

Uncalibrated Photometric Stereo by Stepwise Optimization Using Principal Components of Isotropic BRDFs

Keisuke Midorikawa Toshihiko Yamasaki Kiyoharu Aizawa
The University of Tokyo, Japan
{k.midor, yamasaki, aizawa}@hal.t.u-tokyo.ac.jp

Abstract

The uncalibrated photometric stereo problem for non-Lambertian surfaces is challenging because of the large number of unknowns and its ill-posed nature stemming from unknown reflectance functions. We propose a model that represents various isotropic reflectance functions by using the principal components of items in a dataset, and formulate the uncalibrated photometric stereo as a regression problem. We then solve it by stepwise optimization utilizing principal components in order of their eigenvalues. We have also developed two techniques that lead to convergence and highly accurate reconstruction, namely (1) a coarse-to-fine approach with normal grouping, and (2) a randomized multipoint search. Our experimental results with synthetic data showed that our method significantly outperformed previous methods. We also evaluated the algorithm in terms of real image data, where it gave good reconstruction results.

1. Introduction

Photometric stereo is a technique for shape estimation from the appearance of an object illuminated by various light sources and captured by a fixed camera. The original photometric stereo technique [32], proposed by Woodham, is able to treat only Lambertian targets and assumes that lighting positions are calibrated. Various types of extensions have been proposed such as the application to non-Lambertian objects [3, 2, 28, 29, 11, 13, 10], overcoming the necessity for lighting calibration [36, 5, 26, 22, 8, 15, 17, 16], adopting the environment of the perspective projection model [24, 20], and assuming near point lighting sources [25, 31, 35].

In this paper, we tackle uncalibrated photometric stereo for non-Lambertian reflection. Because the lighting directions and reflectance functions are involved as unknown parameters, this problem setting is quite challenging, with few studies reported to date. In [26, 15], it is assumed

that the similarity of the intensity profiles is in proportion to one of the normal orientations, which enables this problem to be solved up to a binary convex/concave ambiguity. Some methods [34, 17] enforce the symmetry of isotropic bidirectional reflectance distribution functions (BRDFs) about a half vector between a camera and lighting direction, which is observed in many materials. The most recent method [16] utilizes a reference data and formulates the uncalibrated photometric stereo as the matching problem. While these existing methods work well in some conditions, they strongly depend on various assumptions such as half-vector symmetry, good initial estimation, and particular lighting distribution, and they fail when these assumptions do not hold.

It is a common technique in calibrated photometric stereo to represent specular and retroreflection by parametric models such as the biquadratic model [28] and bivariate Bernstein polynomials [11]. Although they can be effective with known lighting directions, they are difficult to apply to uncalibrated cases because the solution space expands enormously with the number of unknowns.

In this paper, we propose a new method solving uncalibrated non-Lambertian photometric stereo, that formulates a regression problem using an image formation model based on principal component analysis (PCA). This is the first work that incorporates the image formation model into uncalibrated photometric stereo for general isotropic surfaces to our best knowledge. This method estimates the surface normals, lighting vectors, and reflectance properties simultaneously and accurately. We apply PCA to various BRDFs in a database and represent any unknown BRDF as a linear combination of principal components as in [19], thereby formulating the uncalibrated photometric stereo as a regression problem. We propose this strategy for two main reasons. First, we can approximate various kinds of materials well by using the unique model. Second, although it also has unknowns of the same number of basis functions, this approach enables stepwise optimization. We can roughly estimate surface normals, lightings, and basis coefficients by using only a few principal components and then

refine the estimations using more basis functions. By using this strategy, we can reduce the search space appropriately thanks to the property that principal components are ranked according to the order of their influence. In order to solve the minimizing problem, we have to deal with two difficulties: (1) discreteness, and (2) strong nonlinearity. For problem (1), we prepare an additional parametric model separate from the evaluation of the reprojection error and carry out an approximate gradient estimation. For problem (2), we propose a strategy based on a coarse-to-fine approach and multipoint searching. In experiments, we compare our proposal with recent techniques and our method shows state-of-the-art performance.

2. Related Work

The original photometric stereo [32] assumes known directional lighting and Lambertian reflectance, which is one of the simplest BRDF models. Subsequent methods have aimed at expanding photometric stereo for more complicated scenes and photographing environments.

Photometric stereo targeting a non-Lambertian object is challenging because its reflectance properties are too complicated and variable for accurate representation by a unique parametric model. Some methods treat non-Lambertian effects as outliers and remove them using techniques such as sparse regression [12], rank minimization [33], and color-space rotation [18]. Although these methods are not capable of treating nonsparse effects such as rough specularity, they can remove sharp specularity, digital noise, calibration error, inter-reflections, and other obstacles. Other methods formulate parametric models that can represent specularity, such as the Torrance–Sparrow model [9], the biquadratic model [28], and bivariate Bernstein polynomials [11]. These flexible models are capable of treating various reflections such as rough specularity or retroreflection.

Photometric stereo with uncalibrated lighting is also a challenging problem. Most of the methods proposed so far assume Lambertian surfaces because of their simplicity. In [36], it was shown that uncalibrated photometric stereo for a Lambertian object can be solved up to the generalized bas-relief (GBR) ambiguity [6] by enforcing integrability. Many subsequent methods have tackled the problem of GBR disambiguation, such as using color profiling [27], local maxima [8], and entropy minimization [3].

Fewer methods exist for uncalibrated photometric stereo with non-Lambertian objects. Instead of modeling BRDF functions parametrically, they use other information such as intensity profiles [26, 15], attached shadow [22], and the symmetry of isotropic BRDFs [34, 17]. [16] employs reference BRDFs and solves the matching problem. They do not represent BRDFs parametrically partly because it is difficult to achieve both an accurate approximation of a BRDF and an efficient optimization at the same time with the large

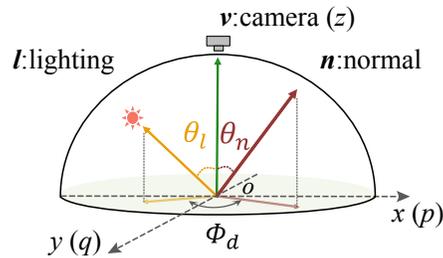


Figure 1: View centred coordinate system.

number of unknowns.

Our proposal method utilizes a BRDF dataset as in [16] but in an entirely different way. It is based on a representation of a BRDF by using principal components of items in a dataset as proposed in [19]. [19] conducts a dense measurement and constructs a data-driven BRDF model. They also apply PCA to the measured BRDFs in an effort to analyse BRDFs. We utilize this semi-parametric model for photometric stereo and formulate a regression problem, and recover surface normals by solving it. Our contributions are listed below.

1. We propose a new method solving uncalibrated non-Lambertian photometric stereo, that utilizes PCA and formulates a regression problem.
2. We propose an algorithm to enhance the global convergence of discrete and highly nonlinear minimization problems.
3. We expand our method for (1) calibrated photometric stereo and (2) uncalibrated photometric stereo for a target with spatially varying reflectance.
4. In our experiments we show that our proposal outperforms the state-of-the-art methods for both uncalibrated and calibrated photometric stereo for non-Lambertian BRDFs.

3. Photometric Stereo Based on PCA

In this section, we formulate the uncalibrated photometric stereo as a regression problem based on PCA. Our problem setting is described as follows:

1. The positions of the camera and the object are fixed across all images. We define the target’s position as the origin and the camera direction as the z -axis of the orthogonal coordinates (see Fig. 1). Although isotropic BRDFs depend on the camera direction v in addition to surface normals and lighting vectors, we treat v as a known fixed vector $([0, 0, 1]^T)$ and therefore do not specify it in the equations in this paper.
2. The object is illuminated by unknown and varying light sources. These lightings are directional.
3. The camera view is orthographic and the radiometric response function is linear.

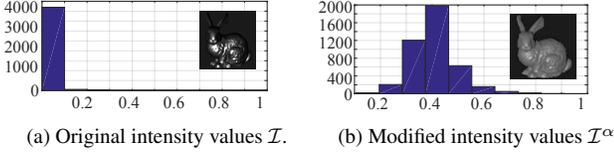


Figure 2: Histograms of intensity values (not including the shadow) and appearances of (a) original and (b) modified images respectively (in this example we used *alum-bronz* BRDF data in the MERL database [19], and set $\alpha = 0.2$). We can observe that the bias of the intensities is reduced.

4. In the preprocessing phase, we take

$$I = \mathcal{I}^\alpha, \quad (1)$$

where \mathcal{I} means an observation intensity and α ($0 < \alpha < 1$) is a correction factor that reduces the variance of intensity values (see Fig. 2). We adopt it because our method does not work well with the original intensity values. This is because of the large differences in order between intensities in the diffuse and specular regions. In [19], Matusik used the natural logarithm before applying PCA to the measured BRDFs and magnified the importance of the diffusion, for the same reason. Our preprocessing has a similar effect to the logarithmic approach but is more useful because it can be applied after normalizing the pixel values. Hereafter, we will describe the modified intensities simply as I or “intensities.”

3.1. Linear Combination Representation of Image Formation Model Based on PCA

The factor that complicates photometric stereo, regardless of whether lighting sources are calibrated or not, is that the reflectance of the target material is unknown. Therefore, the reprojection error cannot be defined uniquely. To overcome this ill-posedness we model the image formation model ($I = f(\mathbf{n}, \mathbf{l}) = \rho(\mathbf{n}, \mathbf{l})\mathbf{n}^T\mathbf{l}$, where ρ is a BRDF, \mathbf{n} is a surface normal vector, and \mathbf{l} is a lighting vector. \mathbf{n} and \mathbf{l} are normalized.) as a linear combination of basis functions generated by applying PCA to MERL database [19]. To apply PCA, which requires a discrete data structure, we prepare intensity set vectors for each material in the dataset. These are normalized column vectors whose elements are intensities parameterized by θ_n , θ_l , and ϕ_d (see Fig. 1 for these parameters).

$$F_b \triangleq \begin{bmatrix} I_b(\theta_n^0, \theta_l^0, \phi_d^0) \\ I_b(\theta_n^0, \theta_l^0, \phi_d^1) \\ \vdots \\ I_b(\theta_n^{\Theta_n-1}, \theta_l^{\Theta_l-1}, \phi_d^{\Phi_d-1}) \end{bmatrix} / \max(I_b(\theta_n, \theta_l, \phi_d)), \quad (2)$$

where b is a BRDF index and $I_b(\theta_n, \theta_l, \phi_d)$ is an intensity given by these parameters. $(\theta_n^0, \theta_l^0, \phi_d^0)$, $(\theta_n^0, \theta_l^0, \phi_d^1)$, \dots , $(\theta_n^{\Theta_n-1}, \theta_l^{\Theta_l-1}, \phi_d^{\Phi_d-1})$ are the sampling points, and Θ_n , Θ_l , Φ_d are the numbers of samples for each parameter. In other words, the length of F_b is $\Theta_n \cdot \Theta_l \cdot \Phi_d$. Then, by applying PCA to vectors of F_0, F_1, \dots, F_{B-1} we obtain B basis vectors that are of the same form as F_b . We represent them as U_0, U_1, \dots, U_{B-1} , where U_0 is an average of F_b , and U_1, \dots, U_{B-1} are the principal components in descending order of eigenvalue. We use the first M bases, and construct the model as

$$f(\mathbf{n}, \mathbf{l}) = \sum_{m=0}^{M-1} a_m u_m(\mathbf{n}, \mathbf{l}), \quad (3)$$

where a_m ($m=0, 1, \dots, M-1$) are the coefficients that determine the reflection properties and $u_m(\mathbf{n}, \mathbf{l})$ is a function which returns the element of U_m corresponding to the nearest sampling points from $(\theta_n, \theta_l, \phi_d)$. Under the assumption that any isotropic reflectance is represented well by this linear combination model, we formulate the uncalibrated photometric stereo problem as an energy minimization:

$$\min_{\mathcal{N}, \mathcal{L}, \mathbf{a}_M} E_M \triangleq \sum_{i=0}^{N-1} \sum_{j=0}^{L-1} \left| \sum_{m=0}^{M-1} a_m u_m(\mathbf{n}_i, \mathbf{l}_j) - I_{i,j} \right|^2, \quad (4)$$

where \mathcal{N} , \mathcal{L} , and \mathbf{a}_M denote sets of N surface normals, L lighting vectors, and M basis coefficients, respectively. $I_{i,j}$ is an intensity value given by the i -th surface normal \mathbf{n}_i and the j -th lighting vector \mathbf{l}_j .

Note that, although the result of minimizing Eq. (4) remains a rotation/flip ambiguity, it can be easily solved up to a convex/concave ambiguity by enforcing an integrability constraint. The detail about this disambiguation is described in our supplementary.

3.2. Approximating Gradient with a Piecewise Hyperplane Function

We minimize Eq. (4) by a gradient descent method that uses a Jacobian matrix. Because f is a discrete function, and therefore $\partial f / \partial \mathbf{n}$ and $\partial f / \partial \mathbf{l}$ cannot be calculated directly, we construct a parametric image formation model \tilde{f} in addition to f and calculate the Jacobian part corresponding to \mathbf{n} and \mathbf{l} by differentiating \tilde{f} instead of f to approximate the Jacobian.

We model a piecewise linear function $\tilde{u}(\mathbf{n}, \mathbf{l})$ parameterized by $\mathbf{n}^T \mathbf{l}$, $\mathbf{l}^T \mathbf{v}$, $\mathbf{n}^T \mathbf{v}$ and we obtain this hyperplane by parameter fitting, using only the first basis function u_0 :

$$u_0(\mathbf{n}, \mathbf{l}) \approx \tilde{u}(\mathbf{n}, \mathbf{l}) = \mathbf{A}\mathbf{n}^T \mathbf{l} + \mathbf{B}\mathbf{l}^T \mathbf{v} + \mathbf{C}\mathbf{n}^T \mathbf{v} + D, \quad (5)$$

as illustrated in Fig. 3. Using this hyperplane we approximate as $f(\mathbf{n}, \mathbf{l}) \approx a_0 u_0(\mathbf{n}, \mathbf{l}) \approx a_0 \tilde{u}(\mathbf{n}, \mathbf{l}) \triangleq \tilde{f}(\mathbf{n}, \mathbf{l})$,

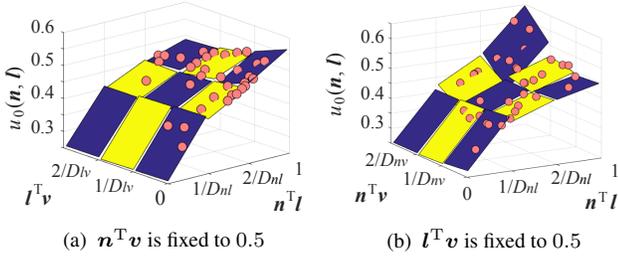


Figure 3: Our piecewise hyperplane model is fitted to collections of $\{\mathbf{n}^T \mathbf{l}, \mathbf{l}^T \mathbf{v}, \mathbf{n}^T \mathbf{v}, u_0(\mathbf{n}, \mathbf{l})\}$ in each region. For visualization, we plotted for two cases: (a) $\mathbf{n}^T \mathbf{v}$ is fixed, but $\mathbf{n}^T \mathbf{l}$ and $\mathbf{l}^T \mathbf{v}$ vary. (b) $\mathbf{l}^T \mathbf{v}$ is fixed, but $\mathbf{n}^T \mathbf{l}$ and $\mathbf{n}^T \mathbf{v}$ vary. In this case, $D_{nl} = D_{lv} = D_{nv} = 3$.

thereby performing the Jacobian approximation. This is because the influence of u_0 covers most of the approximation of Eq. (3), and u_0 is monotonic for $\mathbf{n}^T \mathbf{l}, \mathbf{l}^T \mathbf{v}, \mathbf{n}^T \mathbf{v}$, as shown in Fig. 3. Although this is a coarse approximation, because overfitting can be avoided with this low dimensionality, the gradient direction estimate is unlikely to be mistaken.

We divide the domains of each parameter ($\mathbf{n}^T \mathbf{l}, \mathbf{l}^T \mathbf{v}, \mathbf{n}^T \mathbf{v}$) into D_{nl}, D_{lv}, D_{nv} segments, respectively. The partitions of each parameter domain are set to d_{nl}/D_{nl} ($d_{nl}=1, \dots, D_{nl}-1$), d_{lv}/D_{lv} ($d_{lv}=1, \dots, D_{lv}-1$), and d_{nv}/D_{nv} ($d_{nv}=1, \dots, D_{nv}-1$). In every combination of segments we approximate u_0 with a hyperplane by linear least-squares. In other words, we construct $D_{nl} \cdot D_{lv} \cdot D_{nv}$ hyperplanes. The coefficients A, B , and C are fitted in every region and stored beforehand. They are just looked up and utilized to the Jacobian approximation.

Note that, for the purpose of enforcing a unit-size constraint on surface normals and lighting vectors, we optimize them in the gradient space (p, q) , defined as

$$(p_{n,i}, q_{n,i}) = \left(\frac{n_{x,i}}{n_{z,i}}, \frac{n_{y,i}}{n_{z,i}} \right), \quad (p_{l,j}, q_{l,j}) = \left(\frac{l_{x,j}}{l_{z,j}}, \frac{l_{y,j}}{l_{z,j}} \right), \quad (6)$$

where $n_{x,i}, n_{y,i}, n_{z,i}$ and $l_{x,j}, l_{y,j}, l_{z,j}$ are the three elements of the i -th normal vector \mathbf{n}_i and the j -th lighting vector \mathbf{l}_j , respectively.

Then, a row of the Jacobian corresponding to $e_{i,j} \triangleq f(\mathbf{n}_i, \mathbf{l}_j) - I_{i,j}$ of Eq. (4) comprises

$$\begin{aligned} \begin{bmatrix} \partial e_{i,j} \\ \partial p_{n,i} \end{bmatrix}, \begin{bmatrix} \partial e_{i,j} \\ \partial q_{n,i} \end{bmatrix} &\approx a_0 \frac{\partial \tilde{u}}{\partial \mathbf{n}_i} \begin{bmatrix} \partial \mathbf{n}_i \\ \partial q_{n,i} \end{bmatrix} \\ &= a_0 \left(A_{nl,lv,nv} \mathbf{l}_j^T + C_{nl,lv,nv} \mathbf{v}^T \right) \begin{bmatrix} \partial \mathbf{n}_i \\ \partial p_{n,i} \end{bmatrix}, \end{aligned} \quad (7)$$

$$\begin{aligned} \begin{bmatrix} \partial e_{i,j} \\ \partial p_{l,j} \end{bmatrix}, \begin{bmatrix} \partial e_{i,j} \\ \partial q_{l,j} \end{bmatrix} &\approx a_0 \frac{\partial \tilde{u}}{\partial \mathbf{l}_j} \begin{bmatrix} \partial \mathbf{l}_j \\ \partial q_{l,j} \end{bmatrix} \\ &= a_0 \left(A_{nl,lv,nv} \mathbf{n}_i^T + B_{nl,lv,nv} \mathbf{v}^T \right) \begin{bmatrix} \partial \mathbf{l}_j \\ \partial q_{l,j} \end{bmatrix}, \end{aligned} \quad (8)$$

$$\begin{bmatrix} \partial e_{i,j} \\ \partial a_0 \end{bmatrix}, \dots, \begin{bmatrix} \partial e_{i,j} \\ \partial a_{M-1} \end{bmatrix} = [u_0(\mathbf{n}_i, \mathbf{l}_j), \dots, u_{M-1}(\mathbf{n}_i, \mathbf{l}_j)], \quad (9)$$

where $A_{nl,lv,nv}, B_{nl,lv,nv}, C_{nl,lv,nv}$ are the coefficients looked up depending on the domain of the current estimation. Because the regions corresponding to $\mathbf{n}^T \mathbf{l} \leq 0$ and $0 < \mathbf{n}^T \mathbf{l} \leq 1/D_{nl}$ are included in the same domains, $\partial \tilde{u}/\partial \mathbf{n}$ and $\partial \tilde{u}/\partial \mathbf{l}$ have nonzero values even when $\mathbf{n}^T \mathbf{l}$ is estimated as negative, which is an advantage of this method of calculation. We optimize Eq. (4) by using the Levenberg–Marquardt method [21]. Although our method even works well from the initial surface normals which all are set as $\mathbf{n} = [0, 0, 1]^T$ by combining the methods presented in Sections 3 and 4, we initialize the surface normals by using the method in [15] for more robust estimation. We also initialize the lighting vectors to $\mathbf{l}_j = [0, 0, 1]^T$ ($j = 0, 1, \dots, L-1$) and the basis coefficients to $\mathbf{a}_M = [1, 0, \dots, 0]^T$.

3.3. Refinement by Increasing the Number of Basis Functions

As the number of basis functions increases, Eq. (3) becomes a good approximation of the image formation model but it has strong nonlinearity and the optimization is therefore likely to get stuck in local minima. We therefore propose to use stepwise optimization. We first find an intermediate solution by minimizing Eq. (4) with $M=1$ (only u_0 is being used). We then use the coarsely estimated surface normals, lightings, and a_0 as the next starting point to solve Eq. (4) with $M=2$. In this way, we can refine the estimation sequentially while adding basis functions one at a time.

4. Optimization

Although we can obtain good estimation by using the method presented in Section 3, the method can suffer from local-minima problems, depending on the initial estimation or various parameter values. To obtain robust and accurate estimation, we propose a simple but very powerful strategy that enhances convergence to the better solution.

4.1. Coarse-to-fine Using Normal Grouping

We can use a coarse-to-fine approach, given that it can be redundant and costly to use all terms of Eq. (4) in a one-time optimization. To prune out similar surface normals, we adopt the normal grouping technique proposed in [27] at the beginning of the minimization process. This reduction in the number of unknowns gives improved convergence and shorter computational time, which enhances the multipoint

search described in Section 4.2. We first perform K-means clustering using the correlations of the intensity profiles, as in [27], and select the nearest pixels to the center of each cluster. Second, we estimate the normals of selected points and lighting vectors by minimizing Eq. (4), and set the surface normals that were not selected at the normal grouping step to the estimated normal vectors of the representative points of each cluster. Third, we refine all the normal vectors with lighting vectors fixed by the results of the second step. Finally, we refine all the normals and lighting vectors simultaneously.

4.2. Randomized Multipoint Search

In addition to the stepwise refinement approach, we adopt a multipoint search strategy in each optimization process (Fig. 4). At the beginning of the minimization with M basis functions, we scatter P starting points by randomly shifting all the normals and lighting vectors that have been estimated provisionally. We then carry out the optimization from each starting point and choose the result that has the lowest energy as the revised estimation. In the random-shifting process, the i -th normal vector of the p -th starting point $\mathbf{n}_i^{p,M}$ is calculated as

$$\mathbf{n}_i^{p,M} = \text{vec} \left(\tilde{\theta}_{n,i}^{M-1} + \mathcal{R}(r_{\theta_n}), \tilde{\phi}_{n,i}^{M-1} + \mathcal{R}(r_{\phi_n}) \right), \quad (10)$$

where $\text{vec}(\theta, \phi)$ denotes a three-dimensional normalized vector whose elevation angle and azimuth angle are θ and ϕ , respectively. $\tilde{\theta}_{n,i}^{M-1}$ and $\tilde{\phi}_{n,i}^{M-1}$ are the elevation and azimuth angles of the i -th normal vector which was chosen in the optimization with $M - 1$ basis functions, and $\mathcal{R}(r)$ is an uniform random number whose range is $[-r, r]$. The lighting vector setup is performed similarly. To guarantee at least one result for the optimization at M with energy less than the best estimation at $M - 1$, we set one starting point to be the same as the best estimation at $M - 1$.

This strategy increases the computational cost in proportion to the number of initial points, and the number of points required increases with the solution space of the problem. Therefore, this strategy works harmoniously with the normal-grouping method and they can be adopted together.

5. Extension of Our Method

Calibrated Photometric Stereo

One of the advantages of our method is that it is capable of solving photometric stereo problems in cases where the lighting sources are calibrated simply by fixing the lighting vectors of Eq. (4) to the true directions. Whereas it is possible to estimate the surface normals either simultaneously or per pixel, the former approach works well only when the target object has uniform reflectance.

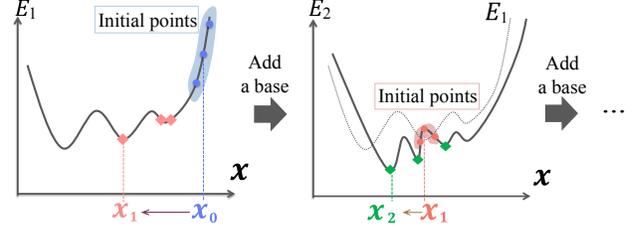


Figure 4: When adding the next basis function, we scatter multiple initial points, and perform an optimization from each point. We then simply select the result with the lowest energy.

The pixelwise calibrated photometric stereo problem is solved by optimizing:

$$\min_{\mathbf{n}, \mathbf{a}_M} E_M^{cali} \triangleq \sum_{j=0}^{L-1} \left| \sum_{m=0}^{M-1} a_m u_m(\mathbf{n}, l_j) - I_j \right|^2, \quad (11)$$

where I_j is the j -th intensity value of the target pixel.

Uncalibrated Photometric Stereo for Spatially Varying Reflectance

The weak aspect of our method is that the model in Eq. (3) is based on the assumption that the target has uniform reflectance. To address this point, we combine the albedo grouping introduced in [27] and Eq. (11). First, we perform the albedo grouping and select a cluster for which most pixels are included, to collect pixels having the same reflectance. Second, we solve Eq. (4) only for selected pixels and lighting vectors by using the methods presented in Sections 3 and 4. Third, the remaining pixels are estimated by optimizing Eq. (11) using the estimated lighting vectors.

6. Experimental Results

In this section, we evaluate our proposal and compare it with previous methods using both synthetic and real-image data. All experiments were performed on an Intel Xeon X5650 (2.67 GHz) machine with 64 GB RAM and were implemented in MATLAB. For the synthetic experiments, we generated 32-bit high-dynamic-range images of Stanford bunny (96×96) with foreground masks rendered by using different BRDFs and lighting sources placed at random. In Sections 6.1, 6.2 and 6.4 we performed the experiments with 45 images. The correction factor α in Eq. (1) was determined manually and set as 0.2. In Eq. (2), we set the number of samples for each parameter as $\Theta_n = \Theta_l = 91$, $\Phi_d = 181$, and defined the sampling points as $(\theta_n^r, \theta_l^s, \phi_d^t) = (r, s, t)$ (degrees). The source material used for the PCA was all 100 items in the MERL dataset in all experiments except as described in Section 6.4, and we used ten basis functions for the approximation Eq. (3) in all experiments except as described in Section 6.2. The number of division for each parameter

of the hyperplane (Section 3.2) were $D_{nl}=D_{lv}=D_{nv}=3$. In Eq. (10) we set the range of scattering searching points to $r_{\theta_n}=r_{\phi_n}=r_{\theta_l}=r_{\phi_l}=1^\circ$ in all experiments except as described in Section 6.5. In Sections 6.1, 6.2, 6.3 and 6.5 we use five common parametric BRDF models: Cook-Torrance [7], Ward [30], Lafortune [14], Ashikhmin-Shirley [4], and Oren-Nayar [23]¹. In Sections 6.1, 6.2, 6.3 and 6.4 we tested our method eight times for each condition and used the averages (we also recorded the standard deviations, worst scores, and best scores for some cases).

6.1. Parameter Test: The Number of Clusters for the Normal Grouping and Search Points

We performed an experiment to evaluate differences in the accuracy of the final results and the computational time for various numbers of normal clusters K (see Section 4.1) and searching points P (see Section 4.2). For each combination of K and P , we tested our method eight times with each BRDF and took the average of these 40 mean angular errors (MAEs) and computational times (5 BRDF models \times 8 trials gives 40 results). The MAE means the average of errors between ground-truth normal maps and the recovered ones. Here, the computational time includes the times for both the normal grouping and the minimization, but does not include preparing the initial estimation by using the method in [15].

The experimental results are illustrated in Fig. 5. Note that the right-hand-side ends of both contour graphs correspond to the results without the normal grouping. It is of interest that the MAE decreases with K when K is relatively large, even with $P=1$. We conjecture that the normal grouping results in a reduction of not only the computational time but also the solution space. On the other hand, the accuracy is poor when K is small because of the overfitting.

We set parameter values $P=20$ and $K=500$ for the experiments described in Sections 6.2, 6.3, 6.4, and 6.6, as a trade-off between accuracy and computational time.

6.2. Parameter Test: the Number of Basis Function

In this section we evaluated our method for various number of the basis functions M for the five parametric models so as to find the valid number of the basis functions for an accurate reconstruction.

The results are shown in Fig. 6. It is observed that the accuracy converges with the number of basis functions M beyond ten. Thus, we set M to ten in our experiments.

6.3. Evaluation with Parametric BRDF

To check if the basis functions we trained by using the MERL database were useful in other applications, together with the dependency on the number of input images, we

¹ The concrete forms and parameter values of these models are the same as in the experiments in [11]. Please refer their supplementary material.

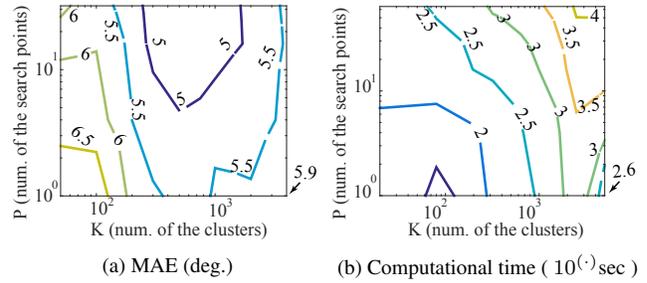


Figure 5: Evaluation of the accuracy and computational time of our method for various values of K and P .

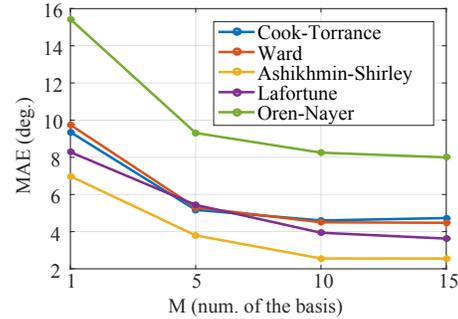


Figure 6: Evaluation of performance against the number of basis functions.

tested our method with various images rendered via parametric BRDF models. In this experiment we compared our method with those in Sato2007 [26], Lu2013 [15], and Lu2015 [17].² Here, we used results of [26] as the initial estimations of the method in [17]. Because [17] does not work well with a Lambertian-like BRDF, they proposed a method of skipping such cases (in this case, the Lafortune and Oren-Nayar models). However, for comparison, we tested [17] in all cases without skipping.

The experimental results are illustrated in Fig. 7. The figure also shows the position of the light sources (in the upper right corner of each bar graph). From these results, it is clear that our method outperformed previous methods in all cases. The MAEs of the other methods are more than 10 degrees in almost all cases because they work well only with lighting sources that are spaced uniformly and widely.

6.4. Evaluation with MERL BRDFs

To evaluate our method for various cases, we examined the performance of our method in terms of images rendered by 100 different BRDFs in the MERL database. In this experiment we used 99 BRDFs for PCA, omitting the one which was used as test images (leave-one-out cross valida-

² We executed MATLAB codes provided by the authors without changing their parameters.

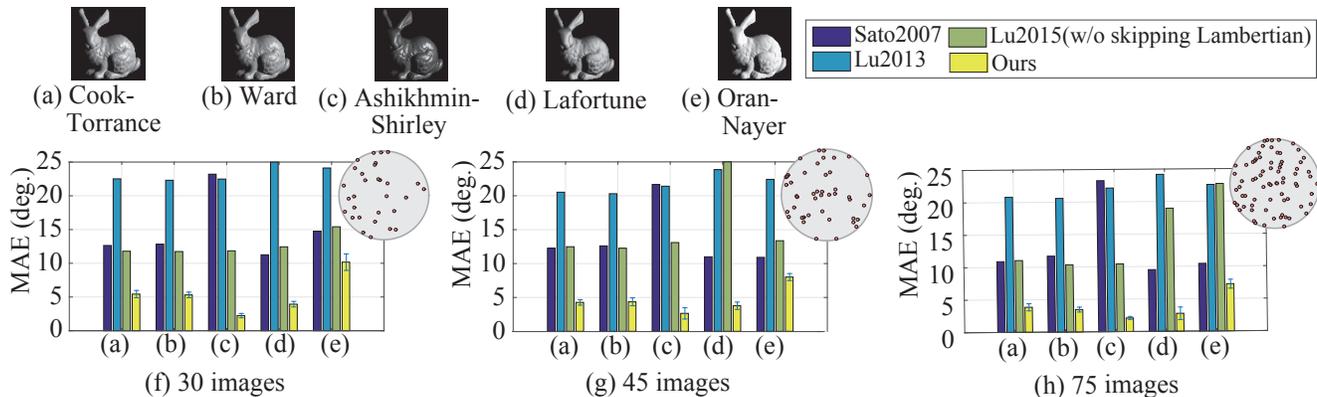


Figure 7: Comparison of the MAEs between some previous methods and our method for various BRDF models and numbers of images. Errors larger than 25 degrees are cut. The blue error bars denote the standard deviations.

tion). In this experiment, our method was compared with those in [26, 15, 17].²

The experimental results are shown in Fig. 8. The average MAEs over the 100 cases were 20.2° (Sato2007 [26]), 22.8° (Lu2013 [15]), 13.8° (Lu2015 [17]) and 4.03° (mean of the eight trials of our method), respectively. Our method outperformed the other methods.

It is reported that the average error of [16] was 4.25° with 89 images of *sphere* dataset.³ Although it is competitive with ours, their method has the assumptions of the uniform light source positions and the presence of one lighting which is parallel to the viewing vector. We emphasize that our method works well, even with the crude initial estimation and randomly placed light sources, because of our efficient optimization process and fewer assumptions.

6.5. Evaluation for Calibrated Photometric Stereo

We evaluated our method for the case when the lighting sources are calibrated. We compared it with a baseline method of Ikehata2014 [11], which is a highly accurate method that adopts a purely pixelwise approach.² Here we used 100 and 300 input images rendered by using five parametric BRDF models. To make a fair comparison, we solved this problem per pixel by solving Eq. (11), and took the MAEs of the foreground pixels. Instead of using previous methods, we initialized all the surface normals to $\mathbf{n} = [0, 0, 1]^T$. We set the parameter values $P = 5$ and $r_{\theta_n} = r_{\phi_n} = 10^\circ$.

As shown in Fig. 9, our method outperformed [11] for four BRDFs except for the Lafortune model with both 100 and 300 images. These results show the approximation Eq. (3) is accurate enough to apply to photometric stereo.

We also performed an evaluation with 45 images rendered by using 100 MERL BRDFs as Section 6.4. The aver-

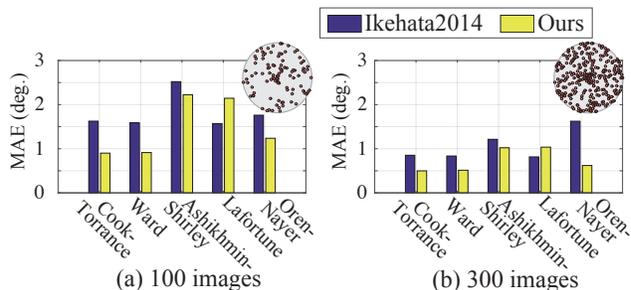


Figure 9: Experimental results for the calibrated case.

age MAEs over 100 BRDFs were 4.79° (Ikehata2014 [11]) and 1.32° (ours), respectively. The detail of this experiment is described in the supplementary material.

6.6. Evaluation with Real Image Data

We conducted experiments for real images. Here we used 50 images of *knight*, 60 images of *alien*, and 44 images of *fatguy* (this dataset is from [12]), illuminated by uncalibrated directional lighting. For every dataset, we adopted the strategy for spatially varying reflectance presented in Section 5. The number of clusters for the albedo grouping and the threshold for shadow removal were selected manually for each dataset.

Estimated normal maps and the reconstructed meshes are shown in Fig. 10. Here, we used the method in [1] to reconstruct the depth maps. For comparison, we also show the results of [15, 17].² It is clear that our method can work well for real scenes with different kinds of materials and spatially varying albedos.

7. Conclusion

In this paper, we have presented a method for uncalibrated photometric stereo that is based on PCA and additional techniques for enhancing the convergence of

³ Because we could not access the code of [16] and compare it with ours, we only describe a qualitative analysis.

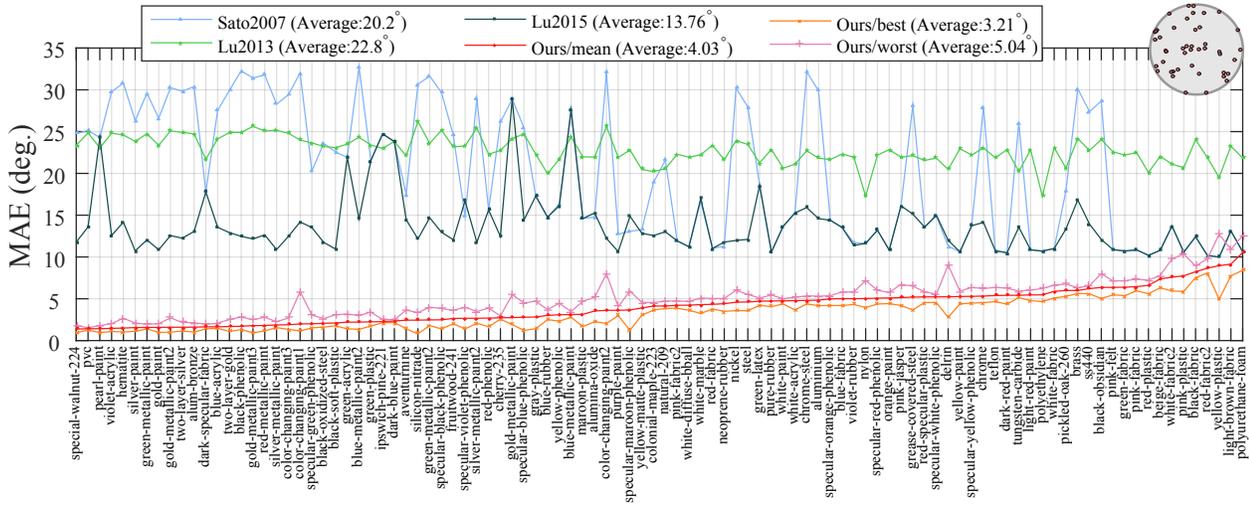


Figure 8: Experimental results with 100 MERL BRDFs (45 images each, *bunny*). For each BRDF, we tested our method eight times and plotted the mean, the worst, and the best result. The results are listed in ascending order of “Ours/mean”.

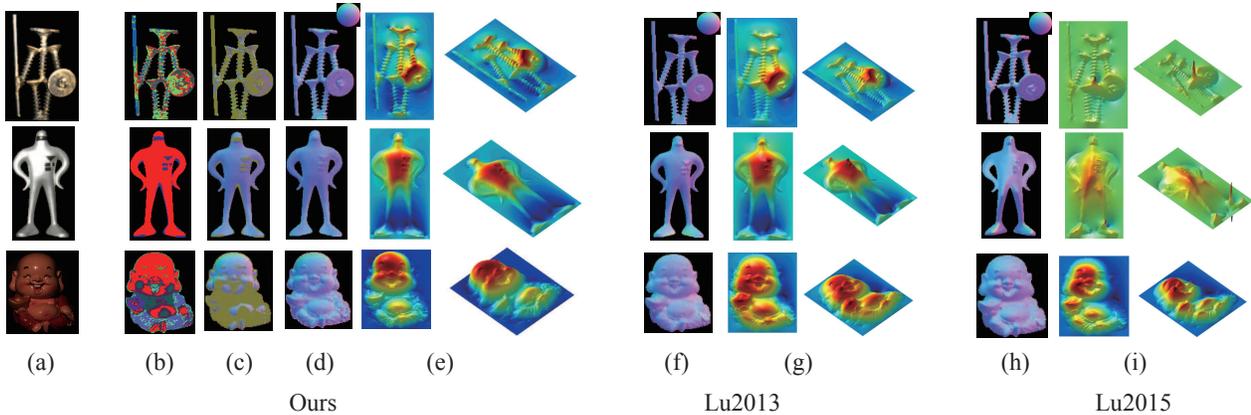


Figure 10: Experimental results for real images. We illustrate (a) an example of input images, (b) the result of albedo grouping (the red regions were selected for the first calculation and the remainder were estimated pixelwise), (c) the result of optimization for selected regions, (d) the eventually recovered normal map, and (e) surface meshes generated from the normal maps. (f), (g), (h), (i) For a comparison we also illustrate the results of [15, 17].

the nonlinear regression problem. This is a derivation of formulating a regression problem based on an image-formation model for non-Lambertian, uncalibrated photometric stereo. We have shown that our method works well for various cases involving isotropic materials. A current limitation is that our method depends on the bias of training database, and we have not performed sufficient tests with measured datasets other than MERL.

8. Acknowledgments

We wish to thank Feng Lu for making his CVPR13, 15 codes available and Imari Sato for giving us access to a photographic studio and equipment.

References

- [1] A. Agrawal, R. Raskar, and R. Chellappa. What is the range of surface reconstructions from a gradient field? In *Proc. ECCV*. 2006. 7
- [2] N. Alldrin, T. Zickler, and D. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In

- Proc. CVPR*, 2008. 1
- [3] N. G. Alldrin, S. P. Mallick, and D. J. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In *Proc. CVPR*, 2007. 1, 2
- [4] M. Ashikhmin and P. Shirley. An anisotropic phong brdf model. *JGT*, 5(2):25–32, 2000. 6
- [5] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *IJCV*, 72(3):239–257, 2007. 1
- [6] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *IJCV*, 35(1):33–44, 1999. 2
- [7] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *Computer Graphics*, 15(3):307–316, 1981. 6
- [8] P. Favaro and T. Papadimitri. A closed-form solution to uncalibrated photometric stereo via diffuse maxima. In *Proc. CVPR*, 2012. 1, 2
- [9] A. S. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. ICCV*, 2003. 2
- [10] Z. Hui and A. C. Sankaranarayanan. A dictionary-based approach for estimating shape and spatially-varying reflectance. In *Proc. ICCP*, 2015. 1
- [11] S. Ikehata and K. Aizawa. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In *Proc. CVPR*, 2014. 1, 2, 6, 7
- [12] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *Proc. CVPR*, 2012. 2, 7
- [13] C. Inoshita, Y. Mukaigawa, Y. Matsushita, and Y. Yagi. Surface normal deconvolution: Photometric stereo for optically thick translucent objects. In *Proc. ECCV*. 2014. 1
- [14] E. P. Lafortune, S.-C. Foo, K. E. Torrance, and D. P. Greenberg. Non-linear approximation of reflectance functions. In *Proc. SIGGRAPH*, 1997. 6
- [15] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato. Uncalibrated photometric stereo for unknown isotropic reflectances. In *Proc. CVPR*, 2013. 1, 2, 4, 6, 7, 8
- [16] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato. From intensity profile to surface normal: photometric stereo for unknown light sources and isotropic reflectances. *TPAMI*, 37(10):1999–2012, 2015. 1, 2, 7
- [17] F. Lu, I. Sato, and Y. Sato. Uncalibrated photometric stereo based on elevation angle recovery from brdf symmetry of isotropic materials. In *Proc. CVPR*, 2015. 1, 2, 6, 7, 8
- [18] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *Proc. CVPR*, 2005. 2
- [19] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A data-driven reflectance model. *ToG.*, 22(3):759–769, 2003. 1, 2, 3
- [20] R. Mecca, G. Rosman, R. Kimmel, and A. M. Bruckstein. *Perspective photometric stereo with shadows*. Springer, 2013. 1
- [21] J. J. Moré. *The Levenberg-Marquardt algorithm: implementation and theory*. Springer, 1978. 4
- [22] T. Okabe, I. Sato, and Y. Sato. Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions. In *Proc. ICCV*, 2009. 1, 2
- [23] M. Oren and S. K. Nayar. Generalization of lambert’s reflectance model. In *Proc. SIGGRAPH*, 1994. 6
- [24] T. Papadimitri and P. Favaro. A new perspective on uncalibrated photometric stereo. In *Proc. CVPR*, 2013. 1
- [25] F. Sakaue and J. Sato. A new approach of photometric stereo from linear image representation under close lighting. In *Proc. ICCV Workshops*, 2011. 1
- [26] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape reconstruction based on similarity in radiance changes under varying illumination. In *Proc. ICCV*, 2007. 1, 2, 6, 7
- [27] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *Proc. CVPR*, 2010. 2, 4, 5
- [28] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. A biquadratic reflectance model for radiometric image analysis. In *Proc. CVPR*, 2012. 1, 2
- [29] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances. In *Proc. ECCV*. 2012. 1
- [30] G. J. Ward. Measuring and modeling anisotropic reflection. *Computer Graphics*, 26(2):265–272, 1992. 6
- [31] A. Wetzler, R. Kimmel, A. M. Bruckstein, and R. Mecca. Close-range photometric stereo with point light sources. In *Proc. 3DV*, 2014. 1
- [32] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Opt. Engg.*, 19(1):139–144, 1980. 1, 2
- [33] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proc. ACCV*. 2011. 2
- [34] Z. Wu and P. Tan. Calibrating photometric stereo by holistic reflectance symmetry analysis. In *Proc. CVPR*, 2013. 1, 2
- [35] W. Xie, C. Dai, and C. C. L. Wang. Photometric stereo with near point lighting: A solution by mesh deformation. In *Proc. CVPR*, 2015. 1
- [36] A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur. Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability. *IJCV*, 35(3):203–222, 1999. 1, 2