Self-calibrating Photometric Stereo

Boxin Shi[1*] Yasuyuki Matsushita[2] Yichen Wei[2] Chao Xu[1] Ping Tan[3]

[1]Key Lab of Machine Perception (MOE), Peking University
[2]Microsoft Research Asia
[3]Department of Electrical & Computer Engineering, National University of Singapore

Abstract

We present a self-calibrating photometric stereo method. From a set of images taken from a fixed viewpoint under different and unknown lighting conditions, our method automatically determines a radiometric response function and resolves the generalized bas-relief ambiguity for estimating accurate surface normals and albedos. We show that color and intensity profiles, which are obtained from registered pixels across images, serve as effective cues for addressing these two calibration problems. As a result, we develop a complete auto-calibration method for photometric stereo. The proposed method is useful in many practical scenarios where calibrations are difficult. Experimental results validate the accuracy of the proposed method using various real-world scenes.

1. Introduction

Photometric stereo [23] recovers surface orientations and albedos from a set of images taken from a fixed viewpoint under different lighting directions. There are two necessary calibrations for a standard photometric stereo algorithm to work correctly. First, the camera needs to be radiometrically calibrated to obtain the scene irradiance from measured pixel values. Second, lighting directions and intensities need to be known to uniquely determine the surface. With these two calibrations, surface orientations and albedos can be estimated uniquely from three images for Lambertian scenes.

Typically, these calibrations are performed separately before applying photometric stereo algorithms. The radiometric camera calibration is often performed as a completely independent preprocessing step by, for example, using multiple images [15] captured from a fixed viewpoint under a static lighting configuration with different exposure times. In practice, however, this involves an additional data measurement that is not always possible. The lighting direc-

tions and intensities are often recorded by putting a mirror sphere and a Lambertian sphere in the scene. These additional calibration objects make the process complicated and could even introduce interreflections that eventually increase error. On the other hand, if the lighting parameters are *unknown*, surface normals of a Lambertian scene can only be recovered up to a 3×3 linear ambiguity [10]. If the surface is known to be integrable, this ambiguity can be reduced to the generalized bas-relief (GBR) ambiguity [3]. Without additional prior knowledge about the scene, or the lighting, this shape/lighting ambiguity cannot be resolved. Hence, these two calibrations limit the practical applications of photometric stereo.

In this paper, we address these two calibration problems without using any additional input data, or calibration objects. From a set of images taken from a fixed viewpoint under *unknown* and varying lightings with an *unknown* camera response function, our method automatically recovers both the camera's radiometric response function and the unknown lighting directions and intensities. Our method uses *color* and *intensity profiles* to achieve this goal. A color profile is defined as the set of measured RGB colors at a pixel across images. A color profile draws a 3D curve in the RGB color space. An intensity profile is defined as an ordered sequence of scene irradiance at a pixel across images. An intensity profile corresponds to a 2D curve in the irradiance-time space, similar to the appearance profiles shown by Koppal and Narasimhan [12]. Our method uses color profiles to determine the radiometric response function, and intensity profiles to resolve the GBR ambiguity. Putting them together, we build a self-calibrating photometric stereo method.

The rest of the paper is organized as follows. After discussing prior work, we describe the problem and overview of our approach in Section 2. Section 3 describes the proposed radiometric calibration method based on the color profiles, and Section 4 explains the method for GBR disambiguation using intensity profiles. Section 5 shows the experimental results followed by discussions.
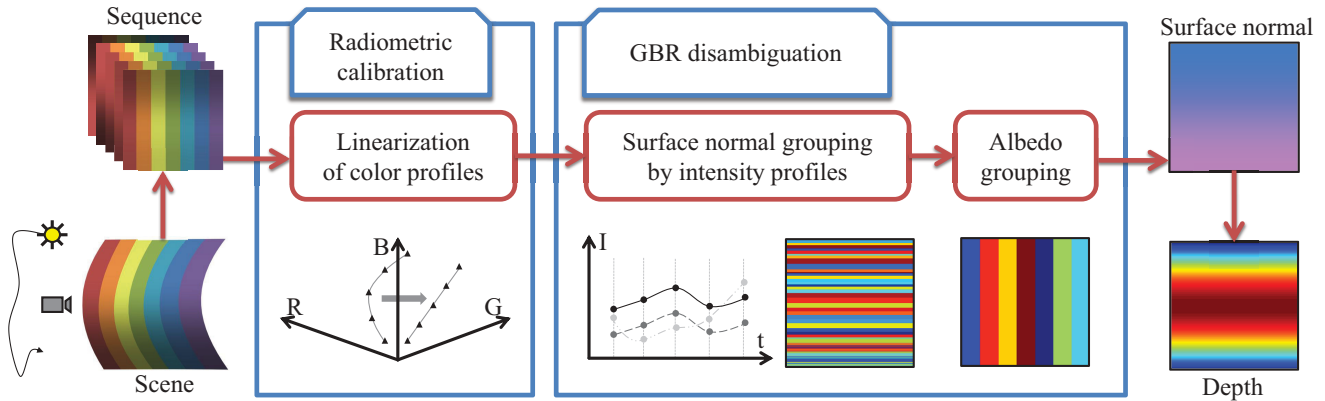
Figure 1. Overview of the proposed method. The proposed method automatically performs radiometric self-calibration and GBR disambiguation from the input images and gives estimated surface normals and depth maps as output.

1.1. Related works

Our work is related to both radiometric calibration and GBR disambiguation. Several methods have been proposed to estimate camera response functions. One of the most widely used methods is Mitsunaga and Nayar's method [15], which uses multiple images of a static scene with multiple different known exposure times. More recent methods work under more general imaging conditions. Lin *et al*. [13] proposed a method that only requires a single color image. Later they extended it to work with a single gray-scale image [14]. These methods are based on color/intensity mixtures on edge boundaries. There are a few works that allow scene/camera movements during data capturing. Kim and Pollefeys [11] computed pixel correspondence across image frames for radiometric calibration. Grossberg and Nayar [8] used intensity histograms of two images of different exposure rather than exact pixel correspondence. Wilburn *et al*. [22] showed the response function can be recovered from a motion blur in a single image. Unlike all of these methods, our radiometric calibration method uses photometric stereo images captured by a static camera under varying lighting conditions.

The GBR ambiguity is an intrinsic shape/lighting ambiguity for uncalibrated photometric stereo with unknown lighting parameters. This ambiguity arises because different combinations of shapes and lightings can produce the same appearance. To resolve the GBR ambiguity, additional cues are necessary. Prior knowledge about surface albedo can be used to resolve the GBR [10, 1]. Chandraker *et al*. [4] used interreflections in a Lambertian scene to resolve the ambiguity. It has been shown that the GBR can also be resolved if the surface reflectance is known to be Lambertian compounded with a Torrance-Sparrow specular reflection [7], or with specular reflections containing specular spikes [6, 5]. More recently, Tan *et al*. [20, 21] showed that the GBR can be reduced and resolved in a stratified manner for isotropic surfaces. Besides resolving the GBR ambiguity, Basri *et al*. [2] proposed a general solution under environment lighting.

Furthermore, to handle various BRDFs, Sato *et al*. [19] suggested using the similarity in radiance changes and Okabe *et al*. [17] utilized the attached shadow. Our method takes a different approach from these. As we will see later in this paper, our method can achieve a high-accuracy estimation using the intensity profile as a cue.

The radiometric calibration and GBR disambiguation have been discussed separately in prior works. To achieve a complete auto-calibration in photometric stereo, we solve the mixture of these problems. We take a fundamentally different approach to achieve this goal using the color and intensity profiles.

2. Problem statement

We perform photometric stereo with images taken from a fixed viewpoint under varying and unknown directional lighting (both intensities and directions are unknown) using a camera where the response function is also unknown. We address the two calibration problems in this setting, *i.e.*, radiometric camera calibration and GBR disambiguation.

In this paper, we use the following imaging model:

$$M = f(I) = f\big((\rho\boldsymbol{n}) \cdot (E\boldsymbol{l})\big), \tag{1}$$

where M is the recorded pixel value, and ρ and $\boldsymbol{n}$ are the surface reflectance (or albedo) and the surface normal vector at a pixel. E and $\boldsymbol{l}$ are the intensity and direction of the directional lighting. I is the irradiance, and f is the camera's radiometric response function. Since the albedo ρ and surface normal vector $\boldsymbol{n}$ are location dependent and lighting intensity E and direction $\boldsymbol{l}$ are image dependent, we use subscripts i and j to index pixels and images respectively. For multiple images under varying directional lighting, the imaging model can be written as

$$M_{ij} = f(I_{ij}) = f\big((\rho_i\boldsymbol{n}_i) \cdot (E_j\boldsymbol{l}_j)\big). \tag{2}$$

Our goal is to estimate $\rho_i, \boldsymbol{n}_i, E_j, \boldsymbol{l}_j$ and f from the observations M_{ij}.

Radiometric response function The goal of radiometric calibration is to estimate the radiometric response function f that maps the irradiance I to a pixel intensity M. As the response function f is a monotonic function, there is always a unique inverse mapping $g(= f^{-1})$. Radiometric calibration amounts to estimating the inverse response function g.

The generalized bas-relief ambiguity As in previous work [10], the photometric stereo problem is formulated in a matrix form:

$$\mathbf{I} = \mathbf{SL}, \tag{3}$$

where $\mathbf{I}$ is a $P \times F$ matrix containing irradiance values from all images, and P and F are the number of pixels and images respectively. Each row of $\mathbf{I}$ corresponds to a pixel position in an image, and each column corresponds to a different image. The surface matrix $\mathbf{S} \in \mathbb{R}^{P \times 3}$ represents surface properties, *i.e.*, the i-th row of $\mathbf{S}$ encodes the surface normal at the i-th pixel scaled by its albedo as $\mathbf{S}_{i*} = \rho_i \boldsymbol{n}_i$. The lighting matrix $\mathbf{L} \in \mathbb{R}^{3 \times F}$ represents the lighting directions and intensities, *i.e.*, the j-th column of the matrix $\mathbf{L}$ corresponds to the lighting direction in the j-th image scaled by its intensity as $\mathbf{L}_{*j} = E_j \boldsymbol{l}_j$. When $\mathbf{L}$ is unknown, $\mathbf{S}$ and $\mathbf{L}$ can be estimated only up to a GBR transformation $\mathbf{G}$ by enforcing integrability over the surface as $\mathbf{I} = \mathbf{SL} = (\mathbf{S}'\mathbf{G})(\mathbf{G}^{-1}\mathbf{L}')$. The GBR transformation $\mathbf{G}$ is a three-parameter matrix:

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{pmatrix}, \quad \mathbf{G}^{-1} = \frac{1}{\lambda} \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ -\mu & -\nu & 1 \end{pmatrix}. \tag{4}$$

The overview of the proposed method is illustrated in Fig. 1. Our method uses color profiles for radiometric calibration (Section 3) and intensity profiles for resolving GBR ambiguity (Section 4).

3. Radiometric calibration using color profiles

This section describes our method for determining the radiometric response function using color profiles. A color profile is a set of RGB color values at the same coordinate in a set of photometric stereo images. Our method capitalizes on the fact that color profiles form straight lines in the RGB color space when the radiometric response function f is linear. On the other hand, if f is nonlinear, color profiles form nonlinear curves in the RGB space.

To understand this relationship between color profiles and the radiometric response function f, let us consider the ratio between red, green and blue colors at a pixel. In a color image, Eq. (2) is evaluated at each color channel, *i.e.*, $M^r = f(I^r)$, $M^g = f(I^g)$ and $M^b = f(I^b)$, where superscripts r, g, and b correspond to the RGB color channels. If the radiometric response function f is linear, *i.e.*,
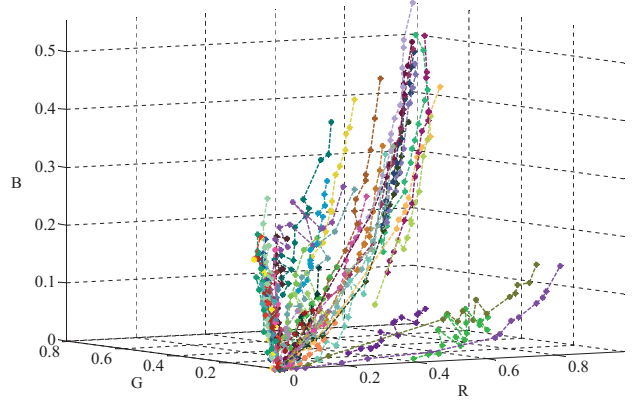

Figure 2. Color profiles of a real-world scene taken by a camera with a non-linear response function. The profiles show clear curves in the RGB space.

$f(x) = kx$, the color ratio at a Lambertian point is

$$(M^r : M^g : M^b) = (kI^r : kI^g : kI^b) = (\rho^r : \rho^g : \rho^b), \tag{5}$$

which is constant over the photometric stereo image sequences. Hence, in this case, a color profile is a straight line with direction (ρ^r, ρ^g, ρ^b) passing through the origin in the RGB space. On the other hand, if f is nonlinear, this ratio will change over the image sequence. As a result, the color profile bends to a nonlinear curve in the RGB space. This result is consistent with the empirical observation in [18]. Examples of color profiles obtained from a real-world scene are shown in Fig. 2.

The correct inverse response function g ($g(M) = I$) should linearize the curved color profiles. Using this property as a constraint, we estimate the inverse response function g that maps curved color profiles to straight lines. We first define a nonlinearity measurement for color profiles to assess how the profiles are bended. We then formulate the estimation problem in an energy minimization framework to derive the inverse response function g.

Degree of nonlinearity We assess the degree of nonlinearity of a transformed color profile by its line fitting error. Let us denote the RGB triplet of the measured color value as $\boldsymbol{M}$ and use $g(\boldsymbol{M})$ to represent that the inverse response function g is applied to each element of $\boldsymbol{M}$ independently. For a color profile at the i-th pixel, it is computed as

$$D_i(g) = \sum_j \frac{\|(\boldsymbol{X}_i^0 - \boldsymbol{X}_i^1) \times (g(\boldsymbol{M}_{ij}) - \boldsymbol{X}_i^1)\|}{\|\boldsymbol{X}_i^0 - \boldsymbol{X}_i^1\|}, \tag{6}$$

where $\boldsymbol{X}_i^0$ and $\boldsymbol{X}_i^1$ are two points in the RGB space that define the line best fitting to the transformed color profile $g(\boldsymbol{M}_{i*})$. Intuitively, D measures the summed Euclidean distance between $g(\boldsymbol{M}_{ij})$ to the fitted line. It approaches to zero when g is the correct inverse function.

Energy function We compute the inverse response function g from a set of color profiles Ω by minimizing the nonlinearity. We use a polynomial representation for the inverse response function to reduce the computational complexity as in [15]. Thus, the inverse response function g is described by a small number of unknown coefficients c_k in the form of K-order polynomial as

$$I = g(M) = \sum_{k=0}^{K} c_k M^k. \tag{7}$$

We estimate the coefficients $\boldsymbol{c} = \{c_0, ..., c_K\}$ and fitted lines $\mathcal{X} = \{\boldsymbol{X}_i^0, \boldsymbol{X}_i^1 | i \in \Omega\}$ by minimizing the summed nonlinearity for each color profile $i \in \Omega$:

$$\{\hat{\boldsymbol{c}}, \hat{\mathcal{X}}\} = \operatorname*{argmin}_{\boldsymbol{c}, \mathcal{X}} \frac{1}{|\Omega|} \sum_{i \in \Omega} D_i(\boldsymbol{c}), \tag{8}$$

where $|\Omega|$ is the number of color profiles used for the computation. The minimization is performed with normalized color values in the range $[0, 1]$.

For minimization, we put two constraints on the inverse radiometric response function g, namely monotonicity and boundary conditions, which are general for any inverse response functions. The monotonicity constraint is represented as $\partial g / \partial M > 0$, and the boundary conditions are described as $g(0) = 0$ and $g(1) = 1$. We use these boundary conditions as a hard constraint by setting $c_0 = 0$ and $c_K = 1 - \sum_{k=1}^{K-1} c_k$ because $g(1) = \sum_{k=1}^{K} c_k = 1$. The final formulation of the energy minimization becomes

$$\{\hat{\boldsymbol{c}}, \hat{\mathcal{X}}\} = \operatorname*{argmin}_{\boldsymbol{c}, \mathcal{X}} \frac{1}{|\Omega|} \sum_{i \in \Omega} D_i(\boldsymbol{c}) + \lambda \sum_t H\left(-\frac{\partial g(t)}{\partial M}\right). \tag{9}$$

Here, $H(\cdot)$ is the Heaviside step function ($H(x) = 1$ when $x \geq 0$, and $H(x) = 0$, otherwise). The derivative $\partial g / \partial M$ are assessed at various t in the range of $[0, 1]$.

Optimization of Eq. (9) is performed with the Nelder-Mead simplex method [16] due to its simplicity, but other similar methods could be used as well. We initialize g as a linear function and $\mathcal{X}$ as least square fitted lines from color profiles.

4. GBR disambiguation using intensity profiles

This section describes the proposed method for resolving the GBR ambiguity using intensity profiles. By Hayakawa's singular value decomposition (SVD) method [10], the irradiance $\mathbf{I}$ in Eq. (3) can be factorized into a surface component $\mathbf{S}$ and a lighting component $\mathbf{L}$, up to a linear ambiguity. If albedos are known at 6 different normals or lighting intensity is known at 6 images, this linear ambiguity can be reduced to a rotational ambiguity. On the other hand,
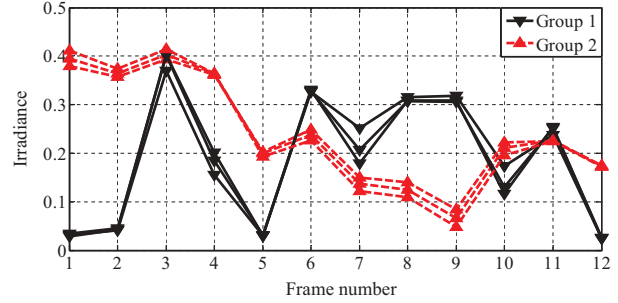

Figure 3. Intensity profiles from a real-world scene. Group 1 and Group 2 represent two different surface normal groups. Within each group, the intensity profiles bear the similar shapes.

Yuille *et al.* [25] showed the original linear ambiguity can be reduced to a GBR ambiguity by enforcing the integrability constraint.

We observe that the intersection of the rotational group and the GBR transformation group contains only the identity matrix. Hence, if both constraints from [10] and [25] can be used, the original linear ambiguity is fully resolved. To do so, we first apply the integrability constraint to derive a result up to the GBR ambiguity. We then resolve the GBR ambiguity by identifying pixels that have different normals but with the same albedo.

Suppose we are given pixels that have the same albedo a but different normals. The GBR transformation should satisfy $||\boldsymbol{s}_i \mathbf{G}|| = a$, *i.e.*,

$$\mathbf{s}_i \mathbf{C} \mathbf{s}_i^T = a^2, \tag{10}$$

where

$$\mathbf{C} = \mathbf{G}\mathbf{G}^{\mathbf{T}} = \begin{pmatrix} 1 & 0 & \mu \\ 0 & 1 & \nu \\ \mu & \nu & \mu^2 + \nu^2 + \lambda^2 \end{pmatrix}, \tag{11}$$

and $\boldsymbol{s}_i$ is the i-th row vector of $\mathbf{S}$. Let $\widetilde{\lambda} = \mu^2 + \nu^2 + \lambda^2$ be an auxiliary variable. The constraint in Eq. (10) is linear in all 4 unknowns $\{\mu, \nu, \widetilde{\lambda}, a^2\}$. Therefore, given at least 4 pixels with different normals but with the common albedo, the GBR ambiguity can be resolved. Our goal is to automatically identify these pixels to resolve the GBR ambiguity.

Surface normal grouping Following Koppal and Narasimhan's work [12], we use intensity profiles to cluster surface normals. As shown in Fig. 3, pixels with the same surface normal show strong correlation in their intensity profiles, while pixels with different surface normals become less correlated. Since we work in the irradiance domain, instead of using the extrema of the profiles as in [12], we simply use the correlation of the entire profiles. This allows us to use fewer images for surface normal clustering. The Pearson's correlation $r_{ii'}$ between two intensity profiles $\mathbf{I}_i$

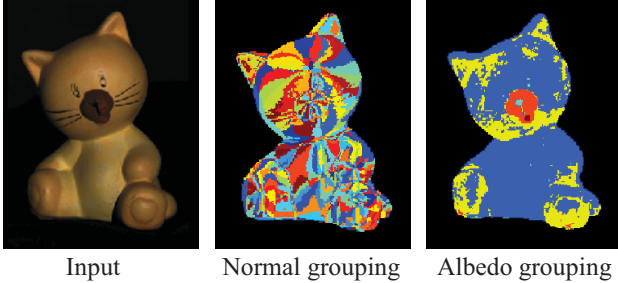| Input | Normal grouping | Albedo grouping |

Figure 4. Surface normal grouping (middle) and albedo grouping (right) on the CAT scene. One of the input images is shown on the left. Different colors indicate different cluster labels.

and $\mathbf{I}_{i'}$ is assessed by

$$r_{ii'} = \frac{\sum_{j=1}^{F} (I_{ij} - \bar{I}_i)(I_{i'j} - \bar{I}_{i'})}{(F-1)\sigma_i \sigma_{i'}}, \qquad (12)$$

where $\bar{I}_i$ is the mean value of the i-th row vector of matrix $\mathbf{I}$, and σ_i is its standard deviation. Using the correlation as a distance, we perform K-means clustering to partition images into a set of groups that have similar surface normals. Hence, by selecting from different groups, we can obtain points with different normal directions. For this purpose, we use a large K. An example of the surface normal grouping result is shown in Fig. 4 (middle).

Albedo grouping   Our albedo grouping aims at identifying pixels with the same albedo. As we have discussed in Section 3, the RGB ratio of a Lambertian scene point should be constant ($\rho^r : \rho^g : \rho^b$) across images in the irradiance domain. Hence, we can approximately measure the similarity of pixel albedos by their difference of chromaticity. To reduce the influence of shadows and noise, we compute the chromaticity of each pixel from its averaged irradiance value $\bar{I}$. The chromaticity is defined as a vector of two elements $[\bar{I}^r/(\bar{I}^r + \bar{I}^g + \bar{I}^b), \ \bar{I}^g/(\bar{I}^r + \bar{I}^g + \bar{I}^b)]$. We then cluster pixels in the chromaticity domain to obtain pixels with similar albedo. Similar to the surface normal grouping, we apply K-means clustering using the Euclidean distance of chromaticity vectors. As a result, we obtain albedo groups as shown in Fig. 4 (right).

Pixel selection   We combine the two grouping results to automatically select pixels with the same albedo but different surface normals.

First, our method selects reliable albedo groups. All the albedo groups are sorted by their variances, and half of the groups with larger variances are eliminated. After that, in each albedo group, we further remove pixels that have a large distance from the group center in the chromaticity domain. Specifically, pixels that have a distance larger than the average are discarded. The selected groups are split into subgroups based on the surface normal grouping result. The resulting subgroups become the groups of pixels with the

Figure 5. Images of the real-world scene for experiments with the number of images used.

same albedo and the same surface normal. Since we aim at finding pixels with the same albedo but different surface normals, only one pixel from each subgroup is selected.

To implement this pixel selection algorithm, we have tested two different approaches. One is random selection from subgroups, and the other is an optimization approach. The second approach selects pixels that minimize the summation of the surface normal similarity among selected pixels. This optimization problem is discrete and can be efficiently solved using a belief propagation algorithm [24]. Although the optimization-based method theoretically guarantees to select pixels with diverse surface normals, we find that the random sampling produces as good result as the optimization-based method when the number of surface normal groups is high. For this reason, we use the random sampling method in this work.

Finally, a set of pixels are selected from R albedo groups of unknown albedo values $\{a_1, a_2, \ldots, a_R\}$. From these pixels, we obtain multiple linear equations Eq. (10) with unknowns $\{\mu, \nu, \widetilde{\lambda}, a_1^2, \cdots, a_R^2\}$ to resolve the GBR ambiguity by a least square solution.

5. Experiments

We use four real-world scenes, CAT, OWL (courtesy of Dan Goldman and Steven Seitz[1]), PILLOW and SHEEP scenes, to verify our method. The scenes are shown in Fig. 5. CAT and OWL are radiometrically calibrated, while PILLOW and SHEEP are recorded by a Nikon D70 and Canon 20D, respectively. Both of these cameras have nonlinear radiometric response functions. In this experiment, we first quantitatively evaluate the radiometric calibration part. Second, we use the CAT and OWL scenes to assess the proposed GBR disambiguation method. Finally, we evaluate the entire pipeline with both radiometric calibration and GBR disambiguation. We implement our method using Matlab. The optimization of Eq. (9) can be implemented simply by a Matlab built-in function "fminsearch". Throughout our experiments, we set $\lambda = 5000$, $|\Omega| = 100$ (100 color profiles), $K = 200$ for the surface normal grouping, and $K = 20$ for the albedo grouping in two K-means stages. With an Intel E6550 (2.33GHz) CPU, the entire process took 10 minutes without optimization.

Inverse response function (Nikon D70)

Nikon D70

RMSE:
0.0252

Disparity:
0.0387

Std. dev.:
0.0131

Inverse response function (Canon 20D)

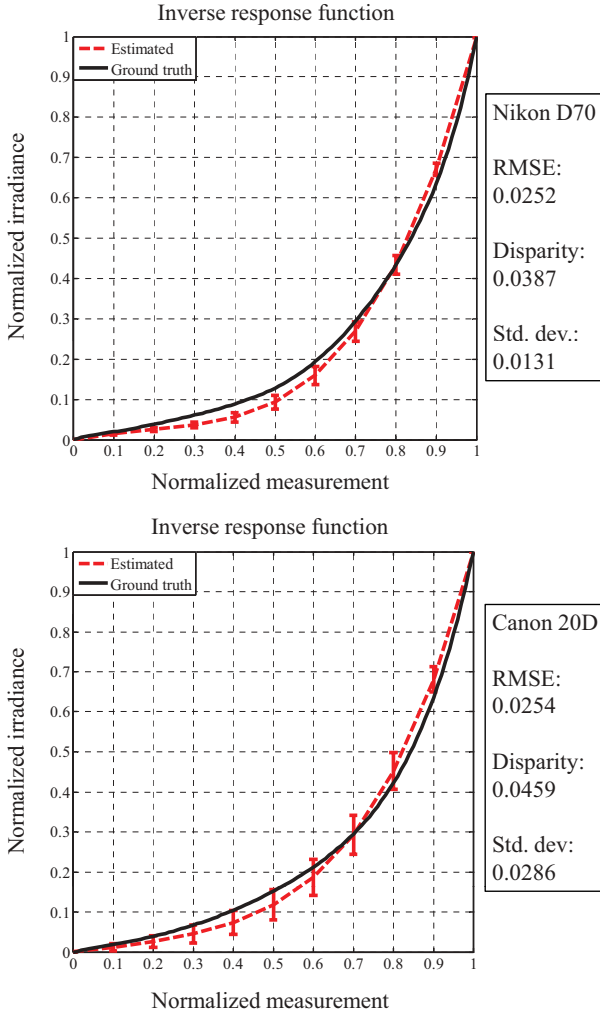Canon 20D

RMSE:
0.0254

Disparity:
0.0459

Std. dev:
0.0286

Figure 6. Result of radiometric calibration. The mean curve of estimated inverse response functions with standard deviation are shown in comparison with the ground truth.

Radiometric calibration result Using a Nikon D70 and Canon 20D, we take 50 datasets for testing the radiometric calibration method. From each dataset, we randomly select pixels to produce color profiles. The result is shown in Fig. 6 in comparison with the ground truth obtained by Mitsunaga and Nayar's method [15]. The averages of root mean squared error (RMSE) and disparity (the maximum difference) are computed from all the 50 datasets. The statistics show that our method consistently works well with various scenes and different cameras.

Surface normal and albedo grouping result We show an example of the grouping result of the SHEEP scene in Fig. 7. It depicts the accuracy of the grouping result for the following GBR disambiguation.

GBR disambiguation result We evaluate the proposed GBR disambiguation method using the CAT and OWL

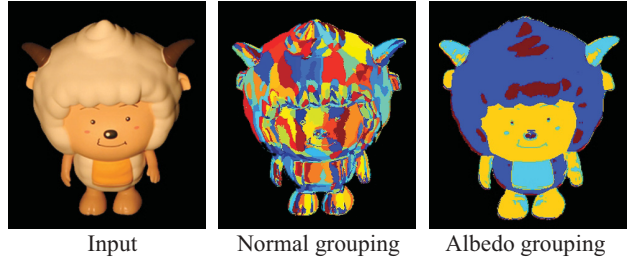| Input | Normal grouping | Albedo grouping |

Figure 7. Surface normal grouping (middle) and albedo grouping (right) on the SHEEP scene. One of the input images is shown on the left. Different colors indicate different cluster labels.

	Angular error	CAT	OWL
Our method	Mean	**6.15**	**10.47**
	Std. dev.	**2.83**	**4.75**
Minimum	Mean	15.82	27.10
entropy [1]	Std. dev.	6.15	6.08

Table 1. GBR disambiguation results. Mean and standard deviation of the angular error of the estimated normal maps. Our method gives lower errors compared to the minimum entropy method in these two scenes.

scenes. Fig. 8 and Fig. 9 show the results.[2] On the top row, the color coded normal maps (x, y, z components are linearly mapped in the RGB channel) are shown. On the bottom row is a depth map integrated from the normal map by Poisson solver. For a validation, we show the ground truth obtained by the calibrated photometric stereo on the left. We compare our method with the minimum entropy method [1]. Because our method uses surface normal groups in addition to the albedo information, it performs well in GBR disambiguation. The average angular errors are summarized in Table 1. Comparing the results between the CAT and OWL scenes, because of the smoothly varying albedo of OWL scene, the performances of both our method and the minimum entropy method degrade.

Self-calibrating photometric stereo result To test the entire pipeline, we use the PILLOW and SHEEP scenes taken by a Nikon D70 and Canon 20D, respectively. The results are shown in Fig. 10 and Fig. 11. In the figures, from left to right, we show the calibrated result (known lighting and response function), the result of our method with and without the radiometric calibration step. These results show that our method produces results that are close to the calibrated data, although we only use the input images taken under unknown lighting parameters with an unknown response function. The importance of radiometric calibration can also be seen clearly in the side-by-side comparison. Table 2 shows the quantitative comparison. Our self-calibration method significantly improves the accuracy. Fig. 12 shows the rendering of the reconstructed surface.

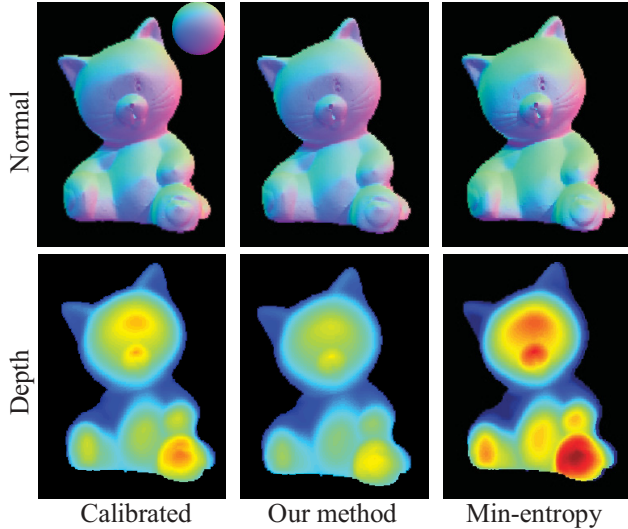[2]We recommend viewing the electronic version of this paper for better visualization.

Figure 8. GBR disambiguation result for CAT scene. From left to right, calibrated, our method and minimum entropy method [1] are shown. Top row shows the computed surface normal, and the bottom row shows the integrated depth maps. In the depth maps, the red color represents nearer, and blue color corresponds to further depth values. The reference sphere shows the normal directions.
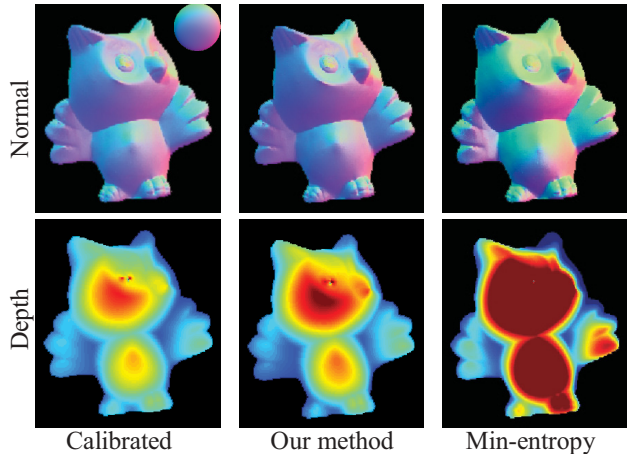

Figure 9. GBR disambiguation result for OWL scene. From left to right, calibrated, our method and minimum entropy method [1] are shown. Top row shows the computed surface normal, and the bottom row shows the integrated depth maps.

	Angular error	PILLOW	SHEEP
Our method	Mean	**5.63**	**7.30**
	Std. dev.	**3.71**	**3.02**
Without	Mean	16.08	20.18
radiometric calib.	Std. dev.	7.35	6.86

Table 2. Mean angular error and standard deviation of self-calibrating photometric stereo result, and the result without radiometric calibration.

6. Conclusion and discussion

Auto-calibration for photometric stereo is important for photometric stereo to be applicable in various settings.
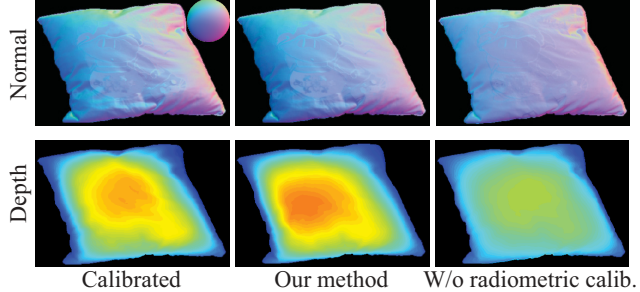

Figure 10. Result of the proposed method for PILLOW scene. From left to right, calibrated, our method with and without radiometric calibration.
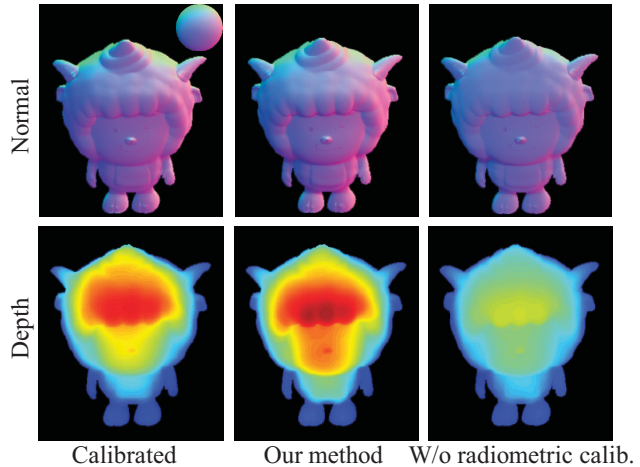

Figure 11. Result of the proposed method for SHEEP scene. From left to right, calibrated, our method with and without radiometric calibration.
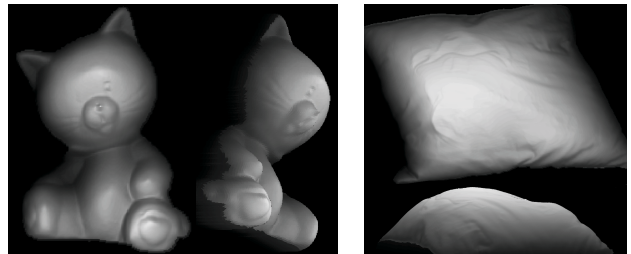

Figure 12. The surfaces of CAT and PILLOW scenes.

In this paper, we propose a self-calibrating photometric stereo method that handles both radiometric calibration and GBR disambiguation. The mixture of these problems is solved by analyzing color/intensity profiles in the RGB and irradiance-time domains. We show that radiometric calibration can be performed from images from a fixed viewpoint under varying lightings. We also develop a method to automatically resolve the GBR ambiguity from at least four pixels with different normals but the same albedo.

Limitations Our radiometric calibration method has a couple of limitations. First, it cannot work with gray-scale images and scenes. Second, it cannot handle a certain class of response functions f where $f(x)/f(ax) = \mathrm{const.}$ holds.

This is a special case when the response function is non-linear while the color profiles remain straight. This class of function includes $f(x) = x^n, n \in \mathbb{R}$, when boundary conditions $f(0) = 0$ and $f(1) = 1$ is imposed.

To assess the practical limitation of our method, we tested our method using the database of response functions (DoRF) [9] with a simulation dataset. We have confirmed that all the response functions in the database, including gamma correcting functions, bend the color profiles in the RGB color space. Among the degenerate case where the response function is in the form of $f(x) = x^n$, only the linear response function $f(x) = x$ is likely dominant in practice.[3] Therefore, by setting the initial guess of the response function as a linear function, our method works well for a wide variety of real-world camera response functions.

Acknowledgement

References

[1] N. Alldrin, S. P. Mallick, and D. J. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In *Proc. of Computer Vision and Pattern Recognition*, pages 1–7, 2007.

[2] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *Int'l Journal of Computer Vision*, 72(3):239–257, 2007.

[3] P. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *Int'l Journal of Computer Vision*, 35(1):33–44, 1999.

[4] M. Chandraker, F. Kahl, and D. J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *Proc. of Computer Vision and Pattern Recognition*, pages 788–795, 2005.

[5] O. Drbohlav and M. Chantler. Can two specular pixels calibrate photometric stereo? In *Proc. of Int'l Conf. on Computer Vision*, pages 1850–1857, 2005.

[6] O. Drbohlav and R. Sara. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. of European Conf. on Computer Vision*, pages 46–62, 2002.

[7] A. S. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. of Int'l Conf. on Computer Vision*, pages 816–823, 2003.

[8] M. D. Grossberg and S. K. Nayar. Determining the camera response from images: What is knowable? *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(11):1455–1467, 2003.

[9] M. D. Grossberg and S. K. Nayar. Modeling the space of camera response functions. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(10):1272–1282, 2004.

[10] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America*, 11(11):3079–3089, 1994.

[11] S. Kim and M. Pollefeys. Radiometric alignment of image sequences. In *Proc. of Computer Vision and Pattern Recognition*, pages 645–651, 2004.

[12] S. J. Koppal and S. G. Narasimhan. Clustering appearance for scene analysis. In *Proc. of Computer Vision and Pattern Recognition*, pages 1323–1330, 2006.

[13] S. Lin, J. Gu, S. Yamazaki, and H. Shum. Radiometric calibration from a single image. In *Proc. of Computer Vision and Pattern Recognition*, pages 938–945, 2004.

[14] S. Lin and L. Zhang. Determining the radiometric response function from a single grayscale image. In *Proc. of Computer Vision and Pattern Recognition*, pages 66–73, 2005.

[15] T. Mitsunaga and S. K. Nayar. Radiometric self-calibration. In *Proc. of Computer Vision and Pattern Recognition*, pages 374–380, 1999.

[16] J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7(4):308–313, 1965.

[17] T. Okabe, I. Sato, and Y. Sato. Attached shadow coding: estimating surface normals from shadows under unknown reflectance and lighting conditions. In *Proc. of Int'l Conf. on Computer Vision*, pages 1693–1700, 2009.

[18] I. Omer and M. Werman. Color lines: Image specific color representation. In *Proc. of Computer Vision and Pattern Recognition*, pages 946–953, 2004.

[19] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape reconstruction based on similarity in radiance changes under varying illumination. In *Proc. of Int'l Conf. on Computer Vision*, pages 1–8, 2007.

[20] P. Tan, S. P. Mallick, L. Quan, D. J. Kriegman, and T. Zickler. Isotropy, reciprocity and the generalized bas-relief ambiguity. In *Proc. of Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[21] P. Tan and T. Zickler. A projective framework for radiometric image analysis. In *Proc. of Computer Vision and Pattern Recognition*, pages 2977–2984, 2009.

[22] B. Wilburn, H. Xu, and Y. Matsushita. Radiometric calibration using temporal irradiance mixtures. In *Proc. of Computer Vision and Pattern Recognition*, pages 1–7, 2008.

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

[24] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. In *Proc. Int'l Joint Conf. on Artificial Intelligence*, pages 239–269, 2001.

[25] 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. *Int'l Journal of Computer Vision*, 35(3):203–222, 1999.

---

[3] The linear response as well as other response functions in the form of $f(x) = x^n$ does not exist in the DoRF.