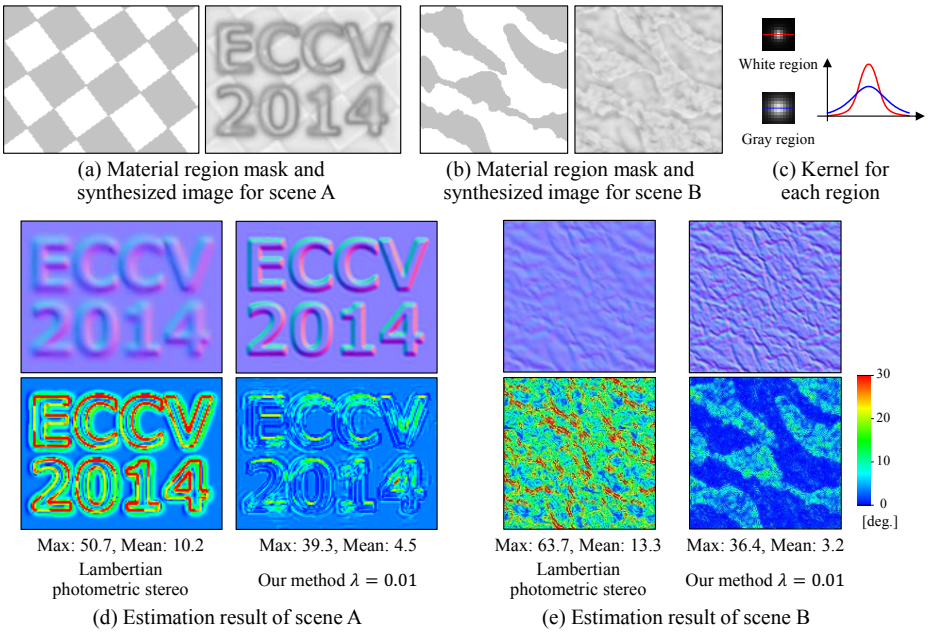


(a) Levin *et al.* [24] (b) Krishnan *et al.* [25]

Fig. 5. Surface normal estimates of scene B using blind deconvolution methods

5.2 Real-World Data

We also tested our method using the real-world translucent objects. Figure 7 (a) shows our experiment setting. We used a Nikon D90 camera with a linear radiometric response function (RAW mode) and with a telescopic lens to approximate an orthographic projection. The target scenes are illuminated under directional lightings, and the light directions are calibrated using a dark specular sphere. In addition, to avoid specular reflections from the scene, we placed polarizers in front of both the light source and camera. We used three target objects: a soap as a homogeneous medium, angel and unicorn ornaments as inhomogeneous media as shown in Fig. 7 (b). Each scene is recorded under 12 different lighting directions. The image size of the soap, angel, and unicorn scenes are 232×164, 206×257, and 158×230 pixels, respectively. Prior to the measurement, the convolution kernels are measured using the procedure described in Sec. 4.1. For the inhomogeneous objects, we measured two distinct kernels which depend on different material regions, one for the white region and the other for the pink region. Examples of the recorded intensity images are shown in the left-most



(a) Material region mask and synthesized image for scene A
(b) Material region mask and synthesized image for scene B
(c) Kernel for each region
(d) Estimation result of scene A
(e) Estimation result of scene B

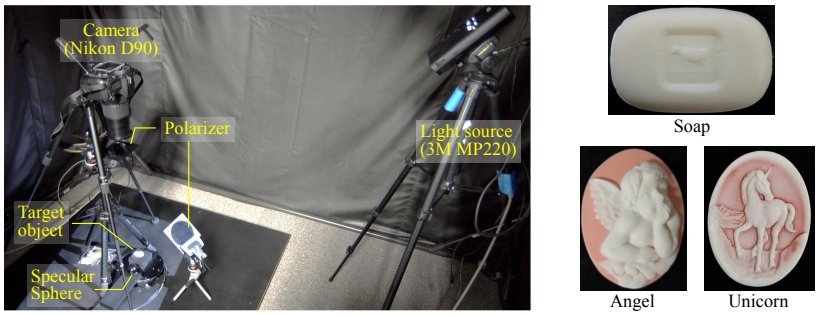
Fig. 6. Results with optically inhomogeneous media using scenes A and B. (a) and (b) show the masks that indicate different material regions and one of the synthesized images. We use two types of convolution kernels shown in (c) for these distinct regions. (d) and (e) show the smoothed surface normals obtained by Lambertian photometric stereo and our results, respectively.

column of Fig. 8. These images are not significantly blurry, but the details are smoothed out due to subsurface scattering.

The top row of Fig. 8 shows the result for the soap scene. While the surface normal recovered by Lambertian photometric stereo is significantly smoothed out, our method produces a much sharper estimate. The middle and bottom rows show the results of the angel and unicorn scenes, respectively. To assess the reconstruction accuracy, we created replicas of those by making molds with plaster. Assuming that the plaster reflectance is close to Lambertian, we obtained the reference surface normal by applying Lambertian photometric stereo to the replicas. The surface normal of plaster replicas exhibits details of the original shape, while the result of Lambertian photometric stereo is smoother. Our method makes the blurry surface normal much sharper and closer to the reference surface normal.

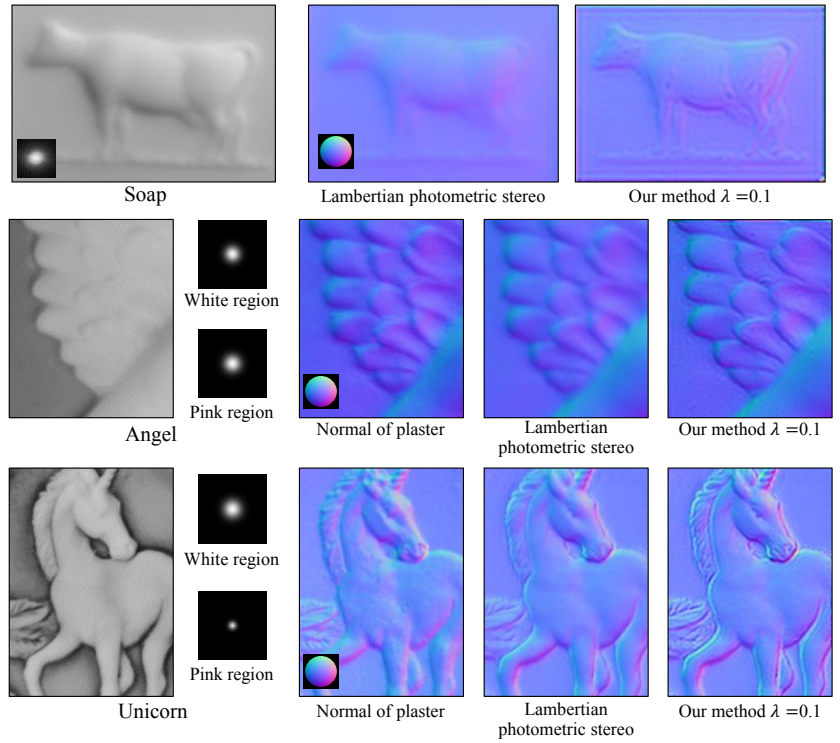
5.3 Discussion

Computation Time. Previous experiments show, in the case of optically homogeneous materials, we can apply various fast deconvolution methods for image deblurring to recover surface normal. However, in the case of inhomogeneous media, we have to solve Eq. (16) to deal with spatially variant convolution kernels.



(a) Experiment setting (b) Target objects

Fig. 7. Experiment setting and target objects. We use a projector as a light source. The camera is equipped with a telescopic lens. Polarizers are used to reduce the effects of specular reflection on the target object.



(a) Example of measured image and observed PSF (b) Estimated surface normals

Fig. 8. Result of the real-world scenes. The soap scene in the top row is a homogeneous medium, while the middle and bottom rows (angel and unicorn scenes) are made of inhomogeneous media.

Our matlab implementation on Intel Core i7 CPU (3.5 GHz) takes about 17.6, 39, and 3.5 seconds to recover surface of soap, angel, and unicorn scenes, respectively. The density of non-zero elements of matrix $\mathbf{F}^T\mathbf{F} + \lambda\mathbf{W}^T\mathbf{W}$ in Eq. (16) is about 2.5%. The computation time depends on the size and the number of non-zero elements of matrix $\mathbf{F}^T\mathbf{F} + \lambda\mathbf{W}^T\mathbf{W}$, which are determined by the input image size and apparent sizes of PSFs in the image coordinates.

Limitations. Our method has a couple of limitations. First, we have ignored the influence of Fresnel transmissions, thus our method is restricted to optically thick materials. As the material shows more directional scattering, the accuracy of our method may gradually decrease. We are interested in exploring an iterative estimation framework to adaptively update the convolution kernels for incorporating the Fresnel transmission effects. The second limitation is that our method in practice relies on known convolution kernels, especially when dealing with optically inhomogeneous materials. Although a sophisticated blind deconvolution method may resolve this issue, but at this point, the knowledge of the convolution kernel plays an important role in obtaining accurate surface normal estimates. We are interested in investigating a good prior for surface normal fields that may potentially improve the blind deconvolution.

6 Discussions and Summary

In this paper, we proposed a photometric stereo method for optically thick translucent objects. We have extended the previous study on a convolutional approximation of subsurface scattering and developed a surface normal deconvolution technique, which consists of a conventional photometric stereo and image deconvolution. Our experiment shows that the surface normals of translucent objects are reliably estimated by our method. As depicted in the experiment section, our method can benefit from a large body of image deblurring methods in the literature including blind deconvolution methods. In addition, we have shown that our method is able to handle optically inhomogeneous media.

References

1. Woodham, R.J.: Photometric Method For Determining Surface Orientation From Multiple Images. Optical Engineering 19, 139–144 (1980)
2. Lambert, J.H.: Photometria sive de mensure de gratibus luminis. Eberhard Klett (1760)
3. Ikehata, S., Wipf, D., Matsushita, Y., Aizawa, K.: Robust Photometric Stereo using Sparse Regression. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2012)
4. Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y., Ma, Y.: Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010, Part III. LNCS, vol. 6494, pp. 703–717. Springer, Heidelberg (2011)

5. Moore, K.D., Peers, P.: An empirical study on the effects of translucency on photometric stereo. The Visual Computer 29(6-8), 817–824 (2013)
6. Nayar, S.K., Krishnan, G., Grossberg, M.D., Raskar, R.: Fast separation of direct and global components of a scene using high frequency illumination. ACM Trans. on Graph. (ToG) 25(3), 935–944 (2006)
7. Munoz, A., Echevarria, J.I., Seron, F.J., Gutierrez, D.: Convolution-based simulation of homogeneous subsurface scattering. Computer Graphics Forum 30(8), 2279–2287 (2011)
8. Dong, B., Moore, K., Zhang, W., Peers, P.: Scattering Parameters and Surface Normals from Homogeneous Translucent Materials using Photometric Stereo. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2014)
9. Alldrin, N., Zickler, T., Kriegman, D.: Photometric stereo with non-parametric and spatially-varying reflectance. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2008)
10. Goldman, D.B., Curless, B., Hertzmann, A., Seitz, S.M.: Shape and spatially-varying BRDFs from photometric stereo. IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI) 32, 1060–1071 (2010)
11. Nayar, S.K., Ikeuchi, K., Kanade, T.: Shape from interreflections. In: Proc. of International Conference on Computer Vision (ICCV) (1990)
12. Chen, T., Lensch, H.P.A., Fuchs, C., Seidel, H.P.: Polarization and Phase-shifting for 3D Scanning of Translucent Objects. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2007)
13. Chen, T., Seidel, H.P., Lensch, H.P.A.: Modulated phase-shifting for 3D scanning. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2008)
14. Gu, J., Kabayashi, T., Gupta, M., Nayar, S.K.: Multiplexed Illumination for Scene Recovery in the Presence of Global Illumination. In: Proc. of International Conference on Computer Vision (ICCV) (2011)
15. Gupta, M., Agrawal, A., Veeraraghavan, A., Narasimhan, S.G.: A Practical Approach to 3D Scanning in the Presence of Interreflections, Subsurface Scattering and Defocus. International Journal of Computer Vision (IJCV) 102, 33–55 (2012)
16. Donner, C., Lawrence, J., Ramamoorthi, R., Hachisuka, T., Jensen, H.W., Nayar, S.: An empirical BSSRDF model. ACM Trans. on Graph. (ToG) 28(3) (2009)
17. Jensen, H.W., Marschner, S.R., Levoy, M., Hanrahan, P.: A practical model for subsurface light transport. In: Proc. ACM SIGGRAPH, pp. 511–518 (2001)
18. d'Eon, E., Irving, G.: A quantized-diffusion model for rendering translucent materials. ACM Transactions on Graphics (TOG) 30 (2011)
19. Borshukov, G., Lewis, J.: Realistic human face rendering for "the matrix reloaded". In: Proc. ACM SIGGRAPH Sketches and Applications (2003)
20. Wiener, N.: Extrapolation, Interpolation, and Smoothing of Stationary Time Series:With Engineering Applications, 1st edn. MIT Press (1949)
21. Richardson, W.H.: Bayesian-based iterative method of image restoration. Journal of the Optical Soceity of America (JOSA) 62, 55–59 (1972)
22. Cho, S., Wang, J., Lee, S.: Handling outliers in non-blind image deconvolution. In: Proc. of International Conference on Computer Vision (ICCV) (2011)
23. Dong, W., Zhang, L., Shi, G.: Centralized sparse representation for image restoration. In: Proc. of International Conference on Computer Vision (ICCV) (2011)
24. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Efficient marginal likelihood optimization in blind deconvolution. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2011)
25. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2011)