Consensus Photometric Stereo

Tomoaki Higo
The University of Tokyo
Tokyo, Japan
higo@cvl.iis.u-tokyo.ac.jp

Yasuyuki Matsushita
Microsoft Research Asia
Beijing, China
yasumat@microsoft.com

Katsushi Ikeuchi
The University of Tokyo
Tokyo, Japan
ki@cvl.iis.u-tokyo.ac.jp

Abstract

This paper describes a photometric stereo method that works with a wide range of surface reflectances. Unlike previous approaches that assume simple parametric models such as Lambertian reflectance, the only assumption that we make is that the reflectance has three properties; monotonicity, visibility, and isotropy with respect to the cosine of light direction and surface orientation. In fact, these properties are observed in many non-Lambertian diffuse reflectances. We also show that the monotonicity and isotropy properties hold specular lobes with respect to the cosine of the surface orientation and the bisector between the light direction and view direction. Each of these three properties independently gives a possible solution space of the surface orientation. By taking the intersection of the solution spaces, our method determines the surface orientation in a consensus manner. Our method naturally avoids the need for radiometrically calibrating cameras because the radiometric response function preserves these three properties. The effectiveness of the proposed method is demonstrated using various simulated and real-world scenes that contain a variety of diffuse and specular surfaces.

1 Introduction

Photometric stereo estimates surface orientation from a set of images taken from a fixed viewpoint under different lighting directions. After the early work of Woodham [24] and Silver [19], many researchers have studied the approach to make it work under more generalized conditions. Many photometric stereo methods are built upon specific parametric reflectance models, and are therefore naturally restricted to limited classes of reflectances. One of the most important directions for making photometric stereo practical is to handle various surface reflectances.

In this work, we present a new approach for solving the photometric stereo problem for handling a wide range of surface reflectances. Instead of assuming a specific parametric reflectance model, such as Lambertian, we assume only three reflectance properties that are often observed in real-world scenes: *monotonicity*, *visibility*, and *isotropy* of reflectance with respect to the cosine of surface normal and light direction. In fact, these properties are observed in many common materials such as plastics, ceramics, rubber, opaque glasses, and smooth glossy paints.

Our method uses about fifty images of a static scene composed of spatially varying reflectances, taken from a fixed viewpoint under varying and known directional lightings. From the intensity observations per pixel, we establish a set of inequalities derived from the monotonicity, visibility, and isotropy properties. These inequalities specify convex cones in the solution space of surface orientations. By taking the intersection of the convex cones, our method obtains a smaller solution space for the surface orientation. To efficiently estimate surface orientation, we estimate the orientation using an energy minimization.

Our consensus approach avoids imposing strict assumptions on surface reflectances and expands the applicability of photometric stereo. We show that the method can also deal with surfaces with only specular reflections using the same scheme by assuming the monotonicity and isotropy properties with respect to the cosine of surface normal and the bisector between the lighting direction and viewing direction. In addition, our method is naturally free from radiometric calibration. Because radiometric response functions are monotonic, the monotonicity, visibility, and isotropy properties are maintained in the observation with any non-linear radiometric response functions. This allows our method to work without knowing the camera response function.

1.1 Previous work

Photometric stereo has a long history. Early methods made strong assumptions on the surface reflectance, often the Lambertian model. There have been many studies to weaken the constraints on the reflectance model.

Coleman and Jain [7] use four images to detect highlights. Barsky and Petrou [3] extend the method to handle highlights as well as shadows by using four color images. These methods treat non-Lambertian effects as outliers. Solomon and Ikeuchi [21] recover surface roughness using the similar four-light setup. Provided there are enough images, non-Lambertian reflectance parameters can be esti-

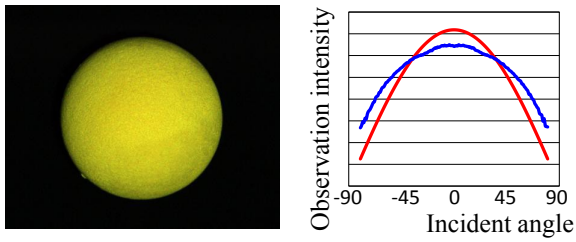


Figure 1. Measured reflectance of a diffuse yellow sphere painted with a poster color containing gum Arabic (blue line) and Lambertian fitting (red line).

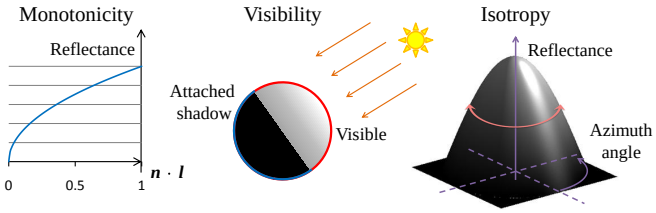


Figure 2. Monotonicity, visibility, and isotropy properties of reflectances. Left: the reflectance r monotonically increases with $\boldsymbol{n} \cdot \boldsymbol{l}$. Middle: the reflectance becomes zero when $\boldsymbol{n} \cdot \boldsymbol{l} \leq 0$. Right: the reflectance r gives the same value when $\boldsymbol{n} \cdot \boldsymbol{l}_i = \boldsymbol{n} \cdot \boldsymbol{l}_j$.

mated with their method. Nayer *et al*. [15] apply photometric stereo using a hybrid reflectance model that is a linear combination of Lambertian and specular components. Tagare and de Figueiredo [22] consider diffuse non-Lambertian surfaces and solve the problem using an m-lobed reflectance map. Georghiades [8] considers both diffuse and specular reflections and estimates surface normals as well as reflectance parameters based on the Torrance-Sparrow model with unknown light directions.

Some early works [12, 13] use a reference object for photometric stereo. Recently, Hertzmann and Seitz [10] proposed an example-based surface reconstruction method with arbitrary BRDFs. Goldman *et al*. [9] consider object surfaces modeled by a linear combination of two fundamental materials and remove the need for a reference object by iteratively estimating the basis BRDFs and surface normals. There are other approaches for the generalization of reflectance properties that are based on monotonicity [20], isotropy [14], Helmholtz stereopsis [25], bilateral symmetry [1, 2], and reflective symmetry of the halfway vector [11].

There have been some diffuse reflection models for dielectric materials. Reichman [18] derives diffuse reflection and transmission from the media of arbitrary optical thickness. Wolff *et al*. [23] provide an azimuth-independent diffuse reflection model with accounting isotropic subsurface scattering and Fresnel boundary effects. Oren and Nayar [17] propose a generalized diffuse reflectance model by taking surface roughness into account. Once the surface roughness is known, it is reported that it works well for estimating surface orientations. However, in practice, it is difficult to know the surface roughness beforehand. While our model theoretically does not entirely cover the Oren-Nayer model because of forward and backward scattering effects, as shown later in our experimental results, our method is still able to handle rough surfaces such as the one shown in Fig. 1.

Our approach is close to Chen *et al*.'s work [6] in that both methods do not use a specific parametric reflectance model. Their method determines surface orientations by taking the bisector of view and lighting directions using specular highlight. Unlike Chen *et al*.'s approach, our method can handle diffuse surfaces as well as specular surfaces using monotonicity, visibility, and isotropy properties.

The rest of this paper is as follows. We first describe the key ideas of the proposed method in Section 2. We formulate the solution method in Section 3 and present results in Section 4 followed by discussions and conclusions.

2 Proposed approach

Let us begin with our image formation model. An intensity observation o_i is described using the light source intensity E, ambient lighting a, surface normal $\boldsymbol{n}$, incident light direction $\boldsymbol{l}_i$, reflectance function r and radiometric response function f as

$$o_i = f\big(E\, r(\boldsymbol{n} \cdot \boldsymbol{l}_i) + a\big), \tag{1}$$

where $\boldsymbol{n} \cdot \boldsymbol{l}_i$ is a dot product of $\boldsymbol{n}$ and $\boldsymbol{l}_i$.

In this work, we assume three properties about the surface reflectance r: monotonicity, visibility, and isotropy (Fig. 2). These reflectance properties are observed in a wide range of diffuse reflectances. In fact, it is pointed out that many existing diffuse materials deviate from the Lambertian model in prior study [23, 17]. Fig. 1 shows an actual measurement that deviates from the Lambertian model. Our reflectance model covers such diffuse reflections as well as the Lambertian model as a special case where $r(\boldsymbol{n} \cdot \boldsymbol{l}) = \rho \boldsymbol{n} \cdot \boldsymbol{l}$, where ρ is surface albedo.

Using these properties of monotonicity, visibility, and isotropy, we derive three constraints in the form of inequalities that specify possible solution spaces of the surface orientation. Each of these three constraints independently gives a solution space. Our method estimates surface orientations by taking the intersection of these solution spaces. We only use illuminated pixels for these constraints. The solution space of surface normal $\boldsymbol{n}$ is initialized on a Gaussian sphere because the surface normal $\boldsymbol{n}$ is a unit vector.

Monotonicity constraint We assume the following monotonicity of the reflectance function r:

$$\boldsymbol{n} \cdot \boldsymbol{l}_i > \boldsymbol{n} \cdot \boldsymbol{l}_j \Leftrightarrow r(\boldsymbol{n} \cdot \boldsymbol{l}_i) > r(\boldsymbol{n} \cdot \boldsymbol{l}_j). \tag{2}$$

This monotonicity represents that as the dot product of surface normal $\boldsymbol{n}$ and lighting direction $\boldsymbol{l}$ increases, the reflectance r increases.

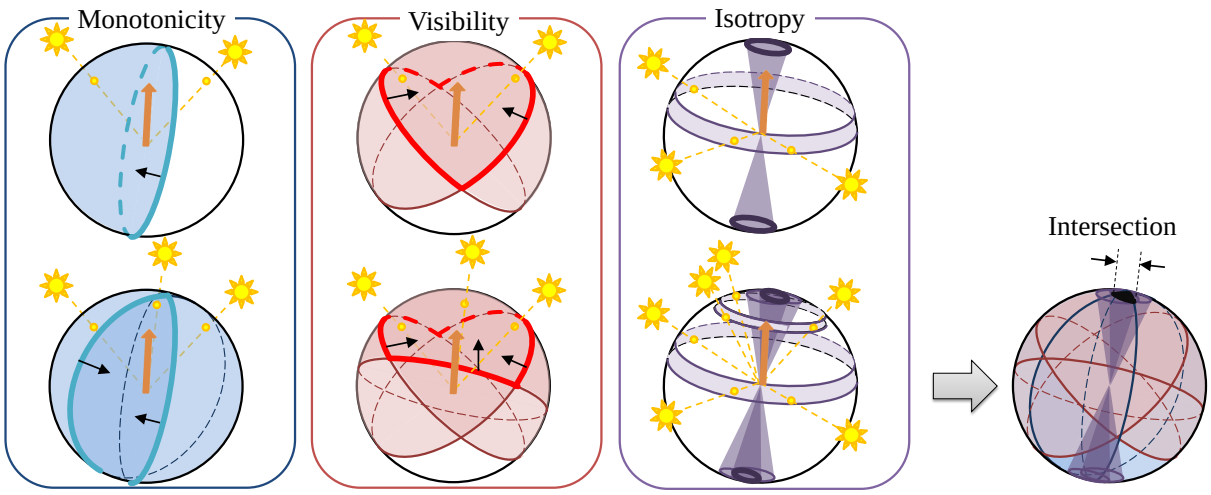


Figure 3. Monotonicity, visibility, and isotropy constraints. Each of these three constraints gives a solution space of the surface orientation. By taking the intersection of the solution spaces, our method obtains a smaller solution space of the surface orientation. The narrow arrows represent the solution space, and the bold ones correspond to the true surface orientation. The two rows show how the solution space becomes smaller as the number of observations increases.

These constraints hold even for unknown radiometric response functions. In Eq. (1), the radiometric response function f is also monotonically increasing, and E and ρ are non-negative. Therefore, $f(r(x))$ is also monotonically increasing, so the intensity observation o monotonically increases as $\boldsymbol{n} \cdot \boldsymbol{l}$ increases. This property eliminates the necessity of radiometric calibration for our method and allows us to directly use the following relationship regardless of the shape of the response function f:

$$\boldsymbol{n} \cdot \boldsymbol{l}_i > \boldsymbol{n} \cdot \boldsymbol{l}_j \Leftrightarrow o_i > o_j. \tag{3}$$

Using inequalities (3) obtained from multiple observation pairs (o_i, o_j), the solution space $\mathcal{N}_1$ of the surface orientation $\boldsymbol{n}$ can be determined by taking the intersection of multiple observations as

$$\mathcal{N}_1 = \Big\{\boldsymbol{n} \in \mathbb{R}^3 \mid \bigcap_{i,j} \big((\boldsymbol{l}_i - \boldsymbol{l}_j) \cdot \boldsymbol{n} > 0\big)\Big\}, \tag{4}$$

for pairs of $\boldsymbol{l}_i$ and $\boldsymbol{l}_j$ that satisfy $r(\boldsymbol{n} \cdot \boldsymbol{l}_i) > r(\boldsymbol{n} \cdot \boldsymbol{l}_j)$. The pair of $(\boldsymbol{l}_i, \boldsymbol{l}_j)$ makes the solution space specified on the north hemisphere whose pole is $(\boldsymbol{l}_i - \boldsymbol{l}_j)$ as illustrated in Fig. 3 (Left).

Visibility constraint When a scene point is illuminated by a light source $\boldsymbol{l}$, the surface normal $\boldsymbol{n}$ should lie in the hemisphere $\boldsymbol{n} \cdot \boldsymbol{l} > 0$. When $\boldsymbol{n} \cdot \boldsymbol{l} \leq 0$, the scene point is in the attached shadow, *i.e.*, the scene point is not visible from the light source. The visibility is defined as

$$\mathcal{N}_2 = \Big\{\boldsymbol{n} \mid \bigcap_{i} (\boldsymbol{n} \cdot \boldsymbol{l}_i > 0)\Big\}, \tag{5}$$

for all lighting directions $\boldsymbol{l}_i$ that illuminate the scene point. A similar constraint is used by Belhumeur and Kriegman [5] for describing possible light source directions. Because our method only uses illuminated pixels, it is not necessary to identify whether the pixel is in an attached or cast shadow.

Isotropy constraint In addition to the above two constraints, we use the isotropy constraint when multiple similar intensity observations are obtained. Suppose an ideal case where more than two observations under different lighting directions show the same intensity value. In this case, given the different lighting directions $\boldsymbol{l}_i$, $\boldsymbol{l}_j$, and $\boldsymbol{l}_k$, the surface normal $\boldsymbol{n}$ should fall on the direction that is perpendicular to the plane spanned by the lighting vectors. It can be determined up to a sign ambiguity by taking the cross-product $\pm(\boldsymbol{l}_i - \boldsymbol{l}_j) \times (\boldsymbol{l}_i - \boldsymbol{l}_k)$. Such ideal situations are rare in practice, so we use a relaxed near-equality constraint:

$$\boldsymbol{n} \cdot \boldsymbol{l}_i \simeq \boldsymbol{n} \cdot \boldsymbol{l}_j \Leftrightarrow o_i \simeq o_j. \tag{6}$$

This represents that when similar intensity observations o are obtained, the cosines of incident lighting direction and surface normal are also similar.

When we have similar observations o_i, $(i = 1, 2, \ldots, k)$ under different lighting directions $\boldsymbol{l}_i$, we can expect that the surface normal $\boldsymbol{n}$ lies near to the direction where the variance of $\boldsymbol{n} \cdot \boldsymbol{l}$ is minimized:

$$\boldsymbol{n} = \pm \min_{\boldsymbol{n}} \sum_{i=1}^{k} \big(\boldsymbol{n} \cdot \boldsymbol{l}_i - \overline{\boldsymbol{n} \cdot \boldsymbol{l}}\big)^2, \tag{7}$$

where $\overline{\boldsymbol{n} \cdot \boldsymbol{l}}$ is the mean of the dot products. Using m such normal directions $\boldsymbol{n}_m$ obtained from m-sets of lighting directions, our method determines the solution space $\mathcal{N}_3$ that is represented by a convex cone spanned by $\boldsymbol{n}_m$ as

$$\mathcal{N}_3 = \Big\{\boldsymbol{n} \mid \boldsymbol{n} = \sum_{m} a_m \boldsymbol{n}_m,\ a_m \geq 0\Big\}. \tag{8}$$

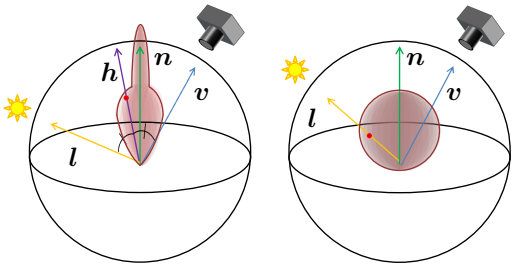


Figure 4. Monotonicity and isotropy constraints for the case of specular reflection. Left: The case of specular reflection. The reflectance r monotonically increases with $\boldsymbol{n}\cdot\boldsymbol{h}$ and gives the same value when $\boldsymbol{n}\cdot\boldsymbol{h_i} = \boldsymbol{n}\cdot\boldsymbol{h_j}$. We use the bisector $\boldsymbol{h}$ replacing the light vector $\boldsymbol{l}$ in Eq. (1) for the specular lobes. Right: The case of diffuse reflection.

Consensus solution Each of the monotonicity, visibility, and isotropy constraints independently gives a solution space (Fig. 3). Our method takes the intersection of these to form a smaller solution space $\mathcal{N}$as

$$\mathcal{N} = \mathcal{N}_1 \cap \mathcal{N}_2 \cap \mathcal{N}_3. \tag{9}$$

As the number of images increases, it is expected that the solution space $\mathcal{N}$ becomes smaller.

2.1 Specular lobes

Our method can be extended naturally to handle specular reflections by assuming monotonicity and isotropy for specular lobes. Here, we assume only specular reflection and no diffuse reflection, like for metallic surfaces. In this case, as shown in Fig. 4, the monotonicity and isotropy are assumed with respect to the cosine of surface orientation $\boldsymbol{n}$ and the bisector $\boldsymbol{h}$ $(= (\boldsymbol{l}+\boldsymbol{v})/|\boldsymbol{l}+\boldsymbol{v}|)$ between the light direction $\boldsymbol{l}$ and the view direction $\boldsymbol{v}$. The right-hand side of the figure depicts the diffuse case for reference. In this way, by replacing the light vector $\boldsymbol{l}_i$ with the bisector $\boldsymbol{h}_i$ in Eq. (1), the previous discussion holds for the specular lobes. For the visibility constraint, since we do not know the width of specular lobes, we still use $\boldsymbol{n}\cdot\boldsymbol{l} > 0$ as the constraint.

3 Implementation

To efficiently estimate surface orientation $\boldsymbol{n}$, we cast the consensus approach to an energy minimization problem. We develop energy terms for monotonicity, visibility, and isotropy constraints, respectively. These energy terms are computed at each pixel.

Monotonicity term From Eq. (4), we develop an energy term that favors $\boldsymbol{n}\cdot(\boldsymbol{l}_i - \boldsymbol{l}_j) > 0$ being satisfied for observations $o_i > o_j$. Using a sigmoid-like function, we formulate this constraint as

$$E_1(\boldsymbol{n}) = \frac{1}{N_1}\sum_{i,j}\frac{1 - k\boldsymbol{n}\cdot(\boldsymbol{l}_i - \boldsymbol{l}_j)}{1+\exp\big(t\boldsymbol{n}\cdot(\boldsymbol{l}_i - \boldsymbol{l}_j)\big)}, \tag{10}$$

for all pairs of (i,j) where $o_i > o_j$. We use sigmoid-like function to equally give a small cost when $\boldsymbol{n}\cdot(\boldsymbol{l}_i - \boldsymbol{l}_j) > 0$. In the energy term, t is a gain, and N_1 is the number of pairs (i,j) that are used for the computations. The numerator is designed to form a slope that is determined by the factor k so that more deviations from the constraint are penalized. With such a slope, the optimization becomes more efficient and quickly converges. Fig. 5 shows the form of the function $s(x) = (1-kx)/(1+e^{tx})$ that is used for E_1.

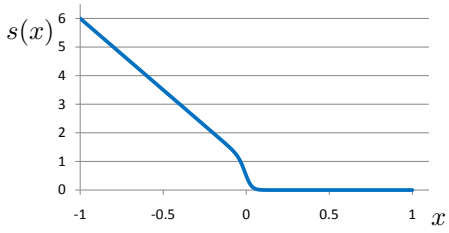


Figure 5. Plot of a function $s(x) = (1-kx)/(1+e^{tx})$ used to design energy terms. $(k,t) = (5,50)$ is used for the plot.

For efficient computation, we resample all possible combinations (i,j) to reduce the number of pairs. We select N_M observations o_j that are close to o_i while satisfying $o_i > o_j$, because the similar intensity observation pairs (o_i, o_j) tend to give smaller solution spaces. However, the combinations that are used for the isotropy constraint are excluded because of the condition $o_i > o_j$.

Visibility term We formulate the visibility constraint of Eq. (5) in a similar manner with the monotonicity constraint Eq. (10). Using a sigmoid-like function, the visibility term E_2 is formulated as

$$E_2(\boldsymbol{n}) = \frac{1}{N_2}\sum_{i}\frac{1 - k\boldsymbol{n}\cdot\boldsymbol{l}_i}{1+\exp(t\boldsymbol{n}\cdot\boldsymbol{l}_i)}, \tag{11}$$

where N_2 is the number of observations that are illuminated, *i.e.*, the number of observations used for the estimation.

Isotropy term The isotropy constraint Eq. (8) gives a solution space from a set of lighting directions that produce similar intensity observations. The more similar the intensity observations are, the smaller the solution space becomes. We formulate this as an energy function E_3 that favors smaller variances of the dot product $\boldsymbol{n}\cdot\boldsymbol{l}$ in each set of similar observations. Given m sets of similar observations S, the energy term is defined as

$$E_3(\boldsymbol{n}) = \frac{1}{\sum_i^m |S_i|}\sum_{i}^{m}\sum_{j\in S_i}\big(\boldsymbol{n}\cdot\boldsymbol{l}_j - (\overline{\boldsymbol{n}\cdot\boldsymbol{l}})_i\big)^2, \tag{12}$$

where S_i is the i-th set of observation indices, $|S_i|$ represents the number of elements in the set, and $(\overline{\boldsymbol{n}\cdot\boldsymbol{l}})_i$ is the mean of the dot product $\boldsymbol{n}\cdot\boldsymbol{l}$ in S_i.

Energy function The energy function E is defined by combining the above three constraints and an additional constraint of a unit normal length as

$$E(\boldsymbol{n}) = \lambda_1 E_1(\boldsymbol{n}) + \lambda_2 E_2(\boldsymbol{n}) + \lambda_3 E_3(\boldsymbol{n}) + (1-|\boldsymbol{n}|^2)^2, \tag{13}$$

where λ_i represents a weighting factor. We use a Levenberg-Marquardt method [4] to minimize the multivariate function to estimate a surface normal vector per pixel. For initialization, we use the lighting direction vector that shows the highest intensity (without saturation) as the initial guess of the normal vector. Before the optimization, we exclude low intensity observations as shadow pixels and use only illuminated observations o_i.

How many lighting directions are required? As mentioned in Section 2, as the number of light directions increases, the solution space becomes smaller. We analyze the statistical relationship between the accuracy of the surface normal estimate $\triangle\theta$ and the number of light directions N_l based on the analysis of Okabe *et al*. [16]. They pointed out that the number of segments N_{sgm} with a sphere divided up by N_{gc} great circles becomes $N_{sgm} = 0.9 \times N_{gc}^{1.8}$. Assuming that N_{sgm} equal-size disks with the area $\pi(\triangle\theta)^2$ cover the entire solid angle of a sphere (4π), we obtain an optimistic estimate of the accuracy $\triangle\theta \simeq \sqrt{4/N_{sgm}}$. This means that it requires $N_{gc} > 200$ for obtaining an accuracy of less than one degree.

Suppose a half of N_l are illuminated observations among the entire N_l observations, *i.e*., $N_2 = N_l/2$. In this case, the visibility constraint provides $N_l/2$ great circles. The monotonicity constraint provides N_M pairs for each observation $o_i\,(i = 1, 2, \ldots, N_2)$. Therefore, the monotonicity constraint gives $N_M N_l/2$ great circles. In our experiments, we use $N_M = 8$, and the total number of great circles is $N_{gc} = N_l/2 + 4N_l = 4.5N_l$. Even without the isotropy constraint, this analysis indicates that given about 50 input images, our method can achieve an accuracy of less than one degree.

4 Experiments

To evaluate the effectiveness of the proposed method, we performed experiments using both simulation and real-world scenes. We first show a quantitative evaluation using the simulation data in Section 4.1. Second, we evaluated our method using five real-world scenes in Section 4.2. Throughout the experiments, we used parameters $k = 5$, $t = 50$, $N_M = 8$, $\lambda_1 = 8$, $\lambda_2 = 1$, and $\lambda_3 = 300$ for diffuse objects. For specular objects, we only changed the weighting factor of isotropy term to $\lambda_3 = 30$.

4.1 Simulation results

The simulation experiment is designed to quantitatively examine the performance of the proposed method. We use combinations of different settings; (1) linear/non-linear camera response functions, (2) Lambertian/non-Lambertian reflectances, and (3) with/without ambient lighting. We represent these settings by **Y**es or **N**o of {linear response function, Lambertian surface, ambient lighting}. For example {Y, N, Y} means the combination of linear response function, non-Lambertian surface, and with ambient illumination. A synthetic scene is rendered using these settings, and our method is applied to each of these datasets. Fig. 6 (left) shows the reference spheres rendered under these settings. In the middle, the shapes of the non-linear response function and non-Lambertian diffuse reflectance are shown. The right-hand side table shows the summary of the estimation results. As shown in the table, our method is not susceptible to ambient lighting, non-Lambertian diffuse reflection, and non-linear response function, while the standard photometric stereo method suffers under these non-ideal conditions.

4.2 Real-world results

We applied our method to various diffuse objects and specular objects and compared with the standard photometric stereo method. We used five different scenes under various conditions: (1) yellow sphere scene (Fig. 1, non-Lambertian), (2) terracotta scene (non-linear response function), (3) statue scene (with ambient illumination), (4) relief scene (non-linear response function with ambient illumination), and (5) clip scene (specular lobes).

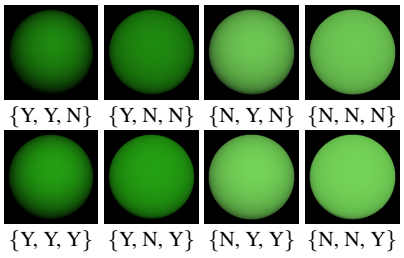
We recorded the scenes using two different cameras: a Sony color digital camera XCD-X710CR that has a linear response function, and a Nikon D1x camera with a non-linear response function. The scenes were illuminated by a moving LED point light source and recorded from a fixed viewpoint. To obtain the light source directions, we used a mirror sphere placed in the scene. We compare our method with the standard photometric stereo method based on the Lambertian model (referred to as 'standard PS' in the following). To use the same input to both methods, shadow pixels were excluded when the standard PS is applied.

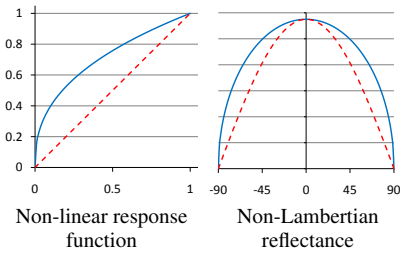
Fig. 7 shows the result of the yellow sphere scene recorded by a Sony XCD-X710CR. Our method recovers surface normals from a non-Lambertian diffuse reflectance scene more accurately than the standard photometric stereo. The error maps in the fourth and fifth figure in Fig. 7 clearly show the deviation.

In Fig. 8, we show the result of the terracotta scene taken with the Nikon D1x camera with a non-linear response function. The results of our method and the standard photometric stereo appear to be similar, but we can still see that our method can obtain accurate orientations, *e.g*., on top of the left hand of the terracotta soldier.

Fig. 9 shows the result of the statue scene under ambient illumination taken by a Sony XCD-X710CR. Our method produces faithful surface orientations, while the standard photometric stereo produce overly smooth surface normals. On computational cost, it takes 592[s] for the statue dataset (47 images, 480×490[pixel]) using a Core2Duo CPU (2.33GHz, single thread).

The relief scene of Fig. 10 was taken with a Nikon D1x camera with a non-linear response function, under ambient lightings. Our method recovers surface orientations even when the real scenes significantly deviate from classical as-



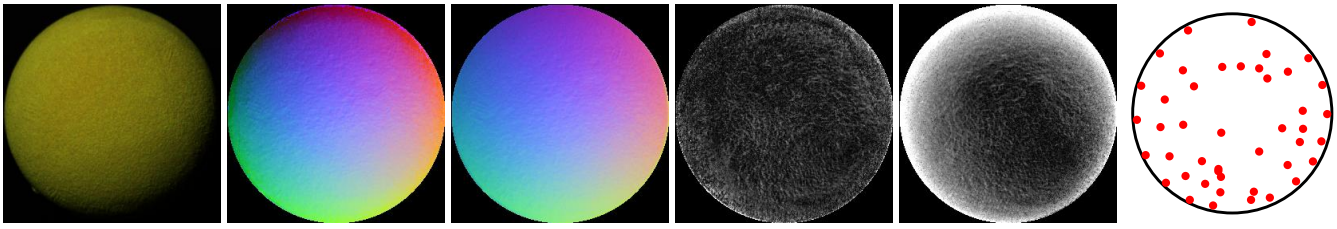


Non-linear response function

Non-Lambertian reflectance

	Our method		Standard PS	
	mean	med	mean	med
{Y, Y, N}	0.708	0.617	**0.000**	**0.000**
{Y, N, N}	**0.740**	**0.651**	8.479	7.999
{N, Y, N}	**0.719**	**0.634**	3.866	3.640
{N, N, N}	**0.737**	**0.647**	8.541	7.855
{Y, Y, Y}	**0.705**	**0.622**	5.320	4.920
{Y, N, Y}	**0.741**	**0.658**	8.412	7.793
{N, Y, Y}	**0.721**	**0.633**	7.105	6.576
{N, N, Y}	**0.723**	**0.627**	8.709	8.020

Figure 6. Simulation setup and results. Left shows the reference spheres rendered with the combinations of {(1) linear response function **Y**es/**N**o, (2) Lambertian: **Y**es/**N**o, (3) ambient illumination: **Y**es/**N**o }. In the middle, the shapes of the non-linear response function and non-Lambertian reflectance that are used in this simulation are shown. The right table shows the mean and median RMSE [deg.] evaluation of the estimated surface normals under corresponding rendering settings.



Input | Our method | Standard PS | Error of our method | Error of standard PS | Light directions

Figure 7. Result of our method applied to the yellow sphere with a non-Lambertian surface. From left to right, one of the input images, the estimated normal map with our method, that with standard photometric stereo method, the corresponding errors from the ground truth, and the sampled light directions are shown. The higher intensity in the error maps indicates the greater errors. 43 images are used as input.

sumptions, *i.e.*, Lambertian reflectance, no ambient illumination, and a linear camera response.

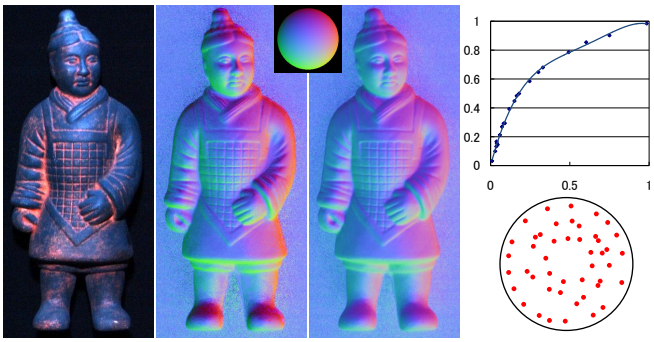
The effect of the three constraints are evaluated in Fig. 11. Monotonicity is a strong constraint, but large errors are found in places. Isotropy alone also gives good estimates; however, inaccurate surface normals are observed where the zenith angle of surface normals is large, *i.e.*, outward-looking surface normals. The visibility constraint prevents large errors, especially for the outward-looking surface normals. These three constraints work in a complementary manner, and the combination of all the three constraints gives the best result.

Fig. 12 shows the result of the clip scene containing specular lobes. Our method can estimate surface normals from the specular lobes as well.

Finally, Fig. 13 shows the rendering of 3D surfaces and relighting. The relief scene (left) is the reconstructed 3D, and the others are relighting results. The reference spheres depict the lighting directions.

5 Discussion

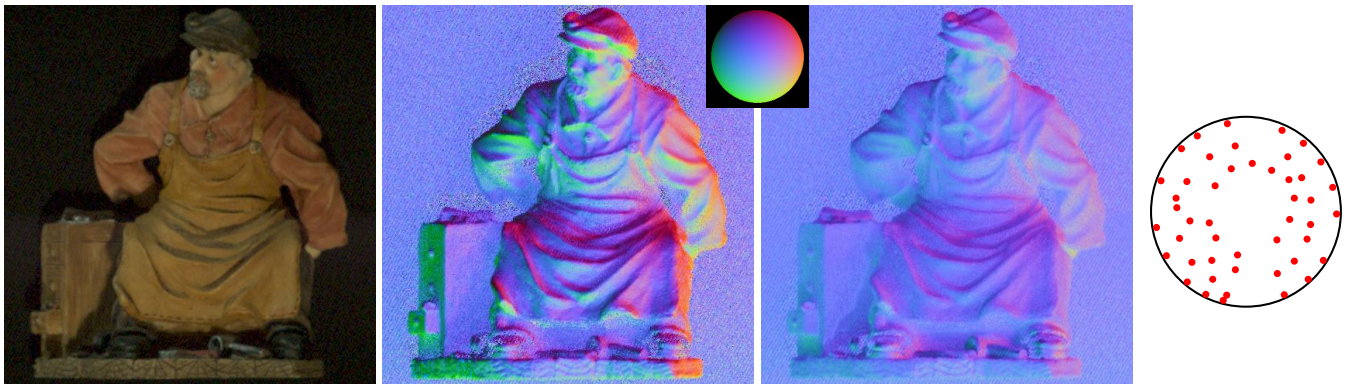
We present a consensus approach for photometric stereo for a generalized reflectance model that holds three properties: monotonicity, visibility, and isotropy. These properties are naturally observed in a wide variety of diffuse reflection as well as in specular lobes. In addition, our method eliminates the necessity of radiometric calibration and any dependency on the ambient illumination.



Input | Our method | Standard PS

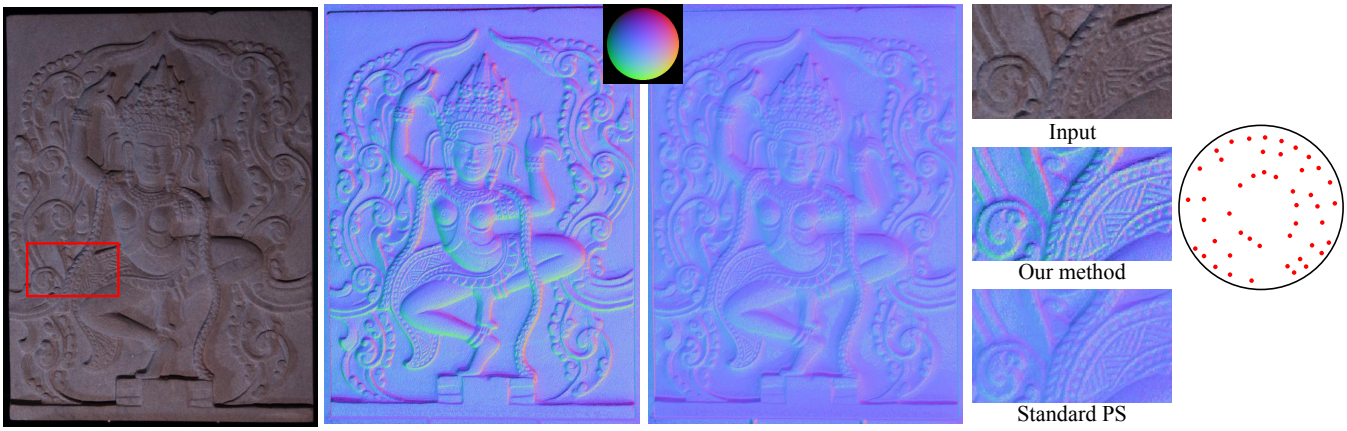
Figure 8. Result of the terracotta scene taken with a Nikon D1x camera with a non-linear response function without ambient illumination. From left to right, one of the input images, the estimated normal map with our method, that with the standard photometric stereo method, the measured response function, and lighting directions are shown. 46 images are used as input.

Currently, our method is limited to work with surfaces that show either diffuse or specular reflection. To handle surfaces that have both diffuse and specular reflections, we are interested in applying a color subspace method [26] for separating these reflections. We are also interested in using shadowed pixels. In our current method, we only use illuminated pixels for estimation; however, it has been shown in the previous work [16] that shadow can be used as a cue for estimation.



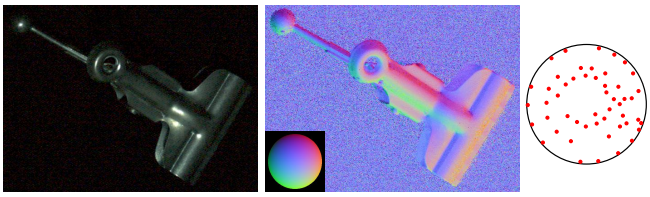
Input Our method Standard PS Light directions

Figure 9. Result of the statue scene recorded by a Sony XCD-X710CR camera with a linear response function under ambient illumination. From left to right, one of the input images, the estimated normal map with our method, and that of the standard photometric stereo method are shown. The reference sphere is overlaid in the middle of the second and third figures. The light source directions are shown on the right. 47 images are used as input.



Input Our method Standard PS Close up Light directions

Figure 10. Result of the relief scene taken with a Nikon D1x camera with a non-linear response function under ambient illumination. From left to right, one of the input images, the estimated normal map with our method, that with standard photometric stereo method, and the light directions. 47 images are used as input.



Input Normal map Light directions

Figure 12. Result of the clip scene captured with a Sony XCD-X710CR camera. From left to right, one of the input images, the estimated normal map, and the light directions are shown. 50 images are used as input.

References

[1] N. Alldrin and D. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance: a stratified photometric stereo approach. *Proc. of Int'l Conf. on Computer Vision*, 2007.

[2] N. Alldrin, T. Zickler, and D. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. *Proc. of Computer Vision and Pattern Recognition*, 2008.

[3] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003.

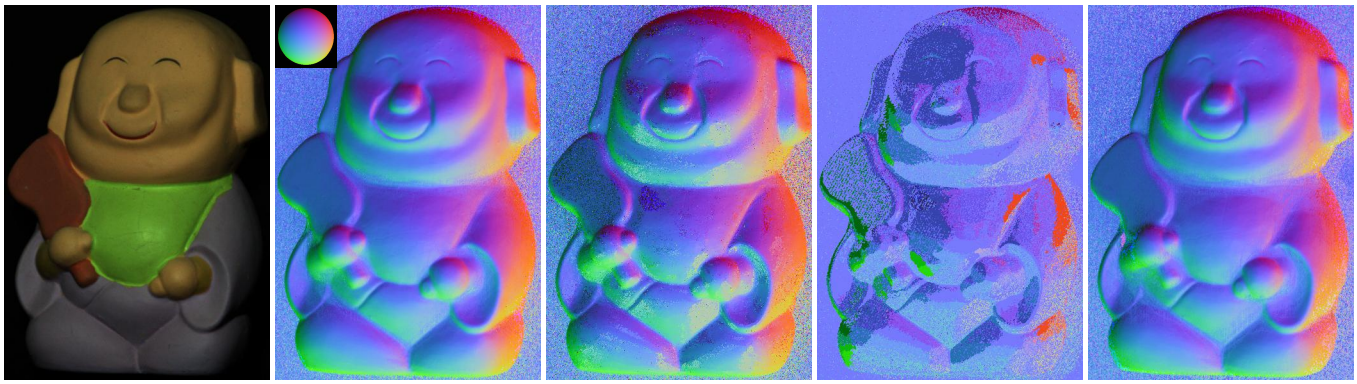
[4] D. M. Bates and D. G. Watts. Nonlinear regression and its applications. 1988.

[5] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible illumination conditions? *Int'l Journal of Computer Vision*, 28(3):245–260, 1998.

[6] T. Chen, M. Goesele, and H. Seidel. Mesostructure from Specularity. *Proc. of Computer Vision and Pattern Recognition*, 2:1825–1832, 2006.

[7] E. N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18:309–328, 1982.

[8] A. Georghiades. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. *Proc. of Int'l Conf. on Computer Vision*, pages 816–823, 2003.



Input | All the three constraints | Only monotonicity | Only visibility | Only isotropy

Figure 11. Results with all the three and individual constraints. Captions below the figures indicate the constraints that are used.

3D reconstruction | Relighting results

Figure 13. 3D reconstruction and relighting results. From left to right, 3D reconstruction of the relief scene from top view, close-up side view, relighting result of terracotta, statue, and clip scenes. The reference spheres show rendering parameters.

[9] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz. Shape and spatially-varying BRDFs from photometric stereo. *Proc. of Int'l Conf. on Computer Vision*, 1:341–348, 2005.

[10] A. Hertzmann and S. Seitz. Example-based photometric stereo: shape reconstruction with general, varying BRDFs. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.

[11] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler. A photometric approach for estimating normals and tangents. *Int'l Conference on Computer Graphics and Interactive Techniques*, 2008.

[12] B. Horn and K. Ikeuchi. The mechanical manipulation of randomly oriented parts. *Scientific American*, 251(2):100–111, 1984.

[13] K. Ikeuchi. Determining a depth map using a dual photometric stereo. *The Int'l Journal of Robotics Research*, 6(1):15–31, 1987.

[14] J. Lu and J. Little. Reflectance and shape from images using a collinear light source. *Int'l Journal of Computer Vision*, 32(3):213–240, 1999.

[15] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. on Robotics and Automation*, 6(4):418–431, 1990.

[16] T. Okabe, I. Sato, and Y. Sato. Attached shadow coding: estimating surface normals from shadows under unknown reflectance and lighting conditions . *Proc. of Int'l Conf. on Computer Vision*, 2009.

[17] M. Oren and S. Nayar. Generalization of the Lambertian model and implications for machine vision. *Int'l Journal of Computer Vision*, 14(3):227–251, 1995.

[18] J. Reichman. Determination of absorption and scattering coefficients for nonhomogeneous media. 1: Theory. *Applied Optics*, 12(8):1811–1815, 1973.

[19] W. Silver. Determining shape and reflectance using multiple images. *Master's thesis, MIT*, 1980.

[20] W. Smith and E. Hancock. Estimating Facial Reflectance Properties Using Shape-from-Shading. *Int'l Journal of Computer Vision*, 86:152–170, 2010.

[21] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(4):449–454, 1996.

[22] H. Tagare and R. de Figueiredo. A theory of photometric stereo for a class of diffuse Non-Lambertian surfaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(2):133–152, 1991.

[23] L. Wolff. Diffuse-reflectance model for smooth dielectric surfaces. *Journal of the Optical Society of America A*, 11(11):2956–2968, 1994.

[24] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[25] T. Zickler, P. Belhumeur, and D. Kriegman. Helmholtz stereopsis: exploiting reciprocity for surface reconstruction. *Int'l Journal of Computer Vision*, 49(2):215–227, 2002.

[26] T. Zickler, S. Mallick, D. Kriegman, and P. Belhumeur. Color subspaces as photometric invariants. *Int'l Journal of Computer Vision*, 2008.