# Photometric Stereo Under Non-uniform Light Intensities and Exposures

Donghyeon Cho[1(✉)], Yasuyuki Matsushita[2], Yu-Wing Tai[3], and Inso Kweon[1]

[1] KAIST, Daejeon, South Korea
cdh12242@gmail.com, iskweon@kaist.ac.kr
[2] Osaka University, Suita, Japan
yasumat@ist.osaka-u.ac.jp
[3] SenseTime Group Limited, Beijing, China
yuwing@gmail.com

Abstract. This paper studies the effects of non-uniform light intensities and sensor exposures across observed images in photometric stereo. While conventional photometric stereo methods typically assume that light intensities are identical and sensor exposure is constant across observed images taken under varying lightings, these assumptions easily break down in practical settings due to individual light bulb's characteristics and limited control over sensors. Our method explicitly models these non-uniformities and develops a method for accurately determining surface normal without affected by these factors. In addition, we show that our method is advantageous for general photometric stereo settings, where auto-exposure control is desirable. We compare our method with conventional least-squares and robust photometric stereo methods, and the experimental result shows superior accuracy of our method in this practical circumstance.

Keywords: Photometric stereo · Shape estimation · Unknown light intensity and exposure · Surface normal

1 Introduction

Non-uniform light intensities and exposures across observed images are a practical and common circumstance in data acquisition for photometric stereo that uses multiple images under distinct light directions. For example, different light bulbs with different intensity characteristics may be used for illuminating a scene. Even with identical light bulbs, due to that scene radiance is determined by surface normal and light directions, auto-adjusted sensor exposure is desirable depending on the light directions to avoid over-/under-exposures, which results in non-uniform exposures (equivalently, non-uniform light intensities). Therefore, the capability of properly handling varying and unknown light intensities and exposures across observed images is an important feature for making photometric stereo practical.

Part of this work was done while the first author was an intern at Microsoft Research Asia.

The setting can be regarded as a "semi-calibrated" photometric stereo, where the light directions are known but their intensities are unknown. We argue that accurate light intensity calibration is practically a hard task to perform due to that the light bulb's luminous efficiency varies over time and quantization error in the measurement even with high-dynamic range imaging. This paper provides a way to bypass the difficult intensity calibration in photometric stereo.

In the Lambertian image formation model, a measured intensity m is written as

$$m_{i,j} = E_i \rho_j \mathbf{n}_j^\top \mathbf{l}_i, \tag{1}$$

where i and j are indices of light direction and pixel location, $\mathbf{l}_i, \mathbf{n}_j \in \mathbb{R}^{3 \times 1}$ are unit vectors of light direction and surface normal, $\rho_j \in \mathbb{R}$ is a Lambertian diffuse albedo, and $E_i \in \mathbb{R}$ is a light source intensity. In a matrix form for representing all pixels and light directions at a time, it can be written as

$$\mathbf{M} = \mathbf{E}\mathbf{L}\mathbf{N}^\top \mathbf{P}, \tag{2}$$

where $\mathbf{M} \in \mathbb{R}^{f \times p}$ is an observation matrix, $\mathbf{E}$ is an $f \times f$ diagonal light intensity matrix, $\mathbf{L} \in \mathbb{R}^{f \times 3}$ is a light direction matrix, $\mathbf{N} \in \mathbb{R}^{p \times 3}$ is a surface normal matrix, $\mathbf{P}$ is a $p \times p$ diagonal diffuse albedo matrix, and f and p are the number of images and pixels, respectively. Conventional photometric stereo [1] assumes that light source intensities are identical across images, where the matrix $\mathbf{E}$ becomes a scaled identity matrix ($\mathbf{E} = e\mathbf{I}$), and computes albedo-scaled surface normal $\mathbf{B}(= \mathbf{P}^\top \mathbf{N})$ by

$$e\mathbf{B}^{*\top} = \mathbf{L}^\dagger \mathbf{M}, \tag{3}$$

up to a scale ambiguity e, where the superscript † indicates a generalized inverse when $f \geq 3$.

Clearly, when the light source intensities are non-uniform or camera exposures vary across images, the assumption $\mathbf{E} = e\mathbf{I}$ does not hold, but instead its diagonal elements have individual scales. When this non-uniformity is present, the surface normal estimates by Eq. (2) naturally becomes biased by greater scales as illustrated in Fig. 1. While there are recently various robust estimation techniques used for photometric stereo [2–6], because the effect of non-uniform $\mathbf{E}$ neither increases the rank of the observation matrix nor sparsifies outliers, robust techniques such as rank minimization or ℓ_0-norm minimization techniques cannot resolve this issue. In the rest of the paper, we collectively call this problem setting, non-uniform light intensities and exposures across images, a *varying light intensity condition*, because they are both considered intensity scaling on individual images.

This paper considers a method to effectively deal with the non-uniform light intensities and exposures. The problem that we deal with in this paper is a bilinear problem written as following.
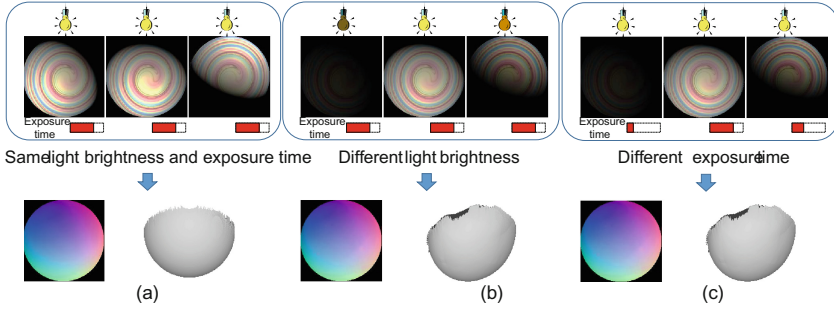
Fig. 1. (a) Conventional photometric stereo setting where constant light intensities and exposures are used, (b)(c): varying lighting intensity/exposure conditions. Estimated surface normal are biased toward brighter light source or images captured with longer exposures with a conventional solution method.

Problem (Photometric Stereo Under Varying Light Intensity Condition). *Given observations* **M** *and light directions* **L**, *estimate a diagonal light intensity matrix* **E** *and an albedo-scaled surface normal matrix* **B** *from the following relationship:*

$$\mathbf{M} = \mathbf{E}\mathbf{L}\mathbf{B}^\top. \tag{4}$$

We first show that there exists a linear closed-form solution method, which simultaneously estimates scales of light intensities (or exposures) **E** and albedo-scaled surface normal **B**. We call this method a linear joint estimation method. This method is straightforward to implement; however, inefficient in terms of computation time and memory consumption. We then introduce a factorization based method for determining only surface normal **B** in Eq. (4) without being affected by **E**. It bypasses the estimation of **E** using algebraic distance minimization (or, cross product minimization) by making the problem independent of vector magnitudes. Finally, we show that this bilinear problem can be efficiently solved by an alternating minimization technique that determines **E** and **B** in each step. We discuss details and characteristics of each method later in this paper.

We further show that our method is advantageous in improving signal-to-quantization-noise ratio (SQNR) in comparison to a standard photometric stereo method when an auto-exposure control is used, and as a result more accurate surface normal estimates can be obtained. Experimental results show the effectiveness of the proposed method in practical settings. In this paper, we assumes a directional light setting where radiance from a light source to a scene is constant except for shadowing, *i.e.*, spatially varying incident radiance within a scene is not assumed.

2   Related Works

Photometric stereo was first introduced by Woodham [1] in 1980's for determining surface normal from images taken under known and varying light directions with a Lambertian reflectance assumption. After Woodham's work, there have been various techniques proposed for making photometric stereo more practical.

Their main focuses are to relax the assumptions of (1) calibrated light sources and (2) Lambertian image formation model.

The first class of the methods, called uncalibrated photometric stereo, tries to eliminate the need for calibrating light directions. When the light directions are unknown, it is understood that the solution can be obtained up to a 3×3 linear ambiguity [7]. If the integrability [8] of the surface is assumed, it has been shown that the linear ambiguity can be reduced to a generalized bas-relief (GBR) ambiguity [9], which only has three parameters. To fully resolve these ambiguities, various types of external clues have been used. For example, there are methods that use the entropy of albedo distributions [10], specular observations [11], shadows [12], and groups of color and intensity profiles [13]. Our problem setting has a similarity to the uncalibrated photometric stereo scenario in that we relax the assumptions of known light intensities and *constant light intensity* across varying light directions. And there has not been uncalibrated PS works that derive disambiguated solution without external assumptions such as albedo entropy [10] and pixel profiles [13].

The second class of the methods tries to make photometric stereo applicable to non-Lambertian scenes. There are methods that use more sophisticated reflectance models than Lambertian reflectance model, such as the works that use Torrance-Sparrow [14,15], Cook-Torrance [16], Phong [17], Blinn-Phong [18]. More recently, Shi *et al.* [19] propose a bi-polynomial reflectance model that produces successful results for non-Lambertian diffusive scenes.

There are approaches that use robust estimation techniques by treating non-Lambertian reflectances and shadows as outliers. In [2], the robustness against outliers is achieved by capturing hundreds of input images coupled with Markov Random Field (MRF) to maintain neighborhood smoothness. Verbiest and Van Gool [3] use a confidence approach to reject outliers in input images of photometric stereo. Wu *et al.* [4] proposed a robust method based on low-rank matrix factorization. Oh *et al.* [5] introduced a partial sum of singular values for rank minimization, and showed good performance in photometric stereo. Ikehata *et al.* [20] used a sparse Bayesian regression for effectively neglecting sparse outliers (specularities and shadows). While these techniques are effective, they are built upon the assumption of *constant light intensity*, and cannot directly address the issue of varying light intensities and exposures.

3   Photometric Stereo Under Varying Light Intensity Conditions

As discussed in Eq. (4), we are interested in determining albedo-scaled surface normal **B** with unknown non-uniform scalings of light intensities or exposures **E**. In a least-squares framework, the problem can be written as

$$\{\mathbf{E}^*, \mathbf{B}^*\} = \operatorname*{argmin}_{\mathbf{E},\mathbf{B}} \|\mathbf{M} - \mathbf{E}\mathbf{L}\mathbf{B}^\top\|_{\mathrm{F}}^2 \qquad (5)$$

given the observations **M** and light directions **L**.

We first present a linear estimation method that simultaneously estimates $\mathbf{B}$ and $\mathbf{E}$ in Sect. 3.1. We then describe a factorization based method in Sect. 3.2, which bypasses the estimation of unknown scalings $\mathbf{E}$. Finally, we describe an efficient alternating minimization method in Sect. 3.3.

3.1   Linear Joint Estimation Method

The original form $\mathbf{M} = \mathbf{ELB}^\top$ can be re-written as $\mathbf{E}^{-1}\mathbf{M} = \mathbf{LB}^\top$, because $\mathbf{E}$ is always invertible as it is a positive diagonal matrix. Given known $\mathbf{M}$ and $\mathbf{L}$, it can be viewed as a variant of a Sylvester equation [21]:

$$\mathbf{E}^{-1}\mathbf{M} - \mathbf{LB}^\top = \mathbf{0}. \tag{6}$$

By vectorizing unknown variables $\mathbf{E}^{-1}$ and $\mathbf{B}^\top$, Eq. (6) can be written as

$$\mathrm{diag}(\mathbf{m}_1)|\cdots|\mathrm{diag}(\mathbf{m}_p)^\top \mathbf{E}^{-1}\mathbf{1} - (\mathbf{I}_p \otimes \mathbf{L})\mathrm{vec}(\mathbf{B}^\top) = \mathbf{0}, \tag{7}$$

where $\mathrm{diag}(\cdot)$, $\mathrm{vec}(\cdot)$ and $\otimes$ are diagonalization, vectorization, and Kronecker product operators, respectively. $\mathbf{I}_p$ is a $p \times p$ identity matrix, and $\mathbf{1}$ indicates a vector whose elements are all one. By concatenating matrices and vectors in Eq. (7), a homogeneous equation can be obtained:

$$\underbrace{\left[-\mathbf{I}_p \otimes \mathbf{L}\middle|\left[\mathrm{diag}(\mathbf{m}_1)|\cdots|\mathrm{diag}(\mathbf{m}_p)\right]^\top\right]}_{\mathbf{D}} \underbrace{\begin{bmatrix} \mathrm{vec}(\mathbf{B}^\top) \\ \mathbf{E}^{-1}\mathbf{1} \end{bmatrix}}_{\mathbf{y}} = \mathbf{0}, \tag{8}$$

where $\mathbf{D} \in \mathbb{R}^{pf \times (3p+f)}$ is a sparse design matrix and $\mathbf{y} \in \mathbb{R}^{(3p+f) \times 1}$ is an unknown vector. The homogeneous system always has a trivial solution $\mathbf{y} = \mathbf{0}$. To have a unique (up to scale) non-trivial solution, the matrix $\mathbf{D}$ should have a one dimensional null space, i.e., when rank of $\mathbf{D}$ is $(3p + f - 1)$, a unique solution can be obtained via singular value decomposition (SVD). The minimum condition to have a unique solution up to scale is $f \geq 5$ and $p \geq 3$, or $f = 4$ and $p \geq 2$. Unlike conventional photometric stereo, increasing the number of light directions does not necessarily make the problem easier in this setting, because it also increases the unknowns about light intensities.

3.2   Factorization Based Method

Although the linear joint estimation method is simple to implement, it has practical limitations in terms of its computational time and memory requirement when the sparse matrix $\mathbf{D}$ is large; not only constructing $\mathbf{D}$ but also computing SVD of $\mathbf{D}$. This limitation can be relaxed by dividing the observation matrix into small groups and deriving solutions for each group. However, this grouping should be performed carefully to avoid the condition numbers of divided submatrices to be high. The condition number increases when observations within each divided group are similar to each other, and as a result, the numerical error

becomes greater. To avoid these issues, we develop a factorization based method described in this section.

Like solution methods of uncalibrated photometric stereo, light directions and surface normal can be solved directly via matrix factorization:

$$\mathbf{M} = \hat{\mathbf{S}}\hat{\mathbf{B}}^{\top}, \tag{9}$$

where $\hat{\mathbf{S}}$ and $\hat{\mathbf{B}}$ are *biased* intensity-scaled light direction and albedo-scaled surface normal, respectively. With an arbitrary 3×3 non-singular matrix $\mathbf{H}$, Eq. (9) can be re-written as

$$\mathbf{M} = (\hat{\mathbf{S}}\mathbf{H})(\mathbf{H}^{-1}\hat{\mathbf{B}}^{\top}). \tag{10}$$

In our setting, since we know the light directions $\mathbf{L}$, we can find an appropriate non-singular matrix $\mathbf{H}$ for resolving the biases. Regardless of the effect of light intensities, direction of $\hat{\mathbf{S}}\mathbf{H}$ should be the same with $\mathbf{L}$. Thus, we can use this constraint, $(\hat{\mathbf{S}}\mathbf{H}) \times \mathbf{L} = \mathbf{0}$ where $\times$ indicates a cross product, for determining $\mathbf{H}$ as

$$\begin{bmatrix} 0 & -l_{i,3}\hat{\mathbf{s}}_i & l_{i,2}\hat{\mathbf{s}}_i \\ l_{i,3}\hat{\mathbf{s}}_i & 0 & -l_{i,1}\hat{\mathbf{s}}_i \end{bmatrix} \begin{bmatrix} \mathbf{h_1}^{\top} \\ \mathbf{h_2}^{\top} \\ \mathbf{h_3}^{\top} \end{bmatrix} = \mathbf{0}, \tag{11}$$

where $\mathbf{H} = [\mathbf{h_1}|\mathbf{h_2}|\mathbf{h_3}]$, $l_{i,*}$ and $\hat{\mathbf{s}}_i$ are the i-th row of $\mathbf{L}$ and $\hat{\mathbf{S}}$, respectively. The solution of Eq. (11) is unique up to scale when there are more than 4 distinct light directions. Using estimated $\hat{\mathbf{H}}$, we can compute unbiased albedo-scaled surface normal $\mathbf{H}^{-1}\hat{\mathbf{B}}^{\top}$. Interestingly, this factorization based method can naturally bypass the light intensity estimation; thus, it is suitable for our setting. Compared to the linear joint estimation method, the computational cost of the factorization based method is lower, even without dividing observations $\mathbf{M}$ into small groups.

3.3   Alternating Minimization Method

While the previous two methods are effective in ideal settings, they are prone to large errors due to un-modelled observations, such as shadows and pixel saturations. To avoid this problem, we develop a robust method that is based on alternating minimization for solving Eq. (5).

Our method computes albedo-scaled surface normal $\mathbf{B}^{(t)}$ and non-uniform scalings $\mathbf{E}^{(t)}$ in an alternating manner using their intermediate estimates from the previous iteration. Using $\mathbf{E}^{(t)}$ from the previous iteration and by fixing it, albedo-scaled surface normal $\mathbf{B}^{(t+1)}$ is updated by

$$\mathbf{B}^{(t+1)} = \underset{\mathbf{B}}{\mathrm{argmin}} \left\| \mathbf{M} - \mathbf{E}^{(t)}\mathbf{L}\mathbf{B}^{\top} \right\|_{\mathrm{F}}^{2}. \tag{12}$$

The above problem is a linear problem with respect to $\mathbf{B}$ and can be solved efficiently. Once matrix $\mathbf{B}^{(t+1)}$ is determined, $\mathbf{E}^{(t+1)}$ is then updated by solving

$$\mathbf{E}^{(t+1)} = \operatorname*{argmin}_{\mathbf{E}} \left\| \mathbf{M} - \mathbf{ELB}^{(t+1)^\top} \right\|_F^2. \tag{13}$$

Since matrix $\mathbf{E}$ is diagonal, each element $E_i^{(t+1)}$ is simply determined by

$$E_i^{(t+1)} = \frac{\sum_j m_{i,j}(\mathbf{l}_i^\top \mathbf{b}_j^{(t+1)})^\top}{\sum_j (\mathbf{l}_i^\top \mathbf{b}_j^{(t+1)})(\mathbf{l}_i^\top \mathbf{b}_j^{(t+1)})^\top}. \tag{14}$$

The initial scaling matrix $\mathbf{E}^{(0)}$ is set to an identity matrix, and the convergence criteria is defined by the magnitude of variation of matrix $\mathbf{B}$, i.e., $\|\mathbf{B}^{(t+1)} - \mathbf{B}^{(t)}\|_F < \epsilon$, where ϵ is set to a small value (in our implementation, $\epsilon = 1.0e\text{-}8$).

If we consider $\mathbf{E}$ as weights, this alternating minimization is similar to iteratively re-weighted least squares (IRLS) [22] except that weights are defined row-wise (each image has same weight). We show how the alternating method operates in the following. Let us consider updating $\mathbf{B}^{(t+1)}$ with fixing $\mathbf{E}^{(t)}$, then Eq. (12) becomes

$$\mathbf{B}^{(t+1)} = \operatorname*{argmin}_{\mathbf{B}} \left\| \mathbf{M} - \mathbf{E}^{(t)}\mathbf{LB}^\top \right\|_F^2 \tag{15}$$

$$= \operatorname*{argmin}_{\mathbf{B}} \left\| \mathbf{M} - \mathbf{E}^*\mathbf{LB}^\top - \mathbf{E}^r\mathbf{LB}^\top \right\|_F^2,$$

where $\mathbf{E}^{(t)} = \mathbf{E}^* + \mathbf{E}^r$, $\mathbf{E}^*$ is the ground truth (that we do not know), and $\mathbf{E}^r$ is the error from t-th iteration. It shows that the smaller the scaling error $\mathbf{E}^r$ is, the smaller objective cost becomes. The elements of $\mathbf{E}^{(t)}$ can also be written as

$$
\begin{aligned}
E_i^{(t)} &= \frac{\sum_j m_{i,j}(\mathbf{l}_i^\top \mathbf{b}_j^{(t)})^\top}{\sum_j (\mathbf{l}_i^\top \mathbf{b}_j^{(t)})(\mathbf{l}_i^\top \mathbf{b}_j^{(t)})^\top} \\
&= \frac{\sum_j m_{i,j}\mathbf{l}_i^\top \mathbf{b}_j^{*\top} + \sum_j m_{i,j}\mathbf{l}_i^\top \mathbf{b}_j^{r\top}}{\sum_j (\mathbf{l}_i^\top (\mathbf{b}_j^* + \mathbf{b}_j^r)^\top)(\mathbf{l}_i^\top (\mathbf{b}_j^* + \mathbf{b}_j^r)^\top)},
\end{aligned} \tag{16}
$$

where $\mathbf{b}^{(t+1)} = \mathbf{b}^* + \mathbf{b}^r$, $\mathbf{b}^*$ is the ground truth, and $\mathbf{b}^r$ is the error from t-th iteration. Since the denominator is fixed for all images, and the left-hand side of the numerator is proportional to the ground truth scaling $\mathbf{E}^*$, the smaller the error $\mathbf{b}^r$ becomes, the better scaling elements E becomes. To summarize, if the current estimate of albedo-scaled surface normal $\mathbf{B}^{(t)}$ is better than the previous one, $\mathbf{E}^{(t)}$ is better updated. In our experiments, above improvements are always observed since updated $\mathbf{E}^{(1)}$ becomes closer to the ground truth than $\mathbf{E}^{(0)}$. Then, $\mathbf{B}^{(t)}$ and $\mathbf{E}^{(t)}$ are alternately updated. The minimum condition for obtaining a stable solution is experimentally found to be $f \geq 5$ and $p \geq 3$.

4   Signal-to-Quantization-Noise Ratio Analysis

One of the important benefits of our method is its compatibility to the sensor's auto-exposure function that makes non-uniform scaling of observations. With auto-exposure, SQNR of observations is effectively increased by avoiding over-/under-exposures. As a result, the surface normal estimates are less suffered from quantization noise, and thus, a greater accuracy can be obtained. Based on the previous study of quantization noise [23], SQNR is written as

$$\text{SQNR} = \frac{\text{signal}}{\text{noise}} \propto \frac{C\mu}{\frac{CR}{Q}} = \frac{Q\mu}{R} = \frac{Q\mu}{(V_h - V_l)}, \tag{17}$$

where μ is the expectation of the signal, Q is the number of quantization levels, and C is a scaling factor representing the amount of exposure. Also, $R = V_h - V_l$, where V_l and V_h are the minimum and maximum scene irradiance. Thus, R and μ are both the functions of exposure time. From Eq. (17), we can observe that SQNR without saturation is dependent of the number of quantization levels Q; thus, better exposed signals produce higher SQNR.

When the signals are over-exposed, the SQNR expression becomes more complicated due to saturation, as

$$\text{SQNR} = \frac{\text{signal}}{\text{noise}} \propto \frac{C_o\mu - \alpha}{\frac{(\lambda - C_o V_l)}{Q} + \alpha}, \tag{18}$$

where λ, α, and C_o are saturation threshold, expectation of error within saturation sub-interval, and scaling factor of the over-exposure case, respectively. Here, $C_o V_h$ is replaced by λ due to saturation.

Let us assume that not all signals are saturated. Then, the condition that the well-exposed case has a greater SQNR than the over-exposed case is following:

$$\frac{Q\mu}{(V_h - V_l)} \geq \frac{C_o\mu - \alpha}{\frac{(\lambda - C_o V_l)}{Q} + \alpha} = \frac{C_o Q\mu - Q\alpha}{\lambda - C_o V_l + Q\alpha}. \tag{19}$$

Above can be simplified by some algebraic operations into:

$$\frac{Q\mu}{(V_h - V_l)} \geq \frac{Q\alpha}{(C_o V_h - \lambda) - Q\alpha}. \tag{20}$$

The condition to satisfy Eq. (19) with respect to Q is

$$Q \leq \frac{(C_o V_h - \lambda)}{\alpha} - \frac{(V_h - V_l)}{\mu}, Q > \frac{(C_o V_h - \lambda)}{\alpha}, \tag{21}$$

where $(C_o V_h - \lambda)$ is the maximum error. Mathematically, over-exposed case can produce a higher SQNR than the well-exposed case. However, in general situations, SQNR of well-exposed case is better than over-exposed case because the number of quantization levels Q is usually larger enough than maximum error $(C_o V_h - \lambda)$ over expectation error α. Therefore, well-exposed signals have

higher SQNR than over- or under-exposure cases in terms of quantization if the number of quantization levels is sufficient. Our method is beneficial with auto-exposure to increase SQNR since it can effectively handle non-uniformity caused by auto-exposure.

If there are quantization noise in the images, the observation matrix $\mathbf{M}$ becomes

$$\mathbf{M} = \mathbf{M}^* + \zeta = \mathbf{ELB}^\top + \zeta, \tag{22}$$

where $\mathbf{M}^*$ and ζ are the ideal observation and quantization noise, respectively. Using the noisy input in Eq. (22) , the objective function in Eq. (5) becomes

$$\{\mathbf{E}^*, \mathbf{B}^*\} = \underset{\mathbf{E}, \mathbf{B}}{\operatorname{argmin}} \|\zeta\|_{\mathrm{F}}^2, \quad \text{s.t.} \quad \zeta = \mathbf{M} - \mathbf{ELB}^\top. \tag{23}$$

Therefore, in the cases of high SQNR data, we can compute surface normal and intensities by optimizing Eq. (23) without biases since ζ is close to zero ($\mathbf{M} \approx \mathbf{M}^*$). However, in low SQNR inputs, minimizing Eq. (23) can produce biased results because ζ is not small anymore ($\mathbf{M} \neq \mathbf{M}^*$). As a result, auto-exposure can help to estimate surface normal by increasing SQNR of images, and our method is suitable for dealing with the exposure variations.

5   Light Intensity Calibration Analysis

One may consider that light intensity calibration is an easy task, but it actually requires both careful control over the environment and explicit knowledge about the reflectance of a calibration target. To show this, we perform light intensity calibration using a diffuse sphere[1]. Assuming a Lambertian reflectance model and known surface normal $\mathbf{N}$, the scaled light matrix $\mathbf{S}$ can be estimated from a set of measurements $\mathbf{M}$ as

$$\mathbf{S}^* = \underset{\mathbf{S}}{\operatorname{argmin}} \left\| \pi_{\Omega^c}(\mathbf{M}) - \pi_{\Omega^c}(\mathbf{SN}^\top) \right\|_{\mathrm{F}}^2, \tag{24}$$

where Ω denotes the locations of shadowed entries in the observation $\mathbf{M}$, and π_{Ω^c} represents an operator that extracts entries that are not shadowed (Ω^c). Since $\mathbf{S} = \mathbf{EL}$, with known light directions $\mathbf{L}$, we can determine the light intensities by

$$\mathbf{E} = \mathbf{S}^* \mathbf{L}^\dagger. \tag{25}$$

We recorded images of a diffuse sphere by changing the light directions of an identical light source with retaining its distance to the target object approximately the same. The camera response function is linear and uncompressed RAW images are used. Exposure times are kept constant with making sure that there is no under- or over-exposures. In addition, to neglect the perspective effect, a

---

[1] Due to the presence of saturations, a chrome sphere with specular highlights is not a proper calibration object for this task.

camera is placed far enough from the target object so that we can assume an orthographic projection model. Figure 2 shows some of the recorded images, and the light intensity matrix $\mathbf{E}$ is obtained by Eqs. (24) and (25).

As summarized in the numbers in Fig. 2, the estimated light intensities have variations while they are supposed to be uniform under this setting. The variations may be caused due to that (1) although the sphere is carefully selected, it still deviates from the Lambertian assumption, and (2) the assumed surface normal directions may be different from the truth due to errors of circle fitting. As such, even with a careful procedure, the light source intensity calibration is not a straightforward task. And in our setting, it had a non-negligible spread of estimated intensities (maximum 0.052 when the intensities are normalized to one, corresponding to 5 % error). Therefore, it is needed to directly model the variations of light intensities in the photometric stereo formulation.

Direction-1	Direction-2	Direction-3	Direction-4
(a) 1.0	(b) 0.948	(c) 0.939	(d) 0.988
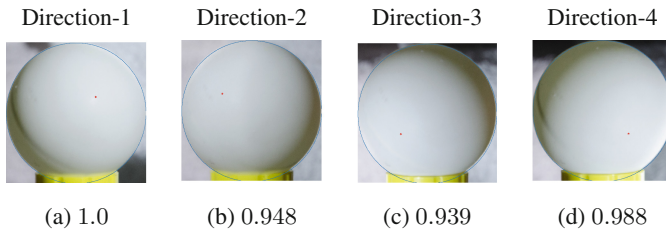
Fig. 2. Light intensity calibration. A diffuse sphere is illuminated under different light directions by moving an identical light source. The red point indicates the lighting direction, and a blue circle is a circle fitting to the image of a sphere. The numbers under photographs are the estimated light source intensities, that are relative to that of Direction 1. (Color figure online)
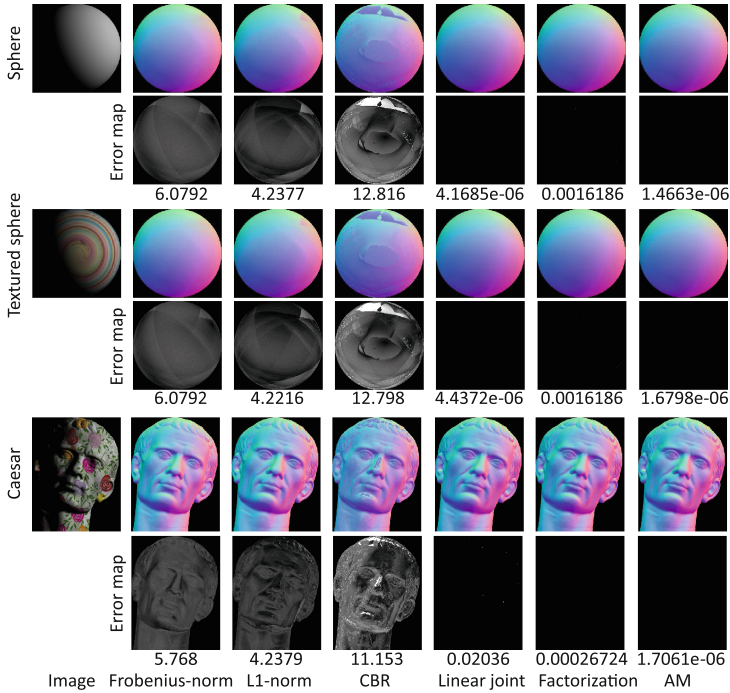
6   Experiments

We evaluate the proposed methods, linear joint estimation, factorization based, and alternating minimization (AM) methods, using synthetic (Sect. 6.1) and real-world (Sect. 6.2) scenes in the setting of non-uniform intensities and exposures. Although none of the previous techniques are designed for the non-uniform intensity setting, as previous methods to compare, we use standard Frobenius-norm minimization [1], robust L1-norm minimization used as a baseline method in [24], and the state-of-the-art photometric stereo method based on constrained bivariate regression (CBR) [24].

6.1   Synthetic Data

We first test our methods using synthetic examples that are textured and rendered with a Lambertian reflectance model with shadows. For qualitative and quantitative comparisons, we analyze the effects of non-uniform light intensities and auto-exposure.

Non-uniform Light Intensities: We first test the setting of non-uniform light intensities. The scenes are rendered under 20 varying light directions with their intensity variance 0.05. The qualitative visualization of surface normal estimates and error maps are summarized in Fig. 3 with comparison to other previous methods, *i.e.*, Frobenius-norm, L1-norm, CBR. Our methods, namely, linear joint, Factorization, and AM methods correspond to the ones described in Sects. 3.1, 3.2 and 3.3, respectively. Our proposed methods produce results close to the ground truth compared to other techniques that do not explicitly consider the non-uniform light intensities. The quantitative results are reported under each error map. The superior performance is consistently observed under varying numbers of images and light intensity variances as shown in Fig. 4.

Fig. 3. Photometric stereo experiment under non-uniform light intensities. The scenes are rendered under 20 distinct light directions with their intensity variance 0.05. Our methods (linear joint, factorization and AM) effectively handle the condition of non-uniform light intensities. Error maps are scaled by 4. The numbers indicate the mean angular errors in degree.

Auto-Exposure Case: Auto-exposure allows us to obtain measurements with a higher Signal-to-Quantization-Noise ratio (SQNR). To assess the benefit of auto-exposure in photometric stereo and effectiveness of our methods in this setting, we render two datasets; one with auto-exposure and the other with fixed-exposure. In the auto-exposure dataset, the sensor irradiances are stretched to
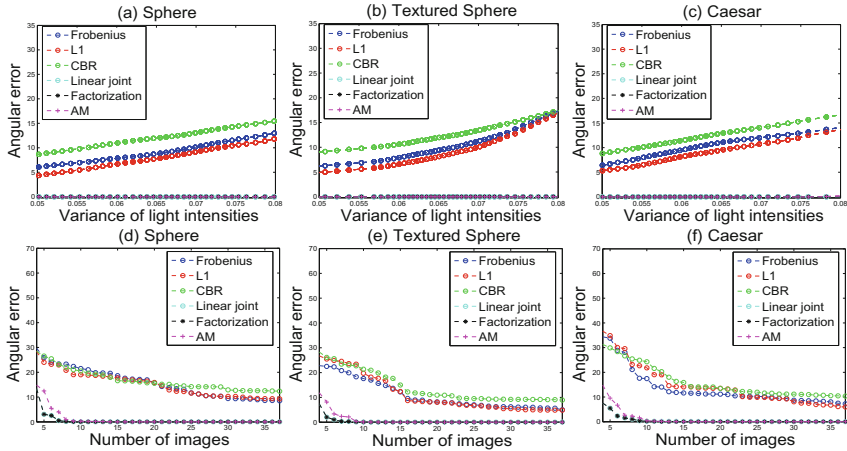
Fig. 4. Variations of mean angular errors of surface normal estimates over variance of light intensities (top row) and the number of images (bottom row) for the three datasets. (a, d) Sphere, (b, e) Textured Sphere, (c, f) Caesar. Our methods consistently yield favorable results across these variations.

properly include the most of dynamic range before quantization. For the fixed-exposure dataset, sensor irradiances are quantized without stretching. From the two types of dataset, we apply the same set of photometric stereo methods for performance evaluation. The results are summarized in Table 1. While the fixed-exposure setting suffers from a low SQNR (which leads to lower accuracy of surface normal estimates), the auto-exposure retains a higher SQNR. And with our methods, this setting is properly handled and accurate surface normal estimates are obtained.

Table 1. Comparison under auto-exposure (Auto) and fixed-exposure (Fixed) settings. SQNR and the mean angular errors of surface normal estimates in degree are shown.

SQNR	Sphere		Textured		Caesar	
	Fixed	Auto	Fixed	Auto	Fixed	Auto
	42.52	129.8	33.68	135.6	33.56	133.3
Frobenius	2.383	24.35	27.93	23.61	8.415	23.86
L1	2.507	23.85	28.05	23.41	8.320	21.03
CBR [24]	21.47	14.54	47.15	16.53	33.02	21.49
Linear joint	20.49	0.886	59.63	1.561	43.72	1.725
Factorization	6.569	0.715	42.12	1.381	26.16	1.504
AM	3.970	**0.256**	29.74	**0.565**	9.684	**0.564**

6.2   Real Data

We design three different settings for the real-world experiment; (A) non-uniform light source intensities across images, (B) with auto-exposures under identical light intensities (by moving the same light source), and (C) use of an uncontrolled mobile phone camera for imaging where auto-exposure is turned on under varying light source intensities. For all real-world examples, we use a shiny sphere to calibrate the light directions. To suppress other un-modelled factors, our experiments are carried out in a dark room.

Non-uniform Light Source Intensities: To record images under different light intensities and directions, we use controllable light sources whose brightnesses can be manually controlled by the gain of power supply. The camera setting, such as shutter speed and aperture, are all fixed in this experiment, and a linear sensor response is used. In this experimental setting, we recorded 20 images for each static scene. The results are summarized in Fig. 5, in which the estimated surface normal and their 3D reconstruction using [25] are presented. As shown in the figures, our methods properly handle the varying light source intensities compared to Frobenius-norm, L1-norm and CBR methods, with which severe distortions are observed in their reconstructed surfaces.
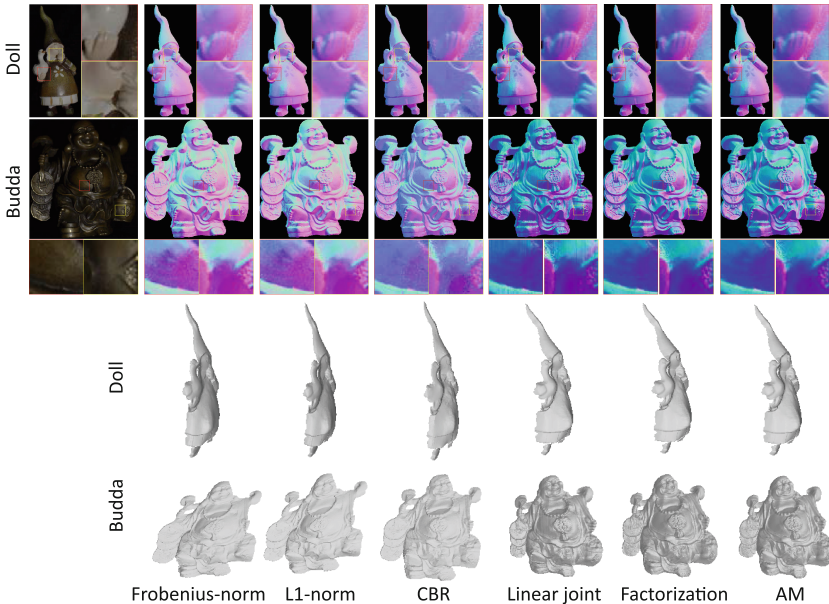
Fig. 5. Result of varying light source intensities case. From left to right, one of input images, results from Frobenius-norm, L1-norm, CBR [24], linear joint estimation, factorization and alternating minimization (AM) methods are shown.

Auto-exposure: When auto-exposure is used, the shutter speed and/or aperture size of a camera is automatically adjusted to record well-exposed images according to the amount of incoming light. While it increases SQNR, it results in the non-uniform intensity setting.

For this experiment, we recorded 20 images of each static scene with auto-exposure. Figure 6 shows the comparative result. As shown in the figure, our methods consistently yield higher quality outputs than the other methods because our method explicitly accounts for the non-uniform exposures.
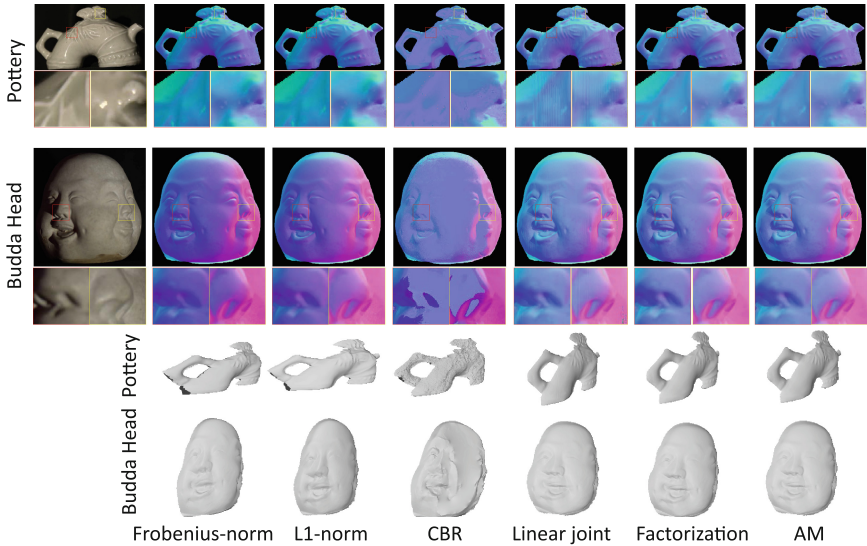
Fig. 6. Result of auto-exposure case. From left to right, one of input images, results from Frobenius-norm, L1-norm, CBR [24], linear joint estimation, factorization and alternating minimization (AM) methods are shown.

Mobile Phone Cameras: Our method is suitable for uncontrollable cameras like many of mobile phone cameras, where we cannot turn off the auto-exposure setting. With such cameras, recorded images are in the condition of non-uniform exposures across images. From recorded images from a mobile phone camera, we linearize the intensity observations using the method of [26] as preprocessing. Figure 7 shows the surface normal estimates and their 3D reconstructions. While the 3D reconstructions of conventional methods are severely deformed, our methods show better reconstructions in general. The linear joint estimation method suffered from the outliers in this case, but that is not observed in factorization based and AM methods.
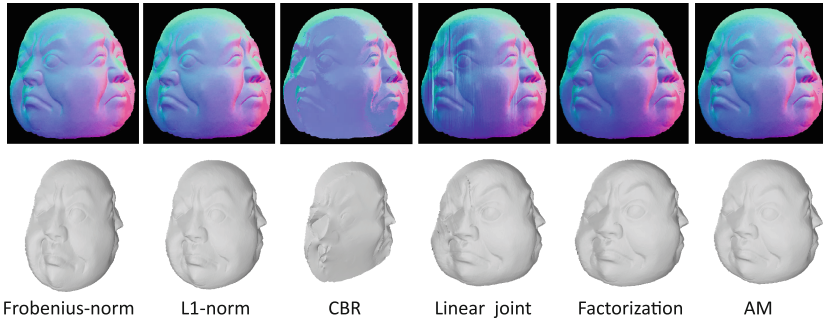
Fig. 7. Result using a mobile phone camera. Top: estimated surface normal, bottom: 3D reconstruction. Our methods (linear joint estimation, factorization, and alternating minimization (AM) methods) produce more faithful results than the conventional methods.

7   Conclusion

This paper described photometric stereo methods that can handle the non-uniform light source intensities and exposures across images. We showed the effect of *varying light intensity conditions* in photometric stereo that is relevant in practical settings. We then developed solution methods that explicitly account for the non-uniform light intensities and exposures; namely, linear joint estimation, factorization based, and alternating minimization methods. The linear joint estimation and factorization based methods are simple and easy to implement, they occasionally suffer from numerical instability due to un-modelled observations. The alternating minimization method showed a greater robustness over these techniques, while retaining the efficiency in computation. They are all effective in the non-uniform intensities setting compared to methods that neglect the effect of the setting. We further illustrate that our proposed methods can benefit from auto-exposure, with which measurements with a greater SQNR can be obtained. Our experiments on synthetic and real-world examples show the importance of properly handling varying light intensities and exposures.

References

1. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Opt. Eng. **19**(1), 139–144 (1980)
2. Wu, T., Tang, K., Tang, C., Wong, T.: Dense photometric stereo: a Markov random field approach. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **28**(11), 1830–1846 (2006)

3. Verbiest, F., Van Gool, L.: Photometric stereo with coherent outlier handling and confidence estimation. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 2886–2893 (2008)

4. Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y., Ma, Y.: Robust photometric stereo via low-rank matrix completion and recovery. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010. LNCS, vol. 6494, pp. 703–717. Springer, Heidelberg (2011). doi:10.1007/978-3-642-19318-7_55

5. Oh, T.H., Kim, H., Tai, Y.W., Bazin, J.C., Kweon, I.S.: Partial sum minimization of singular values in RPCA for low-level vision. In: Proceedings of International Conference on Computer Vision (ICCV), pp. 145–152 (2013)

6. Zhang, Y., Mu, C., Kuo, H.W., Wright, J.: Toward guaranteed illumination models for non-covex objects. In: Proceedings of International Conference on Computer Vision (ICCV), pp. 937–944 (2013)

7. Hayakawa, H.: Photometric stereo under a light-source with arbitrary motion. J. Opt. Soc. Am. (JOSA) **11**(11), 3078–3089 (1994)

8. Yuille, A.L., Snow, D.: Shape and albedo from multiple images using integrability. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 158–164 (1997)

9. Belhumeur, P.N., Kriegman, D.J., Yuille, A.L.: The bas-relief ambiguity. Int. J. Comput. Vis. (IJCV) **35**(1), 33–44 (1999)

10. Alldrin, N.G., Mallick, S.P., Kriegman, D.J.: Resolving the generalized bas-relief ambiguity by entropy minimization. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 1–7 (2007)

11. Drbohlav, O., Chantler, M.: Can two specular pixels calibrate photometric stereo? In: Proceedings of International Conference on Computer Vision (ICCV), pp. 1850–1857 (2005)

12. Sunkavalli, K., Zickler, T., Pfister, H.: Visibility subspaces: uncalibrated photometric stereo with shadows. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6312, pp. 251–264. Springer, Heidelberg (2010). doi:10.1007/978-3-642-15552-9_19

13. Shi, B., Matsushita, Y., Wei, Y., Xu, C., Tan, P.: Self-calibrating photometric stereo. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 1118–1125 (2010)

14. Solomon, F., Ikeuchi, K.: Extracting the shape and roughness of specular lobe objects using four light photometric stereo. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **18**(4), 449–454 (1996)

15. Georghiades, A.: Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In: Proceedings of International Conference on Computer Vision (ICCV), pp. 816–823 (2003)

16. Lin, S., Lee, S.W.: Estimation of diffuse and specular appearance. In: Proceedings of International Conference on Computer Vision (ICCV), pp. 855–860 (1999)

17. Iwahori, Y., Woodham, R.J., H.T., Ishii, N.: Neural network to reconstruct specular surface shape from its three shading images. In: Proceedings of International Joint Conference on Neural Networks, pp. 1181–1184 (1993)

18. Malzbender, T., Wilburn, B., Gelb, D., Ambrisco, B.: Surface enhancement using real-time photometric stereo and reflectance transformation. In: Proceedings of the Eurographics Symposium on Rendering Techniques, pp. 245–250 (2006)

19. Shi, B., Tan, P., Matsushita, Y., Ikeuchi, K.: Bi-polynomial modeling of low-frequency reflectances. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **36**(6), 1078–1091 (2014)

20. Ikehata, S., Wipf, D.P., Matsushita, Y., Aizawa, K.: Photometric stereo using sparse Bayesian regression for general diffuse surfaces. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **36**(9), 1816–1831 (2014)
21. Sylvester, J.: Sur lequations en matrices px = xq. C. R. Acad. Sci. Paris **99**(2), 67–71,115–116 (1884)
22. Holland, P.W., Welsch, R.E.: Robust regression using iteratively reweighted least-squares. Commun. Stat. Theor. Methods **6**(9), 813–827 (1977)
23. Kamgar-Parsi, B., Kamgar-Parsi, B.: Evaluation of quantization error in computer vision. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **11**(9), 929–940 (1989)
24. Ikehata, S., Aizawa, K.: Photometric stereo using constrained bivariate regression for general isotropic surfaces. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 2187–2194 (2014)
25. Xie, W., Zhang, Y., Wang, C.C.L., Chung, R.C.K.: Surface-from-gradients: an approach based on discrete geometry processing. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 2203–2210 (2014)
26. Lee, J.Y., Matsushita, Y., Shi, B., Kweon, I.S., Ikeuchi, K.: Radiometric calibration by rank minimization. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **35**(1), 144–156 (2013)