

Revisiting Outlier Rejection Approach for Non-Lambertian Photometric Stereo

Kevin H. M. Cheng, Ajay Kumar

Abstract—Photometric stereo offers a single camera based approach to recover 3D information and has attracted wide range of applications in computer vision. Presence of non-Lambertian reflections in almost all the real-world objects limits the usage of the Lambertian model for surface normal vector estimation. Previous methods proposed to address such non-Lambertian phenomena employ an outlier rejection approach while more recent methods introduce BRDF models which can generate more accurate results. However, results with comparable accuracy can also be achieved by simply filtering the observed intensity values. This paper presents two novel outlier rejection techniques which attempt to identify the data which are more reliable and likely to be Lambertian. In the first technique, observed intensity values with less reliability are automatically eliminated. This reliability is determined by the responses from a newly introduced inter-relationship function. In the second technique, those photometric ratio equations which are less likely to be Lambertian are identified by observing the residue of the equations. By eliminating the data which is unreliable and likely to be non-Lambertian, surface normal vectors are more accurately estimated. Our comparative and reproducible experimental results using both real and synthetic datasets illustrate superior performance over the state-of-the-art methods, which validates our theoretical arguments presented in this paper.

Index Terms—Photometric stereo, surface normal, non-Lambertian, general reflectance

I. INTRODUCTION

PHOTOMETRIC stereo is a widely studied technique in computer vision for estimating surface normal vectors, which constitutes key sources of information for 3D reconstruction and analysis. Woodham [1] introduced the first classical photometric stereo method that acquired stereo images of a static surface under different illuminations. By assuming Lambertian surfaces, i.e. whose bidirectional reflectance distribution function (BRDF) [2] is independent of illumination directions, the products of light vectors and surface normal vectors are proportional to the observed intensity values for each pixel. Thus, the pixel-wise surface normal vectors can be estimated using a least square error approach. There are many extensions of this work for objects with non-Lambertian surfaces [4], uncalibrated light sources [15], images with environment lightings [16], cameras with non-linear responses [17], and images with subpixel resolution [18]. Some

TABLE I
COMPARATIVE SUMMARY ON RECENT PS METHODS FOR REAL OBJECTS

Approach	Objective	Average Angular Error
Conventional Outlier Rejection [5]	Remove specular and shadow pixels	13.3
BRDF Modeling [10]	Compute the BRDFs	10.3
<i>This Paper</i>	Identify representative Lambertian data	9.1

researchers have also investigated photometric stereo approaches for the multi-view reconstruction [19], and dynamic objects [20].

For real object surfaces encountered in real-world applications, the presence of non-Lambertian observations induce several challenges for accurate estimation of surface normal vectors. Such non-Lambertian phenomena can result from specular reflections, shadow pixels, varying BRDFs for different illumination directions, sensor noises, and inter surface reflections. Previous outlier rejection methods introduced in the literature [5, 6, 24-33] attempted to address such non-Lambertian phenomena by filtering specular and shadow pixels. However, from more recent evaluations on real object dataset (DiLiGenT) [3], it can be observed that the two evaluated methods using the outlier rejection approach [5, 6] do not produce superior results. It shows that previous methods using this approach fail to accurately identify the outliers, when their objectives focus on finding specular and shadow pixels. In addition, the limitation resulting from the varying BRDF for different illumination directions are yet to be addressed. Recent research efforts have attempted to address the non-Lambertian phenomena by developing BRDF models [4, 7-11, 34-38].

The two evaluated methods using the BRDF modeling approach [10, 11] achieve the best performance for the non-Lambertian photometric stereo evaluation. However, it is worth noting that by incorporating a simple filtering method (position threshold method) on the observed intensity values with the traditional least square error approach can also achieve comparable performance. There are two important insights from this finding. One is that the existing BRDF modeling methods do not compute BRDF accurately, which limits the accurate estimation of surface normal vectors. For example, the method [10] has reported the average BRDF fitting errors (RMSE $\times 10^{-4}$) range from 6.71 to 174.88 for variant levels of

additive noise. This approach also fails when only a few samples are available, and is computationally expensive. Another insight is that by only rejecting the observed intensity values, surface normal vectors can be estimated quite accurately. Therefore, there is a strong motivation to revisit the effectiveness of outlier rejection approach for a more accurate non-Lambertian photometric stereo method.

A. Related Work

The first photometric stereo method [1] was introduced with the assumption of Lambertian reflections. However, non-Lambertian phenomena exist in almost all imaging of real objects. Therefore, many studies have attempted to address the limitations caused by non-Lambertian phenomena. A stream of research suggested using an outlier rejection approach for removing specular and shadow pixels, which cause non-Lambertian phenomena. Earlier work [24-26] attempted to select the most reliable intensity values from four light sources. With more light sources, reliable observations can be extracted in a more robust manner by employing various promising mathematical techniques [27-33] to detect and reject the outliers. More recent state-of-the-art methods [5, 6] assumed that the non-Lambertian phenomena are local and sparse. Observations are modeled using a sparse outlier matrix plus a low rank observation matrix. Reference [5] minimized the rank of the matrix while reference [6] enforced rank-3 by employing sparse Bayesian regression. The optimization approach allows searching for the optimal surface normal vectors.

Another stream of research attempted to develop sophisticated BRDF models [34-38] for surface normal estimation. Reference [4] adopted the *Ward* model [47] to explicitly represent the BRDF. Other research utilized general properties of BRDF such as isotropy, reciprocity, and monotonicity [7-11]. A recent research [45] suggested to also model specular reflections in addition to diffuse reflections, while another work [46] suggested to transform the photometric stereo problem into eigen decomposition problem.

Although there have been many promising efforts [4-11, 24-33, 45-46], the comparative effectiveness on real objects remains unclear. Reference [3] developed a photometric stereo dataset (DiLiGenT) of real objects with ground truth surface normal vectors, which is publicly accessible. Comparative experimental results, based on the statistics of angular errors, of various state-of-the-art methods were presented. Some methods (e.g. [4-9]) only perform well on a part of materials. The two representative methods using outlier rejection approach [5, 6] did not produce compelling overall performance, while methods using the BRDF modelling approach [10, 11] achieved the best overall performance for the calibrated non-Lambertian photometric stereo evaluation. A recent method [46] also demonstrated comparable performance on the same dataset.

It is worth noting that results with comparable accuracy can also be produced by incorporating a simple filtering method (position threshold method) on the observed intensity values [3]. The filtering method removes outlying values (i.e. specular and shadow pixels) by sorting the intensity values. The traditional least square error approximation produced

comparable results with the currently best performing methods [10, 11, 46] when only the central 20% of the sorted intensities are used. This suggests that identifying outliers is a critical step for an accurate estimation of surface normal vectors.

B. Our Work and Contributions

Unlike the earlier methods to address only specific limitations (e.g. specular reflections, shadow pixels, varying BRDF for different illumination directions, sensor noises, inter surface reflections) from non-Lambertian surfaces, we are motivated to address most of these limitations at the same time, by identifying the data which are more likely to be influenced by such non-Lambertian phenomena. Following summarized the key contributions in this paper:

1. A new inter-relationship function is developed to identify reliable pixel intensity values from the input data, which allows more accurate estimation of surface normal vectors. State-of-the-art methods typically employ a ranking approach (position threshold method) with the objective of removing specular and shadow pixels, which is achieved by retaining the median range intensity values. Instead, the objective of the inter-relationship function is to remove unreliable values, which is achieved by computing the distances between intensity values. Estimation of surface normal vectors is more accurate when the less reliable observed intensity values are discarded. Systematic experiments are conducted on a real object dataset (DiLiGenT) [3]. The experimental results presented in section IV of this paper indicate superior performance and validates the effectiveness of the proposed function, *i.e.* the average angular error in degrees when using (a) all observations is 11.0; (b) the position threshold method, which is adopted in state-of-the-art methods, is 9.5; (c) our approach is 9.1. This function can also be employed in existing photometric stereo methods.

2. A new truncated photometric stereo method is introduced for a more accurate estimation of surface normal vectors. The photometric equations are extended to photometric ratio equations by assuming constant BRDF values, which is similar to existing formulations. However, the existence of inconstant BRDF values often poses a major limitation for the photometric ratio approach. We attempt to address such limitation by removing the equations which are likely to be non-Lambertian. The photometric equations with large discrepancy in BRDF values generate inconsistent photometric ratio equations, which are detected by the residues and are removed. Remaining equations are solved by the least square error approximation for estimating surface normal vectors. Comparative experiments are performed using a recent real object dataset (DiLiGenT) [3]. Our reproducible [48] experimental results presented in section IV of this paper indicate superior performance and validates the effectiveness of our approach. The average angular error in degrees when using currently available best performing method is 10.3 while our approach can significantly reduce this to 9.1. In addition, experiments are also performed using synthetic images available from a public dataset (MERL BRDF) [22]. The experimental results also achieve superior performance and again validate the effectiveness of our approach, despite our focus being to address the non-Lambertian phenomena in real

situations. The average angular error in degrees when using current best performing method is 0.62 while using our approach is 0.59. Table I summarizes the major difference between our method and the state-of-the-art approaches. The average angular errors are computed from experimental results on DiLiGenT [3], which are detailed in section IV. The errors from the best performing method are also presented.

Rest of the paper is organized as follows: Section II presents the analysis on observed intensity values and introduce the inter-relationship function. Section III presents the analysis on photometric ratio equations and suggests the truncated photometric ratio method. Section IV presents the systematic and comparative experimental results. Section V concludes the paper.

II. ANALYSIS ON OBSERVED INTENSITY VALUES

With the advancement of high speed and low-cost sensors, it is convenient to acquire many images for a more accurate estimation of surface normal vectors using photometric stereo methods. However, observed intensity values in some images can be unreliable and may not be useful for the accurate estimations of surface normal vectors. Therefore, by removing unreliable values, it is expected that estimation of surface normal vectors will be more accurate. In this section, we describe the existing filtering method for removing outlying values and introduce a new inter-relationship function for this problem.

A. Filtering Observed Intensity Values

It can be observed from the study in [3] that, when the position threshold method is incorporated with the traditional least square error approximation approach, it produces a much better result than using all the observed intensity values. With the usage of a tight threshold, the result is even comparable with the currently best performing state-of-the-art methods. This suggests that filtering observed intensity values is a critical step for a more accurate estimation of surface normal vectors when the Lambertian reflectance model is used. Sometimes, the intensity of a pixel is affected by strong specular reflections, or when the pixel is in a shadow region. Those situations strongly affect the observed intensity values, and therefore the photometric stereo equations using those intensities are unreliable. In general, there are also some other factors (e.g. different BRDF for various illumination directions/materials, sensor noises, inter surface reflections) which can adversely affect the observed intensity values, and result in inconsistent system of photometric stereo equations. Since not all observed intensity data are reliable, removing those data values can significantly reduce the number of unreliable samples, which can enable a more accurate estimation of surface normal vectors. The key problem therefore lies in how to determine the observed intensity values to be discarded or accepted.

The position threshold method is described in [3] and is also widely adopted in recent research, including [10] and [11]. It is a simple and efficient method for identifying the potential outliers. The observed intensity values with relatively high intensities are considered as specular points while those with

the relatively low intensities are considered as shadow points. Intensity values of different illuminations are sorted for each pixel. The highest and the lowest intensity values are considered as outliers and are discarded. Those remaining intensities in a specified median range are accepted. Although it is quite effective to remove the extreme observations using this simple approach, there are some drawbacks. The position threshold method only considers the rank and fails to utilize the information from the distribution of the intensity values. There is a possibility that using more advanced methods may achieve better performance. There is some work on this problem such as in [12]. However, those methods lack scientific justification and systematic evaluation on real world datasets with ground truth surface normal vectors such as DiLiGenT [3].

B. Inter-Relationship Function

The observed intensity values which do not closely follow the Lambertian reflectance model are expected to represent outlining values. For example, if an intensity value is associated with strong specular reflections, it is expected that the value is far from other values. Therefore, the set of intensity values which is close to each other is relatively more reliable than the set of intensity values which is far from other values. Furthermore, low intensity values are more likely affected by sensor noise, which causes the observed intensity values to be inaccurate. To evaluate the distance between the observed intensity values, we have adopted a geometric approach with the capability to emphasize the unreliability of the very low intensity values instead of an arithmetic approach. In order to handle the two possibilities that either intensity values in a pair is larger than the other, we incorporate an arithmetic mean so that either case returns the same response. With the above principles, we have introduced a function for determining the reliability of the intensity values. Let m be the number of available images, and x_1, x_2, \dots, x_m be the set of stereo intensity values of a pixel. Any values x equals to 0 or 1 are first filtered out. Let r be the number of remaining intensities, and x_1, x_2, \dots, x_r be the set of remaining intensities. An inter-relationship Function (IRF) is defined as follows:

$$f(i) = \frac{1}{2(r-1)} \sum_{j=1, j \neq i}^r \left(\frac{x_i}{x_j} + \frac{x_j}{x_i} \right) \quad (1)$$

where $i = 1, 2, \dots, r$.

The IRF is a non-linear function which measures the inter-relationship among the observed intensities. If an observation value is far from most of the other values, the IRF response of that observed intensity value will be high. Meanwhile, if an observed intensity value is very small, the responses will be higher than an observed intensity value which is very large. IRF responses of all values are then ranked. Intensity values with high IRF responses are considered to be unreliable. Those with low IRF responses are selected as the reliable observations. Splitting RGB channels can further enhance the selection accuracy. Let (I_r, I_g, I_b) be the set of intensities for RGB channels of a pixel. An inter-relationship function with splitting RGB channels is defined as follows:

$$g(i) = \frac{1}{6(m-1)} \sum_{j=1, j \neq i}^m \sum_{c=r, g, b} \left(\frac{I_{c,i}}{I_{c,j}} + \frac{I_{c,j}}{I_{c,i}} \right) \quad (2)$$

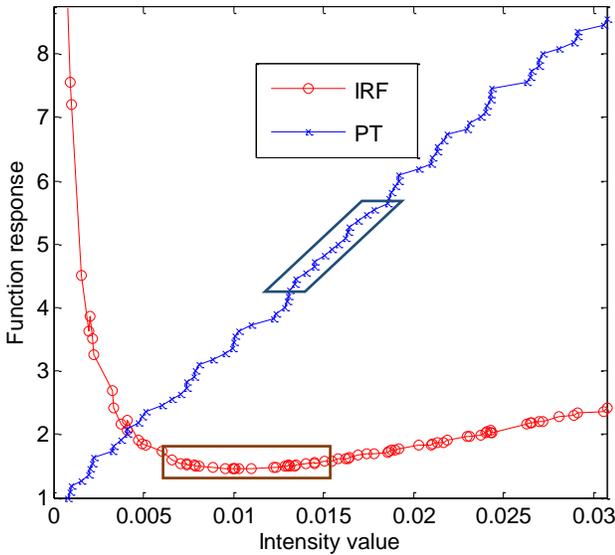


Fig. 1. Illustration of IRF and the position threshold method

where $i = 1, 2, \dots, m$.

The IRF responses are then again ranked. Intensity values with low IRF responses are selected as the reliable observations. Since equation (2) is more computationally expensive, we can simplify the equation so that the implementation can be much simpler and more efficient. IRF function can be approximated as follows:

$$\tilde{g}(i) = \sum_{c=r,g,b} (I_{c,i} \bar{I}_c^* + \frac{1}{I_{c,i}} \bar{I}_c) \quad (3)$$

where $i = 1, 2, \dots, m$, $\mathbf{x}^* = [\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_m}]$, $\bar{\mathbf{x}} = \frac{1}{m} \sum_{j=1}^m x_j$.

Figure 1 illustrates the difference between the IRF method and the position threshold method. A set of stereo intensity values are selected from a generally darker pixel from the object ‘Cat’ in DiLiGenT [3]. The red curve shows the IRF responses for the corresponding intensity values. The blue curve is a scaled rank of the intensity values, which represents the selection criteria for the position threshold method. The position threshold method chooses the data points of the median range which is the middle range of the y-axis (blue box). It can be observed that the intensity values around the minimum range of the IRF responses are assembled together. Those values are expected to be more reliable (brown box). We selected 20 data points using both methods and estimated the surface normal vector using the traditional least square error approximation. The estimated surface normal vectors are compared with the ground truth surface normal vector. The angular error obtained by using the position threshold method is 6.65 while that from the IRF method is 3.21. Comprehensive experimental results are shown in Section IV. The usage of two thresholds (low and high) for the position threshold method may be able to generate the same range of selection criteria as the IRF method. However, the range of selected observations computed from IRF will be different for various combinations of the set of observed intensity values. Furthermore, it is more difficult to automatically determine two parameters.

C. Parameter Selection

For the inter-relationship function, the number of observed intensity values selected, referred to as p here, is the sole parameter. It must be greater than or equal to 3. For example, with 100 available stereo images, the parameter space of p will be from 3 to 100. From our experiments, p is generally in the range of 10% to 30% of the number of available stereo images for achieving optimal performance. It is shown in section IV that smaller p achieves higher efficiency.

III. ANALYSIS ON PHOTOMETRIC RATIO EQUATIONS

Traditional photometric stereo method can also produce a reasonably accurate estimation of surface normal vectors even though the observed intensity values are distorted by non-Lambertian phenomena. This suggests that the major components of the observed intensity values are contributed from the Lambertian reflections (i.e. the dot products between the light vectors and the surface normal vectors). It is believed that there also exists some accurate data which coincide with the Lambertian reflectance model (equation (4)). Therefore, by removing equations which are likely to be non-Lambertian, it is expected that the estimation of surface normal vectors will be more accurate. In this section, we describe a technique for filtering photometric equations with the help of a system of photometric ratio equations. First, traditional system of photometric linear equations is expanded to a system of photometric ratio equations. Inconstant BRDF values is the major limitation of the photometric ratio approach. We attempt to address such limitation by removing the equations which are likely to be non-Lambertian using an iterative filtering approach.

A. Traditional Photometric Equations

The pixel intensities are observed during image acquisitions, while the light vectors are calibrated. The objective of the photometric stereo problem is to estimate surface normal vectors using the following system of linear equations. For each pixel, we can define $\mathbf{i} = [i_1, i_2, \dots, i_m]^T$ be the intensity values corresponding to m different light sources; $\mathbf{l} = [l_x, l_y, l_z]^T$ be the unit vector of a light source; $\mathbf{L} = [\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_m]^T$ be the matrix of the light sources; $\mathbf{n} = [n_x, n_y, n_z]^T$ be the unit normal vector and ρ be the BRDF. A system of linear equations is formulated as follows:

$$\mathbf{i} = \mathbf{L} \cdot \mathbf{n} \cdot \rho(\mathbf{n}, \mathbf{L}) \quad (4)$$

This system of linear equation can be solved by least square error approximation. In the following, we first present the system of photometric ratio equations with an incorrect assumption. Then, a truncated photometric stereo approach is introduced for removing the equations which are likely to be non-Lambertian by handling the incorrect assumption.

B. Formulation using Photometric Ratio Equations

Photometric ratio is a concept for considering photometric invariants, which is discussed by many references [12-14, 40-44]. There are different variations of the photometric ratio equations formulation. Here, we briefly describe our

formulation which allows simple computation for surface normal vectors. When a pair of stereo values (e.g. 2 images) are available, a pair of equations can be rewritten as follows:

$$i_a = \rho_a \cdot (\mathbf{l}_a \mathbf{n}) \quad (5)$$

$$i_b = \rho_b \cdot (\mathbf{l}_b \mathbf{n}) \quad (6)$$

First, assume $\rho_a = \rho_b$ for the initialization. Theoretically, these two BRDFs are different. However, there is no perfect way to accurately compute the BRDF. Inaccurate estimation of BRDF also induce errors in the system for the estimation problem. Therefore, we suggest incorporating this assumption to minimize the errors induced. Evaluation results in Section IV also justified our arguments. Using equation (5) and (6), we can jointly obtain:

$$\frac{i_a}{i_b} = \frac{(\mathbf{l}_a \mathbf{n})}{(\mathbf{l}_b \mathbf{n})} \quad (7)$$

The BRDF values can be eliminated and Equation (7) can be rewritten as follows:

$$i_a \mathbf{l}_b \mathbf{n} = i_b \mathbf{l}_a \mathbf{n} \quad (8)$$

$$(i_a \mathbf{l}_b - i_b \mathbf{l}_a) \mathbf{n} = 0 \quad (9)$$

We also have the constraints that

$$n_x^2 + n_y^2 + n_z^2 = 1 \quad \text{and} \quad n_z > 0 \quad (10)$$

Equation (9) can be rewritten as

$$(i_a \mathbf{l}_b - i_b \mathbf{l}_a) \tilde{\mathbf{n}} = 0 \quad (11)$$

where $\tilde{\mathbf{n}} = \left[\frac{n_x}{n_z}, \frac{n_y}{n_z}, 1 \right]^T$.

With all the possible combinations of the observations, we can formulate a system of linear equations:

$$\mathbf{A} \mathbf{x} = \mathbf{b} \quad (12)$$

where

$$\mathbf{A} = \begin{bmatrix} i_a l_{xb} - i_b l_{xa} & i_a l_{yb} - i_b l_{ya} \\ \vdots & \vdots \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \frac{n_x}{n_z} \\ \frac{n_y}{n_z} \\ \frac{n_z}{n_z} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -(i_a l_{zb} - i_b l_{za}) \\ \vdots \end{bmatrix} \quad (13)$$

The above equations (11) - (13) may become unstable when n_z approaches zero. Since this situation mainly exists in the boundary cases, the proportion of such pixels are small and can be negligible. The traditional system of m linear equations is expanded to a system of C_2^m photometric ratio equations. This system is equivalent to the traditional system only if the all the observed data follows Lambertian model. Since real observations are usually associated with different BRDFs, observed intensity values associated with more similar BRDFs produce more accurate photometric ratio equations. It is shown in section IV that using this system slightly outperforms using the traditional system. However, the incorrect assumption $\rho_a = \rho_b$ for the initialization limits the enhancement on the performance. In the next sub-section, we introduce a truncated photometric ratio approach to address such problem.

Algorithm 1: Surface normal estimation using iterative filtering approach

Input: \mathbf{n}_{in} : initial surface normal vector;
 r : number of equations to be removed;
 i : number of iterations

Output: \mathbf{n}_{out} : enhanced surface normal vector

repeat

1: compute residues by LHS of equation (11);

2: sort the residues

3: remove r equation(s) with largest residue;

4: compute \mathbf{n}_{out} with the new system using Least square error estimation;

until the iteration number is i .

C. Truncated Photometric Ratio Approach

The photometric ratio approach has been briefly introduced in the previous subsection. However, inconstant BRDF values is the major limitation of such photometric ratio approach. Some researchers recommend to solve the homogeneous system from equation (9) by SVD [13]. Others suggest using a variational approach to address this problem [14]. Those methods have not been systematically evaluated with publicly available datasets of real objects with ground truth. Since we have assumed $\rho_a = \rho_b$ during the initialization of the system of photometric ratio equations, some equations can be inaccurate and hinder the accurate estimation of surface normal vectors. We attempt to address this limitation by using an iterative filtering approach to remove the equations which are likely to be non-Lambertian. In the best of our knowledge, there has been no work to incorporate the effectiveness of using an iterative filtering approach to solve the system of photometric ratio equations.

This approach firstly estimates the surface normal vectors using the least square error approximation method, which is one of the most reliable method in the literature. However, the drawback of using such approach is that outliers significantly influence the estimated result. Therefore, we attempt to remove the outlying photometric ratio equations which is likely to be non-Lambertian. When substituting the estimated surface normal vectors into equation (11), residues can be computed as the absolute value of the left-hand side for each equation. If an equation is likely to be Lambertian, the residue should be caused by only the less accurate surface normal vector. However, if an equation is likely to be non-Lambertian, the residue may also be caused by other non-Lambertian factors. Therefore, it can be used as an indicator for whether the equation is likely to be Lambertian. Equations having high residue values are discarded from the system and a new surface normal vector is computed with the remaining equations. With the updated surface normal vector, the indication of residues will be more accurate. Therefore, iteratively inspecting the updated residues, instead of completing this process at once, can allow more accurate estimation of surface normal vectors. This approach significantly improves the accuracy of surface normal vectors estimation. Algorithm 1 summarizes the iterative filtering method.

TABLE II
MEAN ANGULAR ERRORS OF USING THE TRADITIONAL APPROACH WITH DiLiGenT

Methods \ Objects	Ball	Cat	Pot1	Bear	Pot2	Buddha	Goblet	Reading	Cow	Harvest	Average
All observations	4.096	8.413	8.894	8.389	14.650	14.921	18.500	19.803	25.600	30.625	15.389
Position Threshold	1.723	6.609	7.265	6.287	10.434	10.669	12.154	13.663	12.588	21.327	10.272
IRF (RGB)	1.494	6.387	7.076	5.945	9.189	11.171	11.030	13.729	11.540	22.692	10.025

TABLE III
MEAN ANGULAR ERRORS OF USING PHOTOMETRIC RATIO EQUATIONS ON DiLiGenT

Methods \ Objects	Ball	Cat	Pot1	Bear	Pot2	Buddha	Goblet	Reading	Cow	Harvest	Average
All observations	4.061	8.459	8.941	8.370	14.451	14.698	17.853	15.850	24.007	26.605	14.330
Position Threshold	1.719	6.493	7.184	6.291	10.423	10.257	11.962	13.411	12.544	21.177	10.146
IRF (RGB)	1.508	6.240	6.859	5.934	8.960	10.567	10.565	13.354	11.332	22.055	9.737

TABLE IV
MEAN ANGULAR ERRORS OF USING TPR APPROACH ON DiLiGenT

Methods \ Objects	Ball	Cat	Pot1	Bear	Pot2	Buddha	Goblet	Reading	Cow	Harvest	Average
All observations	2.801	6.256	6.629	5.107	10.101	10.439	11.814	13.546	17.347	26.094	11.014
Position Threshold	1.502	5.995	6.764	5.040	9.871	9.048	11.488	11.929	12.544	21.144	9.533
IRF (RGB)	1.500	5.738	6.237	4.969	8.642	8.858	9.996	11.436	11.332	21.902	9.061

TABLE V
MEAN ANGULAR ERRORS OF USING PHOTOMETRIC RATIO EQUATIONS ON DiLiGenT (FIXED PARAMETERS)

Methods \ Objects	Ball	Cat	Pot1	Bear	Pot2	Buddha	Goblet	Reading	Cow	Harvest	Average
IRF (RGB)	1.518	6.240	6.941	7.303	8.992	11.264	11.071	14.495	14.706	22.739	10.527
IRF (Gray Scale)	1.549	6.715	7.103	7.475	9.556	12.815	11.680	16.909	14.918	22.495	11.122
Position Threshold	1.744	6.513	7.206	10.079	10.720	10.378	12.077	15.040	15.105	21.840	11.070
Darkest	8.406	10.694	12.049	11.430	12.477	20.588	18.708	22.281	18.074	31.334	16.604
Brightest	8.529	8.601	11.096	20.126	18.721	13.769	23.244	35.226	36.278	38.612	21.420
Nearest (RGB)	1.800	6.312	7.151	10.142	10.656	9.900	11.945	13.863	14.896	21.501	10.817
Nearest (Gray Scale)	1.795	6.445	7.218	10.355	10.930	10.242	12.132	15.122	15.087	21.576	11.090

D. Parameter Selection

For the truncated photometric ratio method, there are two parameters: the number of equations to be removed in each round, defined as r ; and the number of iterations, defined as i . We suggest setting r to be 1 for better accuracy. However, a larger value of r allows better efficiency. p observed intensity values produce C_2^p photometric ratio equations. The remaining number of equations for estimating a surface normal vector must be greater than or equal to 3. Therefore, if r is 1, the parameter space of i can be from 3 to C_2^p . From our experiments, i is generally in the range of 30% to 50% of the value C_2^p for achieving optimal performance. It is shown in section IV that smaller i achieves higher efficiency.

IV. EXPERIMENTS AND RESULTS

In this section, the proposed approaches are systematically evaluated to ascertain the effectiveness of each individual method. The experiments are performed on both real and synthetic datasets.

A. Evaluation with Real Objects

Our proposed methods are evaluated using the latest benchmark dataset of real objects with ground truth surface normal vectors, DiLiGenT [3]. In terms of shapes, there are objects with a sphere and smooth surfaces. Some objects are with complicated geometry with local details and concave parts. In terms of BRDF, the objects include diffuse and specular materials with rough surfaces, strong and sparse specular spikes, and soft specular lobes. Some materials are also spatially varying, and with metallic paint. For each object, there

are 96 intensity images with corresponding light vectors, a mask image with 15 to 57 thousand white pixels, and ground truth surface normal vectors for each pixel. The angular errors are computed as the arc cosine of the dot products between the estimated surface normal vectors and the ground truth surface normal vectors. Similar to as in the state-of-the-art methods [10, 11, 45, 46], the evaluation criteria for each object is based on the mean angular errors (MAE) over all pixels.

(i) Systematic evaluations

Both the proposed IRF for filtering observed intensity values and the Truncated Photometric Ratio (TPR) approach are systematically evaluated. We have compared three observations selection methods for three different approaches, which results in nine sets of evaluation results. The three observations selection methods include using all observations, the position threshold and the IRF method. The three different approaches include the traditional system of linear equations, the photometric ratio equations, and the TPR approach.

Table II-IV present the evaluation results with mean angular errors of using various approaches. The parameters (number of selected observations and the number of iterations) were optimized for each of the object and each method, i.e. the parameters are optimized for each entry in the tables for producing the best possible performance. Results with bold represents the best performing method for that object. Our proposed IRF method generally outperforms both the position threshold method and using all observed intensity values, for all three approaches. In order to further investigate the effectiveness of the IRF method over other possible observations selection methods, additional experiments were

performed using photometric ratio equations with fixed parameters ($p=20, i=1$), which the results are presented in Table V. The IRF method can accept either color (RGB) pixels or gray scale pixels for the computation of the function response. ‘Position Threshold’ method selects the middle range of the sorted intensity values; ‘Darkest’ method selects the lowest range of the sorted intensity values; ‘Brightest’ method selects the highest range of the sorted intensity values. ‘Nearest’ method computes the distances between observed intensity values. Each intensity value is compared with all other intensity values by computing the absolute difference. The mean of the differences is the function response and the intensity values with the lowest range of the function response are selected. This method can also accept either color (RGB) pixels or gray scale pixels for the computation of the function response. These comparative experimental results validate the theoretical arguments in Section II.B. Filtering observed intensity values using IRF (RGB) has been shown to be more robust than using other potential observations selection methods for the real objects.

The photometric ratio equations approach has not been adopted as the state-of-the-art evaluation of this dataset. We have also evaluated the photometric ratio equations approach in this dataset to verify its effectiveness. It can be observed from the comparison between the traditional approach (Table II) and the photometric ratio approach (Table III) that the photometric ratio approach generally outperforms the traditional approach.

Using the photometric ratio equations approach itself does not enhance the performance significantly, since the BRDF of different illuminations are not equal. When comparing our proposed TPR approach (Table IV) with the basic photometric ratio approach (Table III), the performance of TPR approach has significantly improved. These results verify the improved effectiveness of the TPR approach for more accurate estimations of surface normal vectors.

(ii) Evaluation on the effects of parameters

The effects of two main parameters (the number of selected observations p ; the number of iterations i) on the computational time and mean angular errors (using the object ‘Ball’) are evaluated. These experiments were performed on a machine with Intel Core i7-6700HQ (2.60GHz) and were implemented

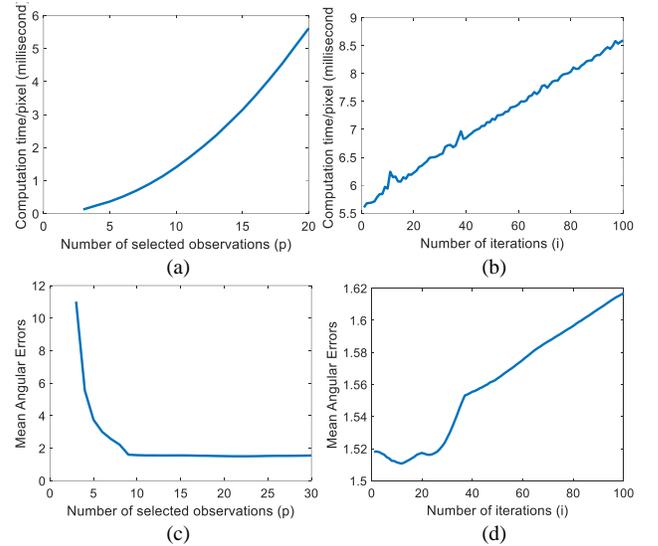


Fig. 2. Effect of the number of selected observations p and the number of iterations i on: (a)/(b) computational time; (c)/(d): mean angular errors

in MATLAB. For the process of selecting observed intensity values, simple position threshold method requires 32 microseconds per pixel while the IRF (RGB) method requires 645 microseconds per pixel. For evaluation the TPR approach, IRF method is adopted for selecting the observed intensity values. For evaluating the effect of p , we set i to be 1. For evaluating the effect of i , we set p to be 20. We also set the number of equations to be removed in each iteration, r to be 1. Figure 2 (a)/(b) show the actual computation time per pixel while Figure 2 (c)/(d) show the mean angular errors. These results also indicate that the number of selected observations p has the time complexity of $O(p^2)$ while the number of iterations i has the time complexity of $O(i)$. Furthermore, the lowest mean angular errors can be obtained using about 20 selected observations with about 10 iterations. The time complexity of our method is larger than the classical method, because there are more equations when the photometric ratio approach is adopted. However, the usefulness of the TPR approach can be justified by the outperforming results, i.e. the accurate estimation of surface normal vectors.

(iii) Comparison with state-of-the-art methods

In this section, we present comparative experimental results obtained from our method and ten state-of-the-art methods [4-

TABLE VI
COMPARISON OF MEAN ANGULAR ERRORS WITH THE STATE-OF-THE-ART RESULTS ON DiLiGENT

Methods \ Objects	Ball	Cat	Pot1	Bear	Pot2	Buddha	Goblet	Reading	Cow	Harvest	Average
LS	4.096	8.413	8.894	8.389	14.650	14.921	18.500	19.803	25.600	30.625	15.389
ECCV12 [8]	13.576	12.338	10.369	19.444	9.841	18.369	17.796	17.170	7.617	19.304	14.582
CVPR12 [6]	2.544	7.210	7.739	7.315	14.088	11.114	16.249	16.166	25.701	29.257	13.738
ACCV10 [5]	2.061	6.726	7.177	6.496	13.123	10.908	15.700	15.394	25.888	30.001	13.348
CVPR10 [9]	3.549	8.402	10.849	11.478	16.371	13.053	14.889	16.824	14.951	21.789	13.215
CVPR08 [7]	2.706	6.529	7.235	5.964	11.032	12.539	13.925	14.175	21.481	30.504	12.609
TPAMI10 [4]	3.211	8.216	8.534	6.620	7.898	14.846	14.222	19.067	9.548	27.839	12.000
TIP15 [45]	2.730	6.660	6.890	5.120	9.800	12.290	11.710	14.560	17.200	25.250	11.220
CVPR14 [11]	3.337	6.743	6.642	7.107	8.768	10.468	9.712	14.189	13.054	25.949	10.597
TIP17 [46]	2.170	5.640	7.280	5.310	8.430	9.300	10.520	13.000	16.790	24.590	10.300
TPAMI14 [10]	1.743	6.116	6.508	6.122	8.777	10.600	10.090	13.629	13.933	25.436	10.295
Ours	1.500	5.738	6.237	4.969	8.642	8.858	9.996	11.436	11.332	21.902	9.061

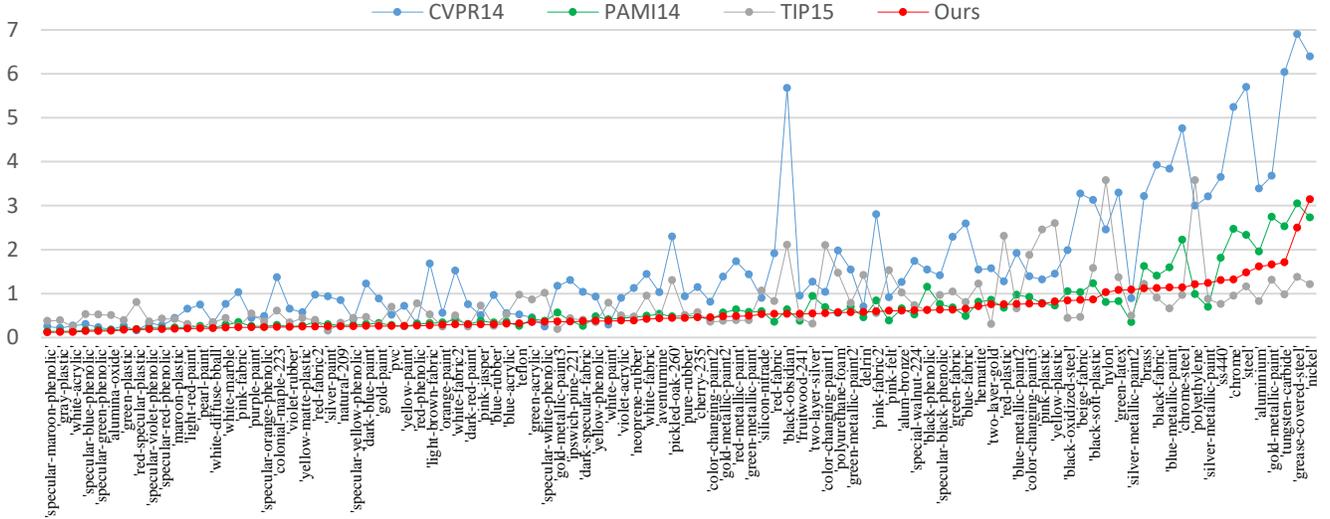


Fig. 4. Comparison of mean angular errors on MERL BRDF dataset

TABLE VII
COMPARISON OF AVERAGE ANGULAR ERRORS WITH THE STATE-OF-THE-ART RESULTS ON MERL BRDF DATASET

Methods	Average Angular Error
CVPR14 [11]	1.6231
TIP15 [45]	0.8065
PAMI14 [10]	0.6835
TIP17 [46]	0.6200
Ours	0.5866

[11, 45, 46]. Our evaluation results for the methods in [4-11] and [46] are as same as those reported in [3] and [46] respectively. Method [45] was evaluated by using the original implementation codes provided along with that paper. The respective parameters for all the evaluated methods in this paper have been provided in the original codes or have been stated in the respective papers. We present our best results with optimized parameters as described in Section IV. A. (i) to ensure fairness in the performance comparisons. Methods [10, 11, 46] achieved the best performance on this dataset. Table VI present summaries of comparative experimental results using mean angular errors. Ours method outperforms the best three methods for seven out of ten objects. Our method also achieves the best performance for five out of ten objects for all evaluated methods, and similar performance to the best performing methods for the remaining objects. It can be observed that our method results in a significant improvement for the average mean angular errors. Besides numerical results, we also present samples of reconstructed images for visual analysis. Figure 3 shows angular error images of the object ‘Cat’, ‘Bear’, ‘Buddha’, ‘Reading’, ‘Cow’, and ‘Harvest’. Since the implemented codes of [46] are not yet available, we have compared our results with the remaining two best performing method in general [10, 11]. Our proposed method is especially effective for the wrinkle regions (with ellipses). In summary, outperforming results presented in this section justify our theoretical arguments presented in section III and IV.

B. Evaluation with Measured BRDF

The truncated photometric ratio approach is evaluated using

the synthetic images of a sphere generated from the MERL BRDF database [22]. Light vectors for generating the synthetic images should be chosen carefully, because they can significantly affect the experimental results. Reference [45] suggested that using uniformly distributed light directions covering the whole hemisphere achieves optimal performance. Therefore, we adopt the same 100 illumination vectors as in the references [45, 46] for rendering the synthetic images. This also allows fairness in comparison with the experimental results presented in [45, 46]. 100 synthetic images are generated for 100 materials. Our method is compared with methods in references [10, 11, 45, 46]. Since the intensity images are synthetically generated, sensor noise, shadows, and unknown distortions such as inter-reflections do not exist. Unlike real images, deep dark pixels are relatively unreliable. Therefore, our proposed IRF method cannot be fairly compared for such synthetic imaging situation. For methods [10, 11] and our method, we have adopted a simple filtering method on observed intensity values, which select the non-zero darkest pixels. This is the best method to filter out specular reflections effectively. The model parameters of methods [10, 11] are the same as detailed in their paper, which has been optimized for the same dataset. The observations selection parameter is optimized individually for both methods and each material. For method [45], we have reproduced the results using the available implemented codes and parameters. For method [46], since the implemented codes are not yet available, the performance are simply compared using the average errors presented in their paper to ensure fairness in the comparison (Table VII). The mean angular errors between the estimated and ground truth surface normal vectors are compared for the 100 different materials. Figure 4 shows the evaluation results with mean angular errors. Since this dataset is synthetic, the effectiveness of the proposed method is limited by having no unknown distortions which exist in the real world. Despite such limitations, the proposed truncated photometric ratio approach still outperforms all state-of-the-arts methods.

C. Qualitative Evaluation with Real Fingerprints

This sub-section presents supportive experimental results via visual inspections. Fingerprints contains fine patterns of ridges

and valleys and the surface properties of human skin are non-Lambertian. Therefore, an accurate reconstruction of 3D fingerprints is a challenging task. Similar to as in [10] and [11],

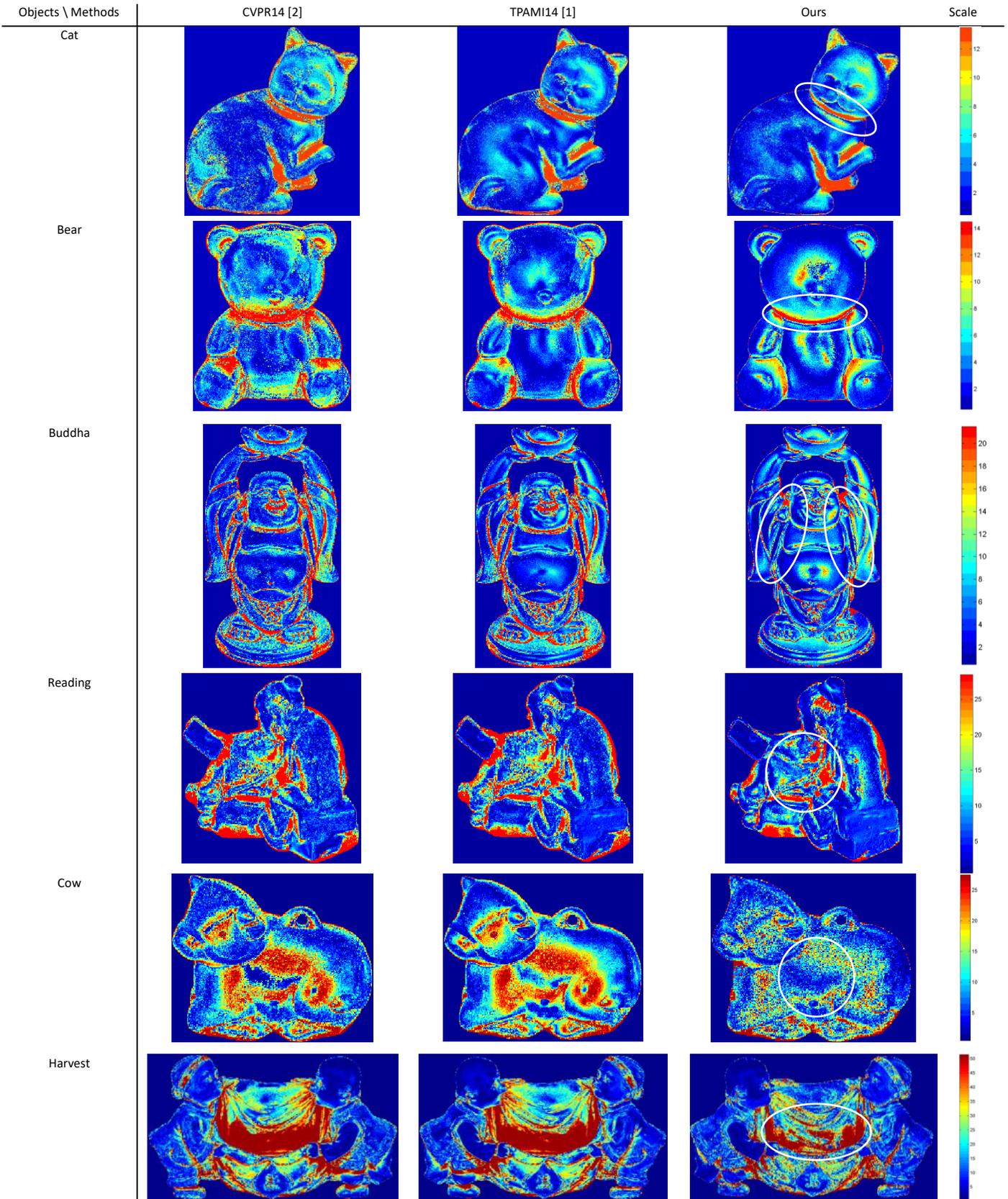


Fig. 3. Images of angular errors for the best performing methods for the ten objects in DiLiGenT [3]

we also performed qualitative experiments on a real dataset without ground truth surface normal vectors to examine the effectiveness of our method. We have comparatively evaluated the TPR approach with both the traditional least square approach (LS), which is also adopted in a recent fingerprint

research [23], [39] and CBR [11] on a fingerprint dataset [39]. Since there are only six intensity values for a pixel, the effect of filtering observed intensity values is not visible through visual analysis. Therefore, all observed intensity values are used. The estimated surface normal vectors are then integrated using the

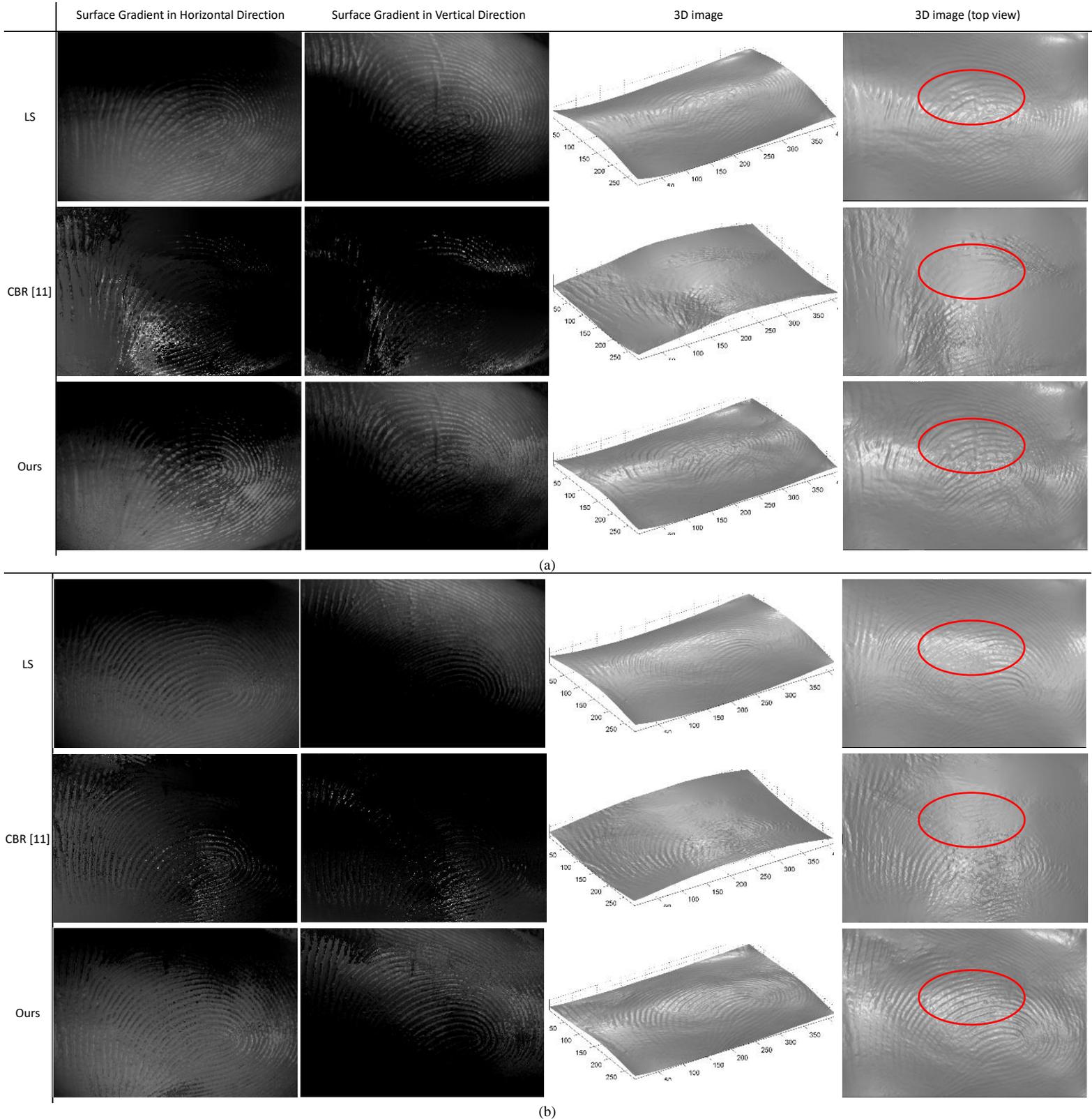


Fig. 5. Sample images of reconstructed surface gradient and 3D images on a 3D fingerprints dataset [39] (two samples)

algorithm in [21]. Since ground truth surface normal vectors are unavailable, we present the surface gradient images and 3D images for visual analysis. Figure 5 shows some images using three different methods (two samples). It can be observed that CBR (one BRDF modelling method) degrades the results with smaller number of intensity values. Our method illustrates outperforming results over the baseline by presenting more details (ridge patterns) in the reconstructed images.

V. CONCLUSIONS AND FURTHER WORK

Non-Lambertian phenomena during the photometric stereo based imaging of real objects can result from specular reflections, shadow pixels, varying BRDFs for different illumination directions, sensor noises, or even from inter surface reflections. Unlike the earlier methods that only address specific limitations from non-Lambertian surfaces, our work has been motivated to simultaneously account for such undesirable limitations, by identifying the data which are more likely to be influenced by such non-Lambertian phenomena. Firstly, the observed intensity values with less reliability are automatically eliminated. The reliability is determined by the responses from a newly introduced inter-relationship function in Section II of this paper. Secondly, those photometric ratio equations which are less likely to be Lambertian are identified. Whether the equations are likely to be Lambertian or not is decided by the residue of the equations. By eliminating non-Lambertian data, surface normal vectors are more accurately estimated. Systematic experiments for both techniques (inter-relationship function and truncated photometric ratio method) are performed on a publicly available dataset of real objects (DiLiGenT) [3] and synthetic images generated from a popular MERL BRDF database [22]. The reproducible [48] experimental results presented in Section IV validates our theoretical arguments for the effectiveness of the proposed methods. When the number of available images is small, using the photometric ratio equations approach can offer more robust and accurate results. The computational time is also short, which can offer an attractive alternative for real-world applications on mobile platforms. The approach presented in this paper can be further improved by adaptive selection of parameters, reduction in computational complexity and is part of further work in this area.

REFERENCES

- [1] R. J. Woodham. Photometric Method For Determining Surface Orientation From Multiple Images. *Optical Engineering*, vol. 19(1), 191139, February 1980
- [2] N. Fred. Directional reflectance and emissivity of an opaque surface, *Applied Optics*. 4 (7): 767–775, 1965
- [3] B. Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo. In *Proc. CVPR*, 2016
- [4] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. *IEEE TPAMI* 32(6):1060–1071, 2010
- [5] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proc. ACCV*, 2010
- [6] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *Proc. CVPR*, 2012
- [7] N. G. Alldrin, T. Zickler, and D. J. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Proc. CVPR*, 2008
- [8] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances. In *Proc. ECCV*, 2012
- [9] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *Proc. CVPR*, 2010
- [10] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE TPAMI* 36(6):1078–1091, 2014
- [11] S. Ikehata and K. Aizawa. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In *Proc. CVPR*, 2014
- [12] W. Smith and F. Fang. Height from photometric ratio with model-based light source selection. *Computer Vision and Image Understanding*, 145(2016) 128–138
- [13] T.-P. Wu and C.-K. Tang. Dense Photometric Stereo Using a Mirror Sphere and Graph Cut. In *Proc. CVPR*, 2005
- [14] Y. Queau, R. Mecca, and J. D. Durou. Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach. In *Proc. CVPR*, 2016
- [15] N. G. Alldrin, S. P. Mallick, and D. J. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In *Proc. CVPR*, 2007
- [16] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *IJCV* 72(3):239–257, 2007
- [17] J. Ackermann, F. Langguth, S. Fuhrmann, and M. Goesele. Photometric stereo for outdoor webcams. In *Proc. CVPR*, 2012
- [18] P. Tan, S. Lin, and L. Quan. Subpixel photometric stereo. *IEEE TPAMI* 30(8):1460–1471, 2008
- [19] C. Hern'andez, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *IEEE TPAMI* 30(3):548–554, 2008
- [20] L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multiview stereo. In *Proc. ICCV*, 2003
- [21] Frankot, R.T., and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE TPAMI*, 10(4):439–451, 1988
- [22] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A data driven reflectance model. *ACM Trans. on Graph.*, 22(3):759–769, 2003
- [23] A. Kumar, and C. Kwong. Towards contactless, low-cost, and accurate 3D fingerprint identification. *IEEE TPAMI*, vol. 37:681–696, 2015
- [24] E. N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *CGIP* 18(4):309–328, 1982
- [25] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE TPAMI* 18(4):449–454, 1996
- [26] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE TPAMI* 25(10):1239–1252, 2003
- [27] M. Chandraker, S. Agarwal, and D. J. Kriegman. ShadowCuts: Photometric stereo with shadows. In *Proc. CVPR*, 2007
- [28] Y. Mukaigawa, Y. Ishii, and T. Shakunaga. Analysis of photometric factors based on photometric linearization. *JOSA A* 24(10):3326–3334, 2007
- [29] K. Sunkavalli, T. Zickler, and H. Pfister. Visibility subspaces: Uncalibrated photometric stereo with shadows. In *Proc. ECCV*, 2010
- [30] F. Verbiest and L. Van Gool. Photometric stereo with coherent outlier handling and confidence estimation. In *Proc. CVPR*, 2008
- [31] C. Yu, Y. Seo, and S. W. Lee. Photometric stereo from maximum feasible lambertian reflections. In *Proc. ECCV*, 2010
- [32] D. Miyazaki, K. Hara, and K. Ikeuchi. Median photometric stereo as applied to the segonko tumulus and museum objects. *IJCV* 86(2):229–242, 2010
- [33] T.-P. Wu and C.-K. Tang. Photometric stereo via expectation maximization. *IEEE TPAMI* 32(3):546–560, 2010
- [34] Q. Zheng, A. Kumar, G. Pan, "On accurate recovery of 3d surface normal using minimum 2d images," Technical Report No. COMP-K-20, Department of Computing, The Hong Kong Polytechnic University, Sep. 2015. <http://www.comp.polyu.edu.hk/~csajaykr/COMP-K-20.pdf>
- [35] N. G. Alldrin and D. J. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach. In *Proc. ICCV*, 2007
- [36] P. Tan, L. Quan, and T. Zickler. The geometry of reflectance symmetries. *IEEE TPAMI* 33(12):2506–2520, 2011
- [37] M. Chandraker, J. Bai, and R. Ramamoorthi. On differential photometric reconstruction for unknown, isotropic BRDFs. *IEEE TPAMI* 35(12):2941–2955, 2013
- [38] F. Romeiro, Y. Vasilyev, and T. Zickler. Passive reflectometry. In *Proc. ECCV*, 2008

- [39] The Hong Kong Polytechnic University 3D Fingerprint Images Database Version 2.0, <http://www.comp.polyu.edu.hk/~csajaykr/3Dfingerv2.htm>
- [40] C. Hernández, G. Vogiatzis, and R. Cipolla. Overcoming shadows in 3-source photometric stereo. *IEEE TPAMI* 33(2):419–426, 2011
- [41] S. Lee and M. Brady. Integrating stereo and photometric stereo to monitor the development of glaucoma. *Image and Vision Computing*, 9(1):39–44, 1991
- [42] L. B. Wolff and E. Angelopoulou. 3-d stereo using photometric ratios. In *Proc. ECCV*, 1994
- [43] R. Mecca and M. Falcone. Uniqueness and approximation of a photometric shape-from-shading model. *SIAM Journal on Imaging Sciences*, 6(1):616–659, 2013
- [44] R. Mecca and Y. Qu'eu. Unifying Diffuse and Specular Reflections for the Photometric Stereo Problem. In *IEEE Workshop on WACV*, 2016
- [45] T. Q. Han and H. L. Shen. Photometric Stereo for General BRDFs via Reflection Sparsity Modeling. *IEEE TIP*, 24(12): 4888-4903, 2015
- [46] H. L. Shen, T. Q. Han, and C. Li. Efficient Photometric Stereo Using Kernel Regression. *IEEE TIP*, 26(1): 439-451, 2017
- [47] H. S. Chung and J. Jia. Efficient photometric stereo on glossy surfaces with wide specular lobes. In *Proc. CVPR*, 2008
- [48] Weblink for downloading codes for all the algorithms in this paper, available from: <http://www.comp.polyu.edu.hk/~csajaykr/PSReal.zip>