

# Photometric Stereo for General Isotropic Reflectances by Spherical Linear Interpolation

Si Li<sup>a</sup>, Boxin Shi<sup>b</sup>

<sup>a</sup>Brandeis University, 415 South St, Waltham, MA, United States, 02453

<sup>b</sup>Singapore University of Technology and Design, 8 Somapah Road, Singapore, 487372

**Abstract.** We present a practical photometric stereo method that works with general isotropic reflectances. Unlike previous approaches that use special hardware setups or dense measurements, our method only requires dozens of random yet known lighting directions. By spherically interpolating the light source directions to approximate the scene appearance under desired lighting directions, our method separately computes azimuth and elevation angles of the surface normal for each pixel. The effectiveness and accuracy of the proposed method is validated using 100 isotropic materials from MERL database and various real-world objects.

**Keywords:** photometric stereo, reflectance, isotropic, surface normal, shadow.

**Address all correspondence to:** Boxin Shi, Singapore University of Technology and Design, 8 Somapah Road, Singapore, 487372; Tel: +65-8408-8676; Fax: +65-6770-5161; E-mail: [boxin.shi@gmail.com](mailto:boxin.shi@gmail.com)

**Copyright notice:** Copyright 2015 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited. Please access SPIE's official online version at <http://dx.doi.org/10.1117/1.OE.54.8.083104>.

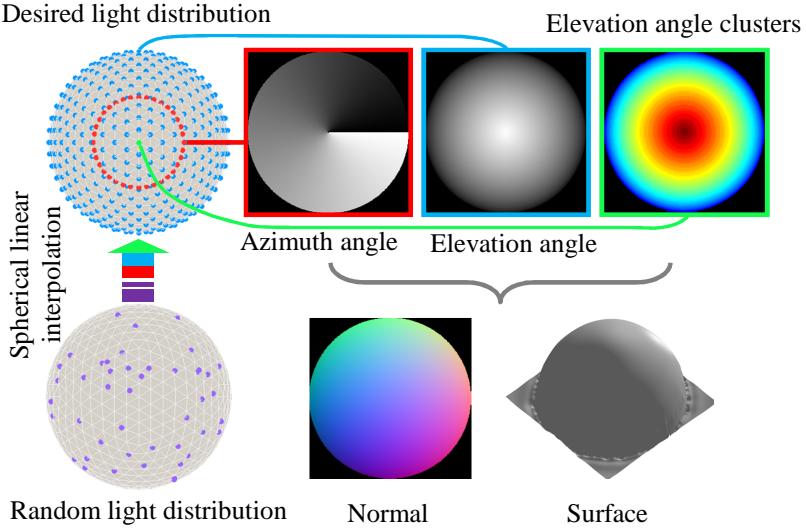
**Citation format:** Si Li and Boxin Shi, "Photometric stereo for general isotropic reflectances by spherical linear interpolation", Opt. Eng. 54(8), 083104 (Aug 21, 2015).

## 1 Introduction

Photometric stereo estimates a pixel-wise surface orientation given a set of images taken under varying illumination from a fixed viewpoint.<sup>1</sup> Conventional photometric stereo is a well-defined problem, but it has strong restrictions on the reflectance property of a target object, i.e., the Lambertian reflectance model. Although the Lambertian model has been broadly adopted in many vision problems because of its simplicity, the majority of real-world materials and objects have reflectance properties beyond its description.

This paper presents a simple and practical photometric stereo method that works with various real-world materials. Unlike previous approaches that use special lighting configurations or dense lighting distributions, our method requires only dozens of images taken under random yet known lighting directions. By capitalizing on the fact that off-specular signals of most real-world Bidirectional Reflectance Distribution Functions (BRDFs) are smooth,<sup>2</sup> our method spherically interpolates the lighting directions to approximate scene appearances under desired lighting configurations. From the interpolated appearances, we compute azimuth and elevation angles of the surface normal independently, as illustrated in Fig. 1.

Our method is partially built upon the conclusion from Alldrin and Kriegman,<sup>3</sup> which uses a ring light (the lighting directions indicated as red in Fig. 1) for recovering the azimuth angle of the surface normal from scenes with isotropic BRDFs. From the random yet known lighting directions,



**Fig 1** Illustration of our approach. Our method uses the spherical linear interpolation for obtaining desired appearances from the original input. From the interpolated appearances, our method computes azimuth and elevation angles of surface normals.

our method uses the spherical linear interpolation to approximate the appearances under the ring light distribution for computing the azimuth angles. In addition, we produce appearances under the uniform lighting distribution (the lighting directions indicated as blue) by the same interpolation method for estimating the elevation angles. Based on the newly proposed shadow ratio approach, we estimate the elevation angle in a closed-form solution. We also introduce a method for refining the elevation angle estimates by clustering using an interpolated appearance that approximates the image under the lighting direction which is the same as viewing direction (the lighting direction indicated as green).

We make a thorough evaluation for the validity of the proposed interpolation approach and the accuracy of normal estimation in comparison with the conventional method<sup>1</sup> and a robust photometric stereo method<sup>4</sup> using 100 real-world isotropic BRDFs in the MERL database<sup>2</sup> as well as various real-world objects. Quantitative comparisons also verify that our method outperforms state-of-the-art photometric stereo methods<sup>5-7</sup> for several general isotropic BRDFs. While our method assumes that there is one dominant isotropic BRDF on the object for clustering elevation angles, as we will see in the real-world experiments, it can work with various objects that do not contain significant and abrupt changes of BRDFs across space.

## 2 Related Work

There are many previous approaches for making photometric stereo work with various real-world materials that are beyond the description of the Lambertian model. One class of approaches uses a composite representation for reflectance, namely “Diffuse + Specular (D + S)”, where the diffuse component is typically represented by a Lambertian term. Sato and Ikeuchi’s method<sup>8</sup> examines the diffuse and specular subspaces in the temporal-color space and extracts the two components by principal component analysis. Hansson and Johansson<sup>9</sup> use crossed polarizers to eliminate the specular part. Barsky and Petrou<sup>10</sup> use the spectral information and directional cues to exclude shadows and highlights from the normal estimation process. There is a class of methods that

uses parametric representations of specular reflections. Georghiades<sup>11</sup> uses the Torrance-Sparrow model. Chung and Jia<sup>12</sup> and Goldman et al.<sup>6</sup> use the Ward model to describe specular lobes for surface normal estimation. All these three methods optimize the reflectance parameters and surface normals alternatively with complicated nonlinear optimizations.

A class of techniques uses robust approaches for removing the specularity with the “D + S” assumption. There are approaches based on RANSAC,<sup>13</sup> median filtering,<sup>14</sup> maximum feasible subsystem,<sup>15</sup> rank minimization,<sup>4</sup> sparse Bayesian regression,<sup>16</sup> and structural light sources.<sup>17</sup> Spatial information has also been used for robustly solving the problem. Prior works include a method that uses belief propagation by Tang et al.,<sup>18</sup> graph cut and expectation maximization based methods proposed by Wu and Tang.<sup>19,20</sup> These methods rely on the statistical information from varying illuminations, therefore the input images usually cover a dense sampling (from 100 to 1000 different images) of spherical lighting distribution.

There are methods that handle more realistic reflectances by exploring the general BRDF properties such as isotropy and monotonicity. Alldrin and Kriegman<sup>3</sup> estimate the azimuth angle of surface normal using a ring light for materials with isotropic BRDFs by exploring their symmetric properties. Later, their method is extended to solve for both shape and reflectance in an alternating optimization framework.<sup>5</sup> More general symmetric properties of BRDFs are used by Holroyd et al.<sup>21</sup> to estimate surface normals and tangents simultaneously for anisotropic materials. Higo et al.’s method<sup>22</sup> uses monotonicity and isotropy properties of general diffuse reflectances. Other BRDF invariants are also used to solve the problem, such as the radiance similarity<sup>23</sup> and attached shadow codes<sup>24</sup> under dense lighting distribution given hundreds of images. Chandraker et al.<sup>25</sup> propose a method based on image derivatives in photometric stereo by using a special light rig for data capture. Based on the bivariate BRDF approximation,<sup>26</sup> recent methods focus on low-frequency reflectance by discarding the high-frequency specular observations, and they use bi-polynomial modeling<sup>27</sup> and bivariate regression<sup>28</sup> to model the general diffuse part of isotropic BRDFs for accurate normal estimation. The uncalibrated photometric stereo is also extended to general isotropic BRDFs by analyzing the geodesic distance of intensity profiles<sup>29</sup> or the BRDF symmetry of separated specular components.<sup>30</sup>

One of the closest exiting works to ours is the method proposed by Shi et al.<sup>7</sup> They propose to project the general isotropic BRDF to a one-dimensional space and estimate the elevation angles of surface normal from reflectance monotonicity. Both their method and our method are two-step approaches built on the azimuth angle estimation method of Alldrin and Kriegman,<sup>3</sup> and both methods theoretically require dense and uniform lighting distribution as the ideal lighting condition for elevation angle estimation. Comparing to Shi et al.,<sup>7</sup> our main advantages are: 1) We could deal with some BRDFs that do not follow the monotonicity property of one-dimensional BRDFs; 2) Due to the spherical linear interpolation employed here, our method only needs dozens of images with randomly distributed lighting directions to obtain accurate results; 3) Instead of a brute-force search for elevation angles, our elevation angle estimation method based on shadow ratio is a closed-form solution.

### 3 Background

The image formation model for conventional Lambertian photometric stereo is described as

$$I = (\rho \mathbf{n}) \cdot \mathbf{l}, \quad (1)$$

where  $I$  is the irradiance,  $\rho$  and  $\mathbf{n}$  are the surface reflectance (or albedo) and surface normal, respectively. The lighting vector  $\mathbf{l}$  encodes the intensity and direction of the incoming light. With the Lambertian assumption, reflectance  $\rho$  becomes constant.

In the non-Lambertian case, the reflectance  $\rho$  becomes a 4D BRDF in general. When the reflectance is assumed to be isotropic, it becomes a 3D function denoted as  $\rho(\theta_i, \theta_r, |\phi_i - \phi_r|)$ , where  $\theta$  and  $\phi$  are the elevation and azimuth angles of incident and reflected (subscript  $i$  and  $r$ ) lighting directions, in the normal-centered coordinate system. The three attributes in  $\rho$  can be represented by three vectors: the surface normal  $\mathbf{n}$ , lighting vector  $\mathbf{l}$ , and viewing direction  $\mathbf{v}$ . Using these, an isotropic BRDF can be parameterized as  $\rho(\mathbf{n} \cdot \mathbf{l}, \mathbf{n} \cdot \mathbf{v}, \mathbf{v} \cdot \mathbf{l})$ . By considering a shadowing effect, the image formation model becomes

$$I = \max \{\rho(\mathbf{n} \cdot \mathbf{l}, \mathbf{n} \cdot \mathbf{v}, \mathbf{v} \cdot \mathbf{l}) \mathbf{n} \cdot \mathbf{l}, 0\}. \quad (2)$$

There exists an inherent reflectance-normal ambiguity: Surface normal  $\mathbf{n}$  cannot be uniquely determined without knowing the BRDF  $\rho$ .

While the surface normal  $\mathbf{n}$  is a 3D vector, it can be represented by its azimuth and elevation angles  $(\phi_n, \theta_n)$  in the spherical coordinate system. Our method exploits invariant properties of reflectance to determine the azimuth and elevation angles, respectively.

## 4 Spherical Linear Interpolation based Approach

We use the spherical linear interpolation,<sup>31</sup> also known as “*Slerp*”, to approximate the appearances under desired lighting directions from an input random lighting distribution. The desired image under the interpolated lighting direction is a weighted average of two or three selected images from all those input. Here we describe the interpolation process in three steps:

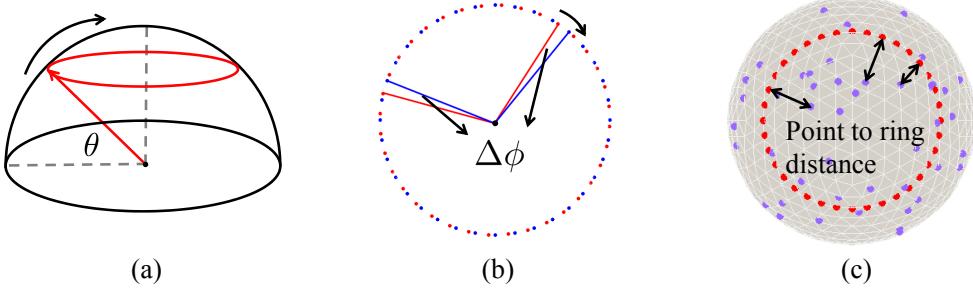
1. Project the input lighting directions onto a plane defined by the interpolation direction through a stereographic projection mapping:  $(x, y, z)$  to  $(\frac{y}{1+z}, \frac{x}{1+z})$ .
2. Apply Delaunay triangulation<sup>32</sup> to the projected data.
3. Determine the weight factor as follows. If the interpolation direction can be found inside a triangle composed by the input directions, we perform spherical linear interpolation using triangle vertices. The barycentric coordinate<sup>33</sup> of interpolation direction is calculated with respect to the vertices as the weights. If the desired interpolation direction is outside the convex hull of the input directions, the spherical linear interpolation is performed along the arc defined by the two nearest neighbors to the interpolation direction.

With these steps, images under random lights are interpolated to form appearances under desired lights as shown in Fig. 1 for estimating the azimuth and elevation angles.

### 4.1 Azimuth angle estimation

Our method uses Alldrin and Kriegman’s approach<sup>3</sup> for estimating azimuth angles. Their method exploits the bilateral symmetry property of isotropic BRDFs and uses a ring light setup for determining the azimuth angles of the surface normal.

Given the random lighting directions, our method synthetically creates a ring light distribution by the aforementioned interpolation method. We define the ring light distribution as 36 lighting



**Fig 2** Illustration of ring light parameter estimation. Our method finds the optimal ring light parameters, i.e., (a) elevation angle  $\theta$  and (b) rotation  $\Delta\phi$  within the interval; (c) depicts the definition of point to ring distance.

directions located at the same elevation angle and with a ten-degree interval of rotation angle, i.e.,  $\mathbf{r}_i = (\cos(\theta) \cos(\Delta\phi + \frac{i\pi}{18}), \cos(\theta) \sin(\Delta\phi + \frac{i\pi}{18}), \sin(\theta))$  with  $i = \{0, 1, \dots, 35\}$ , where  $\theta$  is the elevation angle, and  $\Delta\phi$  is the rotation within the lighting direction interval (Fig. 2(a), (b)). In synthesizing the ring light distribution, we compute the optimal  $(\theta, \Delta\phi)$ . For each random light  $\mathbf{l}_j$ , the distance from the input lighting direction to the ring is defined as the minimum point to point distance from  $\mathbf{l}_j$  to any point  $\mathbf{r}_i$  on the ring  $(\theta, \Delta\phi)$  (Fig. 2(c)). Our method computes the optimal parameters  $(\theta, \Delta\phi)$  for the ring light distribution by minimizing the following objective function:

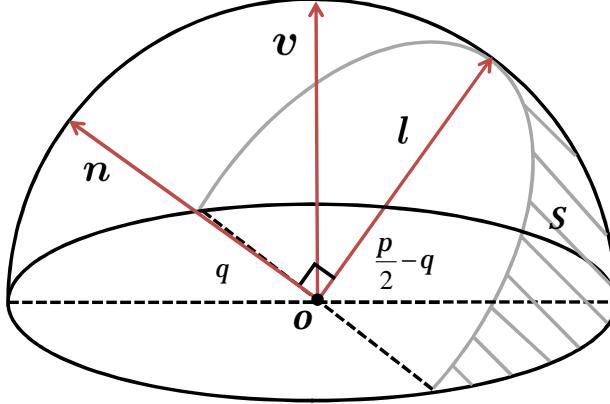
$$\underset{(\theta, \Delta\phi)}{\operatorname{argmin}} \sum_j \min_i \|\mathbf{r}_i - \mathbf{l}_j\|^2. \quad (3)$$

The objective function is optimized by an exhaustive search because the possible values of  $\theta$  and  $\Delta\phi$  are in small ranges. We search  $\theta$  from 30 to 60 degrees and  $\Delta\phi$  from 0 to 9 degrees with a step of one degree. A ring nearby the north pole or equator is excluded because these elevation angles are inadequate for the estimation of azimuth angles.  $\theta$  and  $\Delta\phi$  are properties of the ring light distribution, which need to be determined according to each input lighting distribution and they are irrelevant of surface properties. Once the elevation and rotation angles of the ring light are determined, the spherical linear interpolation is performed, and the method by Alldrin and Kriegman<sup>3</sup> is applied to obtain the azimuth angles of the surface normal.

#### 4.2 Elevation angle estimation

For estimating elevation angles of the surface normal, our method interpolates from the random lighting distribution to approximate the image appearances under uniformly distributed lights. Consider an upper hemisphere with a viewpoint direction  $\mathbf{v}$  looking towards the north pole, as illustrated in Fig. 3. If a surface normal  $\mathbf{n}$  at a scene point with any azimuth angle has an elevation angle  $\theta$  (subscript  $n$  omitted for simplicity), only the lighting directions  $\mathbf{l}$  contained in the spherical surface area  $S$  can cause attached shadows. By identifying the light source directions that make the pixels of interest be in attached shadow (for which we describe the method in the next section) and computing the surface integration of such directions over the hemisphere, we can calculate the area covered by  $S$  as follows:

$$S = 2 \int_0^{\frac{\pi}{2} - \theta} \int_0^{\frac{\pi}{2}} \sin \theta d\theta d\phi = 2 \left( \frac{\pi}{2} - \theta \right). \quad (4)$$



**Fig 3** Illustration of shadow ratio. The area in shadow  $S$  has a linear relationship with the elevation angle of surface normal  $\theta$ .

Dividing Eq. (4) by the surface area of a hemisphere, we can obtain the shadow ratio  $r_S$  as

$$r_S = \frac{S}{2\pi} = \frac{1}{2} - \frac{\theta}{\pi}. \quad (5)$$

Finally, the elevation angle  $\theta$  of this surface normal is calculated as

$$\theta = \pi \left( \frac{1}{2} - r_S \right). \quad (6)$$

Eq. (6) shows that the elevation angle of the normal has a linear relationship with the shadow ratio. Approximating the surface area with the number of discrete points covered by the area, the shadow ratio is represented by the number of lights causing the attached shadows divided by the total number of lights.

To define the uniform distribution of lightings, we use vertices determined by an icosahedron with a tessellation order of three, which results in 337 points defined on the visible hemisphere. This could be simply done by first creating a 20-sided regular solid and then subdividing the faces (by connecting the midpoint on each edge) recursively for three times. Using the spherical linear interpolation described above, we approximate the appearances under the uniform distribution of lightings using the input images.

**Refining elevation angles by clustering.** When the number of input lighting directions is limited, the accuracy of the spherical linear interpolation becomes lower. This effect becomes more significant in the estimation of elevation angles, while not in azimuth angle estimation because the elevation angle estimation requires far more directions than the input lighting directions. We use a special case of Tan and Zickler's approach<sup>34</sup> of identifying the iso-slope contour for refining the elevation angle estimates.

From the definition of isotropic BRDFs  $\rho(\mathbf{n} \cdot \mathbf{l}, \mathbf{n} \cdot \mathbf{v}, \mathbf{v} \cdot \mathbf{l})$ , if we have a scene illuminated from the same direction as the viewing direction, i.e.,  $\mathbf{l} = \mathbf{v} = (0, 0, 1)$ , the BRDF  $\rho$  becomes a single variable function of the elevation angle of surface normal  $\mathbf{n}$ . The same applies to  $\mathbf{n} \cdot \mathbf{l}$  and irradiance  $I$ . Therefore, for pixels with the same BRDF  $\rho$ , pixels with the same irradiance value share the same elevation angle of the surface normal. We again use the spherical linear

interpolation to produce the scene appearance that is illuminated from the viewing direction. In fact, such an appearance has already been included in one of the appearances under the uniform lighting distribution. Using such an image, we use  $K$ -means clustering of the image intensities to group pixels with the same elevation angle. Such a grouping strategy has also been applied to diffuse albedo for removing the ambiguities in uncalibrated photometric stereo.<sup>35</sup> The clustering result is then used to average the estimated elevation angles that are obtained by the aforementioned shadow ratio within each cluster.

## 5 Attached Shadow Detection

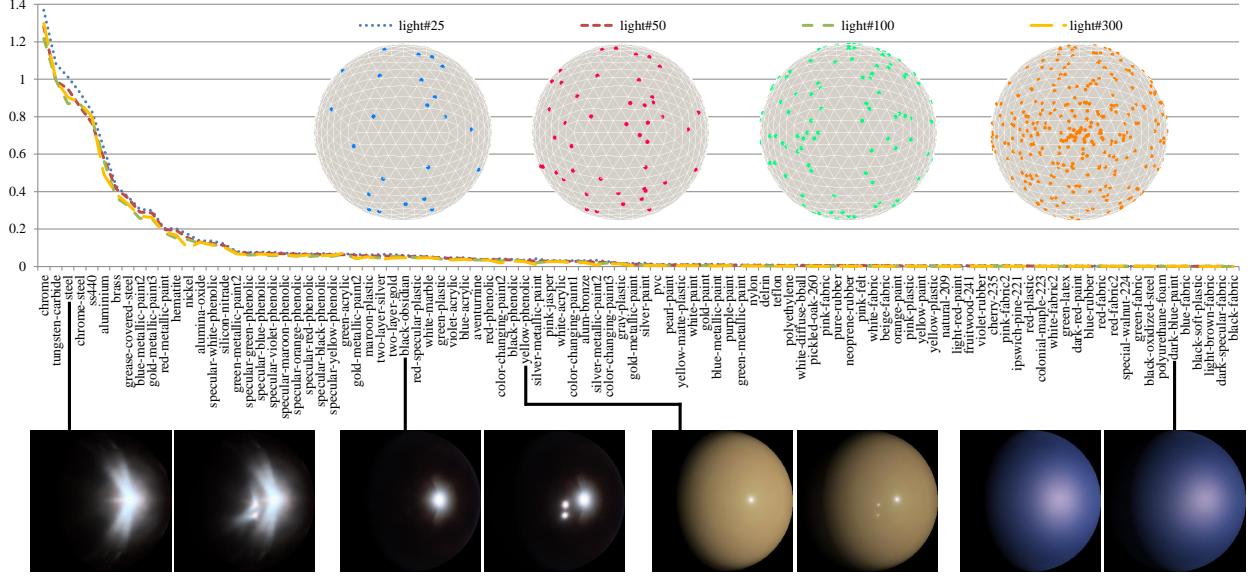
Our method for estimating elevation angles of the surface normal uses attached shadow detection. To reliably detect attached shadows from images, we use the information of interpolated lighting directions and estimated azimuth angles to compute a shadow threshold per image.

From Fig. 3, we know that only the lighting directions in the area of  $S$  can cause the attached shadow. This indicates that when a lighting direction creates the attached shadow at a pixel, the azimuth angle of the lighting direction  $\phi_l$  and that of surface normal  $\phi_n$  differ greater than or equal to  $\frac{\pi}{2}$ , i.e.,  $|\phi_l - \phi_n| \geq \frac{\pi}{2}$ . Thus, given a lighting direction  $l$ , we collect the pixels that satisfy  $|\phi_l - \phi_n| \geq \frac{\pi}{2}$  for initial identifying pixels in the attached shadows. Using the elevation angle of the light source direction  $\theta_l$ , we determine the shadow threshold as the intensity value at the top  $(1 - \frac{2\theta_l}{\pi}) \times 100\%$  percentile of all those collected pixels, with their intensity values ordered in an ascending sequence. This approach is motivated by the fact that the lower  $\theta_l$  will contribute to more shadows.

**Cast shadow removal.** Using the computed threshold, we identify shadowed pixels that include both attached and cast shadows. To eliminate the cast shadow pixels, we further use a similar geometric constraint: For the shadowed pixels satisfying  $|\phi_l - \phi_n| < \frac{\pi}{2}$ , we exclude them as cast shadow pixels.

## 6 Experiments

We use both synthetic data and real-world data to evaluate the performance of our method. We first verify the validity of our interpolation before showing the accuracy of estimated surface normals. There are almost no parameters required to be adjusted through the whole pipeline. For the azimuth angle estimation we use exactly the same setting as it was done in the original paper.<sup>3</sup> The number of clusters of  $K$ -means in elevation angle clustering is set as 150 for an over-segmentation. Because of the simplicity of our method, the computation is very efficient. In our experiments, we used an unoptimized Matlab implementation on a PC with 2.4 GHz i7 CPU and 16 GB RAM. The running time on a single core is about 18 seconds for 50 images with a resolution of  $256 \times 256$  (more than half of the time is spent on azimuth angle estimation), which is much more efficient than the optimization based solutions.<sup>5,6,21</sup> Because the number of input images for our method is significantly different from methods that require dense input<sup>20,21,24</sup> and we also need fewer images than the methods that only use partial observations satisfying their low-frequency assumptions,<sup>27,28</sup> we do not make direct comparisons with these methods. Instead, we compare our method with a standard photometric stereo based on linear least squares ( $\ell^2$ ),<sup>1</sup> a robust approach based on Robust Principal Component Analysis (R-PCA),<sup>4</sup> a recent similar work on elevation angle estimation from



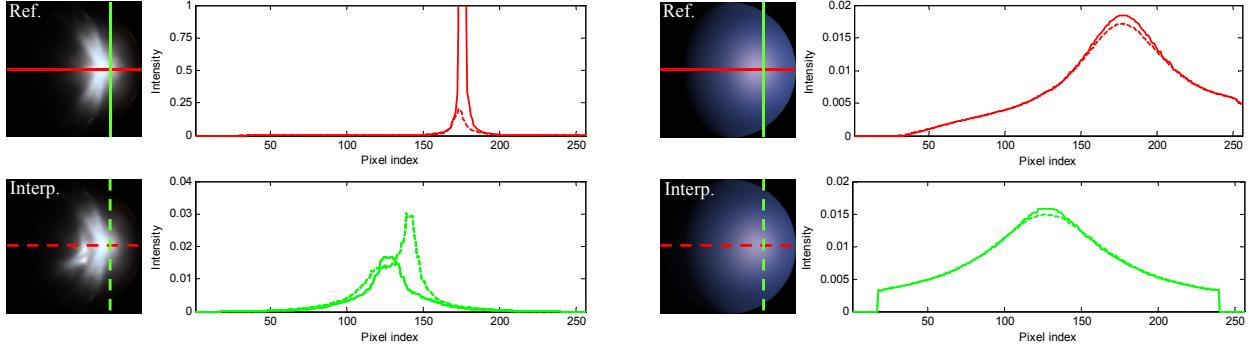
**Fig 4** Verification of our linear interpolation. The RMSE ( $Y$ -axis) between the interpolated and reference images are plotted for 100 different materials and four different lighting conditions. The  $X$ -axis shows various material names in a descending sequence ordered according to the errors of 50 random lights. Below the material names, four examples of reference (left) and interpolated (right) images are illustrated.

reflectance monotonicity,<sup>7</sup> and another two approaches that relying on BRDF fitting using multiple pixels.<sup>5,6</sup>

### 6.1 Validity of the proposed interpolation approach

The linear interpolation will only work perfectly with Lambertian materials, but the key assumption in this paper is that our linear interpolation should work reasonably well for many real-world materials. Since real world BRDFs could be represented more precisely using the measured data than using any analytic model, we use the MERL database<sup>2</sup> to verify the validity of our interpolation. The BRDFs in the database were measured from real-world scenes covering a great diversity of materials. For each BRDF, we render images under the desired lighting distribution (as shown in Fig. 1) as reference images. We then render images under  $\{25, 50, 100, 300\}$  randomly distributed lights (as shown in Fig. 4) as source images. The source images are interpolated using our method to approximate the image appearances under the desired distribution. Then the Root-Mean-Square Error (RMSE) is calculated between the interpolated and reference images for each material and lighting distribution, and the results are summarized in Fig. 4.

From the RMSE distribution and examples of reference images (left) and interpolated images (right) at the bottom of the figure, we can tell: 1) For the number of lights, using more than 50 random lights does not improve the interpolation accuracy much, but too few lights such as 25 will make the interpolation very inaccurate, so we use 50 random lights for most of the experiments in the following; 2) For the distribution of lights, it is intuitive that closer distribution to the desired one can produce better approximation, but the random distribution can also produce small enough errors for most BRDFs; 3) For material properties, more diffuse materials are more accurately interpolated (such as diffuse paint, fabric, etc.) and some materials with sparse highlights can also be reasonably interpolated (such as phenolic, metallic paint, etc.), but very shiny or dark



**Fig 5** Comparison of pixel intensities in the reference (solid line) and interpolated (dashed line) images. Two examples of one difficult (leftmost example in Fig. 4) and one easy (rightmost example in Fig. 4) materials are analyzed. The upper plot shows the pixel values of the horizontal scan line (red) and the lower plot shows the pixel values of the vertical scan line (green).

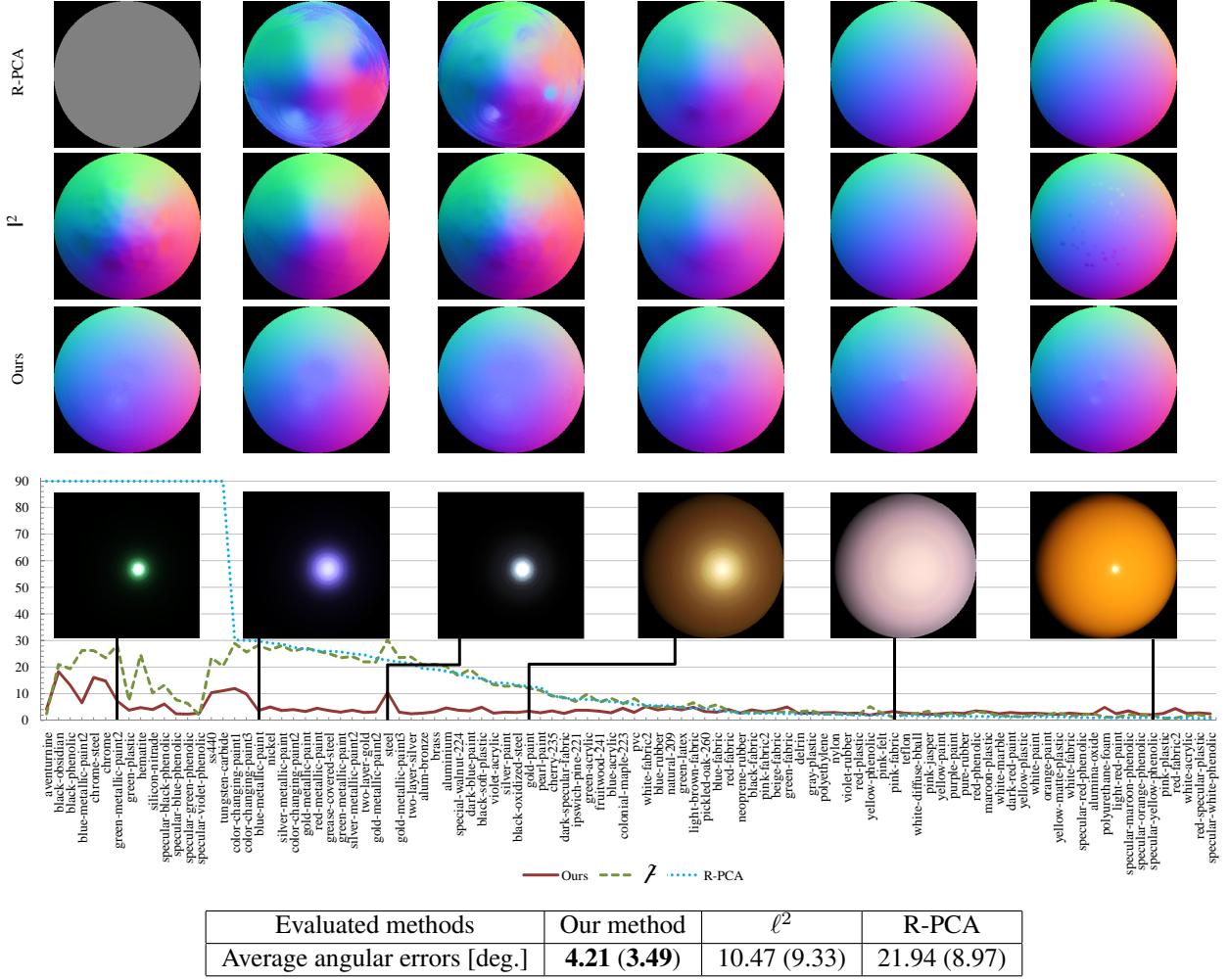
materials (such as steel, black-obsidian, etc.) show strong nonlinearity so that our linear interpolation is not accurate enough. Fortunately, the challenging materials only take less than 30% among all the evaluated BRDFs, therefore our interpolation works sufficiently well for many real-world materials.

How the proposed interpolation meets the requests of the proposed photometric stereo pipeline could be understood by a close check at the pixel values of the reference and interpolated images, as plotted in Fig. 5. The observations with lower intensities are usually close to Lambertian reflectance, and they are densely observed on most objects. Therefore, the interpolation performs better for pixels with lower intensities, which provides accurately interpolated pixels in shadows so that the elevation angle estimation could work. Though the interpolated values at higher intensities are approximated with lower quality, the symmetric properties of isotropic BRDFs are well maintained so that azimuth angle estimation could also work. Since both elevation angle and azimuth angle estimation rely on densely observed values instead of sparse ones and our interpolation works well for majority of pixels except for some sparse highlight, the accuracy of normal estimation is guaranteed using interpolated images. An intuitive example is when the diffuse reflectance is purely Lambertian while the specular term is sparse (e.g., described by the Ward lobe), the interpolation will have little impact on the normal estimation even if such a non-Lambertian BRDF violates the linear interpolation assumption.

## 6.2 Quantitative evaluation using synthetic data

We again use the MERL database<sup>2</sup> to create synthetic images for quantitatively evaluating the normal estimation accuracy of our method.

**Performance varying with materials.** We evaluate the performance of the proposed method using all the BRDFs contained in the MERL database in comparison with the  $\ell^2$ -based<sup>1</sup> and R-PCA<sup>4</sup> methods. The results are summarized in Fig. 6. Our method performs consistently well even for materials that are particularly difficult with the  $\ell^2$  and R-PCA approaches. While the R-PCA approach effectively ignores specular effects, the diffuse component in the database is far from the Lambertian model, and this fact results in large errors in both the  $\ell^2$  and R-PCA methods. By comparing Fig. 4 and Fig. 6, it is easy to tell that the normal estimation accuracy by our method

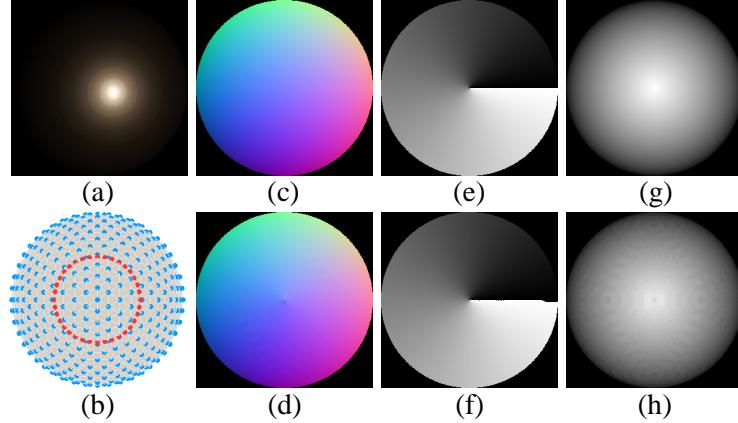


**Fig 6** Performance evaluation of our method on a synthetic sphere scene (50 images under a random lighting direction) using all 100 materials. The normal map encodes the XYZ components of surface normal in RGB channels respectively. The X-axis shows various material names in a descending sequence ordered according to the errors of R-PCA, and the Y-axis is the angular error of the surface normal in degrees. The table shows the average angular error of each method. The values in () show the average errors after removing BRDFs with which R-PCA return trivial results.

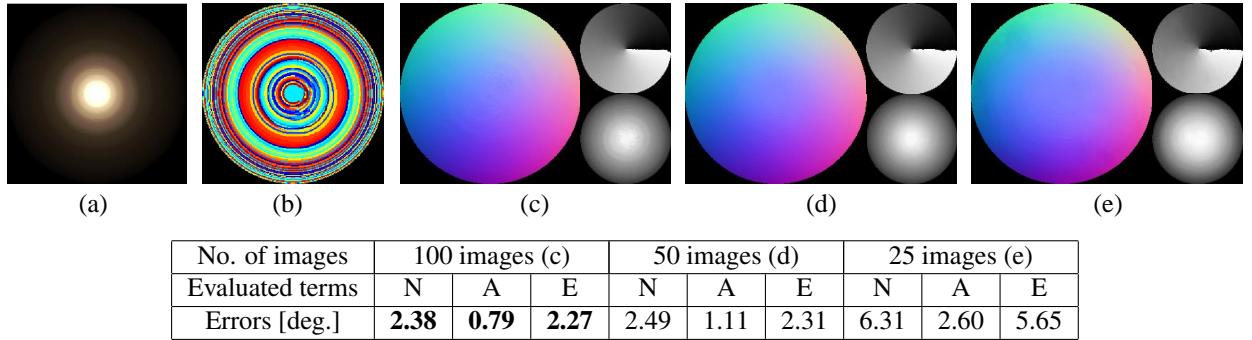
is highly correlated with the interpolation accuracy: Those challenging materials for interpolation, e.g., “chrome”, “tungsten-carbide”, “ss440”, etc., are also difficult for normal estimation.

**Performance of the interpolation-based approach.** Before showing the performance with interpolation, we show the ideal performance *without* interpolation, where the input lighting distribution is exactly the desired “ring + uniform” lighting distribution. The result using the “alum-bronze” BRDF in the MERL database with a sphere model is shown in Fig. 7. The angular errors of the estimated Normal/Azimuth angle/Elevation angle (N/A/E, hereafter) are 2.20/0.26/2.18 degrees, respectively. These are considered as the lower bounds of the error.

Then we test our method using the same BRDF and scene, but with a varying number of random light sources to verify the performance of our method with interpolation. The result is summarized in Fig. 8. When using 100 images, the angular errors in degrees for N/A/E are 2.38/0.79/2.27, while



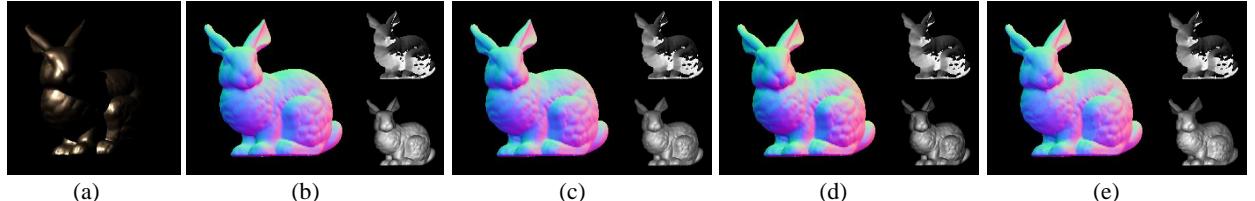
**Fig 7** Surface normal, azimuth and elevation angle estimation under an ideal lighting distribution without interpolation. (a) an input image; (b) lighting configuration (red: ring; blue: uniform); (c), (e), (g) the ground truth N/A/E (they also serve as the reference color maps for normal, azimuth, and elevation angle maps in all the tests in this paper); (d), (f), (h) the estimated N/A/E with angular errors of 2.20/0.26/2.18 degrees.



**Fig 8** Performance variation of our method with respect to the number of random lighting directions. (a) input image; (b) elevation angle clustering in the 50 image case; (c), (d), (e) the estimated N/A/E using 100, 50, and 25 images, respectively.

it slightly increases to 2.49/1.11/2.31 for 50 images but is still as small as the result with an ideal lighting distribution (Fig. 7). When the number of images becomes too few, the error increases as shown by the case of 25 images due to insufficient lighting directions for interpolation, which is consistent with the verification in Fig. 4. We also show an image of elevation clusters next to the image illuminated from the viewing direction to show the correctness of elevation angle clustering in Fig. 8(b) and (a).

**Effectiveness of additional constraints.** We verify the effectiveness of additional constraints that are used in our method. One is cast shadow removal discussed in Sec. 5, and the other one is elevation angle refinement by clustering described in Sec. 4.2. We use a bunny model with the “alum-bronze” BRDF rendered under 50 random lighting directions. The result in Fig. 9 shows that both cast shadow removal and elevation angle clustering are effective in reducing the error of elevation angle estimates.



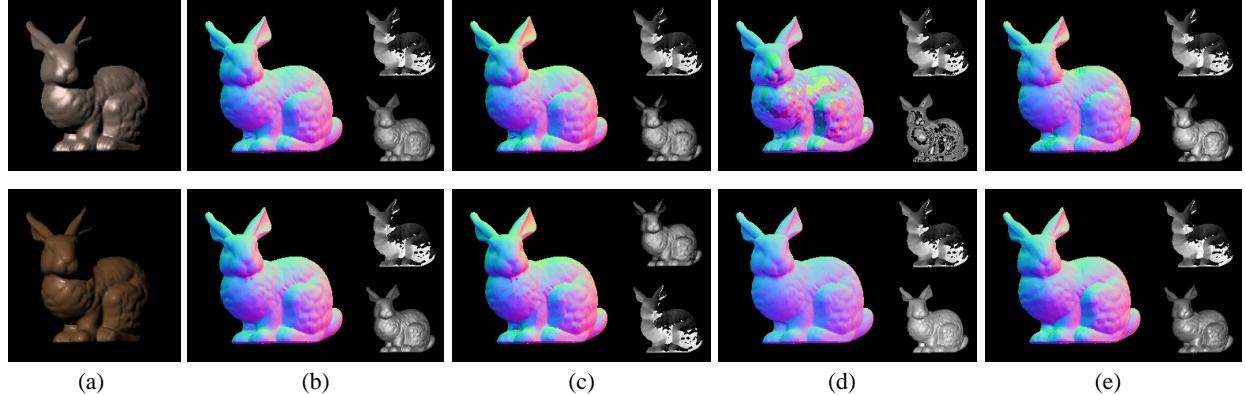
**Fig 9** Effectiveness of cast shadow removal and elevation angle clustering. (a) input image; (b) the ground truth N/A/E; (c) estimated N/A/E using our complete method; (d) result without cast shadow removal; (e) result without refining elevation angles by clustering.

**Comparison with recent methods.** We make a comparison with recent methods dealing with general BRDFs in Fig. 10. We first compare with Shi et al.<sup>7</sup> and Alldrin et al.<sup>5</sup> since both methods estimate elevation angles given azimuth angles but the former works in a per-pixel manner and the latter considers the multi-pixel dependency for BRDF fitting. We use synthetic bunny models with the “silver-paint” and “specular-yellow-phenolic” BRDFs \* rendered under 50 random lighting directions as examples. One of the input images for these two datasets are shown in Fig. 10(a), and the ground truth normal has been shown in Fig. 9(b). The estimated N/A/E maps and the average angular errors show that our method outperforms both Shi et al.<sup>7</sup> and Alldrin et al.<sup>5</sup> for these examples in terms of elevation angle estimation accuracy. The method by Alldrin et al.<sup>5</sup> performs worst partially due to the fragile fitting to general BRDFs with inaccurate normals. We also test other BRDFs in the MERL database by comparing with Shi et al.<sup>7</sup> For those materials that exactly follow the one-dimensional and monotonic BRDF assumptions, such as some acrylic and metallic paint materials (please refer to their papers for materials that are easy or difficult for their methods), the method by Shi et al.<sup>7</sup> shows better accuracy than ours, but for multi-lobe BRDFs that violate their assumptions, such as some fabric and diffuse paint materials, our method is more accurate. We further make a comparison with the method by Goldman et al.<sup>6</sup> Their method estimates the surface normal as a whole by fitting several Ward lobes to address the non-Lambertian effect. When the Ward lobes well explained the input BRDF (e.g., the “silver-paint” BRDF), their method shows higher accuracy on the estimated normals (due to more accurate elevation angles), otherwise our method produces better results. Due to the alternating BRDF fitting and normal estimation that involves highly nonlinear optimization, the methods by Alldrin et al.<sup>5</sup> and Goldman et al.<sup>6</sup> take much longer computation time (10 – 20 hours) than our method (18 second) and the method by Shi et al.<sup>7</sup> (37 seconds). But due to the closed-form solution of elevation angles, our method is theoretically simpler and computationally more efficient than Shi et al.<sup>7</sup>

### 6.3 Qualitative evaluation using real-world data

We use six datasets that contain various materials like plastic, fruit skin, porcelain and metallic paint to qualitatively evaluate our method. These datasets are named “gourd”, “apple”, “orange”, “postbox”, “mask”, and “pear”. All of them contain 50 images under randomly distributed lights.

\*We choose these two materials because they rank in the middle (not very easy but also not the most challenging) for interpolation difficulty according to Fig. 4 to make a fair comparison.



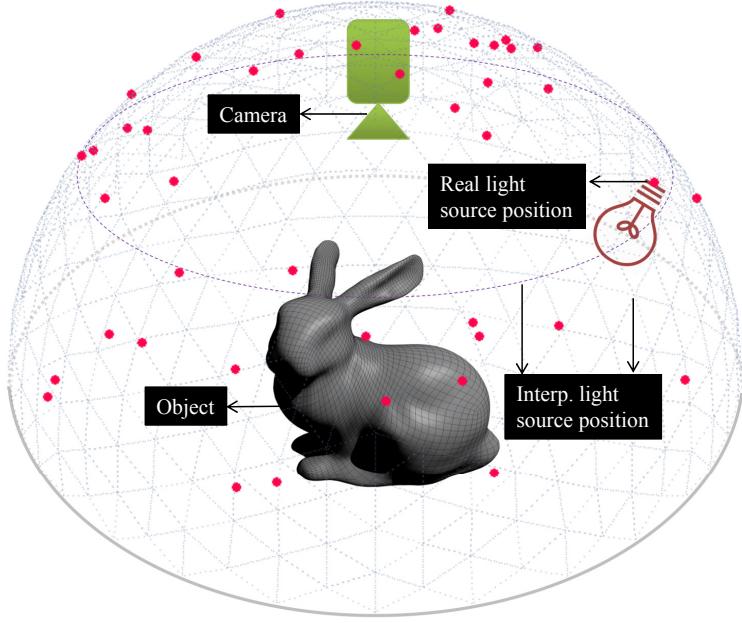
**Fig 10** Comparison with recent methods. The first row is the results from “silver-paint” BRDF and the second row is the result from “specular-yellow-phenolic” BRDF. (a) input images; (b), (c), (d), and (e) are estimated N/A/E using our method, the methods from Shi et al.,<sup>7</sup> Alldrin et al.,<sup>5</sup> and Goldman et al.<sup>6</sup> respectively.

The first two datasets are provided by Alldrin et al.,<sup>5</sup> and we randomly select 50 images from their original datasets that contain more than 100 images. Other datasets are captured using a Sony XCD-X710CR camera with a linear radiometric response. We place the camera and light source far enough from the objects to approximate the orthographic projection and distant lighting assumption, and the lighting directions are calibrated using a mirror sphere. Our data capture setup is illustrated in Fig. 11. In Fig. 12 we show the appearances of all data and the results of our method in comparison with the R-PCA method.<sup>4</sup> We have also evaluated the method by Shi et al.<sup>7</sup> using these data, but the results visually look quite similar to ours, so we omit the visualization of such results.

To visually verify the accuracy of the estimated surface normals, we use the shapelet-based surface reconstruction method<sup>36</sup> to build a surface from normal. The parameters for the reconstruction using the shapelet are fixed for all the tests as (6, 1, 3) for (shapelet scales,  $\sigma$  of smallest scale Gaussian shapelet, scaling factor between successive shapelets). In the second column (next to examples of input images), we show the rendered results from Alldrin et al.<sup>5</sup> or real objects from a different viewpoint as references for the estimated surfaces. The recovered surfaces are aligned to the reference views for making qualitative evaluation, and our reconstruction agrees well with them. Here we integrate the 3D surfaces solely from estimated normals for illustration purpose, and the surfaces could possibly be improved for depth-discontinuous regions by combining with laser sectioning as suggested by McGunnigle et al.<sup>37</sup>

## 7 Conclusion

We proposed a practical photometric stereo method for general isotropic BRDFs. To achieve the goal without relying on the special lighting configuration and dense input images, we make use of the spherical linear interpolation to approximate desired appearances that are used for the estimation of the azimuth and elevation angles per pixel. The newly proposed elevation angle estimation

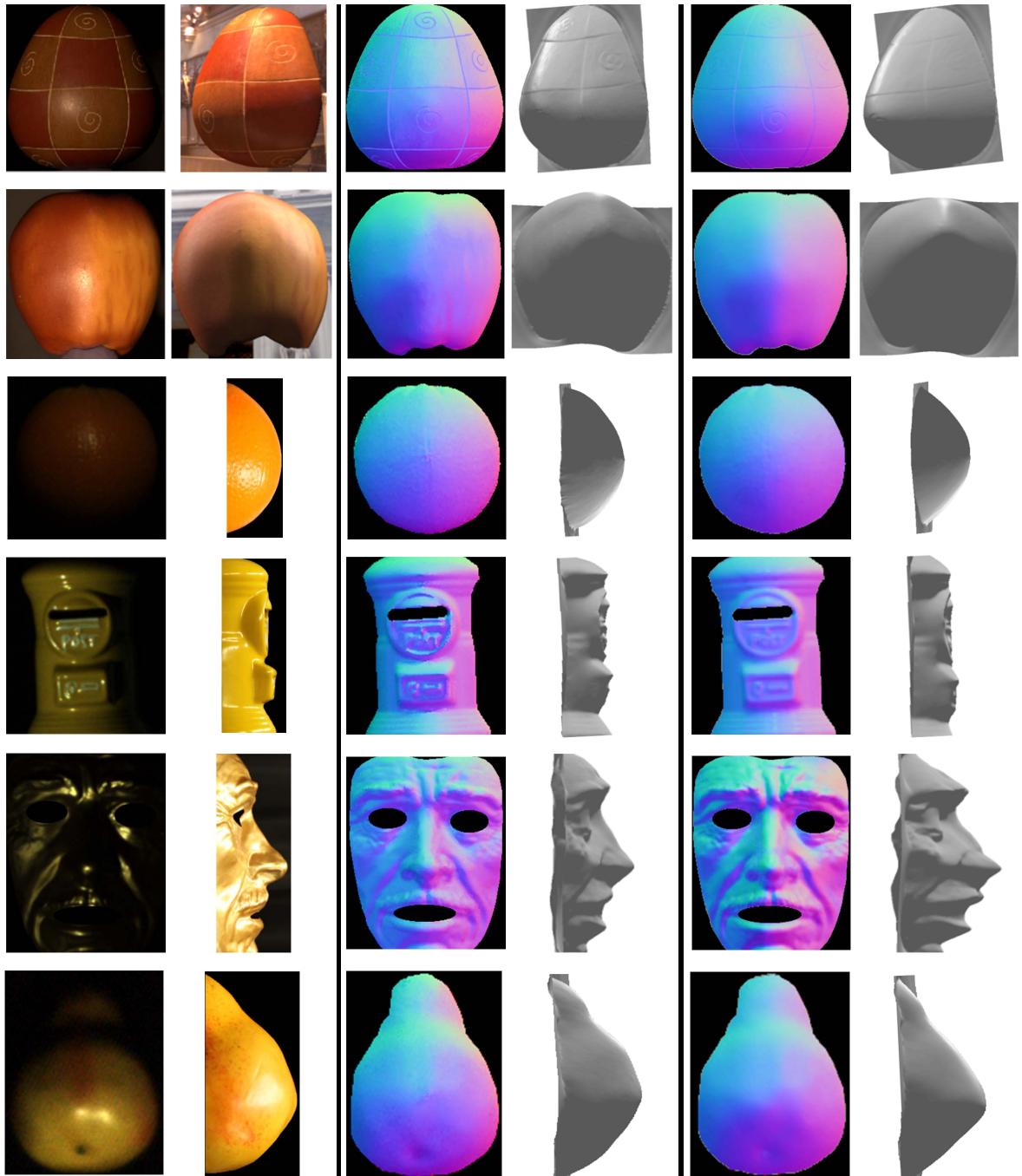


**Fig 11** Data capture setup of our method. The camera is located at  $(0, 0, 1)$ , and the lights are randomly distributed on the visible hemisphere. The input images (captured when lights are located at the red dot, “Real light source position”) are interpolated as if they were illuminated from the “Interp. light source position” (the vertices of triangles and the ring position).

method is a closed-form solution and various considerations to attached and cast shadows have been taken in the proposed framework. We showed that the proposed method is able to reliably estimate surface normals in comparison with traditional and recent robust approaches. We tested the proposed method using the MERL database that contains 100 measured BRDFs and various real-world objects with different materials. The results showed that our method consistently works well with various BRDFs and outperforms many recent methods.

**Limitations.** Strictly speaking, our method is limited to work with a scene of a uniform material. The accuracy of interpolation on a shadowed area and determination of the shadow threshold suffers from spatially varying BRDFs, especially when their BRDFs are severely different (e.g., very high contrast). However, as shown in the real-world results, our method works reasonably well for the objects with a dominant single material and some gradually changing textures (e.g., “pear” data).

The main error of surface normal estimation with our method appears around the north pole of the sphere, where elevation angles are large. This is due to the fact that the shadow ratio becomes tiny when the elevation angle is large, and it results in numerical instability. So we have to rely on some multi-pixel dependency to refine the estimated elevation angles and this issue cannot be relieved by using more input images. We will investigate a robust approach for dealing with this problem by, for example, using an MRF-based approach for regions where the shadow ratio becomes very small. Due to the limited power of linear interpolation, materials that dominantly contain specularities are also beyond the capability of the proposed method.



**Fig 12** Result of our method using real-world data. The left column shows one example of input images and reference pictures for shape verification; the middle column shows the estimated normals and corresponding surfaces of our method; and the right column shows the results from the R-PCA method.<sup>4</sup>

#### Acknowledgments

The authors would like to thank Prof. Yasuyuki Matsushita and Prof. Katsushi Ikeuchi for insightful discussions in early stage of this work, and Dr. Neil Alldrin and Dr. Dan B. Goldman for sharing their codes for comparison.

## References

- 1 R. Woodham, “Photometric method for determining surface orientation from multiple images,” *Optical Engineering* **19**(1), 139–144 (1980).
- 2 W. Matusik, H. Pfister, M. Brand, and L. McMillan, “A data-driven reflectance model,” *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* **22**(3), 759–769 (2003).
- 3 N. Alldrin and D. Kriegman, “Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach,” in *Proceedings of International Conference on Computer Vision*, (2007).
- 4 L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, “Robust photometric stereo via low-rank matrix completion and recovery,” in *Proceedings of Asian Conference on Computer Visoin*, (2010).
- 5 N. Alldrin, T. Zickler, and D. Kriegman, “Photometric stereo with non-parametric and spatially-varying reflectance,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2008).
- 6 D. Goldman, B. Curless, A. Hertzmann, and S. Seitz, “Shape and spatially-varying BRDFs from photometric stereo,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(6), 1060–1071 (2010).
- 7 B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, “Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances,” in *Proceedings of European Conference on Computer Vision*, (2012).
- 8 Y. Sato and K. Ikeuchi, “Temporal-color space analysis of reflection,” *Journal of the Optical Society of America A* **11**(11), 2990–3002 (1994).
- 9 P. Hansson and P. Johansson, “Topography and reflectance analysis of paper surfaces using a photometric stereo method,” *Optical Engineering* **39**(9), 2555–2561 (2000).
- 10 S. Barsky and M. Petrou, “The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**(10), 1239–1252 (2003).
- 11 A. Georghiades, “Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo,” in *Proceedings of International Conference on Computer Vision*, (2003).
- 12 H. Chung and J. Jia, “Efficient photometric stereo on glossy surfaces with wide specular lobes,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2008).
- 13 C. Hernandez, G. Vogiatzis, and R. Cipolla, “Multi-view photometric stereo,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(3), 548–554 (2008).
- 14 D. Miyazaki, K. Hara, and K. Ikeuchi, “Median photometric stereo as applied to the segonko tumulus and museum objects,” *International Journal of Computer Vision* **86**(2), 229–242 (2010).
- 15 C. Yu, Y. Seo, and S. Lee, “Photometric stereo from maximum feasible Lambertian reflections,” in *Proceedings of European Conference on Computer Vision*, (2010).
- 16 S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, “Photometric stereo using sparse Bayesian regression for general diffuse surfaces,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36**(9), 1816–1831 (2014).

- 17 T. Han, Y. Cheng, H. Shen, and X. Du, “Robust photometric stereo using structural light sources,” *Journal of Electronic Imaging* **23**(3), 033004 (2014).
- 18 K. Tang, C. Tang, and T. Wong, “Dense photometric stereo using tensorial belief propagation,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2005).
- 19 T. Wu and C. Tang, “Dense photometric stereo using a mirror sphere and graph cut,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2005).
- 20 T. Wu and C. Tang, “Photometric stereo via expectation maximization,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(3), 546–560 (2010).
- 21 M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler, “A photometric approach for estimating normals and tangents,” *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* **27**(5), 133 (2008).
- 22 T. Higo, Y. Matsushita, and K. Ikeuchi, “Consensus photometric stereo,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2010).
- 23 I. Sato, T. Okabe, Q. Yu, and Y. Sato, “Shape reconstruction based on similarity in radiance changes under varying illumination,” in *Proceedings of International Conference on Computer Vision*, (2007).
- 24 T. Okabe, I. Sato, and Y. Sato, “Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions,” in *Proceedings of International Conference on Computer Vision*, (2009).
- 25 M. Chandraker, J. Bai, and R. Ramamoorthi, “On differential photometric reconstruction for unknown, isotropic BRDFs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**(12), 2941–2955 (2013).
- 26 F. Romeiro, Y. Vasilyev, and T. Zickler, “Passive reflectometry,” in *Proceedings of European Conference on Computer Vision*, (2008).
- 27 B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, “Bi-polynomial modeling of low-frequency reflectances,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36**(6), 1078–1091 (2014).
- 28 S. Ikehata and K. Aizawa, “Photometric stereo using constrained bivariate regression for general isotropic surfaces,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2014).
- 29 F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato, “Uncalibrated photometric stereo for unknown isotropic reflectances,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2013).
- 30 Z. Wu and P. Tan, “Calibrating photometric stereo by holistic reflectance symmetry analysis,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2013).
- 31 K. Shoemake, “Animating rotation with quaternion curves,” in *Proceedings of SIGGRAPH*, (1985).
- 32 J. Shewchuk, “Triangle: engineering a 2D quality mesh generator and Delaunay triangulator,” *Applied Computational Geometry: Towards Geometric Engineering* **1148**, 203–222 (1996).
- 33 H. Coxeter, “Barycentric coordinates,” *Introduction to Geometry*, 2nd ed. New York: Wiley **13**(7), 216–221 (1969).

- 34 P. Tan and T. Zickler, "A projective framework for radiometric image analysis," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2009).
- 35 B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan, "Self-calibrating photometric stereo," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (2010).
- 36 P. Kovesi, "Shapelets correlated with surface normals produce surfaces," in *Proceedings of International Conference on Computer Vision*, (2005).
- 37 G. McGunnigle, J. Dong, L. Su, and Y. Fang, "Topography and reflectance analysis of paper surfaces using a photometric stereo method," *Optical Engineering* **50**(5), 053605 (2011).

**Si Li** received the B.E. and Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2007 and 2012, respectively. She is currently a Postdoctoral Fellow at Department of Computer Science, Brandeis University, Waltham, Massachusetts, USA. Her research interests include multimedia information analysis and natural language processing.

**Boxin Shi** received the B.E. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2007, and the M.E. degree from Peking University, Beijing, China, in 2010, and Ph.D. degree from the University of Tokyo, Tokyo, Japan, in 2013. He is currently an SUTD-MIT Joint Postdoctoral Fellow with Singapore University of Technology & Design and Massachusetts Institute of Technology. His research interests include computer vision and computational photography.