Real-Time 3D Shape of Micro-Details

Maryam Khanian\textsuperscript{1}\textsuperscript{(a)}
\begin{quote}
khaniimar@b-tu.de
\end{quote}

Ali Sharifi Boroujerdi\textsuperscript{1}
\begin{quote}
shariali@b-tu.de
\end{quote}

Michael Breuß\textsuperscript{1}
\begin{quote}
michael.breuss@b-tu.de
\end{quote}

\textsuperscript{1}Chair of Applied Mathematics, Brandenburg University of Technology (BTU), 03046 Cottbus, Germany

Abstract

Motivated by the growing demand for interactive environments, we propose an accurate real-time 3D shape reconstruction technique. To provide a reliable 3D reconstruction which is still a challenging task when dealing with real-world applications, we integrate several components including (i) Photometric Stereo (PS), (ii) perspective Cook-Torrance reflectance model that enables PS to deal with a broad range of possible real-world object reflections, (iii) realistic lightening situation, (iv) a Recurrent Optimization Network (RON) and finally (v) heuristic Dijkstra Gaussian Mean Curvature (DGMC) initialization approach. We demonstrate the potential benefits of our hybrid model by providing 3D shape with highly-detailed information from micro-prints for the first time. All real-world images are taken by a mobile phone camera under a simple setup as a consumer-level equipment. In addition, complementary synthetic experiments confirm the beneficial properties of our novel method and its superiority over the state-of-the-art approaches.

1. Introduction

3D reconstruction can be considered as one of the major areas for computer vision which provides not only 3D interpretation, but also the surface feature extraction. Photometric Stereo (PS) is a potential technique for this purpose. In PS, several input images taken from a fixed viewpoint under different illumination directions are used for 3D reconstruction. The pioneers of the problem are Woodham [66] and Horn et al. [21]. As it is shown by Woodham [67], the orientation of a Lambertian surface can be uniquely determined when at least three input images are given. During last decades, PS has attracted wide attention in different practical domains such as industrial quality control [11, 74, 42, 25], face recognition [72, 75, 14, 73], medical science [44, 63], texture classification [51, 57, 34] and estimation of weather conditions [55]. Another benefit of PS is to generate invariant features utilized in different kinds of recognition tasks [61, 58]. However, most of the later works on PS kept two basic building blocks of Woodham’s approach, namely Lambertian surface reflection and orthographic camera projection. For this reason, Lambert’s law [29] is commonly used in classic computer vision models, which implies that the surface reflects light equally in all directions and ignores the specular component. However, it may easily fails for surfaces whose reflection is non-Lambertian. Corresponding effects (e.g. specularities as occurring on shiny objects) are quite readily observable in real-world images and surfaces with not pure isotropic luminance. So, two major light reflection components in-
cluding specular reflection as the high frequency variations and diffuse reflection as the low frequency component should be combined to provide a successful reflectance model \cite{47, 40, 65}.

In order to circumvent difficulties of specularities, commercial instruments for 3D reconstruction use sophisticated techniques such as white light interferometry or scanning focal microscopy, while these laboratory-based devices tend to be large, slow, or at the expense of $100,000 or more. In many works, non-Lambertian effects are considered as outliers. In this context, Miyazaki et al. \cite{38} suggested a median filtering to evade the influence of specular reflection and a preprocessing technique is presented by Yang et al. \cite{70} to remove specularities. A maximum-likelihood estimation \cite{64} and an expectation maximization \cite{69} are other approaches based on assuming specularities as the outliers. While Mukai gawa et al. \cite{39} suggested a random sample consensus based approach in which only diffuse reflection is employed as the underlying model, Tang et al. \cite{60} proposed a coupled Markov random field treating specularities as noise. An approach relying on a maximum feasible subsystem method is introduced by Yu et al. \cite{72}, where input images satisfying a Lambertian constraint are selected for reconstruction. In addition, a convex optimization technique is developed by Wu et al. \cite{68} in which specularities are considered as deviations from the Lambertian assumption in the objective function. Zickler et al. \cite{23} provides an analyze to a subspace which is free from highlights for image transformation. Finally, Ikehata et al. \cite{23} suggested a regression procedure that deals with the specular component as a sparse error of an underlying purely diffuse reflectance equation. Mallick et al. \cite{32} considered a transformation of the RGB color space to avoid specularity using the light source color which is not always available. An inpainting technique for the highlights correction is presented by Tan et al. \cite{49} using the color of illuminant and the diffuse component. In \cite{59} and \cite{54} the pixel chromaticity direct analysis and specular-free image are used respectively to specularity correction in which the diffuse component is adjusted to a criterion. This criterion is an indicative of smooth transition of color between diffuse and specular regions and evaluated using a nonlinear shifting of specular pixels intensity and chromaticity. Johnson et al. \cite{24} presented an elastomer using a sensor skin constructed with a metal-flake pigment which is attached to the surface to change its BRDF. However, in all these works, the shape information that is evidently contained in non-Lambertian effects (e.g. highlights) is discarded. The orthographic projection as the second mentioned ingredient is realistic when objects are far away from the camera, but not if they are close since then perspective effects grow to be important. In addition, when a surface is projected orthographically, it can be reconstructed only up to a generalized bas-relief transformation. However, the perspective projection as proved in \cite{43}, can remove the bas-relief ambiguity and provide the uniqueness of the reconstructed surface. There is only a small number of previous papers dealing with perspective models in PS. To our knowledge, the first of such models has been developed by Lee and Kuo \cite{30}. They approximated the depth function as a linear combination of individual element basis functions. Other perspective methods are applied by Galo and Tozzi \cite{13}, Tankus and Kiryati \cite{62} and Mecca et al. \cite{37}, while the Lambertian reflectance is applied in all these researches. Concerning more general non Lambertian reflectance models based on perspective PS, Khanian et al. proposed the application of perspective Blinn-Phong model \cite{5, 46} for the signal reconstruction and corresponding lightening conditions in \cite{26} and for the 3D reconstruction with different perspective techniques in \cite{27}. An empirical approach is also considered in \cite{36, 35}. However, for the derivation of the final applied model, purely diffuse and purely specular reflections should be decoupled in \cite{36, 35}. Moreover, the proposed approach in \cite{36} relies on the Fast Marching method \cite{52} where the unknown depth value of the center point should be provided in advance which makes difficulty for real-world applications. Some state-of-the-art in PS methods in which the information of depth should be provided in advance, can be pointed as \cite{7, 50} which require the depth to be known on the boundary, and \cite{35} that needs the mean value of depth. In contrast, we do not need any information of depth to be given in advance in our approach as also proved in Figure 1.

1.1. Our contributions

In this paper, we provide the following innovations:

1. We proposed an accurate hybrid real-time 3D reconstruction approach which merges the benefits of perspective projection and non-Lambertian reflectance model motivated by Cook-Torrance reflectance \cite{8} to handle complex reflecting characteristics. To the best of our knowledge, this is the first work for the real-time PS which also regarded the application of perspective Cook-Torrance model.

2. Introducing a Recurrent Optimization Network (RON) which provides a robust PS framework to obtain a highly-detailed 3D shape from even micro-print texts where their visibility is hard for the human visual system as can be seen in Figure 1. In this Figure, we could reconstruct 3D shape of two regions of A and B with high details, while these regions are very tiny microtexts. The information hidden in these areas became completely clear by our approach. As a result, our technique can be used in different applied areas.

3. Since the performance of many optimization techniques is highly-dependent on the initialization process, we provided a Dijkstra Gaussian Mean Curvature (DGMC) technique to find anchor points which can be applied as a
key point to offer more effective optimization approach.

4. Furthermore, we extended our proposed approach to deal with realistic lightening situation using spatially dependent lighting.

5. We show that our approach is able to provide the real-time 3D shape of objects as captured by an ordinary mobile phone camera without employing controlled laboratory conditions (cf. Figure 2). So, our PS process is conducted by the consumer-level equipment where specialized equipment and professional laboratory settings are not required. Our instruments are easily available for general consumers as should be expected for an ideal solution. Consequently, our method can be readily employed in real-world situations. We demonstrate that this paper provides the basis for a potentially useful 3D reconstruction technique. This improvement breaks inhibitor limitations of a controlled setup and the necessity of working with specific scenes (e.g. Lambertian materials) and turns mobile phone into powerful 3D shape reconstruction tools.

6. Finally, our approach also provides non-uniform colorful albedo from images with diverse color intensities as shown in Figure 12. All these attempts greatly advance the applicability of photometric stereo to various applications, especially for reconstructing 3D surfaces from very tiny structures like microprints.

2. Cook-Torrance reflectance model

The Cook-Torrance reflectance as a physically based model designed to account for the spectral composition of specularity and cope with the problem of modeling reflections of realistic materials [8]. This physically based model agrees well with actual measured BRDFs [28]. Like recent popular theoretical models, this reflectance model considers this assumption that the large scale BRDF is a collection of microfacets as shiny V-grooves with random sizes and orientations [28]. This hypothesis leads to the microfacet theory for non-perfectly-smooth surfaces. These surfaces are essentially height fields, where the distribution of facet normals. This property enables them to effectively model many real-world surfaces [45] in which, mirror-like facets are modeled by the Cook-Torrance reflectance to retrieve the local shape from specularity. The Cook-Torrance image irradiance equation is defined as follows:

\[
I(x, y) = k_d \frac{L \cdot N(x, y)}{\|L\| \|N(x, y)\|} I_d + \frac{G(x, y, z)D(x, y, z)F(x, y, z)}{4 \|V(x, y, z)\| \|N(x, y)\| \|L\| \|N(x, y)\|} \tag{1}
\]

where

\[
G(x, y, z) = \min \left\{ 1, T(x, y, z), R(x, y, z) \right\} \tag{2}
\]

\[
T(x, y, z) = \frac{2H(x, y, z) \cdot N(x, y)}{\|H(x, y, z)\| \|N(x, y)\|} \tag{3}
\]

\[
R(x, y, z) = \frac{2H(x, y, z) \cdot N(x, y)}{\|H(x, y, z)\| \|N(x, y)\|} \tag{4}
\]

\[
Q(x, y, z) = \frac{V(x, y, z) \cdot H(x, y, z)}{\|V(x, y, z)\| \|H(x, y, z)\|} \tag{5}
\]

\[
D(x, y, z) = \frac{1}{\pi m^2 \cos^4(\alpha)} \exp \left( -\frac{\tan^2(\alpha)}{m^2} \right) \tag{6}
\]

\[
\alpha = \arccos \left( \frac{H(x, y, z) \cdot N(x, y)}{\|H(x, y, z)\| \|N(x, y)\|} \right) \tag{7}
\]

and

\[
F(x, y, z) = f_\lambda \left[ 1 - \frac{H(x, y, z) \cdot V(x, y, z)}{\|H(x, y, z)\| \|V(x, y, z)\|} \right] \tag{8}
\]
In this notation, $I(x, y)$ and $N(x, y)$ are the intensity and surface normal at pixel $(x, y)$, respectively. The parameter $k_d$ is the diffuse material parameter (albedo) and $m$ is the root mean square slope of the surface microfacets (i.e. the surface roughness).

It is worth to mention that in expression (8), $H$ is used instead of $N$ as suggested in [18]. Let us emphasize that as opposed to an orthogonal viewer direction $(0, 0, 1)^\top$ which is commonly used, we model here a pinhole camera. Therefore, the viewing vector $V$ depends explicitly on $(x, y)$ and also on the depth $z$. This model is powerful enough to capture the geometrical properties of the scene as recorded by the camera. Moreover, its images resemble what would be seen by human visual system. Hence, pinhole camera is also referred to as perspective camera model in graphics and vision literatures [71].

The vector $H$ is defined to be the halfway vector between the viewing vector $V$ and the light vector $L$, as the direction to which the microfacet normals need to be oriented to reflect $L$ into $V$. The light source intensity is denoted by $l_d$. $F$ is the Fresnel term including $f_\lambda$ which is the reflection coefficient (the reflectance of normal incidence). Here, the distribution of facets is described by $D$. It should be noticed that in PS, it is necessary to apply a set of non-coplanar light vectors $L_i$ (corresponding to the input images $I_i$) as also investigated in [67, 20, 19, 9].

It is worthwhile noting that in all our experiments (including synthetic and real-world scenes and even for specular images), we will still follow Woodham [66] and use only three input images $I_i, i = 1, 2, 3$. We consider it an indication of the robustness and potential practical usefulness of our method that it works effectively with such a low number of input images. Considering (1) for multiple images leads to a highly non-linear system of equations for the unknown surface normal vectors that should be solved.

3. Perspective projection and intrinsic parameters of pinhole camera model

Under the perspective projection model, the surface parametrization is performed as follows:

$$S(x, y) = \begin{bmatrix} \frac{-xz(x, y)}{f} \\ \frac{-yz(x, y)}{f} \\ z(x, y) \end{bmatrix}$$

(9)

The normal vector as the cross product of the partial derivatives of the surface is computed as:

$$N(x, y) = \begin{bmatrix} \frac{z(x, y)}{f} \nabla z(x, y) \\ z(x, y) \nabla z(x, y) \cdot (x, y) + z(x, y) \end{bmatrix}$$

(10)

where $f$ is the focal length and $\nabla z(x, y) = (z_x, z_y)$ is the gradient field. On the other hand, to implement the perspective projection of the pinhole camera model, the intrinsic camera parameters should be involved which leads to the CCD camera [17]. To this aim, we apply the following transformation to convert the homogeneous representations pixel coordinate $X = (x, y, 1)^\top$ to image coordinate $\chi = (\mu, \nu, f)$:

$$X = \Gamma \left( \frac{1}{f} \right) \chi,$$

(11)

where $\Gamma$ is the matrix of intrinsic parameters of the camera expressed as:

$$\Gamma = \begin{bmatrix} \psi_x & \xi & \delta_x \\ 0 & \psi_y & \delta_y \\ 0 & 0 & 1 \end{bmatrix}$$

(12)

Here, $\psi_x$ and $\psi_y$ are the focal length in x and y directions. $\xi$ describes the skew effect and $(\delta_x, \delta_y)$ denotes the principal point or focal point. We neglect the skew parameter ($\xi = 0$) as it is zero for most of normal cameras [17].

4. Lightening sensitivity analysis

Different types of light sources can be defined based on the energy emission from a single point in space with some angularly varying distribution of outgoing lights. At first, we considered the distant light, also known as the directional light or sparse light. This source of energy is an emitter which deposits illumination from the same direction at every point in space. It can be considered also as a light at infinity (e.g. sun). Furthermore, in this kind of light sources, energy strength does not depend on the distance between the light source and the object [45]. In addition, we also equipped our perspective Cook-Torrance model with spatially varying lightening in which the point-wise light source implies how the light varies from one point to another. This assumption facilitates the independent control of intensity. To this aim, a three dimensional vector field of the light $L$ should be imposed instead of a directional light vector. In this case, the rate at which the light falls off w.r.t. the distance should be induced to make the model more consistent with the actual physical aspects. Integrating all the mentioned points, the new spatially dependent light vector $L$ is defined as:

$$L(x, y) = \Pi \frac{\Upsilon - S(x, y)}{\|\Upsilon - S(x, y)\|}$$

(13)

with the light attenuation $\Pi$:

$$\Pi = \frac{l_d}{\|\Upsilon - S(x, y)\|^2}$$

(14)

where $\Upsilon$ is the light source position. Since we developed the perspective projection, the surface parametrization as in
BFGS formula for updating expression is known as Quasi-Newton condition or the BFGS where the difference between various optimization approaches is caused by the definition of $d_k$.

5. Optimization process

Using equation (1) for three input images (as the minimum number of input images in PS), we constitute a system of equation $F(X) = 0$, where $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is defined as $F(X) = [f_h]$, $h = 1, 2, 3$ and formed by equations given by (1) corresponding to each input image $I_h(x, y)$, $h = 1, 2, 3$ as:

$$f_h = I_h(x, y) - (\text{diffuse term}) - (\text{specular term}) = 0$$

(15)

To estimate the normal field, we should solve the optimization problem as follows:

$$X^* = \arg\min_{X} \frac{1}{2} \|F(X)\|^2 = \frac{1}{2} \sum_{h=1}^{3} f_h^2$$

In optimization process, we update the solution using an iterative procedure:

$$X^{k+1} := X^k + d^k,$$

(16)

Here, $d^k$ is the search direction in iteration $k$. The main difference between various optimization approaches is caused by the definition of $d^k$.

5.1. Quasi-Newton with BFGS updating

In this section, we investigate various procedures that we applied for the nonlinear system of equations described by $F(x) = 0$. Among many iterative methods, Newton’s method is one of the popular approaches. But this method has some drawbacks which can be removed by its modified versions. Here, we used the updated formula known as Quasi-Newton with BFGS suggested independently by [6, 12, 15, 53], which as explained in [48], has much better performance than other modified versions of Newton’s method. In this approach, the Quasi-Newton condition (secant condition) is used to find the search direction $d^k$:

$$d^k := -(B^{-1})k J(X^k)$$

(17)

where $B$ and $J(X)$ are the Hessian matrix approximation and the Jacobian matrix of $F(X)$, respectively. The following expression is known as Quasi-Newton condition or the BFGS formula for updating $B$:

$$B^{k+1} = B(X^k) + \frac{Y^k(Y^k)^T}{Y^k Y^k} - \frac{U^k(U^k)^T}{U^k U^k}$$

(18)

with

$$Y^k = J(X^{k+1}) - J(X^k),$$

(19)

$$\Theta^k = X^{k+1} - X^k,$$

(20)

$$U^k = B^k \Theta^k$$

(21)

5.2. Levenberg-Marquardt

Another strategy that we considered is the Levenberg-Marquardt method [31, 33] with the search direction determined as:

$$d_k = -(J(X^k) J(X^k)^T + \lambda_k I)^{-1} J(X^k) F(X^k)$$

(22)

where $\lambda_k > 0$ and the $(J(X^k) J(X^k)^T + \lambda_k I)^{-1}$ is positive definite.

5.3. Powell’s Dog Leg

In this technique, a strategy is proposed to choose optimization steps based on a parameter called radius of trust region $\Delta$. Three search directions are applied as follows:

$$h_{gn}^k := -J(X^k)^{-1} F(X^k),$$

(23)

$$h_{sd}^k := -\alpha J(X^k)^T F(X^k),$$

(24)

$$h_{dl}^k := \begin{cases} h_{gn}^k, & \text{if } \|h_{gn}^k\| \leq \Delta \\ \frac{\Delta}{\|h_{sd}^k\|} h_{sd}^k, & \text{if } \|\alpha h_{sd}^k\| \geq \Delta \end{cases}$$

(25)

(26)

$$h_{dl}^k := \alpha h_{sd}^k + \beta (h_{gn}^k - \alpha h_{sd}^k), \quad \text{a.w.}$$

(27)

where

$$\alpha := \frac{\|J(X^k)^T F(X^k)\|^2}{\|J(X^k) J(X^k)^T F(X^k)\|^2}$$

(28)

Here, we have also used the strategy applied in [41] to choose $\beta$ and $\Delta$. The general process of mentioned algorithms is illustrated in the last section.

5.4. Dijkstra Gaussian Mean Curvature (DGMC) technique

As a heuristic search to find a proper initial value for the normal field, we have devised a technique that consists of two phases. At the first stage, we employ the Dijkstra algorithm [10] to find the nearest point in the outer highlight boundary $S$ to each specular pixel $p$. This point is denoted as $q$.

In the second step, we compute the Gaussian Curvature (GC) and the Mean Curvature (MC) properties of $p$ and $q$, separately:

$$GC = K_1 \times K_2, \quad MC = \frac{K_1 + K_2}{2}$$

(29)

where $K_1$ is the smallest and $K_2$ is the biggest eigenvalues of the Hessian matrix of the pixel local neighboring area. Having those parameter for both points allows us to define their local geometric similarity conditions as follows:

$$GC_q = GC_p \pm 5\%$$

(30)

$$MC_q = MC_p \pm 5\%$$

(31)

In case that $q$ does not meet similarity conditions, the next nearest member of $S$ to $p$ will be nominated.
5.5. Recurrent Optimization Network (RON)

We will introduce an intermittent optimization strategy which allows updating not only the surface normal but also the albedo and roughness parameter. At first, we consider a constant albedo and also a constant roughness so that we can obtain the surface normal. After obtaining the normal field, we update the albedo values. Furthermore, the roughness parameter is updated as well. To this aim, we use the following formulation proposed in [16] to update the roughness parameter $m$, where we already obtained the depth map $\bar{z}$ by integrating the normal field.

$$m = \frac{1}{n} \sum_{x=1}^{n} \sum_{y=1}^{n} (z(x, y) - \bar{z}(x, y))$$  \hspace{1cm} (32)

Here, $z(x, y)$ is the mean surface depth, and $n$ is the number of pixels. The whole algorithm of mentioned RON procedure is illustrated in Algorithm 1.

Thanks to this concise optimization network with back-tracing steps and DGMC technique which are integrated into the photometric stereo algorithm, our method enables PS to work in diverse applications by removing restrictions such as diffuse objects, orthographic projection, laboratory setup and professional equipment.

5.6. Integration

As the integrator of the normal field, the energy minimization of the following functional should be considered:

$$\mathcal{F}(u) = \int_{\Omega} \| \nabla z(x, y) - N(x, y) \|^2 dx \ dy$$  \hspace{1cm} (33)

which leads to the Euler-Lagrange equation $\Delta z = \nabla \cdot N$.

In order to solve the mentioned minimization problem, we applied Generalized Minimal RESidual (GMRES) with the initial solution of Simchony [56].

6. Experiments

In the case of synthetic experiments, to perform a quantitative evaluation, we will make use of Mean Angular Error of Normal vectors (MAEN) in degrees and also Mean Squared Error (MSE) of depth.

6.1. Tests of accuracy on synthetic data

In the following, we discuss our results obtained for the synthetic test images as displayed in Figures 6, 7 and 8. All images are rendered using Blender software. The 3D model of Oldman and Woman images are obtained from [1] and [2] respectively. 3D models of the other synthetic images are publicly available at [3]. All images except Coping $(256 \times 200)$ are in the size of $256 \times 256$. All implementations are performed on an Intel Core i7 processor with 8 GB of RAM.

Comparison with other BRDFs:
In Figure 3 documenting the first set of experiments, we represent the results of comparison between the recent approach suggested by Khanian et al. [27] and also the results of Lambertian reflectance model (as the most common model applied in PS). The model presented in [27] is based on the Blinn-Phong reflectance. We extended our experiments to a wide range of the specular material parameter $k_s = 1 - k_d$ to evaluate how the varying $k_s$ influence the error of 3D reconstruction. It can be noticed that the output of our method achieves higher rate of accuracy over [27] and Lambertian model as $k_s$ increases. Whereas, there are some fluctuations when $k_d \gg k_s$ or $k_d \ll k_s$. As illustrated, error rates of presented models are almost similar for low values of $k_s$. However, as $k_s$ increases, the superiority of our model over Lambertian and [27] is significant. This observation agrees well with our motivation for focusing on the high frequency variations and model these components explicitly. One particular interpretation for the sharp ascending trend of the error rate for Coping input image (in the case of higher values of $k_s$ parameter) is the presence of the flat background region which leads to high amount of shiny highlighted areas in input images. As for the difference between [27] and Lambertian model, we find that the Lambertian still performs a little better on very low values of $k_s$; however, when $k_s$ is larger than 0.3, [27] improves accuracy much more than Lambertian model. As indicated, the MSE comparing our 3D reconstruction and ground truth data, demonstrates that our method is capable of producing accurate 3D reconstructions for the indicated images with specularity. The qualitative evaluations can be also seen as recovered surfaces shown in Figure 6. Our faithful 3D reconstructions with a high amount of details even for tiny surfaces containing low-depth details (e.g. Coping) represent the proficiency of our approach.

Optimization procedures:
As the next experiment shown in Figure 5, we compared various optimization procedures (applied in first and second steps of our RON technique) for several surfaces with diverse $k_s$. It turns out that Dog Leg method achieves higher accuracy in all cases. Although, BFGS and Levenberg approaches follow the same trend for lower values of $k_s$ (Levenberg method can slightly outperform BFGS technique for low values of $k_s$) BFGS performs better as $k_s$ increases.

Lightening analysis:
To investigate the 3D reconstruction sensitivity w.r.t. the lightening model, we applied different PS frameworks.
Algorithm 1: Recurrent optimization network

output: Normal vectors, colorful albedo, roughness value and the depth map

while stopping criterion is not satisfied do

1. Compute normal vectors from eq (1) using the proposed DGMC technique and one of Algorithm 2, Algorithm 3 or Algorithm 4 by the constant albedo and roughness

2. Update albedo from eq (1) by the obtained normal field using the same choice of Algorithm 2, Algorithm 3 or Algorithm 4

3. Integrate the normal field using the proposed approach in section 5.6. to obtain depth map \( z \)

4. Update the roughness value by applying equation (32)

5. Update the normal field from eq (1) using the recent albedo and roughness

end

Figure 3. Comparison with [27] and results obtained by Lambertian reflectance (as the most common model applied in PS) in presence of specularity. For a fair evaluation, these experiments are performed against various material parameter \( k_s \) for different input images.

Quantitative and qualitative results of this investigation are shown in Table 3 and Figure 7, respectively. To this end, we evaluated our approach regarding different available lightening conditions. In the first experiment, we consider the distant light source leads to perform directional-light-based PS referred as DLPS. The output images of this simulation resemble the real-world images that are illuminated by sun. The second evaluation is conducted to simulate spatially dependent light resulting in point-wise-light-based PS denoted as PLPS. We observed that the higher accuracy is achieved by applying the point-wise light source and adjusting lightings based on spatially dependent light as presented in (13) and (14). Figure 7 reveals that DLPS has problem to recover some areas since capturing images by directional-light leads to more shadows as shown in Figure 7. On the other hand, using the point-wise light source in the scene and formulating PLPS can provide the higher amount of accuracy (cf. Table 3).
Coping

| Methods | MAEN | MSED | Time |
|---------|------|------|------|
| Our     | 0.4163 | 0.0070 | 0.0485 |
| [22]    | 22.5992 | 0.1879 | 40.9955 |
| [4]     | 12.9628 | 0.0424 | 1.3645 |
| [36]    | 15.2238 | 0.0495 | 1.4385 |
| [67]    | 11.9886 | 0.0409 | 1.4385 |

Soldier

| Methods | MAEN | MSED | Time |
|---------|------|------|------|
| Our     | 0.7241 | 0.0138 | 0.0470 |
| [22]    | 24.1228 | 0.0511 | 37.5942 |
| [4]     | 14.2859 | 0.0306 | 5.2871 |
| [36]    | 10.3619 | 0.0356 | 1.2781 |
| [67]    | 4.2274  | 0.0424 | 1.3470 |

Militaryman

| Methods | MAEN | MSED | Time |
|---------|------|------|------|
| Our     | 0.2045 | 0.0007 | 0.0320 |
| [22]    | 19.9092 | 0.1205 | 45.9023 |
| [4]     | 10.7417  | 0.0107 | 1.0775 |
| [36]    | 5.1195 | 0.0290 | 1.1693 |
| [67]    | 3.9751 | 0.0222 | 1.3908 |

Horseman

| Methods | MAEN | MSED | Time |
|---------|------|------|------|
| Our     | 0.5948 | 0.0079 | 0.0460 |
| [22]    | 15.4457 | 0.0279 | 28.2307 |
| [4]     | 10.7417 | 0.0107 | 1.0775 |
| [36]    | 5.1195 | 0.0290 | 1.1693 |
| [67]    | 3.9751 | 0.0222 | 1.3908 |

Soldier

| Methods | MAEN | MSED | Time |
|---------|------|------|------|
| Our     | 0.2659 | 0.0053 | 0.0429 |
| [22]    | 24.3848 | 0.06356 | 40.4892 |
| [4]     | 19.9172 | 0.0430 | 57.1407 |
| [36]    | 3.9674 | 0.0133 | 0.9929 |
| [67]    | 5.4562 | 0.0409 | 1.4496 |

Table 1. Comparison of several PS methods w.r.t Mean Square Error of Depth (MSED), Mean Angular Error of Normal vectors (MAEN in degrees), and Time in seconds. Innovations aggregated in our approach leads to a significant improvement regarding obtained normal vectors and depth as well as computational Time.

Figure 4. Left: Total MSED and Right: Total MAEN gained by our approach and compared methods on all images mentioned in Table 1. As it can be seen in the graph, the proposed method can be also favored due to its higher accuracy w.r.t both obtained normal vectors and depth map. To illustrate, our technique can provide accuracy improvement in normal field extraction over [22] by 97.78 % and over [4] by 96.87 %. Furthermore, this improvement in accuracy of depth reconstruction by our scheme over [22] is 89.51 % and over [4] is 76.17 %.

Table 2. Our total speed improvement over other methods for all images of Table 1. It can be seen that our scheme can decrease consuming time dramatically as well as providing reliable results.

Comparison with several PS methods:
As another evaluation shown in Table 1, we compared our method from different aspects on the set of images with the state-of-the-art approaches: constrained bivariate regression [22], differential ratios [36], dimensionality reduc-

tion based on PCA [4] and also least squares-regression [67]. For a fair comparison, normal map are obtained by the finite-difference of depth map for the methods suggested in [22, 36] (these schemes provide depth, but not normal vectors) and the depth map is obtained by our integrator (section 5.6.) for [4, 67] (these techniques obtain normal vectors, but not depth map). All methods show worse performances than our approach. One reason is that high-frequency variations are mis-classified and leads to the unpredictable errors. As can be seen in Table 1, the proposed strategy dramatically reduces error in computing normal
vectors and depth map in comparison with other approaches thanks to our effective techniques of DGMC and RON and capability of handling specularities. The total accuracy of obtained normal field and depth map for all images applied

![Figure 5. MSE of our 3D reconstructions with different optimization procedures for specular surfaces with varying material parameter (k_s) for different input images.](a) Oldman (b) Militaryman (c) Carpenter (a) Woman]

Table 3. MSE of our 3D shape reconstruction with perspective Cook-Torrance reflectance model in the presence of specularities using different lightening models.

| Lightening Model | MSE for Oldman | MSE for Woman | MSE for Carpenter |
|------------------|----------------|--------------|------------------|
| PLPS for Oldman  | 0.007700       | 0.008439     | 0.002580         |
| DLPS for Oldman  | 0.012408       | 0.011801     | 0.003085         |
| PLPS for Woman   | 0.008439       | 0.011801     | 0.002580         |
| DLPS for Woman   | 0.012408       | 0.011801     | 0.003085         |
| PLPS for Carpenter | 0.002580   | 0.003085     |                  |
| DLPS for Carpenter | 0.002580 | 0.003085     |                  |

in Table 1 is shown in Figure 4. For example, we could improve the accuracy in extracted normals by 93.15% over [36] and 90.43 % over [67] and provide higher accuracy in depth reconstruction by 73.29% over [36] and 70.46 % over [67]. Furthermore, the first row in Table 1 demonstrates the efficiency of our proposed method. Our method is the first one with significant high speed providing all information of normal vectors, colorful albedo, roughness value and depth map. Our speed improvement over mentioned techniques is represented in Table 2. The running time of [36, 67] are comparable, yet their results appear both quantitatively and qualitatively different. The qualitative comparison of our approach with the mentioned methods is shown in Figure 8. The performance ranking of methods is consistent with qualitative results. Our results demonstrate the ability of the proposed scheme to provide a complete reconstruction without any deviation at specularities which outperforms other comparing approaches.

6.2. Test of applicability on real-world images

Finally, we evaluated the capability of our approach on a set of real-world images shown in Figures 9, 10, 11, 12 as a test of the practical applicability of our method. All the images used in these experiments are captured by the camera of a Samsung Galaxy S5 smartphone and using a flash. For image acquisition the photographed figurine is standing on a table covered by a black tablecloth as shown in Figure 2. The light directions are measured by considering the camera as the origin of the coordinate system. The recovered 3D shape for the first real-world image is presented in Figure 9. In this case, we created a complex scene composed of numerous small statues with different sizes as a challenging set. However, we could reconstruct the surface with correct geometry where relevant details and main features are recovered faithfully. Let us stress that our real-world images are taken without any laboratory equipment or controlled setups. Nevertheless, the reconstruction result is in our opinion very reasonable. This test demonstrates the robustness of our method and its potential for real-world applications even using consumer-level equipment. Moreover, again for this real-world scene, we applied different PS-based lightnings models for a same light
source of a flash. Improved quality of the 3D reconstruction with spatially dependent lights (PLPS) shown in Figure 9 (in terms of the height field of depth map and details) can be an advantage offered by this model. More real-world experiments are presented in Figures 10 and 11. As can be seen, our approach succeeds in producing accurate 3D reconstruction results even for high frequency details of tiny structure of Berlin souvenir statues. The colors of these images are extracted in Figure 12 where the diversity of colors is recovered as well in our non-uniform colorful albedo reconstruction. Let us also point out that our results show the capability of reliable 3D reconstruction for specular materials by using a minimum number of input images (for all synthetic and real-world experiments, we used three input images). Thus we document here that our method works under reasonable practical conditions with inexpensive instruments so that it can be used for many potential applications.

7. Conclusion

We presented a real-time robust PS which is also benefiting from perspective Cook-Torrance reflectance model to explicitly handle specularities and remove the limitations of working with diffuse materials or orthographic projection. A RON optimization technique based on DGMC approach is introduced to obtain accurate information of normal vectors, depth, roughness and colorful albedo. Our proposed RON and DGMC techniques can be applied in any highly non-linear optimization framework as a key point for providing multi-variable solutions and proper initialization. These innovations provide important steps towards a reliable PS. Furthermore, we equipped our model with the spatially dependent lightening offering more reasonable reconstructions. We have demonstrated the applicability of our method by applying minimum number of input images without any laboratory conditions or facilities. Furthermore, we stretch our approach to its full potential by extracting 3D reconstruction of micro-prints. Our method has the potential to be useful as the basis of future developments.
Figure 6. Left: input images including specularity. Right: our 3D reconstruction results using perspective Cook-Torrance reflectance model. These results illustrate the capability of the proposed method for providing faithful reconstructions with high frequency details even for the fine details of a tiny surface (e.g. first input image).

Figure 7. First row: input specular images with directional light. Second row: our 3D output using DLPS. Third row: input images with pointwise light. Last row: our 3D output with PLPS. Images produced by directional light are imposed by more specularity and shadows in comparison with images with pointwise light. These shadow parts lead to losing some reconstructed area.
Figure 8. Comparison of several PS methods. From left: 1) input image (Militaryman) including high specularities, and the 3D results of 2) our reconstruction applying perspective Cook-Torrance reflectance model without any deviation or artifacts, 3) reconstruction of [67] that could not handle specularity areas, 4) reconstruction of [36] which shows deviations specially around the highly specular areas 5) [4] shows problem in producing complete 3D reconstruction in presence of specularity and 5) [22] which still includes distortions.

Figure 9. From left: 1) input real-world image including several tiny statues captured without laboratory setting cf. Figure 2, 2) our 3D reconstruction with Cook-Torrance reflectance model, 3) our 3D reconstruction using DLPS for the face of the smallest statue in the scene and 4) our 3D reconstruction using PLPS for the same statue.

Figure 10. More real-world experiments with specularity, first row) input images including Einstein statue, two Berlin souvenir statues with comprehensive structure and a metallic coin, second row) our 3D reconstruction using DLPS and last row) our 3D reconstruction using PLPS. High quality details and structures recovered by the proposed approach confirm offered advantages and agrees well with them.
Figure 11. Some qualitative evaluations for the proposed method: from left) close-up views of our reconstruction results corresponding to the red, blue, green and magenta rectangular areas in Coping as synthetic input and Einstein, Coin and first Berlin statue as the real scenes. It can be seen that our method achieves desirable reconstruction quality even for very fine details.

Figure 12. Reconstructed albedo of two color input images shown in Figure 10. As can be seen, our RON approach can produce non-uniform albedo from images with a variety of colors. Since the albedo is not same across the entire real-world surfaces, our variant albedo extraction is more adjusted for real-world situations. It should be noticed that these statues are very tiny surfaces.
Algorithm 2: Quasi-Newton with BFGS updating

Primary input : $X^0$ that should be computed using our proposed "DGMC" technique explained in section 5.4. of the paper

Secondary inputs: $\vartheta = 0$, $k = 0$

output : Normal vectors for photometric stereo

while $\|J(X^k)\| > \varepsilon$, and $k < k_{max}$ and $\vartheta < \vartheta_{max}$ do

Compute $d^k := - (B(X^k)^{-1}) J(X^k)$

$[X^{k+1}, d\vartheta] \leftarrow$ line search $(X^k, d^k)$ as applied in reference [33] of the paper

$\vartheta \leftarrow \vartheta + d\vartheta$

$\Theta \leftarrow X^{k+1} - X^k$

$Y \leftarrow J(X^{k+1}) - J(X^k)$

if $(\Theta)^T Y > \sqrt{\varpi} \|\Theta\|_2 \|Y\|_2$ then

$U \leftarrow B(X^k) \Theta$

$B(X^{k+1}) := B(X^k) + \frac{Y^T(Y^k)^T Y}{(d^k)^T \Theta^T Y} - \frac{U^T(U^k)^T U}{(d^k)^T U^T}$

end

$k \leftarrow k + 1$

end

Where $\varpi$ is the computer accuracy and $\varepsilon = 1e-12$ is considered.

---

Algorithm 3: Levenberg-Marquardt

Primary input : $X^0$ that should be computed using our proposed "DGMC" technique explained in section 5.4. of the paper

Secondary inputs: $\vartheta = 2$, $\lambda_0 = \tau \max\{a_{ii}\}$, where $\{a_{ii}\}$ is the set of the diagonal elements of $A = J(X)J(X)^T$, $\Psi(X) = \frac{1}{2} \|F(X)\|^2_2$, $k = 0$

output : Normal vectors for photometric stereo

while $\|J(X^k)^T F(X^k)\|_\infty > \varepsilon_1$, and $k < k_{max}$ do

$d^k := -(J(X^k)^T J(X^k) + \lambda_k I)^{-1} J(X^k)^T F(X^k)$

if $\|d^k\| > \varepsilon_2 (\|X^k\| + \varepsilon_2) \text{ then}$

$X^{k+1} \leftarrow X^k + d^k$

$N = \frac{1}{2} (d^k)^T (\lambda_k (d^k)^T - J(X^k)^T F(X^k))$

if $N > 0$ then

$\lambda_{k+1} = \lambda_k \max\left\{\frac{1}{4}, 1 - (2N - 1)^3\right\}$, $\vartheta = 2$

else

$\lambda_{k+1} = \lambda_k \vartheta$, $\vartheta = 2\vartheta$

end

else

Exit

end

$k \leftarrow k + 1$

end

Where $\tau = 9e-2$ and $\varepsilon_1 = \varepsilon_2 = 1e-15$ are considered.
Algorithm 4: Powell’s Dog Leg

**Primary input**: $X^0$ that should be computed using our proposed “DGMC” technique explained in section 5.4. of the paper

**Secondary inputs**: $\Delta = \Delta_0$, $\Psi(X) = \frac{1}{2} \| F(X) \|^2_2$, $k = 0$

**output**: Normal vectors for photometric stereo

while $\| F(X^k) \|_\infty > \epsilon_1$ and $\| J(X^k)^T F(X^k) \|_\infty > \epsilon_2$ and $(k < k_{max})$ do

\[ \alpha = \frac{\| J(X^k)^T F(X^k) \|^2_2}{\| J(X^k) J(X^k)^T F(X^k) \|^2} \]

\[ h_{sd} = -\alpha J(X^k)^T F(X^k) \]

\[ h_{gn} = -J(X^k)^{-1} F(X^k) \]

\[ h_{dl} = \begin{cases} h_{gn} & \| h_{gn} \| \leq \Delta \\ \frac{h_{sd}}{\| h_{sd} \|} & \| \alpha h_{sd} \| \geq \Delta \\ \alpha h_{sd} + \beta (h_{gn} - \alpha h_{sd}) & \text{o.w.} \end{cases} \]

if $\| h_{dl} \| > \epsilon_3 (\| X^k \| + \epsilon_3)$ then

\[ X^{k+1} \leftarrow X^k + h_{dl} \]

\[ N = \frac{\Psi(X^k) - \Psi(X^{k+1})}{L(0) - L(h_{dl})} \]

if $N > 0.75$ then

\[ \Delta = \max \{ \Delta, 3 \| h_{dl} \| \} \]

else if $N < 0.25$ then

\[ \Delta = \frac{\Delta}{2} \]

end

if $\Delta \leq \epsilon_3 (\| X^k \| + \epsilon_3)$ then

\[ \text{Exit} \]

end

else \text{Exit}

end

end

where

\[ L(0) - L(h_{dl}) = \begin{cases} \Psi(X) & \text{if } h_{dl} = h_{gn} \\ \Delta \left( 2\| \alpha J(X^k)^T F(X^k) \| - \Delta \right) \frac{2\alpha}{2\alpha} & \text{if } h_{dl} = \frac{-\Delta J(X^k)^T F(X^k)}{\| J(X^k)^T F(X^k) \|} \\ \frac{1}{2} \alpha (1 - \beta)^2 \| J(X^k)^T F(X^k) \|^2 + \beta (2 - \beta) F(X) & \text{o.w.} \end{cases} \]

and for choosing $\beta$, the strategy proposed by reference [33] is applied.

In addition, $\Delta_0 = 5e-1$ and $\epsilon_1 = \epsilon_2 = \epsilon_3 = 1e-12.$
References

[1] msr-zbethel-tu, github repository, d. matu-rana. https://github.com/dimatura/msr-zbethel-tu/tree/master/MSR-Driver/Media/statue2.obj. Accessed: 2016. 6

[2] @roxlu, jadi, github repository. https://github.com/roxlu/jadi/blob/master/demos/ambient-occlusion/bin/data/statue.obj. Accessed: 2016. 6

[3] @roxlu, jadi, github repository. https://www.yobi3d.com/. Accessed: 10.9.2017. 6

[4] 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. 8, 9, 12

[5] J. F. Blinn. Models of light reflection for computer synthesized pictures. In ACM SIGGRAPH Computer Graphics, volume 11, pages 192–198. ACM, 1977. 2

[6] C. G. Broyden. The convergence of a class of double-rank minimization algorithms 1. general considerations. IMA Journal of Applied Mathematics, 6(1):76–90, 1970. 5

[7] 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. 2

[8] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. ACM Transactions on Graphics (TOG), 1(1):7–24, 1982. 2, 3

[9] E. R. Davies. Machine vision: theory, algorithms, practicalities. Elsevier, 2004. 4

[10] E. W. Dijkstra. A note on two problems in connexion with graphs. Numerische mathematik, 1(1):269–271, 1959. 5

[11] A. R. Farooq, M. L. Smith, L. N. Smith, and S. Midha. Dynamic photometric stereo for on line quality control of ceramic tiles. Computers in industry, 56(8-9):918–934, 2005. 1

[12] R. Fletcher. A new approach to variable metric algorithms. The computer journal, 13(3):317–322, 1970. 5

[13] M. Galo and C. L. Tozzi. Surface reconstruction using multiple light sources and perspective projection. In International Conference on Image Processing, volume 2, pages 309–312, 1996. 2

[14] A. S. Georgiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE transactions on pattern analysis and machine intelligence, 23(6):643–660, 2001. 1

[15] D. Goldfarb. A family of variable-metric methods derived by variational means. Mathematics of computation, 24(109):23–26, 1970. 5

[16] C. Gullon. Height recovery of rough surfaces from intensity images. PhD thesis, Heriot-Watt University, 2003. 6

[17] R. Hartley and A. Zisserman. Multiple view geometry in computer vision. Cambridge university press, 2003. 4

[18] N. Hoffman. Background: physics and math of shading. Physically Based Shading in Theory and Practice, 24(3):211–223, 2013. 4

[19] B. K. P. Horn. Robot Vision. The M.I.T. Press, 1986. 4

[20] B. K. P. Horn and M. J. Brooks. Shape from shading. MIT press, 1989. 4

[21] B. K. P. Horn, R. J. Woodham, and W. M. Silver. Determining shape and reflectance using multiple images. M.I.T. Artificial Intelligence Laboratory, Memo 490, 1978. 1

[22] S. Ikehata and K. Aizawa. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2179–2186, 2014. 8, 9, 12

[23] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pages 318–325. IEEE, 2012. 2

[24] M. K. Johnson and E. H. Adelson. Retrographic sensing for the measurement of surface texture and shape. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 1070–1077. IEEE, 2009. 2

[25] D. Kang, Y. J. Jang, and S. Won. Development of an inspection system for planar steel surface using multispectral photometric stereo. Optical Engineering, 52(3):039701, 2013. 1

[26] M. Khanian, A. S. Boroujerdi, and M. Breuß. Perspective photometric stereo beyond lambert. In Twelfth International Conference on Quality Control by Artificial Vision 2015, volume 9534, page 95341F. International Society for Optics and Photonics, 2015. 2

[27] M. Khanian, A. S. Boroujerdi, and M. Breuß. Photometric stereo for strong specular highlights. accepted for publication in Computational visual media (springer), 2017. 2, 6, 7
[28] M. Kurt and D. Edwards. A survey of BRDF models for computer graphics. *ACM SIGGRAPH Computer Graphics*, 43(2):4, 2009. 3
[29] J. Lambert and D. L. DiLaura. Photometry, or, on the measure and gradations of light, colors, and shade: Translation from the Latin of photometria, sive, de mensura et gradibus luminis, colorum et umbrae. *Illuminating Engineering Society of North America*, 2001. 1
[30] K. M. Lee and C. C. J. Kuo. Shape from photometric ratio and stereo. *Journal of Visual Communication and Image Representation*, 7(2):155–162, 1996. 2
[31] K. Levenberg. A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 5:164–168, 1944. 5
[32] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 619–626. IEEE, 2005. 2
[33] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963. 5
[34] G. McGunnigle. The classification of textured surfaces under varying illuminant direction. PhD thesis, Citeseer, 1998. 1
[35] R. Mecca and Y. Quéau. Unifying diffuse and specular reflections for the photometric stereo problem. In *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*, pages 1–9. IEEE, 2016. 2
[36] R. Mecca, E. Rodola, and D. Cremers. Realistic photometric stereo using partial differential irradiance equation ratios. *Computers and Graphics*, 51:8–16, 2015. 2, 8, 9, 12
[37] R. Mecca, G. Rosman, R. Kimmel, and A. M. Bruckstein. Perspective photometric stereo with shadows. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 258–269. Springer, 2013. 2
[38] 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(2-3):229, 2010. 2
[39] Y. Mukai, Y. Ishii, and T. Shakunaga. Analysis of photometric factors based on photometric linearization. *JOSA A*, 24(10):3326–3334, 2007. 2
[40] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: physical and geometrical perspectives. *IEEE Transactions on pattern analysis and machine intelligence*, 13(7):611–634, 1991. 2
[41] H. B. Nielsen and K. Madsen. Introduction to optimization and data fitting. *Month*, page 176, 2010. 5
[42] P. S. Ogun, M. R. Jackson, and R. M. Parkin. In-process surface profile assessment of rotary machined timber using a dynamic photometric stereo technique. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 226(6):823–830, 2012. 1
[43] T. Papadimitri and P. Favaro. A new perspective on uncalibrated photometric stereo. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1474–1481. IEEE, 2013. 2
[44] V. Parot, D. Lim, G. González, G. Traverso, N. S. Nishioka, B. J. Vakoc, and N. J. Durr. Photometric stereo endoscopy. *Journal of biomedical optics*, 18(7):076017, 2013. 1
[45] M. Pharr, W. Jakob, and G. Humphreys. Physically based rendering: From theory to implementation. Morgan Kaufmann, 2016. 3, 4
[46] B. T. Phong. Illumination for computer generated pictures. *Communications of ACM*, 18(6):311–317, 1975. 2
[47] M. D. Plumbley. Algorithms for nonnegative independent component analysis. *IEEE Transactions on Neural Networks*, 14(3):534–543, 2003. 2
[48] M. J. D. Powell. A hybrid method for nonlinear equations. *Numerical methods for nonlinear algebraic equations*, 1970. 5
[49] L. Quan, H. Y. Shum, et al. Highlight removal by illumination-constrained inpainting. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 164–169. IEEE, 2003. 2
[50] Y. Quéau, F. Lauze, and J. D. Duro. A 11-tv algorithm for robust perspective photometric stereo with spatially-varying lightings. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 498–510. Springer, 2015. 2
[51] B. Sathyabam, S. Raju, and A. Varadhan. Quaternion photometric stereo for rotation invariant surface texture classification. *American Journal of Applied Sciences*, 8(10):992, 2011. 1
[52] J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595, 1996. 2
[53] D. F. Shanno. Conditioning of quasi-newton methods for function minimization. *Mathematics of computation*, 24(111):647–656, 1970. 5
[54] H. L. Shen and Q. Y. Cai. Simple and efficient method for specularity removal in an image. *Applied optics*, 48(14):2711–2719, 2009. 2

[55] L. Shen and P. Tan. Photometric stereo and weather estimation using internet images. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1850–1857. IEEE, 2009. 1

[56] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving poisson equations in computer vision problems. *IEEE transactions on pattern analysis and machine intelligence*, 12(5):435–446, 1990. 6

[57] M. L. Smith. The analysis of surface texture using photometric stereo acquisition and gradient space domain mapping. *Image and vision computing*, 17(14):1009–1019, 1999. 1

[58] W. Smith and F. Fang. Height from photometric ratio with model-based light source selection. *Computer Vision and Image Understanding*, 145:128–138, 2016. 1

[59] R. T. Tan and K. Ikeuchi. Separating reflection components of textured surfaces using a single image. *IEEE transactions on pattern analysis and machine intelligence*, 27(2):178–193, 2005. 2

[60] K. L. Tang, C. K. Tang, and T. T. Wong. Dense photometric stereo using tensorial belief propagation. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 132–139. IEEE, 2005. 2

[61] Y. Tang, R. Salakhudinov, and G. Hinton. Deep lambertian networks. *arXiv preprint arXiv:1206.6445*, 2012. 1

[62] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 611–616. IEEE, 2005. 2

[63] N. Tsumura, Y. Miyake, and F. H. Imai. Medical vision: measurement of skin absolute spectral-reflectance image and the application to component analysis. In *Proceedings of the 3rd International Conference on Multispectral Color Science (MCS’01)*, pages 25–28, 2001. 1

[64] F. Verbiest and V. G. L. Photometric stereo with coherent outlier handling and confidence estimation. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 2

[65] L. B. Wolff. Relative brightness of specular and diffuse reflection. *Optical Engineering*, 33(1):285–294, 1994. 2

[66] R. J. Woodham. Photometric stereo: a reflectance map technique for determining surface orientation from image intensity. In *Image Understanding Systems and Industrial Applications, SPIE*, volume 0155, pages 136–143, 1978. 1, 4

[67] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):134–144, 1980. 1, 4, 8, 9, 12

[68] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Asian Conference on Computer Vision*, pages 703–717. Springer, 2010. 2

[69] T. P. Wu and C. K. Tang. Photometric stereo via expectation maximization. *IEEE transactions on pattern analysis and machine intelligence*, 32(3):546–560, 2010. 2

[70] Q. Yang, S. Wang, and N. Ahuja. Real-time specular highlight removal using bilateral filtering. In *European conference on computer vision*, pages 87–100. Springer, 2010. 2

[71] J. Ye and J. Yu. Ray geometry in non-pinhole cameras: a survey. *The Visual Computer*, 30(1):93–112, 2014. 4

[72] C. Yu, Y. Seo, and S. W. Lee. Photometric stereo from maximum feasible lambertian reflections. In *European Conference on Computer Vision*, pages 115–126. Springer, 2010. 1, 2

[73] S. Zafeiriou, G. A. Atkinson, M. F. Hansen, W. A. P. Smith, V. Argyriou, M. Petrou, M. L. Smith, and L. N. Smith. Face recognition and verification using photometric stereo: The photoface database and a comprehensive evaluation. *IEEE Transactions on Information Forensics and Security*, 8(1):121–135, 2013. 1

[74] J. Zheng, K. Fraser, J. C. Ho, J. V. Mendez, E. Natalia, P. Mandal, C. C. Huang, and R. a. o. Chiou. E-quality control in solar panel inspection using photometric stereo reconstruction and 3d measurement. In *IEEE Annual Conference. Proceedings*, page 3430. Institute of Industrial and Systems Engineers (IISE), 2013. 1

[75] S. K. Zhou, G. Aggarwal, R. Chellappa, and D. W. Jacobs. Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):230–245, 2007. 1