Efficient Photometric Stereo on Glossy Surfaces with Wide Specular Lobes

CHUNG, Hin Shun

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

in

Computer Science and Engineering

Thesis/Assessment Committee

Prof. WONG Tien Tsin (Chairman)

Prof. JIA Jiaya Leo (Thesis Supervisor)

Prof. WONG Kin Hong (Committee Member)

Prof. TANG Chi Keung (External Examiner)

論文評審委員會

黃田津教授(主席)

賈佳亞教授 (論文導師)

黃健康教授(委員)

鄧智強教授(校外委員)

Abstract

This thesis presents a new photometric stereo method aiming to efficiently estimate BRDF and reconstruct object shapes from glossy surfaces. Rough specular surfaces exhibit wide specular lobes under most lighting configurations. They are ubiquitous and usually bring difficulties in both specular pixel removal and surface normal recovery. In our approach, we neither apply unreliable highlight separation nor adopt unrealibale direct specularity fitting method. An important visual cue, i.e. the cast shadow silhouette of the object, is employed to optimally recover global BRDF parameters. These parameters are then taken into our reflectance model for robustly computing the surface normals and local parameters using the iterative optimization. Using only a small number of input images, our method can faithfully reconstruct both the surface normals and textures even when the surface exhibits large overlapping hight-light regions. Our method is also efficient due to the reduced search space using the information from the cast shadow silhouette. Moreover, our method does not make any assumption of the light color and texture. So it works well even when the light color and the texture color are similar. Within the unified framework, our method can also be extended to handling object surfaces assembled with multiple materials. Using graph cuts, we can segment the surface according to material variations, and the corresponding BRDF parameters and surface shape can be recovered at the same time.

摘要

這篇文章介紹了一種全新的光度立體方法，用以快速地對光亮物件計算雙向反射分布函數(BRDF) 參數並恢復物件形狀。粗糙光亮表面在任何光暗環境都會顯示寬闊鏡面反射波瓣現象。常見的粗糙光亮表面會對光亮像素去除方法和表面法向量恢復方法帶來很大的困難。而我們的方法並不運用光亮去除步驟，也不使用不穩定的直接光亮擬合法。我們首先運用陰影輪廓的資料去優化全局 BRDF 參數，繼而再用迭代的方法去計算局部參數。我們只需要少量的輸入圖像，在大光亮區域表面情況下也能正確恢復表面法向量和紋理。我們的方法也較直接光亮擬合法有高效能。並且，因爲並沒有假定光線的顏色和紋理的顏色，我們的方法能適用於所有不同的困難情況。我們的方法同時能處理多物料表面，我們用圖割的方法去根據物料去分割表面，也能同時計算 BRDF 參數和物件法向量。在文章中的多個困難的例子將足以說明我們方法的效能。

Acknowledgments

Firstly, I would like to thank my Supervisor, Prof. Leo Jiaya Jia. If it had not been for his continual guidance, inspiration and encouragement, I would not be able to finish this project. He is an excellent mentor, not only support my research at full steam, but also care about me as a friend. Two years ago, I was starting as a fresh graduate student, Professor Jia has taught me how to do research, he has also led me to the door of the Computer Vision field. I would like to express my full gratitude to him.

Secondly, I would like to thank the present and past members of our lab : Wei Xiong, Leilei Xu, Chenjun Tao, Renting Liu, Qi Shan, Xu Li, Jianing Chen and Zhaorong Li for sharing ideas and answering my questions. And I am specially obliged to Qi Shan for his help in the formulation of the proof.

I would also like to thank Prof. Wong Tien Tsin for all his helpful advices in my term presentation and paper.

Finally, I thank my parents and family for their love and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Shape reconstruction from different lightings has long been a fundamental problem in machine vision. Photometric stereo method is the most common technique in recovering surface shape and texture from multiple input images in different lightings. In this section, we are going to introduce the photometric stereo problem on the Lambertian surfaces and non-Lambertian surfaces. Then the problem of wide specular lobes will be discussed.

1.1 Lambertian photometric stereo

Photometric stereo is first proposed in [31] to deal with Lambertian surfaces. Lambertian surface is the kind of surface from which the reflected light is independent of the view angle. In other words, the observer will perceive the same amount of light in different view angles. Lambertian property can be used to model surfaces like the cotton cloth and matte paper. At least three input images are required in photometric stereo to reconstruct the surface normal and albedo. We first describe the previous Lambertian photometric stereo algorithms:

Assume we have H images of the same object with the fixed viewpoint but in different lightings. The light vectors are $l_1, l_2 \cdots l_H$ respectively. The images are $i_1, i_2 \cdots i_H$ where $i_m(x, y)$ is the pixel intensity at position (x, y) in image m. k_m is the photometric parameter of the camera which maps the irradiance to the pixel

intensity.

$$i_m(x, y) = k_m \rho_m(x, y) N_m(x, y) \cdot L_m. \tag{1.1}$$

k and L have the same values for all images m. $\rho_m(x, y)$ and $N_m(x, y)$ only depend on the position (x, y) in images. i.e. $\rho(x, y)N(x, y) = \rho_m(x, y)N_m(x, y)$ for all m.

Therefore we can rewrite Equation(1.1) as

$$
\begin{align}
i_m(x, y) &= \rho(x, y)N(x, y) \cdot kL_m, \tag{1.2} \\
&= g(x, y) \cdot v_m, \tag{1.3}
\end{align}
$$

$$\tag{1.4}$$

where

$$
\begin{align}
g(x, y) &= \rho(x, y)N(x, y), \\
v_m &= kL_m, \\
k &= k_m.
\end{align}
$$

To recover normal and albedo, we have to recover ρ and N for all pixel locations (x, y). We first construct the linear system:

$$I(x, y) = VG(x, y), \tag{1.5}$$

where

$$
\begin{align}
I(x, y) &= \{i_1(x, y), i_2(x, y) \cdots i_H(x, y)\}^T, \\
V &= \{v_1, v_2 \cdots v_H\}^T.
\end{align}
$$

Since the light vectors, i.e. L_m can be calculated from the experimental setup, we assume k is one. Then V is known. $G(x, y)$ in Equation(1.5) can therefore

be solved if $H >= 3$. This is the least number of images required for conventional photometric stereo. When $H > 3$, the least square solution of G(x,y) can be obtained.

For convenience, we express $I(x,y)$ as I,$G(x,y)$ as G. From Equation(1.5), we get

$$I = VG, \tag{1.6}$$

$$V^T = V^TVG, \tag{1.7}$$

$$G = V^*I, \tag{1.8}$$

where $V^* = (V^TV)^{-1}V^T$ is the pseudoinverse of V.

G(x,y) is evaluated for every image position (x,y) based on Equation(1.8). Unit normal and albedo reflectance for gray scale images can be estimated from the following equations:

$$N(x,y) = \frac{G(x,y)}{\|G(x,y)\|}. \tag{1.9}$$

$$\rho(x,y) = \|G(x,y)\|. \tag{1.10}$$

Lambertian photometric stereo is well-established and widely used. But simple Lambertian model fails to model many natural objects, limiting the practical use of it. Therefore many research effort has been put into non-Lambertian photometric stereo in recently years.

1.2   Non-Lambertian photometric stereo

Most surfaces can be modeled as a combination of a Lambertian component and a specular component. This can be illustrated by the simplest Phong model:

$$I = \rho_d n \cdot l + \rho_s (v \cdot s)^n \tag{1.11}$$

where ρ_d and ρ_s are the diffuse and specular albedo respectively, l, v and s are light direction, viewing direction, and specular direction respectively. $\rho_d n \cdot l$ is the

Lambertian component which is independent of the view angle, $\rho_s(n \cdot s)^n$ is the specular component which is nonlinear and depends on the view angle. n is the roughness coefficient which controls the size of the specular lobe.

Many recent photometric stereo methods [16, 23, 2, 10] were proposed to deal with non-Lambertian surfaces. In order to handle complex reflection properties, most previous approaches either require that highlight regions in input images are relatively small and can be easily removed, or assume that the highlights in different input images do not overlap such that Lambertian rule can be applied to the surface normal recovery. The details of these non-Lambertian photometric stereo algorithms can be found in Chapter 2.

1.3   Large specular lobe problems

An ideal specular surface is perfectly smooth and behaves like a mirror. It reflects light only alone the specular direction, obtained by reflecting the incident radiation direction about the surface normal. But most surfaces are not perfectly smooth, and can only be modeled as a layer of microfacets with varying normals in microscopic level. Each microfacet reflects light in a direction, and the distribution of the normals of microfacets determines the roughness of this specular surface, i.e. rough surfaces have more microfacets whose normals deviate with large angles from the surface normal. Figure 1.1(b) shows an microfacet illustration of a smooth specular surface. The normals of the microfacets point to the similar directions. Thus the reflected light will be in the similiar directions, causing a narrow specular lobe. On the other hand, the microfacets of the rough specular surface in Figure 1.1(a) have very dissimilar normals. The reflected lights spread within a wider range of angles. Therefore the wide specular lobe is produced.

When the specular surface is relatively smooth, the specular lobe is narrow and there is a sharp spike of reflected radiation, we call it specular spike. Figure 1.1(d) shows a smooth specular ball. The non-highlighted pixels can be easily separated
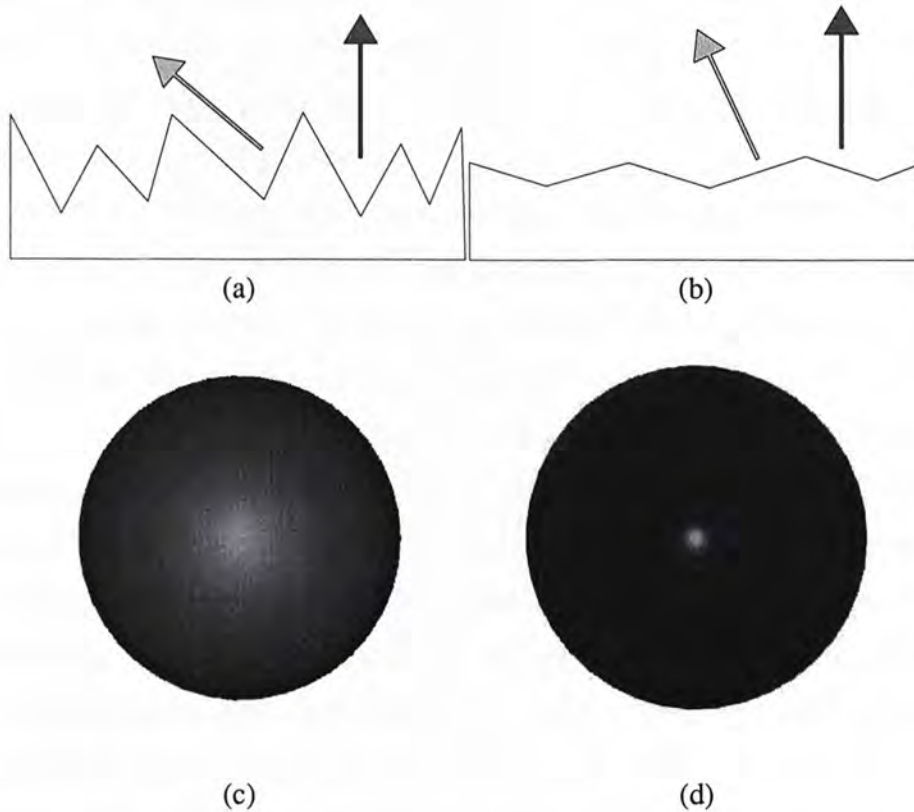
Figure 1.1: (a) and (b) illustrate the microfacets of rough specular surface and that of smooth specular surface respectively. The red arrow points to the direction of the surface normal, the green arrow points to the direction of the micro-facet normal. (c) and (d) show a rough specular black sphere and a smooth specular black sphere. Note that there is a large highlight area with fuzzy boundary on the rough specular sphere surface.

from the highlighted ones in the input images, and the non-highlighted ones then can be used in photometric stereo. Most previous work like [15], [16], [6], [23], [2], [10] is based on this idea to handle specular surfaces in photometric stereo. When the specular surface becomes rougher, and the specular lobe gets larger like the ball shown in Figure 1.1(c) , it will be problematic to those methods. Large specular lobe means a large highlight area and there is a blurred boundary between the non-highlighted area and highlighted one. Large highlight will overlap in some input images. In photometric stereo, at least three corresponding unshadowed

non-highlighted pixels (without any specular components) in three input images are needed to reconstruct one pixel normal. In a set of rough specular object images, three non-highlighted pixels may not be obtained since some points are under high-light in most light configurations.

Obviously, specular component removal is not easy on objects with rough specular surfaces. The micro-facets on rough surface reflect light differently, producing a wide specular lobe with highly blurred boundary under directional light. Four examples are shown in Figure 1.2. Moreover, it is usually not easy to determine whether the reflected light from the surface contains any specular components. An imprecise specularity segmentation makes the surface normal recovery error-prone. In order to remove the wide specular lobe on surface, Mallick *et al.*[14] proposed a data-dependent specular separation method using color space. It requires that the diffuse color of the surface and the color of incident light are not similar. Georghiades [9] fits specularities from the input images to reflectance model in order to directly estimate the normal and BRDF parameters. This method does not work well in estimating complex object surfaces due to the inherently huge search space of the surface reconstruction problem.

In this thesis, we present an efficient photometric stereo method to robustly estimate both surface normals and BRDF parameters. Our method does not require specularity separation in order to recover surface normals, thereby is capable of automatically and precisely reconstructing convex object surfaces with different level of roughness from a small number of images. The object surface is allowed to have complex textures and colors without influencing the computation accuracy.

We adopt a two-step optimization to precisely compute the normals and all reflectance parameters. In the first step, we use the visual information from shadow to estimate the normals of some surface points. These recovered normals play a key role in estimating the global reflectance parameters. In the second step, with the reduced number of unknowns, a robust iterative optimization process is applied to estimate the surface normals and diffuse albedo for all surface points. This two-step

(a) (b)

(c) (d)
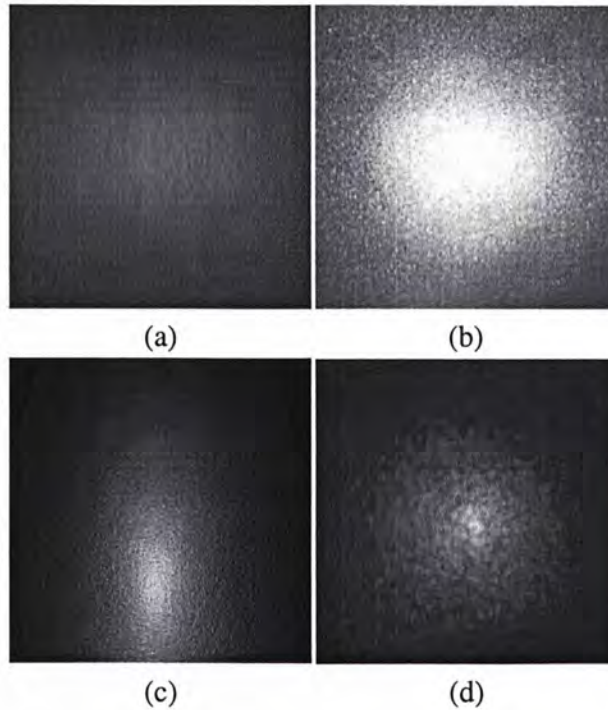
Figure 1.2: Rough surface examples. The specular lobes on the surfaces are wide under point light source and there is no clear boundary of the lobes.

approach is proven to be more efficient and robust than direct specularities fitting methods.

In the unified framework, our method can also be applied to dealing with surfaces assembled with several different materials. The material segmentation, parameter estimation and normal computation can be achieved automatically in our approach. Moreover, our method is the first attempt to incorporate the sparse shadow information in the surface normal and BRDF parameter estimation. We will show that the shadow boundary provides useful information for computing the surface normals.

The rest of this thesis is organized as follows. Section 2 first reviews the related work of both Lambertian and non-Lambertian photometric stereo. Then, in Section 3 we present our photometric stereo method to construct single material surface and

to estimate BRDF parameters. We compare our method and the direct specularity fitting method in Section 4. The multiple managerial surfaces reconstruction is described in Section 5, then followed by the conclusion in Section 6.

Chapter 2

Related Work

2.1 Lambertian photometric stereo

In pioneer work [31], Woodham first proposed photometric stereo method to estimate surface normal and diffuse albedo of Lambertian surface in 1980. It started the wide research field of photometric stereo. While the Lambertian photometric stereo is widely used, many work have been done to tackle surfaces with specularities.

2.2 Non-Lambertian photometric stereo

Several different approaches have been investigated by researchers to reconstruct non-Lambertian surfaces using photometric stereo. They are described in the following subsections.

2.2.1 Analytic models to reconstruct non-Lambertian surface

[24] developed the theory for a class of non-Lambertian surfaces.Most related work on non-Lambertian surface reconstruction uses various bi-directional reflection distribution function (BRDF) models. [15] used the Torrance-Sparrow model to recover both the surface normal and parameters of the reflectance models, but it assumed that surfaces are smooth and specular pixels can be treated as outliers and

9

can be separated from the pure Lambertian ones. In this case, the separation of Lambertian pixels and the specular ones is not difficult since the specular lobe is narrow.

Tagare and deFigueiredo [25] estimated both the shape and reflectance map for diffuse surfaces which consist of forescatter lobes, normal lobes and backscatter lobes. In our work, we estimate the shape and reflectance of surfaces containing both specular and diffuse components.

[6] used four combination of three light sources to compute four albedo values, the difference in one of these albedo values helps to detect which pixel is in specular lobe. This method only need four images. [23] extended [6] and used Torrance-Sparrow model to compute the surface roughness. [2] also refined [6] to handle the presence of shadows and used color information to help the detection of specular pixels. But all these methods assumed that there is no specular component in at least three observations for each surface point, and this assumption is usually valid in surfaces with narrow specular lobes.

Recently, [10] computed both shape and spatially-varying BRDFs of target object using photometric stereo. It assumed that the object surface can be modeled by a linear combination of two fundamental materials, they used the shape obtained by conventional photometric stereo as the initialization, and alternatively optimized the BRDF parameters and shape until convergence. Again, they assumed the specular lobe is narrow so they can remove specular highlight pixels by intensity thresholds.

[32] used a set of more than one thousand of photometric stereo images to compute the surface normal using graph cut algorithms. The thesis demonstrated recovery of complex surfaces with discontinuities, highlight and transparency, but the application of this method is limited by its huge number of input images.

2.2.2   Reference object based

[12] reconstructd the shape of object with arbitary BRDF with the help of a reference object with the same material. Orientation consistency enabled them to establish the correspondence between the known shape reference object and the target object.

2.2.3   Highlight removal before shape reconstruction

In order to robustly reconstruct surface, the specular components, i.e., the highlight, are usually required to be removed such that traditional photometric stereo method on Lambertian surface can be applied. Most highlight removal methods are based on dichromatic reflection model [22].[13] separated highlights from diffuse components by considering RGB color space.

Sato *et al.*[20] proposed a metod based on the color difference between diffuse and specular components to separate components and estimate reflectance.

[14] introduced a data-dependent SUV color space, it is obtained by rotating the RGB color space such that R channel is aligned with the color of the light source (it assumed light source color is the same as the highlighted color), then the BG channels contain only diffuse color components which can be used in photometric stereo. But it does not work on textured objects whose body color is similar to the light source color.

[21] separated diffuse and specular components of a homogeneous dichromatic surface by using color histogram analysis. [19] used more light sources direction to obtain color histogram for each surface point and then work on inhomogenous surface, but they are all limited by the problem of similarity of light source color and body color. They also assumed the specular lobe is narrow with respect to the surface curvature.

In [26, 28, 27, 34], highlight components were removed on single image. Most of these methods assume that the diffuse color does not change too much inside

and outside the highlight region or the specular color is different from the surface texture. When these assumptions are violated, the recovered diffuse color is noisy or incorrect.

Most of previous image-based highlight removal methods have assumptions in colors, shape, or brightness. When the specular surface becomes rough, the specular lobe is getting wide, largely blurring its boundary. This causes problems in determining the specular component in the reflected light. Consequently, the highlight removal methods may fail on these objects especially when the surface has complex textures.

2.2.4 Polarization based method

[30],[17] make use of different polarizing angles to separate diffuse and specular components, but specular lobe is assumed to be narrow and filters are needed to get the polarization images.

2.2.5 Specularity fitting method

Georghiades [8] showed that the specularities of surfaces can be used to disambiguate uncalibrated photometric stereo. In [9], he proposed to recover shape and reflectance by fitting the specularities to Torrance and Sparrow reflection model. This method iteratively optimizes over a very large set of variables, including reflectance parameters, light directions, and surface normals. To avoid local optimum due to the large search space, a good initialization of all unknowns are generally required. Direct specularity fitting also does not work well on complicated surfaces with wide specular lobes, which we shall show examples in Chapter 4.1. Besides, the efficiency is also an issue for high-dimensional optimization.

2.2.6   Photometric stereo with shadow

In [4] and [1], shadow clue was used to constrain the depth construction of object surface. In our method, the shadow information is explored in a novel way, that is using the cast shadow boundary normal to assist recovery of surface normal and BRDF measurement.

Chapter 3

Our System

Similar to the configuration of calibrated photometric stereo methods in previous work, our input image set $\mathbf{I} = \{I_1, I_2, ..., I_n\}$ are taken by a static camera where the target object is illuminated by varying directional light source. We put light source far from the object such that parallel light can be assumed, in our experiments, the distance between the object and light source is about 2.5m. Input images are assumed to be taken under orthographic projection. The photometric calibration is carried out such that the pixel intensities represent the radiance. The incident radiance direction in each image is measured using a chrome sphere.

For simplicity's sake, we first assume that the object is built of a single material. Our method can be naturally extended to the objects assembled with different materials, which will be described in Chapter 5.

We model the reflectance property of the object's glossy surface using the Ward BRDF model [29]:

$$f(\mathbf{i}, \mathbf{o}) = \frac{\rho_d}{\pi} + \frac{\rho_s}{4\pi\alpha^2\sqrt{\cos\theta_i \cos\theta_o}}\exp^{\frac{-tan^2\beta}{\alpha^2}}, \qquad (3.1)$$

where i and o denote the incident and outgoing light directions. θ_i and θ_o are polar angles of i and o respectively. α is the roughness coefficient, determining the size of the specular lobe on object surface. Larger α value causes wider specular lobe with severely blurred boundary. ρ_d and ρ_s measure the diffuse and the specular reflectance respectively. β is the angle between surface normal and halfway vector

14

[29]. Note that other parametric reflectance models can also be employed similarly in our method.

Ward model is a combination of two terms. $\frac{\rho_d}{\pi}$ is the diffuse term. Other parameters form the specular term which is highly non-linear with respect to the unknowns. Directly optimizing all the unknowns in (3.1) is usually a slow and unstable process. In our method, we separate the unknowns into two classes and estimate them separately:

- The global parameters ρ_s and α have fixed values for all the surface points and they can be estimated in the first pass. For computation efficiency, a few selected samples are sufficient to robustly estimate these parameters. This step is described in Chapter 3.1.

- The estimated global parameters are taken into (3.1) to simplify the specular term. We then apply iterative optimization to computing the surface normals and other local parameters respectively on each pixel. This step is described in Chapter 3.2.

3.1 Estimation of global parameters

We first show that if the normal of one surface point is known on the object, the global parameters can be computed. In order to use the normal information, we first study the BRDF measurement on object surface using single directional light source. The amount of outgoing light, L_o, is given by:

$$L_o = BRDF(\theta_i, \phi_i, \theta_o, \phi_o)L_i \cos \theta_i, \tag{3.2}$$

where L_i is the incoming light. Substituting (3.1) into (3.2), and generally assuming the unit intensity of the incoming light, we obtain

$$L_o = \frac{\rho_d}{\pi} \cos \theta_i + \frac{\rho_s}{4\pi\alpha^2} \sqrt{\frac{\cos \theta_i}{\cos \theta_o}} \exp^{\frac{-\tan^2 \beta}{\alpha^2}}. \tag{3.3}$$

Note that if the corresponding normal of one surface point q is known (the details of deriving this normal will be depicted in Chapter 3.1.3 and 3.1.4), the formulation of L_o can be greatly simplified since θ_o, θ_i, and β are determined. We can estimate other BRDF parameters by minimizing

$$g(q) = \sum_{0 \leq i < n} (I_i(q) - L_o(q, \rho_s, \alpha, \mathbf{n}, \rho_d))^2, \tag{3.4}$$

where I_i is one of the input images. In the rest of the thesis, without causing ambiguity, we simplify the notation of $L_o(q, \rho_s, \alpha, \mathbf{n}, \rho_d)$ to $L_o(q)$. We propose the following iterative optimization algorithm to minimize g and estimate the BRDF parameters.

1. **Initialization of BRDF parameters.**

   α is initialized to a value between 0 and 0.4 while ρ_d and ρ_s are initialized to values between 0 and 1.

2. **Optimize ρ_d by fixing other parameter values.**

   In the diffuse term, ρ_d is the only parameter to be estimated on the pixel. It can be exactly computed by setting $\frac{\partial g}{\partial \rho_d} = 0$.

3. **Optimize ρ_s and α by fixing ρ_d.**

   In the specular term, ρ_s and α are optimized together using the Levenberg-Marquet method.

4. **Termination.**

   Step 2 and 3 are iteratively performed until the energy of objective function does not decrease in successive iterations.

Since the energy of g in step 2 and 3 are monotonically decreasing, so the whole process is guaranteed to converge to an optimal solution. This optimization process converges rapidly given the small number of unknowns.

The above optimization process shows that the global BRDF parameters can be efficiently estimated from one or a set of known surface normals. In the following,

we describe our method of estimating the surface normals using the shadow visual clue.

3.1.1 Shadow separation

Since directional light source is used in photometric stereo methods, we notice that the lighting in each image not only forms a specular lobe on the object but also produces cast shadow of the object on the background. These shadows are caused by the occlusion of incident light.

In order to capture the shadow of the object under different lighting, we place a small Lambertian board on the back of the object. To reduce the possible inter-reflection, the board is painted with mid-grey color to absorb most of the incident light while the shadow can still be faithfully detected. We show one example in Figure 3.1 (a). To verify that using the board does not influence our result much, we carried out several experiments with and without the grey board and found that the differences are constantly very small. One comparison is given in Figure 3.1 (a) and (b) where the root mean square (RMS) difference on the object pixels including image noise is only 0.0024.

3.1.2 Separation edges of shadow and edges of foreground object

For automatically separating the shadow and the foreground object in the input images, we observe that the shadow region changes among the images due to the change of lighting directions. Statistically, one shadow edge pixel in one image has small chance to be an edge point in all other images. The object in the image has fixed position. So its boundary and textures constantly appear in multiple images. We detect the edges combining multiple images using the method presented in [33] where a two-component GMM is constructed for each pixel to robustly reject outliers. In the following, the method to robustly extract the edges of the shadow and
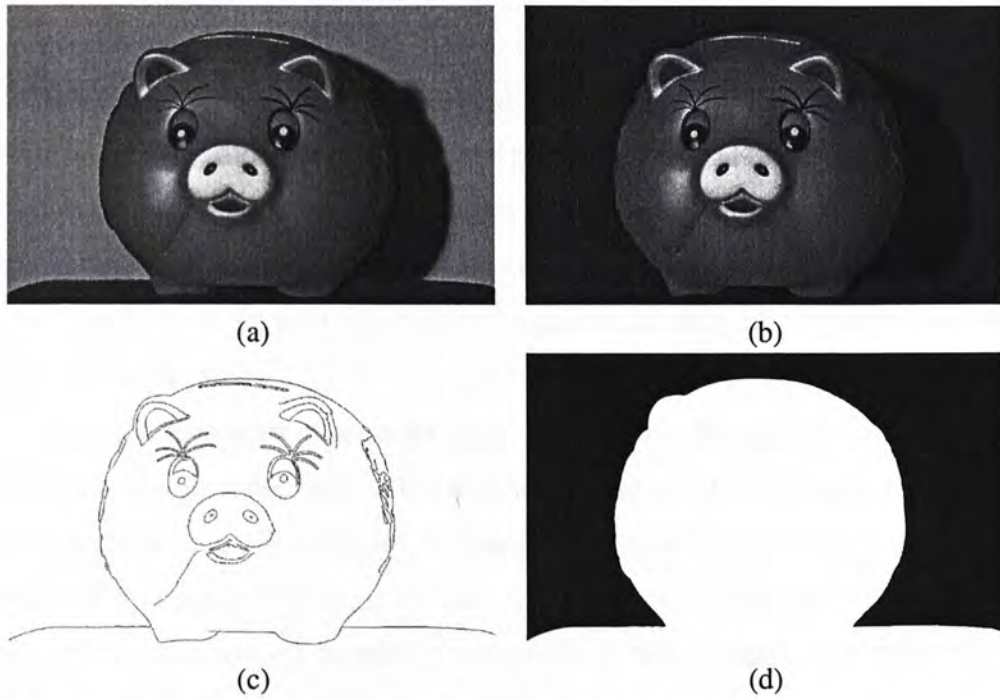
(a)   (b)

(c)   (d)

Figure 3.1: Mid-grey backboard configuration. (a) A mid-grey board is placed be-
hind the object to capture the cast shadow. The reflection from the board is low. (b)
Same scene captured with black background. The object color difference between
(a) and (b) is very small. (c) and (d) show the automatic foreground extraction. (c)
The edge detected using input image information. Most shadows boundary pixels
are removed. (d) Using level set method, the foreground map is formed on the edge
map by curve evolving from the background.

that of the foreground object is introduced.

First, to remove high frequency noise from input image I, the input image is
convolved with a low pass Gaussian filter G_σ to obtain the smoothed image I'.

$$I' = G_\sigma * I \tag{3.5}$$

The gradients along x-axis and y-axis of the smoothed image I' can be obtained by
:

$$Q_x = \frac{\partial G_x}{\partial x} * I' \tag{3.6}$$

$$Q_y = \frac{\partial G_y}{\partial y} * I' \tag{3.7}$$

And the magnitude w and orientation θ of the gradient for each pixel position can be computed by $w = \sqrt{Q_x^2 + Q_y^2}$ and $\theta = tan^{-1}(\frac{Q_y}{Q_x})$.

Two component gaussian mixture model (GMM) is used to model the local gradient orientation distribution at each pixel position. The main Gaussian distribution models the edges of the foreground object (inliers) while the outliers are the edges due to shadow or image noise, which are captured by the second Gaussian distribution. The Expectation Maximization can be used to decompose the mixture model and identify the inliers.

Let the total number of orientation bins be K and denote bin i as (θ_i, M_i) where θ_i and M_i are the center orientation and the magnitude of bin i respectively. For a pixel position (x, y), we collected N gradient samples within the sampling window centered at (x, y) from all input images. And the gradient magnitude w and orientation θ for each gradient sample are computed, gradient sample j belongs to bin i if $\theta_i - W_\theta/2 < \theta_j < \theta_i + W_\theta/2$ where W_θ is the width of the bin. Membership function m_{ij} of sample j with respect to bin i is set to 1 if sample j belongs to bin i, else set to 0. Then magnitude of bin i is then computed by $M_i = \sum_{j=1}^{N} w_i m_{ij}$.

The two-component GMM can be formulated as:

$$g(\theta) = \sum_{k=1}^{2} \alpha_k q_k(\theta|m_k, \sigma_k) \qquad (3.8)$$

where α_k is the mixing coefficient, q_k is the probability distribution function of the k'th component Gaussian distribution and it is defined as:

$$q_k(\theta|m_k, \sigma_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} exp(-\frac{min(|\theta - m_k|, 360 - |\theta - m_k|)^2}{2\sigma_k^2}) \qquad (3.9)$$

EM algorithm is used to iteratively update the parameters in the mixture model until convergence. The parameters are updated at each iteration according to these functions:

$$q_i(k|\theta_i, \Theta^{(t)}) = \frac{\alpha_k^{(t)} q_m(\theta_i|m_k^{(t)}, \sigma_k^{(t)})}{\sum_{l=1}^{2} \alpha_l^{(t)}, m_l^{(t)}, \sigma_l^{(t)}} \qquad (3.10)$$

$$\alpha_k^{(t+1)} = \sum_{i=1}^{K} \frac{M_i}{\sum_{l=1}^{K} M_l} q_i(k|\theta_i, \Theta^{(t)}) \tag{3.11}$$

$$m_k^{(t+1)} = \frac{\sum_{i=1}^{K} M_i q_i(k|\theta_i, \Theta^{(t)}) cos\theta_i}{\sum_{i=1}^{K} M_i q_i(k|\theta_i, \Theta^{(t)}) sin\theta_i} \tag{3.12}$$

$$\sigma_k^{(t+1)} = \frac{\sum_{i=1}^{K} M_i q_i(k|\theta_i, \Theta^{(t)})(D_{ik}^{(t)})^2}{\sum_{i=1}^{K} M_i q_i(k|\theta_i, \Theta^{(t)})} \tag{3.13}$$

where k represent the k'th component Gaussian distribution and k can be 0 or 1. $\Theta^{(t)} = \{\alpha_k^{(t)}, m_k^{(t)}, \sigma_k^{(t)}\}$ is the parameter set at iteration t. $D_{ik}^{(t)}$ is the circular distance between θ_i and $m_k^{(t)}$ and is defined as:

$$D_{ik}^{(t)} = min(|\theta - m_k^{(t)}|, 360 - |\theta - m_k^{(t)}|) \tag{3.14}$$

The EM algorithm produces two set of parameters $(m_1, \alpha_1 \sum_{i=1}^{K} M_i)$, $(m_2, \alpha_2 \sum_{i=2}^{K} M_i)$ for each pixel, suppose $\alpha_1 > \alpha_2$, we say component 1 is the main component at this pixel, then the gradient orientation mean of the main component m_1 captures the orientation of the foreground object edge in each pixel, its corresponding magnitude is $\alpha_1 \sum_{i=1}^{K} M_i$. The composite gradient map of the foreground object can be obtained from extracting the gradient orientation mean of all main components at all pixel positions.

After that, the postprocessing techniques of Canny edge detector are employed to improve the result. Non-maximum suppression is first applied to the gradient map, such that each pixel is set to zero unless it is a local maximum along a line oriented along the gradient direction. Then hysteresis is then used to eliminate streaking. Two thresholds are used, T_h and T_l where $T_h > T_l$. Any pixel has a value greater than T_h is set to edge pixel immediately, then any pixels connected to edge pixels and have a value greater than T_l are set to edge pixels.

After these postprocessing steps, the output is the edge pixels detected on the object as shown in Figure 3.1 (c).
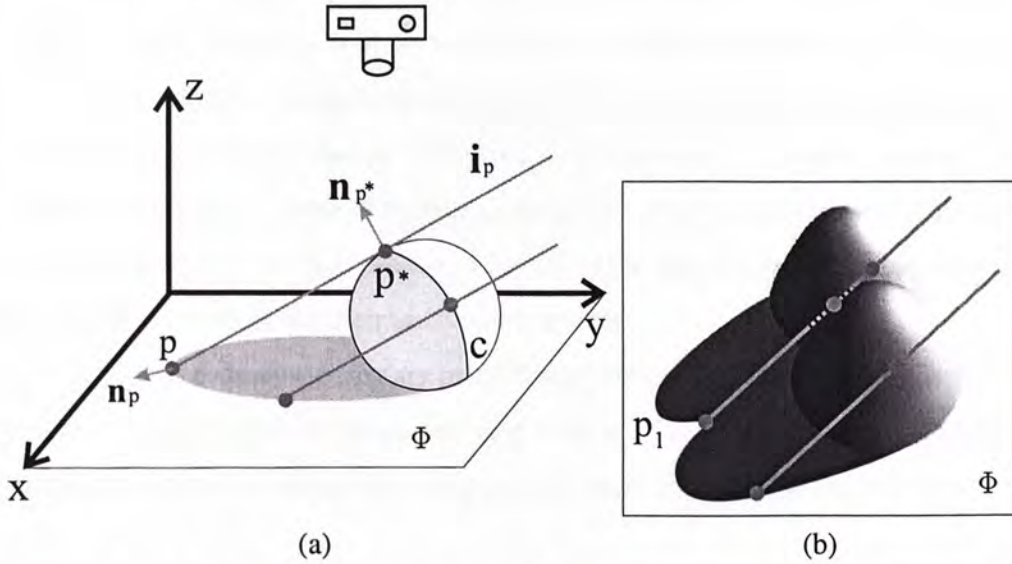
Figure 3.2: (a) Shadow formation. $\mathbf{i}_p$ shows the lighting direction. The object casts shadow on the back plane Φ. The camera is fixed as shown. $\mathbf{n}_p$ is the shadow boundary normal of p in Φ. p^* is the corresponding surface point and $\mathbf{n}_{p^*}$ is the surface normal. (b) Shadow silhouette. p_1 is at the junction formed by two hills' cast shadow. The gradient computed on p_1 does not map to any surface normal.

To extract the object in each image, the smallest closure of the edge points can be computed using flood fill or the geometric level set method [18]. One example is shown in Figure 3.1 (d).

When the object has been separated, shadow remains in the images. Due to the use of directional light, the shadow penumbra is narrow. We, thus, can detect shadow boundary using edge detection.

3.1.3 Normal estimation using shadow boundary

Once the shadow boundary pixels are detected, we compute image gradients on these pixels. As shown in Figure 3.2 (a), each of the shadow boundary pixels p in back plane Φ can be mapped to at least one object surface point p^*. We study the following two classes of cast shadow boundary in order to compute the surface normal $\mathbf{n}_{p^*}$ at p^*. The illustration is shown in Figure 3.2.

Class 1: The shadow boundary pixel, e.g., p in Figure 3.2 (a), maps to a single surface point p^* along the light ray and the first order derivative exists at p^*. It can be easily proven that the surface normal $\mathbf{n}_{p^*}$ at p^* is projected to the shadow boundary normal $\mathbf{n}_p$ in continuous space. The proof is included in the Appendix section. In discrete image space, there may exist quantization errors in the computed shadow boundary gradient, we shall propose a robust estimation which will be depicted in Chapter 3.1.4 to diminish their unfavorable effects.

Class 2: The shadow boundary pixel, for example, p_1 in Figure 3.2 (b), is at the junction formed by two or more hills' cast shadow. The normal of the surface point cannot be determined since there exist at least two surface points projected to that pixel.

In order to estimate the surface normal using cast shadow, the pixels in class 2 should be rejected. We employ the feature detection method to filter out the shadow boundary pixels at junctions. We construct the Harris matrix $M(x, y)$ [11] on each shadow boundary pixel $I(x, y)$:

$$M(x, y) = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{3.15}$$

where $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$ and $w(u, v)$ is a Gaussian filter with $\sigma = 1$.

The junctions are detected when the corner strength $S(x, y) = \frac{\det M(x,y)}{\operatorname{tr} H(x,y)}$ is larger than a threshold.

After feature removal, we compute the gradients on the remaining cast shadow boundary pixels. We denote the remaining shadow gradient set as $\{\mathbf{n}_1^i, \mathbf{n}_2^i, ... \mathbf{n}_{m_i}^i\}$ in each image I_i. The corresponding surface normal for $\mathbf{n}_p^i$ is $\mathbf{n}_{p^*}^i$. $\mathbf{n}_{p^*}^i$ is computed by applying the following constraints.

- $\mathbf{n}_{p^*}^i$ is a surface normal on the self-shadow boundary, its direction must be perpendicular to the lighting direction $\mathbf{i}_p$:

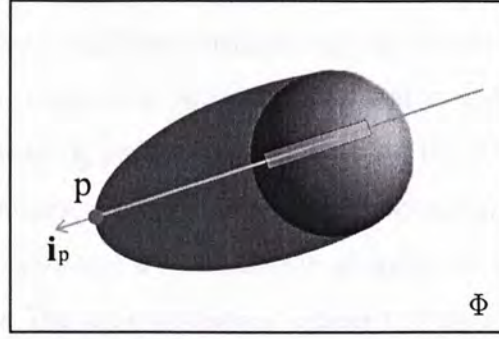$$\mathbf{i}_p \cdot \mathbf{n}_{p^*}^i = 0. \tag{3.16}$$

Figure 3.3: Determining the surface points. On object surface, the point p^*, on which the normal $\mathbf{n}_p^*$ is computed, is only known along the ray $\mathbf{i}_p$, as shown in green.

- $\mathbf{n}_{p^*}^i$, $\mathbf{n}_p^i$ and $\mathbf{i}_p$ are coplanar, we thus have

$$(\mathbf{i}_p \times \mathbf{n}_p^i) \cdot \mathbf{n}_{p^*}^i = 0. \tag{3.17}$$

- $\mathbf{n}_{p^*}^i$ is a unit vector

$$||\mathbf{n}_{p^*}^i|| = 1. \tag{3.18}$$

Combining (3.16), (3.17), and (3.18), we solve a set of linear equation systems to compute the surface normals corresponding to the shadow boundary pixels.

Note that our feature removal may not be sufficient to reject all shadow pixels in class 2 in computing surface normals. In addition, the possible noise produced by inaccurate image gradient and shadow boundary computation may also contaminate the results. In the next Chapter, we introduce a robust clustering method to further refine the parameter estimation.

3.1.4 Global parameter estimation and refinement

We have recovered a set of possible surface normals corresponding to the cast shadow boundary pixels. In order to robustly estimate the BRDF parameters in (3.1), the positions of corresponding surface points should also be known so that (3.4) can be optimized.

In most cases, the attached shadow on object fades gradually and no clear boundary exists between the illuminated area and the attached shadow region. One example is shown in Figure 3.3. We only know that p^* corresponding to normal n_{p^*} is along the light ray i_p passing through p. It is difficult to determine its exact position. In our approach, a small number of surface candidates S_p of p^* along i_p are collected. Each candidate is in the transition region between fully illuminated and shadowed pixels. The region is shown in green in Figure 3.3.

Notice that using shadow boundary, we compute several surface normals in image i, each normal $n_{p^*}^i$ has a set of candidates S_p^i of its possible position. To simultaneously find the true normal position and to estimate the global BRDF parameters, we introduce the following optimization process.

1. For each surface normal $n_{p^*}^i$:

   (a) We select one candidate normal position from S_p^i and take it into the iterative optimization described in Chapter 3.1 to minimize (3.4). The residual error is also recorded.

   (b) If there exists candidate in S_p^i that is not used in the above computation, repeat (a).

   (c) If one surface point from all candidates in S_p^i produces the smallest residual error, we record the BRDF parameters $\rho_s(i, p)$ and $\alpha(i, p)$ produced by using that surface point.

2. For all recorded BRDF parameters ρ_s and α, we apply the mean shift clustering [7] to robustly estimate parameter values.

Regarding step 1, among all candidates for a normal position, we regard the one producing the least residual error in the optimization as the true surface point. It is proven to be a reliable measurement in our experiments. The computation is also efficient.

Step 2 is based on observation that there may exist small noise and outliers. We use the statistical clustering method to refine the parameter estimation. The correctly estimated parameter values form high density in the parameter space. For a few normal outliers, the correspondingly estimated parameters have small probability to be similar to each other. So the point density around these values is small.

3.2 Surface shape and texture reconstruction

After estimating the specular parameters, surface normals and diffuse albedo for all surface points can be reconstructed without highlight separation. The surface texture is the recovered diffuse albedo. We first rewrite (3.3) as

$$L_o(p) = \frac{\rho_d(p)}{\pi} \mathbf{i} \cdot \mathbf{n}(p) + \frac{\rho_s}{4\pi\alpha^2} \sqrt{\frac{\mathbf{i} \cdot \mathbf{n}(p)}{\mathbf{o} \cdot \mathbf{n}(p)}} \exp^{\frac{1}{\alpha^2}\left\{1 - \frac{1}{[\mathbf{h}(p) \cdot \mathbf{n}(p)]^2}\right\}},$$

where $\mathbf{i}$ and $\mathbf{o}$ are the incident and outgoing light directions respectively. $\mathbf{h}$ and $\mathbf{n}$ are the half-way vector and surface normal respectively. We minimize the energy function $g(p)$ similar to (3.4) to reconstruct surface normals:

$$g(p) = \sum_{0 \le i < n} (I_i(p) - L_o(p))^2. \tag{3.19}$$

Comparing to previous methods constructing normals using only Lambertian component [31, 6, 23, 14], our formulation has the same number of unknowns. We employ an iterative optimization to estimate them:

1. For each pixel q on object surface:

   (a) The normal $\mathbf{n}_q$ of q is initialized as the interpolation of the recovered normals in neighborhood.

   (b) **Optimize $\rho_d(p)$.** $\rho_d(p)$ can be exactly computed by setting $\frac{\partial g(p)}{\partial \rho_d(p)} = 0$.

   (c) **Optimize normal $\mathbf{n}_q$.** The normal vector is estimated using gradient descending optimization. Given the small number of unknowns, it converges rapidly and produces optimal result.

(d) Repeat (b) and (c) until convergence.

The energy in steps (b) and (c) is monotonically decreasing, so the convergence is guaranteed. In all our experiments, the process converges within 10 iterations.

3.3 Single material results

We show experimental results in this section. Most of the surfaces contain wide specular lobes in input images. Since our algorithm does not rely on separation of specular components, we can faithfully recover surface textures even if the highlight color is similar to the texture color.
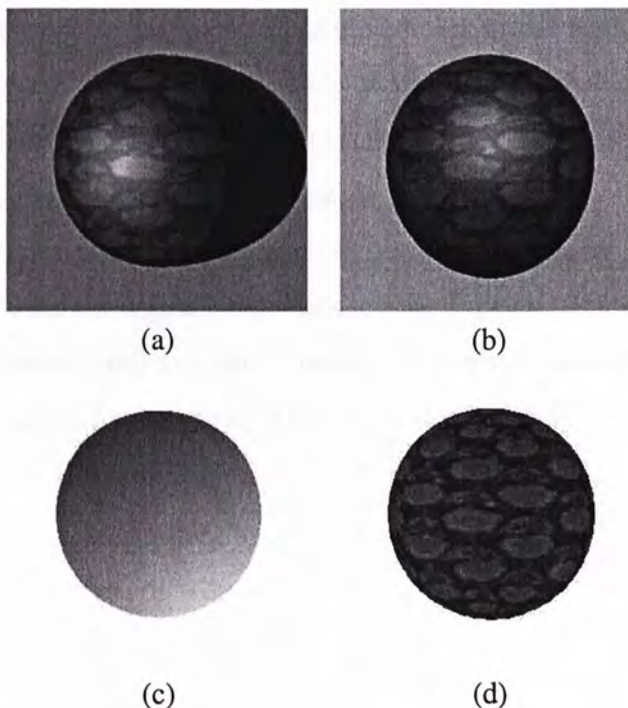

(a)   (b)

(c)   (d)

Figure 3.4: (a),(b) are two of six sphere images for reconstruction, note the large area of overlap of hightlight between them even though the lights come from very different directions. (c) Reconstructed surface normal encoded by RGB-channels. (d)Reconstructed surface texture.

The first example is a synthetic rough specular sphere with complicated texture

as shown in the top row of Figure 3.4. Note that the color of light source is similar to that of the texture on the sphere, it is an error-prone step to separate the highlight using highlight removal methods mentioned in Section 2.2.3. The large specular lobe makes a large area of hightlight and the highlight regions overlap in most of the input images. Using our method, the BRDF parameters and surface normal are faithfully recovered. The surface normal and the texture image are shown in (c) and (d). Comparing to the ground truth surface, the surface normal is very accurate, with only 0.94 degree in average error.

Figure 3.5 shows a "mouse" example with large highlight area. (a) and (b) illustrate two out of the six input images where the object and the shadow boundary can be automatically extracted. The surface normal and texture shown in (c) and (d) are faithfully reconstructed. We evaluate the efficacy of our method by rendering the object under a novel lighting condition as shown in (e) and comparing it with the ground truth image (f), the RMS error is only 0.0326, proving that our recovered BRDF parameters and surface normals are accurate.

Figure 3.6 shows another object in which the reflection is complex and highlight can be seen in different regions in the image. We only use 5 input images in this example. We faithfully recover surface normals and texture as shown in (b) and (c). The rendered image (d) under novel lighting is visually satisfying.
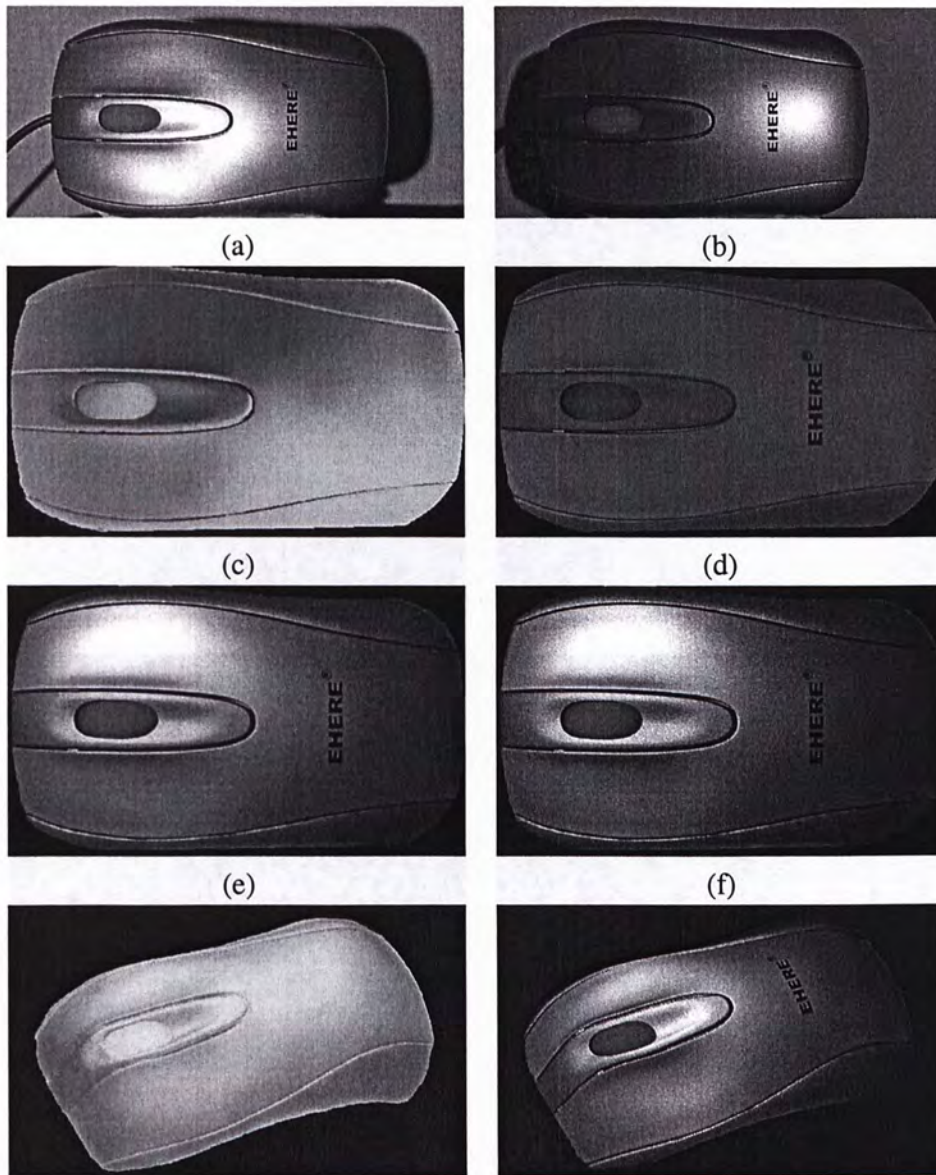
Figure 3.5: Mouse. (a) and (b) are two out of six input images contain large regions of highlight, (c) The recovered surface normal map encoded in RGB channels. (d) Surface textures. (e) Rendered mouse under a novel lighting direction. (f) The camera captured image using the same lighting direction, which is similar to our rendering result (e). (g) Reconstructed shape in a synthetic view. (h) Rendered image under novel lighting with the same view as (g).
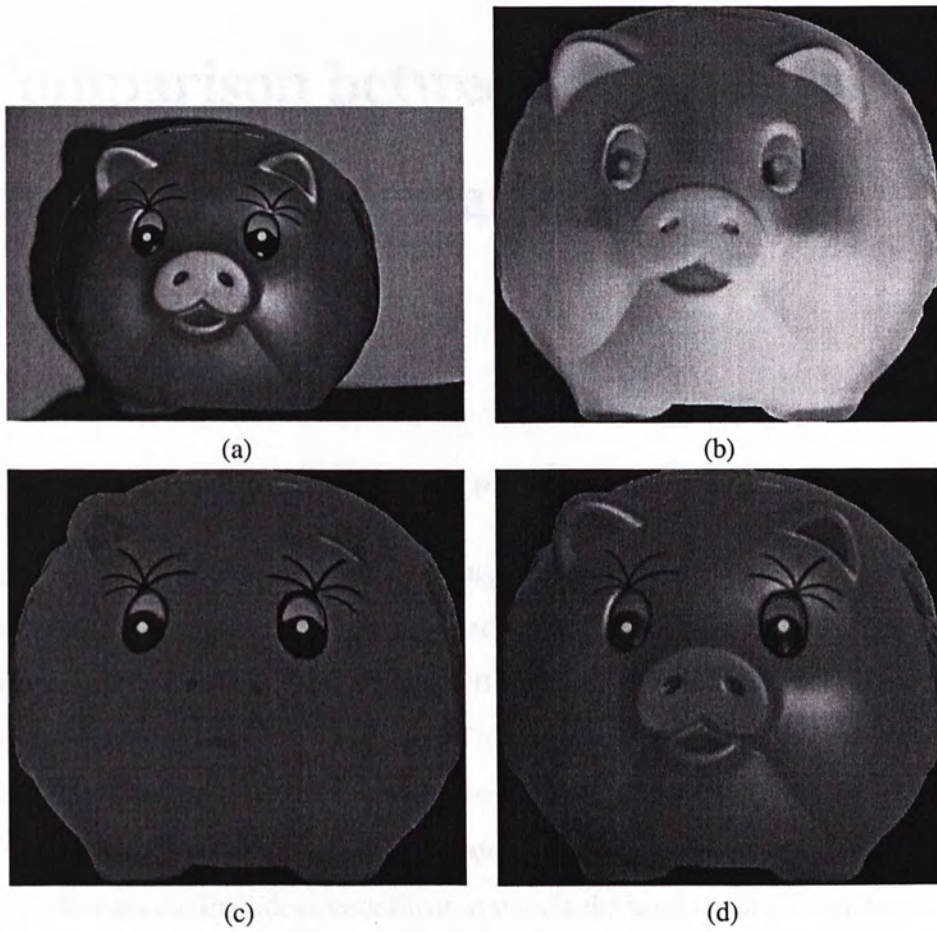
Figure 3.6: Saving box. (a) One of the input images, (b) and (c) Reconstructed normals and textures. (d) The rendered image under novel lighting.

Chapter 4

Comparison between Our Method and Direct Specularity Fitting Method

4.1 Summary of direct specularity fitting method [9]

[9] introduces the direct specularity fitting method to incorporate the Torrance and Sparrow reflectnace model in uncalibrated photometric stereo. It demonstrates that the specularity can help to resolve the GBR ambiguity in uncalibrated photometric stereo. The details in GBR ambiguity of uncalibrated case can be found in [5]. The following is the summary of this specularity fitting method:

Let the surface of the object be modeled by a height function $z(x, y)$, where (x, y) denotes the Euclidean coordinate system in the image plane. Then the surface normal is

$$n(x, y) = \frac{(z_x(x, y), z_y(x, y), -1)}{\sqrt{x_x^2(x, y) + x_y^2(x, y) + 1}} \qquad (4.1)$$

$$z(x, y; c(w)) = \sum c(w)\psi(x, y; w) \qquad (4.2)$$

To enforce continuity of the surface, the surface $z(x, y)$ can be expanded using a set

of discrete cosine basis functions $\psi(x, y; w)$ in 4.2. $c(w)$ is the discrete consine transorm coefficients of $z(x, y)$. The summation is performed over the two dimensional index w where $w = (y, v)$.

The partial derivatives of $z(x, y)$ can be expressed in this expansion form, so does the surface normal.

$$z_x(x, y; c(w)) = \sum c(w)\psi_x(x, y; w) \tag{4.3}$$

$$z_y(x, y; c(w)) = \sum c(w)\psi_y(x, y; w) \tag{4.4}$$

Using Torrance and Sparrow model, the intensity of a surface point at (x, y) lighted by light source l_i can be expressed as

$$I(n(x, y; c(w)), \rho_d(x, y), \rho_s, \omega, l_i) = \rho_d n^T l_i + \frac{\rho_s exp(-\omega^2[arccos^2(K)])}{n^T v} \tag{4.5}$$

where

$$K = \frac{n^T(\frac{l_i}{\|l_i\|} + v)}{\|(\frac{l_i}{\|l_i\|} + v)\|} \tag{4.6}$$

ω is the surface roughness and v is the viewing direction. $\rho_d(x, y), \rho_s$ are the diffuse and specular coefficients.

Direct fitting algorithm solve the following minimization problem using iterative scheme, the optimization steps are listed in Table 4.1:

$$min_{c(w), \rho_d(x,y), \rho_s, \omega, l_i} P(c(w), \rho_d(x, y), \rho_s, \omega, l_i) \tag{4.7}$$

$$= \sum_{ij} |M_{ij} - I(n(x, y; c(w)), \rho_d(x, y), \rho_s, \omega, l_i)|^2 \tag{4.8}$$

Direct specularity fitting method depends heavily on the result of the initialization step, due to the fact that the search space is inherently huge in this problem, unsatisfactory initialization makes the algorithm easily trap in local minimum. Moreover, the algorithm is very slow, more than 2000 iterations are needed before convergence for complicated examples.

1. Run the algorithm for a number of iterations using Lambertian model by setting ρ_s fixed to zero. This provides the algorithm a reasonable initialization. Initialize the DCT coefficients $c(w) = 0$, $\rho_s = 0$ and $\omega = 2$.

2. Perform SVD on matrix X to get the light light matrix L, each column of L is used as initialization of the light sources.

3. Update $c(w)$ using steepest descent by calculating $\frac{\partial P(c(w))}{\partial c(w)}$.

4. Update albedo $\rho_d(x, y)$ using linear least squares while keeping other optimization parameters fixed.

5. Update each of light source directions l_i using Newton's method. The partial derivative $\bigtriangledown P(l_i)$ and Hessian $\bigtriangledown^2 P(l_i)$ can be estimated.

6. Repeat steps 3-5 until $P(c(w), \rho_d(x, y), \rho_s, \omega, l_i)$ is smaller than a threshold.

7. Perform inverse DCT on the $c(w)$ to obtain the surface $z(x, y)$.

Table 4.1: Algorithm of Direct Specularity Fitting Method

4.2 Comparison results

In Figure 4.1, we compare our method with the one introduced in [9] using the "mouse" example. This direct specularity fitting process involves a large set of unknowns, the result is shown in 4.1(a). The large scale of artifact can be seen in the recovered surface, showing that the overlapping large highlight problem prohibit the algorithm to converge to a satisfactory result. The huge search space is also another significant problem, this method run more than 4.5 hours before convergence, which is very inefficient.

In our implementation, we assign known values to lighting variables. So their method is simplified by including less unknowns. The search space of the problem is reduced, and it work as a calibrated photometric stereo method just similar to our method. As shown in Figure 4.1(b), the reconstructed surface using this simplified version of method still contains problems . Most of the errors occur in pixels where strong highlight overlaps in the input images. It takes 2.6 hours to produce the
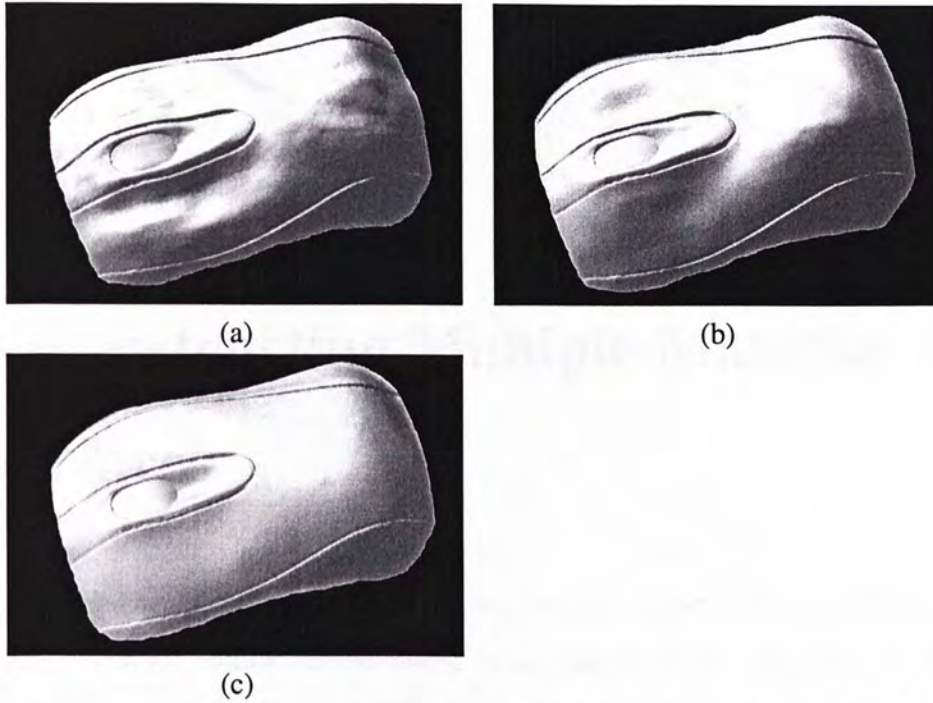
(a)　　　　　　　　　　　　　　　　(b)

(c)

Figure 4.1: (a) shows the result from [9], (b) is result by our implementation of modifying the method in [9] using known light directions and fixing the corresponding variables. So the optimization is simplified with less unknowns. (c) shows the result by our method for comparison.

surface result.

Figure 4.1(b) shows the result computed by our method. Since our method uses shadow visual clue to estimate the global parameters, the complexity of normal estimation is reduced. Moreover, our method runs very fast, using only 50 minutes in computation for this example.

Chapter 5

Reconstructing Multiple-Material Surfaces

Our method can also be naturally applied to surface reconstruction of multiple materials. With the similar configuration, we can obtain object silhouettes projected on the backboard. Similarly applying the method described in Section 3.1, a set of surface normals are computed. Then we use the method described in Section 3.1.4 to compute several optimal specular parameter sets from these normals. We place these values in the parameter space. Assuming there are M different materials with different specular properties where $M > 1$, rather than using mean shift clustering to find a unique parameter configuration, here we employ mean shift segmentation to initially form M partitions in the parameter space. Then the center of each partition has high density in which M different specular parameter sets can be robustly obtained.

After specular estimation, we use the parameters to recover the surface normal for each pixel. Using the optimization described in Section 3.2, we are able to compute a set of optimal $\mathbf{n}^*(j)$ and $\rho_d^*(j)$, where $0 \le j < M$, in terms of minimizing the residual error $\xi_{p,j}$ where

$$\xi_{p,j} = \sum_{0 \le i < n} (I_i(p) - L_o(p, \rho_s(j), \alpha(j), \mathbf{n}^*(j), \rho_d^*(j)))^2.$$

There should exists only one true surface normal for each pixel p, but simply comparing all $\xi_{p,j}$ for different j and selecting normal corresponding to the smallest $\xi_{p,j}$ is unreliable. Considering the material continuity property of the object surface, we treat the selection of optimal specular parameter set as a labeling problem in Markov Random Field where the parameter set of each pixel can be selected from M different assignments. Then we minimize the energy defined as

$$U(f) = U_{data}(f) + U_{smooth}(f) \qquad (5.1)$$

$$= \sum_{p \in \mathcal{P}} \xi_{p,f(p)} + \alpha \sum_{\{p,q\} \in \mathcal{N}} \mathcal{H}(s_{f(p)}, s_{f(q)}), \qquad (5.2)$$

where $f : \mathcal{P} \to \mathcal{S}$, and $\mathcal{P}$ and $\mathcal{S}$ denote the pixel set and label set respectively. So $f(p)$ returns the label at p. $s_j = [\rho_s(j), \alpha(j)]^T$, denoting the jth set of specular parameters, where $0 \leq j < M$. $\mathcal{N}$ is the set of neighborhood. α is a weight. $\sum \xi_{p,f(p)}$ is the data term. $\sum \mathcal{H}(a,b)$ is the smoothness term defined by $\mathcal{H}(a,b) = 1 - exp[-\frac{1}{2}(a-b)^T \Sigma^{-1}(a-b)]$, encoding the labeling smoothness between adjacent pixels. Σ denotes the covariance matrix between vector a and b. We solve (5.2) using graph cuts [3]. The whole surface is thereby segmented into partitions with computed optimal normal and albedo for each pixel.

Note that in order to faithfully reconstruct normals for object surface consisting of M different materials, M sets of global parameters must be recovered. Therefore, there should exist surface points from each material projecting to at least one of the cast shadow boundaries. This condition, in our experiments, can be generally satisfied when several input images under different lighting are used.

5.1 Multiple material results

Figure 5.1 shows our result on a synthetic kettle example. Eight input images are used. The kettle consists of two material, one is painted wood on the lid and the handle, the other is the metal body of the kettle. Our computed composition map in (b) correctly separate the two material components. The computed surface normal
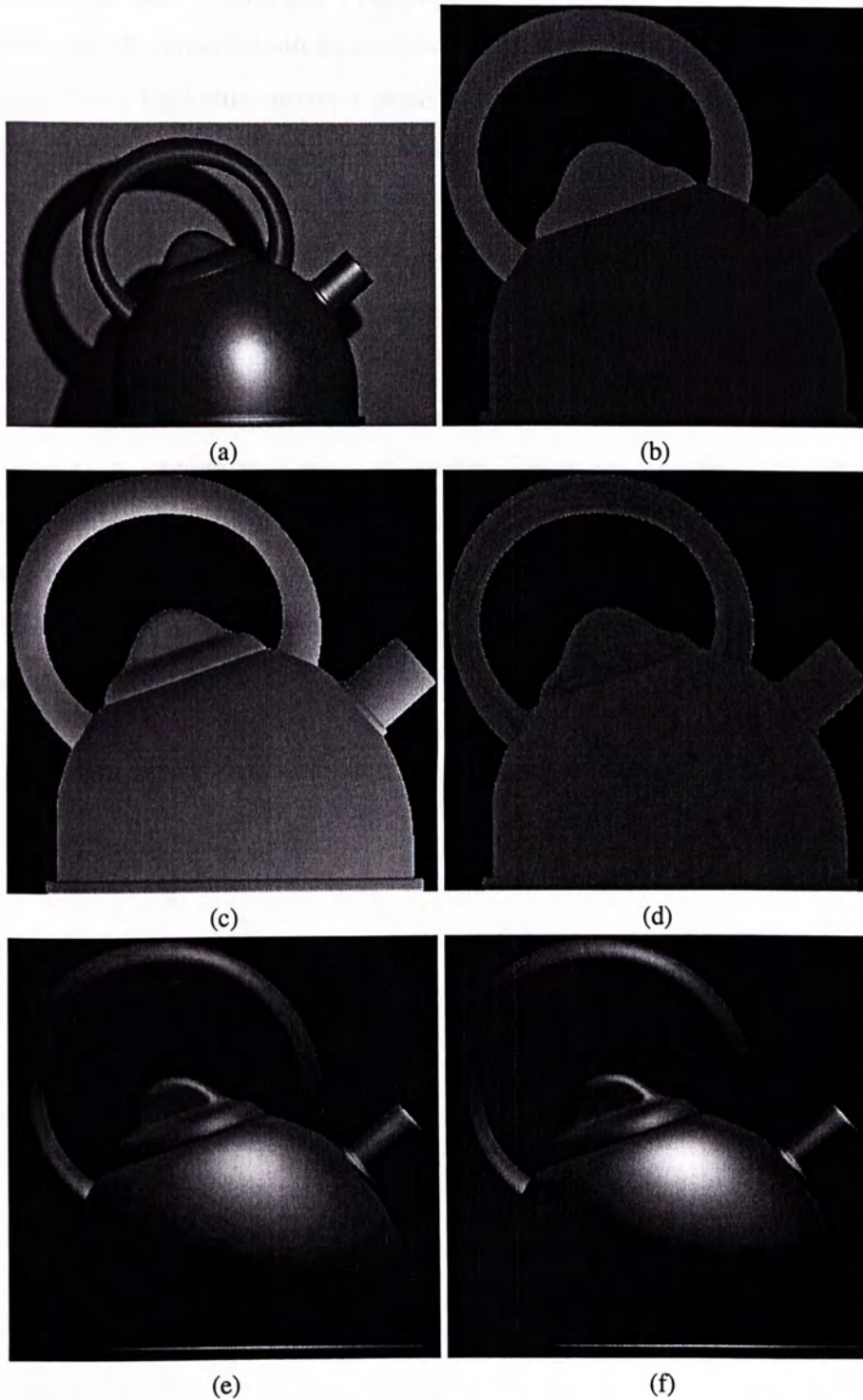
Figure 5.1: Kettle. (a) The kettle is made of two materials. (b) Recovered material composition map. (c) and (d) The computed surface normals and textures. (e) We re-render the kettle under a novel lighting condition. (f) The ground truth image for comparison.

and texture are shown in (c) and (d) respectively. We compare the recovered surface normal with the ground-truth surface model, the average difference is only 1.14 degree. Using the kettle surface normal and texture, we can relight the kettle in different lighting (e), which is very comparable to the the ground truth one in (f).

Figure 5.2 illustrates a very challenging daily-life example. It is a toy model with multiple materials on surface. Each material has its own roughness and specularity properties. We use 10 input images and the reflections in the input images are complex. Using our method, different materials can be successfully recovered as shown in (b). The material shown in green is a kind of hard plastic which has strong and overlapping highlight in all images. The material in blue shows smaller highlight whereas the red material is near Lambertian. The recovered normals in (c) preserve sufficient fine details and structures. Comparing our rendered image with the image taking by the camera under the same lighting condition, the RMS error is only 0.0423.

(a)

(b)

(c)

(d)

(e)

(f)

Figure 5.2: Toy. (a) The toy surface consists of multiple materials, and the reflection is very complex. (b) Recovered material composition map. (c) and (d) The computed surface normals and textures. (e) We re-render the toy under a novel lighting condition. (f) The ground truth image taken by the camera for comparison.

Chapter 6

Conclusion

To conclude, in this thesis, we have proposed an efficient photometric stereo method to estimate the BRDF parameters and to reconstruct surface normals. We focus on the examples that wide specular lobes can be seen in the input images. This kind of surfaces is problematic to all related work on non-Lambertian photometric stereo. Our method is the first work to use the cast shadow silhouette information to help estimate the BRDF parameters and surface normal. Our method does not require general specular component separation and shows better performance than direct specularity fitting methods, in terms of accuracy and convergence rate. Our method uses a new configuration to capture the cast shadow of the object on back planes. Using the small number of input images, the shadow is applied to estimating the normal samples and computing the global parameters. Once the global parameters are estimated, the local reflectance parameters and normals can be optimized separately on each pixel. This largely reduces the computation complexity and improves the quality of the computed object surface and texture map. All our experimental results shows that our algorithm is an efficient and accurate method to solve wide specular lobe problem, allowing the photometric stereo method applicable to more complicated surfaces.

Bibliography

[1] N. G. Alldrin and D. J. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance:a stratified photometric stereo approach. *ICCV*, 2007.

[2] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(10):1239–1252, 2003.

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.

[4] M. K. Chandraker, S. Agarwal, and D. J. Kriegman. Shadowcuts: Photometric stereo with shadows. *CVPR*, 2007.

[5] M. K. Chandraker, F. Kahl, and D. J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1*, pages 788–795, Washington, DC, USA, 2005. IEEE Computer Society.

[6] E. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Comp. Vision, Graphics and Image Processing*, 1982.

[7] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5):603–619, 2002.

[8] A. S. Georghiades. Incorporating the torrance and sparrow model of re-flectance in uncalibrated photometric stereo. *iccv*, 02:816, 2003.

[9] A. S. Georghiades. Recovering 3-d shape and reflectance from a small number of photographs. In *EGRW*, pages 230–240, 2003.

[10] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *ICCV*, pages 341–348, 2005.

[11] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey Vision Conf.*, pages 147–151, 1988.

[12] A. Hertzmann. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(8):1254–1264, 2005. Member-Steven M. Seitz.

[13] G. Klinker, S. Shafer, and T. Kanade. The measurement of highlights in color images. *International Journal of Computer Vision*, 2(1):7–32, 1988.

[14] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *CVPR*, pages 619–626, 2005.

[15] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of lambertian, specular, and hybrid surfaces using extended sources. In *Workshop on Industrial Applications of Mach. Intell. and Vision*, pages 169 – 175, 1989.

[16] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. on Robotics and Automation*, 6(4):418–431, 1990.

[17] S. K. Nayar, X.-S. Fang, and T. Boult. Separation of reflection components using color and polarization. *Int. J. Comput. Vision*, 21(3):163–186, 1997.

[18] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.

[19] Y. Sato and K. Ikeuchi. Temporal-color space analysis of reflection. Technical Report CMU-CS-92-207, 1992.

[20] Y. Sato, M. D. Wheeler, and K. Ikeuchi. Object shape and reflectance modeling from observation. In *SIGGRAPH*, pages 379–387, 1997.

[21] K. Schluns and O. Wittig. Photometric stereo for non-lambertian surfaces using color information. In *CAIP*, pages 444–451, 1993.

[22] S. A. Shafer. Using color to separate reflection components. *Color Research and Application*, 10(4):210–218, 1985.

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

[24] H. Tagare and R. deFigueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):133–152, 1991.

[25] H. D. Tagare and R. J. P. deFigueiredo. Simultaneous estimation of shape and reflectance map from photometric stereo. *CVGIP: Image Underst.*, 55(3):275–286, 1992.

[26] P. Tan, S. Lin, L. Quan, and H.-Y. Shum. Highlight removal by illumination-constrained inpainting. In *ICCV*, pages 164–169, 2003.

[27] P. Tan, L. Quan, and S. Lin. Separation of highlight reflections on textured surfaces. In *CVPR*, pages 1855–1860, 2006.

[28] R. T. Tan and K. Ikeuchi. Separating reflection components of textured surfaces using a single image. *IEEE Trans. Pattern Anal. Mach. Intell*, 27(2):178–193, 2005.

[29] G. J. Ward. Measuring and modeling anisotropic reflection. *SIGGRAPH Comput. Graph.*, 26(2):265–272, 1992.

[30] L. B. Wolff and T. E. Boult. Constraining object features using a polarization reflectance model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):635–657, 1991.

[31] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[32] T.-P. Wu and C.-K. Tang. Dense photometric stereo using a mirror sphere and graph cut. In *CVPR*, pages 140–147, 2005.

[33] W. Xu, M. Jenkin, and Y. Lesperance. A multi-channel algorithm for edge detection under varying lighting. In *CVPR*, pages 1885–1892, 2006.

[34] K. J. Yoon, Y. Choi, and I. S. Kweon. Fast separation of reflection components using a specularity-invariant image representation. In *International Conference on Image Processing*, pages 973–976, 2006.

Appendix A

Proof of Surface Normal Projecting to Gradient of Cast Shadow Boundary
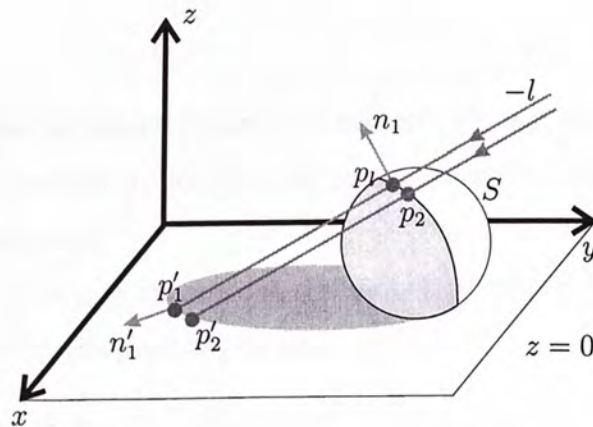

Figure A.1: Illustrative figure shows the configuration with surface S and a shadow plane.

As shown in Figure A.1, let $S(x, y, z) = 0$ denote a 3D surface and $S \in C^1$, i.e. S is differentiable with respect to x, y, z and its first order derivative is continuous. $z = 0$ is the shadow plane, $-l$ is the direction of the light rays, p_1 is the examined surface point on S, p'_1 is the point on the shadow boundary which is produced by

p_1 blocking the incident light, n_1 is the surface normal of S at point p_1. We define the 3D position of p_1 as (x_1, y_1, z_1). According to the basic theorem of differential geometry, we have

$$n_1 \;=\; (S_{x_1}, S_{y_1}, S_{z_1}),$$

where S_{x_1}, S_{y_1} and S_{z_1} are the first order derivatives of S with respect to x, y and z respectively at point p_1. Then the projection of n_1 onto the shadow plane can be written as

$$n^* \;=\; \frac{1}{\sqrt{S_{x_1}^2 + S_{y_1}^2}}(S_{x_1}, S_{y_1}, 0) \tag{A.1}$$

Obviously, p_1 and p_1' should be on the same light, thus

$$
\begin{aligned}
p_1' &= p_1 - \|p_1 - p_1'\|_2 \cdot l \\
p_1' &= p_1 + p_{1_z} q \cdot l, \\
q &= -\frac{1}{l_z}
\end{aligned}
\tag{A.2}
$$

We assume that there exists a point p_2 on the surface S such that $\|p_2 - p_1\|_2 \to 0$, and it produces a shadow point p_2' on the shadow boundary. This assumption is reasonable in most cases.

Let $p_2 = (x_1 + dx, y_1 + dy, z_1 + dz)$, obviously $\|(dx, dy, dz)\|_2 \to 0$. Therefore p_2 must be on the tangent plane of S at point p_1. So

$$(dx, dy, dz) \perp n_1 \Rightarrow S_{x_1} dx + S_{y_1} dy + S_{z_1} dz = 0$$

Similar to p_1,

$$p_2' \;=\; p_2 + p_{2_z} q \cdot l \tag{A.3}$$

Since p_1' lies on the shadow boundary, the light ray producing p_1' has only one intersection with the surface S. So

$$n_1 \perp l \;\Rightarrow\; S_{x_1} l_x + S_{y_1} l_y + S_{z_1} l_z = 0$$

45

Since $\|p_2 - p_1\|_2 \to 0$, combining Eqn. (A.2) and (A.3),

$$
\begin{aligned}
p_2' - p_1' &= p_2 - p_1 + dz \cdot q \cdot l \\
\|p_2' - p_1'\|_2 &\leq \|p_2 - p_1\|_2 + \|dz \cdot q \cdot l\|_2 \quad\quad (A.4) \\
\|p_2' - p_1'\|_2 &\to 0
\end{aligned}
$$

Let n_1' be the normal direction of the shadow boundary on point p_1', since $\|p_2' - p_1'\|_2 \to 0$, we have

$$
n_1' \perp (p_2' - p_1') \quad\quad (A.5)
$$

On the other hand, using Eqn. (A.1), we have

$$
\begin{aligned}
(n^*, (p_2' - p_1')) &= S_{x_1} \cdot (dx + dz \cdot q \cdot l_x) + S_{y_1} \cdot (dy + dz \cdot q \cdot l_y) \\
&= dx S_{x_1} + dy S_{y_1} - dz \frac{l_x}{l_z} S_{x_1} - dz \frac{l_y}{l_z} S_{y_1} \\
&= -dz S_{z_1} - dz \frac{l_x}{l_z} S_{x_1} - dz \frac{l_y}{l_z} S_{y_1} \\
&= -\frac{dz}{l_z} (l_z S_{z_1} + l_x S_{x_1} + l_y S_{y_1}) \\
&= 0
\end{aligned}
$$

So

$$
n^* \perp (p_2' - p_1') \quad\quad (A.6)
$$

Since n^*, n_1' and $(p_2' - p_1')$ are on the same shadow plane, combining Eqn. (A.5) and (A.6), we have

$$
n_1' = n^*.
$$

□

The above proof shows that the shadow boundary normal is the same as the projection of surface normal on the shadow plane, the only assumptions here are the surface should be C^1 and the neighboring pixels on the shadow boundary should be produced by neighboring points on the surface, which are reasonable for natural objects.

46