PHOTOMETRIC STEREO WITH APPLICATIONS IN MATERIAL CLASSIFICATION

RAKESH SHIRADKAR

(B.Tech (Electrical Engineering), IIT Roorkee)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

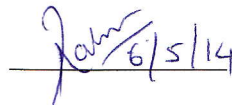DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2014

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Rakesh Shiradkar

May 6, 2014

i

Acknowledgements

There are several people without whose support this thesis would not have been possible.

Firstly, I would like to thank my advisor Dr. Ong Sim Heng for always being a helpful and supportive guide. He has given me the freedom to explore different ideas.

I am grateful to Dr. Tan Ping from whom I have learnt the basics of doing research, writing papers. He is very enthusiastic and passionate in pursuing research. Through my interactions with him, my appreciation and interest in computer vision and image processing has increased. I also express my thanks to my thesis committee members Dr. Yan Shuicheng and Dr. Cheong Loong Fah for their constructive comments.

I would like to thank my lab in-charge Mr. Francis Hoon Keng Chuan for being friendly and helpful in several practical aspects while conducting experiments. I am also grateful to Dr.Shen Li and Dr.George Landon for their helpful discussions. Thanks are also due to my labmates Zhou Zhenglong, Wu Zhe, Cui Zhaopong, Zuo Zhaoqi, Zhang Yinda, Tay Wei Liang Dr. Nianjuan, Dr. Loke and Dr. Csaba for their company at the lab and helpful discussions.

Most importantly, I would like to thank Dr. P V Krishnan, who has inspired me to pursue the direction of research. His personal example and precepts have inspired many people including myself. I am also grateful to Dr. Ankush Mittal, Dr. Sujoy Roy and Dr. Vipin Narang for their guidance and support. I am also

grateful to Dr. Karthik, Dr. Sivanand and Dr. Badarinath for being good friends.

Finally, I thank my parents and sister for their continuous trust and support without which I wouldn't have come this far.

Contents

Abstract

Recovering the shape and appearance of a scene are important problems in computer vision. Of all the methods developed towards solving these problems, photometric stereo is unique in terms of estimating the fine details in the geometry of the scene based on information from shading. In this thesis, different aspects of photometric stereo are explored and newer methods are presented to increase its scope.

Firstly, a method for resolving the ambiguities associated with the estimated shapes from uncalibrated Lambertian photometric stereo is presented. It is shown that a ring of light sources can explicitly determine the ambiguities and provide accurate shape estimates.

Next, an algorithm for estimating depth in the case of non-Lambertian surfaces is presented, building on previous methods which determine partial shape estimates using physical properties of the BRDFs.

Lastly, the aspects photometric stereo for reflectance based material classification are explored. It is observed that a slight modification in the camera and light configuration can significantly improve the performance of material classification. Additionally, a simple handheld device is presented which captures important discriminative information, although sampling a smaller BRDF space. This is applied to the problem of ink classification, an important area of forensics, introducing reflectance for the first time to the area.

x

List of Figures

xii

xiii

List of Publications

Journal articles

1. Rakesh Shiradkar, Ping Tan, Sim Heng Ong, "Auto-calibrating Photometric Stereo using Ring Light Constraints", *Machine Vision and Applications*, vol. 25, no. 3, pp. 801-809, 2014.

Conference proceedings

1. Rakesh Shiradkar, Sim Heng Ong, "Surface Reconstructions using Isocontours of Constant Depth and Gradient", in *Proc. of the IEEE International Conference on Image Processing (ICIP)*, pp. 360-363, 2013.

2. Rakesh Shiradkar, George Landon, Shen Li, Sim Heng Ong, Ping Tan, "A New Perspective on Material Classification and Ink Identification", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014. *(accepted)*

Chapter 1

Introduction

Computer vision concerns the extraction of information of a scene from its images. The information can be of various types such as recovering the shape of the objects and obtaining an estimate of useful features and this can be applied to desirable tasks such as computer assisted surgery, surveillance, preserving historical artefacts and developing navigational aids. With the advance of technology in recent times, we have high quality cameras widely available even to the common man. Therefore, computer vision systems can offer several services to professionals and the common man alike. Some of the commercially valuable applications of computer vision are in animation, computer assisted surgery where 3D reconstructions of objects and their renderings are especially important. For these applications, estimation of shape and reflectance are very important.

While several vision algorithms already exist, such as shape from stereo, structure from motion, shape from shading, shape from texture etc., the desirable goal is to develop robust, economically viable solutions for commercial applications. Among these algorithms, shading is one of the most powerful cues used in shape and appearance capture of real objects. A significant advantage of shading algorithms

over other vision algorithms is that shading can capture refined details on the surface which other algorithms such as stereo, SfM cannot capture.

Photometric stereo belongs to the class of vision algorithms which uses shading cues for shape and appearance capture. The present state of the art photometric stereo algorithms can recover very high quality 3D surfaces of objects [1, 2]. At the same time, photometric stereo requires a very simple arrangement for image capture: a camera and a set of light sources at different positions to illuminate the object. As the camera is fixed, the correspondence between the images is already established. As a consequence, with high quality shape estimates and low economic overhead, photometric stereo is a popular choice for shape estimation and appearance capture. The applications vary from 3D animation, gaming, cultural heritage preservation, computer assisted surgeries to material identification and classification.

1.1   Lambertian Photometric Stereo

Since photometric stereo mainly relies on shading cues, accurate estimates of intensities is very important. Therefore, one of the most important aspects of photometric stereo is calibrating the lighting directions and the intensity of the light sources. Calibration of light sources is usually done by calibration objects such as a mirror sphere. Alternatively, calibration can be done directly from pixel intensities and these algorithms are usually classified as uncalibrated or auto-calibrated photometric stereo algorithms. As the problem is highly constrained, these are typically designed for surfaces with *Lambertian* reflectance.

Lambertian surfaces are those which reflect light equally in all directions. Some examples of Lambertian surfaces are matte or diffuse objects such as chalk, terra
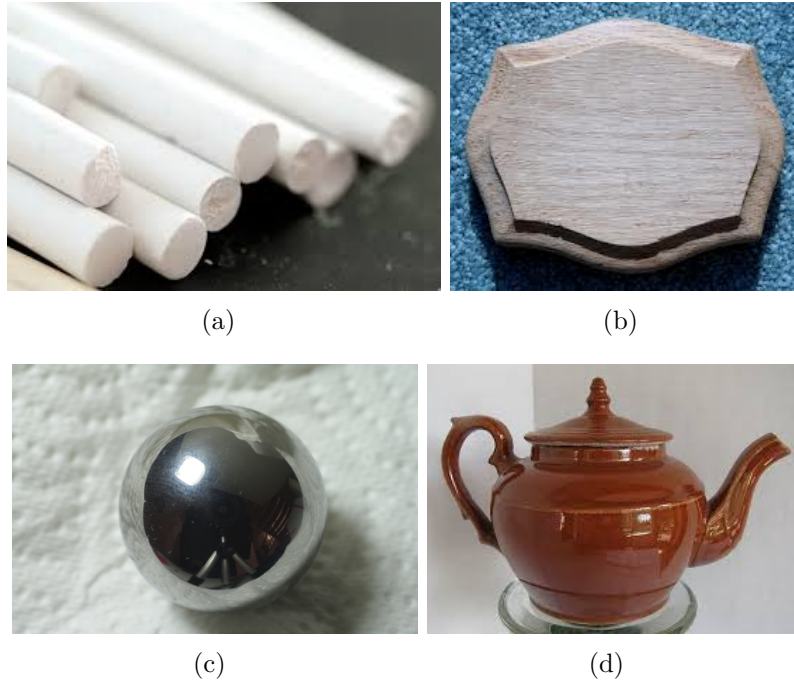
Figure 1.1: Examples of diffuse (top row) and specular (bottom row) surfaces. (a) Chalk [3], (b) unfinished wood [4], (c) steel ball [5], (d) glazed ceramic teapot [6]

cotta and unfinished wood as shown in Figure 1.1(a),(b). On the other hand, specular surfaces are those which reflect light more in a specific direction than any other direction. Examples of highly specular surfaces are stainless steel and glazed ceramic as shown in Figure 1.1(c),(d). There are no perfectly diffuse or perfectly specular materials in the real world; some materials exhibit a high degree of specularity while others may have a greater degree of diffuseness.

Shape estimation from photometric stereo with the Lambertian reflectance assumption is a well known problem in computer vision. It was introduced by Woodham [7] nearly three decades back. However, it is still a challenging problem, especially when the light directions from which the images are illuminated are unknown. This case is known as *uncalibrated* photometric stereo, as stated previously. This problem is interesting for the fact that shape estimates can be computed directly from the images. Unlike other vision methods such as stereo and structure-from-motion (SfM), we do not have the problem of finding the point

correspondences. At the same time, a per-pixel estimate of the shape can be recovered from photometric stereo.

We resume our discussion of auto-calibration of photometric stereo for Lambertian surfaces. Auto-calibration is an attractive solution as no extraneous calibration of light sources is necessary. However, the estimates of lighting direction and the shapes are not uniquely determined and additional user intervention is necessary. In Chapters 3 and 4, we present a method for the auto-calibration of photometric stereo using constraints on the light source positions.

1.2   Non-Lambertian Photometric Stereo

The assumption of Lambertian reflectance does not always hold for surfaces of real objects. In cases when the surface is relatively less diffuse or less shiny, the specular or shiny portions in the images are generally removed as outliers and Lambertian photometric stereo algorithms are applied. However, when the surface material exhibits a greater degree of specularity, different approaches are adopted to estimate the shape from images using photometric stereo. In this context, certain properties of the surface reflectance are exploited for shape estimation, for example, the property of *isotropy*, which is exhibited by most materials. *Isotropy* implies that the appearance of the surface remains unchanged when the object is rotated about the surface normal. In practise, a large number of materials (such as plastics, rubbers, most fabrics) exhibit isotropy. There also exist a few *anisotropic* materials whose appearance changes under rotation about the surface normal, such as brushed metal and velvet. The difference is illustrated in Figure 1.2

With the assumption that the surface reflectance follows the property of isotropy, non-Lambertian surfaces can be handled efficiently [11, 12, 13, 14]. Important

4
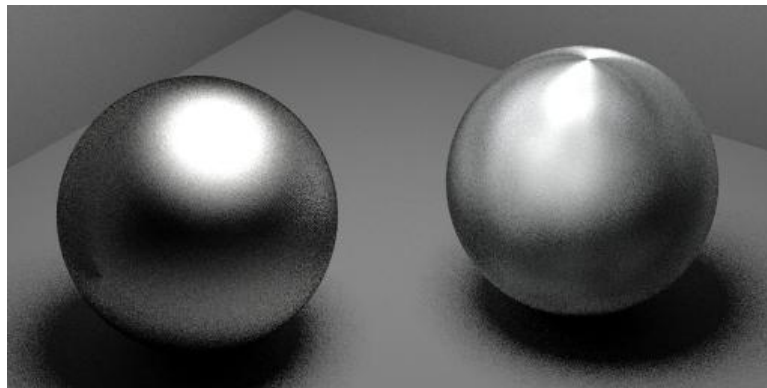
(a)

(b)

(c)

Figure 1.2: (a) Comparison of isotropic (left) and anisotropic BRDF (right)[8], (b) Isotropic material (rubber) [9], (c) Anisotropic material (brushed steel)[10]

information related to the object shape can be recovered by photometric stereo under this assumption. In Chapter 5, we propose an algorithm for surface reconstruction of non-Lambertian objects using the results of the previous algorithms.

1.3   Reflectance based Material Classification

While recovering the 3D shape of the object is one of the important problems of computer vision, it is also important to recover the appearance such as colour, texture, reflectance of the surface to produce accurate renderings of the objects. Typically, the reflectance at a point on the surface of an object is described by its *bi-directional reflectance distribution function* (BRDF). BRDF specifies behaviour of the incident light when it interacts with a surface.

The conventional capture of the BRDF is done using specialised devices called *gonioreflectometer*. However, they require significant effort to sample the BRDF; for a given set of incident and viewing directions, a single measurement of BRDF is made. Therefore, measurement of the BRDFs from images, such as Matusik et al. [15], is being commonly used in recent times.

Besides, shape estimation, photometric stereo is also used for recovering the surface reflectance. Images captured under varying illumination at a constant view point reveal information regarding the reflectance of the material on the object's surface. The knowledge of the reflectance of the object's surface is especially used in generating the renderings of 3D objects. As discussed previously, assuming certain characteristics of the BRDF (such as isotropy) helps in shape estimation using photometric stereo. Besides, such an assumption also helps in simplifying the BRDF representation. In Chapter 6, we give further insights into the problem of BRDF capture and apply it to the problem of ink classification, which is an

important field of study in forensics.

1.4 Contributions

In this thesis, we explore the various aspects of photometric stereo: (1) shape estimation from Lambertian photometric stereo, (2) shape estimation from non-Lambertian photometric stereo, (3) reflectance capture from photometric stereo, and (4) application of photometric stereo to ink classification.

The uncalibrated photometric stereo problem gives an ambiguous solution when only the pixel intensities from images are used. While assuming the surface to be continuous and integrable, the estimated surface shape is up to a so called *generalized bas-relief* (GBR) ambiguity. There are many methods to disambiguate the estimated shape completely [16, 17, 18, 19]. However, none of these methods explore the possibility of imposing constraints on the light sources for resolving the ambiguity. Recent works [12] have shown that parametrically constrained light sources can offer significant advantages. We observe that constraining the light sources to a ring can offer sufficient information so that the ambiguity parameters can be estimated with a closed form solution.

While photometric stereo produces high quality normal maps, the depth estimates from photometric stereo suffer from low frequency noise. The depth maps from depth sensors suffer from high frequency noise. There is a body of works [20, 21] which combine the normal maps from photometric stereo and the depth maps from depth sensors to create high quality surfaces. We propose a method to automatically calibrate photometric stereo towards normal and depth fusion techniques using rectangularly placed light sources. Such a configuration is easy to implement since conventional monitor screens inherently have this geometry.

Using characteristic properties of the BRDFs, non-Lambertian surfaces can be handled by photometric stereo algorithms without explicitly recovering their BRDFs. Alldrin et al. [12] recover the shape up to a set of contours of constant depth assuming that the surface is isotropic. Chandrakar et al. [13] estimate the surface shape upto a set of contours of constant depth gradient with the assumption of isotropic BRDFs. These methods only give a partial result which can be fully estimated with additional user information.We show that these contours together can give us sufficient information such that per pixel estimates of depth can be estimated.

Image based methods for BRDF acquisition often capture a subset of the BRDF (2D BRDF slice) [15, 22] by making assumptions on the nature of the surface (for example, isotropy). We observe that, most of the capture systems [23] used for the BRDFs generally capture an even smaller subset of the BRDF than what is proposed. We show that a small modification in the capture setup can significantly improve the performance on material classification. Next, we intend to develop a convenient and simple arrangement, a handheld flashlight-camera device, for capturing images for the purpose of material classification. This arrangement captures only a small portion, a 1D BRDF slice, of the complete BRDF. However, this small portion contains sufficient discriminative information to differentiate a class of materials. We test this arrangement on a collection of inks and obtain promising results. It is the first attempt at using reflectance for use in ink classification, one of the important areas of forensics.

1.5 Organization of the thesis

This thesis is organized as follows. The following Chapter 2 introduces the principles of photometric stereo and its various aspects. This chapter also reviews

significant works in this area to provide the reader with sufficient background to appreciate the contribution of the thesis. In Chapter 3, an auto-calibration solution to Lambertian photometric stereo is presented. It is shown that lights constrained on a ring can fully calibrate the photometric stereo. In Chapter 4, rectangularly placed light sources are used for auto-calibration of photometric stereo is presented. This forms a preliminary portion of the method aimed at generating high quality depth maps by integrating the depth estimates from depth sensors and normal map from photometric stereo. In Chapter 5, an algorithm is presented which builds on the previous works using parametric properties of the BRDFs for non-Lambertian Photometric stereo. An iterative algorithm is presented which can offset the use of additional constraints used in the previous works for shape estimation. In Chapter 6, a subset of the full BRDF is captured for material classification. While suggesting configurations for improved performance of material classification, the method is applied to the problem of ink classification. In Chapter 7, conclusions and suggestions for future work are presented.

Chapter 2

Photometric Stereo

Photometric stereo belongs to the class of methods which recover the geometry of a scene based on variations in image intensities observed due to changing illumination. The result usually is a normal map of a scene which is recovered from images acquired under a fixed view point and varying illumination. Besides shape estimation, reflectance and illumination are also recovered in many cases. Photometric stereo is closely related to the shape from shading (SfS) problem. Although both methods use shading information for shape estimation, shape from shading uses only a single image while photometric stereo makes use of multiple images. The idea was developed based on the assumption of a Lambertian surface. Although most real objects deviate from this assumption to various degrees, much of the work in photometric stereo has been developed on the Lambertian assumption owing to its simple definition and linearity in mathematical modelling. In this chapter, we review various aspects of photometric stereo and its course of development over the years. This chapter is organised as follows. In Section 2.1, we provide a brief background of the basic concepts upon which photometric stereo is built. Section 2.2 introduces the classic photometric stereo problem followed by the uncalibrated photometric stereo problem in Section 2.3. Section 2.4 deals

11

with photometric stereo for non-Lambertian surfaces. We discuss the recovery of surface from the normal map produced by the photometric stereo in Section 2.5 and the recovery of reflectance in Section 2.6.

2.1 Background

Photometric stereo is built upon the intensity estimates obtained from images. Therefore, an understanding of the physical process behind image formation and the mathematical models is important. In this section, we provide relevant background on the understanding of radiometry, reflectance and models for representing surfaces.

2.1.1 Radiometry

Radiometry is a field which concerns measurement of electromagnetic radiation (including visible light) and its transport within matter. Since image formation is based on the interaction of light with surfaces, the concepts of radiometry are very often used in computer vision.

The fundamental quantity in radiometry is *radiant energy Q*, measured in *Joules*. An associated quantity is the *radiant power or the radiant flux* Φ which is defined as the radiant energy flowing through a surface in unit time.

$$\Phi = dQ/dt \tag{2.1}$$

The radiant flux passing through a unit surface area is called the *radiant flux*
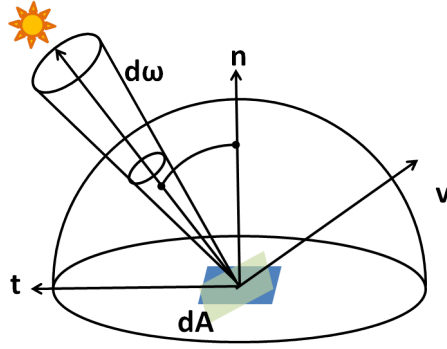
Figure 2.1: Diagram showing the solid angle and the projected area. The projected area (light green) actually receiving the light is smaller than the actual area (blue)

density u, measured in $Watts/m^2$.

$$u = d\Phi/dA \qquad (2.2)$$

where A represents the area of the surface patch through which the flux is measured.

If the radiant energy is arriving at the suface, the radiant flux density is termed *irradiance E* and if it is exiting from the surface, it is termed *radiant exitance* or *radiosity M*.

One of the most important radiometric quantity often used in computer vision is the *radiance*. Radiance L is defined as the flux arriving or leaving from a surface per unit solid angle per unit projected area.

$$L = \frac{d^2\Phi}{dAd\omega} \qquad (2.3)$$

where dA is the projected area and $d\omega$ is the solid angle. We know that an angle is a measure of the space spanned between two line segments. Similarly, a solid angle (measured in *steradians sr.*) is defined by the the space spanned by a surface patch on a unit sphere with respect to its centre. Projected area is that portion of the area perpendicular to the light direction as shown in Figure 2.1.

Table 2.1: Radiometric quantities

Quantity	Symbol	Units
Radiant energy	Q	J(Joule)
Radiant flux	Φ	$W \, or \, J/s$
Radiant flux density	u	J/m^3
Irradiance	E	W/m^2
Radiant exitance	M	W/m^2
Radiance	L	$W/(sr \cdot m^2)$

The radiometric quantities discussed are summarized in the Table 2.1

2.1.2 Reflectance

In the previous section, we have defined radiometric quantities which describe the transport of light in a medium. The light now interacts with the surface material of the object. When light interacts with matter, it undergoes different phenomenon - a portion of light gets reflected, another portion may get transmitted, another portion may get absorbed or scattered or re-emitted in a different direction. Only the portion of light that is reflected reaches the camera and contributes towards the formation of image. Therefore, we will discuss the concept of reflectance in this section.

Reflectance is the phenomenon when light arrives and exits on the same side at the boundary of the light and surface material. Reflectance properties of a surface are studied by measuring the *Bidirectional Reflectance Distribution Function* BRDF of a surface. The term was introduced by Nicodemus [24]. It is defined as the ratio of the reflected radiance dL in the reflected direction to the incident irradiance dE. Mathematically,

$$f(\mathbf{l}, \mathbf{v}) = f(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{dL(\theta_o, \phi_o)}{dE(\theta_i, \phi_i)} \tag{2.4}$$

Here, $\mathbf{l}, \mathbf{v}$ represent the incident and reflected directions; (θ_o, ϕ_o) describes the
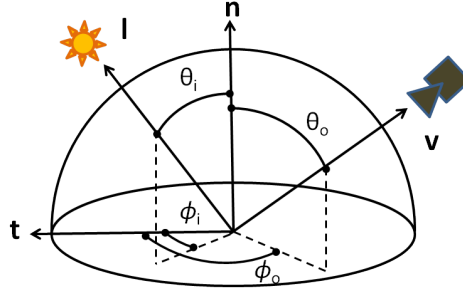
Figure 2.2: The angles and directions for describing the BRDF

reflected direction $\mathbf{v}$ and (θ_i, ϕ_i) describes the incident direction $\mathbf{l}$. θ and ϕ indicate the azimuth and the elevation angle. The directions and the angles are illustrated in the Figure 2.2.

The BRDF is usually measured using a device called *gonioreflectometer*. BRDFs are known to exhibit certain properties which are briefly summarized below.

The units of BRDF is inverse steradians sr^{-1}.

Physically based BRDFs are always positive.

$$f(\theta_i, \phi_i, \theta_o, \phi_o) \geq 0$$

BRDFs obey the energy conservation principle.

$$\int_\Omega f \cos \theta_o d\omega_o \leq 1$$

Certain BRDFs are invariant under rotation about the surface normal $\mathbf{n}$. Such BRDFs are called isotropic BRDFs. Mathematically, it means

$$f(\theta_i, \phi_i, \theta_o, \phi_o) = f(\theta_i, \theta_o, |\theta_i - \theta_o|)$$

BRDFs exhibit reciprocity. If the incident and reflected directions are

15

swapped, the value of the BRDF remains unchanged.

$$f(\theta_i, \phi_i, \theta_o, \phi_o) = f(\theta_o, \phi_o, \theta_i, \phi_i)$$

2.1.3 Reflectance models

In this section, we review some of the commonly used reflectance models.

Lambert's model This is the most fundamental and simplest reflectance model which describes a perfectly diffuse surface. It was proposed by Lambert [25] . According to this model, light is reflected uniformly in all directions for any given incident lighting direction. This implies that the appearance of the object remains the same for any viewing direction. The value of this BRDF is a constant. It is written as,

$$f(\mathbf{l}, \mathbf{v}) = \rho \tag{2.5}$$

where, ρ is the diffuse albedo in the range $[0, 1]$.

Because of its simplicity, Lambert's reflectance is the most widely used model in vision algorithms such as stereo, segmentation, shape from shading.

Phong model Phong [26] proposed a model to deal with non-Lambertian surfaces. Lambert's model assumes that the surface is perfectly diffuse and cannot model phenomenon such as specular highlights. The Phong model is a linear combination of a diffuse and a specular component,

$$f(\mathbf{l}, \mathbf{v}) = d\rho_d + s\rho_s \frac{(\mathbf{r} \cdot \mathbf{v})^n}{(\mathbf{n} \cdot \mathbf{l})} \tag{2.6}$$

16

where, d, s are the weights describing the surface and specular component such that $d + s = 1$; $\mathbf{r}$ is the mirror direction of $\mathbf{l}$, $n \in [1, \infty)$ represents the sharpness of the specular highlight. ρ_s, ρ_d are the specular and diffuse albedo.

Cook-Torrance model The Cook-Torrance BRDF model also gives linear combination of a diffuse part and specular part. Mathematically, it is given by

$$f(\mathbf{l}, \mathbf{v}) = d\frac{\rho_d}{\pi} + sF(\mathbf{l}, \mathbf{v})\frac{D(\mathbf{h})G(\mathbf{l}, \mathbf{v})}{4\pi(\mathbf{n} \cdot \mathbf{l})(\mathbf{n} \cdot \mathbf{v})} \tag{2.7}$$

where,$D(\cdot)$ is the distribution function of microfacet normals, $G(\cdot)$ is geometrical attenuation factor which represents the masking and shadowing effects of between facets, $F(\cdot)$ is the Fresnel term computed by Fresnel's equation, $d, s, \rho_d, \mathbf{h}$ carry the same meaning as in the Phong model, and $d + s = 1$. For more detailed explanation, please refer to the original work [27].

Several other improved reflectance models are proposed and the reader is advised to refer to existing literature for a more comprehensive view of radiometry.

2.2   Classical Photometric Stereo

Photometric stereo, as first introduced by Woodham [7], is a method to determine surface orientation from images captured under varying incident illumination at a fixed view point. As the camera position is fixed, the correspondence between points is already established. The early works in photometric stereo directly used the idea of reflectance maps. The reflectance map determines the reflectance as a function of the surface gradients p, q. Different lighting directions require appropriate reflectance maps showing isocontours of constant reflectances. When the object is imaged from two lighting directions, we obtain two reflectance maps
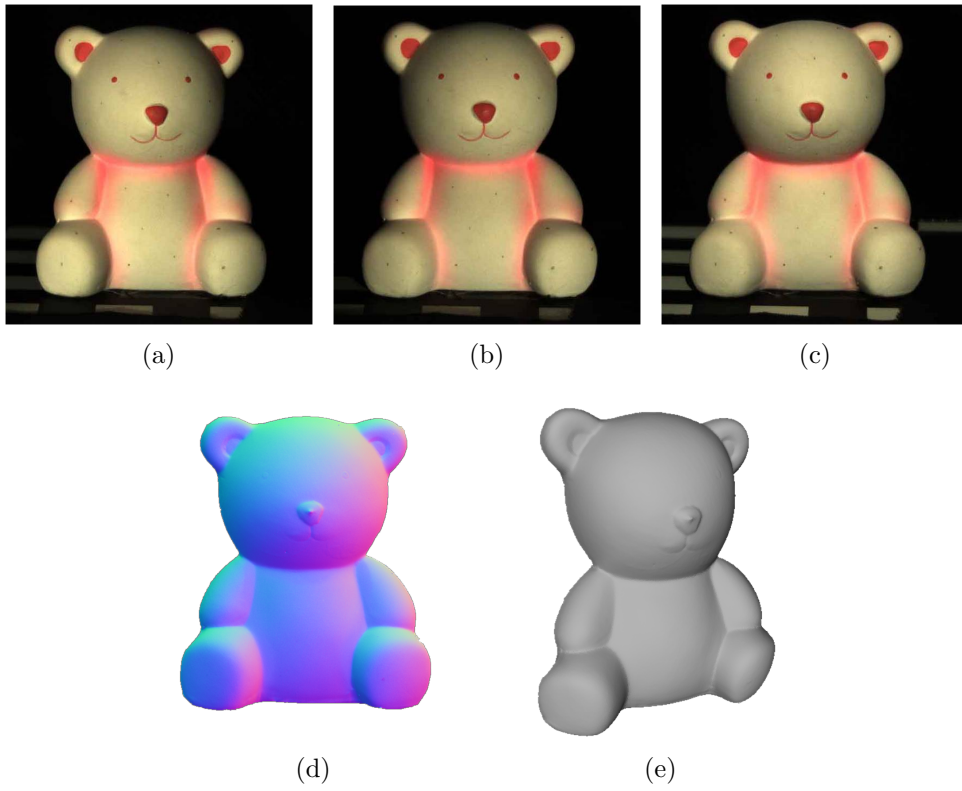
Figure 2.3: Photometric stereo.(a) - (c) Images captured under varying illumina-tion, (d) normal map, (e) depth map
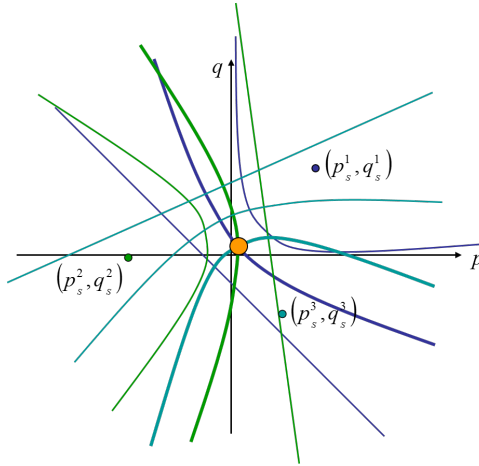
Figure 2.4: Reflectance map of a Lambertian object[28]

for each lighting direction. The surface orientation at a point can be estimated uniquely by the intersection of three isocontours, each determined from respective reflectance maps. With three measurements, the surface orientation can be uniquely determined as shown in Figure 2.4.

The classical photometric stereo algorithm assumes that the surface is Lambertian, the illumination direction from point light sources is known, albedo is uniform and the object is viewed from an orthographic camera. According to the Lambertian reflectance model, the observed intensity is given by,

$$I = \rho \mathbf{n} \cdot \mathbf{l}$$

Here, $\mathbf{n}$ is the surface normal and $\mathbf{l}$ is the illumination direction. To uniquely determine the normal direction, a minimum of three different lighting directions is required. Using more than three light sources (images) gives the additional advantage of noise reduction in the estimates. However, there are algorithms which use less than three images. For example, Onn and Bruckstein [29] where only two images are used to determine the normals under known lighting directions. They denote the normal in terms of the depth gradients p, q which will be defined later in Section 2.5. They assume surface integrability to uniquely determine the normals.

19
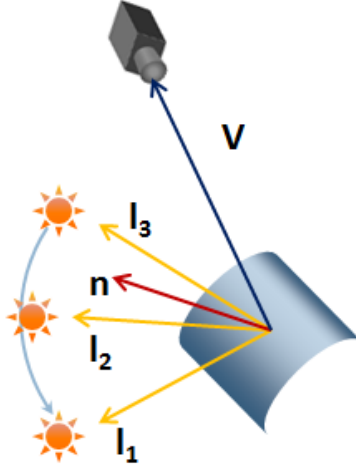
Figure 2.5: Photometric Stereo acquisition setup

Here, we individually determine the 3 components of the normal vector. Let us consider an object illuminated by m light sources, capturing an image of n pixels with each light turned on, one at a time. The image intensity at the ith pixel in an image of n pixels when illuminated by the jth light source is given by

$$I_{ij} = \rho_i \mathbf{n}_i{}^T \mathbf{l}_j \tag{2.8}$$

where, ρ_i and $\mathbf{n}_i$ are the albedo and surface normal at the ith pixel and $\mathbf{l}_j$ is the illumination direction from the jth light source.

By stacking the pixels from all the m images, (2.8) can be expressed in matrix form as

$$\mathbf{I} = \mathbf{BL} \tag{2.9}$$

where, $\mathbf{I} = \{I_{ij}\}_{n \times m}$ is the intensity matrix, $\mathbf{B}$ is an $n \times 3$ matrix representing the surface normals compounded with albedo and $\mathbf{L}$ is a $3 \times m$ matrix representing the lighting directions respectively.

As previously stated, the classical photometric stereo problem assumes that the light directions are known beforehand. They are usually determined by calibration

by placing a shiny mirror sphere in the scene. So, we have an overdetermined system in (2.9). Since $\mathbf{L}$ and $\mathbf{I}$ are already known, the matrix $\mathbf{B}$ can be uniquely determined by a least squares solution as,

$$\mathbf{B} = \mathbf{L}^{\dagger}\mathbf{I} \tag{2.10}$$

where $\mathbf{L}^{\dagger} = \left(\mathbf{L}^T\mathbf{L}\right)^{-1}\mathbf{L}^T$ is the pseudo inverse of $\mathbf{L}$. The normals and albedo can be determined from $\mathbf{B}$ as,

$$\mathbf{n}_i = \frac{\mathbf{b}_i}{\|\mathbf{b}_i\|} \rho_i = \|\mathbf{b}_i\| \tag{2.11}$$

where, $\mathbf{b}_i$ refers to the ith row of $\mathbf{B}$

We observe that the solution is very straightforward. But this requires that we capture the light directions using a mirror sphere and the surface is Lambertian without any outliers such as specular highlights, inter reflections, cast and attached shadows. The capture of light directions using a mirror sphere may introduce undesirable effects such as inter reflections. To overcome the disadvantage of having the light directions calibrated, auto-calibrated or uncalibrated photometric stereo methods can be used.

2.3 Uncalibrated Photometric Stereo

Uncalibrated photometric stereo was pioneered by Hayakawa [30] in 1994. For a Lambertian surface, the image intensities matrix $\mathbf{I}$ (as described in Section 2.2) ideally is a rank 3 matrix. When this matrix is factorized using a singular value decomposition (SVD), we can directly estimate the normal and lighting directions.

This result is up to a 3×3 unknown invertible linear transformation.

$$
\begin{aligned}
\mathbf{I} &= \mathbf{UDV}^T \\
&= \left(\mathbf{UD}^{\frac{1}{2}}\right)\left(\mathbf{D}^{\frac{1}{2}}\mathbf{V}^T\right) \\
&= \tilde{\mathbf{B}}\tilde{\mathbf{L}}
\end{aligned}
\tag{2.12}
$$

where, $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{L}}$ could differ from the true estimates $\mathbf{B}$, $\mathbf{L}$ by the 3×3 linear transformation A such that $\tilde{\mathbf{B}} = \mathbf{BA}^{-1}$ and $\tilde{\mathbf{L}} = \mathbf{AL}$.

Following this attractive solution, there are many works which try to resolve this linear ambiguity $\mathbf{A}$ and estimate the true normal directions. Hayakawa [30] proved that, if at least six light sources of equal intensity are known or atleast six normal directions of a curved surface with uniform albedo are known, this linear ambiguity can be reduced to a rotational transformation. On the other hand, enforcing the integrability constraint can also improve the solution by reducing the ambiguity to a GBR ambiguity [31, 16]. The GBR transformation consists of 3×3 matrices of the form

$$
\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}
\tag{2.13}
$$

where, $\mu, \nu, \lambda \in \mathbb{R}$ are the GBR parameters.

Under GBR transformation, the albedo and surface normal estimates are distorted as

$$
\hat{\rho} = \rho\|\mathbf{n}^T\mathbf{G}^{-1}\| \qquad \hat{\mathbf{n}}^T = \frac{\mathbf{n}^T\mathbf{G}^{-1}}{\|\mathbf{n}^T\mathbf{G}^{-1}\|}
$$

Belhumeur et al.[16] showed that the GBR ambiguity is intrinsic to Lambertian surface. As observed in Figure 2.6, the object's visible surface cannot be distinguished under a GBR transformation. This means that for each image captured under a particular lighting direction, there will be an identical image of the trans-
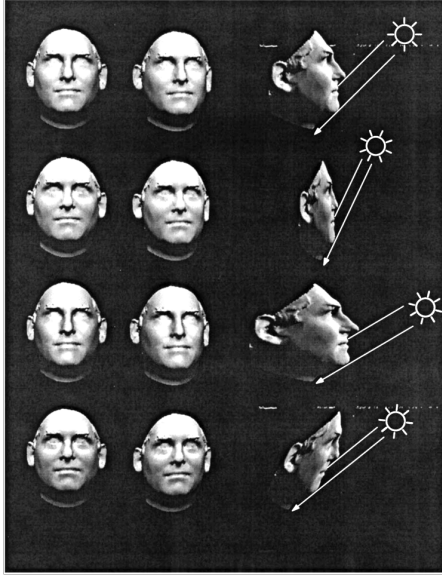
Figure 2.6: The bas relief ambiguity [16]

formed object by similarly transformed light sources. The integrability and equal intensity constraints together reduce the ambiguity to a binary convex/concave ambiguity . In terms of the GBR ambiguity, (2.12) can be written as,

$$\mathbf{I} = \tilde{\mathbf{B}} \mathbf{G}^{-1} \mathbf{G} \tilde{\mathbf{L}} \tag{2.14}$$

The GBR ambiguity can be resolved in several ways by making certain additional assumptions on either or more of geometry, albedo, light configuration. In [18], the GBR ambiguity was resolved by minimizing the entropy of recovered albedos. From (2.10), we know that albedo can be measured as the magnitude of the recovered normal. Under GBR transformation, this albedo is distorted and the entropy of the albedo distribution increases. By relating the entropy of the albedo distribution to the GBR transformation, the GBR parameters are found by minimizing the albedo distribution. This can resolve the GBR ambiguity, but the normals and albedos are assumed to obey a certain distribution which means not all surfaces can be solved by this method.

23

Using color and intensity profiles from images acquired for photometric stereo, the GBR ambiguity can be resolved [32]. By grouping normals with constant albedo, the GBR parameters and the albedo are recovered. From (2.12), we know that the albedo estimates after SVD decomposition contain the GBR ambiguity matrix. Therefore, if we know the normals with constant albedo, the GBR parameters can be estimated using a least squares solution.

Specular highlights can also be used to resolve the GBR ambiguity [19]. If at least two images contain sharp specular highlights, the normal and light direction estimated at the specular pixels and the GBR parameters form a linear system of equations. Solving these system of equations, the ambiguity parameters can be estimated. This method still falls under Lambertian reflectance as the specular highlights are only used to resolve the ambiguity. These highlights are later removed as outliers, and normals are estimated under the Lambertian reflectance assumption.

Other works use inter reflections [33], special lighting configurations [34], complex reflectance models [35], isotropic BRDF assumption [11, 36, 37], smoothly varying surface albedos [17] to resolve the GBR parameters. In Chapter 3, we propose a method to automatically calibrate photometric stereo using a ring light configuration.

Until now, we have reviewed methods related to Lambertian photometric stereo where the surface is assumed to be perfectly diffuse. To handle real world scenes, we need algorithms that can deal with non-Lambertian surfaces which is the subject of the following section.

2.4 Non-Lambertian photometric stereo

Under Lambertian reflectance, the incoming light is equally reflected in all directions so that the object appears the same for the viewer under any direction. This simple law is used very commonly in formulating algorithms for photometric stereo because of its simple, linear model. For non-Lambertian surfaces, which closely model the real scenes, photometric stereo algorithms become more challenging.

Broadly, non-Lambertian surfaces are handled either by (1) removing the non-Lambertian component as outliers, (2) using more detailed reflectance models, or (3) exploiting the reflectance properties. These approaches are briefly reviewed below.

2.4.1 Removing the non-Lambertian components

A majority of non-Lambertian surfaces exhibit near-Lambertian behavior outside the specular regions. It appears an attractive option to apply the linear algorithm developed for Lambertian surfaces to apply for non-Lambertian reflectance by treating highlights as deviations from the Lambertian law.

One of the earliest ideas in this direction was by Coleman and Jain [38]. Three light sources are needed to uniquely determine a normal for a Lambertian surface. They use 4 light sources to compute a set of 4 normal vectors for each point - one normal from the permutation of three intensity values. Using a thresholding procedure, surface normals are estimated from intensity values which do not contain a specular component. This is one of the most common and simple procedures to detect and remove non-Lambertian components. This method can also be used to eliminate attached and cast shadows while estimating the surface normals.

Barsky and Petrou [39] improvise the above 4 source photometric stereo method

using color images. The color information serves as an additional cue in the detection of specularities. Using pixel-wise estimates of color, they detect highlights at each pixel. They also detect the shadows as deviations from the linearity on the input quadruples of pixels. Several other works [40, 41, 42]have been proposed based on a similar idea of removing the non-Lambertian components as outliers for both, calibrated and uncalibrated photometric stereo.

2.4.2 Using sophisticated reflectance models

We have reviewed some of the reflectance models commonly used to represent BRDFs of non-Lambertian earlier in the chapter. This class of methods try to recover the parameters of the reflectance models for shape estimation.

Some of the earlier works assume that the BRDF of a surface is a linear combination of specular and Lambertian component [43]. They term this model as a hybrid model and solve for the reflectance and shape at each surface point using the set of intensity values. The dichromatic reflectance model is a widely used model for modeling non-Lambertian components. Mallick [44] try to recover the surface without explicit separation of diffuse and specular components. A transformation of color space is employed which directly gives diffuse image freed from specular effects. The transformation preserves shading information and thus the obtained images are used for photometric reconstruction.

Lensch et al. [45] assume the BRDF at each point to be a linear combination of a set of basis BRDFs, defined by a set of reference objects. They assume a parametric reflectance model and known shape of the surface. Building further on the idea of linear approximation of basis BRDFs, Goldman et al. [46] define a Wards model for each basis material, and iteratively estimate surface normals and parameters for the reflectance models.

(a)   (b)
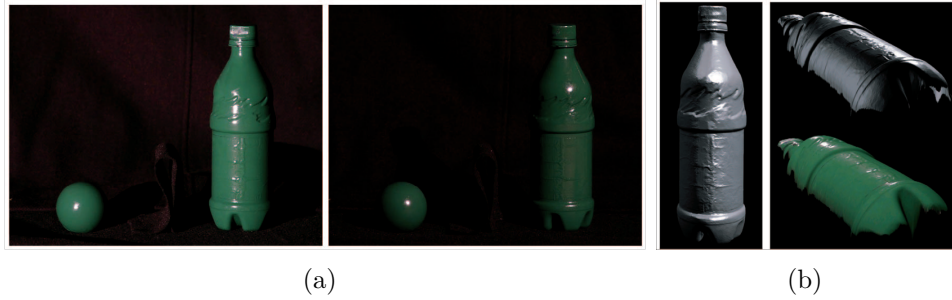
Figure 2.7: Example based photometric-stereo [47]

Example based methods have been made popular to obtain surface reconstruction for arbitrary BRDFs by Hertzmann et al. [47]. By introducing a reference object of a known shape and of same material as the object, images are obtained for different illumination directions with the same view point. The normal map of the object is computed by matching the intensity of the object at each pixel in image with a similar corresponding pixel of the reference object. If the intensity matches with the intensity of reference pixel, along every illumination direction, the known normal at the reference pixel is attributed to the corresponding pixel in the image of the object (Figure 2.7). It is one of the simplest algorithms in the field of photometric reconstruction. Although this method provides a very intuitive and simple solution, the major limitation is the necessity of the reference object in each case.

Although, there are methods which [48] estimate the parameters of reflectance models,they are not often used in photometric stereo for shape estimation because of their complexity.

2.4.3 Using reflectance properties

Another direction for reconstructing surfaces with arbitrary BRDF's is to exploit properties of the BRDF such as symmetry, isotropy and reciprocity. In this way, arbitrary BRDFs can be handled efficiently without the complexity of estimating

the parameters of the reflectance models. We know that the BRDF $f(\theta_i, \phi_i, \theta_o, \phi_o)$ is defined as the ratio of reflected radiance along the outgoing direction (θ_o, ϕ_o) to the received irradiance along the incoming direction (θ_i, ϕ_i).

Reciprocity ensures that the BRDF is symmetric along the incoming and outgoing directions,(i.e. $f(\theta_i, \phi_i, \theta_o, \phi_o) = f(\theta_o, \phi_o, \theta_i, \phi_i)$).

Isotropy implies that the BRDF is invariant with respect to rotation of the surface around the surface normal (i.e. $f(\theta_i, \phi_i, \theta_o, \phi_o) = f(\theta_i, \theta_o, |\phi_i - \phi_o|)$)

One of the early and important contributions in this direction was by Magda et.al [49] who introduced Helmholtz reciprocity and radiance constancy along a ray of light. The reciprocity constraint eliminates the BRDF term and geometric information of the surface is obtained from the known measurements. This was extended by Zickler et al. [50] to obtain reconstruction using Helmholtz stereopsis for the multiple view case.

Alldrin et.al. [12] proposed surface reconstruction for arbitrary BRDFs by exploiting the isotropy property of the BRDF. In [12], a dense configuration of lights on a circle is used to exploit symmetry in measurements of irradiance from the image. Such symmetry arises from the isotropy exhibited by the BRDF. The symmetry in intensities observed at each pixel is exploited to constrain the normal to a plane. This gives the surface gradient direction which is used to estimate the shape up to isocontours of constant depth. With further user interaction, Euclidean reconstruction can be achieved.

A subsequent work by Alldrin et al. [51] proposed a method for shape and reflectance recovery for non-parametric BRDFs. They however assume a non parametric model even for the basis BRDFs (unlike [46]), and assume a bivariate approximation for the BRDF.

The properties of isotropy and reciprocity in the BRDFs are also capable of resolving the GBR ambiguity (see (2.12)) of the uncalibrated photometric stereo [11]. Under a GBR transformation, the symmetry in the radiance functions along the isotropic and reciprocal curves is lost. By finding the transformation which restores the symmetrical structure, the GBR parameters can be estimated and Euclidean reconstruction can be established. This transformation is derived when at least a single isotropic pair in two images and a single reciprocal pair in a single image is known.

Recently, Chandraker et al.[13] presented a theory of differential photometric stereo in which the information from spatial and temporal gradient is used to obtain a photometric constraint which gives a linear relationship between the gradients. This is independent of the light source directions and the BRDF. When the light source is constrained to move on a circle, they recover isocontours of constant depth and gradient, and given a little additional information (normals along a boundary, depth at a point), they recover the Euclidean shape.

2.5   Recovering the surface

In the previous sections, we have reviewed both Lambertian and non-Lambertian photometric stereo. The direct result obtained from photometric stereo is a normal map while the final, desired result is to recover the 3D surface. The normal map essentially is a depth gradient field and therefore, it can be used for finding the 3D surface as a depth/height field. The simplest method to recover the depth field is by directly integrating the normal map. The depth gradients p, q can be derived from the estimated normal $\mathbf{n} = \begin{bmatrix} n_x & n_y & n_z \end{bmatrix}$ as $p = \frac{\partial z}{\partial x} = \frac{n_x}{n_z}$ and $q = \frac{\partial z}{\partial y} = \frac{n_y}{n_z}$. In terms of a unit normal $\hat{\mathbf{n}}$, we can relate the estimated normal and the depth

gradients as,

$$\hat{\mathbf{n}} = \frac{\begin{bmatrix} n_x & n_y & n_z \end{bmatrix}}{\sqrt{n_x{}^2 + n_y{}^2 + n_z{}^2}} = \frac{\begin{bmatrix} p & q & -1 \end{bmatrix}}{\sqrt{p^2 + q^2 + 1}} \tag{2.15}$$

In order to integrate the normal map for depth estimation, the surface normal has to satisfy the condition of integrability. Integrability ensures that the estimated depth remains the same irrespective path used for integration. Mathematically, this means

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x} \tag{2.16}$$

Assuming that the surface is integrable, the depth at a point $z(x, y)$ can be estimated from integration as

$$z(x, y) = z(x_0, y_0) + \int_{(x_0, y_0)}^{(x, y)} (p\mathrm{d}x + q\mathrm{d}y) \tag{2.17}$$

where, $z(x_0, y_0)$ is the initial depth.

In practical cases, the obtained gradient field is very rarely integrable due to noise present in the images or in the estimation process. There are a series of works that deal with this problem and propose methods to generate accurate depth maps. Frankot and Chellappa [52] project the non-integrable gradient field on to a set of integrable slopes using the Fourier basis functions. Several variants of this approach have been proposed by either choosing a different basis function or using a redundant non-orthogonal set of basis functions (for example, shapelets). Agrawal et al.[53] propose a generalized theory based on applying anisotropic weights to the gradient field while integration. They also propose different solvers to achieve trade-offs between different desired features in the reconstructed depth field. More recently, Harrison and Joseph [54] use a zero-mean Gaussian model for image noise and propose a method to estimate a maximum likelihood solution for depth map. Instead of treating the noise in the gradient field, they begin with

handling the noise in the images and continue throughout the process of surface generation.

2.6   Recovering surface reflectance

Photometric stereo methods are used not only for estimating the shape but also the reflectance of the object surface from the images. Most of the methods assume a parametric reflectance model (most commonly Lambertian). For Lambertian reflectance model, the BRDF is constant and equal to the albedo which is recovered in a straightforward manner as observed in the Sections 2.2,2.3. To capture the reflectance of objects which have a different reflectance, more complex approaches are adopted. This section also contains references to some of the works described in Section 2.4.2, but the emphasis here is on recovering the reflectance while Section 2.4.2 focuses on shape estimation.

Goldman et al. [46] make the assumption that most materials in the real world can be can be modeled as a weighted combination of few fundamental materials. They use Ward's reflectance model for each material. In an iterative framework, they estimate the normal map, weight maps and the BRDF parameters for each material at every pixel. Although naïve because of making many assumptions such as known lighting, no shadows and interreflections, this work is one of the well established works. Zhou et al. [2] recover the spatially varying BRDFs in a similar manner by inferring a set of basis BRDFs and their mixing weights at each pixel (see result in [2]). However, their method assumes that the basis materials are isotropic which can handle more general cases than other methods which assume a paramteric model for the BRDF.

While these methods use the estimated BRDF for rendering purposes, the esti-
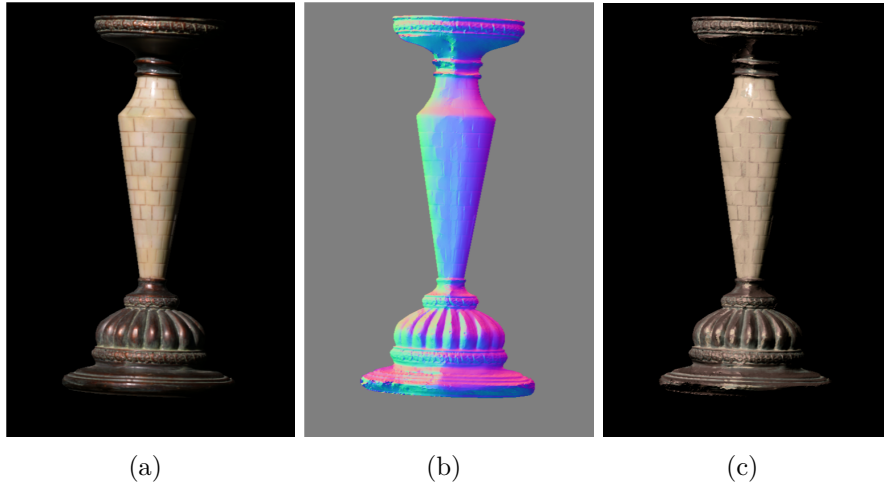
Figure 2.8: Rendering using the reflectance obtained from photometric stereo [46]. (a) input image, (b) estimated normal map and (c) rendered image

mated BRDF can also be used in the problem of material classification. There are several works which use BRDF for material classification [22, 23]. We capture a subset of the 4D BRDF and apply it for the problem of ink classification in Chapter 6.

Chapter 3

Auto-Calibrating Photometric Stereo using Ring Light Constraints

Uncalibrated photometric stereo algorithms use factorization to estimate the normals and the light directions from the pixel intensities. However, these estimates are obtained up to a 3×3 linear ambiguity. By imposing integrability constraints, the ambiguity is further reduced to the generalized bas-relief (GBR)ambiguity [16]. Various additional assumptions such as known pixels of equal intensity [32], separable isotropic highlight [55], specular spikes [56, 19], diffuse-maximum point [17], or constraints on light source positions [34], have been made on the scene to resolve this ambiguity. Uncalibrated photometric stereo is also referred to as auto-calibrated photometric stereo because the algorithm simultaneously estimates the light directions while estimating the normal directions. The main advantage of auto-calibration is that no explicit calibration object (such as a shiny mirror sphere) is required in the scene.

While there are various choices for resolving the GBR ambiguity, constraints on light source positions are known to offer additional advantages. Zhou and Tan [34] exploit geometric constraints derived from multiple (>5) light sources located on the same circle to automatically calibrate photometric stereo. They also show that by constraining the light sources to a ring, more general geometry (such as non-integrable surfaces) and weak perspective effects in lighting can be dealt with. Besides solving the GBR amiguity, the circular light configuration has also been used elsewhere in photometric stereo. Alldrin et al. [12] make use of circular light constraints to recover the azimuth angle of normals on a surface with isotropic materials. This can recover the isogradient contours of the object and recover the full Euclidean structure with some more additional assumptions. Jones et al. [57] use a circular light configuration to illuminate the face in a head mounted arrangement for real time photometric stereo.

We explore the circular light arrangement for solving the ambiguity in uncalibrated photometric stereo. While we also use light sources located on a circle, our focus is on the case of uniformly distributed lights. This simple and special case helps us to obtain stronger results in both theory and practice. More specifically, there are two key differences in our approach when compared to [34]. First, Zhou and Tan [34] show that uniformly distributed lights on a circle can recover the surface normals up to a planar rotation and scaling ambiguity. These ambiguities are resolved using two pairs of corresponding normals in two views with a 1D exhaustive search. In comparison, we show the rotation and the scaling ambiguities are independent of each other. Therefore, we can solve them separately with just one pair of corresponding normals from two views. Furthermore, our method has a closed-form solution, which is more efficient and precise than the 1D exhaustive search.

In this chapter, we propose an auto-calibration method for photometric stereo. Our method exploits constraints placed on light sources to recover their positions

and the surface normals of an object up to a scaling and planar rotation ambiguity. The ambiguity is resolved with multi-view consistency constraints leading to a Euclidean reconstruction of the geometric shape of the object. Our main contribution in this chapter is a closed form solution to the scaling and planar rotation ambiguities using the ring light constraints and estimate the normals maps. Additionally, we also demonstrate that using multi-view constraints, we can recover an extended reconstruction of the object surface beyond 2.5 D (beyond a single camera's field of view) unlike the conventional photometric stereo algorithms.

3.1   Related Work

We briefly review works related to uncalibrated photometric stereo and multiple-view photometric stereo in this section.

A major breakthrough in photometric stereo was made with Hayakawa's [58] proposition of uncalibrated photometric stereo. Hayakawa demonstrated that light source directions and surface normals can be recovered from the intensity information itself by factorizing the intensity matrix. The normals recovered by this method were up to a linear ambiguity, which can be resolved up to a bas-relief ambiguity [16]. Several works were proposed later to resolve this ambiguity[58, 34, 32, 55, 56, 19, 18, 37]

Recent works suggest that constraints placed on the light sources can offer a significant advantage in solving the photometric stereo problem under more general conditions. Alldrin et al. [12] propose a dense configuration of light sources on a circle to recover the azimuth angle and thus constrain the normal at each point to a known plane. In this case, the light source positions were assumed to be known. Zhou and Tan [34] propose ring light constraints to solve the photometric

stereo problem on Lambertian object with unknown light source positions. Zhou and Tan[34] make use of a ring of light sources to obtain the shape up to a mirror transformation, scaling and rotation ambiguity. These ambiguities can be further resolved by various constraints, such as uniformly distributed lights and two corresponding normals in two views. Further, they require a 1D exhaustive search to solve for the scaling and rotation parameters. We focus on the special case of uniformly distributed lights. We derive analytical solutions of the remaining ambiguities with just one pair of corresponding normals in two views.

We also review some relevant works which combine photometric stereo with multi-view stereo to recover the full 3D geometry of the object. Hernandez et al. [1] designed a multiple view photometric stereo algorithm for Lambertian surfaces. Wu et al. [59] propose a method to fuse multi-view stereo and photometric stereo with uncalibrated illumination. Zhou et al. [2] propose a multi-view photometric stereo that works on surfaces with spatially varying isotropic materials. Our method can be straightforwardly combined with structure-from-motion techniques, to provide a complete 3D reconstruction.

3.2 Formulation

3.2.1 A ring of point light sources

Let us consider point light sources located on a circle of radius r at a regular interval of t radians with respect to the centre of the circle. These light sources, one at a time, are used to illuminate an object. A camera is located at the centre of the ring, and the camera axis is perpendicular to the plane of the ring, as shown in Figure 3.1. We assume the object to be Lambertian and the light sources located far away from the object to assume directional illumination. If the object is located

36
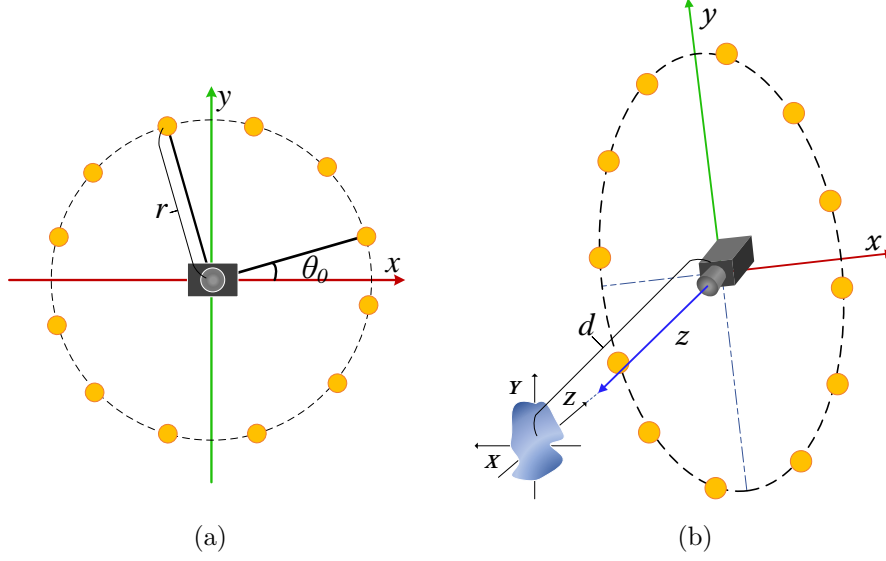
Figure 3.1: Schematic diagram of the acquisition setup. (a) The top-down view of the ring-light and camera arrangement. (b) The setup is used to capture the image of an object. The xyz coordinates represent the camera coordinate system and XYZ coordinates represent the global coordinate system.

at the origin of the world coordinates and the circular light–camera configuration is located at a distance d from the origin, the position of an arbitrary light source can be written as $\mathbf{l} = \begin{bmatrix} r\cos\theta & r\sin\theta & d \end{bmatrix}$, where θ is the angular displacement of the light source as shown in Figure 3.1. Let us assume that the angular position of the first light source is θ_0. Since the lights are equally spaced, the angular positions of the rest of the N light sources can be determined parametrically as $\theta = \theta_0 + (i-1)\,t$, where $i = 1, 2, ..., N$. If the normal at any point P on the object is denoted by $\mathbf{n} = \begin{bmatrix} n_1 & n_2 & n_3 \end{bmatrix}^T$, the intensity for a Lambertian object is

$$I = \rho\,\mathbf{l} \cdot \mathbf{n} = \rho(n_1 r \cos(\theta) + n_2 r \sin(\theta) + n_3 d) \tag{3.1}$$

where ρ denotes the albedo of the object at point P.

The parameters to be determined in order to recover the light source positions are the radius r, angular interval t, the distance d and the initial angular position θ_0. The parameters r and t can be measured easily and we can safely assume them

37

to be known. However, d and θ_0 are unknown. θ_0 is established when the camera is mounted manually at the centre of the ring light while d changes when the ring light and camera are rotated in multiple views.

The light sources are constrained to lie on a circle. This constraint offers an intrinsic parametric structure such that knowledge of a few parameters (d, θ_0) is sufficient to recover the position of all other light sources, and consequently the surface normals. For N light sources, we stack the equations of the form as (3.1) as,

$$\mathbf{L} = \begin{bmatrix} r\cos\theta_0 & r\sin\theta_0 & d \\ r\cos(\theta_0 + t) & r\sin(\theta_0 + t) & d \\ \vdots & \vdots & \vdots \\ r\cos(\theta_0 + (N-1)\,t) & r\sin(\theta_0 + (N-1)\,t) & d \end{bmatrix}. \tag{3.2}$$

We can now write the relationship between the observed intensities, light source direction and the surface normals as,

$$\mathbf{I} = \rho\mathbf{L}(d, \theta_0) \cdot \mathbf{n} \tag{3.3}$$

where $\mathbf{I}$ denotes the matrix of image intensities observed at point P due to the corresponding light sources.

3.2.2 Reconstruction ambiguities

We might assume arbitrary value for the distance d and the initial rotation angle θ_0 to solve (3.3). The solution will be related to the true result by a planar rotation and a scaling, as stated in the following proposition.

Proposition 1 *The true normal $\mathbf{n_g}$ of a Lambertian surface illuminated by a ring-light at a distance d_g can be recovered up to a classical bas-relief ambiguity*

compounded with a planar rotation ambiguity, i.e.

$$
\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} \cos\triangle & \sin\triangle & 0 \\ -\sin\triangle & \cos\triangle & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{d_g}{d} \end{pmatrix} \begin{pmatrix} n_{g1} \\ n_{g2} \\ n_{g3} \end{pmatrix} \tag{3.4}
$$

$$
\text{or} \quad \mathbf{n} = R_\triangle S \mathbf{n}_g \tag{3.5}
$$

Here, $\triangle = \theta_* - \theta_0$, d and θ_* are the assumed distance and initial rotation angles respectively, $\mathbf{n}$ is the recovered normal.

The proof of this proposition is provided in the appendix. We can see from Proposition 1 that the normal recovered in a single view is up to two ambiguities - 1) a scaling ambiguity, and, 2) a planar rotation ambiguity. The scaling or classical bas-relief ambiguity is a function of the distance d. The planar rotation ambiguity is a function of the initial angle θ_0. Mathematically, this implies that many solutions exist for the light direction and normal vectors, such that their scalar product remains unchanged. Intuitively, the distance d affects the light directions and normal vectors as illustrated in Figure 3.2. The images captured of an object using the ring light device from a certain distance (d_1) will be the same as those captured from a smaller distance $(d_2 < d_1)$ of a proportionately flatter version of the same object.

Zhou and Tan [34] solved this compound ambiguity by two corresponding normals from two views with a 1D exhaustive search. In the following, we show that these two ambiguities can be solved step-by-step in closed form solution, which requires only one pair of corresponding normal directions in two views. This separation is possible, since R and S operate on the different components of a normal direction. R changes only the first two components, while S changes only the third component of a normal direction. Due to the independence of these two

ambiguities, the step-by-step solution is possible and is favorable than the method in [34], where a brute force search is used to determine the final solution.
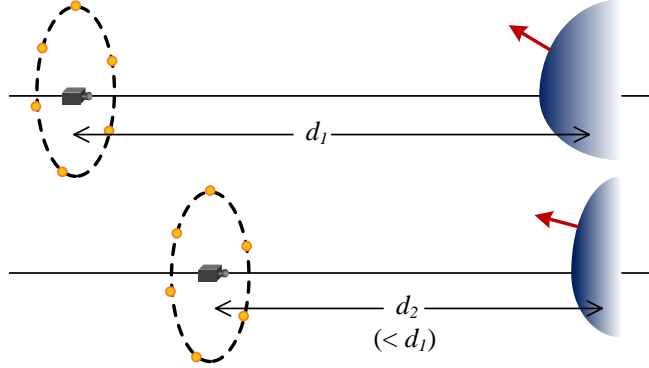
Figure 3.2: Visualisation of the scaling ambiguity

3.2.3 Consistency constraint from two views

The recovered normals are up to a scaling or classical bas-relief ambiguity and a planar rotation ambiguity as noted from Proposition 1. We note that these two ambiguities are independent of each other and therefore, while solving for one of these ambiguities, we do not consider the effect of the other. In this section, we derive an analytical solution for resolving these two ambiguities using constraints from multiple views in a step-by-step manner.

Estimating the scaling ambiguity: If $\mathbf{n}_1$ and $\mathbf{n}_2$ are the estimated normals of a point P in two views and $\mathbf{n}_{1g}$ and $\mathbf{n}_{2g}$ are the corresponding true normals, then

$$\mathbf{n}_1 = S_1\mathbf{n}_{1g} \text{ and } \mathbf{n}_2 = S_2\mathbf{n}_{2g} \qquad (3.6)$$

where S_1 and S_2 are the scaling ambiguities. Note that the object can be at different distances to the camera in the two views such that S_1 and S_2 can be

different. Here, all the normal directions are defined in the local camera coordinate system.

Let the two views be related by a camera transformation matrix T i.e.,

$$\mathbf{n}_{1g} = T\mathbf{n}_{2g} \tag{3.7}$$

where, T can be estimated using structure from motion (SfM). In other words, from (3.6),

$$\mathbf{n}_1 = S_1 T S_2^{-1} \mathbf{n}_2 \tag{3.8}$$

$$\mathbf{n}_1 = E\mathbf{n}_2 \tag{3.9}$$

where $E = S_1 T S_2^{-1}$

We have two unknowns $s_1 = \frac{d_1}{d_{1g}}$ and $s_2 = \frac{d_{2g}}{d_2}$ corresponding to the ambiguities S_1 and S_2. We can rewrite (3.9) as $\mathbf{n}_1 \times E\mathbf{n}_2 = 0$. This vector product gives the following independent equations.

$$A_{11}s_1 + A_{12}s_1 s_2 + A_{13}s_2 + A_{14} = 0 \tag{3.10}$$

$$A_{21}s_1 + A_{22}s_1 s_2 + A_{23}s_2 + A_{24} = 0 \tag{3.11}$$

$$A_{31}s_2 + A_{32} = 0 \tag{3.12}$$

Here, A_{ij} are constants in terms of the normals $\mathbf{n_1}$, $\mathbf{n_2}$ and the elements of the transformation matrix T. These parameters are provided in the appendix. We can solve these set of equations to analytically determine the parameters s_1 and s_2 as,

$$\begin{aligned} s_2 &= \frac{-A_{32}}{A_{31}} \\ s_1 &= \frac{-(A_{13}s_2 + A_{14})}{A_{11} + A_{12}s_2} = \frac{-(A_{23}s_2 + A_{24})}{A_{21} + A_{22}s_2} \end{aligned} \tag{3.13}$$
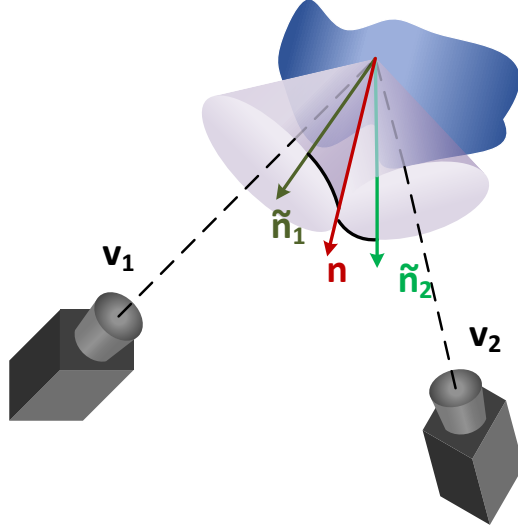
Figure 3.3: The two view constraints used to estimate the scaling and rotation ambiguities. The estimated normals $\tilde{\mathbf{n}}_1$ and $\tilde{\mathbf{n}}_2$ in the two views and the true normal $\mathbf{n}$ lie on the cone centred on the camera axes $\mathbf{v}_1$ and $\mathbf{v}_2$.

Estimating the rotation ambiguity: After determining the scaling ambiguity, we obtain the surface normal in the camera coordinate system up to a planar rotation given by θ_0. We note that θ_0 is fixed when the camera moves. Therefore, we can determine the rotation ambiguity using a simple auto-calibration method as described below.

If $\mathbf{n}_1$ and $\mathbf{n}_2$ are the estimated normals of a point P in two views and $\mathbf{n}_{1g}$ and $\mathbf{n}_{2g}$ are the corresponding true normals, then

$$\mathbf{n}_1 = R\mathbf{n}_{1g} \text{ and } \mathbf{n}_2 = R\mathbf{n}_{2g} \tag{3.14}$$

where R is the planar rotation ambiguity which rotates the true normals about the camera axis in the xy plane of the local coordinate system. Since the same ring light device is used in both the views, the rotation angle is the same in both views. Again, the normal directions are all defined in their local camera coordinate system.

If $\mathbf{v}_1$ and $\mathbf{v}_2$ are the camera view axes in view1 and view2 respectively defined in the global coordinate system, then we have the following relationships,

$$\tilde{\mathbf{n}}_1^T \mathbf{v}_1 = k_1 \qquad \tilde{\mathbf{n}}_2^T \mathbf{v}_2 = k_2 \qquad\qquad (3.15)$$

Here, $\tilde{\mathbf{n}}_1$ and $\tilde{\mathbf{n}}_2$ are the estimated normals in the global coordinate system and k_1 and k_2 correspond to the cosine of the half angle of the cone in view1 and view2. If R_1 and R_2 are the rotation matrices from camera - calibration in the two views, $\tilde{\mathbf{n}}_1$ and $\tilde{\mathbf{n}}_2$ can be computed as $\tilde{\mathbf{n}}_1 = R_1 \mathbf{n_1}$ and $\tilde{\mathbf{n}}_2 = R_2 \mathbf{n_2}$

The planar rotation ambiguity constrains the estimated normal to a cone whose axis is the camera view direction. The true normal $\mathbf{n}$ is unique in the global coordinate system and must lie on the cone in both views as shown in Figure 3.3. If $\mathbf{n}$ is the true normal in the global coordinate system, we have

$$\mathbf{n}^T \mathbf{v}_1 = k_1 \qquad \mathbf{n}^T \mathbf{v}_2 = k_2 \qquad\qquad (3.16)$$

Since k_1 and k_2 are known from (3.15), we have two linear equations in terms of the components of $\mathbf{n}$ from (3.16). We can recover $\mathbf{n}$ up to a 1D linear space from (3.16). Further we have the constraint $\|\mathbf{n}\| = 1$. From this, we can solve all the components to estimate $\mathbf{n}$. After recovering $\mathbf{n}$, we can estimate θ_0 by recovering the angle between the projections of $\mathbf{n}$ and $\tilde{\mathbf{n}}_1$ on the xy plane.

Once the scaling and rotation ambiguity parameters are determined, the normal maps can be estimated accurately.

3.2.4 Multiple view extension

In recovering the light source positions, we have used constraints from multi-view geometry. We have shown that a transformation matrix T between the two views is sufficient to resolve the scaling ambiguity. This transformation aligns the normal vectors of the object in the global coordinate system. Therefore, we can combine the geometry from multiple views to obtain an extended reconstruction of the object. If z_1 is the depth map obtained in the first view, say view1, and z_2 is the depth map obtained in the second view, say view2, the transformation T as described in (3.7) can be used to register the depth maps as

$$z_1 = T z_2 \tag{3.17}$$

In this way, we can combine the depth maps from multiple views and obtain an extended reconstruction of the object beyond 2.5 D.

3.3 Experiments

3.3.1 Experimental setup

We implement our method using 36 equally spaced LED lights on a circular disk forming a ring light source with a Point Grey Grasshopper GRAS-20S4C-C camera located at the centre of the ring of lights. The camera is fixed such that the camera axis is orthogonal to the plane of the ring. (Note that the camera is mounted manually; hence, it might not precisely sit in the circle center and its orientation could be slightly off.) As explained in the previous section, our algorithm resolves the two ambiguities using ring light constraints and two-view constraints. To recover the transformation between the two views, as described in Section 3.2, we

calibrate the camera using a checker board pattern and the camera calibration toolbox for MATLAB [60]. Having estimated the camera matrix from calibration, we use SIFT features, refine them using RANSAC and match them for detecting point correspondences in a given pair of views. To deal with textureless objects, we put our test objects on a checker board pattern to facilitate matching.

3.3.2 Results and discussion

We implement our method on three real world objects, 'Cat', 'Teddybear' and 'Teapot'. 'Cat' and 'Teddybear' are nearly Lambertian while 'Teapot' has a slightly specular surface. We assume an arbitrary value for d and θ_0 to first recover the normal maps up to a scaling and rotation ambiguity. For convenience, we assume $d = 10$ and $\theta_0 = 0$. As described in Section 3.2, we require a single pair of corresponding normals in two views to solve the rotation and scaling ambiguities. For estimating these normals, we compute point correspondences P and P' using SIFT.

The estimated scaling parameter or the distance is observed to vary slightly for different pairs of corresponding points. This could be due to the sensitivity of SIFT to the surface albedo in two views or corrupted intensities due to non-Lambertian behaviour of the surface. To minimise errors due to these factors we average the scaling parameter over 20 corresponding points to solve the scaling ambiguity. The parameter θ_0 is also estimated by averaging the values obtained for different pairs of corresponding normals. The estimated values for d and θ_0 for 'Cat' object are shown in Figure 3.4. We observe from the figure that the estimated distance term is accurate within 4% error. Although the estimated θ_0 seems to indicate a rather large error, practically the estimate is quite reasonable. This is because the angle is estimated as an inverse cosine, which is extremely sensitive for angles

(a)   (b)

Light Positions

(c)

Light Positions
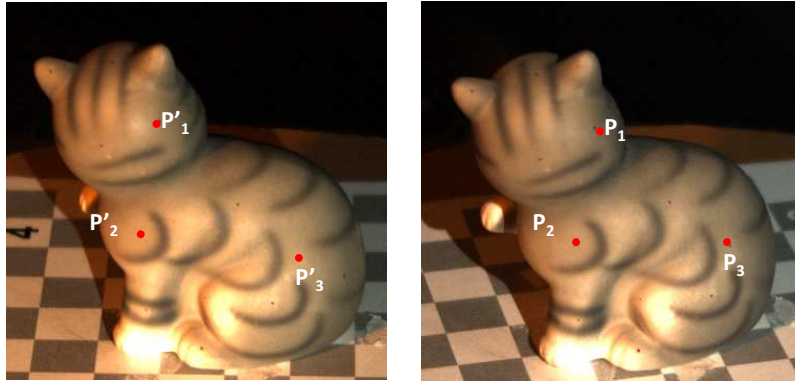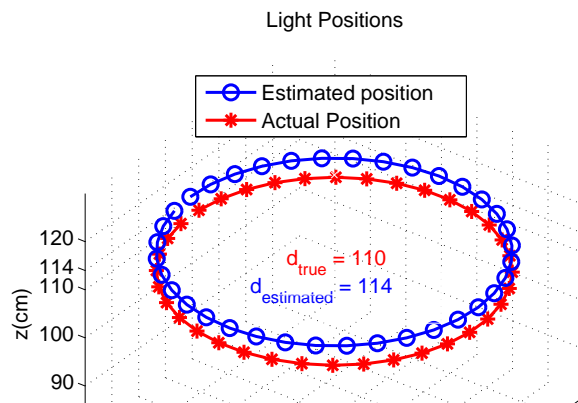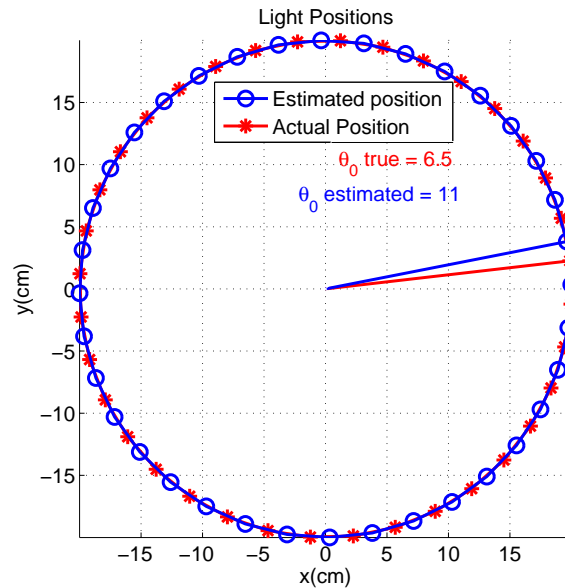
(d)

Figure 3.4: The images of 'Cat' object are captured from two views (a) and (b) and the corresponding points are estimated. We recover the distance and the initial angle using the multiple view constraints.
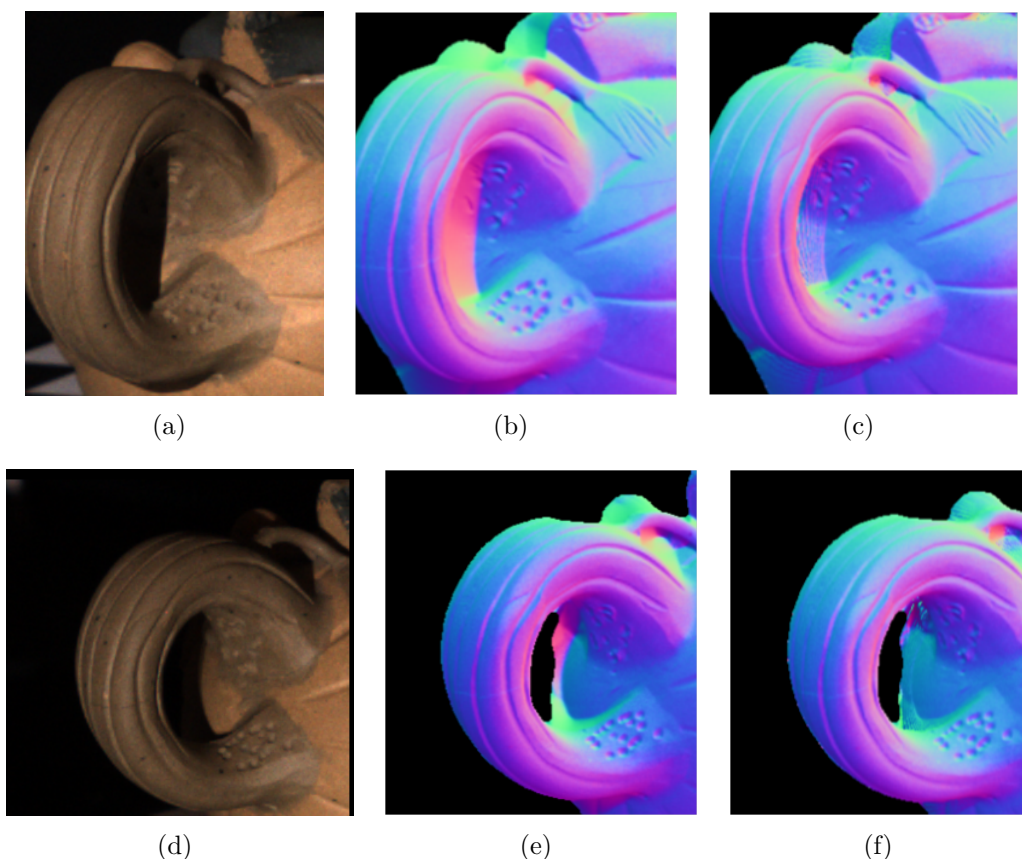
Figure 3.5: Two different views of 'Teapot' are shown here. In the first column, we have the sample image; in the second column, we have the reconstructed normal with shadows;in the third column, we have the reconstruction with outliers removed.

close to $0°$. This slight deviation could be due to issues in image registration while determining the transformation between the two views. Also, the effect of θ_0 is not very significant as long as the estimates are within an error of $\pm 5°$.

We use RANSAC [61] to eliminate the outliers before estimating the normal map. At each pixel, we use RANSAC to estimate normal selecting 3 samples randomly from the available 36 samples. We impose a threshold of 0.01 on the error for the data point to fit the model generated using the chosen samples. The normal maps in Figures 3.5(c),(f) show outliers removed using RANSAC.

After resolving the rotation and the scaling ambiguities, we estimate the normal map of the 'Cat' object. The normal map obtained is compared with the ground
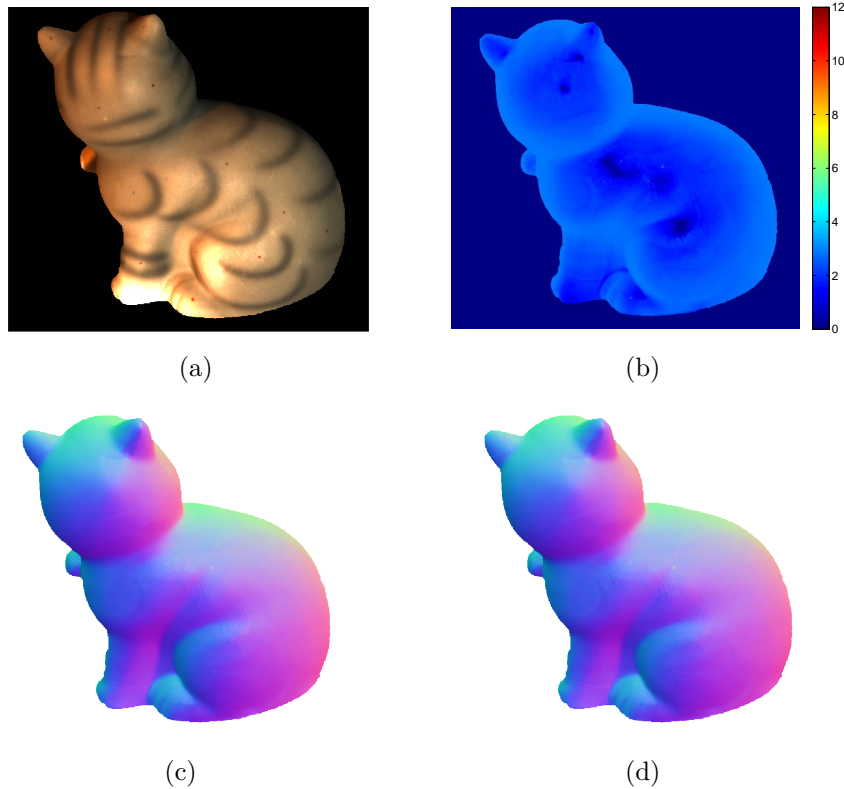
Figure 3.6: The 'Cat' object is nearly Lambertian. We have a sample image of the object in (a), the calibrated and the estimated normal map in (c) and (d) respectively. It can be observed from the error map in (b) show that the estimated normal map is very close to the calibrated normal map.

truth obtained from calibrated light directions using a mirror sphere as shown in Figure 3.6. As observed from the error map, the normals recovered from our method are very close to the ground truth. In uncalibrated photometric stereo, the recovery of normal map and lighting direction requires solving the GBR ambiguity after enforcing the integrability constraint. In comparison, we need multiple view constraints and the integrability constraint to calibrate the light positions completely. As shown in Section 3.2, the scaling parameters for two different views are estimated separately. Therefore, the object need not be located at the same distance to the camera in multiple views while resolving the scaling and rotation ambiguities.
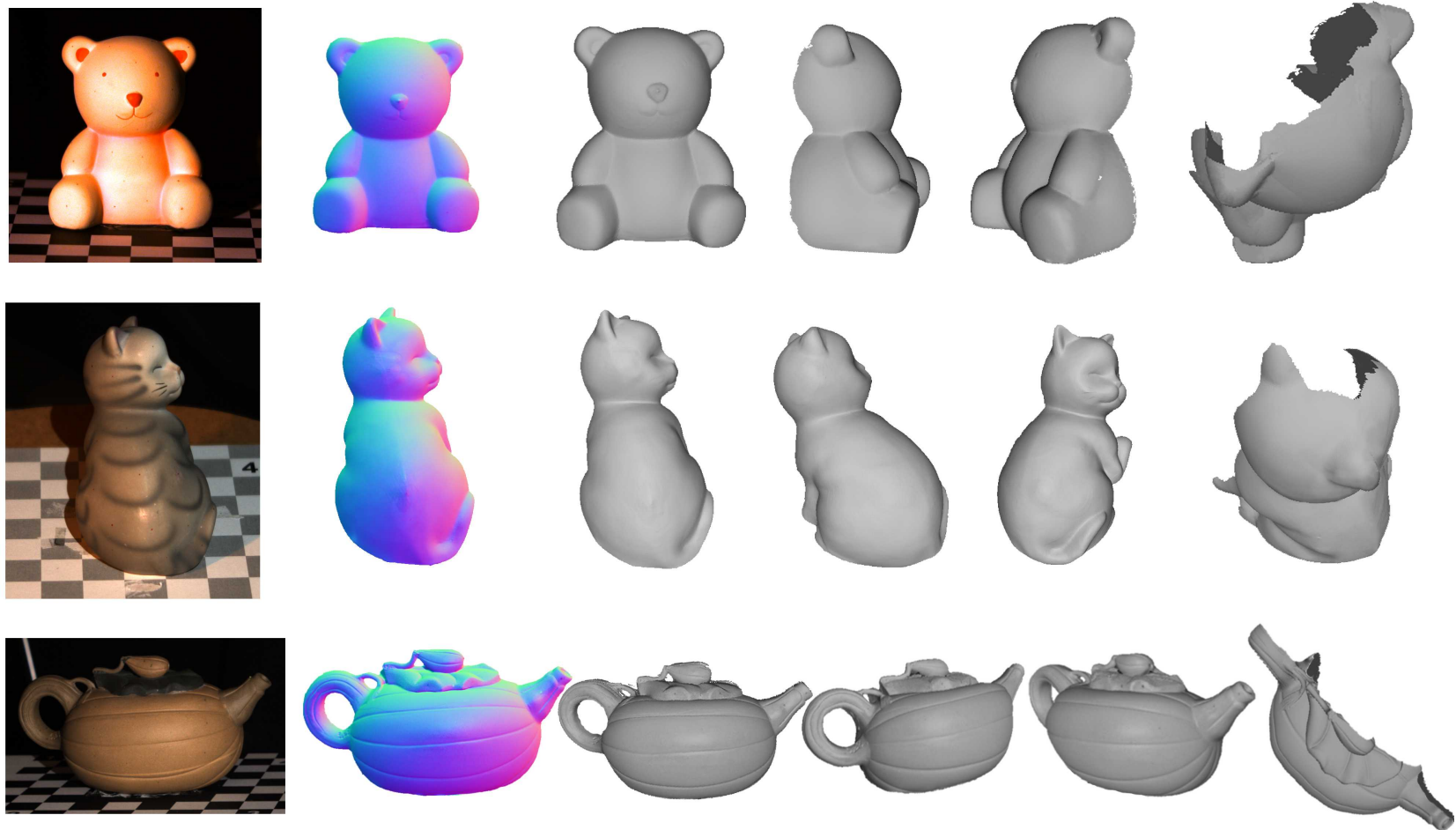
Figure 3.7: Experimental results on objects - 'Teddy bear','Cat' and 'Teapot'. In the first column, we have a sample image of the object; in the second column, we have the normal map; the third, fourth and fifth columns contain the depth maps by integrating the normals from three adjacent views; in the last column, we have the result of the merged depth map.

We next present an interesting application of our algorithm in fusing the depth maps from multiple views. As discussed in Section 3.2.4, we use the multi-view constraint in estimating the light source positions and fuse the depth maps from multiple views. We demonstrate this by combining the depth maps from a set of five adjacent views. We combine the individual depth maps by registering them using the Iterative Closest Point (ICP) algorithm [62]. The combined depth maps contain irregularities in the overlapping region. We average the depth estimates in the overlapping regions and smooth the positions and the normals as proposed by Diego et al. [21]. The individual and combined depth maps are as shown in Figure 3.7.

In this way, we successfully use the ring light constraints to solve the ambiguities and estimate the normal maps. We also use additional constraints from SfM to register depth maps in multiple views and obtain an extended 3D surface of objects beyond 2.5 D (which is usually the result of conventional photometric stereo algorithms). The algorithm, experiments and the results are also discussed in [63].

3.3.3 Limitations

In the proposed method, we assume that the surface is Lambertian. We remove the non-Lambertian components such as cast shadows, specular highlights as outliers using a simple threshold. In the case where the surface material has complex BRDFs or spatially varying BRDFs, the algorithm fails to estimate the true normals. Our method requires that the light sources are uniformly distributed on the ring, camera is located at the centre of the ring light with the camera axis perpendicular to the plane containing the ring lights. While our algorithm can absorb minor deviations, it fails when the camera is oriented differently or not

located concentrically with the ring light. However, as long as the assumptions are met closely, our algorithm can successfully estimate the surface shape using the ring light constraints.

3.4 Summary

We have presented an auto-calibration method for photometric stereo using constrained light sources. Using the ring light constraints and multiple view constraints, we determine the scaling and planar rotation ambiguities in a step-by-step method. Next, the normal maps are estimated after removing these ambiguities. Our results show that the recovered normal maps are very close to the ground truth obtained from manual calibration. We have also shown that multiple view constraint also helps in fusing depth maps from multiple views of the object. However, our method works only on Lambertian material and cannot handle objects with surface material containing significant non-Lambertian components.

Chapter 4

Auto-calibrating Photometric Stereo with Rectangularly Placed Light Sources

With the introduction of low cost depth cameras such as Microsoft's Kinect, it has become simple and commonplace to estimate the depth maps. However, the estimated depth maps are corrupted by high frequency noise. As discussed earlier, photometric stereo generates high quality normal maps, but, the surface estimates from photometric stereo contain low frequency noise. To enhance the result, high resolution surface normals are integrated with the low frequency depth estimates to generate 3D surfaces of high detail [64, 21].

One of the major limitations of these algorithms is that they require calibration of the light sources. In the previous chapter, we have seen that constraints on the light sources can successfully calibrate photometric stereo under unknown illumination. Since the Kinect's camera is usually placed on top or at the bottom of the television or monitor screens, it is an attractive option to use the rectangular

geometry of the screen to calibrate photometric stereo.

In this chapter, we explore rectangularly placed light sources, i.e. light sources placed at the four corners of a screen to automatically calibrate the photometric stereo problem. The geometry is very simple such that a user with no technical expertise can mount the four lights and automatically calibrate photometric stereo. The application of calibrating photometric stereo using rectangular light sources would be in integration of the normal map with the depth map from depth sensors, IR sensors for expanding the scope of photometric stereo. The detailed experimentation will be done in our future work.

In Section 4.1 we review some of the related work to our problem, specifically in the area of depth fusion and auto-calibration of photometric stereo. In Section 4.2, we derive a mathematical formulation to calibrate the photometric stereo problem using the rectangular geometry formed by light sources locations. In Section 4.3, we validate our algorithm on synthetic data and finally summarize our work in Section 4.4. Experiments on real data will be a part of the future work.

4.1   Related Work

Our work is concerned with the problem of automatically calibrating the photometric stereo from its images. The related works are already reviewed in Chapter 3. Here we review works related to fusing the result of photometric stereo with the result from other shape estimation techniques such as stereo, depth sensors, structure-from-motion. One of the important works in this direction is by Hernandez [1]. This work uses silhouettes from multiple view points to determine the object's visual hull and refine it using normals from photometric stereo. Another important work in this direction was by Nehab et al.[21]. In this work, the

normals and depth positions are registered and optimized to produce high quality surface reconstructions. Recently, depth sensors are extensively used in areas such as 3D reconstruction [65] and gaming [66] with the introduction of devices such as the Kinect [67]. However, the depth maps generated contain high frequency noise. To enhance the quality of depth maps, high quality surface normals are used. Anderson et al. [68] present a method to enhance the depth map quality from the Kinect by using high quality normal maps from photometric stereo. They use multispectral photometric stereo (different coloured light sources) to estimate the normal map to make it suitable for dynamic scenes. They perform calibration using constraint from structure-from-motion as proposed in [69]. Zhang et al. [64] propose a fusion technique of the result from the Kinect depth camera and normal map from three light photometric stereo. Their main aim is to preserve the depth discontinuities of the real scene and achieve it by explicitly modelling the depth discontinuities. While they produce promising results, they also need to calibrate the light sources and assume directional illumination. We present a method that exploits the inherent rectangular geometry of monitor screens for automatic calibration in photometric stereo.

4.2   Formulation

Uncalibrated photometric stereo algorithms are designed such that no prior knowledge of illumination directions are assumed. In this section, we briefly review the uncalibrated photometric stereo problem and propose a solution using constraints from rectangularly placed light sources

4.2.1   Uncalibrated photometric stereo

Under unknown illumination, shape estimates from Lambertian photometric stereo are up to a linear transformation. Let $\mathbf{I}$ be the matrix of pixel intensities ordered in such a way that the matrix element I_{ij} represents the intensity of the ith pixel $(i = 1, 2, \cdots n)$ in the jth image $(j = 1, 2, \cdots m)$, where n is the number of pixels and m is the number of images. If $\mathbf{L}$ represents the true lighting directions with the intensities and $\mathbf{B}$ represents the normal directions along with the surface albedo, on SVD decomposition, we obtain

$$\begin{aligned} \mathbf{I} &= \mathbf{UDV}^T \\ &= \left(\mathbf{UD}^{\frac{1}{2}}\right)\left(\mathbf{D}^{\frac{1}{2}}\mathbf{V}^T\right) \\ &= \tilde{\mathbf{B}}\tilde{\mathbf{L}} \end{aligned}$$

(4.1)

where, $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{L}}$ could differ from the true estimates $\mathbf{B}$, $\mathbf{L}$ by the 3×3 linear transformation A such that $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{A}^{-1}$ and $\tilde{\mathbf{L}} = \mathbf{A}\mathbf{L}$. Once A is estimated, the true lighting and normal directions can be determined.

4.2.2   Constraints from Four Rectangularly Place Light Sources

Let us suppose that four light sources are located at the corners of a monitor screen as shown in Figure 4.1. The camera is setup midway on top of the monitor screen and perpendicular to the plane of the screen. We consider the lighting directions on the projective plane. On the projective plane, a ray with orientation (l_x, l_y, l_z) is reduced to a point with coordinates given by $(l_x/l_z, l_y/l_z)$. Let us assume a coordinate system such that the camera axis is perpendicular to the projective plane and is reduced to a point on the plane. The true lighting directions form a rectangle on the projective plane (Figure 4.2).
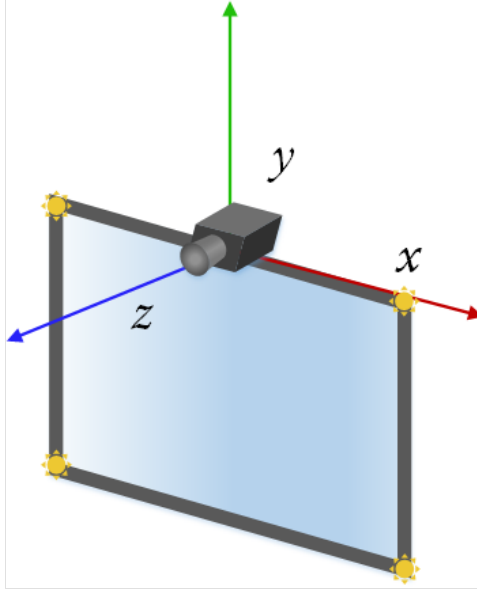
Figure 4.1: Diagram showing the position of four light sources and the camera with respect to the screen

From (4.1), we observe that on SVD decomposition, the lighting directions are distorted by the linear transformation A. These directions form an arbitrary quadrilateral on the projective plane (Figure 4.2). In geometry, two quadrilaterals on a plane can be related by a projective transformation. The projective transformation can be determined when the positions of the 4 corners of these quadrilaterals are known. Consumer monitor and television screens generally specify partial dimensions of the screen. Therefore, we can safely assume that information such as aspect ratio of the screen is known. Under this assumption, we can estimate the true lighting positions up to a scaling ambiguity S. If $\mathbf{L}'$ and $\tilde{\mathbf{L}}$ represent the matrices containing the coordinates of the rectangle formed knowing the aspect ratio and the quadrilateral on the projective plane, then

$$\mathbf{L}' = H\tilde{\mathbf{L}} \tag{4.2}$$

where, $\mathbf{L}'$ represents the scaled rectangle on the projective plane and H is the corresponding homography such that $A = SH$
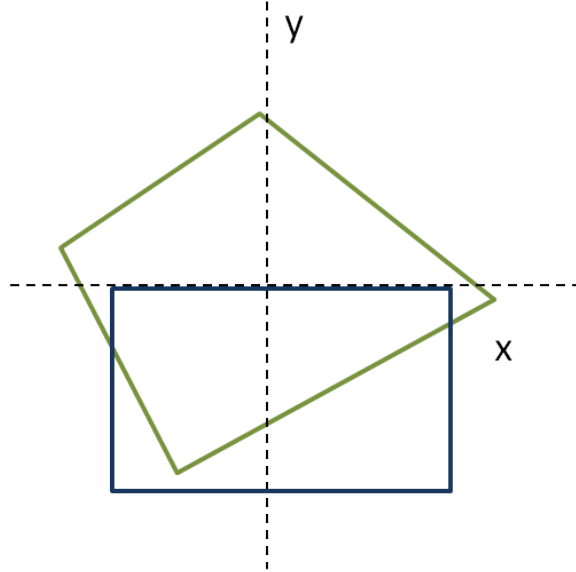
Figure 4.2: The true lighting directions form a rectangle (blue) and the distorted lighting directions (green) form an arbitrary quadrilateral on the projective plane

The scaling ambiguity S can be estimated using additional constraints. Since we use a monitor screen faced by the user, the face of the user itself can be used as an object for autocalibration. In such a case, the additional constraint can be derived from a generic mesh. The overall shape of the human face is generally similar among all human beings. Therefore, a generic face mesh can be used to obtain a few closely aligning normals so that S can be completely resolved using a least squares solution. Once S is determined, A can be fully determined and the normal map can be estimated consequently.

4.3 Experiments

We validate our approach on a synthetic face data. The face mesh shown in Figure 4.3(a) is illuminated by 4 point light sources located on a rectangle. We estimate the normal maps from the factorization of the intensity matrix as explained in (4.1). The aspect ratio is then assumed to be known and the normal map is estimated after determining the homography. From Figure 4.3(c), we observe that
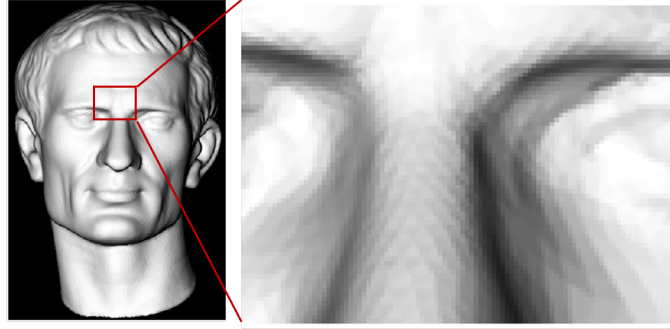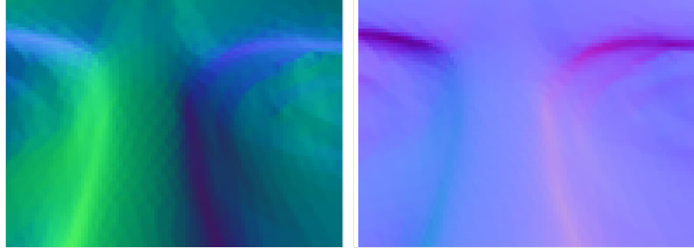
Figure 4.3: (a) Input image, (b) Normal map upto linear transformation, (c) Result after assuming known aspect ratio

the result after assuming known aspect ratio is still ambiguous.

We next use a generic face mesh and align it with the face mesh used in our experiments. We locate landmarks on the face image and warp it for alignment. The scaling ambiguity is determined and finally the normal map is estimated as shown in Figure 4.4(a). However, we observe from Figure 4.4(c) that the result from generic face mesh constraint does not completely solve the ambiguity. This is due to the fact the generic mesh does not fully coincide with the given face mesh at all points.

We further verify if the result is up to a global scaling ambiguity. In that case, a better approach for resolving the scaling ambiguity can provide us with accurate estimates of the normal map. Assuming known ground truth, we perform a brute force search to solve for the scaling ambiguity. From Figure 4.4(b),(d), we observe that a normal map with much less error is obtained suggesting a global scaling ambiguity. This concludes that with a better alignment of normals with the generic
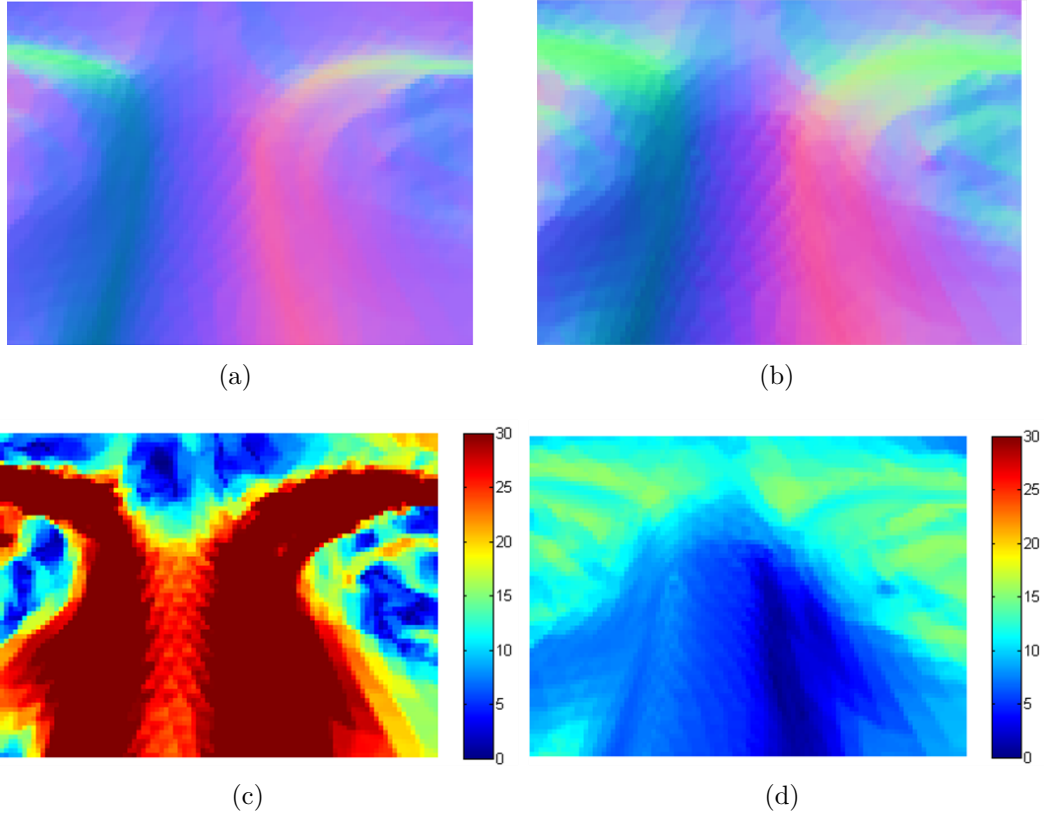
Figure 4.4: (a), (b) Normal maps from generic mesh constraint and brute force search, (c), (d) corresponding error maps.

mesh, we can successfully resolve the global scaling ambiguity.

4.3.1 Future work

The above section describes experiments on synthetic data. To validate our method on real data, we require images of a human face illuminated by four rectangularly placed light sources. While capturing images from such an arrangement, the user's face must be steady. We require light sources synchronized with the camera to quickly capture the images such that no apparent movements of the face are captured. We are in the process of developing such a system and leave it for future work.

4.4 Summary

In this chapter, we present an auto-calibration method for photometric stereo using rectangularly placed light sources. We implement our auto-calibration method on synthetic data and generate promising results. Our method can calibrate photometric stereo and generate high quality normal maps. The proposed pattern of light sources can be used in sensor fusion applications such as depth sensors, IR sensors. An important application can be to fuse high quality normal maps from photometric stereo with depth maps from depth sensors to generate high resolution depth maps.

Chapter 5

Surface Reconstruction using Isocontours of Constant Depth and Gradient

Most works in photometric stereo based 3D reconstruction [58, 70, 17] make the fundamental assumption that the surface is Lambertian. However, surfaces in the real world deviate from Lambertian behaviour to varying degrees. To include such objects, broadly two approaches are adopted - to remove the non-Lambertian observations as outliers [46, 71] or to assume a more complex reflectance model for the surface [43, 72]. The former approach results in a loss of information and is not always reliable while the latter approach increases the computational complexity.

Different from these approaches, there is a recent trend (e.g. [12, 73, 13]) towards exploiting more general reflectance properties of a surface for photometric 3D reconstruction. These properties include bilateral symmetry, isotropy, reciprocity and monotonicity. These approaches do not require assumption of the parametric BRDF of the surface reflectance or removal of highlights, and therefore, is more

general and robust in dealing with real surfaces.

In this chapter, we propose an automatic method to solve the 3D surface from iso-gradient and iso-depth contours up to a linear transformation. The key contribution made is the derivation of an equation system from the known contours and solving it to obtain the 3D surface. Unlike previous methods, our method does not require the user to specify additional information. It is potentially useful in applications of photometric stereo where the objects have isotropic reflectance, such as the internal organs of the body in computer assisted surgery or objects of cultural heritage.

5.1 Isodepth and Isogradient contours

We give a brief description of iso-depth and iso-gradient contours in this section.

5.1.1 Iso-depth Contours

Alldrin et. al. [12] uses the propoerty of bilateral symmetry of isotropic BRDFs to estimate the azimuthal angle of a surface normal. Bilateral symmetry implies that the BRDF is symmetric about the plane spanned by the normal and viewing vector with respect to the incident lighting direction. If we consider a surface patch with normal $\mathbf{n}$ and viewed from direction $\mathbf{v}$, bilateral symmetry ensures that the observed irradiance E is the same in two cases where light sources are placed symmetrically with respect to the plane containing $\mathbf{n}$ and $\mathbf{v}$ as shown in Figure 5.1. The orientation of this plane for each pixel can be determined from images captured by illuminating the scene from a circle of light sources.

This measurement provides the direction of the depth gradient but not the mag-
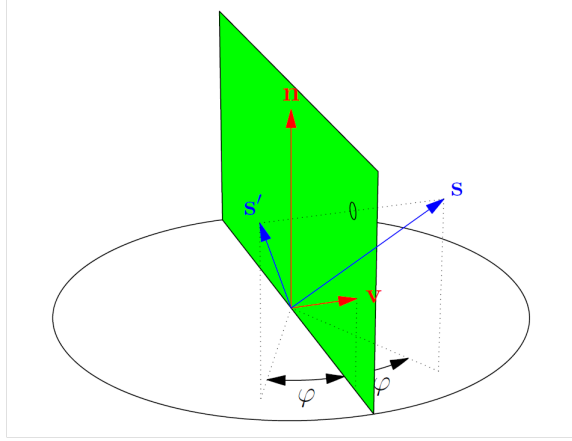
Figure 5.1: Symmetric light source vectors about the plane spanned by surface normal **n** and viewing direction **v** (Alldrin et al. [12])

nitude. For differentiable surfaces, this information is sufficient to recover the isocontour structure. Additional constraints such as uniform BRDF, specular highlights and shadows can be used to fully recover the Euclidean structure.

5.1.2 Iso-gradient contours

The work by Chandraker [13] focuses on using the image gradients (spatial and temporal) for recovering scene structure in the case of isotropic BRDFs. The temporal gradients I_t are obtained by capturing images from a differential pair of light sources (two light sources on a circle separated by a small distance) while the spatial gradients I_x, I_y are obtained by computing the image gradients. They obtain the set of image gradients on ratio images (images which are normalized with respect to a floodlit image) and derive a linear relationship between the image gradients from the fundamental equations of image formation. The photometric constraint is given by

$$I_x - \lambda I_y - \kappa I_t = 0 \tag{5.1}$$

where λ and κ are constants defined purely in terms of geometrical properties.

If we consider a scalar field $g(x,y) = \|\mathbf{n}(x,y)\|$, the gradient associated with this
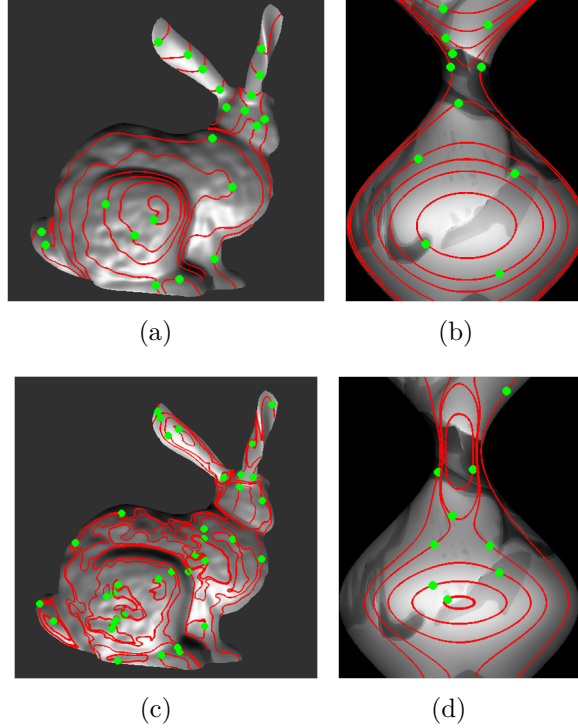
Figure 5.2: Iso-depth contours((a) and (b)) and Iso-gradient contours ((c) and (d)) marked out. (Chandraker et al. [13])

field is orthogonal to the level curves of $g(x, y)$. Therefore, the tangent along the level curves of $g(x, y)$ should be defined in the direction orthogonal to ∇g. This is proven to be equal to $\frac{1}{\lambda}$. This gives us the contours having a constant magnitude of the gradient norm.

5.2   Related Work

Isotropy and reciprocity are two reflectance properties to reconstruct general surfaces. One of the early and important contributions in this direction is [49], which introduced Helmholtz reciprocity to eliminate the dependence on the BRDF in 3D reconstruction. In [50], the work of [49] is extended further to obtain reconstruction using Helmholtz stereopsis for multinocular case. Helmholtz stereopsis is applicable in the case of arbitrary BRDFs and can give a direct estimate of

surface depth and normals while handling depth discontinuities.

Recent works introduce the concept of introducing constraints on lighting while exploiting properties of BRDF for estimating geometric information of the object. Alldrin et al. [12] propose surface reconstruction for arbitrary BRDFs by exploiting their isotropy. They use a dense configuration of lights on a circle to exploit symmetry in measurements of irradiance from an image. Such symmetry arises from the isotropy of BRDF. The symmetry of intensities observed at each pixel is used to constrain the normal at the pixel to a plane. This gives the surface gradient direction, which is then used to recover the isocontours of the constant depth. The true depth is recovered using additional cues in the image data or imposing additional constraints on the surface (e.g., shadows, specular highlights, Helmholtz Reciprocity, uniform BRDF). Tan and Zickler [73] further estimate iso-gradient contours by exploiting isotropy and reciprocity jointly. Iso-gradient contours consist of pixels with the same gradient magnitude. It has been shown [73] that the surface can be determined up to a linear transformation (a depth scaling combined with a shift) of the contours.

Recently, a theory of differential photometric stereo using a pair of differential light sources is proposed by Chandraker et al. [13]. This method employs the spatial and temporal gradient of image intensities to obtain a photometric constraint which gives a linear relationship between the gradients. This is independent of the light source directions and the BRDF. By allowing the light source to move on a circle, they recover isocontours of constant depth and gradient, and given additional information (normals along a boundary and depth at a point), they recover the entire surface.

While the approaches described in [12] and [13] promise to give surface information, it is limited to a set of iso-depth and iso-gradient contours. They require an additional set of constraints to be specified by the user in order to obtain a

full Euclidean reconstruction. These constraints include the normal specified at a point on the object or the depth along the boundary of the object in the image. We observe that given an arbitrary scene, it is not trivial for the user to specify a normal or define a depth along a boundary. Hence, we explore the possibility of using the iso-depth and iso-gradient contours to reconstruct the surface without additional information from the user.

5.3   Reconstruction with Isocontours

We propose an algorithm which makes use of the available information from iso-contours of constant depth [12] and gradient norm [13] to obtain depth at each pixel of the object.

5.3.1   Initial solution by integrating the contours

Let $P(x, y)$ be a point on the image of an object where isocontours of constant depth z and constant gradient $\|\nabla z\|$ intersect transversally. The iso-gradient contour gives us a set of points having constant $\|\nabla z\|$, but its value is unknown. If the value of $\|\nabla z\|$ is known on a contour, the depth z at each point on this contour can be computed by integrating the gradient vector along the curve. If we take z_0 to be the magnitude of depth at a given point $P_0(x_0, y_0)$ on an iso-gradient contour C, and a unit magnitude of the gradient $\|\nabla z\|$, then the depth at any arbitrary point on the curve is obtained by integrating the gradient along the curve as described below.

We know that the depth gradient vector is orthogonal to the iso-depth contour. Therefore, the direction of the depth gradient $\nabla \mathbf{z}$ at each point is obtained from
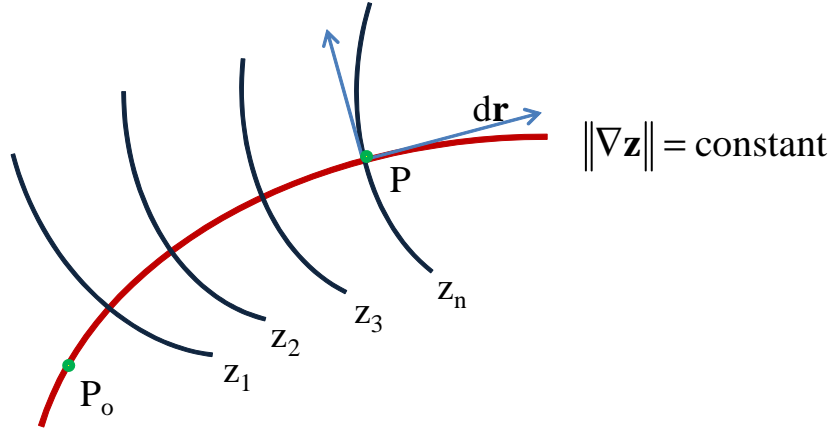
Figure 5.3: Integration of gradient vector along a given iso-gradient contour gives the depth at each point

the azimuthal angle [12].

$$\nabla \mathbf{z} = [z_x z_y] = \nabla \hat{\mathbf{z}} \|\nabla z\| \qquad (5.2)$$

where $\nabla \hat{\mathbf{z}}$ is the unit vector orthogonal to the iso-depth contour obtained from the azimuth angle, $[z_x, z_y]$ are the coordinates of the gradient direction $\nabla \mathbf{z}$.

The depth at arbitrary point P is given by

$$z(x, y) = \int_{P_0}^{P} \nabla \mathbf{z} \cdot \mathbf{dr} + z_0 \qquad (5.3)$$

where $d\mathbf{r}$ is the direction of the tangent vector to the curve.

Given the depth magnitude at each point on the iso-gradient contours, iso-depth contours can be plotted running through each of these points over the entire surface. The given iso-gradient contour need not necessarily intersect every pixel in the image. In such cases, we propagate the above process to the leftover regions by initiating another iso-depth contour from the region where depth is already known from the above and which passes into the leftover regions.

This serves as the initialization for our solution to the non linear system of equa-

tions discussed below. We have assumed a value of 1 for initializing $\|\nabla \mathbf{z}\|$. The reconstruction using iso-depth and iso-gradient contours for any other value will be upto a scaling ambiguity. This is an intrinsic ambiguity and can be solved with information such as BRDF symmetry [73].

5.3.2 Non-linear system of equations

Let us consider a pixel $P(x, y)$ on the image of the object. For every point P, we can define another point $P^*(x^*, y^*)$, such that P^* is along the iso-depth contour passing through P. From the definition of iso-depth contours, we can write $z(x, y) = z(x^*, y^*)$. At the same point P, we can define another point $P'(x', y')$ such that P' is along the iso-gradient contour passing through P. From the definition of iso-gradient contours, we can write $\|\nabla z(x, y)\| = \|\nabla z(x', y')\|$. The gradient at any point $P(x, y)$ can be approximated as

$$\|\nabla z(x, y)\| = [(z(x, y) - z(x + 1, y))^2 + (z(x, y) - z(x, y + 1))^2]^{\frac{1}{2}} \quad (5.4)$$

Using the above definitions, at every point $P(x, y)$, we have the following set of equations,

$$z(x, y) = z(x^*, y^*) \quad (5.5)$$

$$\begin{aligned}
\left[(z(x, y) - z(x + 1, y))^2 + (z(x, y) - z(x, y + 1))^2\right]^{\frac{1}{2}} = \\
\left[(z(x', y') - z(x' + 1, y'))^2 + (z(x', y') - z(x', y' + 1))^2\right]^{\frac{1}{2}}
\end{aligned} \quad (5.6)$$

Equations (5.5) and (5.6), which can be defined at every point $P(x, y)$ on the image plane, form a system of non linear equations. The solution to these system of equations over the entire image plane will give depth z at every point $P(x, y)$ without requiring any additional information.

5.3.3 Solving the non linear system of equations

The main challenges in solving the set of equations (5.5) and (5.6) are the following

:

- Since the number of pixels in an image, say of dimensions (640×480), is generally of the order of 10^5, the number of equations is also of the same order. Hence, we need a solver capable of solving the huge system of non linear equations .

- The nonlinearity associated with (5.6) poses another challenge. The large system of equations can be solved using a sparse matrix representation if the equations are linear. However, for a system of non linear equations, it is not straightforward to solve directly.

We need to reformulate the problem into a system of linear equations. To convert the system of non linear equations in (5.6) into linear equations, we define another set of variables s which refer to $\|\nabla z(x, y)\|$. This requires us to solve the system of equations iteratively and update z and s for each iteration until we achieve convergence. The algorithm is detailed in Algorithm 1.

5.4 Experiments and Results

To validate our iterative algorithm, we first test it on a mathematical surface 'Saddle' defined by $z = x^3 - 3xy^2$. The images of the 'Saddle' are generated by rendering, assuming a Lambertian surface and illuminated by point light sources. The iso-gradient and iso-depth contours are obtained from [12, 13] as mentioned in Sections 5.2 and 5.3. To determine the iso-depth contours from [12], we require

---
Algorithm 1 Compute per-pixel depth from iso-gradient and iso-depth contours
---

1. Input : Set of iso-depth and iso-gradient contours

2. Initialize z at each pixel location as described in Section 5.3.1. Initialize s using the known depth z as defined in (5.4). Initialize $\nabla \mathbf{z}$ as defined in (5.2).

3. Initialize the magnitudes of parameters based on the constraints as follows :

$$s(x, y) = s(x', y') \tag{5.7}$$

$$\left. \begin{aligned} z(x, y) &= z(x^*, y^*) \\ z_x &= z(x, y) - z(x + 1, y) \\ z_y &= z(x, y) - z(x, y + 1) \end{aligned} \right\} \tag{5.8}$$

$\forall \quad x \leq u, y \leq v$; where u, v are the dimensions of the image.

4. Optimize the gradient: Optimize the system of equations (5.7) for one iteration and update the magnitudes of $s(x, y), z_x$ and z_y according to (5.2).

5. Optimize the depth : Optimize the system of equations (5.8) for one iteration and update the magnitudes of depth z. Compute the magnitude of s according to (5.4).

6. Repeat steps 4. and 5. iteratively until convergence.

---

a dense set light sources on a circle at different elevation angles. We use 162 light sources and render images turning on each light at a time. For estimating the iso-gradient contours, we capture two pairs of images from three point light sources located on a circle at an angular difference of $1°$. The iso-depth and iso-gradient contours intersect transversally in most of the regions to give a good initialization. We use MATLAB implementation of LSQR method [74] to solve the sparse linear system of equations.

The initial reconstruction is subject to noise and leaves surface patches without depth information. We observe from Figure 5.4(a)-(d) that our iterative algorithm achieves the final reconstruction free from these outliers. The effect is more pronounced in the case of noisy data as observed in Figure 5.4(e)-(h). We introduce a Gaussian noise of zero mean and standard deviation 0.01 and obtain the initial reconstruction. As in the previous case, we derive the final result after iterative
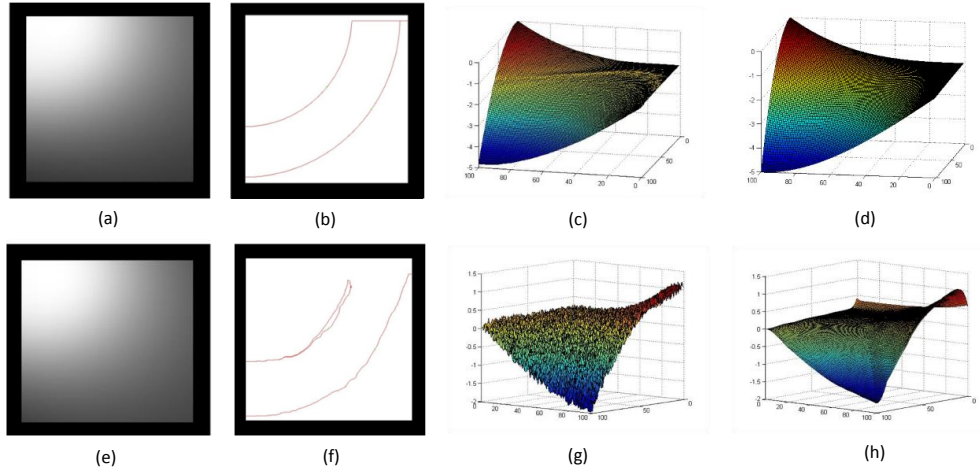
Figure 5.4: Result on 'Saddle' object. The first row (a) - (d) shows the result on an ideal 'Saddle' in the order of a sample image, sample iso-gradient contours, initialization and optimized results. The second row (e) - (h) shows result in the same order on 'Saddle' corrupted with noise.

optimization.

Next, we implement our algorithm on a slightly more complicated object 'Cup' as shown in Figure 5.5. We use a light configuration similar to the 'Saddle' to capture the images of the object. The 'Cup' is again a synthetic object and rendered assuming a Lambertian surface. The iso-depth and iso-gradient contours are determined using the previous algorithms as explained earlier. 'Cup' contains a large ellipsoidal area and an offshoot handle. The initialization is smooth in the ellipsoidal region. However, where there is a transition in the geometry from the ellipsoid to the handle, there are certain regions where the transversality condition is violated leading to empty regions. The optimized result demonstrates the effects of smoothing and filling in empty regions to obtain a complete reconstruction.

We test our algorithm on a real object 'Apple' as shown in Figure 5.6. The surface of 'Apple' is non-Lambertian as we can observe significant highlights in the sample image. However, the surface is still isotropic so we can implement the previous algorithms for deriving the iso-depth and the iso-gradient contours. Here, we use

73
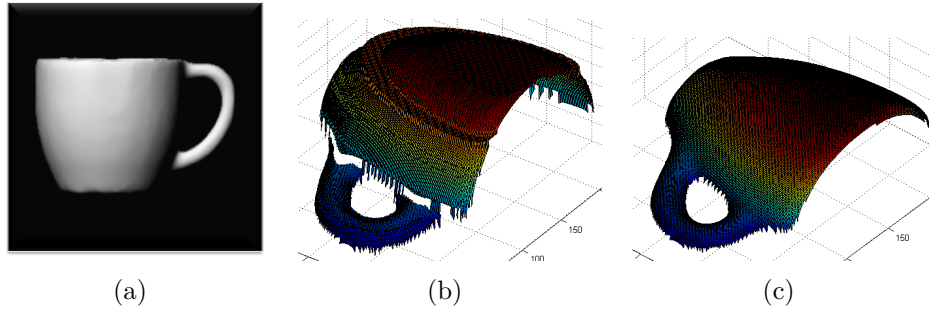
Figure 5.5: Result on 'Cup' object. (a) Sample image (b) Initialization and (c) Optimized result

a set of 72 LED light sources located over two concentric circles with the camera located at the centre of the circle, to capture images for estimating the iso-depth contours. We use a pair of LED lights separated by an angular interval of 3° to capture images for estimating the iso-gradient contour structure. Although the surface of the object is non-Lambertian, the iso-gradient and iso-depth contour structure helps us to estimate the initial reconstruction. This reconstruction is again corrupted by noise and empty patches. Our iterative algorithm determines the final reconstruction free from these outliers.
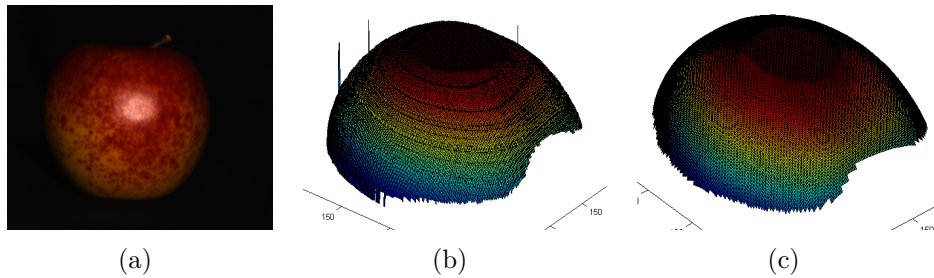
Figure 5.6: 'Apple' data result. (a) Sample image (b) Initialization (c) Optimized result

In all the above examples, we note that the only information used in the reconstruction are the images captured under the light configurations for estimating the iso-depth and iso-gradient structure. There was no additional information required in estimating the final structure. We note that the initialization requires a value for z_0. This result is independent of this value and we use $z_0 = 1$ in all

our experiments to determine the initial structure. The algorithm achieves convergence in 35, 55 and 80 iterations for 'Saddle', 'Cup' and 'Apple' respectively. The algorithm, experiments and the results have been discussed in [75].

Limitations

While our algorithm estimates the per pixel depth of the object of all isotropic materials, it still has the following limitations. Our algorithm requires that the iso-depth and iso-gradient contours should intersect transversally to estimate the initial reconstruction by integration. However, in cases where this condition is violated, which may happen when the object is not isotropic or corrupted by heavy specularities, we cannot use our algorithm which requires an initial depth and gradient estimates. We also note that our algorithm fails when the surface of the object has discontinuities or a complicated shape such that the isocontours cannot be propagated into all the regions of the object in the image.

5.5 Summary

This chapter presents an iterative algorithm which makes use of the iso-gradient and iso-depth contours to directly estimate the per pixel depth. Iso-gradient and iso-depth contours are estimated using previous methods which assume the surface reflectance of the object to be an isotopic BRDF. We do not require additional information from the user to obtain the reconstruction. We have validated our algorithm on synthetic and real data and obtained convergence. However, with the current constraints, our algorithm fails in the case where condition of transversal intersection of the contours is not satisfied.

Chapter 6

A New Perspective on Material Classification and application to Ink Identification

6.1   Introduction

In the previous chapters, we have discussed photometric stereo for automatic calibration and non-Lambertian surfaces. We note that the result of photometric stereo is a normal map of the scene, the images of which are captured under varying illumination. Besides estimation of the normal map of a scene, photometric stereo is also used extensively in appearance or the BRDF capture of the scene. The combined estimation of the normal map and the reflectance produces realistic 3D renderings of the scene. This has extensive applications in animation, cultural heritage and computer assisted surgeries [76, 42, 77]. Although BRDF is widely used for rendering purposes, because of its descriptive capacity of the surface material, it is also used in the area of material classification. Material classification
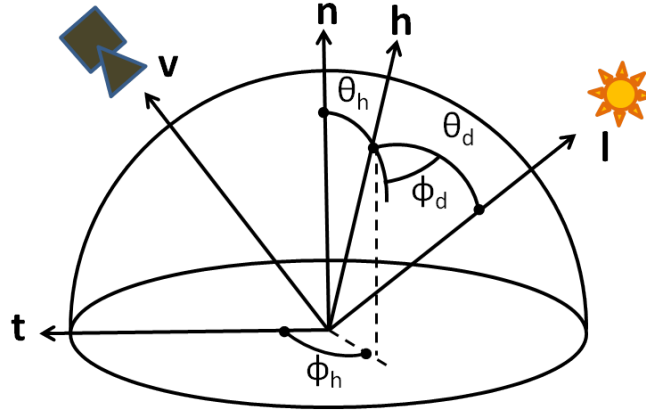
Figure 6.1: The half-vector parameterization of BRDF. The incident and reflected lighting directions **l** and **v** provide a 4D parameterization of a BRDF function. The BRDFs of many isotropic materials can be well approximated as a 2D function of the half-angle θ_h and the difference angle θ_d.

has significant utility in industry and research (in areas such as remote sensing, food inspection and recycling).

Different materials can be distinguished by examining their reflectance properties encoded in the BRDF [22]. Conventionally, BRDFs are represented as 4D functions - two dimensions each for the incident direction **l** and the reflected/viewing direction **v**. Capturing the full BRDF requires changing the incident lighting directions and the viewing directions spanning the entire upper hemisphere of the object at regular intervals. For the purpose of material classification, a subset of the full 4D BRDF was proven to be sufficient for classifying a significant variety of materials [22]. For example, factorization techniques [78, 79] represent an arbitrary BRDF as the sum of products of 2D functions for efficient storage and rendering. Under appropriate parameterization [14, 80], an isotropic BRDF can be well approximated as a 2D function of the 'half-angle' θ_h, and the 'difference angle' θ_d. The details of these angles are shown in Figure 6.1.

The BRDFs of most materials in the real world are near isotropic and are known to be well approximated by a 2D function or a slice of the full BRDF [14]. How-

ever, when the principal axis of the camera is coincident with the surface normal of the material sample, the captured BRDF slice is nearly 1D. In previous material classification works [22, 81, 23], the camera directly faces the sample material to provide a fronto-parallel image for better resolution. However, this causes a degeneration where the 'half-angle' θ_h is always equal to the 'difference angle' θ_d, because the surface normal $\mathbf{n}$ is coincident with the viewing direction $\mathbf{v}$. In other words, even if the sampled incident lighting directions $\mathbf{l}$ cover the entire upper hemisphere, effectively we only capture a 1D BRDF slice. We observe that such a configuration suffers from significant information loss. A considerable improvement in the classification performance can be achieved by a slight modification in the conventional camera and light source assembly. In this chapter, we explore this phenomenon more closely and propose configurations for enhanced material classification using BRDF capture.

Further, we observe that a 1D BRDF slice also possesses sufficient discriminative information although slightly less in comparison to the 2D or the 4D BRDF. This 1D BRDF slice captures important reflectance properties such as specular reflection and retro-reflectance. At the cost of a slight reduction in the performance, a very simple and low cost setup can be arranged for performing material classification. We propose a handheld flashlight-camera arrangement to capture a 1D BRDF slice for material classification. This setting allows us to obtain a 1D BRDF slice where the 'difference angle' θ_d is fixed at 0, but θ_h can change from 0 to $\pi/2$. This slice can capture distinctive reflectance properties such as specular reflection (when θ_h is small) and retro-reflectance (when θ_h is large). It achieves similar classification accuracy as the conventional setting [22, 81, 23] with much lesser input images and more flexible data capture. We apply this proposition for the task of ink classification, which has important applications in forensics and analysis of historical manuscripts [82, 83].

79

Spectral analysis based techniques such as spectroscopy and infrared reflectography [84, 85] are extensively used to classify different inks according to their absorptivity of light at different wavelengths. These methods usually require special devices by which the optical properties of the pigments are examined under illumination beyond the visible spectrum. We observe promising results for ink classification using reflectance up to an accuracy of 79%. To the best of our knowledge, no prior work has used reflectance for ink classification.

The main contributions of the chapter are 1) a new perspective to material classification observed as a result of a slight modification in the camera, light source arrangement for BRDF capture, 2) introducing reflectance for ink classification, and 3) a hand held camera-flashlight arrangement for ink classification.

6.2 Overview

Following this introduction, we review related work in Section 6.3. In Section 6.4, we briefly discuss the reduction in BRDF dimensionality for material classification. We bring out the limitations in the existing approaches and propose a modification in the image capture setup for improved performance. Our proposition is validated on an ink database for material classification. In Section 6.5, we establish that a 1D BRDF slice can be used for material classification. We also propose a simple handheld device for 1D BRDF slice capture. In Section 6.6, we apply the proposed approach to the problem of ink classification and design a working system. We validate our system in Section 6.7.

6.3 Related Work

In this section, we review some of the previous works related to our problem under the following categories.

Reflectance-based Material Classification

Classifying materials based on reflectance is a relatively less explored area. Wang et al. [22] propose a method of material classification using '2D BRDF slices'. They capture a '2D BRDF slice' by acquiring images with a fixed camera and varying illumination. They then fit hemispherical harmonics to the observations and use the coefficients as a basis for classification. Jehle et al. [81] and Gu and Liu[23] further studied the optimal illumination condition to enhance the separation. We study the sample distribution in the BRDF domain and prove that a slanted camera can significantly improve classification accuracy. We further develop a method to capture a 1D BRDF slice with a handheld flashlight camera.

Ink Classification

Documents and manuscripts are analyzed for their authenticity and dating by determining the type of ink material. The existing methods for such an analysis can be broadly characterized into destructive and non-destructive types, the latter being preferred. In this chapter, we focus on the non-destructive types. Most non-destructive methods such as spectroscopy and reflectography [84, 85] use illumination beyond the visible spectrum. Alternatively, image processing based methods offer cost effective and non-destructive solutions by computing models and analyzing ink properties in the visible domain. Chakravarthy et al. [86] showed that statistical properties such as saturation histograms in HSV color space can

differentiate liquid inks and viscous inks. This color analysis based method cannot distinguish inks of similar color. Kokla et al. [87] propose a method to distinguish inks by studying statistical properties of the ink intensities in visible and infrared light. However, the accuracy reported is not very high. Licata et al. [82] propose a near IR based method for distinguishing and restoring inks in historical manuscripts. Hu et al. [88] provide a survey of ink segmentation methods used in forensics. None of the above methods in ink segmentation studied the reflectance properties of inks for segmentation. A recent work by Berger [89] describes an ink segmentation method based on color deconvolution. This method can distinguish inks of similar color based on intrinsic differences in the ink shades. However, it cannot deal with common black inks with almost constant shade. Our method is based on the reflectance properties of manuscript inks, which can identify different inks of similar color. Our method outperforms Berger's method[89].

6.4   BRDFs for Material Classification

6.4.1   Dimensionality of BRDFs

BRDFs of real materials exhibit strong symmetry and redundancy such as isotropy, reciprocity, half-vector symmetry, etc. Thus, the dimension of the BRDF domain is often reduced without a significant loss of information. As shown in Figure 6.1, a BRDF can be parameterized as a 4D function $f(\theta_h, \phi_h, \theta_d, \phi_d)$. There is strong evidence that many BRDFs can be well approximated by a 2D function $f(\theta_h, \theta_d)$. Specifically, the isotropy reduces the BRDF domain to $(\theta_h, \theta_d, \phi_d)$. The half-vector symmetry[37], which suggests that the BRDF is unchanged by a rotation of $\mathbf{l}$ and $\mathbf{v}$ around the bisector $\mathbf{h}$, further reduces the BRDF domain to (θ_h, θ_d). This 2D representation is empirically verified in [90] with the MERL BRDF database [91]
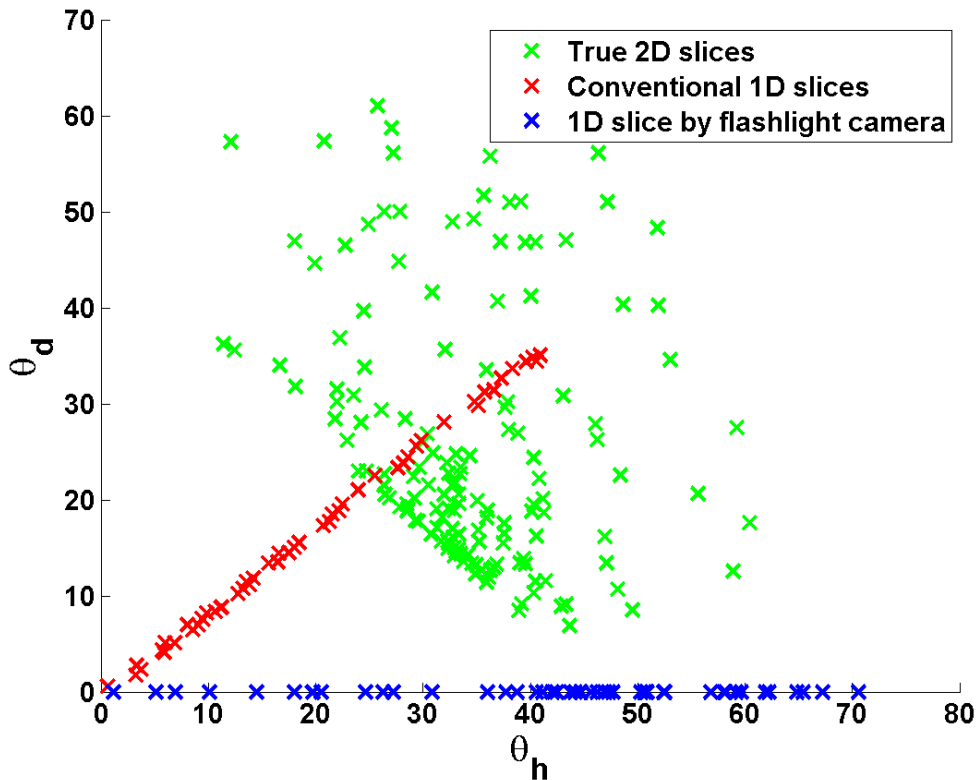
Figure 6.2: The sample distribution in the θ_h- θ_d space. The conventional setting (i.e. $\mathbf{v} \approx \mathbf{n}$) collects red samples along a straight line. By simply setting the camera at a slanted angle, we collect the green samples spanning a larger range of the BRDF domain. The blue samples are collected by a handheld flashlight camera.

and recently applied for photometric stereo [92].

6.4.2 Limitations of Conventional Approaches for 2D BRDF Capture

We now analyse the material classification problem with this 2D BRDF approximation. In previous methods [22, 23], the camera is usually oriented in a way such that $\mathbf{v} \approx \mathbf{n}$. In this case, the angles $\theta_h = \arccos(\mathbf{h}^\top \mathbf{n})$ are approximately equal to $\theta_d = \arccos(\mathbf{h}^\top \mathbf{v})$. Effectively, only a 1D slice of the BRDF can be captured no matter how many lighting samples are collected from the upper hemisphere. Fur-

thermore, the range of θ_d and θ_h is restricted to $[0, \pi/4]$. This problem is clearly demonstrated in Figure 6.2. We move an LED light on the upper hemisphere of a planar sample. Each red cross indicates an observation captured by a camera facing directly to the sample. All these sample points are along a line segment in the θ_h-θ_d space. This set of samples does not capture some important phenomena such as retro-reflectance[1] and Fresnel effects[2] that make certain materials distinctive. Typically, the retro-reflectance is stronger when $\theta_d = 0$ and $\theta_h > \pi/3$. The Fresnel effect is most significant when $\theta_d > \pi/3$.

6.4.3 Our Approach

To achieve improved ranges of θ_d and θ_h, we can simply orient the camera along a slightly slanted direction. Under this configuration of lights and camera, we can collect the samples from a larger BRDF domain as shown in green in Figure 6.2. Naturally, they encode more reflectance properties and will enhance the performance of material classification.

To verify our analysis, we collect 55 different inks of various types (water soluble, gel, ball point, permanent mark) and from different manufacturers (Pilot, Zebra, Uni, Staedtler, Pentel, Faber-Castell, Reynolds, Parker, etc.). We capture multiple images of a flat document containing different ink strokes under varying illumination directions. We collect two sets of images, one set for a slanted camera (i.e. a true 2D BRDF slice), one set for the camera facing the document head-on (i.e. a near 1D BRDF slice because of $\mathbf{v} \approx \mathbf{n}$). The data capture setup and some sample images are provided in Figure 6.3(a),(b). All the lighting intensities and directions are recorded with calibration objects. The incident lighting direction is densely sampled over the upper hemisphere defined by the document surface

---

[1]Retro-reflectance is the phenomenon that light is reflected back towards the incident direction.

[2]Fresnel effects explain the increased specular reflection at the grazing angle.
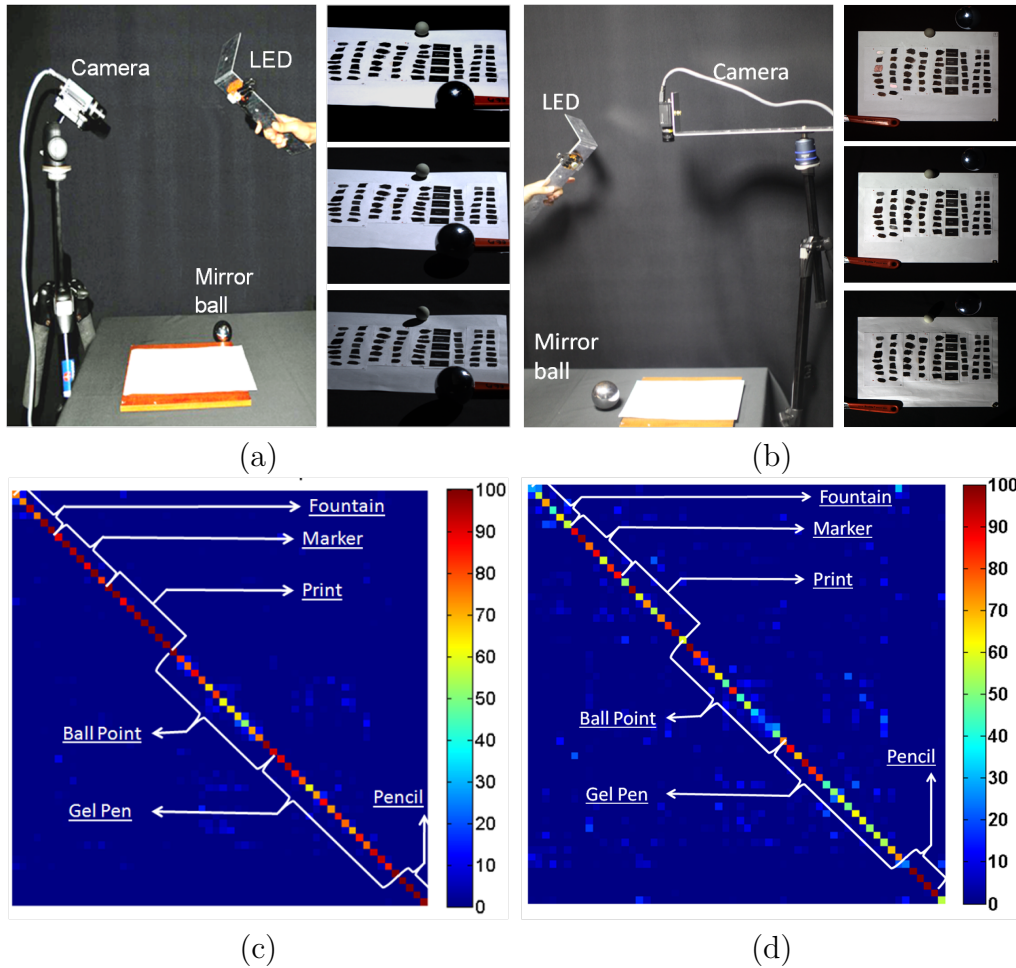
Figure 6.3: Experiment on ink classification based on true 2D BRDF slices and near 1D BRDF slices. a) Acquisition setup for true 2D BRDF data and captured images, b)Acquisition setup for 1D BRDF slices and captured images , c) Confusion matrix for ink classification with true 2D BRDF slices. It achieves average accuracy of 85% over 55 inks. d) Confusion matrix result with near 1D BRDF slices. The average accuracy is 69%.

normal. From the recorded image radiance I and calibrated lighting information, we obtain a sample of the 2D BRDF from each pixel of each image according to the following equation,

$$f\left(\theta_d, \theta_h\right) = I(\mathbf{x})/\left(\mathbf{n} \cdot \mathbf{l}\right). \tag{6.1}$$

We then take this recovered BRDF as a distinguishing feature for material classification. In our images, each ink often covers 4000 pixels. We generate a BRDF profile for each pixel and select 1000 pixels (or profiles) to train a standard support vector machine (SVM) classifier [93]. The classifier is then applied to the remaining pixels for evaluation. The classification results (in terms of a confusion matrix) with the true 2D BRDF slice and the near 1D BRDF slice are shown in Figures 6.3 (c) and (d) respectively. Almost all the inks are successfully classified when the true 2D BRDF slices are used. The average classification accuracy is 85% for the 55 inks. This high accuracy clearly demonstrates that BRDFs provide a strong cue to identify different inks. However, when the near 1D BRDF slices are used, the performance is significantly dropped to 69%. This demonstrates the importance of selecting the right BRDF domain sampling strategy. When the samples cover a larger range, more reflectance properties can be captured and the classification performance will be better.

6.5   1D BRDF Slice for Material Classification

6.5.1   A Handheld Flashlight Camera Arrangement

To allow flexible data capture, we propose a handheld flashlight camera to capture BRDF slices for material classification. When images are captured with a light attached to the camera as shown in Figure 6.4, the incident and outgoing lighting
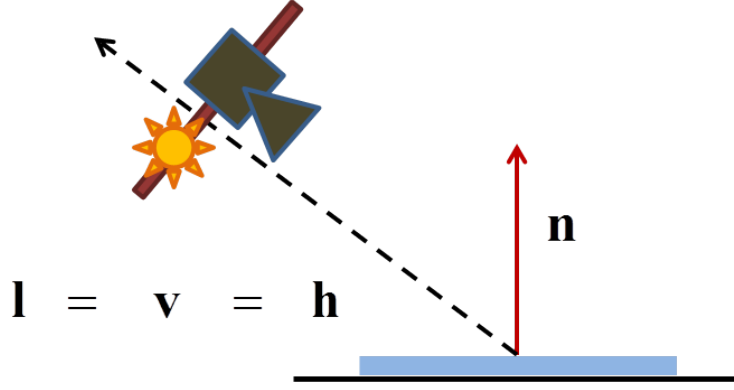
Figure 6.4: Schematic diagram showing the handheld flashlight camera arrangement.

directions are coincident, i.e. $\mathbf{l} = \mathbf{v}$. Thus, the half-vector $\mathbf{h}$ overlaps with $\mathbf{v}$, and θ_d is always 0. By moving the camera around, we can capture images with θ_h varying over the range of $[0, \pi/2]$. This simple setting allows us to capture some important reflectance properties.

Firstly, specular reflectance highly depends on θ_h [14]. Since our method covers the full dynamic range of θ_h, it faithfully captures the characteristics of specular reflections, e.g. the strength and extent of specular lobes. Secondly, retro-reflectance often becomes stronger when $\theta_d = 0$ and $\theta_h > \pi/3$. Therefore, our method can also capture retro-reflectance to help distinguish different materials. Lastly, by fixing θ_d, we cannot capture any Fresnel effects.

A possible solution is to attach multiple lights to the camera, each with a different distance to the camera. By turning these lights on and off during image recording, we can capture Fresnel effects. However, this setting is more complicated and we found empirically our flashlight camera setting works well for ink identification.

To verify the performance of this handheld system, we apply it to the flat document sample with the 55 inks. Figure 6.5 (a) shows the confusion matrix based on this handheld device. The average classification accuracy is 71%. Please notice this is slightly higher than the accuracy (69%) of the conventional setting. Though

87
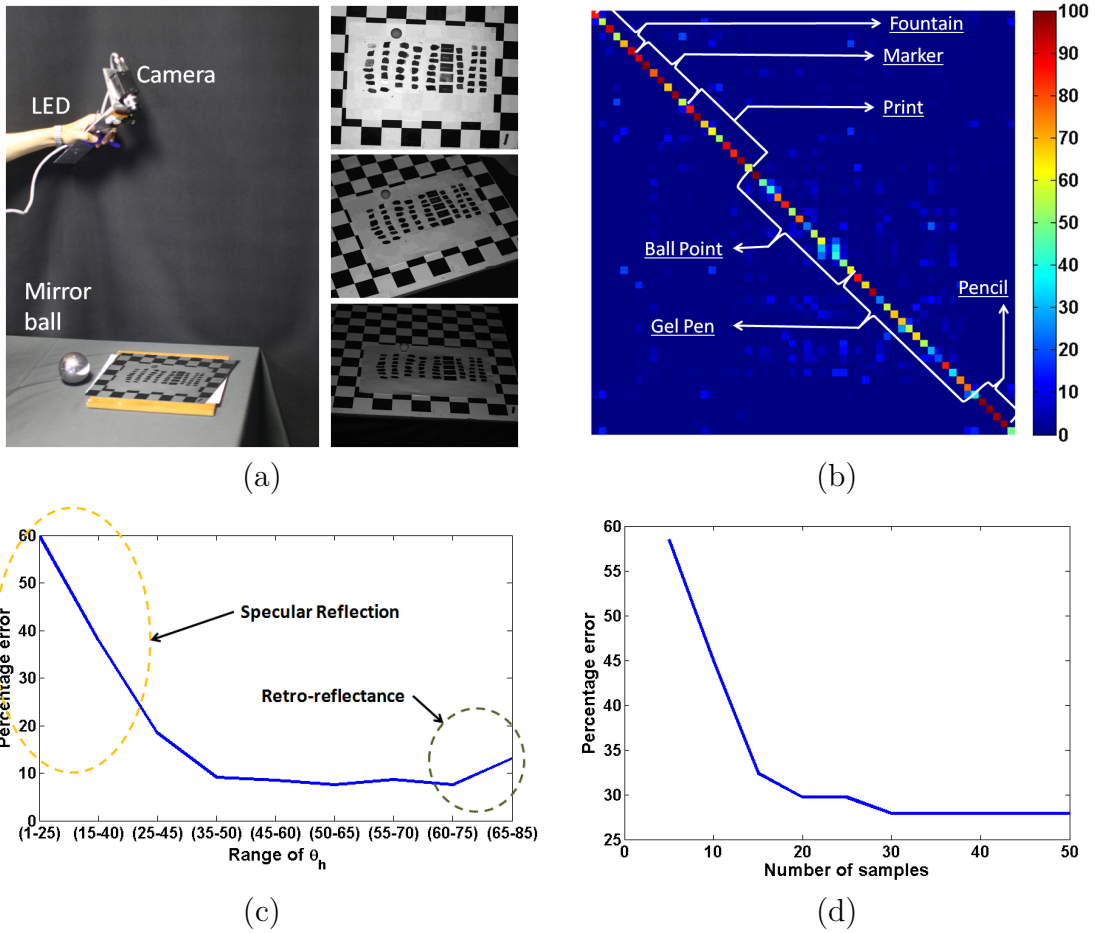
Figure 6.5: (a) Acquisition setup for 1D BRDF slices and captured images, (b) Confusion table showing the separability of the 55 inks using the proposed hand-held flashlight camera setting, c) Sensitivity of θ_h, d) Analyzing the number of samples vs. classification performance.

the classification accuracy is reduced compared to the true 2D BRDF slices, this performance is still good for classification of 55 inks. The reduced accuracy is partly because of the imprecision in registration when capturing with a handheld device. We can observe that performance for Marker, Prints, Pencil and Gel pens have quite high accuracy, while the performance for Ball Point pens and fountain inks have quite low accuracy.

6.5.2 Distinctive Intervals

To further understand which interval of θ_h is more informative for material classification, we cut the 1D BRDF slices into multiple overlapping segments. We perform similar SVM based classification based on each of these segments. The result is summarized in Figure 6.5(b). We observe that: 1) The classification accuracy is highest in the range of $(1° - 25°)$ and gradually falls after that and slightly rises in the range of $(65° - 85°)$. These two ranges correspond to the specular and retro-reflectance phenomena respectively. It confirms our design choice of using a flashlight camera. 2) classification accuracy using the full 1D BRDF slice, 71%, is greater than using any individual range, $< 60\%$. This suggests that a full sampling of θ_h in the range $[0, \pi/2]$ will be useful.

6.5.3 Optimal Number of Images

We further investigate as to what is the appropriate number of input images for this method. We show its performance against the number of randomly sampled images in Figure 6.5(c). Here, samples are randomly selected over the range of θ_h. The overall classification error curve takes a sharp turn at samples 10 to 20 and gradually falls after that. The samples here are uniformly sampled in the range of θ_h ($[0, \pi/2]$). The classification error remains almost constant after the number of samples reaches 30 as observed in Figure 6.5(d). Therefore, samples sampled in 30 bins in the interval of θ_h is sufficient for a reasonable performance of the classifier. We select around 30 samples for later experiments in Section 6.7. For a comparison, we provide a similar error curve for the conventional data capture setting in Figure 6.6. It will take over 100 images to reach similar classification accuracy.
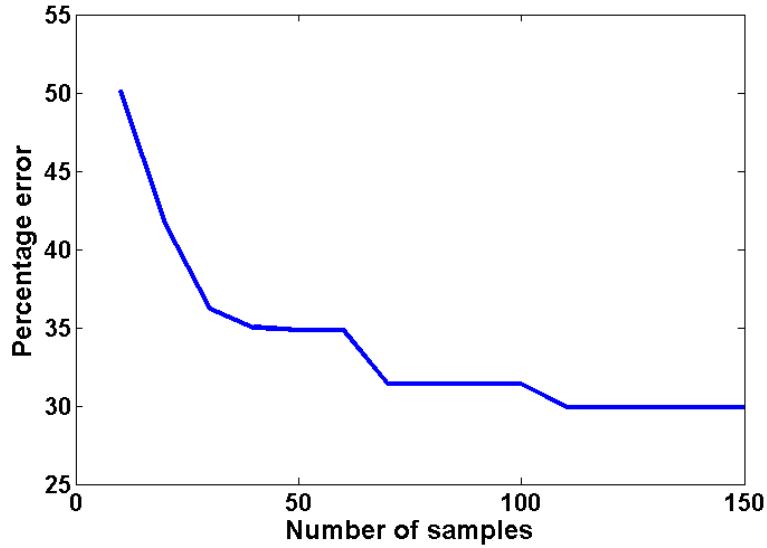
Figure 6.6: Classification error rate of the conventional data capture setting where the camera faces directly to the sample.

6.6 Application to Ink Classification

In the previous sections, we have discussed the constraints for the performance of 1D BRDF based material classification. In this section, we implement the classification method to differentiate inks. As stated previously, ink classification is an important area in forensics for analysing questioned documents. We first introduce the major types of ink available in the market in the following Section.

6.6.1 Ink types

Ink is a material which contains pigments or dyes to impart specific colour to an impacting substrate. Depending on the type of usage, the composition of ink varies. Some of the major types of inks [3] found in the market are categorised below depending on their chemical composition [83]:

---

[3]Although we do not include pencil in the types of inks, in our experiments, we also consider pencil as one of the inks because its appearance in the camera is similar to ink.

1. **Carbon inks** : These inks are a suspension of ground amorphous carbon cake in a water-glue medium. The colour of the ink can be improved by adding pigments to the dye. These inks are insoluble in water, very stable and are not decomposed by factors such as light, air, heat moisture etc. These inks have been used for over 2000 years and are found in historical manuscripts.

2. **Fountain pen inks** : In most cases, these inks are an aqueous solution of synthetic dyes. These inks are bright and attractive in colour but not as stable as carbon ink.

3. **Ballpoint inks** : These consist of synthetic dyes in various glycol solvents or benzyl alcohol. Several other ingredients are added to the ink to impart specific characteristics such as lubrication, binding, viscosity etc.

4. **Rolling ball marker inks** : These inks are water based and contain organic liquids such as glycol to retard the drying of the ball point.

5. **Fiber tip pen inks** : These are usually water or xylene based containing dyes and additives as in ball point and fountain pen inks.

6. **Gel-pen inks** : These inks contain completely insoluble pigments rather than organic dyes. This ink is a gel and not a liquid and insoluble both in water and organic solvents.

7. **Printing inks** : Different types of printers and print jobs require different types of ink. As a result, there are several types of printer ink available [94]. For example ink-jet printers use liquid inks in cartridges, laser printers need powdery toners and some other printers use solid inks. In general, printing inks are made up of four basic components - pigments, resins, solvents and additives. Depending on the application, these components are used in different proportion.

91

Almost all the above inks except carbon ink are widely available.

6.6.2   An Ink Classification System for Curved Documents

In this section, we extend the use of handheld system to a general curved document. The 3D document shape is reconstructed to facilitate image registration and 1D BRDF slice estimation. We then apply a clustering method to group pixels with similar BRDF slices together for ink segmentation. Note that we do not seek to identify the type of ink but only estimate if two pixels have the same ink or not, which is useful for forensics.

3D Reconstruction: The surfaces of documents are often curved. We need a precise 3D reconstruction for image registration so that we can obtain multiple observations for a surface point on the document. We apply the structure-from-motion package VisualSFM [95] to perform 3D reconstruction of the input images. This 3D reconstruction provides the camera locations and a set of sparse 3D points on the document surface. Since the light source is attached to the camera, we obtain the location of the light source directly from this 3D reconstruction.

Surface Fitting: The reconstructed 3D point cloud is often sparse. We fit a parametric surface to the sparse point cloud to recover the 3D shape of the surface. Following the work of Yamashita et al. [96], we make use of Non Uniform Rational B-Spline (NURBS) to fit a smooth parametric surface. This 3D surface also provides an accurate estimation of the normal direction **n** at every point.

Image Registration: Given the 3D surface, we can determine the 3D location of each pixel on to intersecting a ray from this pixel on the reference image with the 3D surface. By reprojecting this 3D point back on to the other images, we can obtain pixel correspondence among images captured from different viewpoints. We take the view where the document is most fronto-parallel to the camera as the

reference view. We then project the 3D position of each of its pixels to all the other images to collect multiple observations of each pixel.

BRDF Slice Estimation: We have collected multiple observations of each pixel as observed in the reference view. Further, we know its surface normal direction from the NURBS surface fit, and the lighting directions from the camera positions. Thus, we can recover a 1D BRDF slice for each pixel from (6.1).

Ink Identification: The estimated BRDF profiles will be used as a discriminative feature for ink classification. We compare the BRDF slices of different points. Noting that θ_h varies in the range 0 to $\pi/2$, we sample this interval to 30 bins, and calculate the average BRDF value in each bin for every pixel. The BRDF slices form a discriminative feature for each ink. We can differentiate the inks using clustering techniques by grouping the pixels according to their associated BRDF slices. This will allow us to tell if the inks at some given pixels on a document are different or not. In our experiments, we apply the affinity propagation method [97] for clustering. Affinity propagation is one of the several methods available for clustering. Any of the available methods can be chosen for performing unsupervised classification of inks.

This is summarised in the Figure 6.7.

6.7 Experiments

6.7.1 Ink Classification for Flat Documents

In Figure 6.8, we examine the SVM based ink classification on a flat document with some representative inks. As shown in Figure 6.8(a), a flat document contains horizontal strokes and text written with 12 different inks of black color. The

93
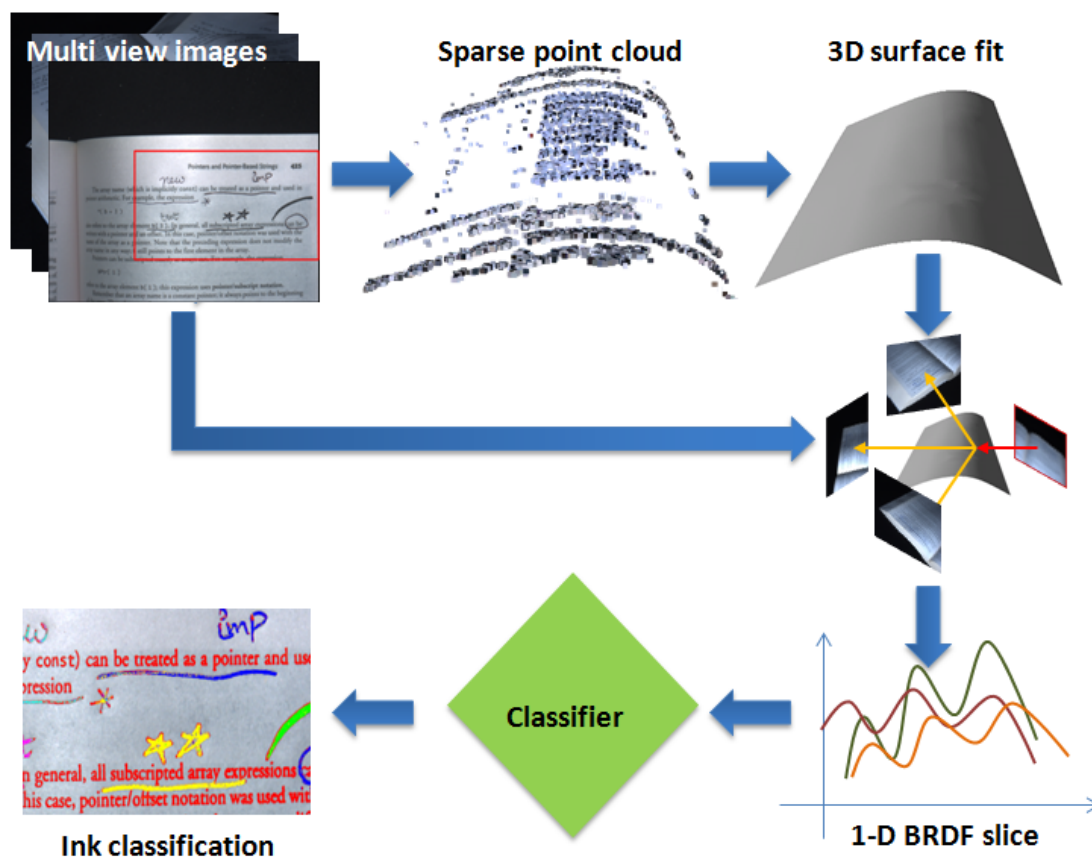
Figure 6.7: Flowchart showing the ink classification using BRDF slices.

horizontal stroke and the text written beside belong to the same ink type. Pixels on the horizontal strokes are used for training and pixels on the texts are used for testing. The 12 inks are selected from the five major types of inks: pencil (Pencil1 (Staedtler 3B), Pencil2 (Staedtler HB)), fountain pen inks (Pilot, Stabilo),ballpoint inks (Ball OM, Ball JS, Zebra), mark pen ink(Sharpie, Perm Ptl, Per Zig), and gel-pen inks(EngerGel, Jimnie).

The classification results are shown as an image in Figure 6.8(b). We also show the zoomed-in classification results of the inks of the best and worst performance. In Figure 6.8(c), we evaluate the separability on these 12 inks in the form of a confusion table. We observe that the carbon ink (Pencil1, Pencil2) and the markers (Perm Ptl, Sharpie) have higher classification accuracy above other inks. The gel pen inks and ballpoint inks have less accuracy but these inks have more confusion among themselves for classification. This observation is consistent the result on 55 inks presented in Figure 6.3.

We must also note that all the inks are black in color and there are almost negligible variations in their shades. Bearing this fact, the classification result demonstrates that BRDF profiles are promising discriminative features for identifying different inks.

The average BRDF profile of each ink is shown in the Figure 6.9. We observe from Figure 6.9 that most of the profiles exhibit distinctive behavior though the inks may have similar color. For example, the BRDF profiles of Pencil1 and Pencil2 exhibit a strong specular reflection that happens at about $\theta_h = \pi/4$, but not at $\theta_h = 0$. This indicates that the lead particles have changed the microfacet normal distribution of the document surface. These variations in the BRDF profiles account for the separability of the inks.
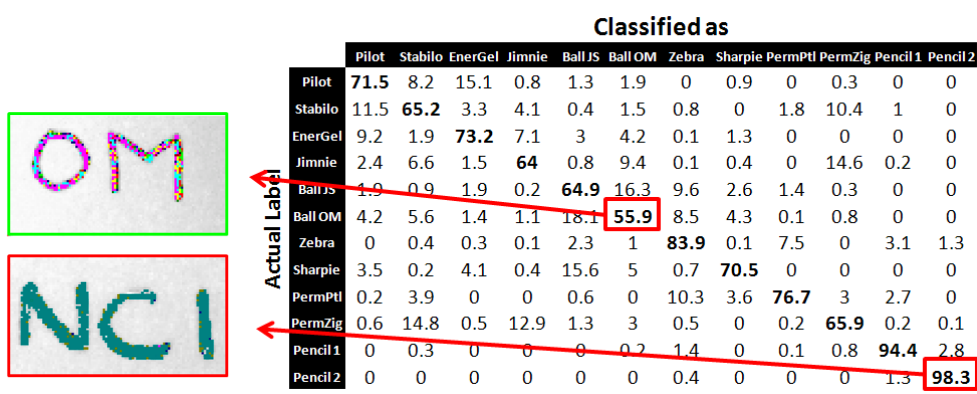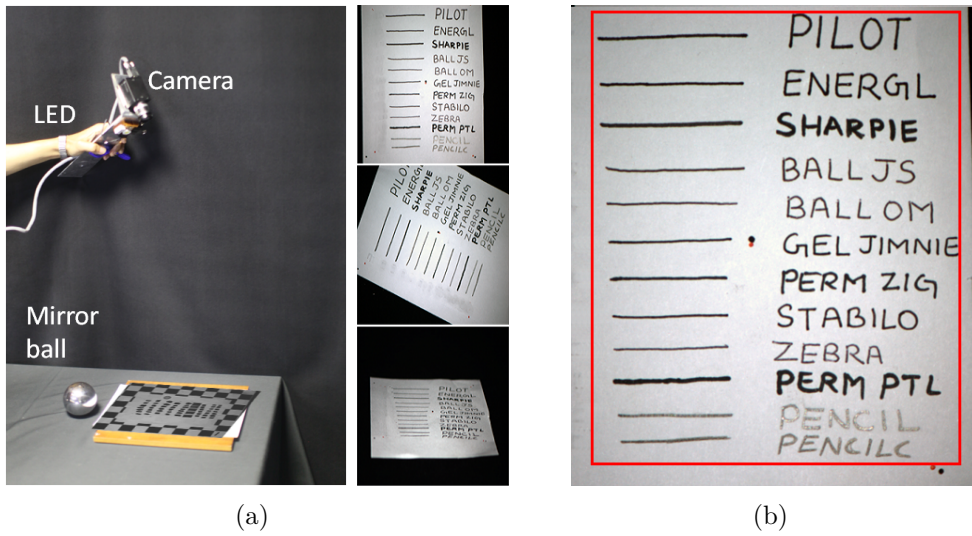
Figure 6.8: (a)Acquisition setup for 1D BRDF slice and the sample images, (b) Selected portion of ink strokes, (c) Classification of ink strokes by an SVM classifier, with zoomed in results in (d), (e) Confusion matrix for classification performance.

Classified as

Actual Label	Pilot	Stabilo	EnerGel	Jimnie	Ball JS	Ball OM	Zebra	Sharpie	PermPtl	PermZig	Pencil 1	Pencil 2
Pilot	71.5	8.2	15.1	0.8	1.3	1.9	0	0.9	0	0.3	0	0
Stabilo	11.5	65.2	3.3	4.1	0.4	1.5	0.8	0	1.8	10.4	1	0
EnerGel	9.2	1.9	73.2	7.1	3	4.2	0.1	1.3	0	0	0	0
Jimnie	2.4	6.6	1.5	64	0.8	9.4	0.1	0.4	0	14.6	0.2	0
Ball JS	1.9	0.9	1.9	0.2	64.9	16.3	9.6	2.6	1.4	0.3	0	0
Ball OM	4.2	5.6	1.4	1.1	18.1	55.9	8.5	4.3	0.1	0.8	0	0
Zebra	0	0.4	0.3	0.1	2.3	1	83.9	0.1	7.5	0	3.1	1.3
Sharpie	3.5	0.2	4.1	0.4	15.6	5	0.7	70.5	0	0	0	0
PermPtl	0.2	3.9	0	0	0.6	0	10.3	3.6	76.7	3	2.7	0
PermZig	0.6	14.8	0.5	12.9	1.3	3	0.5	0	0.2	65.9	0.2	0.1
Pencil 1	0	0.3	0	0	0	0.2	1.4	0	0.1	0.8	94.4	2.8
Pencil 2	0	0	0	0	0	0	0.4	0	0	0	1.3	98.3

Figure 6.9: 1D BRDF slices of the 12 ink examples.

(a)   (b)   (c)
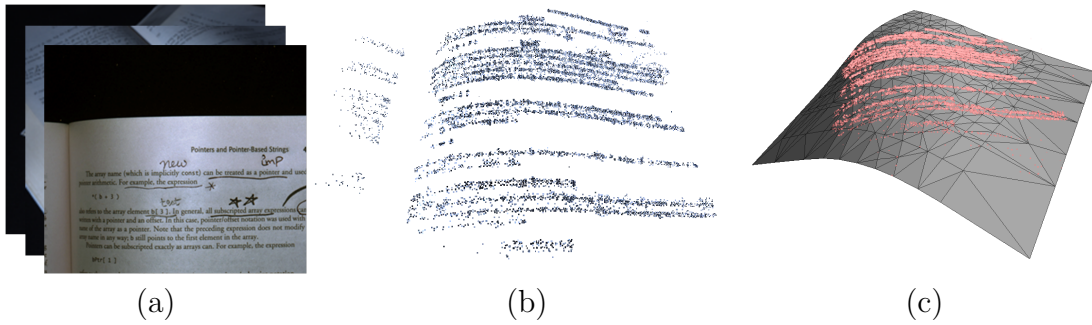
Figure 6.10: 3D reconstruction of the document surface. a) Multiple images of the document captured using the handheld device; b) Reconstructed sparse point cloud; c) NURBS surface fit through thepoint cloud.

6.7.2 Practical Ink Classification

We further evaluate the problem of ink classification for general curved documents. Figure 6.10 shows the recovered point cloud and the 3D surface of the document from multiple view images. For the affinity propagation method, we assume that the number of different inks is known to limit the number of clusters formed. We use the minimum bound on the preference value for determining the required number of clusters. In our experiments, we use the negative Euclidean distance measure to determine the similarity matrix. Figure 6.11 shows our segmentation

97

results, where (a),(b) and (c) are the input image, segmentation result, and ground truth segmentation.
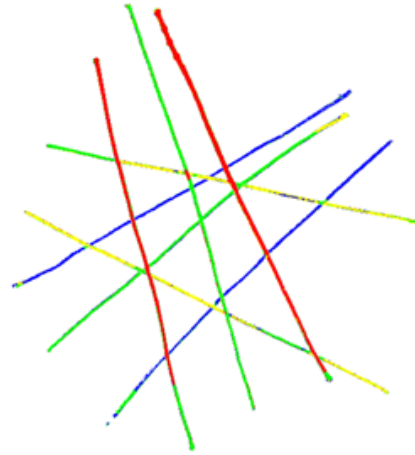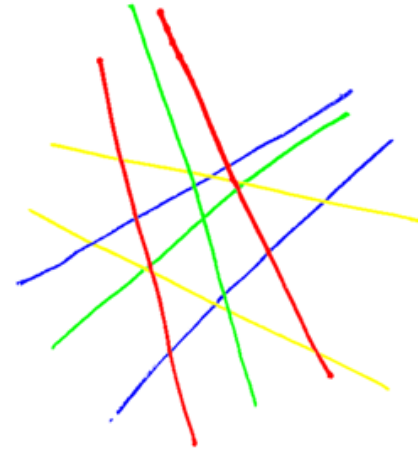
(a)   (b)   (c)

Figure 6.11: Segmentation of ink strokes. a) Sample image. b) Segmentation result. c) Manually marked ground truth segmentation.

As can be observed, the proposed method can correctly group pixels of the same ink in most regions. For the example in the top row, we have a curved document containing different inks (Fountain, Zebra, Perm Zig) of black color. The three inks are successfully identified by the proposed method. For the second example, we have a printed book page with 5 different inks (Sharpie, Omni, Pilot,Fountain, EnerGel). We observe the confusion of EnerGel, a gel ink, with Fountain ink. Liquid inks tend to get absorbed in the paper, which leads to the confusion in identification. For the example of cross lines, we have lines with different inks(Fountain, Ball OM, Zebra, Stabilo) intersecting each other. The strokes appear very similar to each other in their color. Our method can differentiate Fountain, Stabilo and Zebra easily, while is confused with Ball OM and Zebra, which are ball point pens, in some regions. As we have observed previously in our experiments on the ink database of 55 inks, ball point pens tend to have confusion within the same class while they are distinct from other classes. In the last row of Figure 6.11, we show an interesting example of a forged cheque. Using the proposed method, we are able to identify different inks and the attempted forgery. However, we observe misclassifications in the identified inks especially near the signature on the sloping portion of the check. We find that the slight intensity fall off leads to the confusion.

The number of pixels for each ink depends on the length and width of the stroke. Most of the ink strokes have around 1000 pixels in each image. As we have already discussed the optimal number of images in Section 6.5, we capture sufficient number of images such that the entire range of θ_h can be sampled in 30 bins. We capture around 70 images covering the maximum possible range over the upper hemisphere of the document. We must also note that a portion of the ink stroke along the boundary blends into the paper. A wider stroke will have more ink pixels and will result in better performance. However, a high resolution camera that can capture sufficient details of the ink stroke (in approximately 15-20 pixels) can

101

enhance the performance of ink classification.

6.7.3 Comparison

Previous computer vision algorithms for ink classification have basically relied on the color information. As discussed in Section 6.3, Chakravarthy and Dasari [86] use statistical properties of the intensity distribution in the RGB channels of the inks for classification; Berger [89] uses difference in color shades of the inks for differentiating them. However, in fraudulent documents, the inks are almost same in color in most of the cases. In that case, color information is not sufficient to distinguish inks. Reflectance describes the intrinsic characteristic of the material which is not solely dependent on color. In Figure 6.12, we compare our method with Berger's work [89] which separates inks based on color analysis. Here, strokes of the same orientation have the same ink. In Figure 6.12(a), the inks have slightly different colors, so that both methods work well. However, for the examples in (b-d) where the inks are black with almost no difference in their shades, Berger's method fails while our method can correctly distinguish them.

It is also interesting to note that the ink strokes overlap with each other in Figure 6.12. The ink in the overlapping regions is detected either as one of the two inks or a third ink as observed in the rightmost column. We note that in case of viscous inks (such as ballpoint pen ink) the two inks do not mix in the overlapping region, in most cases. Therefore, the ink on the top layer is identified in this case as observed in Figure 6.12(a). However, in Figure 6.12(b),(c), the inks are liquid and mix with each other in the overlapping regions and therefore detected as a third ink different from the other two inks. The inks in Figure 6.12(d) belong to marker inks and these inks also are identified in the same way as the viscous inks. The algorithm, experiments and the results are also discussed in [98].
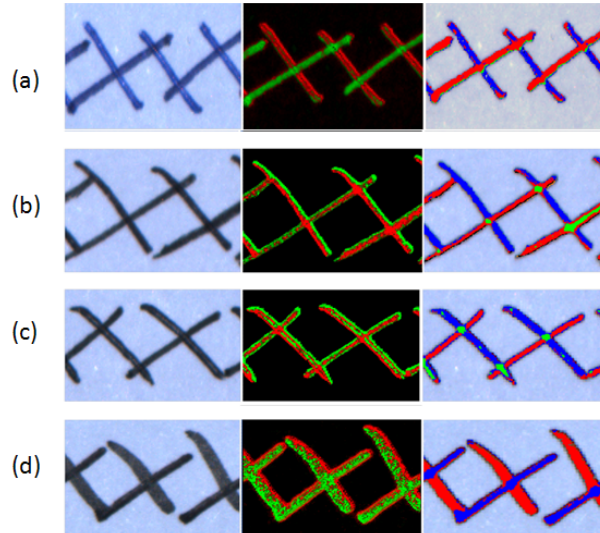
Figure 6.12: Comparison with Berger's work [89]. Left: input image. Middle: Berger's result. Right: Our result.

6.7.4 Limitations

In this chapter, our main goal was to explore and validate the idea of reflectance based material classification to the problem of ink classification and we have achieved the goal. We have demonstrated a working system for ink classification. However, our method works on those inks whose discriminative properties can be captured in a 1D BRDF slice. We observe that the class of ball point inks do not respond favourably to our method. Nevertheless, we believe that by analysing the properties of these inks, we can determine the sensitive portions of the 2D BRDF space to which these inks can respond. Another limitation is that our method is applicable to curved documents so that a sparse point cloud from structure from motion is sufficient to generate the 3D surface of the document. Our algorithm will not work properly when the document is not smooth.

6.8   Summary

In this chapter, we have analysed the issue of sampling in the BRDF domain for reflectance based material classification methods. Our analysis has shown that the performance of the previous methods degrades because they only capture a near 1D BRDF slice. A considerable improvement in classification accuracy can be achieved by making a slight change in the arrangement of the camera and lighting arrangement. Further, we have proposed a flexible method for material classification with a handheld flashlight camera. This arrangement captures a 1D BRDF slice that encodes distinctive reflectance features such as specular reflection and retro-reflectance. We empirically found that our 1D BRDF slice based classification performs comparably with respect to previous methods and requires much less images and allows flexible data capture.

In terms of ink identification, we investigated the BRDFs of inks belonging to various types available in the market and showed that these inks can be successfully identified using BRDF profiles. We have demonstrated the application of our algorithm on flat documents, general curved document, and a realistic case of a fraudulent cheque.

Chapter 7

Conclusion

In this thesis, we have explored three different aspects of photometric stereo :
(1) shape estimation from Lambertian photometric stereo, (2) shape estimation
from non-Lambertian photometric stereo, (3) reflectance capture from photometric
stereo and its application to ink classification.

In Chapter 3, we have presented a method to resolve the ambiguity parameters in
the form of a closed for solution. The light sources were constrained to uniformly lie
on a ring and images were captured using a camera located at the centre of the ring,
oriented orthogonal to the plane containing the ring. These constraints helped to
solve the Lambertian photometric stereo problem without explicitly computing
the light source positions. In Chapter 4, we have presented a preliminary analysis
for autocalibrating photometric stereo using rectangularly placed light sources.
In Chapter 5, we have presented an algorithm which combines the information
from iso-gradient and iso-depth contours to recover a per-pixel depth estimate. In
Chapter 6, we have shown that a camera fixed with its viewing direction orthogonal
to the material sample fails to capture the entire 2D BRDF. When the camera
is slanted at different angles with respect to material sample, the measurements

are sampled from a larger portion of the 2D BRDF space. We have also proposed a simple handheld camera flashlight arrangement for 1D BRDF slice capture. Although the 1D BRDF spans a very small portion of the 2D BRDF, it captures certain important information which can differentiate different materials.

In Chapter 6, we have applied our algorithm to the problem of ink classification. Ink classification is an important area of forensics that analyses questioned documents. Generally, expensive methods such spectroscopy and reflectometry are used for differentiating inks. However, with the growth of imaging technology and easy availability, we can exploit vision based algorithms to aid, if not replace, existing methods. We have shown that BRDF based ink classification can be a promising alternative to classify inks.

The connecting theme of our research has been to improve and expand the scope of photometric methods for computer vision applications and beyond. Photometric methods have demonstrated that many computer vision problems such as shape estimation, BRDF estimation, material classification can lead promising solutions and better alternatives in comparison to other stereo algorithms.

7.1 Future work

We recommend the following possible directions for future work based on the research conducted for this thesis. Firstly, ring light constraints have already been proven to offer useful constraints for both Lambertian [34] and non-Lambertian [12] photometric stereo. In our auto-calibration solution, we have assumed that the surface is Lambertian and integrable. These assumptions were made to simplify the ambiguity and preserve the linearity of Lambertian models. A possible alternative is to extend the auto-calibration for non-Lambertian surfaces, building

on previous works in this direction such as [12].

Secondly, the idea of constraining the light sources can be explored further by trying novel light configurations. Drbohlav and Chantler [99] have shown that optimum lighting configurations can be determined for photometric stereo in the presence of camera noise. There is also a body of works in the area of computational illumination which prove that specific light patterns and positions can be used for material classification [23] and efficient BRDF acquisition [100]. It is an interesting direction to explore novel lighting configurations for problems such as auto-calibration, outlier removal and non-Lambertian photometric stereo.

Thirdly, we have only performed preliminary analysis using rectangularly placed light sources for calibrating photometric stereo. We have not validated our algorithm on real data. In future, we will extend our method to include the challenges of using real data. Our main aim for using rectangular light sources is towards reconstructing high quality depth maps by integrating high quality normals with the noisy depth maps from depth sensors. Most of the previous works in this direction assume zero ambient illumination, especially for photometric stereo. The advantages of this approach can be further enhanced by handling ambient illumination. One of the possible directions is to use infra red lights for illuminating the scene.

Fourthly, in our work on material classification, we have observed that different portions of the 2D BRDF space contain specific information. For example, retro-reflectance is more pronounced when $\theta_d = 0, \theta_h > \pi/3$ and Fresnel effects when $\theta_d > \pi/3$. An interesting study will be to understand the specific regions in the BRDF space where materials are sensitive on a case to case basis. This can improve the performance and efficiency of material classification by designing specific capture setups so that only the most essential portions of the BRDF space can be captured.

Lastly, in our ink classification approach, we have only used the 1D BRDF slice as a feature vector for classifying inks. A possible way to improve the performance is to add additional cues to the feature vector for classification. For example, stroke shape, pressure, velocity of the pen tip can be measured from the images to make the feature vector robust for enhanced classification.

Bibliography

[1] C. Hernandez, G. Vogiatzis, and R. Cipolla, "Multiview photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 548–554, 2008. 2, 36, 54

[2] Z. Zhou, Z. Wu, and P. Tan, "Multi-view photometric stereo with spatially varying isotropic materials," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1482–1489, 2013. 2, 31, 36

[3] http://www.blogcdn.com/www.diylife.com/media/2010/06/ chalk-uses-unusual-590kb060810.jpg. 3

[4] http://www.statuaryplace.com/store/images/Smallwoodbase.jpg. 3

[5] http://upload.wikimedia.org/wikipedia/commons/2/24/Polished_ steel_ball_60mm_diameter.jpg. 3

[6] http://image0-rubylane.s3.amazonaws.com/shops/owensantiques/ 11533.1L.jpg. 3

[7] R. Woodham, "Photometric stereo: A reflectance map technique for determining surface orientation from a single view," *SPIE*, vol. 155, pp. 136 – 143, 1978. 3, 17

[8] http://www.graphics.cornell.edu/~westin/image_gallery.html. 5

[9] `http://upload.wikimedia.org/wikipedia/commons/thumb/9/9a/`
`Rubber_duck.jpg/215px-Rubber_duck.jpg`. 5

[10] `http://thumbs.dreamstime.com/x/brushed-steel-5368208.jpg`. 5

[11] P. Tan, S. Mallick, L. Quan, D. Kriegman, and T. Zickler, "Isotropy, reciprocity and the generalized bas-relief ambiguity," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2007. 4, 24, 29

[12] N. G. Alldrin and D. J. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance : A stratified photometric stereo approach," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 1–8, 2007. 4, 7, 8, 28, 34, 35, 63, 64, 65, 67, 68, 69, 71, 106, 107

[13] M. Chandraker, J. Bai, and R. Ramamoorthi, "A theory of differential photometric stereo for unknown BRDFs," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2505–2512, 2011. 4, 8, 29, 63, 65, 66, 67, 68, 71

[14] S. Rusinkiewicz, "A new change of variables for efficient BRDF representation," in *Eurographics Rendering Workshop*, pp. 11 – 22, 1998. 4, 78, 87

[15] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 759–769, 2003. 6, 8

[16] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1060–1066, 1997. 7, 22, 23, 33, 35

[17] P. Favaro and T. Papadhimitri, "A closed-form solution to uncalibrated photometric stereo via diffuse maxima," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 821–828, 2012. 7, 24, 33, 63

[18] N. Alldrin, S. Mallick, and D. Kriegman, "Resolving the generalized bas-relief ambiguity by entropy minimization," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–7, 2007. 7, 23, 35

[19] O. Drbohlav and M. Chantler, "Can two specular pixels calibrate photometric stereo?," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, vol. 2, pp. 850–1857, 2005. 7, 24, 33, 35

[20] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneaux, S. Hodges, D. Kim, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 127–136, 2011. 7

[21] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi, "Efficiently combining positions and normals for precise 3D geometry," *ACM Transactions on Graphics*, pp. 536–543, 2005. 7, 50, 53, 54

[22] O. Wang, P. Gunawardane, S. Scher, and J. Davis, "Material classification using BRDF slices," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2805 –2811, 2009. 8, 32, 78, 79, 81, 83

[23] G. Jinwei and C. Liu, "Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral BRDF," in

Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR), pp. 797–804, 2012. 8, 32, 79, 81, 83, 107

[24] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis, *Radiometry*, ch. Geometrical Considerations and Nomenclature for Reflectance, pp. 94–145. USA: Jones and Bartlett Publishers, Inc., 1992. 14

[25] J. H. Lambert, *Lamberts Photometrie: (Photometria, Sive De Mensura Et Gradibus Luminis, Colorum Et Umbrae)*, vol. 2. 1760. 16

[26] B. T. Phong, "Illumination for computer generated pictures," *Commun. of the ACM*, vol. 18, no. 6, pp. 311–317, 1975. 16

[27] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *SIGGRAPH Comput. Graph.*, vol. 15, no. 3, pp. 307–316, 1981. 17

[28] S. Narasimhan, "Computer vision lecture notes," 2009. Available: `http://www.cs.cmu.edu/afs/cs/academic/class/15385-s06/lectures/ppts/`. 19

[29] R. Onn and A. Bruckstein, "Integrability disambiguates surface recovery in two-image photometric stereo," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 105–113, 1990. 19

[30] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *Journal of the Optical Society of America*, vol. 11, no. 11, pp. 3079–3089, 1994. 21, 22

[31] A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur, "Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability," *International Journal of Computer Vision*, vol. 35, pp. 203–222, 1999. 22

[32] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan, "Self-calibrating photometric stereo," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1118–1125, 2010. 24, 33, 35

[33] M. Chandraker, F. Kahl, and D. Kriegman, "Reflections on the generalized bas-relief ambiguity," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 788–795, 2005. 24

[34] Z. Zhou and P. Tan, "Ring-light photometric stereo," in *Proc. European Conference on Computer Vision (ECCV)*, pp. 265–279, Springer - Verlag, 2010. 24, 33, 34, 35, 36, 39, 40, 106

[35] A. Georghiades, "Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 816–823, 2003. 24

[36] P. Tan and T. Zickler, "A projective framework for radiometric image analysis," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2977–2984, 2009. 24

[37] Z. Wu and P. Tan, "Calibrating photometric stereo by holistic reflectance symmetry analysis," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1498–1505, 2013. 24, 35, 82

[38] J. Coleman and R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," *Computer Graphics and Image Processing*, vol. 18, no. 4, pp. 309–328, 1982. 25

[39] 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.*, vol. 25, no. 10, pp. 1239–1252, 2003. 25

[40] 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. Asian Conference on Computer Vision (ACCV)* , vol. 6494, pp. 703–717, 2011. 26

[41] K. Sunkavalli, T. Zickler, and H. Pfister, "Visibility subspaces: Uncalibrated photometric stereo with shadows," in *Proc. European Conference on Computer Vision (ECCV)*, pp. 251–264, 2010. 26

[42] D. Miyazaki, K. Hara, and K. Ikeuchi, "Median photometric stereo as applied to the segonko tumulus and museum objects," *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 229–242, 2010. 26, 77

[43] S. Nayar, K. Ikeuchi, and T. Kanade, "Visibility subspaces: Uncalibrated photometric stereo with shadows," in *Workshop on Industrial Applications of Machine Inteligence and Vision*, pp. 169–175, 1989. 26, 63

[44] S. Mallick, T. Zickler, D. Kriegman, and P. Belhumeur, "Beyond lambert: reconstructing specular surfaces using color," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 619–626, 2005. 26

[45] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H. peter Seidel, "Image-based reconstruction of spatially varying materials," in *In Proceedings of the 12th Eurographics Workshop on Rendering*, pp. 104–115, 2001. 26

[46] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz, "Shape and spatially varying BRDF's from photometric stereo," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 341–348, 2005. 26, 28, 31, 32, 63

[47] A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: Shape

reconstruction with general, varying BRDFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1254–1264, 2005. 27

[48] C. Yu, Y. Seo, and S. W. Lee, "Global optimization for estimating a BRDF with multiple specular lobes," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 319–326, 2010. 27

[49] S. Magda, D. J. Kriegman, T. Zickler, and P. N. Belhumeur, "Beyond lambert: reconstructing surfaces with arbitrary BRDFs," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 391–398, 2001. 28, 66

[50] T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *International Journal of Computer Vision*, vol. 49, no. 2-3, pp. 215–227, 2002. 28, 66

[51] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2008. 28

[52] R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 439–451, 1988. 30

[53] A. Agrawal, R. Raskar, and R. Chellappa, "What is the range of surface reconstructions from a gradient field?," in *Proc. European Conference on Computer Vision (ECCV)*, pp. 578–591, 2006. 30

[54] A. P. Harrison and D. Joseph, "Maximum likelihood estimation of depth maps using photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1368–80, 2012. 30

[55] P. Tan, L. Quan, and T. Zickler, "The geometry of reflectance symmetries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2506 – 2520, 2011. 33, 35

[56] O. Drbohlav and R. Sara, "Specularities reduce ambiguity of uncalibrated photometric stereo," in *Proc. European Conference on Computer Vision (ECCV)*, vol. 2, pp. 46 – 60, 2002. 33, 35

[57] A. Jones, G. Fyffe, X. Yu, A. Ma, J. Busch, M. Bolas, and P. Debevec, "Head-mounted photometric stereo for performance capture," in *ACM SIG-GRAPH 2010 Emerging Technologies*, SIGGRAPH '10, (New York, NY, USA), pp. 14:1–14:1, ACM, 2010. 34

[58] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *Journal of the Optical Society of America*, vol. 11, no. 11, pp. 3079–3089, 1994. 35, 63

[59] C. Wu, Y. Liu, Q. Dai, and W. B., "Fusing multiview and photometric stereo for 3d reconstruction under uncalibrated illumination," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 8, pp. 1082–1095, 2011. 36

[60] J.-Y. Bouguet, "Camera calibration toolbox for matlab," Available: `http://www.vision.caltech.edu/bouguetj/calib_doc`. 45

[61] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981. 47

[62] P. Besl and N. D. McKay, "A method for registration of 3D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992. 50

[63] R. Shiradkar, P. Tan, and S. H. Ong, "Auto-calibrating photometric stereo using ring light constraints," *Machine Vision and Applications*, vol. 25, no. 3, pp. 801–809, 2014. 50

[64] Q. Zhang, M. Ye, R. Yang, Y. Matsushita, B. Wilburn, and H. Yu, "Edge-preserving photometric stereo via depth fusion," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2472–2479, IEEE, 2012. 53, 55

[65] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, *et al.*, "Kinectfusion: real-time 3D reconstruction and interaction using a moving depth camera," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pp. 559–568, ACM, 2011. 55

[66] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," *Communications of the ACM*, vol. 56, no. 1, pp. 116–124, 2013. 55

[67] Microsoft, "Kinect camera," 2010. 55

[68] R. Anderson, B. Stenger, and R. Cipolla, "Augmenting depth camera output using photometric stereo.," in *IAPR Conference on Machine Vision Applications*, pp. 369–372, 2011. 55

[69] C. Hernández and G. Vogiatzis, "Self-calibrating a real-time monocular 3D facial capture system," in *International Symposium on 3D Data Processing, Visualization and Transmission*, 2010. 55

[70] C. Yu, Y. Seo, and S. Lee, "Photometric stereo from maximum feasible lambertian reflections," in *Proc. European Conference on Computer Vision (ECCV)*, pp. 115–126, 2010. 63

[71] P. Tan, S. Lin, L. Quan, and H. yeung Shum, "Highlight removal by illumination - constrained inpainting," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 164–169, 2003. 63

[72] S. A. Shafer, "Using color to separate reflection components," *Color Research and Application*, vol. 10, no. 4, pp. 210–218, 1985. 63

[73] P. Tan and T. Zickler, "A projective framework for radiometric image analysis," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2977–2984, 2009. 63, 67, 70

[74] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM Transactions on Mathematical Software*, vol. 8, pp. 43–71, 1982. 72

[75] R. Shiradkar and S. H. Ong, "Surface reconstruction using isocontours of constant depth and gradient," in *Proc. of 20th IEEE International Conference on Image Processing (ICIP)*, pp. 360–363, Sept 2013. 75

[76] C. Wu, B. Jaramaz, and S. G. Narasimhan, "A full geometric and photometric calibration method for oblique-viewing endoscope," *Computer Aided Surgery*, vol. 15, no. 1-3, pp. 19–31, 2010. 77

[77] C. Hernandez, G. Vogiatzis, G. Brostow, B. Stenger, and R. Cipolla, "Nonrigid photometric stereo with colored lights," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 1–8, 2007. 77

[78] J. Kautz and M. D. McCool, "Interactive rendering with arbitrary BRDFs using separable approximations," in *Eurographics*, pp. 247–260, 1999. 78

[79] M. D. McCool, J. Ang, and A. Ahmad, "Homomorphic factorization of BRDFs for high-performance rendering," in *Proceedings of the ACM SIGGRAPH*, pp. 171–178, 2001. 78

[80] M. M. Stark, J. Arvo, and B. Smits, "Barycentric parameterizations for isotropic BRDFs," *IEEE Trans. on Visualization and Computer Graphics*, vol. 11, no. 2, pp. 126–138, 2005. 78

[81] M. Jehle, C. Sommer, and B. Jähne, "Learning of optical illumination for material classification," *Pattern Recognition*, vol. 6376, pp. 563–572, 2010. 79, 81

[82] A. Licata, A. Psarrou, and V. Kokla, "Unsupervised ink type recognition in ancient manuscripts," in *Proc. ICCV, Workshops*, pp. 955–961, 2009. 79, 82

[83] J. A. Siegel, *Ink Analysis*, pp. 375–379. Elsevier, second ed., 2013. 79, 90

[84] B. E. Fau and R. B. Dyer, "Fourier transform hyperspectral visible imaging and the nondestructive analysis of potentially fraudulent documents," *Applied Spectroscopy*, vol. 60, no. 3, pp. 833–840, 2006. 80, 81

[85] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 779–785, 1994. 80, 81

[86] B. Chakravarthy and H. Dasari, "Classification of liquid and viscous inks using hsv colour space," in *Proc. IEEE Intl. Conference on Document Analysis and Recognition (ICDAR)*, pp. 660–664, 2005. 81, 102

[87] V. Kokla, A. Psarrou, and V. Konstantinou, "Ink recognition based on statistical classification methods," in *Proceedings of the International Conference on Document Image Analysis for Libraries*, pp. 254–264, 2006. 82

[88] H. S. Chen, H. H. Meng, and K. C. Cheng, "A survey of methods used for the identification and characterization of inks," *Forensic Science Journal*, 2002. 82

119

[89] C. E. H. Berger, "Objective ink color comparison through image processing and machine learning," *Science and Justice : Journal of the Forensic Science Society*, vol. 53, no. 1, pp. 55–59, 2013. 82, 102, 103

[90] F. Romeiro, Y. Vasilyev, and T. Zickler, "Passive reflectometry," in *Proc. European Conference on Computer Vision (ECCV)*, pp. 859–872, 2008. 82

[91] L. Mcmillan, A. Smith, and W. Matusik, "A data-driven reflectance model," in *Proceedings of the ACM SIGGRAPH*, pp. 759–769, 2003. 82

[92] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi., "A biquadratic reflectance model for radiometric image analysis," in *Proc. IEEE Intl. Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 230–237, 2012. 83

[93] E. Caputo, B. Hayman and P. Mallikarjuna, "Class-specific material categorisation," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 1597–1604, 2005. 86

[94] J. Machay, "Types of printer inks," Available: `http://smallbusiness.chron.com/types-printer-inks-57328.html`. 91

[95] C. Wu, "Visualsfm: A visual structure from motion system," 2011. Available: `http://ccwu.me/vsfm/`. 92

[96] A. Yamashita, A. Kawarago, T. Kaneko, and K. T. Miura, "Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system," in *Proceedings of the IEEE International Conference on Pattern Recognition*, pp. 482–485, 2004. 92

[97] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, pp. 972–977, 2007. 93

[98] R. Shiradkar, L. Shen, G. Landon, S. H. Ong, and P. Tan, "Surface reconstruction using isocontours of constant depth and gradient," in *Proc. IEEE*

Intl. Conference on Computer Vision and Pattern Recognition (CVPR), p. (to appear), 2014. 102

[99] O. Drbohlav and M. Chantler, "On optimal light configurations in photometric stereo," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, vol. 2, pp. 1707–1712, 2005. 107

[100] A. Ghosh, S. Achutha, W. Heidrich, and M. O'Toole, "BRDF acquisition with basis illumination," in *Proc. IEEE Intl. Conference on Computer Vision (ICCV)*, pp. 1–8, 2007. 107

[101] M. P. Knapp, "Sines and cosines of angles in arithmetic progression," *Mathematics Magazine*, vol. 82, no. 5, 2009. 124

Appendix A

Proof for Establishing the Ambiguities following Ring Light Constraints

Proposition *The true normal $\mathbf{n_g}$ of a Lambertian surface illuminated by a ring-light at a distance d_g can be recovered up to a classical bas-relief ambiguity compounded with a planar rotation ambiguity, i.e.*

$$\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} \cos\triangle & \sin\triangle & 0 \\ -\sin\triangle & \cos\triangle & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{d_g}{d} \end{pmatrix} \begin{pmatrix} n_{g1} \\ n_{g2} \\ n_{g3} \end{pmatrix} \tag{1}$$

$$\text{or} \quad \mathbf{n} = R_\triangle S \mathbf{n}_g \tag{2}$$

Here, $\triangle = \theta_* - \theta_0$, d and θ_* are the assumed distance and initial rotation angles respectively. n is the recovered normal.

Proof. The pseudo inverse of a matrix $\mathbf{L}$ is computed as $\mathbf{L}^\dagger = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T$. The

matrix $\mathbf{L}$ as defined in (3.2) in Section 3.2. We can write the matrix $\mathbf{L}^T\mathbf{L}$ as,

$$\mathbf{L}^T\mathbf{L} = \begin{bmatrix} r^2 \sum\limits_{i=0}^{N} \cos^2\alpha & r^2 \sum\limits_{i=0}^{N} \sin\alpha\cos\alpha & dr \sum\limits_{i=0}^{N} \cos\alpha \\ r^2 \sum\limits_{i=0}^{N} \sin\alpha\cos\alpha & r^2 \sum\limits_{i=0}^{N} \sin^2\alpha & dr \sum\limits_{i=0}^{N} \sin\alpha \\ dr \sum\limits_{i=0}^{N} \cos\alpha & dr \sum\limits_{i=0}^{N} \sin\alpha & N+1d^2 \end{bmatrix} \tag{3}$$

where $\alpha = (\theta_0 + it)$ and i is the index in the summation. This can be rewritten as,

$$\mathbf{L}^T\mathbf{L} = r^2 \begin{bmatrix} \frac{1}{2} \sum\limits_{i=0}^{N} (1 + \cos 2\alpha) & \frac{1}{2} \sum\limits_{i=0}^{N} \sin 2\alpha & \frac{d}{r} \sum\limits_{i=0}^{N} \cos\alpha \\ \frac{1}{2} \sum\limits_{i=0}^{N} \sin 2\alpha & \frac{1}{2} \sum\limits_{i=0}^{N} (1 - \cos 2\alpha) & \frac{d}{r} \sum\limits_{i=0}^{N} \sin\alpha \\ \frac{d}{r} \sum\limits_{i=0}^{N} \cos\alpha & \frac{d}{r} \sum\limits_{i=0}^{N} \sin\alpha & (N+1)\frac{d^2}{r^2} \end{bmatrix} \tag{4}$$

From [101], we have the following identities,

$$\sum_{i=0}^{N} \cos(\theta_0 + it) = \frac{\sin\left(\frac{(N+1)t}{2}\right)\cos\left(\theta_0 + \frac{Nt}{2}\right)}{\sin\left(\frac{t}{2}\right)}$$

and

$$\sum_{i=0}^{N} \sin(\theta_0 + it) = \frac{\sin\left(\frac{(N+1)t}{2}\right)\sin\left(\theta_0 + \frac{Nt}{2}\right)}{\sin\left(\frac{t}{2}\right)} \tag{5}$$

Substituting the above identities in (4), the inverse of the matrix $\mathbf{L}^T\mathbf{L}$ is calculated as,

$$(\mathbf{L^T L})^{-1} = \frac{1}{d^2 K} \begin{bmatrix} d^2 A & d^2 B & dC \\ d^2 B & d^2 D & dC \\ dC & dE & F \end{bmatrix} \tag{6}$$

where A, B, C, D, E, F and K are terms which are functions of the parameters θ_0, r, N and t and independent of d.

The pseudo inverse of $\mathbf{L}$ is computed as,

$$\mathbf{L}^{\dagger} = (\mathbf{L^T L})^{-1} \mathbf{L^T} = \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2N} \\ \frac{1}{d}a_{31} & \cdots & \frac{1}{d}a_{3k} & \cdots & \frac{1}{d}a_{3N} \end{bmatrix} \qquad (7)$$

where a_{ij} are a function of θ_0, r, N and t.

The normal is obtained by multiplying this with the intensity matrix $\mathbf{I}$,

$$\rho\mathbf{n} = \mathbf{L}^{\dagger}\mathbf{I} = \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2N} \\ \frac{1}{d}a_{31} & \cdots & \frac{1}{d}a_{3k} & \cdots & \frac{1}{d}a_{3N} \end{bmatrix} \begin{bmatrix} I_1 \\ \vdots \\ I_k \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \frac{1}{d}b_3 \end{bmatrix} \qquad (8)$$

where, b_{ij} are also functions of θ_0, r, N and t.

The error term at each pixel is now computed as,

$$e(d) = \|\mathbf{I} - \mathbf{L}(d) \cdot \rho\mathbf{n}\|$$
$$= \left\| \mathbf{I} - \begin{bmatrix} r\cos\theta_0 & r\sin\theta_0 & d \\ \vdots & \vdots & \vdots \\ r\cos(\theta_0 + kt) & r\sin(\theta_0 + kt) & d \\ \vdots & \vdots & \vdots \\ r\cos(\theta_0 + Nt) & r\sin(\theta_0 + Nt) & d \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \frac{1}{d}b_3 \end{bmatrix} \right\| \qquad (9)$$

or

$$e\left(d\right) = \left\| \mathbf{I} - \begin{bmatrix} c_1 \\ \vdots \\ c_k \\ \vdots \\ c_N \end{bmatrix} \right\| \tag{10}$$

where, c_i are also functions of θ_0, r, N and t.

Hence, we observe that although the normal obtained from pseudo inverse of $\mathbf{L}$ is a function of the distance d, yet the product of $\mathbf{L}$ and $\mathbf{n}$ is independent of d. As the distance is varied, since all other parameters remain constant, the error remains constant while the computed normal changes as a function of d. In other words, we can say that for two different values of the distance term, say d and d_g, and the corresponding light source directions $\mathbf{L}\left(d\right)$ and $\mathbf{L}\left(d_g\right)$,

$$\mathbf{L}\left(d\right) \cdot \rho\mathbf{n} = \mathbf{L}\left(d_g\right) \cdot \rho\mathbf{n_g} \tag{11}$$

where $\mathbf{n}$ and $\mathbf{n_g}$ are the respective normals at the distance d and d_g. From (8), the normals obtained at a distance d and d_g are

$$\mathbf{n} = \begin{bmatrix} b_1 & b_2 & \frac{1}{d}b_3 \end{bmatrix}^T \text{ and } \mathbf{n_g} = \begin{bmatrix} b_1 & b_2 & \frac{1}{d_g}b_3 \end{bmatrix}^T \tag{12}$$

Here the b_i's are independent of the term d. We can now rewrite $\mathbf{n}$ as

$$\mathbf{n} = \begin{bmatrix} b_1 & b_2 & \frac{1}{d}b_3 \end{bmatrix}^T = \begin{bmatrix} b_1 & b_2 & \left(\frac{d_g}{d}\frac{1}{d_g}b_3\right) \end{bmatrix}^T$$

126

Thus, we have,

$$
\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{d_g}{d} \end{bmatrix} \begin{bmatrix} n_{g1} \\ n_{g2} \\ n_{g3} \end{bmatrix} \text{ or } \mathbf{n} = S\mathbf{n}_g \tag{13}
$$

where $\mathbf{n}$ and $\mathbf{n_g}$ are as defined in the statement of Proposition1.

Next, we show that the recovered normals are up to a rotational ambiguity. Note that these two are independent of each other. Let θ_g be the initial angle and θ_* be an arbitrary estimate of the initial angle. We know the position of each light source $\mathbf{l_g}$ is given by

$$
\mathbf{l_g} = \begin{bmatrix} r\cos(\theta_g + (i-1)t) & r\sin(\theta_g + (i-1)t) & d \end{bmatrix} \tag{14}
$$

where, $t = \frac{2\pi}{N}$ and i is the index of the light source $(i = 1, ..., N)$

Now, for an arbitrary θ_*, the light source position $\mathbf{l_*}$at the same index is computed as

$$
\begin{aligned}
\mathbf{l_*} &= \begin{bmatrix} r\cos(\theta_* + (i-1)t) & r\sin(\theta_* + (i-1)t) & d \end{bmatrix} \\
&= \begin{bmatrix} r\cos(\theta_g + (i-1)t + \theta_* - \theta_g) & r\sin(\theta_g + (i-1)t + \theta_* - \theta_g) & d \end{bmatrix} \\
&= \begin{bmatrix} r\cos(\theta_g + (i-1)t) & r\sin(\theta_g + (i-1)t) & d \end{bmatrix} \mathbf{R}_\triangle \\
&= \mathbf{l_g}\mathbf{R}_\triangle
\end{aligned}
$$

where $\triangle = \theta_* - \theta_0$ and

$$
\mathbf{R}_\triangle = \begin{bmatrix} \cos\triangle & \sin\triangle & 0 \\ -\sin\triangle & \cos\triangle & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{15}
$$

From (15), we observe a planar rotation ambiguity (in the xy plane) between

127

the estimated and the true light source positions. We can derive the following relationship between the true normal and estimated normal as,

$$\mathbf{I} = \rho \mathbf{l}_* \mathbf{n}$$

$$= \rho \mathbf{l_g} \mathbf{R}_\triangle \mathbf{n_g}$$

$$= \rho \mathbf{l_g} \mathbf{n_g}$$

We can now say that

$$\mathbf{n} = \mathbf{R}_\triangle \mathbf{n_g} \tag{16}$$

In other words, for any arbitrary θ_*, the recovered normals are rotated by an angle $\triangle = (\theta_* - \theta_g)$. $\qquad\square$

Appendix B

Parameters of the Solution to

Scaling Ambiguity

If $\mathbf{n_i} = \begin{bmatrix} n_{i1} & n_{i2} & n_{i3} \end{bmatrix}$ and $T = \{t_{ij}\}_{3\times3}$, then the constants A_{ij} are,

$$
\begin{aligned}
A_{11} &= n_{12}n_{21}t_{31} + n_{12}n_{22}t_{32} \\
A_{12} &= n_{12}t_{22}n_{23} \\
A_{13} &= -n_{13}n_{23}t_{23} \\
A_{14} &= -n_{13}t_{21}n_{21} - n_{13}n_{21}t_{21} \\
A_{21} &= -n_{21}t_{31}n_{11} - n_{22}t_{32}n_{11} \\
A_{22} &= n_{13}n_{23}t_{13} \\
A_{23} &= -n_{11}n_{23}t_{22} \\
A_{24} &= n_{13}t_{11}n_{21} + n_{13}n_{22}t_{12} \\
A_{31} &= n_{11}t_{23}n_{23} - n_{12}n_{23}t_{13} \\
A_{32} &= n_{11}t_{21}n_{21} + n_{11}t_{22}n_{22} - n_{12}t_{11}n_{21} - n_{12}t_{12}n_{22}
\end{aligned}
\tag{17}
$$

Appendix C

Names of the Inks used to Construct Ink Database

Fountain

1. Pelikan

2. Aurora

3. Monteverde

4. Lamy

5. Hero

6. Parker

Marker

1. MaxFlo

2. IdentiPen

3. Pentel WB

4. Zebra Permanent

5. Sakura Permanent

6. Omni Permanent

7. Pilot Permanent

8. Pilot WB

9. Zig Permanent

10. Pentel Permanent

Print

1. Samsung Laser

2. HP Laser

3. Lexmark Laser

4. Photocopy

5. InkJet

Gel

1. Pentel EnerGel

2. Pilot Signature

3. Jet Stream

4. Uniball Signo

5. Faber Castell Clickball

6. Monami Geller

7. Uni Style Fit

8. Pilot BPS

9. Pop Bazic Gel

10. Zebra Z1

11. Lotus Gel

12. Zebra Jimnie Gel

Ball Point

1. Omni Softball

2. Pentel Star

3. Stabilo Liner

4. M& G Devilsmask

5. Staedtler 422

6. Uni Locknock

7. Faber Castell Super Click

8. Pilot Super Grip

9. Monami BPP

10. Monami 153 Stick

11. BIC Bu2

12. Zebra

13. Anonymous Black 1

14. Anonymous Black 2

Pencils

1. Staedtler 3B

2. Faber Castell 6B

3. Staedtler HB

4. Staedtler Black Coloured Pencil

5. Faber Castell 2B

6. Be Goody A 850 HB

7. Geddes HB