

Robust photometric stereo using structural light sources

Tian-Qi Han, Yue Cheng, Hui-Liang Shen,* and Xin Du

Zhejiang University, Department of Information Science and Electronic Engineering, Hangzhou 310027, China

Abstract. We propose a robust photometric stereo method by using structural arrangement of light sources. In the arrangement, light sources are positioned on a planar grid and form a set of collinear combinations. The shadow pixels are detected by adaptive thresholding. The specular highlight and diffuse pixels are distinguished according to their intensity deviations of the collinear combinations, thanks to the special arrangement of light sources. The highlight detection problem is cast as a pattern classification problem and is solved using support vector machine classifiers. Considering the possible misclassification of highlight pixels, the ℓ_1 regularization is further employed in normal map estimation. Experimental results on both synthetic and real-world scenes verify that the proposed method can robustly recover the surface normal maps in the case of heavy specular reflection and outperforms the state-of-the-art techniques. © 2014 SPIE and IS&T [DOI: 10.1117/1.JEI.23.3.033004]

Keywords: photometric stereo; structural light sources; shadow detection; highlight detection; normal estimation; surface reconstruction.

Paper 13608 received Oct. 29, 2013; revised manuscript received Apr. 8, 2014; accepted for publication Apr. 11, 2014; published online May 13, 2014.

1 Introduction

Photometric stereo¹ estimates the normals of surface points from multiple images which are captured under various lighting conditions using a fixed camera. Since the 1980s, photometric stereo has attracted intensive interest in surface reconstruction as it can produce dense normal maps at the level of detail that can hardly be achieved by other techniques.

If a Lambertian surface is illuminated under three or more light directions, its normal could be well estimated using least squares (LS). In practice, however, the images may be corrupted by non-Lambertian effects, such as shadow and specular highlight. Shadow has the nature of low intensity and thus could be detected by simple thresholding.² The detection of highlights, however, is a much complicated issue because the pixel intensity is determined by both surface orientation and albedo. Consequently, image intensity is not a robust feature for highlight detection and the simple thresholding method is not applicable.

The photometric stereo techniques dealing with this non-Lambertian problem can be approximately classified into two categories, i.e., the modeling ones and the statistical ones. The modeling methods estimate the bidirectional reflectance distribution function (BRDF)³ of the surfaces⁴⁻⁶ or use specified BRDF assumptions.⁷⁻⁹ These techniques perform well when the BRDF models can be precisely established. However, it is always quite difficult, if not impossible, to estimate the BRDF parameters accurately without using a large-scale image set. The statistical methods treat the non-Lambertian surface points as outliers. The most recent techniques in this category include random sample consensus (RANSAC) scheme,^{10,11} median filter,¹² low-rank matrix completion,¹³ and sparse Bayesian learning (SBL).¹⁴

In this work, we follow the statistical way for non-Lambertian photometric stereo and mainly focus on

highlight detection. We position the light sources in a planar grid such that they form a set of three-source collinear combinations, as illustrated in Fig. 1. Note that in this work, the term collinear refers to the positions of light sources, not light directions. By this arrangement, the image intensities corresponding to each collinear combination are linearly correlated with fixed coefficients under the Lambertian assumption. When image intensities deviate from this assumption, the linear relationship will no longer hold. In this way, the intensity deviations implicitly reflect the distribution of highlight pixels. Highlight detection is then cast to a pattern classification problem.¹⁵ In this work, the classification is accomplished by using the support vector machine (SVM) classifiers,^{16,17} which perform well even when the highlights seriously overlap in photometric images. To further deal with the misclassified highlight pixels, the ℓ_1 regularization is employed in surface normal estimation.

1.1 Related Work

The idea of photometric stereo was first introduced by Woodham for surface reconstruction,¹ under the assumption that the surface is of strict Lambertian reflections. Coleman and Jain¹⁸ extended Woodham's method by introducing a fourth light source. Mallick et al.¹⁹ presented a photometric stereo method to reconstruct specular surfaces by removing highlights in a data-dependent color space. Liu and Han²⁰ estimated surface normals of specular objects by using a special configuration in which the light source is rotated around the optical axis of the camera. Barsky and Petrou²¹ employed the linear dependence of four illumination vectors to non-Lambertian surface positions and detected highlight pixels using color. Argyriou and Petro²² extended this technique²¹ to more light sources in a recursive photometric stereo. They ruled out shadow and highlight pixels first and then corrected the erroneous normals using interpolation or robust fitting

*Address all correspondence to: Hui-Liang Shen, E-mail: shenhl@zju.edu.cn

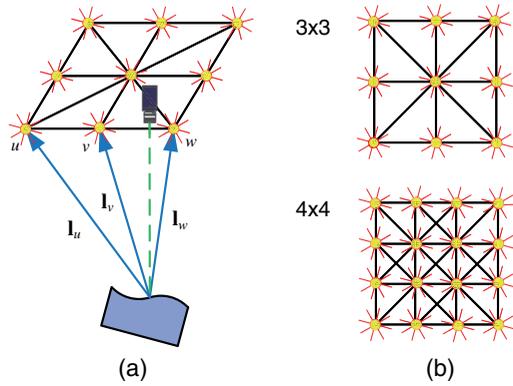


Fig. 1 (a) Photometric stereo system. (b) Two structural arrangements of light sources. The light source positions are equispaced in planar grids such that they form 3×3 and 4×4 structures, respectively. In the 3×3 structure, there are eight (three horizontal, three vertical, and two diagonal) three-source collinear combinations. In the 4×4 structure, there are 44 three-source collinear combinations, as each four-source line contributes four combinations.

before surface reconstruction. These methods can keep the number of photometric images quite low, but at the cost of accuracy degradation.

The above-mentioned modeling and statistical methods detect non-Lambertian reflections using a large number of photometric images. In the modeling methods, the nonlinear reflectance models are introduced to include non-Lambertian reflections. Georgiades⁴ used the Torrance-Sparrow model to estimate both surface normals and BRDF parameters in an uncalibrated photometric stereo by solving many unknown variables. Goldman et al.⁶ assumed that the objects are composed of a limited number of fundamental materials. By representing each pixel with a combination of, at most, two fundamental materials, both shape and material BRDFs are estimated simultaneously. Alldrin and Kriegman⁸ used the general BRDF isotropy to estimate the isocontours of the depth map under a particular light configuration. Following their work, Shi et al.⁹ estimated the elevation angles of surface normal based on the monotonicity of reflections. Besides specular reflections, shadows are also studied as extra information for different applications. Chandraker et al.²³ proposed a fast graph cuts based method to detect shadow and the recovered shadow maps were employed in constrained surface normal integration. Okabe et al.² encoded surface points via attached shadow and embedded the high-dimensional codes into a three-dimensional (3-D) surface normal space to solve the uncalibrated photometric stereo. Chung and Jia⁵ employed the cue of cast shadow to compute surface normal and estimated the global reflectance parameters using iterative optimization.

The statistical methods treat non-Lambertian surface pixels as outliers. Wu et al.²⁴ formulated a Markov network for dense photometric stereo, which was optimized by graph cuts and tensor belief propagation. Miyazaki et al.¹² presented a median-based photometric stereo that iteratively refines the normals by increasing the candidate normal subsets. These methods incorporate the spatial neighborhood information to eliminate noise and non-Lambertian effect, and thus are not quite appropriate for object surfaces with fine details. In this regard, Mukaigawa et al.^{10,11} proposed a method to classify diffuse reflection, specular reflection,

attached shadow, and cast shadow in a purely pixel-wise manner. This method is based on RANSAC and thus is computationally expensive. Wu et al.¹³ proposed a robust principal component analysis method by treating the non-Lambertian effects as missing or corrupted entries, and solved the problem using a rank minimization scheme. Motivated by this, Ikehata et al.¹⁴ explicitly used the rank-3 Lambertian constraint and cast the problem as sparse regression. A hierarchical Bayesian model is thus derived from an SBL framework to solve the problem.

1.2 Motivation of Our Work

To date, most robust photometric methods, though effective, use a large number of images in real scene experiments. However, in practical applications, it is always desired to use a relatively low number of images for the sake of speeding up image acquisition and computational efficiency. In addition, either explicitly or implicitly, these methods assume the highlight regions are sparsely distributed. This assumption makes the arrangement of the light sources difficult in a limited space capacity.

In the proposed photometric stereo method, the light sources are arranged to form a set of three-source collinear combinations, whose structure is fundamental to highlight detection. Our method uses a relatively low number (for example, 9) of light sources and simultaneously keeps the imaging system compact. Compared to the existing techniques that use illumination dependence,^{21,22} our method exploits full correlations among the collinear combinations and solves the highlight detection problem using pattern classification. Thanks to the mentioned attributes, the proposed method can produce satisfactory surface reconstruction and is suitable for practical applications.

2 Problem Statement

Suppose n images of an object surface are captured under variant light directions and each image contains m pixels. Under the Lambertian assumption, the image intensity i_t under a light direction \mathbf{l} is formulated as

$$i_t = \rho \mathbf{n}^T(\mathbf{l}), \tag{1}$$

where ρ denotes the diffuse albedo, $\mathbf{n} = (n_x, n_y, n_z)^T$ denotes the surface normal, and t represents the light intensity. Note that in this work, the light direction \mathbf{l} and light intensity t are available via light source calibration.²⁵ Then the calibrated image intensity i is computed as

$$i = \frac{i_t}{t} = \rho \mathbf{n}^T \mathbf{l}, \tag{2}$$

which is independent of light intensity. By reshaping each image as a vector and stacking them to form a matrix $\mathbf{I} \in \mathbb{R}^{m \times n}$, Eq. (2) can be written in a matrix notation as

$$\mathbf{I} = \mathbf{N}^T \mathbf{L}, \tag{3}$$

where $\mathbf{N} = (\rho_1 \mathbf{n}_1, \dots, \rho_m \mathbf{n}_m) \in \mathbb{R}^{3 \times m}$ denotes the albedo-scaled normal matrix and $\mathbf{L} = (\mathbf{l}_1, \dots, \mathbf{l}_n) \in \mathbb{R}^{3 \times n}$ denotes the light direction matrix.

In the case of non-Lambertian surfaces, Eq. (2) is formulated as

$$\begin{aligned}
 o &= i + e \\
 &= \rho \mathbf{n}^T \mathbf{l} + e,
 \end{aligned} \tag{4}$$

where o is the observed image intensity and e is the non-Lambertian corruption term mainly due to shadow and highlight. The corresponding matrix is now expressed as

$$\begin{aligned}
 \mathbf{O} &= \mathbf{I} + \mathbf{E} \\
 &= \mathbf{N}^T \mathbf{L} + \mathbf{E},
 \end{aligned} \tag{5}$$

where \mathbf{O} is the observation matrix and \mathbf{E} is the corruption matrix. The objective of photometric stereo is to estimate the albedo-scaled surface normal \mathbf{N} from the corrupted observation \mathbf{O} . With the estimated normal map, the 3-D surface can be reconstructed by normal integration.

3 Proposed Method

Figure 2 shows the framework of the proposed robust photometric stereo method. It consists of four main steps, i.e., shadow detection, deviation calculation, highlight detection, and normal estimation. The shadow pixels always have low intensities, and in this work, we propose an adaptive thresholding method for shadow detection. Specular highlight is difficult to detect as the image intensity is determined by both surface normal and albedo, as well as the BRDF. To deal with this problem, we propose to arrange the light sources with a set of collinear combinations as illustrated in Fig. 1. The highlight detection is accomplished by pattern classification using SVM classifiers.

We note that there are numerous arrangements that can produce collinear combinations. In this work, we elaborate the fundamentals of the proposed method based on the 3×3 structure and illustrate the experimental results on both the 3×3 and 4×4 structures. The application of the proposed method to other structures is straightforward.

3.1 Shadow Detection

For real images, a shadow pixel has nonzero intensity due to the presence of slight ambient illumination or inter-reflections. On the other hand, a nonshadow pixel may also have

a relatively low intensity caused by low albedo. Therefore, a global threshold is not appropriate to detect shadow for all pixels. In this work, we identify the shadow pixels of each surface position by using an adaptive threshold T .

$$T = \eta \cdot \text{median}(o_1, o_2, \dots, o_n), \tag{6}$$

where η is a fixed parameter and $\text{median}(\cdot)$ denotes the median filter applied on the observed image intensities o_i under n light sources. A pixel is labeled as shadow if its intensity is less than T and as nonshadow otherwise. Here, we assume that for the majority of surfaces, the median intensity is of Lambertian reflection. In this way, the albedo is incorporated in the adaptive threshold. The optimal value of the parameter η will be determined in experiment.

We note that, as the light sources are arranged relatively compactly, highlight and shadow are unlikely to occur simultaneously on the same surface point. Hence, for surface points with shadow pixels, we estimate surface normals directly without highlight detection. For nonshadowed surface points, we detect highlight pixels using the collinear light source structure before normal estimation.

3.2 Analysis of Intensity Deviation

In the following, we discuss the characteristics of the intensity deviation of a three-source collinear combination that is fundamental to highlight detection. As illustrated in Fig. 1(a), the bottom three light sources, u , v , and w , make a collinear combination. It is evident that their corresponding light directions, \mathbf{l}_u , \mathbf{l}_v , and \mathbf{l}_w , are coplanar. Mathematically, these light directions satisfy

$$\alpha \mathbf{l}_u + \beta \mathbf{l}_v + \gamma \mathbf{l}_w = 0, \tag{7}$$

where α , β , and γ are referred to as collinear coefficients, which can be solved up to a scale. Without loss of generality, we set $\hat{\alpha} = 1$, and then $\hat{\beta}$ and $\hat{\gamma}$ can be uniquely computed. In this work, we further normalize the coefficient such that $(\alpha, \beta, \gamma) = [1/(\sqrt{1 + \hat{\beta}^2 + \hat{\gamma}^2})](1, \hat{\beta}, \hat{\gamma})$.

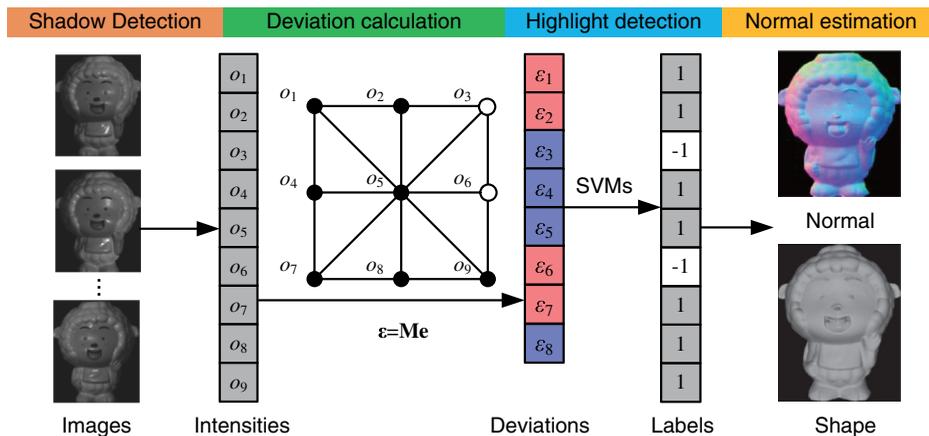


Fig. 2 Framework of the proposed method. The 3×3 light source structure is illustrated as an example. The pixels o_3 and o_6 are assumed to be corrupted by highlight. The intensity deviations corresponding to the collinear combinations are fed to support vector machine (SVM) classifiers for highlight detection. The three-dimensional (3-D) object surface can be finally reconstructed from the estimated normal map by integration.

According to Eq. (2), the image intensity of a diffuse pixel is completely determined by surface normal and light direction scaled by albedo ρ . Hence, the image intensities i_u , i_v , and i_w under the light sources u , v , and w satisfy

$$\begin{aligned} \epsilon &= \alpha i_u + \beta i_v + \gamma i_w \\ &= \alpha \rho \mathbf{n}^T \mathbf{l}_u + \beta \rho \mathbf{n}^T \mathbf{l}_v + \gamma \rho \mathbf{n}^T \mathbf{l}_w \\ &= \rho \mathbf{n}^T (\alpha \mathbf{l}_u + \beta \mathbf{l}_v + \gamma \mathbf{l}_w) \\ &= 0. \end{aligned} \tag{8}$$

In the following, ϵ is referred to as intensity deviation. It is zero when all the three pixels in a collinear combination have pure diffuse reflection.

For highlight pixels, the observed image intensities o_u , o_v , and o_w are corrupted by specular reflection, and consequently, we have

$$\begin{aligned} \epsilon &= \alpha o_u + \beta o_v + \gamma o_w \\ &= \alpha(i_u + e_u) + \beta(i_v + e_v) + \gamma(i_w + e_w) \\ &= (\alpha i_u + \beta i_v + \gamma i_w) + (\alpha e_u + \beta e_v + \gamma e_w) \\ &= \alpha e_u + \beta e_v + \gamma e_w, \end{aligned} \tag{9}$$

which implies that the intensity deviation ϵ is nonzero.

From Eqs. (8) and (9), it is possible to identify highlight pixels using the intensity deviations of collinear combinations. Figure 3 illustrates an example in which o_3 and o_6 are highlight pixels. According to the above discussion, in ideal circumstances, the deviations ϵ_1 , ϵ_2 , ϵ_6 , and ϵ_7 have nonzero values and the others have zero values. However, because of image noise and BRDF variation, the simple inference of highlight pixels from the deviations is always not reliable. In this work, we cast this problem as a pattern classification problem and use all eight deviation errors to locate highlight pixels.

3.3 Highlight Detection

We stack all the collinear coefficients into a matrix \mathbf{M} by padding zeros; then under the Lambertian assumption, the deviation vector ϵ is

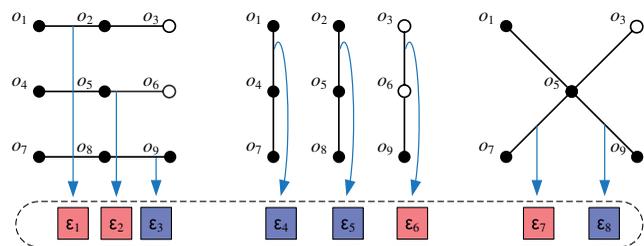


Fig. 3 The three-source collinear combinations and corresponding intensity deviations in the 3×3 structure. Pixels o_3 and o_6 are assumed to be corrupted by highlight. The zero deviations are in blue and the nonzero ones are in red. Nonzero deviations indicate the existence of highlight pixels in the collinear combination.

$$\epsilon = \mathbf{M}\mathbf{i}$$

$$= \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_2 & \beta_2 & \gamma_2 & 0 & 0 & 0 \\ \dots & \dots \\ \alpha_8 & 0 & 0 & 0 & \beta_8 & 0 & 0 & 0 & \gamma_8 \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ \dots \\ i_9 \end{pmatrix} = \mathbf{0}, \tag{10}$$

where \mathbf{i} is the stacked image intensity vector under the nine light sources in the 3×3 structure. The locations of the coefficients α_k , β_k , and γ_k , $k \in \{1, 2, \dots, 8\}$, in matrix \mathbf{M} are determined by the collinear combinations illustrated in Fig. 3.

For pixels corrupted by highlight, the observed intensities are larger than the ideal diffuse components; hence, the deviation vector becomes

$$\begin{aligned} \epsilon &= \mathbf{M}\mathbf{o} \\ &= \mathbf{M}(\mathbf{i} + \mathbf{e}) \\ &= \mathbf{M}\mathbf{e}, \end{aligned} \tag{11}$$

where \mathbf{o} and \mathbf{i} denote the stacked intensity vectors with and without corruption, respectively, and \mathbf{e} represents the corruption vector due to highlight.

According to Eq. (11), the highlight detection problem is to infer the corruption vector \mathbf{e} from the deviation vector ϵ . However, as the matrix \mathbf{M} is rank-deficient and the BRDF varies for different materials, there is no one-to-one mapping from the deviation ϵ to corruption \mathbf{e} . Fortunately, in this work, we are only interested in highlight detection, not the exact value of \mathbf{e} . Hence, we convert highlight detection to a pattern classification problem.

We use ball scenes as training data because a ball contains all possible surface normals. The balls are rendered using the Cook-Torrance reflectance model,³ with both diffuse and specular reflections. Given BRDF parameters, the diffuse components i and observed intensities o of each surface point are known. We designate labels 1 to the diffuse pixels and labels -1 to the highlight pixels. The pattern classification problem is solved using n binary SVM classifiers,^{16,17}

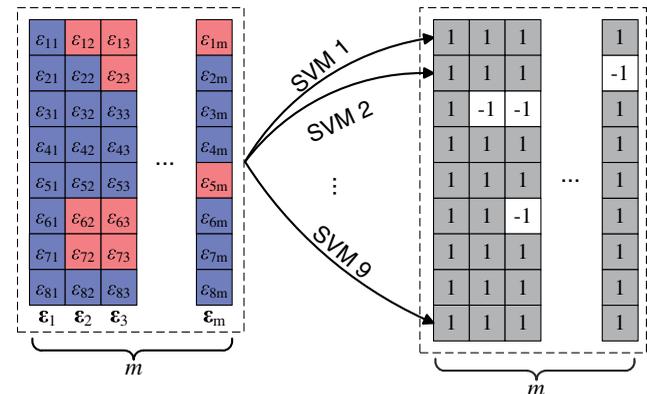


Fig. 4 Illustration of the SVM training procedure in the case of 3×3 structure. The nine SVM classifiers use the same deviation matrix as input data and output the highlight labels for m pixels under the nine different light sources, where m is the number of pixels in an image.

where n is the number of images (or light sources). Figure 4 illustrates the training procedure. Note that although >1 balls can be used for training, we only illustrate the case of a single training ball for clarity purposes. For each SVM classifier, the inputs are m deviation vectors $\boldsymbol{\varepsilon}$ and the outputs are m labels, where m is the number of pixels in an image. In the test procedure, we compute the deviation vector for each surface point and predict the labels of the m pixels using the trained binary SVM classifiers.

In this work, the SVM classifier is implemented using the LIBSVM library.¹⁷ The radius basic function is used as the kernel function, which is defined as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{1}{\tau} \|\mathbf{x}_i - \mathbf{x}_j\|^2\right), \quad (12)$$

where the input vector \mathbf{x}_i is actually the deviation vector $\boldsymbol{\varepsilon}_i$. The parameter τ controls the width of the Gaussian function and is set as the dimension of \mathbf{x}_i . The tolerance of termination criterion is set as $\varepsilon = 0.001$, and the penalty parameter of the error term is set as $C = 10^4$.

We note that the collinearity among light sources may not ideally hold in a practical experimental setup. Nevertheless, as the training balls are rendered using the actual light directions, the possible systematic biases are incorporated in the SVM classifiers and will not affect the classification accuracy.

3.4 Normal Estimation and Surface Reconstruction

After both shadow and highlight detection, the detected shadows and highlights are discarded in normal estimation. However, a few shadow and highlight pixels may still be incorrectly classified because of image noise or the BRDF difference between the actual and training data. In addition, in real circumstances, some image intensities may deviate from the hard constraint in Eq. (5) due to system noises and nonideal diffuse reflection. Therefore, similar to Ref. 14, we use the ℓ_1 regularization in normal estimation to eliminate possible outliers.

With a defined mask matrix $\mathbf{W} \in \mathbb{R}^{m \times n}$ whose elements are 1's for classified diffuse pixels and 0's for highlight and shadow pixels, the surface normal is estimated by

$$\hat{\mathbf{N}} = \arg \min_{\mathbf{N}} \{ \|\mathbf{W} \odot (\mathbf{N}^T \mathbf{L} + \mathbf{E} - \mathbf{O})\|_2^2 + \lambda \|\mathbf{W} \odot \mathbf{E}\|_1 \}, \quad (13)$$

where the operator \odot denotes the element-wise product between two matrices. The regularization parameter λ is set as $\lambda = 10^{-6}$ in this work. Note that the highlight term \mathbf{E} will shrink to 0's when $\lambda \rightarrow \infty$. This eventually becomes a weighted LS problem, $\hat{\mathbf{N}} = \arg \min_{\mathbf{N}} \{ \|\mathbf{W} \odot (\mathbf{N}^T \mathbf{L} - \mathbf{O})\|_2^2 \}$, in which ℓ_1 regularization is not applied.

For surface reconstruction, the gradient fields $(z_x, z_y)^T$ are computed from the estimated normal as $z_x = -n_x/n_z$, $z_y = -n_y/n_z$. The surface depth $z(x, y)$ can be represented in the finite difference form.

$$z_x(x, y) = \frac{z(x+1, y) - z(x-1, y)}{2}, \quad (14)$$

$$z_y(x, y) = \frac{z(x, y+1) - z(x, y-1)}{2}. \quad (15)$$

Equations (14) and (15) can be rewritten in matrix-vector notation with the stacked depth values \mathbf{z} and the gradient vectors \mathbf{z}_x and \mathbf{z}_y .

$$\begin{pmatrix} \mathbf{z}_x \\ \mathbf{z}_y \end{pmatrix} = \begin{pmatrix} \mathbf{D}_x \\ \mathbf{D}_y \end{pmatrix} \mathbf{z}, \quad (16)$$

where \mathbf{D}_x and \mathbf{D}_y denote the gradient operator matrix along the x axis and y axis, respectively. Each row of matrices \mathbf{D}_x and \mathbf{D}_y has two nonzero entries, namely, ± 0.5 , in the positions corresponding to the particular difference operation. Equation (16) can be treated as a gradient linear encoding operation, which translates depth to gradient fields.

Additionally, we incorporate the Laplacian-regularized term in Eq. (16) and apply the ℓ_1 decoding scheme as²⁶

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \{ \|\mathbf{z}_x - \mathbf{D}_x \mathbf{z}\|_1 + \|\mathbf{z}_y - \mathbf{D}_y \mathbf{z}\|_1 + \mu \|\text{Laplacian}(\mathbf{z})\|_1 \}, \quad (17)$$

where μ is the weight of the regularization term, which is set as $\mu = 0.05$ in this work. The operator $\text{Laplacian}(\cdot)$ denotes the finite difference Laplacian operator, which corresponds to the nonzero elements of the following convolution kernel:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (18)$$

Equation (17) can be formulated as a linear programming problem, which can be effectively solved using convex optimization.²⁷

4 Experiments

We evaluate the proposed method on both synthetic and real scenes, and in both circumstances, the imaging condition is kept identical. The proposed method is evaluated on both the 3×3 and 4×4 structures. The planar grids of the two structures are of same size, and, therefore, the light sources in the 4×4 structure are a little more densely positioned. The performance of the proposed method is compared with one baseline LS method¹ and three state-of-the-art techniques, i.e., RANSAC (Refs. 10 and 11), median photometric stereo,¹² and SBL.¹⁴

The synthetic data include two parts, i.e., the scenes generated using a BRDF model and those rendered using the MERL material database.^{28,29} The real scenes were captured using a practical imaging system. The descriptions of the synthetic and real scenes are as follows.

BRDF-model scenes: The images were rendered using the Cook-Torrance reflectance model,³ in which the observed image intensity was computed as

$$o = \rho \mathbf{n}^T \mathbf{l} + \rho_s \frac{DGF}{(\mathbf{n}^T \mathbf{v})}, \quad (19)$$

where $\mathbf{v} = (0, 0, 1)^T$ denotes the viewing direction, G is the geometrical attenuation factor, and F is the Fresnel coefficient. The term D represents the fraction of facets and is defined as

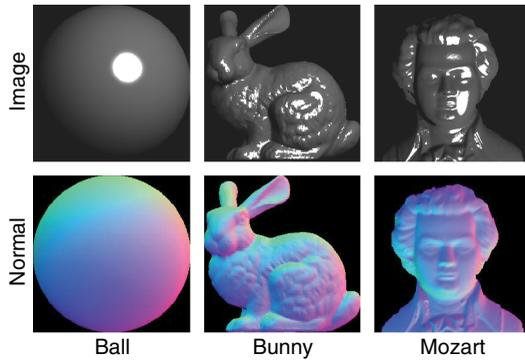


Fig. 5 Example images rendered by the Cook-Torrance reflectance model (top) and the ground truth normal maps (bottom) of the three surfaces *Ball*, *Bunny*, and *Mozart*.

$$D = \frac{1}{\sigma^2 \cos^4 \delta} \exp\left(-\frac{\tan^2 \delta}{\sigma^2}\right), \quad (20)$$

where σ denotes the roughness of surface, \mathbf{h} is the half vector, and δ is the angle between \mathbf{n} and \mathbf{h} . In the experiment, we set the diffuse albedo $\rho = 1$ and the specular albedo $\rho_s = 0.5$. The roughness parameter varies in the range $[0.005, 0.29]$ to simulate different highlight levels. Note that in the BRDF-model scenes, the observed image intensity is the combination of ideal Lambertian diffuse reflection and specular reflection components.

MERL scenes: The MERL database contains the reflectance data of 100 real materials which were captured using a BRDF device. In our experiment, the 100 reflectance data were used to render synthetic images. Note that as the MERL database was acquired from real-world materials, the diffuse reflections of the synthesized object surfaces are not ideally Lambertian. This puts forward challenges to photometric stereo methods.

Real scenes: The real imaging system consists of a QImaging Retiga Exi monochrome camera and a number of light-emitting diode light sources. The height and width of the planar grid is 1.2 m, which is relatively compact. In this setup, the spatial distance between two neighboring light sources is 0.6 m for the 3×3 structure and 0.4 m for the 4×4 structure. The distance between the light source grid and the objects is 1.8 m, and the maximum height of the objects is 15 cm. With this setup, we can assume parallel light direction and orthographic projection, which is required in photometric stereo. For practical consideration, the acquired images are in the low dynamic range $[0, 255]$.

The BRDF-model and MERL image scenes were rendered using three object surfaces, i.e., *Ball*, *Bunny*, and *Mozart*. Figure 5 shows the example images and normal maps of these three surfaces. In these images, both attached and cast shadows were rendered, whose intensity values are 0's. All the images have 256×256 pixels. As the reflectance data in the MERL database are of RGB format, the rendered color MERL images were transformed to gray-scale ones. In order to keep the imaging conditions of the synthetic and real scenes identical, the dynamic ranges of the synthetic images were adjusted as follows. First, we computed the median of all the intensities of the n images. Second, we rescaled the image intensities linearly such that the median intensity was mapped to 0.3. Finally, the rescaled image intensities were

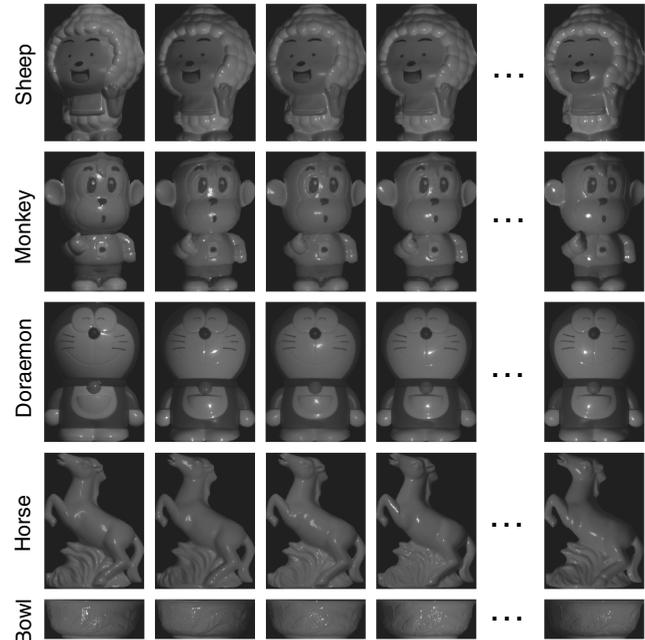


Fig. 6 Example images of the real objects *Sheep*, *Monkey*, *Doraemon*, *Horse*, and *Bowl*.

clipped into the range $[0, 1]$. We note that this manipulation is similar to the exposure adjustment in the real imaging system. It will also eliminate the effect of albedo difference among different materials.

Figure 6 shows the example images of five real objects, namely, *Sheep*, *Monkey*, *Doraemon*, *Horse*, and *Bowl*. As can be observed, in addition to highlights there are also attached and cast shadows in these images. We note that the surface albedos of real objects can be different from those of training balls, which may affect the accuracy of highlight classification. A possible solution to this problem is to cluster the surface points into groups with similar surface albedos and then rescale the image intensities such that the median intensities of individual groups are all mapped to 0.3. Nevertheless, we find that the influence of surface albedo is not obvious in normal estimation, and thus we just linearly map the original image intensity $[0, 255]$ to range $[0, 1]$ in the experiment.

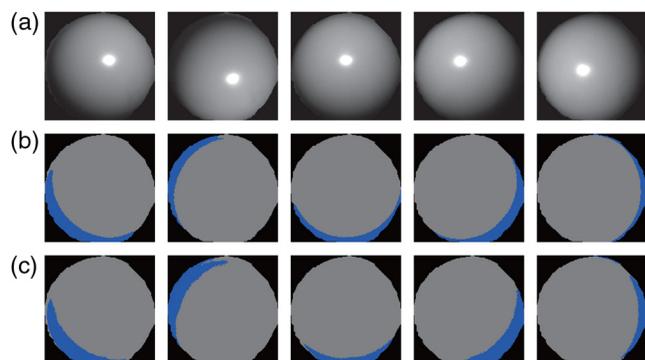


Fig. 7 Shadow detection of the ball scene. (a) Ball images. (b) Manually labeled shadow maps. (c) Detected shadow maps when $\eta = 0.5$. In (b) and (c), the shadow pixels are in blue and the nonshadow pixels are in gray.

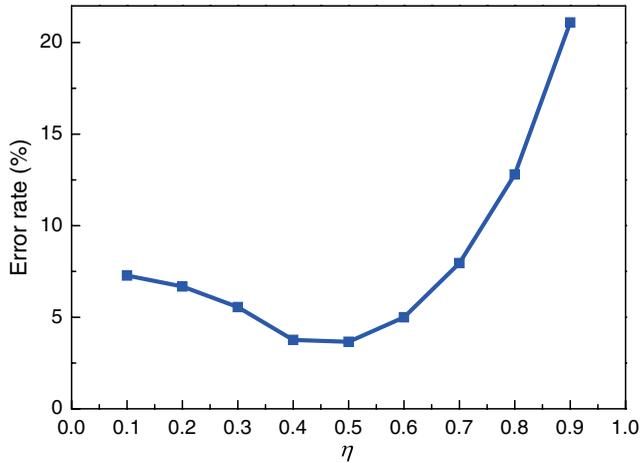


Fig. 8 Error rates of shadow detection with respect to different η values.

4.1 Accuracy of Shadow Detection

Unlike the synthetic images, the image intensities of the shadows in real images are nonzeros due to ambient illumination or inter-reflection. We use real ball images to determine the η value for the adaptive threshold in Eq. (6). Figures 7(a) and 7(b) show the example images and the manually labeled shadow masks, respectively. The appropriate η value is determined according to the error rate, which is defined as

$$\text{error rate} = \frac{\text{number of misclassified pixels}}{\text{number of image pixels}} \times 100\%. \quad (21)$$

Figure 8 shows that $\eta = 0.5$, which results in an error rate $< 5\%$, is a good choice for shadow detection. Figure 7(c) indicates that the resultant shadow maps are visually quite close to the manually labeled ones. In the following experiment, we fix $\eta = 0.5$ for shadow detection in real images.

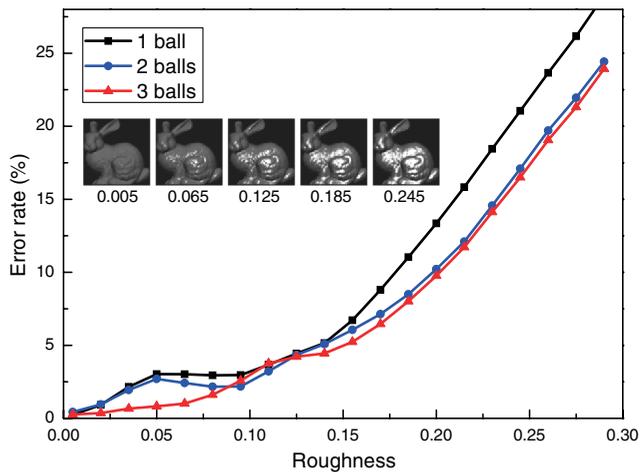


Fig. 9 Error rates of highlight detection with respect to different roughness levels in the cases of one, two, and three training balls.

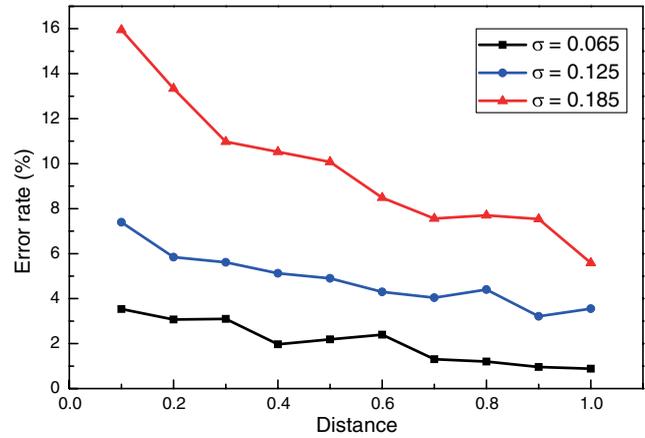


Fig. 10 Error rates of highlight detection with respect to different spatial distances d (in meters) between neighboring lighting sources in cases of different roughness levels σ .

4.2 Accuracy of Highlight Detection

As mentioned above, highlight detection is accomplished using SVM classifiers and the synthetic *Ball* scenes are employed in the training procedure. It is required that the trained classifiers should perform well on object surfaces with various highlight levels. In this regard, we render *Ball* scenes under the 3×3 structure using the Cook-Torrance reflectance model. The roughness parameter σ varies in the range $[0.005, 0.29]$, which is considered to be quite wide.

We evaluate the error rates of highlight detection in the cases of one, two, and three training balls on the synthetic *Bunny* scene. In each case, the training balls are selected according to the criterion that their roughness parameters evenly divide the whole roughness range. Figure 9 shows the error rates of highlight detection with respect to various roughness levels. In all cases, the error rates become high when the roughness increases. This is because large roughness levels correspond to wide specular lobes, which will result in more overlapping highlight regions in photometric images. The two-ball and three-ball cases exhibit obvious improvement over the one-ball case. To balance classification accuracy and computation burden, we use two training balls in the following experiment. Our investigation indicates that, although the error rate does not seem very good in the roughness range $[0.005, 0.075]$, the quality of the estimated normal almost does not degrade due to the incorporation of ℓ_1 regularization in Eq. (13).

We also evaluate the error rates of highlight detection with respect to different distances, d , between neighboring light sources. Figure 10 shows the results in cases of roughness levels $\sigma = 0.065, 0.125, \text{ and } 0.185$. It is observed that

Table 1 Error rates of highlight detection for the linear discriminant analysis (LDA) and support vector machine (SVM) classifiers with respect to various roughness levels.

Roughness	0.05	0.10	0.15	0.20	0.25
LDA (%)	5.65	8.08	11.7	16.2	21.9
SVM (%)	2.70	2.35	5.76	10.2	18.0

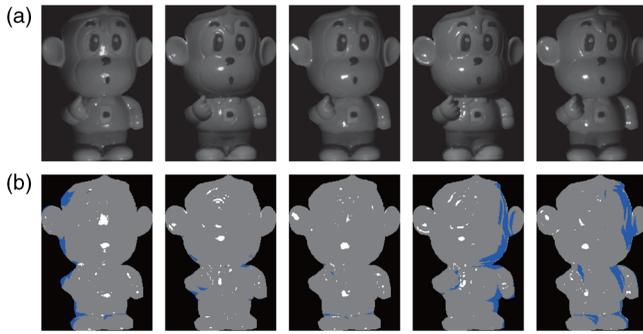


Fig. 11 Detected shadow and highlight pixels of the *Monkey* object. (a) Input images. (b) Detection results. The shadow pixels are in blue, the specular pixels are in white, and the diffuse pixels are in gray.

for all roughness levels, the error rates generally decrease when d increases. This is expected as a smaller distance will generate more serious highlight overlapping among neighboring light sources. In the experiment, we adopt $d = 0.6$, which is an appropriate balance of highlight detection accuracy and system compactness.

It is of interest to investigate if a simple classifier can also accomplish the highlight detection task. Table 1 compares the error rates of highlight detection of linear discriminant analysis (LDA) and SVM classifiers. It is observed that LDA performs worse than SVM on various roughness levels. This indicates that the highlight detection problem is relatively difficult, especially for large roughness levels.

Figure 11 shows both the shadow and highlight detection results on the *Monkey* object. It is observed that the highlight spots in the photometric images are successfully located. The attached shadows, and even a portion cast shadows, are also reliably detected. This verifies that, although the parameter η

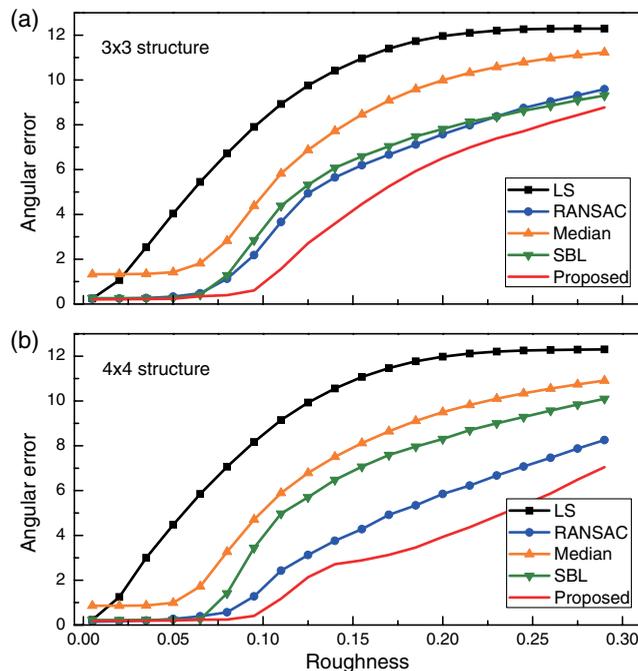


Fig. 12 Angular errors (in degrees) of the photometric stereo methods on the *Bunny* surface when using the 3×3 and 4×4 structures.

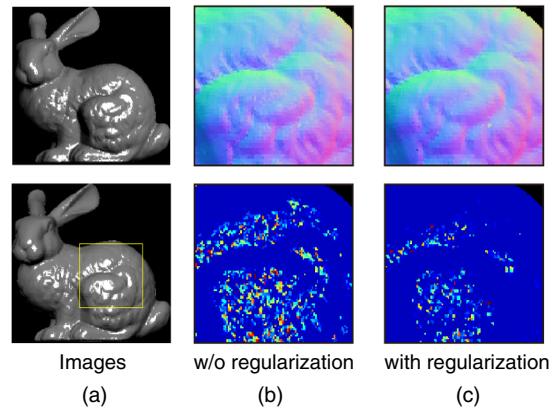


Fig. 13 Normal estimation with and without ℓ_1 regularization. (a) Two images and the marked region. (b) Estimated normal map and its spatial error without regularization. (c) Estimated normal map and its spatial error with regularization.

in Eq. (6) was learned from a simple ball scene, it is applicable to much more complex surfaces.

4.3 Results on BRDF-Model Scenes

We use the BRDF-model scenes to evaluate the photometric stereo methods on synthetic images with varying roughness levels. Figure 12 shows the average angular errors of estimated normal maps of the *Bunny* surface. As shown, the proposed method performs the best, followed by the RANSAC and SBL methods. The median method, although performing better than the LS method, does not bring forth a satisfactory accuracy.

It is worth noting that the angular errors of the proposed method are quite low when the roughness is < 0.1 . In addition, distributions of the angular errors are considerably different from the distributions of the highlight error rate in Fig. 9. This is partially due to the reason that, in addition to highlight detection, the proposed method also employs ℓ_1 regularization in normal estimation. It is verified in Fig. 13 that the quality of the normal map can be improved by applying ℓ_1 regularization.

Table 2 lists the average angular errors of the photometric stereo methods on the *Ball*, *Bunny*, and *Mozart* surfaces for a typical roughness $\sigma = 0.095$. As shown, the proposed method performs the best among all five methods. Furthermore, compared with the 3×3 structure, the proposed method produces improved accuracy under the 4×4 structure. This is because the 4×4 structure contains more collinear combinations, which is essential for the proposed method. It is interesting to note that three methods (LS, median, and SBL) perform slightly worse when adapted from the 3×3 structure to the 4×4 structure. This is because, although more images are available in the 4×4 structure, the highlight spots are more densely distributed. This in turn verifies the superiority of the proposed method in dealing with overlapping highlights.

Figure 14 illustrates the spatial angular errors of the normal maps produced by different methods. As observed, the normal maps of the LS and median methods are greatly biased. The RANSAC method performs better than the SBL method, but still much worse than the proposed method.

Table 2 Average angular errors (in degrees) of the estimated normal maps of the two structures. The surface roughness of the three objects are 0.095. The lower errors are in bold.

	3 × 3 structure					4 × 4 structure				
	LS	RANSAC	Median	SBL	Proposed	LS	RANSAC	Median	SBL	Proposed
<i>Ball</i>	5.63	1.47	3.41	1.97	0.43	5.75	1.00	3.39	2.40	0.29
<i>Bunny</i>	7.90	2.18	4.39	2.85	0.61	8.17	1.28	4.72	3.45	0.41
<i>Mozart</i>	14.39	4.05	7.21	5.50	1.17	14.92	2.43	8.58	6.95	0.60

Note: LS, least squares; RANSAC, random sample consensus; SBL, sparse Bayesian learning.

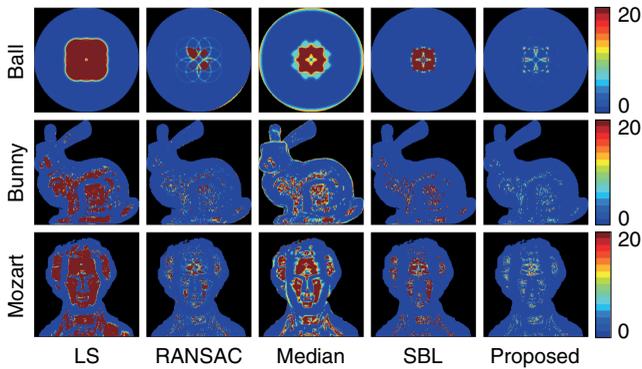


Fig. 14 Angular errors (in degrees) of the three surfaces in the 3 × 3 structure in case of roughness $\sigma = 0.095$.

Table 3 Computational times (unit: seconds) of different methods in case 4 × 4 structure and roughness $\sigma = 0.095$.

Method	LS	RANSAC	Median	SBL	Proposed
<i>Ball</i>	7	707	83	190	103
<i>Bunny</i>	6	548	65	146	81
<i>Mozart</i>	5	498	57	132	78

Table 3 lists the computational times of different methods, including the outlier detection and normal estimation steps. All the methods were implemented using MATLAB® on a personal computer with 2.20 GHz GPU and 8 GB RAM. As observed, the proposed method runs faster than the RANSAC and SBL methods, but lower than the median and LS methods. For images with 256 × 256 pixels, the computational times of the proposed method ranges from ~80 to 100 s, with about half of this time for highlight detection and half for normal estimation. Note that the computational efficiency can be much improved if we implement the proposed method using C++ programming language.

4.4 Results on MERL Scenes

Figure 15 shows the angular errors of the five photometric stereo methods under the 3 × 3 and 4 × 4 structures. The

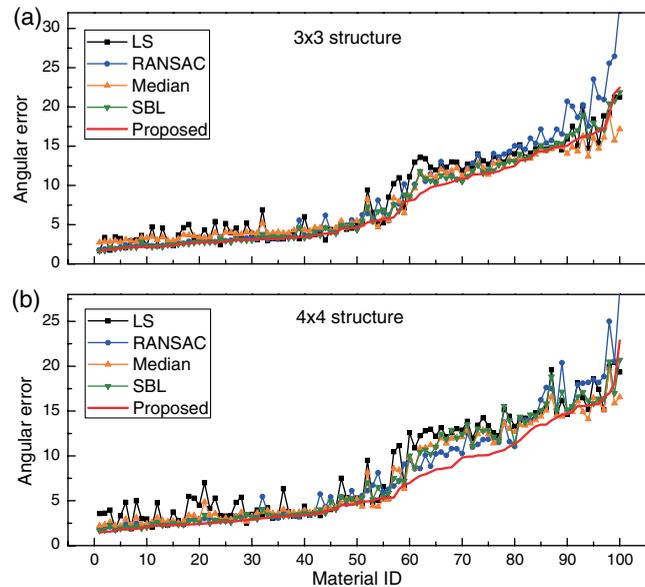


Fig. 15 Angular errors (in degrees) of normal maps on 100 different MERL materials under the (a) 3 × 3 and (b) 4 × 4 structures. The material IDs are independently numbered under the two structures, in the ascending order of angular errors by the proposed method.

material IDs in Figs. 15(a) and 15(b) are, respectively, numbered in the ascending order of angular errors by the proposed method. It is observed that for all those methods, the magnitudes of the angular errors are much larger than those in the BRDF-model scenes (see Fig. 12). This is because the diffuse and specular components are not as ideal as those produced by BRDF models, as the MERL database is acquired from real materials. Overall, the proposed method preforms better than the other methods under both structures. The improvement is relatively obvious in the 4 × 4 structure case, as illustrated in Fig. 15(b).

Table 4 lists the average angular errors of the estimated normals for the 100 MERL materials. For the proposed method, angular errors under the 4 × 4 structure are lower than those under the 3 × 3 structure. Overall, the proposed method performs best, followed by the median, RANSAC, and SBL methods. Figure 16 illustrates the spatial angular error of the *Bunny* surface with material ID = 52 (please refer to Fig. 17 and Table 5 for material information of the material ID) under the 3 × 3 and 4 × 4 structures. As

Table 4 Mean angular errors (in degrees) of the estimated normal maps of the 100 materials in the two structures. The lower errors are in bold.

	3 × 3 structure					4 × 4 structure				
	LS	RANSAC	Median	SBL	Proposed	LS	RANSAC	Median	SBL	Proposed
<i>Ball</i>	8.13	6.95	6.75	7.55	6.68	7.56	6.74	6.23	7.46	6.13
<i>Bunny</i>	8.56	8.68	7.94	7.85	7.44	8.61	8.08	7.74	7.87	6.87
<i>Mozart</i>	12.48	11.35	10.66	10.74	10.14	12.7	10.28	10.96	10.91	9.28

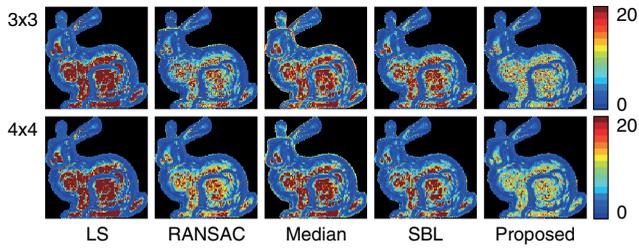


Fig. 16 Angular errors (in degrees) of the *Bunny* surface with material ID = 52.

observed, the proposed method performs better than the other methods under both structures.

4.5 Results on Real Scenes

We also evaluate the five methods on real surfaces under the 3 × 3 and 4 × 4 structures. As ground truth normal maps are not available for real surfaces, we only conduct visual comparisons. Figure 18 illustrates the estimated normal maps of the real object *Sheep*. It is observed that the normal maps produced by the LS method are greatly biased in the hair and hand regions. Although the RANSAC,

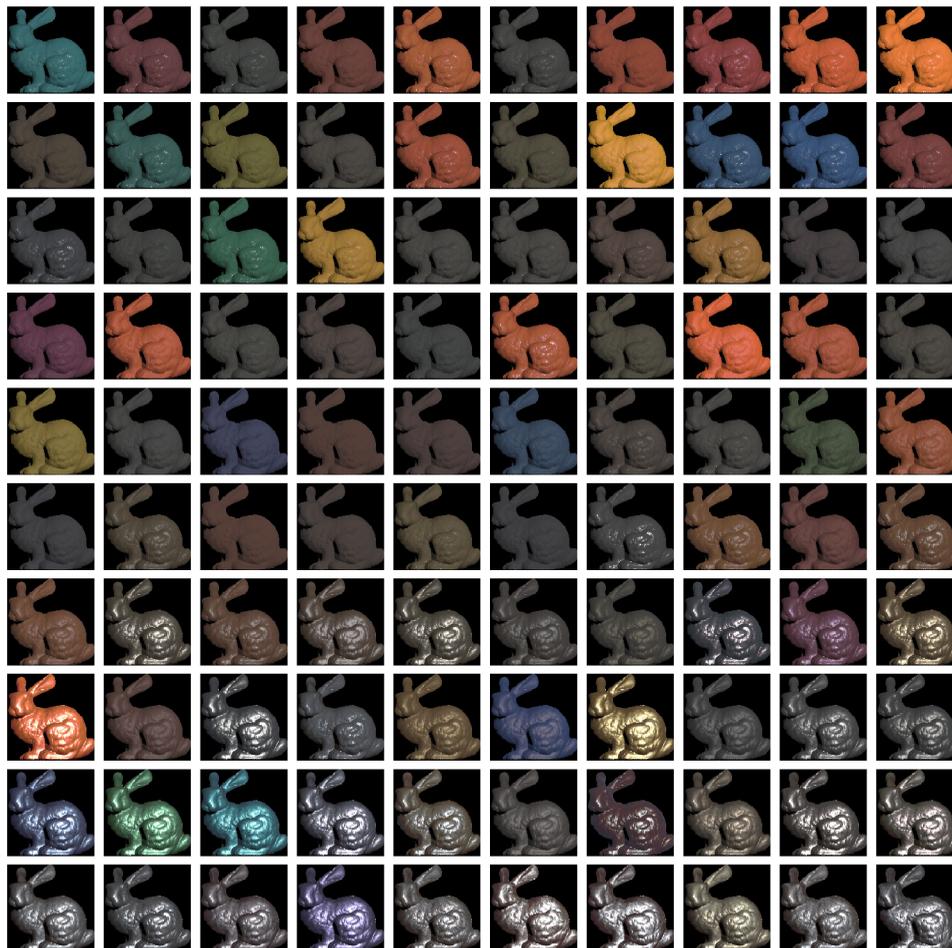


Fig. 17 Material IDs under the 4 × 4 structure [in accordance to Fig. 15(b)] and the corresponding rendered *Bunny* surfaces. The rendered images are tone mapped for visualization.

Table 5 Material IDs under the 4×4 structure [in accordance with Fig. 15(b)] and the corresponding material names.

1	Green-plastic	26	Alumina-oxide	51	Green-fabric	76	Dark-blue-paint
2	Specular-violet-phenolic	27	Pink-jasper	52	Fruitwood-241	77	Gold-metallic-paint
3	Aventurnine	28	Yellow-matte-plastic	53	Light-brown-fabric	78	Black-soft-plastic
4	Pink-plastic	29	Pink-fabric	54	Beige-fabric	79	Two-layer-silver
5	Specular-maroon-phenolic	30	White-paint	55	Yellow-phenolic	80	Silver-metallic-paint2
6	Gray-plastic	31	Violet-rubber	56	White-fabric2	81	Blue-metallic-paint2
7	Red-fabric2	32	Red-fabric	57	Silicon-nitride	82	Green-metallic-paint2
8	Purple-paint	33	Delrin	58	Natural-209	83	Green-metallic-paint
9	Specular-red-phenolic	34	Pink-felt	59	Pink-fabric2	84	Color-changing-paint1
10	Specular-orange-phenolic	35	Polyethylene	60	Ipswich-pine-221	85	Black-phenolic
11	White-fabric	36	Red-phenolic	61	Colonial-maple-223	86	Black-oxidized-steel
12	Specular-green-phenolic	37	White-diffuse-bball	62	Gold-metallic-paint2	87	Black-obsidian
13	Yellow-plastic	38	Light-red-paint	63	Cherry-235	88	Alum-bronze
14	Pure-rubber	39	Dark-red-paint	64	Special-walnut-224	89	Nickel
15	Red-specular-plastic	40	Neoprene-rubber	65	Two-layer-gold	90	Color-changing-paint3
16	Specular-white-phenolic	41	Yellow-paint	66	Silver-paint	91	Hematite
17	Orange-paint	42	Teflon	67	Pearl-paint	92	Aluminium
18	Specular-blue-phenolic	43	Blue-fabric	68	Color-changing-paint2	93	Chrome
19	Blue-acrylic	44	Polyurethane-foam	69	Violet-acrylic	94	Blue-metallic-paint
20	Maroon-plastic	45	Black-fabric	70	Gold-metallic-paint3	95	ss440
21	PVC	46	Blue-rubber	71	Red-metallic-paint	96	Chrome-steel
22	Nylon	47	Pickled-oak-260	72	Dark-specular-fabric	97	Steel
23	Green-acrylic	48	White-marble	73	Silver-metallic-paint	98	Brass
24	Specular-yellow-phenolic	49	Green-latex	74	Specular-black-phenolic	99	Tungsten-carbide
25	White-acrylic	50	Red-plastic	75	Gold-paint	100	Grease-covered-steel

median, and SBL methods produce improved normal estimation, the discontinuous effects caused by seriously overlapping highlights are still obvious. In comparison, the proposed method produces relatively smooth normals in the corresponding regions, especially under the 4×4 structure. The estimated normal maps of the *Monkey* surface are illustrated in Fig. 19. Clearly, the proposed method produces the best normal maps in the face region. As the 4×4 structure provides more collinear combinations, it provides further an improved normal map when compared with the 3×3 structure. The proposed method also exhibits its superiority

over its competitors on the *Doraemon* surface, which is illustrated in Fig. 20. The normal maps produced by the proposed method are very natural, while those by other methods contain obviously visible artifacts.

Figure 21 shows the normal maps of the *Horse* object estimated by the proposed method, as well as the reconstructed 3-D surfaces under two different viewing directions. As observed, the normal maps on the horse body appear quite smooth, almost not affected by highlights. Figure 22 shows the normal maps and reconstructed surfaces of the *Bowl* object. It is found that the estimated normal maps

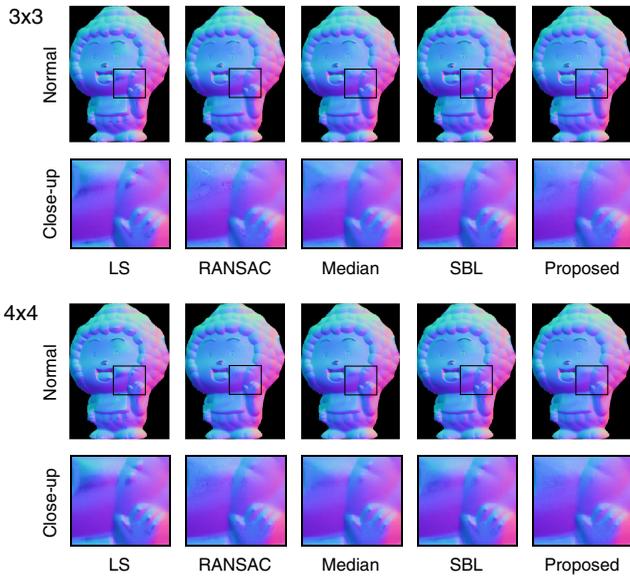


Fig. 18 Estimated normal maps of the real object *Sheep* using different methods.

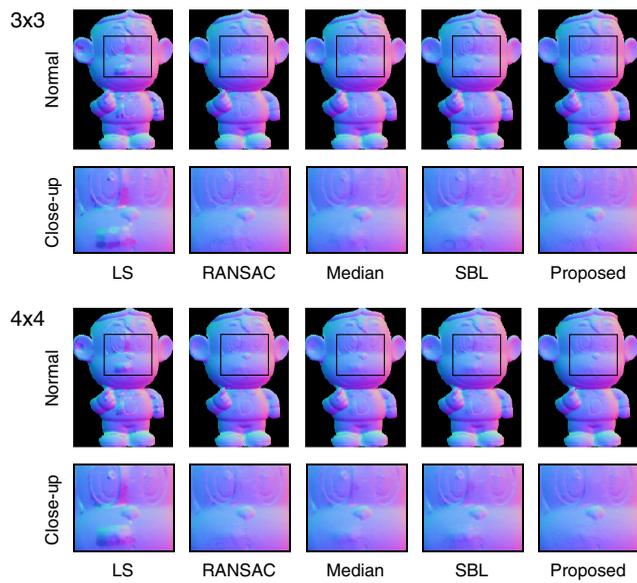


Fig. 19 Estimated normal maps of the real object *Monkey* using different methods.

are unbiased by highlights and the details are well kept on the reconstructed surfaces.

5 Conclusions

This paper proposes a robust photometric stereo method by using a structural arrangement of light sources. The idea of the proposed method is to utilize the intensity correlation among a set of collinear combinations. The shadow pixels are detected using an adaptive thresholding strategy, and the highlight pixels are located using the SVM classifiers. The normal map is further improved using an ℓ_1 regularization scheme. Experimental evaluations were conducted on two typical light source arrangements, i.e., the 3×3

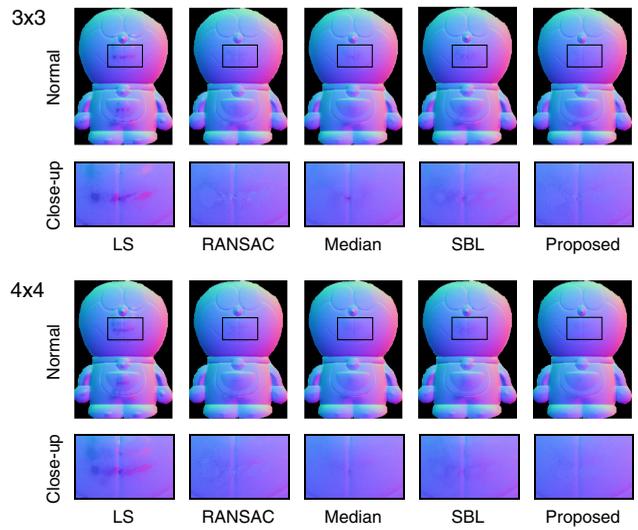


Fig. 20 Estimated normal maps of the real object *Doraemon* using different methods.

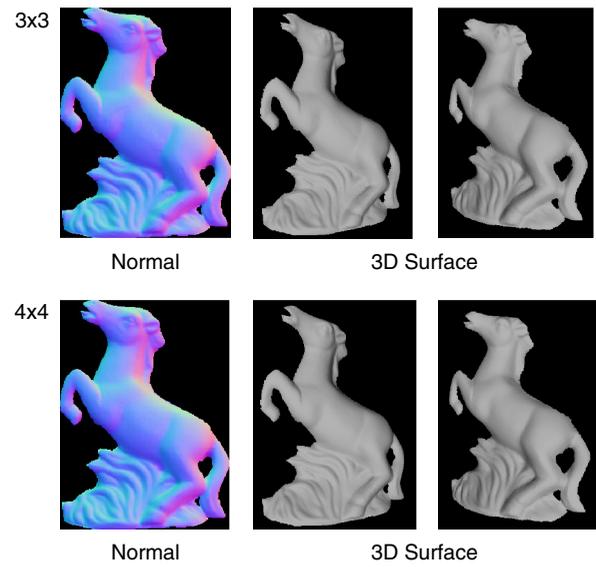


Fig. 21 Estimated normal maps and reconstructed 3-D surfaces of real object *Horse* using the proposed method.

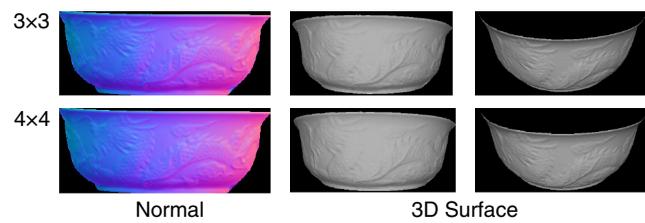


Fig. 22 Estimated normal maps and reconstructed 3-D surfaces of real object *Bowl* using the proposed method.

and 4×4 structures. It is validated that the proposed method outperforms the state-of-the-art techniques on both synthetic and real scenes.

A limitation of the proposed method is that the normal estimation accuracy on the MERL scenes is not quite satisfactory. This is because the diffuse reflection of real materials

is not ideal Lambertian, and consequently, the intensity deviation is not a very discriminable feature for highlight detection. In the future work, we will investigate alternative structural arrangement of light sources, as well as improved techniques for highlight and shadow detection.

Appendix

As described in the Experiment section, the material ID of the MERL database is numbered in the ascending order of angular errors by the proposed method. As reference, we show material IDs and the corresponding rendered *Bunny* surfaces in Fig. 17 and the corresponding material names in Table 5. Note that the material IDs are in accordance with those under the 4×4 structure in Fig. 15(b).

Acknowledgments

We thank the anonymous reviewers for their valuable suggestions. This work was supported by the Natural Science Foundation of China under grants 61371160 and 61271339, and Science Technology Department of Zhejiang Province under grant 2010R50006.

References

- R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.* **19**(1), 139–144 (1980).
- T. Okabe, I. Sato, and Y. Sato, "Attached shadow coding: estimating surface normals from shadows under unknown reflectance and lighting conditions," in *Proc. of the 12th IEEE Int. Conf. on Computer Vision*, pp. 1693–1700, IEEE, Kyoto, Japan (2009).
- R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. Graph.* **1**(1), 7–24 (1982).
- A. S. Georgiades, "Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo," in *Proc. of the 9th IEEE Int. Conf. on Computer Vision*, pp. 816–823, IEEE, Nice, France (2003).
- H. S. Chung and J. Jia, "Efficient photometric stereo on glossy surfaces with wide specular lobes," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, Anchorage, Alaska (2008).
- D. B. Goldman et al., "Shape and spatially-varying BRDFs from photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(6), 1060–1071 (2010).
- M. Chandraker, J. Bai, and R. Ramamoorthi, "On differential photometric reconstruction for unknown, isotropic BRDFs," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(12), 2941–2955 (2013).
- N. Alldrin and D. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance: a stratified photometric stereo approach," in *Proc. of the 11th IEEE Int. Conf. on Computer Vision*, pp. 1–8, IEEE, Rio de Janeiro, Brazil (2007).
- B. Shi et al., "Elevation angle from reflectance monotonicity: photometric stereo for general isotropic reflectances," in *Proc. of European Conf. on Computer Vision*, pp. 455–468, Springer, Florence, Italy (2012).
- Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Classification of photometric factors based on photometric linearization," in *Proc. of the Asian Conf. on Computer Vision*, pp. 613–622, Springer, Hyderabad, India (2006).
- Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Analysis of photometric factors based on photometric linearization," *J. Opt. Soc. Am. A* **24**(10), 3326–3334 (2007).
- D. Miyazaki, K. Hara, and K. Ikeuchi, "Median photometric stereo as applied to the segoonko tumulus and museum objects," *Int. J. Comput. Vis.* **86**(2–3), 229–242 (2010).
- L. Wu et al., "Robust photometric stereo via low-rank matrix completion and recovery," in *Proc. of the Asian Conf. on Computer Vision*, pp. 703–717, Springer, Queenstown, New Zealand (2010).
- S. Ikehata et al., "Robust photometric stereo using sparse regression," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 318–325, IEEE, Providence, Rhode Island (2012).
- C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*, Vol. 1, Springer New York (2006).
- C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.* **20**(3), 273–297 (1995).
- C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Syst. Technol.* **2**(3), 27 (2011).
- E. Coleman, Jr. and R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," *Comput. Graph. Image Process.* **18**(4), 309–328 (1982).
- S. P. Mallick et al., "Beyond Lambert: reconstructing specular surfaces using color," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 619–626, IEEE, San Diego, California (2005).
- R. Liu and J. Han, "Recovering surface normal of specular object by Hough transform method," *IET Comput. Vis.* **4**(2), 129–137 (2010).
- S. Barsky and M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(10), 1239–1252 (2003).
- V. Argyriou and M. Petrou, "Recursive photometric stereo when multiple shadows and highlights are present," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–6, IEEE, Anchorage, Alaska (2008).
- M. Chandraker, S. Agarwal, and K. David, "Shadowcuts: photometric stereo with shadows," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, Minneapolis, Minnesota (2007).
- T. P. Wu et al., "Dense photometric stereo: a Markov random field approach," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(11), 1830–1846 (2006).
- H. L. Shen and Y. Cheng, "Calibrating light sources by using a planar mirror," *J. Electron. Imaging* **20**(1), 013002 (2011).
- Y. Cheng, H. L. Shen, and X. Du, "Robust surface reconstruction from gradient fields," *Electron. Lett.* **48**(7), 375–376 (2012).
- S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge (2004).
- W. Matusik et al., "A data-driven reflectance model," *ACM Trans. Graph.* **22**(3), 759–769 (2003).
- A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of BRDF models," in *Proc. of the Sixteenth Eurographics Conf. on Rendering Techniques*, pp. 117–126, Eurographics Association, Aire-la-Ville, Switzerland (2005).

Tian-Qi Han received his BE degree from Tianjin University, China, in 2011. He is currently pursuing his PhD degree in the Department of Information Science and Electronic Engineering, Zhejiang University, China. His research interest is computer vision, especially in photometric stereo.

Yue Cheng received his BE and PhD degrees in electronic engineering from Zhejiang University, China, in 2008 and 2013, respectively. His research interests are image processing and computer vision.

Hui-Liang Shen received his BE and PhD degrees in electronic engineering from Zhejiang University, China, in 1996 and 2002, respectively. He was a research assistant, research associate, and research fellow at the Hong Kong Polytechnic University from 2001 to 2005. He is currently a professor in the Department of Information Science and Electronic Engineering, Zhejiang University. His research interests include color imaging, image processing, and computer vision.

Xin Du received his BE degree in electronic engineering from Hangzhou University, China, in 1997, and his ME and PhD degrees from Zhejiang University, China, in 2000 and 2003, respectively. He is currently an associate professor in the Department of Information Science and Electronic Engineering, Zhejiang University. His research interests are image/video processing and computer vision.