Surface Shape and Reflectance Analysis Using Polarisation

Gary A. Atkinson

Submitted for the degree of Doctor of Philosophy

Department of Computer Science

THE UNIVERSITY of York

May 2007

Abstract

When unpolarised light is reflected from a smooth dielectric surface, it becomes partially polarised. This is due to the orientation of dipoles induced in the reflecting medium and applies to both specular and diffuse reflection. This thesis aims to exploit the polarising properties of surface reflection for computer vision applications. Most importantly, the thesis proposes novel shape and reflectance function estimation techniques. The methods presented rely on polarisation data acquired using a standard digital camera and a linear polariser. Fresnel theory lies at the heart of the thesis and is used to process the polarisation data in order to estimate the surface normals of target objects.

Chapter 2 surveys the related literature in the fields of polarisation vision, shape-from-shading, stereo techniques, and reflectance function analysis. Chapter 3 commences by presenting the underlying physics of polarisation by reflection, starting with the Fresnel equations. The outcome of this theory is a means to ambiguously estimate surface normals from polarisation data, given a rough estimate of the material refractive index. The first novel technique is then presented, which is a simple single-view approach to shape reconstruction. In this case, efficiency is given priority over accuracy. Chapter 3 ends with a description of a device for measuring refractive indices.

Chapter 4 is concerned with image-based reflectance function estimation. Firstly, the special case of retro-reflection is assumed. An algorithm is described that constructs a histogram of surface zenith angles and pixel intensities. Probability density functions are then robustly fitted to this histogram to obtain a one-dimensional "slice" of the material BRDF. A second algorithm is then presented, which is designed for more general illumination conditions. This uses a three-dimensional histogram that includes the surface azimuth angles, in addition to the zenith angles and pixel intensities. Simulated annealing optimisation is then used to fit surfaces to the histogram, thus estimating a two-dimensional slice of the BRDF. Chapter 4 also contains a method for photometric stereo, which is used by the above two-dimensional BRDF technique to fully constrain the previously ambiguous surface normal estimates.

The most sophisticated and accurate shape reconstruction technique is presented in Chapter 5. The reflectance function algorithm described previously is first applied to enhance the surface normal estimates obtained from the raw polarisation data. This is done for two views. An new algorithm is then described that solves the stereo correspondence problem using the refined surface normals. To do this, a set of patches are extracted from each view and are aligned by minimizing an energy functional based on the surface normal estimates and local topographic properties. The optimum alignment parameters for different patch pairs then gives a matching energy. The combination of pairings that minimises the total matching energy for all patches relates to the correct correspondences. In solving the correspondence problem in this way, two fields of accurate and unambiguous surface normals are acquired which can subsequently be converted into depth.

Our techniques are most suited to smooth, non-metallic surfaces. The multi-view method complements existing stereo algorithms since it does not require salient surface features to obtain correspondence. The use of shading information is shown to significantly improve the accuracy of the reconstructions compared to most previous polarisation-based methods. A set of experiments, yielding reconstructed objects and reflectance functions, are presented and compared to ground truth. The experiments involve a variety of materials and object geometries. The work provides the foundation for the development of a novel sensor that can non-intrusively recover depth information using two cameras. Suggestions for such a device are outlined in Chapter 6.

Contents

Declaration

I declare that the work in this thesis is solely my own except where attributed and cited to another author. Most of the material in this thesis has been previously published by the author. A complete list of publications can be found on page viii.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Edwin Hancock, for his continued support, advice and suggestions during my research and writing up. His breadth of knowledge in the field is considerable. He has been invaluable both as a supervisor and as a friend.

Secondly, appreciation goes to my assessor, Dr. Richard Wilson, for his constructive feedback on this thesis and on various reports and presentations. Thank you also to my external examiner, Prof. Andrew Wallace. For financial support, I thank the EPSRC funding body.

My thanks go to various members of the computer vision group for invaluable technical discussions and exchanges of ideas. Most notably, I thank Dr. Antonio Robles-Kelly and Dr. Hossein Ragheb for their assistance in my first year of study and Dr. William Smith for useful suggestions during the remainder.

Finally, to my parents. For their support, encouragement and love. To them, I will be forever grateful.

List of Publications

Journal Papers

- G. A. Atkinson and E. R. Hancock. Recovery of Surface Orientation from Diffuse Polarization. *IEEE Transactions on Image Processing* 15:1653-1664, June 2006.
- G. A. Atkinson and E. R. Hancock. Shape Estimation using Polarization and Shading from Two Views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. To appear.

Conference Papers

- G. A. Atkinson and E. R. Hancock. Shape from Diffuse Polarisation. In. *Proc. British Machine Vision Conference*, pp. 919-928. University of Kingston, London, UK 2004.
- G. A. Atkinson and E. R. Hancock. Recovery of Surface Height from Diffuse Polarisation. In. *Proc. International Conference on Image Analysis and Recognition*, vol. 1, pp. 621-628. Porto, Portugal 2004.
- G. A. Atkinson and E. R. Hancock. Recovery of Surface Height Using Polarization from Two Views. In. *Proc. Computer Analysis of Images and Patterns*, pp. 162-170. Rocquencourt, Paris, France 2005.
- G. A. Atkinson and E. R. Hancock. Multi-view Surface Reconstruction Using Polarization. In. Proc. *International Conference on Computer Vision*, pp. 309-316. Beijing, China 2005.
- G. A. Atkinson and E. R. Hancock. Analysis of Directional Reflectance and Surface Orientation using Fresnel Theory. In. *Proc. Iberoamerican Congress on Pattern Recognition*, pp. 103-111. Havana, Cuba 2005.

- G. A. Atkinson and E. R. Hancock. Polarization-based Surface Reconstruction via Patch Matching. In *Proc. Computer Vision and Pattern Recognition*, vol. 1, pp. 495-503. New York, USA 2006.

- G. A. Atkinson and E. R. Hancock. Robust Estimation of Reflectance Functions from Polarization. In *Proc. Iberian Conference on Pattern Recognition and Image Analysis*, vol. 2, pp. 363-371. Girona, Spain 2007.

- G. A. Atkinson and E. R. Hancock. Surface Reconstruction using Polarization and Photometric Stereo. In *Proc. Computer Analysis of Images and Patterns*. To appear. Vienna, Austria 2007.

List of Figures

List of Symbols

$\mathcal{BC}$	Set of points in a histogram bin (bin contents)
$\mathcal{D}$	Set of Cartesian pairs, $\{\theta_d, I_d\}$, or triples, $\{\theta_d, \alpha_d, I_d\}$, used in histogram
F	Fresnel function
f	BRDF
$\mathcal{HC}$	Histogram contents ($\mathcal{HC} = \lvert\mathcal{BC}\rvert$)
h	Surface height
I	Pixel intensity
L	Radiance
M	Correspondence indicator matrix
n	Index of refraction
p_x, p_y, p_z	Surface normal components
s	Shape index
$\mathcal{U}, \mathcal{V}, \mathcal{W}$	Sets of points forming patches
x, y, z	Cartesian position
α	Azimuth angle
γ	Angle of the projection of the normal onto the $y - z$ plane (Figure 3.5)
ε	Energy/cost function
$\overline{\boldsymbol{\Theta}}$	Transformation parameter vector
θ	Zenith angle
θ_{pol}	Polariser angle
θ_{rot}	Object rotation angle
κ	Mean curvature
ρ	Degree of polarisation
ϱ	Total diffuse albedo
ϕ	Phase angle (angle of polarisation)

Subscripts

i	Incidence
L	Illuminance
r	Reflectance
t	Transmittance

Chapter 1

Introduction

When people see images, they can usually predict the three-dimensional geometry of the objects in view. On many occasions, the person will have seen the objects before. In this case the brain is effectively performing shape recognition on the objects, before applying a priori knowledge to estimate their shapes. However, the visual system is more advanced than this and uses many independent cues to aid in the perception of depth. Such cues include shading, motion parallax, texture, perspective effects, size, binocular vision, and more [97].

Figure 1.1, shows a rendering of a hemisphere. Without any prior knowledge, most people would immediately conclude that the object is spherical. The ability of humans to deduce shape in this way has, for many years, been the motivation for the field of shape-from-shading [40]. That is, the use of computer algorithms to estimate the three-dimensional geometry of a surface from a single image. The computer vision community reasoned that, since humans are so efficient at deducing shapes from previously unseen shading patterns, it should be possible for computers to do the same. After all, for Figure 1.1 specifically, shading is the only source of information available.

After several decades of research, constructing images based on shading alone has proved to be more difficult than expected, as explained in Chapter 2. This thesis aims to complement existing efforts in the field, by incorporating an additional cue into the shape estimation procedure: the polarisation state of reflected light.

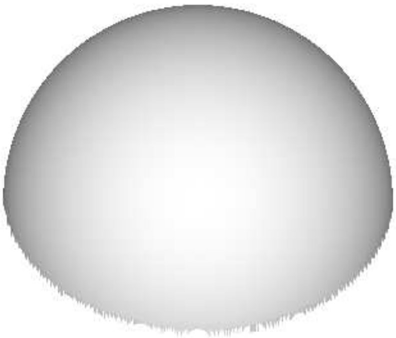

Figure 1.1: A rendering of a hemisphere.

1.1 Vision and Polarisation

After concluding that shading information alone is an insufficient source of shape information in many cases, the computer vision community turned to other shape cues. One of these was polarisation. It is not immediately obvious that polarisation provides any useful knowledge about surface shape. This is because the human eye is insensitive to the phenomenon. Despite this, a great deal of natural light found in everyday life is polarised. There are several mechanisms found in Nature whereby sunlight (which is initially unpolarised) becomes partially or completely polarised. The most ubiquitous of these mechanisms are scattering, refraction and reflection [49].

Polarisation by scattering occurs in the atmosphere and is caused by gas molecules and dust particles. The result is that the sky has a polarisation pattern where the light can be as much as 70% linearly polarised when observed at an angle 90° to the Sun. Scattering also occurs due to water molecules beneath the world's seas and oceans. Polarisation by refraction is mainly found underwater, where the transmission of the light at the air/water interface causes significant polarisation. Finally, when light is reflected from surfaces there is a partial polarising effect. The strength of polarisation by reflection depends upon the reflecting material and other factors.

We now know that many creatures, including some insects, birds and marine animals, have eyes that are sensitive to the polarisation state of light [100]. Bees and ants, for example, use the polarisation pattern in the sky to aid navigation. Several marine creatures use similar patterns found underwater for the same purpose [41]. Surprisingly, the first

Figure 1.2: Photograph of a stomatopod or mantis shrimp (Odontodactylus havanenis) [15].

people to exploit the polarisation patterns of skylight were the Vikings [49, §12]. It is widely believed that they used calcite crystals to predict the direction of the Sun after sunset or before sunrise. Suggestions have recently been made that the patterns could be used in computer vision systems [68].

One of the most sophisticated polarisation-sensitive eyes in Nature is that of the mantis shrimp. A photograph of one of these creatures is shown in Figure 1.2. The special feature of the eye of the mantis shrimp is its high sensitivity to both colour and polarisation. There is evidence [16] that the creature, and a handful of other species, use polarisation to isolate objects of interest within their field of view. It is also believed that some species communicate by reflecting light only of a certain polarisation angle. The polarising properties of reflection are also used by some insects who orient themselves according to the horizontally polarised light reflected from flat water.

Given the abundant use of polarisation in Nature, it is not surprising that people have attempted to emulate its exploitation for industrial applications. This thesis is concerned with the exploitation of the polarising properties of surface reflection. As explained in subsequent chapters, the partially linearly polarised light that is reflected from surfaces can be described by three quantities: the intensity of the light, the angle of the polarisation (the phase), and the extent to which the light is polarised (the degree of polarisation). Refer back to Figure 1.1. This figure shows a rendered image of a surface similar what could be obtained using a normal camera or a human eye. Essentially, it corresponds to the *intensity* of light reflected from the surface. Now consider Figure 1.3, which simu-

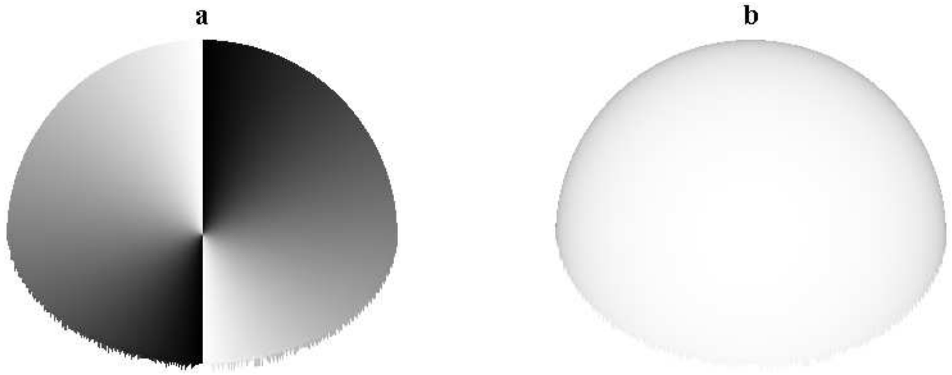

Figure 1.3: Simulation of (a) the angle of the polarisation of light reflected from a hemisphere and (b) the degree of polarisation. The intensity of these images encodes the two quantities respectively.

lates the angle of polarisation and the degree of polarisation using the theory described in Chapter 3. Collectively, the three quantities clearly contain significantly more shape information than the intensity alone. Combining information from each of the quantities for the purpose of shape estimation forms the central aspect of this thesis.

1.2 Motivation and Contribution

As explained in the next chapter, there have been many attempts to incorporate different cues into the shape recovery process for computer vision. One of the most powerful means by which surface geometry can be constrained is to use more than one image of the object. In the case of smooth surfaces, one of the best-known methods is photometric stereo [120]. This method involves imaging the target object using several different light source directions, and can fully constrain the surface orientation at each point. In the case where detectable surface features are present, geometric stereo is possible [32]. Using this method, the light source and object are fixed and two viewing directions are used to estimate depth through triangulation.

The primary goal of this thesis is to develop a shape recovery technique that uses polarisation to supplement shading information. Polarisation has been suggested previously for shape reconstruction but not in conjunction with shape-from-shading. This combination permits the accurate estimation of surface orientation, whilst avoiding the need for the time consuming light source repositioning necessary in photometric stereo. However, the surface geometry can still not be fully determined in this way from a single image. For this reason, this thesis also makes a contribution to the field geometric stereo to constrain the object geometry completely. Most existing geometric stereo methods encounter difficulty when presented with featureless surfaces. In our new method, polarisation information is used to establish correspondence between two views, and can be applied to textureless and featureless surfaces. Our method therefore addresses the above weakness of geometric stereo.

In addition to the development of a novel method for shape reconstruction. The thesis has several secondary contributions. The thesis tests the most basic form of polarisation vision on a wider range of materials and surface geometries than has previously been considered. Also, it shows how photometric stereo can be enhanced using polarisation data (this is not used in the main reconstruction algorithm). Finally, it shows how the reflectance properties of different materials can be estimated. Two methods with this aim are presented, one of which provides the means of incorporating shading information into the shape reconstruction procedure.

1.3 Thesis Overview

In Chapter 2, we present a detailed discussion of related work in the field. The most relevant area of past research is in the subject of shape recovery using polarisation. This is covered in Section 2.1, which also mentions a few other applications of polarisation vision. Section 2.2 is a survey of developments in shape-from-shading and stereo (both geometric and photometric) methods. A few additional shape reconstruction methods are also listed. Section 2.3 is concerned with research into material reflectance properties. Techniques to both model and estimate the properties are reviewed.

Chapter 3 commences by presenting the relevant details of the background theory from the field of physical optics. This is presented in Section 3.1, with Fresnel theory for the reflection of electromagnetic waves from interfaces taken as the starting point. The theory leads to a means to estimate the degree and angle of polarisation of light reflected from a surface. In Section 3.2, a method for obtaining shape using polarisation from a single view is presented. Experiments in shape reconstruction, which includes an analysis of the accuracy of the polarisation measurements, are described in Section 3.3.

Chapter 4 is on the subject of image-based methods for estimating reflectance properties. The background theory and nomenclature for this field is given in Section 4.1. In Section 4.2, a method is proposed that estimates the relationship between the reflected radiance and the surface orientation for the simplified case where the light source and camera lie in the same direction from the object. The method for enhancing photometric stereo is presented in Section 4.3. In Section 4.4 the reflectance properties are estimated for more general illumination conditions. Experiments with various materials are presented for each of the methods.

Chapter 5 uses the estimated reflectance properties to combine shading and polarisation information from two views, thus acquiring more accurate orientation estimates than would otherwise be possible. Sections 5.1 to 5.3 contain details of the two-view method for reconstruction and includes a new patch matching algorithm. A set of reconstructions are presented in 5.4, which also shows a few images rendered using the reflected functions estimated in Chapter 4.

Chapter 6 summarises the contributions of the thesis, highlights the strengths and weaknesses, and suggests avenues for future research.

Chapter 2

Literature Review

The ultimate goal of this thesis is to develop a shape recovery technique for surfaces that are both smooth and featureless. The final outcome, presented towards the end of the thesis, is a technique involving polarisation (Chapter 3), reflectance functions (Chapter 4) and geometric stereo (Chapter 5). The thesis also makes contributions to photometric stereo and shape-from-shading (SFS). Broadly speaking, polarisation is used to obtain initial (and ambiguous) surface normal estimates, and is the foundation for the thesis. The reflectance functions and SFS enhance these estimates, while the stereo algorithms disambiguate the estimated surface normals. Standard integration methods are then used to recover depth.

This chapter presents a survey of the existing literature in the above fields. Since the foundations of the entire thesis are drawn from polarisation, we commence with a summary of related work in the field of polarisation vision. The main emphasis is on shape recovery techniques, but several other applications are also mentioned. Section 2.2 then describes other shape recovery techniques, principally SFS and stereo. In Section 2.3, key contributions to the field of reflectance function modelling and estimation are discussed. Both theoretical and practical approaches are considered. The chapter concludes with a summary of the successes and failures of the existing methods and a statement on the position of this thesis in the context of other research.

2.1 Polarisation Vision

Research into the use of polarisation in computer vision has attracted a significant amount of attention during the last twenty or so years. Most of the research relies on the fact that when light is reflected from a surface, it undergoes a partial polarisation [35]. That is, if the incident light is unpolarised, then the reflected light will be a superposition of unpolarised light and completely polarised light. In general, the polarisation state of light is represented by the Stokes vector, which contains four degrees of freedom. These correspond to the intensity of the light, the extent to which the light is partially polarised (later quantified by the *degree of polarisation*), the angle of the completely polarised component (the *phase* angle) and the circular polarisation. The last component is negligible for most situations however, so the majority research concentrates on one or more of the other three.

As explained below, the information contained in the polarisation state of light arriving at a detector can be processed and interpreted in many different ways. However, the large majority of recent papers on image-based methods have two aspects in common. The first is that the phase of the polarisation takes a central role. This is mainly because it is easy to measure and is largely independent of the material. The second common aspect is that the polarisation state is captured using a linear polariser and by measuring the intensity variation for different transmission axis orientations. The transmission axis can be varied either by manually rotating a standard polariser mounted on a digital camera [117], or electronically using liquid crystals [116]. We now present a review of the related literature that uses polarisation for shape recovery.

2.1.1 Shape Reconstruction

The first attempt to use polarisation analysis in computer vision dates back to Koshikawa in 1979 [51]. Koshikawa analysed the change in polarisation state of circularly polarised light from non-conducting (dielectric) surfaces to provide constraints on surface orientation. The the need for circularly polarised incident light is the obvious disadvantage of this early work. Also, the method was made complicated by the use of Mueller calculus [6] to manipulate the full Stokes vectors.

Both of these issues were resolved when Wolff and Boult applied Fresnel theory to polarisation vision [117]. Fresnel theory (see [35], [6] or Section 3.1) is a quantitative description of the direct reflection of electromagnetic waves from smooth interfaces. It can be used to predict the degree of polarisation of light reflected from surfaces at different angles for a given polarisation state of the incident light. This directly leads to a means to estimate the surface zenith angles (the angles between the surface normals and the viewing direction). Furthermore, Fresnel theory predicts that the surface azimuth angles (the angles of the projection of the surface normal onto the image plane) relate directly to the phase angle of the reflected light. The accuracy of these angle predictions was assessed by Saito et al. [90], using planar and hemispherical samples.

It is therefore possible to estimate surface normals from specularly (directly) reflected light. Unfortunately it turns out that both the degrees of freedom of the surface normals have a two way ambiguity. Wolff [114] attempts to combine the polarisation information from two views in order to fully constrain the orientation of a plane. Miyazaki, Kagesawa and Ikeuchi [64] also use two views to solve the ambiguities mentioned above. They show how to recover the 3D geometry of transparent objects. This is highly desirable since most other techniques for reconstruction, including SFS and stereo, face particular difficulties when faced with transparent media. Miyazaki et al. [65] have also applied Fresnel theory to infrared light in order to overcome the ambiguity in the zenith angle.

A major drawback of the methods that analyse specularly reflected light is that specularities are seldom found over entire object surfaces. For the above papers by Miyazaki, Kagesawa and Ikeuchi and by Saito et al., the target object was placed inside a spherical translucent “cocoon.” Several external light sources were used to illuminate the cocoon, which diffused the light, causing specularities to occur across the entire surface.

Several researchers have consequently investigated the use of *diffusely* reflected light in polarisation analysis. Diffuse reflection is a result of the scattering of light by rough surfaces [3, 74] and/or the light being scattered internally and re-emitted after internal scattering [115]. For the former case, the light becomes depolarised and can cause errors in shape analysis from polarisation. For this reason, most efforts in shape recovery using polarisation information focus on smooth surfaces (although we do consider slightly rough surfaces in this thesis). For smooth surfaces, Fresnel theory can be applied to light being radiated from the surface after the internal scattering mentioned above.

An early use of this was in the Wolff and Boult paper already mentioned [117]. It was shown how polarisation analysis of diffusely reflected light is particularly useful for determining the surface orientation near the limbs of objects, where the zenith angle is usually large. The process was also studied by Drbohlav and Šára [20], to recover the surface normals of a sphere. This demonstrated high potential for shape estimation using diffuse reflection, without relying on specularities. Another advantage of diffuse reflection is that there are no ambiguities in the zenith angle estimates. The two-way ambiguity in the azimuth angle remains however, and the two possibilities are different to those for specular reflection. This means that the reflection type (specular or diffuse) must be known.

One disadvantage of using diffuse reflection, compared to specular reflection, is that the polarising effects are weaker in the former case, increasing the susceptibility to noise. A second problem concerns the refractive index. The direct use of Fresnel theory for either reflection type requires an estimate of the refractive index of the reflecting medium in order to calculate the zenith angle. However, for specular reflection, the dependency is negligible, whereas for diffuse reflection, the dependency is more noticeable. Little effort has been made to estimate the refractive index from image data, although Tominaga and Tanaka [106] estimate the quantity using specular highlights under varying illumination and viewing directions.

Miyazaki et al. [67] overcome the need for the refractive index by assuming that the histogram of zenith angles matches that of a sphere. This also partly overcomes problems of roughness. Indeed, certain parameters describing the roughness of the surface are estimated by their proposed algorithm. The price of these improvements is the hefty assumption about the zenith angle histogram. Rahmann and Canterakis [85, 84] also avoid the problem of unknown refractive indices by using only the phase information of the reflected light. Their method requires two views and the phase is used to establish correspondence. In addition, both specular and diffuse reflection are considered. The weakness here, is the discarding of large amounts of information. Also, the method is yet to be tested on objects with complicated geometry.

Drbohlav and Šára [21, 22] have investigated the idea of improving photometric stereo [120] methods using polarisation. In standard photometric stereo (reviewed in Section 2.2.2) object shapes are reconstructed using greyscale images of an object placed under varying illumination positions. If only two positions are used, which are both unknown, then the surface normals are only recovered up to a regular transformation. Drbohlav and Šára reduce this ambiguity by using polarisation. This is at the expense of introducing the need for linearly polarised incident light.

A major weakness of most polarisation methods is the amount of time needed to acquire the data (the actual processing of the data is often highly efficient). For the method that uses a digital camera and linear polariser, three or more images are required with the polariser at different orientations. This limits applications to static scenes. Wolff and others have improved matters a little by developing *polarisation cameras* [116]. These devices use liquid crystals to rapidly switch the axis of the polarising filter. The disadvantage here is that the data have a greater susceptibility to noise. More recently, PLZT (Polarised Lead Zirconium Titanate) [99] has been used, which can be applied to recover all four components of the stokes vector [66].

2.1.2 Further Applications

Image-based shape recovery has been the most active area of research into polarisation vision to date. In this section we briefly review a few other applications. Wallace et al. [111] use polarisation to improve depth estimation from laser range scanners. Traditional techniques in the field encounter difficulty when scanning metallic surfaces due to inter-reflections. In the Wallace et al. paper, the problems are reduced by calculating all four Stokes parameters to differentiate direct laser reflections from inter-reflections.

Nayar, Fang and Boult [70] note that consideration of colour shifts upon surface reflection, in conjunction with polarisation analysis, can be used to separate specular and diffuse reflections. Umeyama [109] later used a different method to achieve the same goal using only polarisation.

Shibata et al. [101] used polarisation filtering as part of a reflectance function estimation technique (see Section 2.3.2). Chen and Wolff [11] use polarisation to segment images by material type (metallic or dielectric). Schechner, Narashimhan and Nayar [94] show how polarisation can enhance images taken through haze. Schechner and Karpel [93] later extend this work to marine imagery.

2.2 Shape from Shading and Stereo

This section provides an overview of two of the most extensively studied methods for shape recovery. The first method, shape-from-shading [127], attempts to estimate the geometry of a surface using variations in pixel intensities. Typically, this involves making assumptions on the surface geometry, illumination conditions and/or reflectance properties. Many algorithms recover a field of surface normals (needle map), which is then used to calculate the depth.

The second method, stereo [32], aims to determine the surface geometry using more than one image. This can broadly be divided into *geometric stereo* and *photometric stereo*. In geometric stereo, two or more camera positions are needed for a given scene and triangulation is used to determine the distance to the camera. In photometric stereo, a single viewing direction is used but several images are needed, each with different illumination conditions. The amount of literature in both SFS and stereo is huge, so we concentrate here on the most important contributions and those most relevant to this thesis.

2.2.1 Shape from Shading

Many early efforts to obtain surface structure from intensity images were aimed at recovering relief maps of the moon [87]. This research dates back at least as far as the 1960s.

Two Ph.D theses by Horn [38] and Krakauer [53] introduced SFS to computer vision in the early 1970s and research has been extensive for the past few decades [127, 24].

The traditional approach to SFS is to estimate a mapping between the measured pixel intensities and the orientation of the surface at each point. This immediately raises several questions:

1. How can the relationship between the surface orientation and the pixel brightness be determined?
2. Can the two degrees of freedom of a surface orientation be estimated from a single intensity measurement?
3. Can the shape be unambiguously determined if the surface orientations are fully constrained?
4. Is there a *unique* surface that is able to generate a particular image for unknown illumination?

The answers to these questions are all negative, making SFS is a difficult and ill-posed problem.

In answer to the first question, the majority of SFS algorithms assume that the reflecting surface is Lambertian [57]. This means that the reflecting intensity is determined only by the albedo (the ratio of reflected to incident light at normal incidence) and foreshortening (proportional to the cosine of the zenith angle). Note that this implies that image should be independent of the illumination direction(s). Worthington and Hancock [124], developed a SFS algorithm for Lambertian surfaces where the pixel brightnesses were treated as hard constraints on the surface normals. Worthington [123] later shows how albedo changes can be incorporated into the technique. Prados, Camilli and Faugeras [79] show how Lambertian SFS can be extended from orthographic projection (which most papers assume) to the perspective case.

It has long been known however, that the Lambertian assumption is seldom acceptable [118]. Ragheb and Hancock [81] attempt to overcome this problem by using some of the theoretical reflectance models discussed in Section 2.3 to "correct" images; making them appear to be Lambertian. To avoid the difficulties with estimating a global needle map altogether, Dupuis and Oliensis [23] developed an algorithm that calculates depth directly, by propagating away from the brightest points on the image.

Regarding the second question, one must either use more than one view (see Section 2.2.2) or make constraints about the surface geometry to overcome the under-constrained nature of SFS. Ikeuchi and Horn [44] enforce the constraints that (1) the reconstructed surface must produce the same brightness as the intensity image and (2) the surface is smooth and continuous. These constraints were later adopted by Brooks and Horn [7]. Shimishoni, Moses and Lindenbaum [102] have developed an algorithm which is specially constrained for symmetric objects. This overcomes the problem that SFS is under-constrained and has shown promising results for face recognition.

In a seminal paper by Frankot and Chellappa [27], the smoothness constraint is applied by enforcing integrability. That is, the original set of surface gradient estimates were mapped onto the closest set to which Schwartz's theorem can be applied. In other words, the second derivative of the surface found by differentiating first along one axis, then along the other is equivalent to the result if the order of the differentiation operations is reversed. The importance of this paper is due to the fact that it also addresses the third question above: given a field of surface orientations (normals), a global integration method is developed to recover the depth (see also Section 3.2.2). Therefore, whatever method was used to recover a needle map, the Frankot-Chellappa method can be applied to convert it to depth. The method can only be applied *accurately* to a field of *smooth* surface normals, although reasonable reconstructions can be obtained when small discontinuities are present.

In more recent work, Agrawal, Chellappa and Raskar [2] propose a related global method for recovering depth from needle maps by solving a linear system of equations. The technique is non-iterative, avoids any error propagation and, unlike the earlier Frankot and Chellappa method, does not depend upon a choice of basis functions. Another method was proposed by Simchony, Chellappa and Shao [103] where integrability is enforced by seeking a gradient field that minimises the least square error between the original field and the integrable one. Finally, Robles-Kelly and Hancock [88] used the changes in surface normal directions to estimate the sectional curvature on the surface. A graph-spectral method is then applied to locate a curvature minimising path through the field of surface normals. By traversing this path and using the estimated surface orientation, simple trigonometry is used to compute the height offset from a reference level.

Consider now the final question above. For a surface under particular illumination conditions and viewed from a fixed direction, there exists a set of transformations that simultaneously deform the surface and move the light sources such that the surface brightness remains constant from the viewer's perspective [4]. This problem is known as the bas-relief ambiguity. A simple example of such an ambiguity is the concave/convex ambiguity, where it is impossible to deduce whether an image depicts a concavity or convexity, unless information about the illumination direction(s) is to hand. Pentland [77] and Zheng and Chellappa [128] attempt to calculate the light source direction, but general and reliable algorithms to do so remain elusive.

In answering the above questions, it should be apparent that SFS is an ill-posed problem. Several researchers have therefore attempted to combine shading with other cues in order to avoid enforcing hefty geometric constraints. White and Forsyth [113] for example, incorporate texture information so that a field of unambiguous surface normals can be estimated. Prados and Faugeras [80] meanwhile, use the fact that the light intensity from the source decays as the inverse square of distance to provide an extra constraint.

2.2.2 Multi-view Methods

The principles of shape reconstruction using geometric stereo are very different to those used for SFS. The basic idea is very simple and can be described in three steps:

1. Two (or more) images of a scene are taken.
2. An algorithm is applied to determine which pixel pairs correspond to the same point in three dimensional space.

3. Triangulation is applied to calculate the distance between points in the scene and the cameras.

The difficult step is the second one and a wide range of techniques have been suggested to solve this problem (commonly referred to as the *correspondence problem*). Scharstein and Szeliski [92] and, more recently, Seitz et al. [96] have compared the performance of some the leading methods quantitatively.

In their survey paper, Brown and Hager [8] categorise both local and global methods. Local methods include block matching, where small region correspondences are established; gradient based optimisation, usually involving a least squares minimisation over small areas; and feature matching, where correspondence is acquired by measuring the similarity between points of interest. One of the most promising recent developments in geometric stereo is the *shape-invariant feature transform (SIFT)* of Lowe [60]. This is a feature matching technique with the main advantage that the points of interest are chosen to be invariant to both scale and rotation. The features used by the SIFT algorithm are also only weakly affected by variations in the viewing or illumination directions.

The most studied *global* methods for correspondence characterised by Brown and Hager are dynamic programming, where a path cost minimisation algorithm is applied to the corresponding scanlines from two images; intrinsic curves, a related method that uses a vector representation of the scanlines; and graph cut methods, which perform matching by seeking maximum flow in graphs (a comprehensive list of references can be found in [8]).

Most traditional geometric stereo algorithms encounter difficulty when presented with surfaces that are both smooth and featureless. In an attempt to overcome this problem, a few researchers have attempted to combine geometric stereo with SFS. Cryer, Tsai and Shah [18] observe that SFS is well suited to surface estimation for local patches but stereo performs better at recovering the general shape (depending on the amount of texture on the surface). They therefore proposed an algorithm that obtains low frequency information from stereo and high frequency features from SFS. Jin, Yezzi and Soatto [46] have since used shading information from multiple views to formulate stereo vision in terms of region correspondence. They developed an algorithm that uses the calculus of variations for region matching optimisation in an infinite-dimensional space.

Using a somewhat different approach, Zickler, Belhumeur and Kriegman [129] describe a stereo technique where the positions of the single light source and the camera are interchanged for the second view. A major advantage is that the surfaces considered may have arbitrary reflectance properties and may or may not include texture. In the case of textureless surfaces, where the above stereo techniques face difficulty, the technique recovers a set of surface normals. If significant texture is to hand, then the depth can be calculated directly. The main drawback of the technique is that the camera–light source interchange is often unpractical and limits applications.

Gold et al. [29] and Chui and Rangarajan [12] have researched the problem of two- and three-dimensional point matching. That is, given two related sets of points in space, they determine which points correspond between sets. The techniques developed involve estimating a set of transform parameters to align the surfaces using deterministic annealing [91] and softassign [86]. An example application of this is optical character recognition, where points on a template character must be matched to handwriting. The technique can also be used in stereo to match surface features. Cross and Hancock [17] represent surface feature points as nodes on a graph and apply the EM algorithm to perform the stereo matching.

All of the methods discussed so far require more than one camera position. Woodham [120] described an alternative scheme for textureless surfaces where the camera is kept fixed, but the light source is moved to two or more positions between successive images. In 1980 [120], he showed how the surface normals can be calculated in this way for a Lambertian surface and soon after [121] showed that a third image allows the albedo to be estimated. At least four images are necessary if more complicated reflectance functions are present and to deal with shadows and inter-reflections.

The important advantage of photometric stereo compared to geometric stereo is that, since the camera is fixed, then pixel correspondences can be directly inferred from their image location. The drawback is that, like Helmholtz stereopsis, the image acquisition procedure plays an active and time consuming role in the overall process. The images for geometric stereo by contrast can be obtained with separate cameras in an instant. Photometric stereo faces difficulty if shadows are cast on parts of the surface, but does not face problems of occlusion (which can severely affect geometric stereo in some circumstances).

Photometric stereo has attracted a steady stream of interest since its conception. As explained recently by Wu et al. [126], the large body of literature now available is wide ranging in the assumptions required: two views with Lambertian reflectance and known albedo; three views with unknown albedo; four views with more complicated reflectance properties and so on. Wu et al. describe a detailed technique using Markov random fields, which is able to deal with complex geometries, shadows, specularities, variation in light source attenuation and even, to an extent, transparencies. The method assumes light source direction estimates, but these need not be precise.

Hertzmann and Seitz [37] use a very different photometric stereo technique for materials with arbitrary reflectance properties. Their work involved placing a spherical reference object in the scene made from the same material as the target object. In addition to allowing for general reflectance properties, their work offers two other notable advances. Firstly, unlike most earlier papers on photometric stereo, the technique works for arbitrary and unknown light source directions. Secondly, their method also proposes a means to segment the image according to material type. The necessity for a reference object is the obvious weakness of this approach. Similar work, by Treuille, Hertzmann and Seitz [108], use varying camera positions.

2.2.3 Additional Shape Recovery Techniques

In this chapter, we have reviewed shape recovery techniques using polarisation, shading and stereo in some detail. There are, in fact, many other shape recovery techniques which are given the generic name *shape-from-X*. Some of these methods are closely related to SFS. Nayar, Ikeuchi and Kanade [71] for example, consider the inter-reflection problem in SFS. That is, the problem that light can be reflected several times between different surface points before being redirected towards the camera. They show how this phenomenon can actually be used to *aid* shape recovery. Similarly, Kriegman and Belhumeur [54] show how shadows, another common complication for shape-from-X algorithms, can be used to obtain geometry information.

As Koenderink [50] and later Laurentini [58] noted, one of the potentially strongest constraints on surface structure is the object's occluding contours. Of course, a reliable means to segment the target object from the rest of the image is required in order to incorporate this information. Hernández, Schmitt and Cipolla [36] have shown very accurate shape reconstruction from occluding contours by taking a succession of images of an object as it is rotated on a turntable. In related work, Wong and Cipolla [119], show how shape can be estimated in a similar way with unknown camera positions. Using occluding contours as the only shape cue limits the technique to convex surfaces.

We would finally like to mention contributions from Hwang, Clarke and Yuille [43], who use focus information for direct depth estimation; Super and Bovik [105], who devised a method for shape-from-texture; and Soatto [104], who calculates shape from motion.

2.3 Reflectance Functions

The final main section of this chapter is concerned with the reflectance properties of surfaces. More specifically, we shall consider methods to predict the distribution of light reflected from a surface for given distributions of incident illumination. This relationship is most commonly quantified using the *bidirectional reflectance distribution function* or BRDF, which has been used since the 1970s [73]. In simple terms, the BRDF is a quantity that predicts the ratio of the reflected light in a particular direction to the incident irradiance in any other direction. A more precise definition is given in Chapter 4 (Equation 4.1). The BRDF therefore has four degrees of freedom, two each for the incident illumination direction and reflection direction.

It is not always necessary however, to consider all four angles. For instance, if the reflecting surface is known to be isotropic, then only the difference between the illumination azimuth angles is important, so there are only three degrees of freedom. Alternately, it may be that only a one- or two-dimensional "slice" of the BRDF is required for a particular application. This is true for example, if the relationship between the radiance reflected towards a camera is needed as a function of the surface orientation, as in Chapter 4 of this thesis.

The BRDF is an important quantity in both graphics and vision. In graphics, it has been used for many years for image rendering. For vision, it has been used for image understanding to aid in surface reconstruction techniques such as photometric stereo, as discussed earlier.

We shall now review the relevant papers in the field. The literature can be divided into theoretical and empirical approaches. The former involves attempting to describe the interaction of light with the surface mathematically. Methods differ in terms of the inclusion of "real" physics and how plausible the surface structure is modelled. The second approach measures the reflectance properties directly using specialised equipment or by applying image processing techniques from one or more views of an object.

2.3.1 Theoretical Models

The most widely used reflectance model in computer vision is the Lambertian reflectance model [57], which was proposed long before the birth of computer vision. The Lambertian model was discussed in Section 2.2.1 since it has been used extensively for SFS. The model breaks down completely in the presence of roughness or specularities. In addition to this failure, it is not entirely accurate for shiny surfaces, even away from specularities. Despite its shortcomings, it is still used today in some applications for its simplicity. It can also be used to represent a simple diffuse term as part of a more complex model.

The famous Phong reflection model [78] was invented for computer graphics in the 1970s by constructing an equation based on heuristic observations. The model consists of additive diffuse and specular components and an extra “ambient” term. The parameters of the model can be adjusted to fine tune the relative strengths of each component and the broadness of specularities. Despite its simplicity and empirical foundation, the Phong model is surprisingly versatile and provides good agreement with experiment in many situations.

In an attempt to construct a BRDF model on more concrete physics, Wolff [115] proposed a reflectance model based on Fresnel theory. The model predicts the effects of internal scattering on the reflected radiance using results from radiative transfer theory. Fresnel theory is needed to calculate the light attenuation at the medium/air interface. All of the parameters of the model are based on well-established physical quantities and the model agrees with experiments for the diffuse component of smooth surfaces. The model breaks down for rougher surfaces and does not attempt to model specular reflection.

The reflectance properties of rough surfaces have been studied extensively for many years. One of the first efforts was by Torrance and Sparrow [107] who modelled the microscopic structure of the surface as an array of randomly oriented V-shaped grooves (often referred to as “microfacets”). The grooves have a gaussian distribution of slopes and are assumed to be perfect mirrors, each of which obey Fresnel theory. The Torrance-Sparrow model was shown to accurately predict the broadening of specularities caused by roughness. Although the original model is rather complex and unwieldy, it was later shown by Healey and Binford [34] that some of the terms could be simplified without significantly compromising the accuracy of the model. Cook and Torrance [14] later modified the Torrance-Sparrow model for computer graphics by simplifying the Fresnel term.

Oren and Nayar [75] adopt a similar approach to Torrance and Sparrow, but replace the mirror-like microfacets with Lambertian reflectors. This model accurately accounts for the diffuse component of rough surface reflection and complements the Wolff model for smooth surfaces. It is also possible to combine these two methods [118], by assuming that each microfacet follows the Wolff model. This combined model can be used for both rough and smooth surfaces, as well as the intermediate case. The resulting equations however are complicated and have not been extensively studied.

Beckmann and Spizzichino [3] use a very different approach, where the surface is modelled by a Gaussian height distribution and a Gaussian correlation function. The model, commonly referred to as the Beckmann-Kirchhoff model, uses Kirchhoff theory to estimate the distribution of light scattered from a rough surface. Kirchhoff theory essentially provides an integral equation for the light field reflected from a single scatterer. Beckmann and Spizzichino adapted Kirchhoff theory for rough surfaces and derived several simplifications applicable to particular roughness conditions. Two desirable properties of the Beckmann-Kirchhoff model are its solid basis on physical optics and its realistic description of the surface structure. Nayar, Ikeuchi and Kanade [72] present a useful comparison between this physical optics model and the geometric Torrance-Sparrow model.

The assumptions made by the Beckmann-Kirchhoff model mean that it is unsuitable for wide angle scattering. Vernold and Harvey [110] made modifications to the model based on empirical observations. This clearly meant that the highly desirable physical foundation of the model no longer prevailed. Nevertheless, the modification has been shown to give accurate agreement with experiments for a variety of rough surfaces. Indeed, Ragheb and Hancock [82] have recently shown that it gives better results than the Oren-Nayar model in some instances.

The Kirchhoff integral was also used by He et al. [33] who incorporate many physical phenomena into a reflectance model including dependencies on wavelength, the refractive index, polarisation and shadowing/masking. The result is a complicated BRDF involving separate terms for specular, directional diffuse (surface reflection) and uniform diffuse (subsurface scattering) reflection.

Hanrahan and Krueger [31] note that many surfaces consist of multiple layers and are poorly described by the above models. Examples of such surfaces include painted or varnished surfaces and some organic materials such as leaves and, importantly, skin. They therefore developed a model based on linear transport theory to account for this fact. Jensen et al. [45] also model skin (amongst other things) but in a different way. They extend the BRDF to the *bidirectional surface scattering distribution function* or BSSRDF.

This quantity adds two further degrees of freedom to the BRDF so that the incident light reflected from a particular point not only depends on the incident light at that point, but also on the incident light in the local vicinity of the point. This effectively accounts for internal scattering mechanisms. Wu and Tang [125] use the BSSRDF formulation to devise a method to separate the surface and subsurface components of diffuse reflection from an image, in addition to the specular component.

In papers by Schlick [95] and Lafortune et al. [55], an intermediate stance is taken on the use of physics in the proposed reflectance functions. While neither method attempts to model the light interaction in a plausible fashion, both take care to ensure certain physical requirements are met. These include satisfying conditions such as energy conservation and reciprocity. Both contributions were motivated by the graphics community so special care was taken to ensure that the resulting equations were simple and easy to manipulate. The Schlick paper also provides realistic renderings of anisotropic surfaces.

Despite great efforts, the search for a general BRDF (or related function) capable of modelling the wide range of materials found in everyday life remains elusive. One should bare in mind however, that if such a model did exist, then it would probably be too complicated and contain too many parameters to be of any practical use. It is therefore sufficient for most applications to accept the simplifications of the models and select the most suitable theory for the material/application being studied. There are, of course, many materials that are poorly modelled by all existing methods. Furthermore, in some vision applications it is not always clear what the material being viewed should be. In these cases, the reflectance properties must be measured experimentally or estimated from image data. We now complete our literature review with a brief survey of these methods.

2.3.2 Empirical Reflectance Function Acquisition

Most methods to experimentally recover a material BRDF aim to generate a lookup table. Given a particular incoming light direction and the direction of the viewer, the lookup table can then be used to predict the reflected radiance. As with the mathematical models discussed above, the ultimate goal for recovering the BRDF is to estimate the quantity covering the entire domain of all four degrees of freedom. Also in common with the mathematical models however, it is often only necessary to consider one, two or three of the degrees of freedom.

The most direct method to acquire a BRDF experimentally is to use a *gonioreflectometer*. That is, a device consisting of a light source, a material sample and a photometer. The photometer is then used to record the variation in reflected radiance as the various degrees of freedom in the BRDF are adjusted. A recent example of such a device is presented in a paper by Li et al. [59], who also list previous papers on gonioreflectometers.

Dana et al. [19] used similar apparatus to a traditional gonioreflectometer, but replaced the photometer with a camera. This allowed both featureless and textured surfaces to be imaged and then rendered by combining images from different illumination and viewing angles. Textured surfaces were also considered by Han and Perlin [30], who used a kaleidoscope to simultaneously image textured samples from many different angles.

BRDF acquisition using a gonioreflectometer is a very time consuming process and may necessitate $10^4 \sim 10^5$ independent measurements [59]. An alternative is to use a camera and a curved sample of known shape to make recordings. Marschner et al. [61] used two cameras as detectors (the second camera is for calibration purposes) and cylindrical and spherical objects as samples. In the case of a sphere, two of the degrees of freedom are effectively varied across a single measurement (image), while a third is varied between successive images. The fourth degree of freedom, which is associated with anisotropy, is not considered. In an earlier contribution by Ward [112], an opposite technique was used. The sample was planar but a hemispherical mirror was used to allow simultaneous analysis of two degrees of freedom. Ward also attempted to model the measured data for both isotropic and anisotropic surfaces.

Zickler et al. [130] have suggested an alternate means to reduce BRDF acquisition times. They used relatively few images and illumination/viewing angle combinations and fit the resulting sparse data sets to radial basis functions. Their method can be regarded as an interpolation between separate image recordings.

Shibata et al. [101] use the separation of reflectance types using polarisation as part of a process to recover reflectance functions. They fit the data to the Torrance-Sparrow reflection model to obtain all its parameters. The method is not entirely image-based however, as it requires surface reconstruction using a range scanner to obtain the surface orientation at densely sampled surface points.

Robles-Kelly and Hancock [89] have recently proposed a purely image-based method for reflectance function estimation from a single view. Their idea was to determine a mapping from a single image onto a Gauss sphere using the cumulative distribution of intensity gradients. The technique is general in terms of reflectance properties since it requires no model parameter fitting. Another advantage is that no calibration is needed. The main weakness is that the method can only be applied under conditions where the light source and viewer directions are identical.

2.4 Conclusions

It should be clear from this literature review that (1) extensive research in shape and reflection analysis has been conducted during the last few decades, and (2) there are many shortcomings of the current state of the art. We conclude this chapter with a summary of the above literature, and state the position of this thesis in the context of the broad field.

One of the earliest classes of shape reconstruction methods to be conceived was SFS. Many of these methods work best for constant-albedo Lambertian surfaces with a single known point-source illumination. More recent work in the field has relaxed some of these restrictions to an extent by estimating the illumination distribution or reflectance function. Most researchers have accepted that single-view SFS is an ill-posed problem and have devised algorithms specific to particular domains such as face reconstruction. Photometric stereo methods are more constrained and, depending on the precise method and number of light sources, are more general with respect to the surface reflectance properties. The need for multiple light sources and images however, makes the method unsuitable for many applications.

Geometric stereo is, in many ways, opposite in its strengths compared to SFS and photometric stereo. While the latter methods often break down in the presence of texture or complex albedo variations, geometric stereo *relies* on these features to solve the correspondence problem. On the other hand, geometric stereo breaks down for featureless surfaces since there is no means to establish correspondence. Unlike SFS and photometric stereo which typically calculate a depth map from estimated surface normals, geometric stereo uses a more direct depth acquisition based on triangulation.

Polarisation methods have proven to be of great use in providing additional constraints on surface normals. The azimuth angles of the normals are particularly accurately estimated using polarisation data, albeit ambiguously and subject to the specular/diffuse reflection type. The zenith angles are accurately but ambiguously estimated for specular reflection and crudely but unambiguously estimated for diffuse reflection. Collectively, these constraints are strong and come at a low computational cost without the need for any light source calibration. The main weaknesses are the need for an estimate of the material refractive index and the gradual breakdown of the technique for rougher surfaces.

In the field of reflectance function modelling, success has been achieved from three approaches:

1. Using physics to model the interaction of light (as an electromagnetic wave) with a surface
2. Modelling the surface as an array of microfacets and applying geometrical optics
3. Fitting measured data to ad hoc mathematical functions.

Despite the undesirable ad hoc nature of the last of these approaches, results are simple and relatively general. The other two approaches are more accurate but tend to be restricted to certain cases (smooth, rough, layered surfaces etc.). In a completely different approach, measured reflectance data for various materials has been stored in lookup tables. This gives best results when the time required to obtain the BRDF is not important. Finally, a small number of techniques have been proposed to estimate BRDFs (or portions of BRDFs) from a single image. These methods are of use in supplementing other computer vision applications such as SFS.

This thesis addresses some of the open questions raised by the above literature. Chapter 3 is largely concerned with the underlying techniques used in polarisation vision. The chapter describes a simple shape recovery method and an analysis of the results. An important novel aspect of the contribution is that the analysis is more extensive than most previous efforts in polarisation vision and covers a wider range of materials and geometries. The contribution of Chapter 4 is to advance current techniques on image-based reflectance functions. Relatively little work has been conducted with this goal from few ($\lesssim 3$) images. We show how polarisation and robust statistics can be used to estimate portions of an arbitrary BRDF under known, but arbitrary, light source directions. Polarisation has not been used for the this task until now, except with the aid of a range scanner. Finally, the contribution of Chapter 5 is to use polarisation to address a major weakness of geometric stereo: that most existing methods face difficulty when seeking correspondence for featureless surfaces. Previous efforts to address this weakness involved dynamic programming or SFS methods.

Chapter 3

Polarisation Vision

The aim of this chapter is to assess the practical aspects of polarisation vision and to evaluate the accuracy of the technique for various materials. The main emphasis of this chapter is on the underlying theory of polarisation and its experimental exploitation in computer vision using minimal computation. The novel contributions of the chapter are a simple single-view shape reconstruction algorithm that uses Fresnel theory and polarisation information, and an uncertainty analysis of the recovered surface orientation estimates. More sophisticated novel algorithms will be developed in Chapters 4 and 5.

The physics that underpins most of the work presented in the thesis (and other research in polarisation vision) is based on Fresnel theory [35], which predicts the reflected and transmitted intensities of a ray of light as it impinges upon a surface. The most important results of this theory are the Fresnel coefficients. These give the ratio of the reflected and transmitted light intensities to the incident light intensity as functions of the angle of incidence and the polarisation state of the light. The fact that the reflected light depends on both these factors makes polarisation vision possible. The dependency of the reflected light intensity on the polarisation state results in initially unpolarised light becoming partially polarised by the reflection. The dependency on the angle of incidence then means that polarisation measurements can be used to deduce constraints on the reflection angles and, therefore, the surface orientation. The polarisation state of the reflected light can be analysed using a standard linear polariser and a digital camera.

There is also a third dependency on the polarisation state of the reflected light: the refractive index of the reflecting material. We do not currently make any attempt to estimate the refractive index from the acquired *images*, but use typical values throughout the thesis. Nevertheless, the success of our algorithms prove that precise knowledge of the quantity is not essential. We have included a short aside on non-image based refractive index estimation towards the end of this chapter.

Polarisation vision has been used previously to extract information from both diffusely reflected light and specularly reflected light. As explained previously in Section 2.1, and later in Section 3.1.5, there are advantages and disadvantages of exploiting each. For the bulk of this thesis, we focus on diffuse reflection, mainly because it can be found over entire surfaces, while specular reflections cannot. In the diffuse case, an estimate of the surface orientation can be obtained from raw polarisation data, up to a single two-way ambiguity.

The main contributions of this chapter are presented in Sections 3.2 and 3.3 and include a series of experiments. The aim is to assess the ability to estimate surface normals experimentally, using results from Fresnel theory. Specifically, experiments with the following goals were conducted:

1. Verify that the theory and the adopted data acquisition technique is adequate for shape analysis.
2. Describe a simple shape recovery algorithm using polarisation information, which preludes to the more sophisticated methods presented in later chapters.
3. Assess the limitations of the above algorithm and polarisation vision in general.
4. Evaluate the accuracy of polarisation vision for various materials.
5. Make estimates of the refractive indices of the materials considered.

Chapters 4 and 5 enhance the shape recovery technique using two views and by incorporating shading information.

The remainder of this chapter is organised as follows. We start by presenting an overview of Fresnel theory and show how this leads to measurable quantities that can be utilised in computer vision. The relevant relationships between these quantities and the orientation of the reflecting surface are then derived. The basic experimental arrangement is described, which is used for most of this thesis. Details of the single-view reconstruction technique is presented in Section 3.2. Section 3.3 presents the details of the experiments listed above.

3.1 Fresnel Theory and Polarisation Vision

The main purpose of this section is to present the relevant equations from Fresnel theory and show how they can be used in polarisation vision. Before moving on to the mathematical description, it is worth considering the mechanics of an interaction between light and a reflecting medium qualitatively. The description we give [35] is not entirely complete as it is classical in nature. However, this thesis concentrates specifically on the highly abundant dielectric surfaces found in everyday life (the majority of non-metals) and the resulting equations are sufficient for our needs.

Consider an electromagnetic wave impinging a smooth surface, as in Figure 3.1. A fraction of the energy of this wave penetrates the surface and is refracted in the process. The electric field of the wave inside the medium causes the electrons of the reflecting medium near the surface to vibrate, forming dipoles. These dipoles are aligned with the electric field of the refracted electromagnetic wave. The vibrating electrons then re-radiate, generating the reflected ray. It turns out that the net reflection is only in a single direction due to destructive wave interference in all other directions. The strength of the reflection depends upon the components of the electron oscillations that are orthogonal to the reflected ray.

The incident light can always be resolved into components parallel to, and perpendicular to, the plane of incidence (the plane that contains the incident, transmitted and reflected rays). Consider the component of the electric field that is polarised parallel to

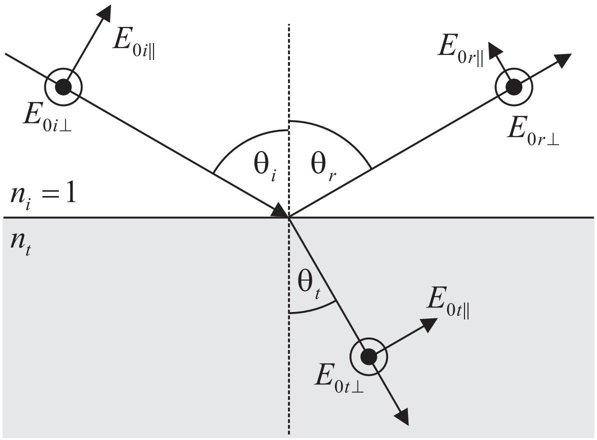

Figure 3.1: Reflection of an electromagnetic wave, with electric field directions and relative amplitudes indicated. Within the medium, the electrons vibrate parallel to the electric field.

the plane of incidence. The induced dipoles are not generally orthogonal to the reflected ray, as Figure 3.1 illustrates. This results in some attenuation due to foreshortening. However, for the component polarised perpendicular to the plane of incidence, the dipoles are always orthogonal to the reflected ray so the attenuation is not present. We can conclude therefore, that the amount of light reflected will generally be different for each of the two orthogonal components.

As a consequence of this, if the incident light is unpolarised, then the reflected light will be *partially polarised*, i.e. will consist of a superposition of an unpolarised component and a completely polarised component. In fact, it is possible to arrange the incident light such that the electron oscillations due to the wave component parallel to the plane of incidence is exactly parallel to the reflected ray. In this case, the parallel component of the reflected ray vanishes. The light is thus completely polarised perpendicular to the plane of incidence. The angle of incidence corresponding to this phenomenon is known as the *Brewster angle*.

As Figure 3.1 suggests, the difference in attenuation between the two orthogonal components of the light depends upon the angle of incidence. We now introduce the equations that can predict the relative attenuations for purely specular reflection. Later, we show how this description can be extended, by considering internal scattering, to account for the partial polarisation of diffusely reflected light.

3.1.1 Fresnel Coefficients

The Fresnel equations give the ratios of the reflected wave amplitude to the incident wave amplitude for incident light components that are polarised perpendicular to, or parallel to, the plane of incidence [35, 6]. For each component, the ratios depend upon the angle of incidence and the refractive index of the reflecting medium.

For the geometry of Figure 3.1, where $E_{0i\perp}$ and $E_{0r\perp}$ are the amplitudes of the incident and the reflected waves respectively, the amplitude reflection coefficient for the component of the light polarised perpendicular to the plane of incidence, $r_\perp$, at a boundary between two media, is given by

$$r_\perp\left(n_i, n_t, \mu_i, \mu_t, \theta_i\right) \equiv \frac{E_{0r\perp}}{E_{0i\perp}} = \frac{\frac{n_i}{\mu_i}\cos\theta_i - \frac{n_t}{\mu_t}\cos\theta_t}{\frac{n_i}{\mu_i}\cos\theta_i + \frac{n_t}{\mu_t}\cos\theta_t} \tag{3.1}$$

The quantities n_i and n_t are the refractive indices of the first and second media, μ_i and μ_t are the magnetic probabilities of the two media, and the angles θ_i and θ_t for the incident and transmitted light are defined in Figure 3.1.

Equation 3.1 is derived by applying the condition at the interface that the tangential components of the electric and magnetic fields are continuous. This is achieved using the plane wave solution [35, §2.7] to the famous wave equation. A similar equation can be derived for the component of light that is polarised parallel to the plane of incidence:

$$r_\parallel\left(n_i, n_t, \mu_i, \mu_t, \theta_i\right) \equiv \frac{E_{0r\parallel}}{E_{0i\parallel}} = \frac{\frac{n_t}{\mu_t}\cos\theta_i - \frac{n_i}{\mu_i}\cos\theta_t}{\frac{n_i}{\mu_i}\cos\theta_t + \frac{n_t}{\mu_t}\cos\theta_i} \tag{3.2}$$

In computer vision however, we are not usually interested in ferromagnetic materials so $\mu_i = \mu_t = \mu_0$, the permeability of free space. Therefore, Equations 3.1 and 3.2 reduce to

$$r_\perp\left(n_i, n_t, \theta_i\right) \equiv \frac{E_{0r\perp}}{E_{0i\perp}} = \frac{n_i\cos\theta_i - n_t\cos\theta_t}{n_i\cos\theta_i + n_t\cos\theta_t} \tag{3.3}$$

$$r_\parallel\left(n_i, n_t, \theta_i\right) \equiv \frac{E_{0r\parallel}}{E_{0i\parallel}} = \frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_t\cos\theta_i + n_i\cos\theta_t} \tag{3.4}$$

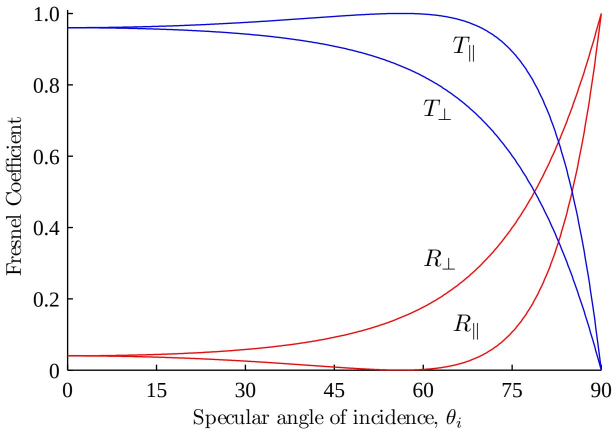

Figure 3.2: Reflection and transmission coefficients for a typical dielectric ($n = 1.5$).

The angle θ_t can be obtained from the well-known Snell's Law:

$$n_i \sin\theta_i = n_t \sin\theta_t \tag{3.5}$$

If the reflecting medium is placed in air, then $n_i \approx 1$.

Generally, it is not the amplitude of the wave that is measured by the detector, but the intensity, which is proportional to the square of the amplitude. With this in mind, it is possible to show that the *intensity coefficients*, which relate the reflected power to the incident power, are $R_{\perp} = r_{\perp}^2$ and $R_{\|} = r_{\|}^2$.

Figure 3.2 shows the Fresnel intensity coefficients for a typical dielectric as a function of the angle of the incidence, θ_i. Both the reflection and the transmission coefficients are shown, where the latter refers to the ratios of transmitted to incident powers and are simply

$$\begin{aligned} T_{\perp} &= 1 - R_{\perp} \\ T_{\|} &= 1 - R_{\|} \end{aligned} \tag{3.6}$$

The point where $R_{\|} = 0$ corresponds to the Brewster angle and is typically about $60°$. Figure 3.2 confirms what was implied by the electron vibration argument above, that the fraction of reflected light polarised perpendicular to the plane of incidence is greater than that polarised parallel to it.

3.1.2 The Polarisation Image

We now consider how the above theory can be used in computer vision. In physics terminology, the reflected light intensity is commonly referred to as the radiance, L_r. It is defined as the power per unit area per unit solid angle. Our image capture system is to measure this quantity using a standard digital camera. Consider the radiance, L_r, reflected from a surface point and arriving at the camera. If the target object subtends a small angle with respect to the optical axis of the camera, then the irradiance on the sensing element in the camera (CCD chip) is [26, §4.2]

$$L_{\text{CCD}} = \frac{\pi}{4}\left(\frac{D_L}{D_\Pi}\right)^2 L_r \propto L_r \;\; (\propto I) \tag{3.7}$$

where D_L is the lens diameter, D_Π is the distance between the lens and the image plane, and I is the measured pixel brightness. For the proportionality relation to I, we are assuming that the camera response function is linear. In fact, this is not quite true and it may be possible to improve future work by taking this into account. For this thesis however we assume linearity, as in most previous work. Equation 3.7 shows that processing greylevel intensities provides a good representation of the physical quantity of radiance, up to an unknown calibration constant.

For all our measurements, a polariser was placed in front of a camera and rotated to different angles. Let θ_{pol} be the polariser angle relative to an arbitrary reference direction (vertical). The measured pixel brightness varies with the polariser angle according to the *transmitted radiance sinusoid*:

$$I\left(\theta_{\text{pol}}, \phi\right) = \frac{I_{\text{max}} + I_{\text{min}}}{2} + \frac{I_{\text{max}} - I_{\text{min}}}{2} \cos\left(2\theta_{\text{pol}} - 2\phi\right) \tag{3.8}$$

Here, I_{min} and I_{max} are the minimum and maximum observed pixel brightnesses as the polariser is rotated and ϕ is the *phase angle*, or the angle of the linearly polarised component of the reflected light. The maximum pixel brightness is therefore observed when the polariser is aligned with the angle of polarisation, i.e. $I\left(\theta_{\text{pol}} = \phi\right) = I_{\text{max}}$, and the

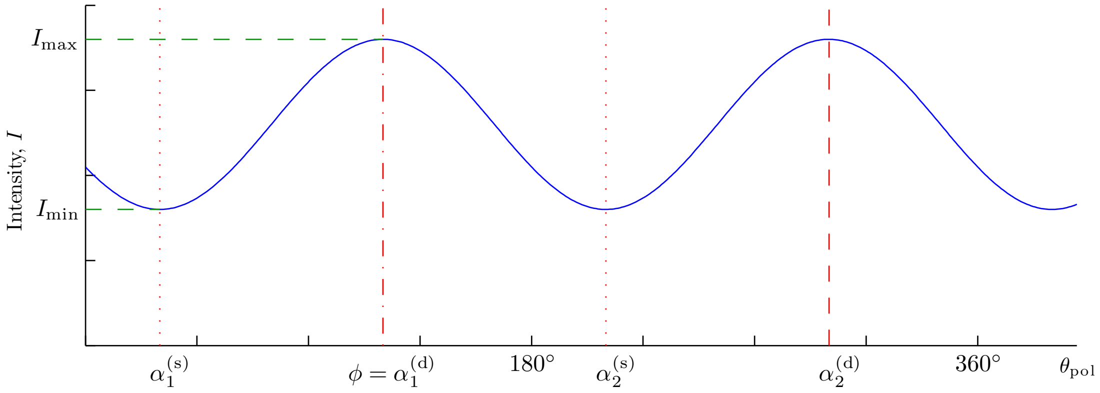

Figure 3.3: Transmitted radiance sinusoid. $\alpha_1^{(s)}$ and $\alpha_2^{(s)}$ are the two possible surface azimuth angles for a given phase angle ϕ, assuming that the reflection is specular. For diffuse reflection, the possible azimuth angles are $\alpha_1^{(d)}$ and $\alpha_2^{(d)}$.

minimum is observed when the polariser is rotated by a further $90°$ in either direction, i.e. $I\left(\theta_{\text{pol}} = \phi \pm 90°\right) = I_{\min}$. Figure 3.3 illustrates this relationship.

The *degree of polarisation* or *partial polarisation*, which is frequently used in polarisation vision, is defined to be

$$\rho = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{3.9}$$

This quantity will be used extensively throughout the thesis for the calculations of the zenith angles.

We are now in a position to introduce the *polarisation image*, which is used by all the methods developed for this thesis. This is a three component image that encodes the full set of polarisation data for a given object or scene. The first component is simply the *intensity image*, i.e. the standard greyscale image that would be obtained using a normal camera without a polariser.

The second component of the polarisation image is the *phase image* (defined as ϕ in Equation 3.8). The phase is the angle of the linearly polarised component of the reflected light, or equivalently, the angle of the polariser that permits maximum transmission, as in Figure 3.3. Note that polarisers cannot distinguish between two angles separated by $180°$, so the range the phase is $[0, 180°)$.

The final component of the polarisation image is the *degree of polarisation image*, as determined by Equation 3.9. The phase and degree of polarisation components of the

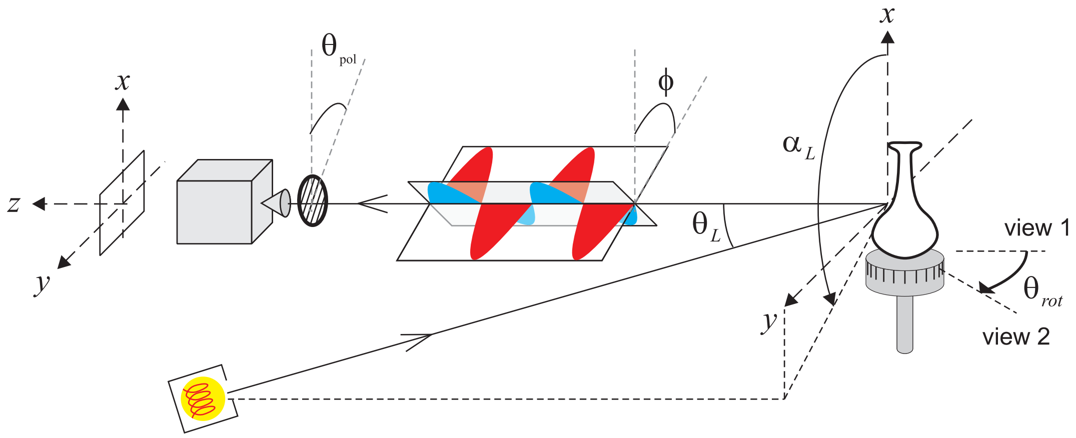

Figure 3.4: Experimental set-up. The sine wave of greater amplitude illustrates the component of the electromagnetic wave in the direction of polarisation, i.e. is at an angle ϕ to the vertical. The other sine wave is the component 90° to this. The axes on the left define co-ordinate system.

polarisation image are usually found by taking images with the polariser at three or more angles and fitting the resulting pixel brightnesses to Equation 3.8.

Figure 3.4 is a schematic of the polarisation image acquisition system. The co-ordinate system is defined to the left of the figure and is used for the remainder of the thesis. We assume that the projection is orthographic. The light source angles are shown and are defined relative to the vertical x-axis (which may be regarded as the illumination azimuth) and the camera direction (the illumination zenith). For this chapter, our experiments were conducted with the light source placed beside the camera, so that $\theta_L \approx 0$. The rotatable table is not used in this chapter, but will be referred to in Chapter 5. Although Figure 3.4 shows the two components of the reflected light to be in phase, this is not generally true. However, since only the *radiance* is needed for shape recovery this does not complicate matters.

3.1.3 Shape from Specular Polarisation

Following the dichromatic reflectance model [98], the reflected light is a superposition of specular and diffuse components. Specularities result from direct surface reflections while diffuse reflection can be due to subsurface scattering or the diffusion of light by rough surfaces [118]. Both reflection types can be used in polarisation vision. The purpose of this section is to show how to estimate the two degrees of freedom that define surface orientation at each point using specular reflection. The degrees are freedom are the zenith and azimuth angles of the surface normals. The surface zenith, θ, is defined as the angle between the surface normal and the viewing direction. The surface azimuth, α is defined as the angle of the projection of the normal onto the image plane relative to the x-axis.

Figure 3.5 illustrates some of the angles and normal vectors used in the thesis. Two unit vectors are shown since polarisation data only allows for an ambiguous estimate, as explained below. We shall see that, if the reflection is specular, then polarisation analysis yields two candidates for the azimuth angle, $\alpha_1^{(s)}$ and $\alpha_2^{(s)}$, where $\alpha_2^{(s)} = \alpha_1^{(s)} + 180°$. In the case of diffuse reflection, there are also two candidates, $\alpha_1^{(d)}$ and $\alpha_2^{(d)}$, which are shifted by $90°$ from those for the specular case. Since we will be concentrating on diffuse reflection only after this section, we will also be referring to the azimuth angle candidates for diffuse reflection simply by α_1 and α_2.

As Figure 3.2 shows, the reflected light is attenuated to a greater extent if it is polarised parallel to the plane of incidence. This means that the polariser allows greatest transmission when oriented at an angle $90°$ from the azimuth angle of the surface, α. We therefore have two azimuth angle candidates, which are simply given by

$$\alpha_1^{(s)} = \phi - 90° \qquad \alpha_2^{(s)} = \phi + 90° \tag{3.10}$$

Possible methods for dealing with this ambiguity include initially directing the surface normal estimates away from the object at the occluding boundary [67], and propagating into the object (see Section 3.2), tracing level curves [83] or using several views [64] (see also Chapter 5). If the light source and viewing directions are different, then the consistent viewpoint constraint [22] can be used to recover information from specularities or photometric stereo methods can be applied (Section 4.3).

The surface zenith angle can be computed by considering the degree of polarisation (Equation 3.9). Assume that the incident light is unpolarised with radiance L_i. The components parallel to, and perpendicular to, the plane of incidence are then both $L_i/2$. The

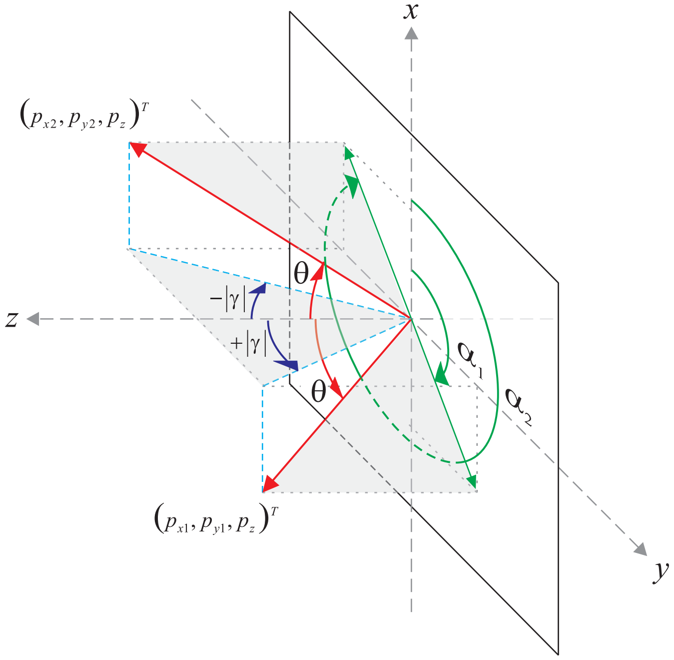

Figure 3.5: Definition of angles and the two possible surface normals. γ is used in Chapter 5. p_{x_1} and p_{y_1} can be calculated using Equation 3.16 with $\alpha = \alpha_1$. For p_{x_2} and p_{y_2}, then $\alpha = \alpha_2$ is used.

magnitude of the reflected radiance is then

$$L_r = \frac{\left(R_\perp(n, \theta_i) + R_\parallel(n, \theta_i)\right) L_i}{2} \tag{3.11}$$

The minimum and maximum radiances transmitted through the polariser are

$$L_{\text{min}} \propto \frac{R_\parallel(n, \theta_i)\, L_i}{2} \qquad L_{\text{max}} \propto \frac{R_\perp(n, \theta_i)\, L_i}{2} \tag{3.12}$$

respectively. We use proportionality instead of equality here, to account for absorption by imperfect polarisers. However, since the radiance only varies by a small amount for all polariser angles, the two proportionality constants are almost exactly equal.

Eliminating L_i from Equations 3.11 and 3.12 and writing in terms of pixel brightnesses, we have

$$I_{\text{min}} = \frac{R_\parallel(n, \theta_i)}{R_\perp(n, \theta_i) + R_\parallel(n, \theta_i)} I_r \qquad I_{\text{max}} = \frac{R_\perp(n, \theta_i)}{R_\perp(n, \theta_i) + R_\parallel(n, \theta_i)} I_r \tag{3.13}$$

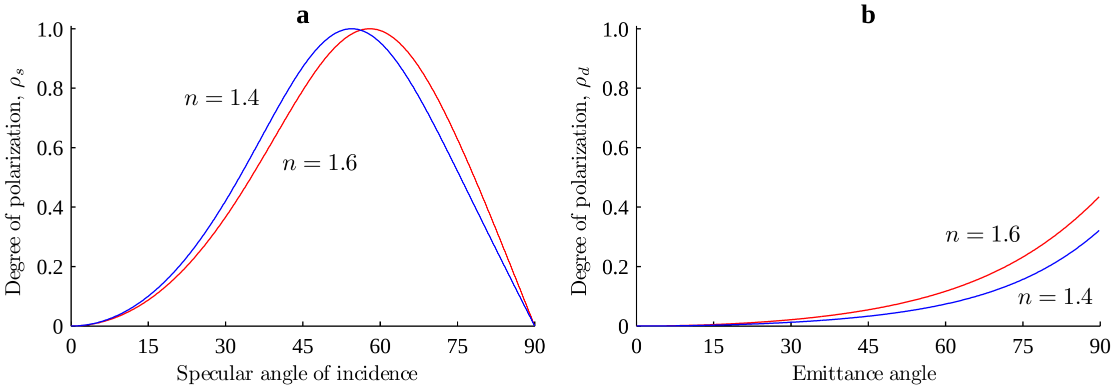

Figure 3.6: Degree of polarisation for (a) specular and (b) diffuse reflection for two different refractive indices. Many opaque dielectrics have refractive indices between these two values.

where $I_r \propto L_r$ is the magnitude of the specularity in the image. Note that the proportionality constants in Equation 3.12, have been absorbed into I_r.

Substituting Equation 3.13 into Equation 3.9 gives the degree of specular polarisation[1] in terms of the Fresnel coefficients:

$$\rho_s\left(n,\theta\right) = \frac{R_\perp(n,\theta_i) - R_\|(n,\theta_i)}{R_\perp(n,\theta_i) + R_\|(n,\theta_i)} \tag{3.14}$$

Using Equation 3.14 with Equations 3.3 and 3.4 gives ρ_s in terms of the refractive index, n, and the zenith angle, θ:

$$\rho_s\left(n,\theta\right) = \frac{2\sin^2\theta\cos\theta\sqrt{n^2-\sin^2\theta}}{n^2-\sin^2\theta-n^2\sin^2\theta+2\sin^4\theta} \tag{3.15}$$

This equation has two real solutions for θ, which are shown in Figure 3.6a. The presence of two solutions means that there is a second ambiguity that must be solved, in addition to the azimuth angle problem. This is not our concern however, since we shall see that there is no zenith angle ambiguity for diffuse reflection. Note that at the Brewster angle for specular reflection $\rho_s = 1$ since the light is totally polarised. In practice ρ_s will be a little less than 1 due to a small, but finite, diffuse component of reflection.

The refractive index is not generally known, but for most dielectrics falls between 1.4

---

[1] We shall refer to polarisation caused by specular reflection as "specular polarisation," and that caused by diffuse reflection as "diffuse polarisation." ρ_s and ρ_d refer to the expected degrees of polarisation for purely specular and purely diffuse reflection respectively. ρ refers to the measured value.

and 1.6 and the dependence of ρ_s on n is weak. Thus, with a known value or estimate of n, and with ρ_s measured using Equation 3.9, the zenith angle, θ, can be determined up to the ambiguity. The components of the surface normal vectors can then be obtained using

$$\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} \cos\alpha\sin\theta \\ \sin\alpha\sin\theta \\ \cos\theta \end{pmatrix} \tag{3.16}$$

where α takes values of either $\alpha_1^{(s)}$ or $\alpha_2^{(s)}$ (see Figure 3.3).

3.1.4 Shape from Diffuse Polarisation

Diffuse polarisation is a result of the following process [117, 118]. A portion of the incident light penetrates the surface and is refracted and partially polarised in the process, as predicted by the Fresnel equations. Due to the random nature of internal scattering however, the light becomes depolarised. Considered another way, the electron dipole orientations are randomised. Some of the light is then refracted back into the air and is, once again, refracted and partially polarised.

When light approaches the surface-air interface from within the medium after scattering, as shown in Figure 3.7, a similar process to that discussed earlier takes place. The Fresnel equations can be applied, but with n_i and n_t in Equations 3.3 and 3.4 interchanged. If the internal angle of incidence is above a critical angle ($\arcsin 1/n$), then total internal reflection occurs. Otherwise, Snell's Law (Equation 3.5) can be used to find the angle of emittance for any given angle of internal incidence. The Fresnel *transmission* coefficients can then be calculated for a given *emittance* angle. Figure 3.8 shows the result of this calculation for a typical dielectric with an additional factor of $1/n$ introduced due to a difference in wave impedance [13, §2.3]. This additional factor cancels in the next step so is not important. It should be pointed out that diffuse reflection also results from multiple scattering from microfacets, to which this theory clearly cannot be applied, although this extra contribution is negligible for smooth surfaces.

The azimuth angle of the normal can then be determined using the same method as

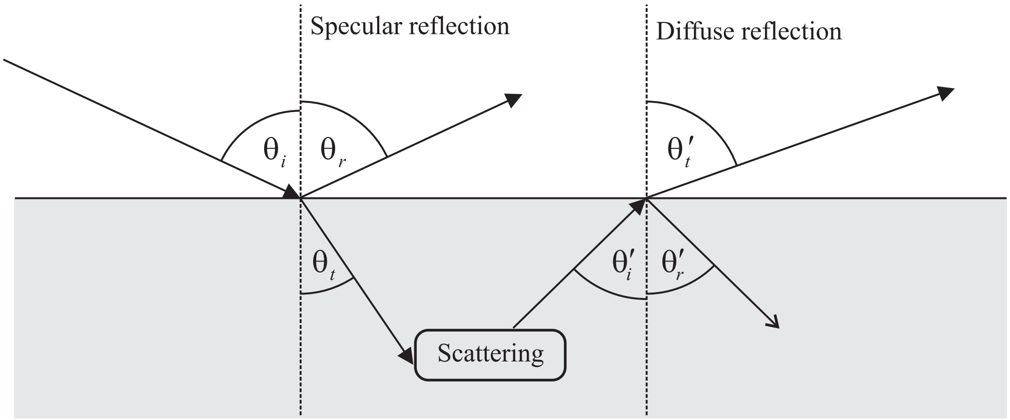

Figure 3.7: Formation of specular and diffuse reflections.

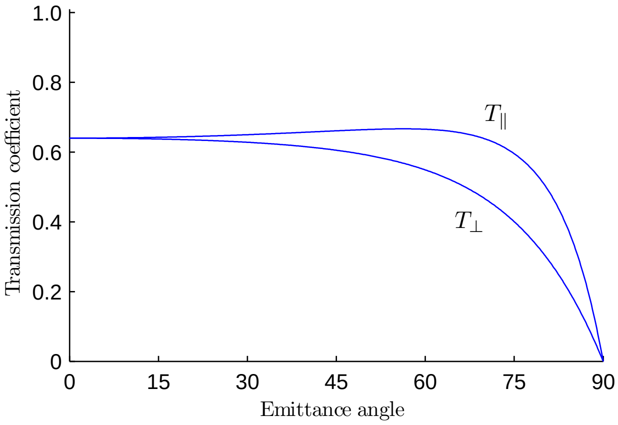

Figure 3.8: Fresnel transmission coefficients for the diffuse reflection in Figure 3.7 ($n = 1.5$).

that used for specular reflection, except that a phase shift of 90° is necessary. The need for a phase shift is illustrated by Figure 3.8, which shows that light polarised parallel to the plane of incidence has the highest transmission coefficient and so greater intensity is observed with a polariser at that orientation. This is in contrast to specular reflection (Figure 3.2). The azimuth angle candidates are therefore

$$\alpha_1^{(\mathrm{d})} = \phi \qquad \alpha_2^{(\mathrm{d})} = \phi + 180^\circ \tag{3.17}$$

Using a similar approach to that used for specular reflection, we can write expressions for the maximum and minimum pixel brightness due to diffuse reflection in terms of Fresnel transmission coefficients:

$$I_{\min} = \frac{T_\perp(1/n, \theta_i')}{T_\perp(1/n, \theta_i') + T_\parallel(1/n, \theta_i')} I_r \qquad I_{\max} = \frac{T_\parallel(1/n, \theta_i')}{T_\perp(1/n, \theta_i') + T_\parallel(1/n, \theta_i')} I_r \tag{3.18}$$

where the angles are defined in Figure 3.7 and $1/n$ is used in place of n, since light is being transmitted from the dense medium into air. Substituting into Equation 3.9, the degree of diffuse polarisation is then

$$\begin{aligned} \rho_d &= \frac{T_\parallel(1/n, \theta_i') - T_\perp(1/n, \theta_i')}{T_\parallel(1/n, \theta_i') + T_\perp(1/n, \theta_i')} && (3.19) \\ &= \frac{R_\perp(1/n, \theta_i') - R_\parallel(1/n, \theta_i')}{2 - R_\perp(1/n, \theta_i') - R_\parallel(1/n, \theta_i')} && (3.20) \end{aligned}$$

Snell's Law (Equation 3.5) can be used to interchange between the internal angle of incidence, θ_i', and the more useful angle of emittance, θ_t'. When the surface is viewed from this angle of emittance, then θ_t' is equivalent to the zenith angle, θ. The relevant Fresnel equations (3.3 and 3.4) can be substituted into Equation 3.20 to obtain ρ_d in terms of the refractive index and the surface zenith angle. The resulting equation for the degree of diffuse polarisation is:

$$\rho_d = \frac{(n - 1/n)^2 \sin^2\theta}{2 + 2n^2 - (n + 1/n)^2 \sin^2\theta + 4\cos\theta\sqrt{n^2 - \sin^2\theta}} \tag{3.21}$$

The dependence of the diffuse polarisation ρ_d on the zenith angle θ is shown in Figure 3.6b. The surface normals can then be calculated using Equation 3.16, where α is either $\alpha_1^{(d)} = \phi$ or $\alpha_2^{(d)} = \phi \pm 180°$ (Figure 3.3).

We therefore have a means to estimate the zenith angle and (ambiguously) the azimuth angle for diffuse reflection using a polarisation image and Fresnel theory. Before describing our experimental methods and results in detail, where we concentrate exclusively on diffuse reflection, we summarise the key differences in the polarising properties of the two reflection types.

3.1.5 Discussion

There are several advantages with the exploitation of diffuse reflection compared to specular reflection. Firstly, diffuse reflection occurs over the entire object surface. Secondly, the relationship between the degree of polarisation and the zenith angle is one-to-one for diffuse reflection but one-to-two for specular reflection. Finally, if we assume that the scattered light becomes completely depolarised after surface penetration, then the polarisation state of incident light can be arbitrary for diffuse reflection, but must be unpolarised for specular reflection.

The main difficulties with using diffuse reflection are as follows. Firstly, the data is generally noisier for diffuse reflection due to the weaker polarizing effects, as Figure 3.6 shows. Secondly, the dependence upon the refractive index is stronger for diffuse reflection. Finally, inter-reflections between points are mainly specular and cause erroneous orientation estimates if the theory for diffuse reflection is applied.

Both diffuse and specular reflections are affected by roughness. At the microscopic scale, a rough surface can be regarded as a set of microfacets (Section 2.3.1) with random orientation. Due to external scattering between these facets (and after subsurface scattering followed by re-emission), light impinging on a single element of the camera CCD chip effectively has a distribution of polarisation angles. The degree of polarisation is therefore observed to be reduced, resulting in an underestimate of the zenith angle. The azimuth angle estimate remains accurate provided that the mean facet orientation matches the local mean surface orientation, i.e. so long as the surface is isotropic.

3.2 A Method for Single View Surface Reconstruction

This section describes a basic method for shape estimation using polarisation information from diffuse reflection. We also discuss some of the practical matters regarding polarisation image acquisition. The method will then be used to reconstruct a selection of objects in Section 3.3. The algorithm is computationally efficient and easy to implement. Although it is unsuitable for precision applications and objects with complicated geometry, results in this chapter show good reconstructions for simple surfaces. The limitations of the technique are discussed afterwards, and are the motivations for subsequent chapters where more accurate reconstruction techniques are developed.

3.2.1 Method Overview

The method can be summarised as follows:

1. Acquire greyscale images of the target object using the arrangement shown in Figure 3.4 with three or more polariser orientations. Only a single view is required for this chapter.
2. Fit the pixel brightnesses at each point to the transmitted radiance sinusoid (Equation 3.8) to obtain the phase and the degree of polarisation images.
3. Calculate the ambiguous surface normal estimates using Equations 3.16, 3.17 and 3.21.
4. Disambiguate the azimuth angles.
5. Recover surface height from the estimated normals.

The raw greyscale images can be taken with the polariser rotated to just three different orientations, $\theta_{\text{pol}} = 0^\circ$, 45° and 90°, to give intensities I_0, I_{45} and I_{90} respectively.

We then employ the following equations [116] (as used by the liquid crystal polarisation cameras mentioned in Section 2.1.1), to determine the intensity I, phase ϕ, and degree of polarisation ρ, at each pixel:

$$I = I_0 + I_{90} \tag{3.22}$$

$$\phi = \begin{cases} \frac{1}{2}\arctan\left(\frac{I_0+I_{90}-2I_{45}}{I_{90}-I_0}\right) & \text{if} \quad I_{90} < I_0 < I_{45} \\ \qquad \text{"} \qquad +180^\circ & \text{if} \quad I_{90} < I_0 \quad \text{and} \quad I_{45} < I_0 \\ \qquad \text{"} \qquad +90^\circ & \text{otherwise} \end{cases} \tag{3.23}$$

$$\rho = \frac{I_{90} - I_0}{(I_{90} + I_0)\cos 2\phi} \tag{3.24}$$

All results presented in this thesis use these equations to estimate the phase and the degree of polarisation unless otherwise stated. We have we also applied the Levenberg-Marquardt non-linear curve fitting algorithm to recover the polarisation image components from images taken with the polariser oriented at 10° intervals. This was obviously more time consuming but less affected by noise. To reach a compromise between these options, our results are based on the average of four frames for each I_0, I_{45} and I_{90} image.

A single source of illumination was used for this chapter. This was placed beside the camera, although there is nothing in the theory to suggest that this must be the case. Indeed, in Chapter 4 we show results for different light source directions. The walls of the room and the table on which the objects lay were black so there were negligible inter-reflections from the environment. The camera used was a Nikon D70 digital SLR with a fixed aperture of $f5.6$, and exposure of 0.25s.

The degree of polarisation is converted to zenith angle by numerically solving Equation 3.21. It is assumed throughout this thesis that the refractive index of the reflecting medium is 1.4 except in the uncertainty analysis (Section 3.3.3). Next, Equation 3.16 was used to calculate the Cartesian components of the surface normal vectors, at this stage assuming that the azimuth angle $\alpha = \alpha_1^{(\mathrm{d})}$ (see Figure 3.3).

The simple disambiguation routine used in this chapter works as follows. First, the pixels of the image are sorted according to the zenith angle with the greatest angles listed first. We know that the vectors at the occluding contours must be pointing away from the body of the object, so we rotate the azimuth vectors by $180°$ to $\alpha_2^{(d)}$ where necessary in order to meet this requirement. For now, the occluding contour is simply taken as the curve in the image where the intensity drops to greylevel 20 or lower (since the laboratory walls were matte black). The occluding contour is not used in the techniques developed in later chapters. Next, using the sorted pixels, the algorithm propagates into the object, and rotates the azimuth angles by $180°$ where this provides better alignment with the local mean normal direction. This simple approach faces difficultly far from object limbs, due to insoluble convex/concave ambiguities. Nevertheless, a reasonable first estimate of object shape is possible, since the method is reliable at object limbs. Also, the method is computationally efficient.

3.2.2 Frankot and Chellappa Method

The final step in processing was to the recover the surface depth from the normal estimates. This was performed using the well known Frankot-Chellappa surface integration algorithm [27], which is used at several points in this thesis. Here, we present a brief description of the method.

The method works by projecting the gradient field (which can be based on the surface normal estimates) onto the nearest integrable solution. That is, the following condition about the surface height, $h(x, y)$, is enforced:

$$\frac{\partial^2 h(x,y)}{\partial x \partial y} = \frac{\partial^2 h(x,y)}{\partial y \partial x} \tag{3.25}$$

Enforcing this constraint globally has two effects. Firstly, it means that the reconstructed surface is independent of the integration path taken through the field of surface normals. Secondly, it smooths the surface normals, since it enforces continuous spatial derivatives.

The *integrable* (target) surface is represented by the sum of a series of basis functions.

Effectively, the surface is composed of a sum of Fourier components, each with a unique coefficient:

$$h(x, y) = \sum_{i,j} \chi_{i,j} \Psi_{i,j}(x, y) \tag{3.26}$$

where i and j are discrete point indices, the basis functions are denoted by Ψ and their respective coefficients are χ. The first partial derivatives of the integrable height are expressed similarly for each of the x and y directions of differentiation. As the surface is integrable, the set of coefficients for each differentiation direction are identical.

The same Fourier representation is used for the original non-integrable surface derivatives. However, this time the coefficients are different for each derivative direction. The Frankot-Chellappa algorithm calculates the integrable surface by minimising the distance between the coefficients, χ, for the integrable and non-integrable surfaces. The final surface estimate is then obtained using an inverse Fourier transform on the resulting basis function components.

3.3 Experiments

This section describes a set of experiments conducted with the aim of illustrating both the possibilities and the limitations of shape recovery using the method described above. Firstly, we present the results of applying the technique to a variety of objects with complex geometries. Secondly, a discussion of some of the difficulties with shape recovery using the method is given. We then focus on cylindrical objects, with the aim to assess the accuracy of the measurements for various materials. Finally, we describe a technique and some experiments to measure refractive indices.

3.3.1 Shape Recovery

The first experiments conducted for this thesis are summarised in Figure 3.9. The aim here was simply to verify the theory described in Section 3.1. Using the arrangement shown in Figure 3.4, a sequence of images of a porcelain cylinder were acquired, with 36

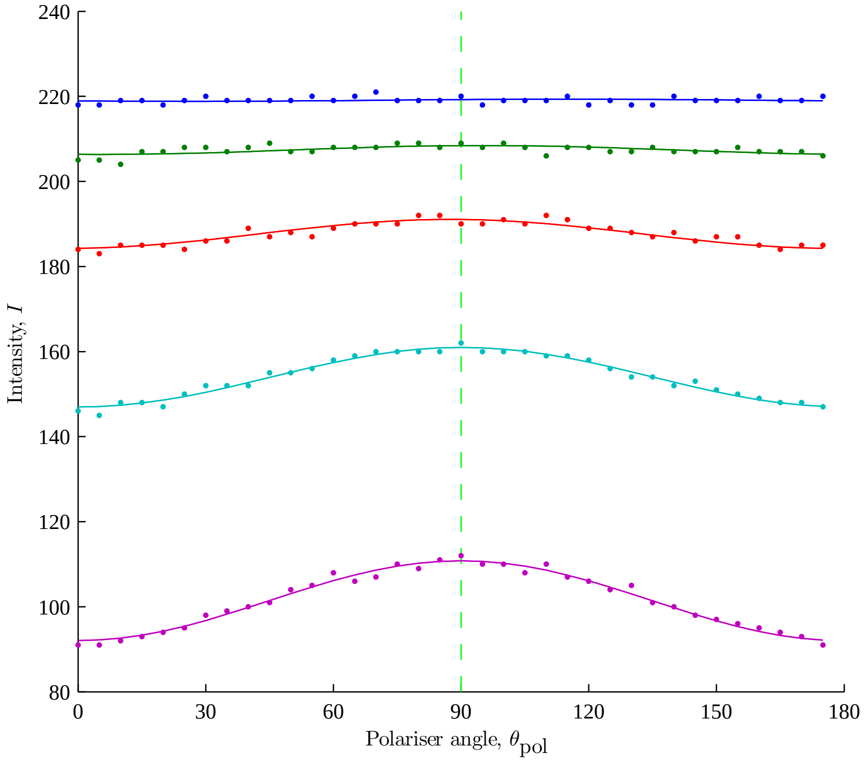

Figure 3.9: Measured intensities (and sine fits) as a function of polariser angle for points on a smooth porcelain cylinder where the zenith angle, $\theta = 15°$ (top curve), 30°, 45°, 60° and 75°.

equally spaced polariser angles, θ_{pol}, between 0 and 180°. The transmitted radiance sinusoid (Equation 3.8) was then fitted to measured pixel brightnesses using the Levenberg-Marquardt algorithm.

Since the surface is a smooth cylinder, some of the complications described later such as inter-reflections and roughness, are not present. The recovered zenith and azimuth angles are shown in Table 3.1. Not surprisingly, given the lack of intensity variations associated with small zenith angles, the first phase estimate was inaccurate. Note however, that the effect of the azimuth angle error on the overall surface normal error is small in such cases. The other estimates are all within a few degrees of the ground truth.

Having demonstrated the possibility of surface orientation estimation, the method described in Section 3.2 was implemented and tested on various objects, as shown in Figure 3.10. The top row shows the raw greyscale images of the objects studied. The first three objects are made from smooth shiny porcelain and the fourth object is made from slightly rough yellow plastic. The second row shows the phase angle of the reflected light. The

Table 3.1: Measured zenith and azimuth angles of points on a porcelain cylinder

Ground truth θ	Measured θ	Ground truth α	Measured α
15	13.4	90	38.9
30	27.1	90	79.4
45	46.6	90	93.9
60	64.0	90	89.6
75	78.4	90	90.2

phase has been accurately recovered for all objects including the rougher plastic duck. The measured degree of polarisation is shown in the third row. The effects of the roughness of the plastic duck are more noticeable here: the degree of polarisation appears slightly lower at the object limbs, compared to the smoother porcelain objects. The recovered fields of surface normals, or needle maps, are shown in the bottom row. Figure 3.11 shows the result of applying the needle map integration algorithm to the field of surface normals. It is clear that the basic shapes have been recovered, but with the shortcomings discussed below.

3.3.2 Limitations

There are three main limitations of the technique and are described in this section. The first of these is that the 180° ambiguity in azimuth angles is not comprehensively solved. Close examination of Figs. 3.10 and 3.11 reveal that not all areas of the surfaces have been correctly disambiguated. This is particularly problematic far from the occluding contours and for objects with complex geometry. A more complete disambiguation process is described in Chapter 5 and involves two views.

The second limitation of the technique is due to the difficulties encountered when reconstructing rough surfaces, as described in Section 3.1.5. The reconstruction of the slightly rough plastic duck in Figure 3.10d, is encouraging since the phase angle still provides a good estimate of the surface azimuth angle (Figure 3.10p). The depth estimate, shown in Figure 3.11d, is not too bad but the zenith angles were underestimated resulting in the reconstruction being too flat.

The third limitation of the method is that caused by specularities. This includes both

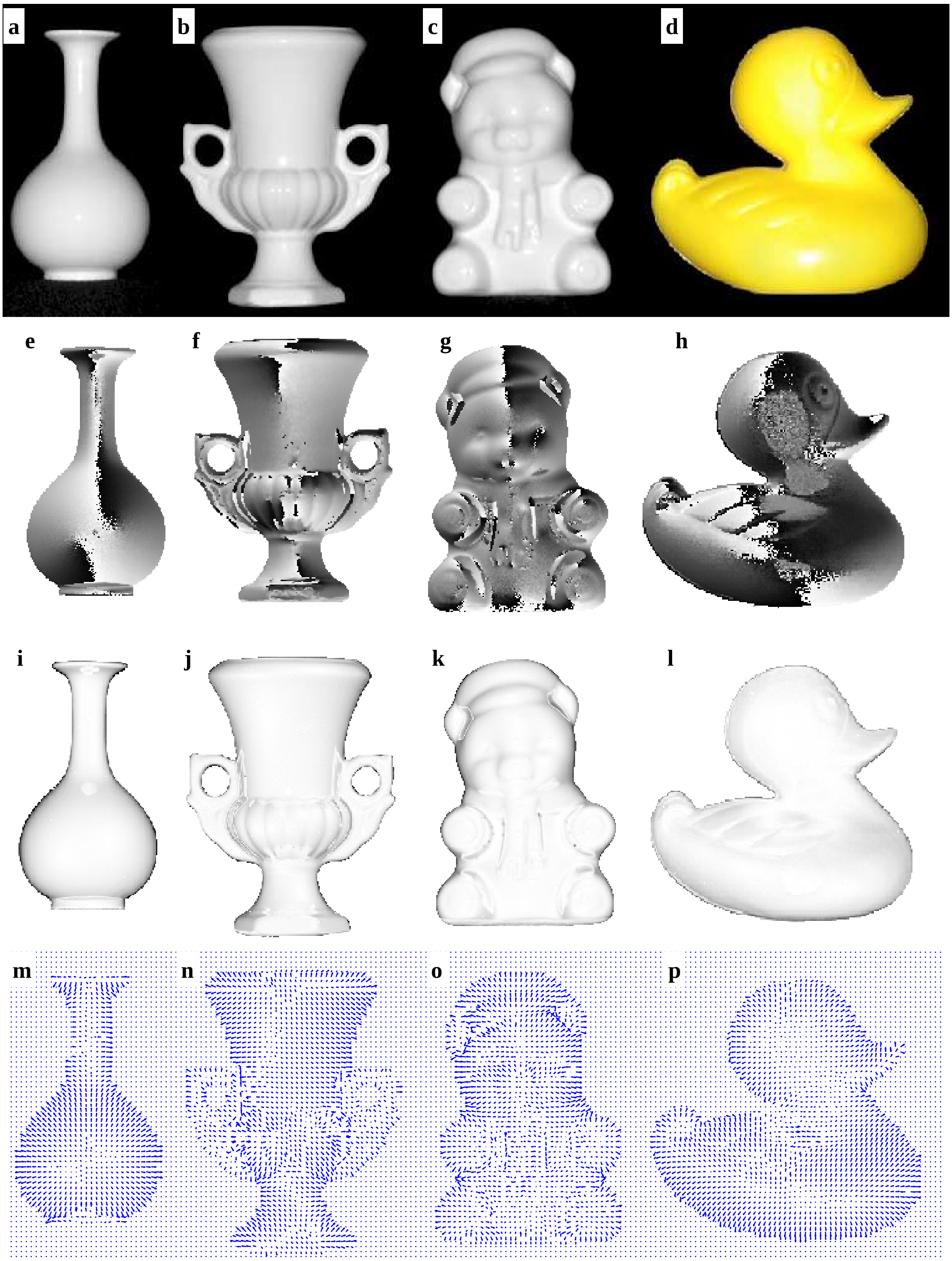

Figure 3.10: Greyscale images of a smooth porcelain (a) vase, (b) urn and (c) bear model and (d) a slightly rough plastic duck. The images are approximately 350 pixels in height. (e-h) Phase images. (i-l) Degree of polarisation images, where dark areas have highest values. (m-p) Estimated surface normals (reduced resolution).

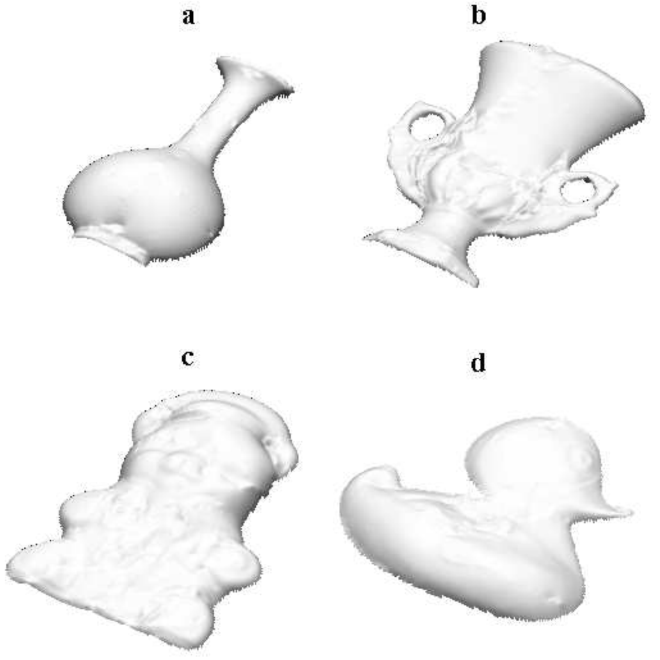

Figure 3.11: Recovered depth maps of the porcelain vase, urn and bear model and the plastic duck.

direct reflections from a light source and inter-reflections between different points on the object. In the case of direct specularities, we should ideally apply the theory for specular reflection to recover the surface orientation (i.e. use $\alpha = \alpha_1^{(s)}$ or $\alpha_2^{(s)}$ and calculate θ from Equation 3.15). Of course, isolating specularities from the rest of the image is not a trivial task [109]. In many situations however, including those described in this thesis, there are only a small number of light sources and each source subtends only a small angle from the object. In such situations, direct specularities are small and have little effect on results. For many of the experiments described here, a single light source is used which lies in the same direction from the object as the camera. In this case, the specularities cause saturation for all polariser angles. This means that $I_{\min} = I_{\max}$, and so the estimated zenith angle is zero, as desired.

For inter-reflections (or direct specularities from weak sources), there are competing diffuse and specular components that may or may not cause large errors in surface orientation estimates, depending on the relative strengths of each component. We demonstrate the detrimental effects of inter-reflections here, but postpone a detailed analysis until Chapter 4, where we also describe a method to minimise their effects. Figure 3.12

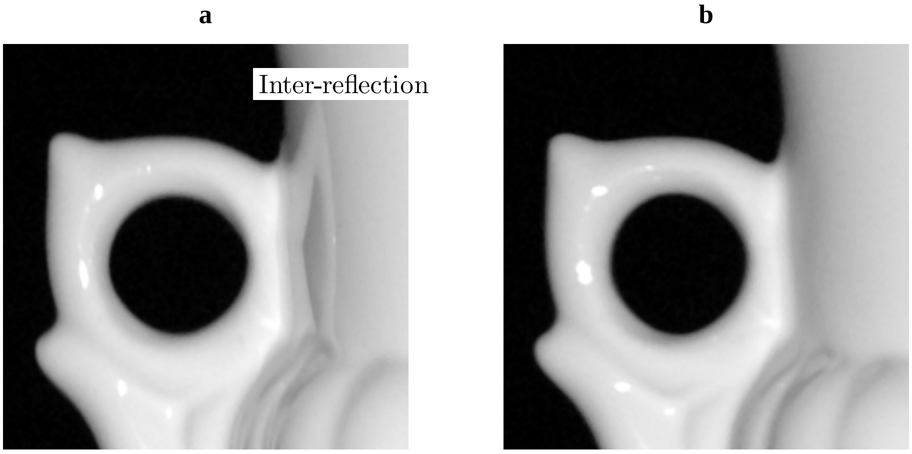

Figure 3.12: Close-up images of the handle of the porcelain urn shown in Figure 3.10b using (a) a vertical polariser ($\theta_{\text{pol}} = 0$) and (b) a horizontal polariser ($\theta_{\text{pol}} = 90°$).

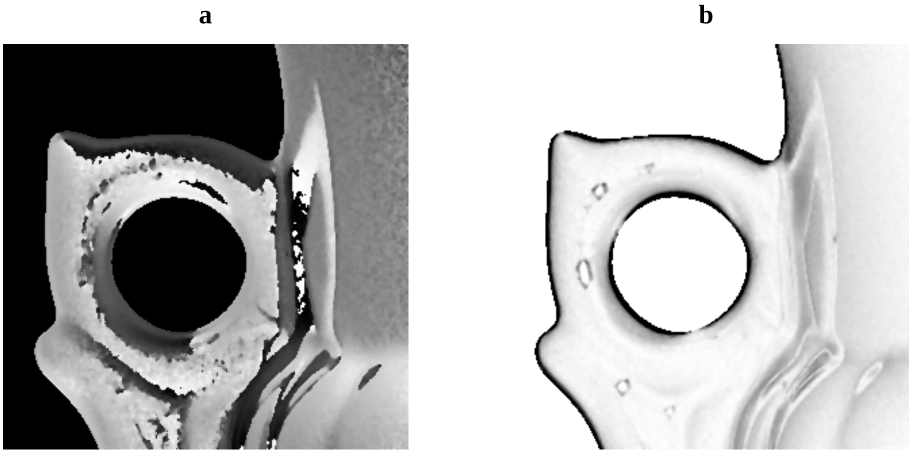

Figure 3.13: (a) phase image and (b) degree of polarisation image of the handle of the porcelain urn. The degree of polarisation image has been rescaled here to aid the figure's clarity.

shows two close-ups of the handle of the porcelain urn from Figure 3.10b. The images were taken using $\theta_{\text{pol}} = 0$ and $90°$. In (a), the inter-reflection of the handle can be seen on the body of the urn since the specularly reflected light is polarised vertically and so is transmitted through the polariser. When the polariser is rotated by $90°$, the inter-reflected light is absorbed by the polariser. The consequences of this can be seen in Figure 3.13, where there is a phase discrepancy of $90°$ in the azimuth angle. The degree of polarisation also differs from surrounding areas.

3.3.3 Evaluation of Accuracy

To assess the accuracy of the method, a set of vertically oriented cylinders of various materials were used. The geometry of a cylinder is convenient for three reasons. First, the structure is simple enough for shape recovery to be performed exactly from a single image[2]. Second, the analysis can easily be performed for all possible zenith angles. Thirdly, noise can be reduced by taking the average image intensity for each column of pixels. Due to this final point, and the fact that no inter-reflections are present with cylinders, this analysis therefore represents the best case scenario for the method.

Fitting the measured intensities to Equation 3.8 as before, we obtained a set of graphs showing the measured and theoretical zenith angles against position across the cylinder for different materials. For this experiment, eighteen different polariser angles were used. Since the azimuth angle of the cylinder is constant, we can also assess how the accuracy of azimuth angle estimates vary with zenith angle, if at all. A sample of these results for porcelain, blank photographic paper, photographic paper coated with cling film and normal paper are shown in Figure 3.14. Here the solid lines show the measurements and the broken lines show the ground truth. The photographic paper is much smoother than normal paper as it is coated to produce glossy photographs. The normal paper on the other hand is matte and much rougher. Several other material samples were also analysed, including different paper types, plastics, wax, terracotta and paper coated with ink. The first point to note about the figures is that, even for normal paper, which at the fine scale is very rough, the azimuth angles have been recovered to a very high degree of accuracy. However, more noise is associated with the rougher surfaces.

There was more variation in the accuracy of the zenith angle estimates. For Figure 3.14, the refractive index used was simply the value that produced greatest similarity between theory and experiment for the material in question. The shiny white porcelain object (the same type of porcelain can be seen in Figs. 3.10a-c) produced excellent agreement with theory down to very small zenith angles.

---

[2] This is simply done by manually isolating the cylinder from the background and calculating semicircles that arch from one side of the object to the other.

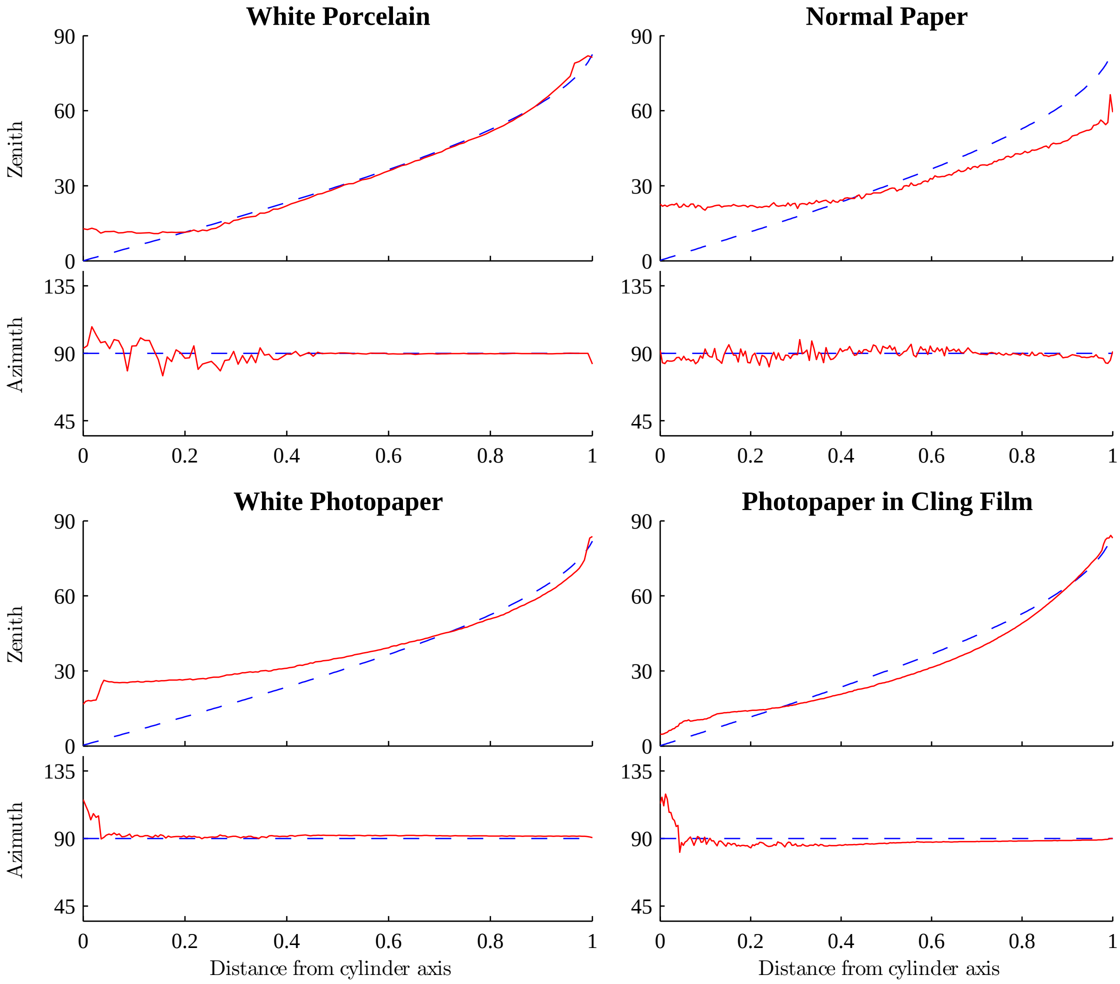

Figure 3.14: Plots of measured zenith and azimuth angles (solid lines) across the surfaces of cylinders of different materials. The exact values are indicated by the broken lines.

The remaining graphs in Figure 3.14 demonstrate the complications that can cause the measured zenith angles to deviate from the expected values. The result for blank white photographic paper, for example, is very accurate for large zenith angles but an increasing discrepancy is present as the zenith angle approaches zero. When the paper is coated in cling film, the discrepancy is less marked. Clearly, this suggests that there is a process occurring that is not accounted for by the theory and may be due to microfacet inter-reflections or imperfect depolarisation of incident light upon surface penetration. In any case, it should be noted that the greylevels differ only by a few greylevels as the polariser is rotated for regions of small zenith angles, meaning that results are generally unreliable for such areas. The results for paper, which is a rough matte surface, also show the phenomenon of finite polarisation at low zenith angles, as well as the depolarising effects of roughness nearer to the limbs.

For comparison, Figure 3.15 shows the zenith angle prediction using the Lambertian reflectance model. This model has been used previously for many SFS algorithms since the conception of the field by Horn [38]. For retro-reflection, the Lambertian model predicts that the intensity is simply given by

$$I \propto \cos\theta \tag{3.27}$$

where the constant of proportionality is the albedo. For Figure 3.15, the albedo was selected manually to give best agreement with the measured data. Clearly, polarisation analysis gave much better results for porcelain, but worse for paper (a classic example of a Lambertian surface). Note however, that polarisation gave a good estimate (albeit ambiguous) of the azimuth angle in both cases, while there is no obvious means to calculate the azimuth using SFS [123].

A final accuracy experiment involved dyeing the cylinder of photographic paper with different colours of ink, as shown in Figure 3.16. The recovered zenith angles are shown in Figure 3.17. It is obvious from the graphs where the transitions between colours occur. It appears that darker inks cause greatest variation, suggesting that good shape recovery

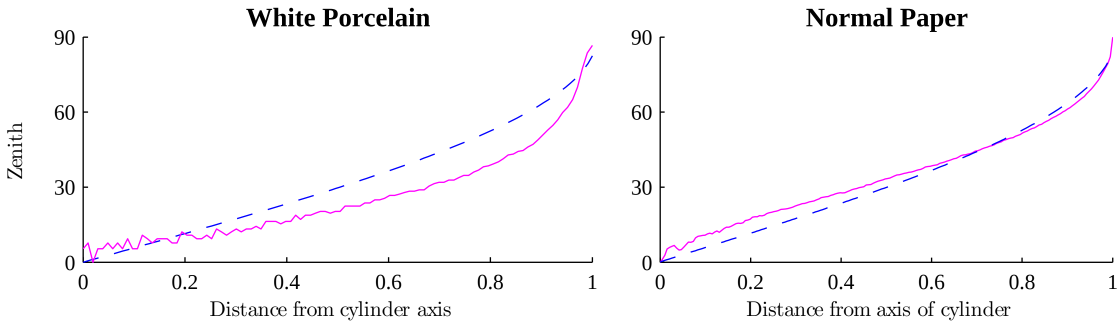

Figure 3.15: Estimation of zenith angles using the Lambertian reflectance model (solid line) compared to ground truth.

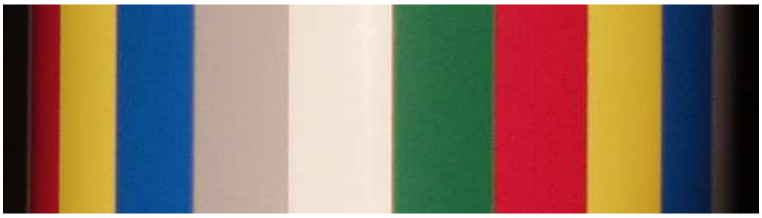

Figure 3.16: Dyed photographic paper wrapped around a cylinder used to test shape recovery for surfaces with different coloured coatings.

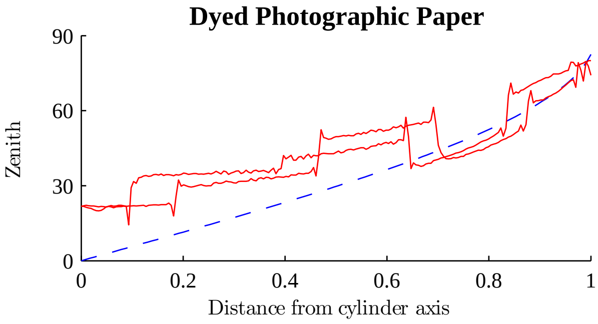

Figure 3.17: Measured (solid line) and exact (broken line) zenith angles across the surface of the dyed photographic paper in Figure 3.16. There are two solid lines, corresponding to the left and right hand sides of the cylinder.

would be possible only if inks of similar albedo were present in an image. This figure assumes that the refractive index is 1.4 for all of the inks — an assumption that is proved invalid in the next section.

3.3.4 Refractive Index Measurements

This short section has been included in the thesis to provide a link between the experiments and the theory. We describe a simple method to estimate the refractive index using a laser and photometer. Estimates for different materials are compared to the values used in Section 3.3.3.

In order to measure the refractive indices, the device shown in Figure 3.18 was constructed. The device works by placing a sample of the material under study on a table that can be rotated by small angles (measured to within 1° if the accuracy of the refractive index is to be within 0.1 of the true value). Laser light is then directed at a plane, vertical surface of the sample, which reflects the light through a polariser with horizontal transmission axis onto a white screen.

The spot of light on the screen is then observed as the sample is rotated about a vertical axis. The vertically polarised component of the reflected light is absorbed by the polariser, so the spot brightness varies as $R_{\|}$ in Figure 3.2. At the Brewster angle, θ_B, we know that $R_{\|} = 0$ and so the spot disappears. A simple relation between this angle and the refractive index can be found by applying Snell's Law (Equation 3.5) to the case where $\theta_i + \theta_r = 90^\circ$, to give:

$$n = \tan \theta_B \tag{3.28}$$

The refractive index can therefore be estimated directly from the Brewster angle. For smooth surfaces, the angle where the spot of light disappears is easy to measure accurately, giving an uncertainty in refractive index measurements of only 4%. This increases for rougher surfaces, since the laser light gets diffused by the reflecting surface. Conducting the experiment in a dark room meant that the disappearance of the spot could be detected by the naked eye, although a photometer can increase accuracy. It should be noted that the refractive index should not be measured using infra-red or any other non-visible light since it is a function of wavelength (we assume that the dependence on wavelength is negligible within the range of visible light). Note also that for coated surfaces, it is possible that a reflection from the underlying substrate affects the measured refractive index. We assume that this effect is negligible for these measurements.

Table 3.2 gives a comparison between the measured refractive indices and the values that produced the best agreement with the experiments in Section 3.3.3. Measurements taken from the refractometer are denoted by n_{ref}, while n_{exp} refers to the values that gave

Figure 3.18: Photograph and schematic of the refractometer. As the sample is rotated, the intensity of the spot of light on the screen changes according to Equation 3.4. The refractive index can be found by orienting the sample such that the spot disappears.

Table 3.2: Measured refractive indices

Material	n_{ref}	n_{exp}	Difference
Photographic paper	1.30	1.30	0.00
Photographic paper in cling film	1.54	1.70	-0.16
Photographic paper in thicker clear film	1.54	1.70	-0.16
Smooth white porcelain	1.28	1.30	-0.02
Smooth white plastic	1.28	1.28	0.00
Slightly rough orange plastic	1.54	1.25	0.29
Slightly rough wax	1.48	1.23	0.25
Cello tape	1.43	1.30	0.13
Grey ink	1.35	1.35	0.00
Blue ink	1.37	1.45	-0.08
Yellow ink	1.30	1.32	-0.02
Red ink	1.37	1.40	-0.03
Green ink	1.40	1.45	-0.05

best agreement with the experimental curves, such as those in Figure 3.14. Some of the roughest materials dispersed the laser light too much to allow reliable measurements to be made, and so are not included. Results for all the smooth, opaque materials, and all but one of the inks fell within experimental error. As one might expect, given the discussion in the previous section, rougher materials, translucent materials and materials coated in transparent media caused greater differences between the two values.

3.4 Conclusions

This chapter has accomplished the following:

1. It has presented Fresnel theory and its extensions to polarisation vision.
2. Described a polarisation image acquisition system.
3. Demonstrated that polarisation information can be used to accurately, but ambiguously, estimate surface azimuth angles for both smooth and slightly rough surfaces.
4. Shown that the zenith angles can also be calculated, but less accurately and are affected by roughness and inter-reflections.
5. Described a simple shape recovery algorithm based on polarisation data from a single view.
6. Conducted experiments to show the accuracy of surface normal estimates in the best case scenario (i.e. little noise, no inter-reflections and known refractive index).
7. Conducted an experiment to measure refractive indices of various materials.

The first two of these points are essentially standard, but form the theoretical and experimental foundations of the thesis. The third point is important, since many shape recovery algorithms are unable to accurately and reliably estimate azimuth angles. It also shows that polarisation vision is not restricted to perfectly smooth objects. Obviously, the ambiguity in the azimuth angles present difficulties for shape recovery and the disambiguation routine described in this chapter was rather simplistic. We shall return to this issue in Section 4.3 and Chapter 5. Point 4 shows that, for smooth surfaces, reasonable zenith angle estimates can be made from the polarisation data. However, the complications caused by inter-reflections and roughness illustrate that there is scope for improvement. In Chapter 4, we address this issue by refining the zenith angle estimates using shading information. Point 5 shows that reasonable shape reconstruction is possible using a simple and efficient algorithm for certain cases. Point 6 shows the extent to which roughness affects the normal estimates in the absence of inter-reflections. Finally, point 7 offers an interesting link between the theory and experiments, but was not included to improve shape recovery techniques.

Chapter 4

Reflectance Analysis

This chapter presents an analysis of the reflectance characteristics of a variety of materials. The aim is to recover material reflectance functions for known illumination conditions with the aid of polarisation data. The recovered reflectance functions can be regarded as “slices” of the full BRDFs corresponding to the given lighting. They are to be represented numerically by relationships between the surface zenith and azimuth angles and the measured image intensities.

Two methods for reflectance function estimation are presented in this chapter. The first is concerned with the special case where the light source and camera directions are identical, i.e. the case of retro-reflection. In this situation, the attainable slice of the BRDF is a one-dimensional function of the zenith angle since the reflected intensity in the direction of the camera is independent of the azimuth angle. This allows a simpler, more accurate and more efficient BRDF estimation procedure than in the more general case. Our proposed algorithm uses robust statistics to fit a curve to the measured polarisation and intensity data.

The second method assumes known but arbitrary light source positions and seeks a two-dimensional BRDF that minimises an energy functional using simulated annealing. This method also makes use of photometric stereo, which is used to disambiguate the azimuth angle estimates. Both methods utilise the histogram of surface orientations (estimated from the polarisation image) and pixel intensities, which forms our data set for curve/surface fitting.

In addition to estimating BRDFs from greyscale images, we have conducted a few experiments with colour. In this case, three BRDFs are recovered from the raw images, one for each of the red, green and blue colour channels. These independent reflectance functions are then combined allowing to render colour images.

The organisation of this chapter is as follows. In the first section, we formally define the BRDF and explain how the physical quantities involved relate to the discrete, dimensionless pixel greylevels associated with an image. The section also introduces the results of a physics-based BRDF model derived by Wolff that we later compare to our estimated BRDFs. In Section 4.2, the method for co-incident viewing and illumination is presented, along with a set of experiments for various materials. Section 4.3 describes a multiple image technique to disambiguate azimuth angle estimates using three different light source directions. Finally, Section 4.4 describes the algorithm to recover BRDF data with non-coincident viewing and illumination. Again, results are presented to support the method.

4.1 Reflectance Function Theory for Computer Vision

The BRDF, $f\left(\theta_i, \alpha_i, \theta_r, \alpha_r\right)$, of particular a material is the ratio of the reflected radiance, $L_r(\theta_r, \alpha_r)$, to the incident irradiance, $L_i(\theta_i, \alpha_i)$, for given illumination and viewing directions. It is measured per unit solid angle per unit foreshortened area and is given by

$$f\left(\theta_i, \theta_r, \alpha_i, \alpha_r\right) = \frac{L_r(\theta_r, \alpha_r)}{L_i(\theta_i, \alpha_i) \cos\theta_i \, \mathrm{d}\omega} \tag{4.1}$$

where θ_i and α_i denote the zenith and azimuth angles of the incident light and θ_r and α_r refer to the zenith and azimuth angles of reflected light respectively [26]. The light source for Equation 4.1 is assumed to subtend a differential solid angle $\mathrm{d}\omega$ (integration can be used for extended light source distributions). As indicated by Equation 3.7, the measured intensity is approximately proportional to the incoming radiance. We do not, therefore, calibrate the light source but work with quantities that are proportional to the true radiance.

We will not be using the reflectance function directly, but the *radiance function*, $L_r(\theta_i, \theta_r, \alpha_i, \alpha_r)$, instead. The radiance function incorporates the $\cos\theta$ term in Equation 4.1, so that $L_r(\theta) \propto f(\theta)\cos\theta$. For the remainder of this chapter, we therefore concentrate on the estimation of the radiance function. It is straightforward to calculate the reflectance function *up to an unknown constant* with the radiance function to hand.

Throughout this thesis, all surfaces are assumed to exhibit isotropic reflectance properties. As a result, the BRDF is independent to rotations about the surface normal and the number of degrees of freedom reduces to three, i.e. $f(\theta_i, \alpha_i, \theta_r, \alpha_r) \to f(\theta_i, \theta_r, \alpha_i - \alpha_r)$. For the special case of retro-reflection, the number of degrees of freedom reduces to one. This is because the two zenith angles in Equation 4.1 are equal $(\theta_i = \theta_r = \theta)$, and the radiance reflected in the viewer direction is independent of azimuth angle. For retro-refection, the reflected radiance can therefore be written as

$$L_r(\theta) = f(\theta)\, L_i(\theta) \cos\theta \,\mathrm{d}\omega \tag{4.2}$$

Section 4.2 focusses on this specific case. In effect, the method presented estimates a unidirectional reflectance distribution function. Since the measured pixel intensities are independent of the azimuth angles under retro-reflection conditions, the azimuth angle ambiguity does not complicate the reflectance function estimation process.

The incident radiance, $L_i(\theta)$, is not generally known since we work with an uncalibrated light source. Consequently, we simply estimate the reflected radiance up to an unknown constant. However, this constant is unimportant for many graphics and vision applications such as surface rendering or shape estimation, since we only require the relationship between the surface normal directions and the pixel brightness, I.

Our approach to BRDF recovery is image-based, since the estimation is derived numerically from the image data alone. In contrast, many attempts have been made previously to fit mathematical functions to the observed intensities, with or without reference to physical considerations. A brief survey of these methods was provided in Section 2.3. Many of the results in this chapter are compared to the physics-based model developed by Wolff [115]. The model is based on Fresnel theory, which was presented in Section 3.1, where light penetrates a surface and is internally scattered. Subsurface refraction is modelled using a Fresnel attenuation factor. Wolff uses a scattering theory, which was originally derived for radiative transfer, to model the distribution of scattered light within a medium. Multiple internal reflections are permitted by the theory, as in the reflected ray in Figure 3.7. The principle result of the Wolff model is the following BRDF prediction:

$$f\left(\theta_i, \theta_r, n\right) = \varrho \ \left(1 - F(\theta_i, n)\right) \left(1 - F\left(\arcsin\left(\tfrac{\sin\theta_r}{n}\right), \tfrac{1}{n}\right)\right) \tag{4.3}$$

where ϱ is the *total diffuse albedo* and $F\left(\theta_i, n\right)$ is the Fresnel function given by

$$F\left(\theta_i, n\right) = \tfrac{1}{2} \tfrac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)} \left(1 + \tfrac{\cos^2(\theta_i + \theta_t)}{\cos^2(\theta_i - \theta_t)}\right) \tag{4.4}$$

where $n = \left(\sin\theta_i\right) / \left(\sin\theta_t\right)$ (Snell's law).

In the case of retro-reflection, the two Fresnel terms in Equation 4.3 are equal. Introducing a $\cos\theta$ factor to account for foreshortening effects, we can therefore write

$$I\left(\theta, n\right) \propto \cos\theta \left(1 - F\left(\theta, n\right)\right)^2 \tag{4.5}$$

For the next section, Equation 4.5 is sufficient for our comparisons with the model. For Section 4.4, when more general illumination conditions are used, we must revert back to using Equation 4.3.

Equation 4.5 is monotonic in θ. This coincides with intuition, since one would expect the intensity to continuously decease as the line of sight approaches the grazing angle of the surface. Not surprisingly, the Lambertian model (Equation 3.27) also predicts this, as do our experiments presented later in this chapter. It is worth noting however, that this is not always the case. Research into light scattering from rough surfaces [74, §4.3] has shown that multiple scattering at the surface can destroy the monotonicity in certain rare cases. The effect is particularly marked for extremely rough surfaces when the roughness correlation length is comparable to the dimensions to the wavelength of the incident light. Some success has been achieved at quantitatively accounting for such effects but the models are not yet comprehensive.

In their experiments, Dana et al. [19] measured BRDF data for a wide range of materials. A few of which were shown to exhibit non-monotonic relationships. For example, the individual strands of a velvet surface can reflect light very strongly if the orientation of the strands and the direction of the incident light are perpendicular. This results in a brightening effect near occluding contours. Since this thesis is only concerned with smooth and slightly rough surfaces however, we shall make assumptions on the monotonicity of the BRDFs to aid our algorithms (details of the constraints that we apply are given in subsequent sections).

4.2 One-dimensional Reflectance Function Estimation

This section presents the method for estimating a one-dimensional slice of the BRDF for the case of retro-reflection. The technique makes use of the 2D histogram of zenith angles and intensities, i.e. the observed distribution of the greyscale values with the zenith angles. The aim is to robustly fit a curve to this histogram, which accurately represents the material reflectance function. The fitting is achieved in two steps. First, an initial estimate is acquired by tracing the peak frequency curve on the histogram. The estimate is then refined by fitting probability density functions to the histogram and forming the BRDF curve from those. At the end of the section, we present results for a series of objects and materials.

4.2.1 Statistical Analysis

The histogram of zenith angles and intensities, which is the essential ingredient for our method, is defined formally as follows. From the polarisation data, we have a set of Cartesian pairs, $\mathcal{D} = \{(\theta_d, I_d)\,; d = 1, 2 ... |\mathcal{D}|\}$, where θ_d is the zenith angle estimated

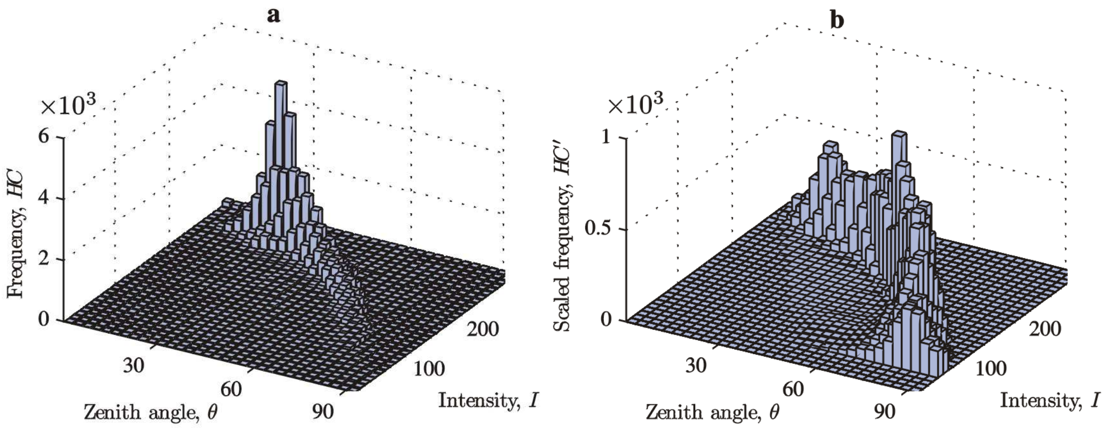

Figure 4.1: (a) Histogram of intensities and zenith angle estimates for the porcelain bear. (b) Scaled histogram.

from Equation 3.21 and I_d is the measured pixel greyscale value at the pixel indexed d. The histogram contents for bin (a, b) are given by $\mathcal{HC}_{a,b} = |\mathcal{BC}_{a,b}|$, where $\mathcal{BC}_{a,b} \subseteq \mathcal{D}$ is the set Cartesian pairs falling into the relevant bin. This subset is given by

$$\mathcal{BC}_{a,b} = \left\{ (\theta_d, I_d)\,; \begin{array}{rcl} \theta_d & \in & \left[\theta_a^{(\text{bin})} - \frac{90^\circ}{2k_{\text{max}}}, \theta_a^{(\text{bin})} + \frac{90^\circ}{2k_{\text{max}}}\right) \wedge \\ I_d & \in & \left[I_b^{(\text{bin})} - \frac{255}{2k_{\text{max}}}, I_b^{(\text{bin})} + \frac{255}{2k_{\text{max}}}\right) \end{array} \right\} \tag{4.6}$$

where $\left\{\theta_a^{(\text{bin})}; a = 1...k_{\text{max}}\right\}$ and $\left\{I_b^{(\text{bin})}; b = 1...k_{\text{max}}\right\}$ are k_{max} equally spaced bins ranging from 0 to 90° and 0 to 255 respectively and k_{max} is the number of bins. For this chapter, we use $k_{\text{max}} = 100$. Figure 4.1a shows the histogram $\{\mathcal{HC}_{a,b}\}$ for the porcelain bear model shown in Figure 3.10c.

Figure 4.2a shows a "footprint" of the histogram. Since the light source and camera directions are approximately identical, one would expect the intensity to decrease monotonically with the zenith angle. Indeed, the general structure of the histogram confirms this. Note however, that the main curve in the histogram is broad and that a significant number of pixels fall into bins far away from this curve. Figure 4.2 highlights image regions corresponding to three different parts of the histogram. Unlike the pixels falling into box 1, the pixels in boxes 2 and 3 do not follow the general trend.

The main reason for the wide spread of the data in Figure 4.2 (in addition to noise)

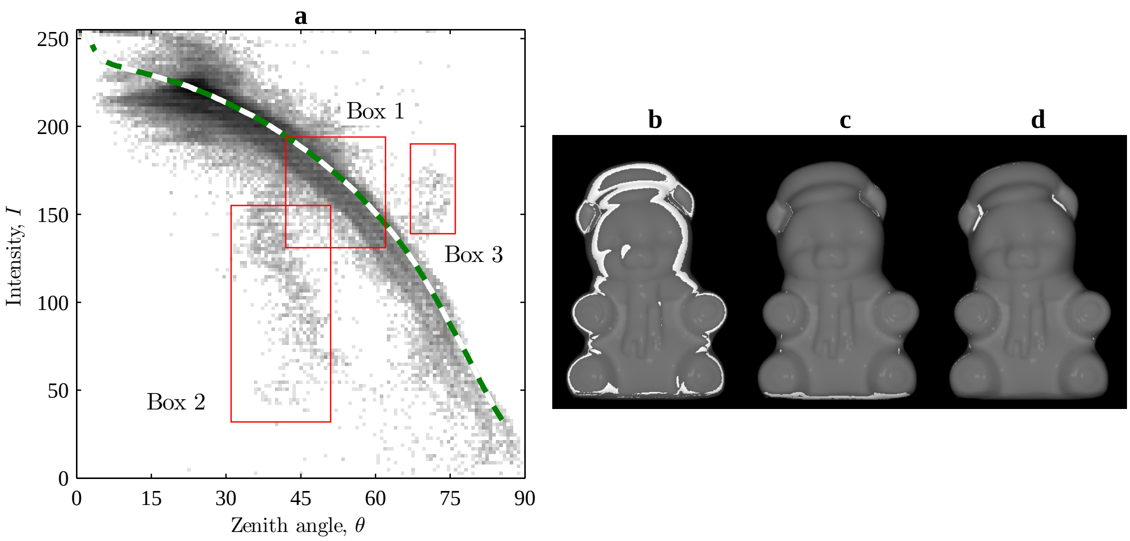

Figure 4.2: (a) Footprint of the histogram in Figure 4.1a. Bins of higher frequency are shown as dark patches (logarithmic scale). The exact reflectance curve is also shown. The greyscale images highlight pixels falling into (b) box 1, (c) box 2 and (d) box 3.

is that inter-reflections are taking place. An inter-reflection occurs where light from a source is *specularly* reflected from one point on the object to another, and then toward the camera. This process therefore satisfies the conditions for specular reflection, although a diffuse component will generally be present also. It can occur at the small scale between the corrugations that constitute roughness, or macroscopically between different regions of the surface or the environment.

The exact effect of an inter-reflection depends upon its strength. If it is weak, then the diffuse component still dominates. The degree of polarisation, and hence the zenith angle estimate, will be reduced since the angle of polarisation of specularly reflected light is perpendicular to that of diffuse reflection (see Chapter 3). This process can be seen in box 2 of Figure 4.2. Here, a small specular component is present due to reflections from the table on which the object rests. For strong inter-reflections, the specular component dominates, as in box 3. Since the polarizing properties of specular reflection are greater than that for diffuse reflection, we have a situation where the degree of polarisation exceeds that which would be expected for purely diffuse reflection.

Notice that, apart from the cases were the strong inter-reflection limit is met (as in box 3), the degree of polarisation is equal to, or less than, the expected value for a given zenith angle. This is substantiated by the fact that the exact curve (measured using an object of the same material but known shape) approximately follows the outer envelope of the histogram.

4.2.2 Proposed Method

The radiance function, $L_r(\theta)$, is to be approximated in terms of a set of discretely sampled Cartesian pairs $\hat{L}_r = \{(\theta_k, I_k); k = 1, 2 \ldots k_{\text{max}}\}$. Next we describe a means to acquire an initial estimate of $\hat{L}_r$, before presenting a method for refining the results.

Initial Estimate

Figure 4.2 suggests that we need to calculate a curve that follows the outer envelope of the histogram. However, we obtain an *initial estimate* of the reflectance function curve derived from the crest of the histogram. One difficulty here is that there are too few data points in some parts of the histogram to obtain a reliable estimate. We overcome this problem by dividing the data into histogram bins of equal frequency, instead of using the uniformly spaced bins mentioned previously. Our algorithm accomplishes this task by sorting the pixels in $\mathcal{D}$ according to intensity and selecting the intensity of every $(|\mathcal{D}|/k_{\text{max}})^{\text{th}}$ pixel as a bin edge. For all pixels falling into a given intensity bin, we apply a similar process to obtain zenith angle bin boundaries.

The frequency for each bin, $\mathcal{HC}$, is then divided by its area to give a scaled frequency $\mathcal{HC}'$ at the centre points of each bin. Two-dimensional linear interpolation is then used between the bin centres to calculate a value at each point on an evenly spaced $k_{\text{max}} \times k_{\text{max}}$ grid to form a resampled and rescaled version of the histogram, $\{\mathcal{HC}'_{a,b}\}$, as shown in Figure 4.1b.

Although, the resultant "histogram" has no direct physical interpretation, it makes it easier to robustly trace the peak frequency curve. The curve is first calculated in polar

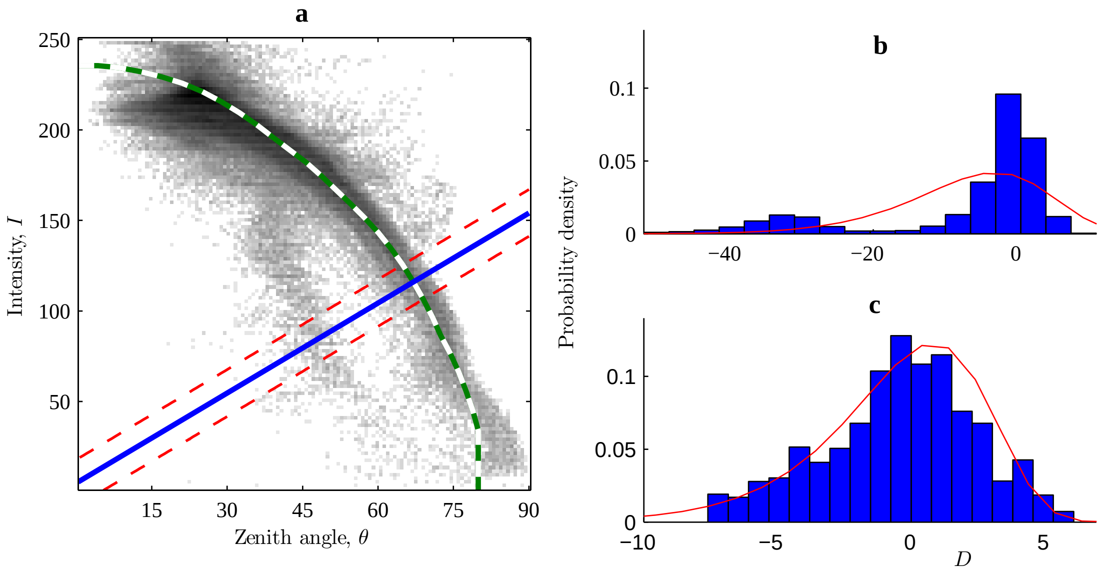

Figure 4.3: (a) Broken curve: initial BRDF estimate. Solid straight line: $I_B(\theta)$. Broken straight lines: $I_L(\theta)$ and $I_U(\theta)$. (b) Histogram of the length $\{D_d\}$ for all pixels bounded by $I_L(\theta)$ and $I_U(\theta)$. (c) Histogram after removal of pixels that do not obey Equation 4.13.

co-ordinates[1] with the origin at $(\theta = 0, I = 0)$. For k_{max} equally spaced polar co-ordinate angles between 0 and $90°$, the scaled frequency, HC', is plotted against the distance from the origin. The distance where HC' reaches its maximum is then taken as a data point on the initial BRDF estimate. These distances are then smoothed before reverting back to Cartesian space. An example of an initial BRDF estimate is shown in Figure 4.3a. The initial estimate is often quite reasonable by itself and takes less than two seconds to calculate for a 512×512 image using MATLAB on a 2.8GHz computer (which is used for all computations mentioned in this thesis).

Final Estimate

Figure 4.4 illustrates the geometry used to calculate the final BRDF estimate. For each pair of consecutive points on the initial estimate, (θ_k, I_k) and (θ_{k+1}, I_{k+1}), we calculate a point on the refined BRDF estimate that falls on the straight line $I_B(\theta, k)$. This is the perpendicular bisector of the second straight line $I_\perp(\theta, k)$, which connects the pair of

---

[1] It may be useful to think of the Cartesian co-ordinate space that contains the histogram and the discretely sampled points of the BRDF in terms of normalised co-ordinates of bin lengths, i.e. $I \leftarrow I k_{\text{max}}/255$ and $\theta \leftarrow \theta k_{\text{max}}/90°$ for the $k_{\text{max}} \times k_{\text{max}}$ histogram.

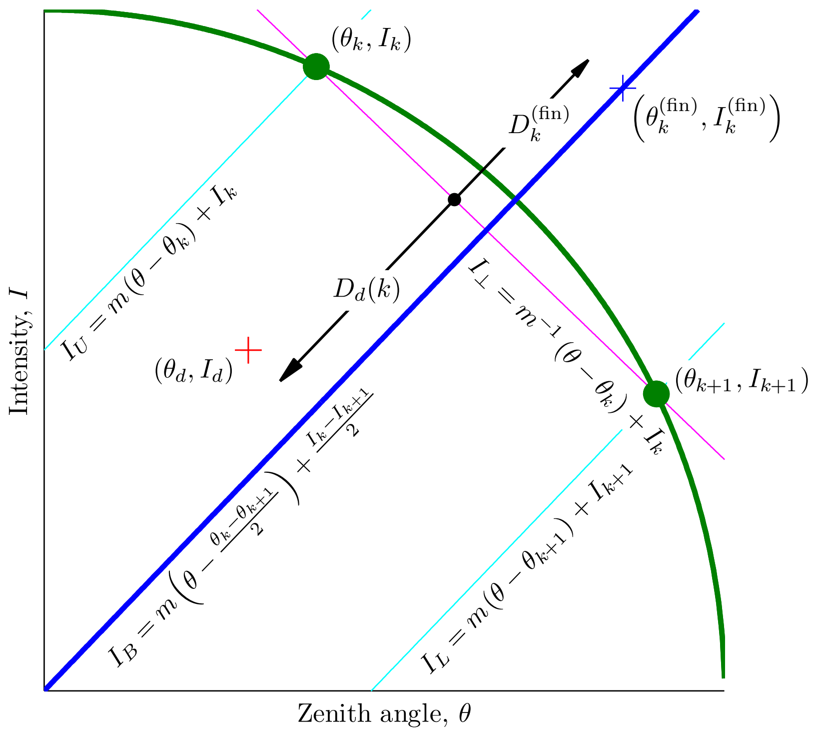

Figure 4.4: Definitions used in the calculation of the reflectance curve. The bold curve is the initial estimate.

points on the initial estimate. The gradient $m(k)$ is simply given by

$$m(k) = \frac{\theta_k - \theta_{k+1}}{I_{k+1} - I_k} \tag{4.7}$$

The two straight lines parallel to $I_B(\theta, k)$ that pass through (θ_k, I_k) and (θ_{k+1}, I_{k+1}) are also calculated. Call these $I_U(\theta, k)$ and $I_L(\theta, k)$ respectively. These lines define the region of the histogram that we use to estimate the point on the final BRDF curve related to point k on the initial estimate. The broken lines in Figure 4.3 show $I_U(\theta, k)$ and $I_L(\theta, k)$ for a particular k in a real-world example. For each k, we therefore extract data points $\{(\theta_d, I_d)\}$ that satisfy

$$\begin{aligned} I_L(\theta_d, k) < \quad & I_d & < I_U(\theta_d, k) \\ \theta_L(I_d, k) < \quad & \theta_d & < \theta_U(I_d, k) \end{aligned} \tag{4.8}$$

where $\theta_L(I, k) = \theta_{k+1} + (I - I_{k+1})/m(k)$ and $\theta_U(I, k) = \theta_k + (I - I_k)/m(k)$ are the inverse functions of I_L and I_U respectively.

We are interested in the distribution of distances $\{D_d(k)\}$ between the location of the data points d and the line $I_\perp(\theta, k)$ (see Figure 4.4). The distance $D_d(k)$ is given by:

$$D_d(k) = \frac{(I_d - I_k)\, m(k) + \theta_d - \theta_k}{\sqrt{m(k)^2 + 1}} \tag{4.9}$$

Negative values of $D_d(k)$ indicate that the point is to the left of the intersection between $I_B(\theta, k)$ and $I_\perp(\theta, k)$.

The broken straight lines in Figure 4.3a indicate the region of the histogram where Equation 4.8 is obeyed for a particular k. Figure 4.3b shows a histogram of $\{D_d(k)\}$ for the points falling into this region. To maintain the robustness of the algorithm, we would like to fit a probability distribution function (PDF) to the data. The most obvious choice of PDF would be a Gaussian. However, the histogram can be distorted due to inter-reflections and roughness, meaning that it is asymmetric. We therefore fit $\{D_d(k)\}$ to a Weibull distribution [52] instead, which takes the form of a skewed Gaussian. The Weibull PDF is given by

$$g\left(t|\psi, \zeta, \nu\right) = \begin{cases} \frac{\zeta}{\psi}\left(\frac{t-\nu}{\psi}\right)^{\zeta-1} \exp\left(-\left(\frac{t-\nu}{\psi}\right)^{\zeta}\right) & t \geq \nu \\ 0 & t < \nu \end{cases} \tag{4.10}$$

where ψ is a scale parameter, ζ is a shape parameter and ν is a location parameter.

In Figure 4.3b, the distances, $\{D_d(k)\}$, bounded by Equation 4.8 for a particular k, have been fitted to a Weibull PDF using a standard maximum likelihood approach. Clearly this is inadequate since there is an inter-reflection taking place that is resulting in a different distribution for some of the points (box 2 in Figure 4.2). However, since there are generally relatively few points in the inter-reflection we can discard data that falls under the tails of the PDF and calculate a new set of parameters. We iterate this until convergence (typically only three or four iterations are necessary). The histogram in Figure 4.3c shows the result. Note that we have used the same number of bins as for the first histogram, but over a smaller range.

In order to determine which points to discard, the algorithm first calculates the mean,

$\overline{\mu}$, and standard deviation, σ, of the PDF, which are given by

$$\overline{\mu} = \psi \Gamma\left(1 + 1/\zeta\right) \tag{4.11}$$

$$\sigma = \psi \sqrt{\Gamma\left(1 + 2/\zeta\right) - \Gamma\left(1 + 1/\zeta\right)^2} \tag{4.12}$$

where $\Gamma(x) = \displaystyle\int_0^\infty e^{-\xi} \xi^{x-1} \mathrm{d}\xi$ is the gamma function. We then discard points that do not fall within two standard deviations of the mean. That is, the algorithm recalculates the PDF parameters using only points that satisfy

$$\overline{\mu} - 2\sigma < D_d(k) < \overline{\mu} + 2\sigma \tag{4.13}$$

After the Weibull parameters have converged, we set

$$D_k^{(\text{fin})} = \overline{\mu} + 2\sigma \tag{4.14}$$

to represent the outer envelope of the histogram, which is shown in Figure 4.4. $D_k^{(\text{fin})}$ is generally positive since the BRDF is expected to lie on the *outer* envelope of the histogram. $D_k^{(\text{fin})}$ can easily be converted to a Cartesian pair $(\theta_k^{(\text{fin})}, I_k^{(\text{fin})})$, which we then accept as a point on the reflectance curve.

Applying the above procedure for each point on the initial estimate is sufficient to arrive at a good estimate of the reflectance function. There are nevertheless, several steps that we used to increase the robustness of the algorithm further:

Additional Steps

1. If there are too few pixels that satisfy Equation 4.8 for a points k and $k+1$ on the initial estimate, then use points k and $k + k_n$. The choice of k_n should be the minimum value such that a sufficient number (e.g. 1% of the total number of pixels) of points satisfy the condition.

2. Apply Gaussian smoothing to $\{D_k^{(\text{fin})}\}$ before calculating $\{(\theta_k^{(\text{fin})}, I_k^{(\text{fin})})\}$.

3. Apply the extra constraint [118] that $\lim_{\theta \to 90^\circ} f(\theta) = 0$

The algorithm is summarised in Pseudocode 1.

Pseudocode 1 One-dimensional BRDF estimation.

```
Initial Estimate
  Calculate the 2D histogram of intensities and zenith angles $\{HC\}$
  Scale the histogram to obtain $\{HC'\}$
  for [$k_{\text{max}}$ equally spaced polar angles] do
      Extract the 1D slice of $\{HC'\}$ for the given polar angle
      Set the peak of the 1D slice to be a point on the initial estimate, $(\theta_k, I_k)$
  Smooth the points on initial estimate, $\{(\theta_k, I_k)\}$

Final Estimate
  for [$k_{\text{max}}$ points on BRDF estimate] do
      Calculate the lines $I_B(\theta)$, $I_\perp(\theta)$, $I_L(\theta)$ and $I_U(\theta)$ (Figure 4.4)
      Extract points $k$ in relevant region of histogram (Equation 4.8)
      if [insufficient number of points in region] then
          Apply additional step 1
      Calculate distances $\{D_d(k)\}$
      Calculate 1D histogram of $\{D_d(k)\}$
      while [Weibull parameters not converged] do
          Calculate maximum likelihood Weibull parameters
          Discard outliers (Equation 4.13)
      Calculate $D_k^{(\text{fin})}$ (Equation 4.14)
  Smooth distances $\{D_k^{(\text{fin})}\}$
  Convert $\{D_k^{(\text{fin})}\}$ to Cartesian pairs $\{(\theta_k^{(\text{fin})}, I_k^{(\text{fin})})\}$
```

4.2.3 Experiments

The estimated reflectance functions for several porcelain objects are shown in Figure 4.5. Each curve took a total of approximately five seconds to compute. Clearly, the curves follow the general trend of the exact curve and are approximately the same as each other. Importantly, this shows that the inter-reflections have been discounted from the BRDF. The algorithm appears to be most reliable at intermediate zenith angles. For the largest zenith angles, less data tends to be available, although additional step 1 above improves matters. Arbitrarily setting the final PDF condition Equation 4.14 to $\overline{\mu} + \sigma$ would give results that more closely match the ground truth.

Figure 4.6 shows the histogram footprint of a partially painted porcelain cat model, which is used to investigate how areas of different albedo affect the algorithm perfor-

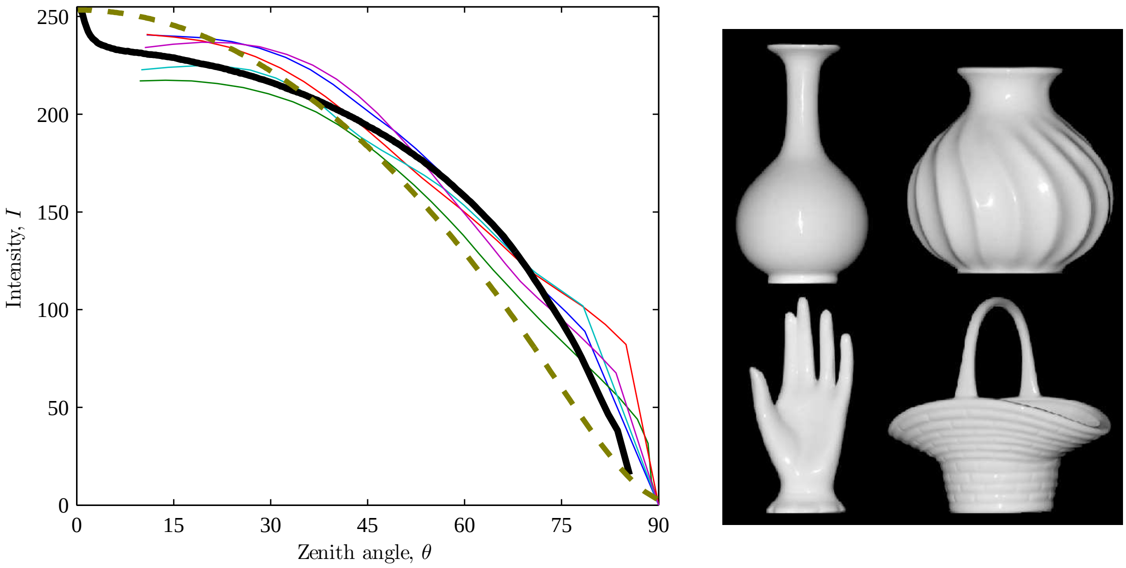

Figure 4.5: Estimated reflectance functions for the four porcelain objects shown and the model bear, compared to the exact (thick) curve and the Wolff model prediction (broken curve).

mance. The figure shows that pixels of the painted areas of the porcelain cat model fall well within the envelope of the histogram. This meant that the algorithm recovered the radiance function of the porcelain, while the smaller painted areas were disregarded by the robust fitting method. The estimated reflectance function is shown in Figure 4.7. Note however, that the algorithm does not always recover the radiance function of the most prevalent colour/material. For example, suppose that most of the cat model was red but that some areas were white. In this case, the results would be strongly affected by both red and white areas.

In Figure 4.7, we have applied the algorithm to an apple and an orange. Again, the results seem reasonable but the larger specular lobe [118] present with these rougher surfaces has reduced the accuracy at small zenith angles to a greater extent than that for the smoother porcelain objects in Figure 4.5. Note however, that the direct specular reflections, which cover a significant part of the orange's pitted surface, have been discarded by the algorithm. This gives us an approximation of the diffuse component of the BRDF, as desired.

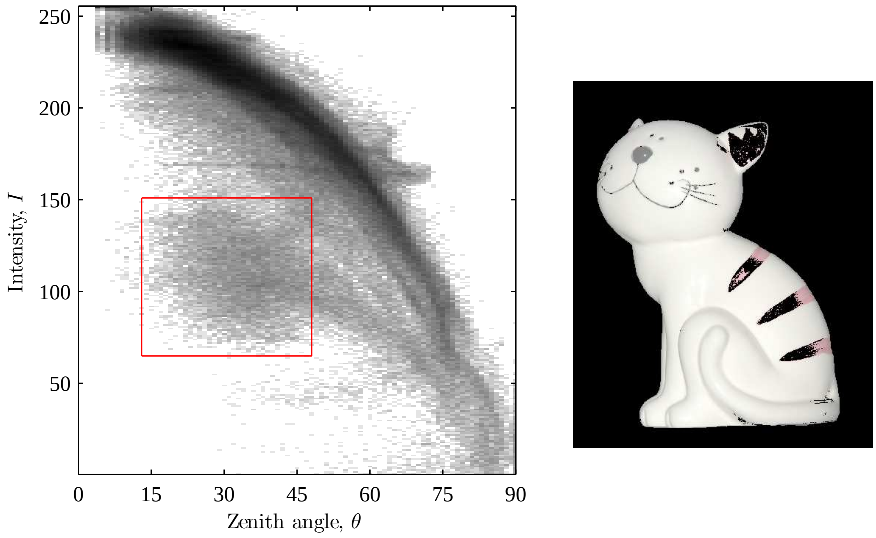

Figure 4.6: Histogram of a partly painted porcelain cat model. The areas falling into the box are shown in black with the rest of the image brightened for clarity.

4.3 Polarisation-based Photometric Stereo

We would like to extend the work from the previous section on retro-reflection to cover more general illumination conditions. The obvious limitation of the method above is the assumption that the BRDF is independent of the azimuth angle and is hence a one-dimensional function of the zenith angle. Clearly, this is only true for the considered case of retro-reflection. To extend the technique to more general light source directions, we need to obtain two-dimensional BRDFs since the reflected intensity becomes a function of both the zenith and the azimuth angles. The direct extension of the previous approach is to fit a surface to the three-dimensional histogram that incorporates azimuth angle bins, in addition to the zenith angle and intensity bins.

For the previous method, the confinement of the technique to retro-reflection resulted in the following two simplifications:

1. The $180°$ ambiguity in azimuth angle did not need to be resolved for the reflectance function estimation.

2. The portion of the BRDF that we were estimating was a monotonic one-dimensional function of zenith angle.

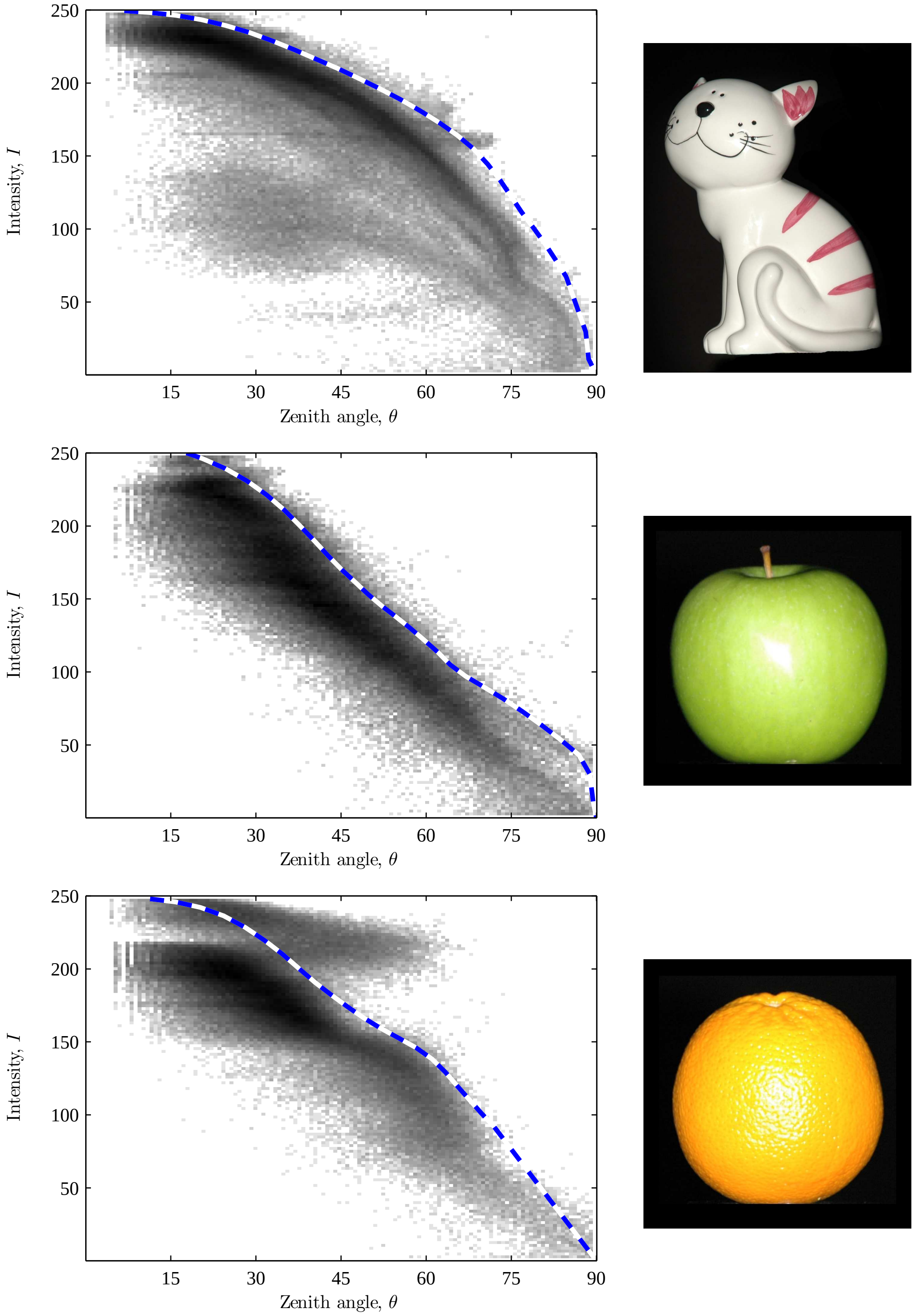

Figure 4.7: Histograms and estimated reflectance functions of the cat model, an apple and an orange.

Consider now the case where the light source and camera directions are not identical. The two points above are no longer valid and the task of reflectance analysis becomes significantly more involved. In this section, we discuss a method to overcome the first of these two issues using several light sources. The simple and efficient algorithm developed for this task is local and uses the variation in pixel intensities for changing light sources. This allows for the determination of the correct choice of azimuth angle, giving us an unambiguous set of Cartesian triples, $\mathcal{D} = \{\theta_d, \alpha_d, I_d\}$. In Section 4.4, the three-dimensional histogram is constructed from $\mathcal{D}$. Simulated annealing is then used to fit the two-dimensional BRDF to this histogram, thus addressing the latter point above.

4.3.1 Disambiguation by Photometric Stereo

Our principle aim here is to obtain a field of unambiguous surface normals using several different illumination directions. We obtain three separate polarisation images, each with a different light source position. This provides three phase angle (ambiguous azimuth angle) estimates and three zenith angle estimates for each pixel, in addition to the measured intensities.

The first task is to select the most reliable phase and degree of polarisation for each pixel. This is done by comparing the measured intensities at each pixel. For any particular pixel, the phase and degree of polarisation were used that correspond to the light source which gave the greatest intensity for that pixel. We make an exception to this rule where the maximum intensity is saturated (which we define as an intensity greater than grey level 250). In this case we use the highest intensity that is not saturated.

We use the experimental arrangement shown in Figure 4.8 for both the above angle selection process and for the disambiguation described below. Assume for now that the angles subtended by the camera and light sources from the object are equal (θ_L) and that the distances between the object and the light sources are also equal (D_L). The distance to the camera is unimportant.

Figure 4.9 shows the object and light sources from the point of view of the camera.

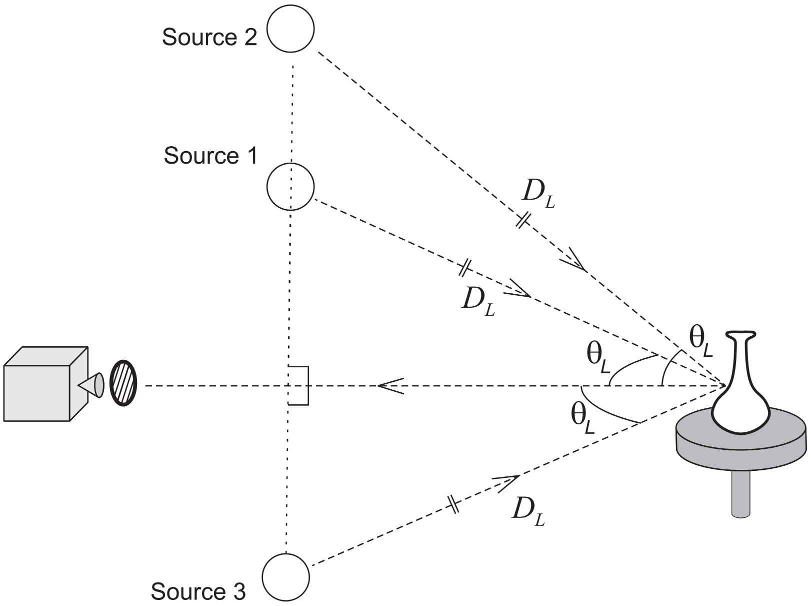

Figure 4.8: Geometry used for disambiguation by photometric stereo.

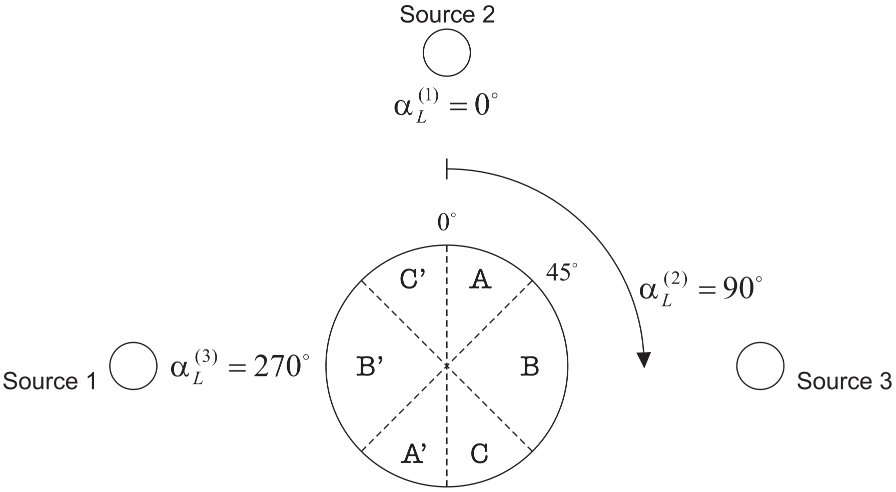

Figure 4.9: View of a spherical target object from the camera viewpoint. Azimuth angles are disambiguated differently depending on whether the phase angle is less than 45° (regions `A` and `A'`), between 45° and 135° (regions `B` and `B'`) or greater than 135° (regions `C` and `C'`).

In this case, the object is a perfect sphere. Let the intensity at pixel k observed under light source 1 be $I_k^{(1)}$. Similarly, under light sources 2 and 3, the intensities are $I_k^{(2)}$ and $I_k^{(3)}$ respectively. Considering the geometry of the photometric stereo rig, we can see that for pixels falling into the region labelled `A`, we would expect $I_k^{(2)} > I_k^{(3)} > I_k^{(1)}$. In region `A'`, which has the same phase angles as region `A`, we would expect $I_k^{(1)} > I_k^{(3)} > I_k^{(2)}$. We therefore, have a simple means of disambiguating azimuth angles in these regions:

$$\text{if} \quad \phi_k < 45^\circ \quad \text{then} \quad \alpha_k = \begin{cases} \phi_k & \text{if} \quad I_k^{(2)} > I_k^{(1)} \\ \phi_k + 180^\circ & \text{otherwise} \end{cases} \tag{4.15}$$

We use these specific inequalities (as opposed to $I_k^{(2)} > I_k^{(3)}$ or $I_k^{(3)} > I_k^{(1)}$) since we expect the difference in intensity to be greatest between $I_k^{(2)}$ and $I_k^{(1)}$ for regions `A` and `A'`.

Similar arguments can be used for the remaining regions of the sphere in Figure 4.9. In region `B` we expect $I_k^{(3)} > I_k^{(2)} > I_k^{(1)}$. Therefore, we have the condition:

$$\text{if} \quad 45^\circ \leq \phi_k < 135^\circ \quad \text{then} \quad \alpha_k = \begin{cases} \phi_k & \text{if} \quad I_k^{(3)} > I_k^{(1)} \\ \phi_k + 180^\circ & \text{otherwise} \end{cases} \tag{4.16}$$

Finally, in region `C`, $I_k^{(3)} > I_k^{(1)} > I_k^{(2)}$. Which gives us our final condition:

$$\text{if} \quad 135^\circ \leq \phi_k \quad \text{then} \quad \alpha_k = \begin{cases} \phi_k & \text{if} \quad I_k^{(3)} > I_k^{(2)} \\ \phi_k + 180^\circ & \text{otherwise} \end{cases} \tag{4.17}$$

Figure 4.10 shows results of azimuth angle disambiguation using the above process for porcelain objects. For this experiment, $\theta_L = 45^\circ$ was used. Note that using too large a value for θ_L would result in problems due to cast shadows. In Figure 4.10a, each pixel is shaded according to which light source resulted in greatest pixel intensity. Figure 4.10b shows the disambiguated azimuth angles. The angles have clearly been disambiguated successfully for most regions. Only for small zenith angles, where the intensity has least variation with light source, are there any artifacts that reveal the light source distribution. Figure 4.11 shows the disambiguated azimuth angles for a slightly rough plastic duck,

Figure 4.10: (a) The light grey areas of these figures show regions where greatest intensity was observed using light source 1. Dark grey areas correspond to source 2 and black areas relate to source 3. (b) Disambiguated azimuth angles.

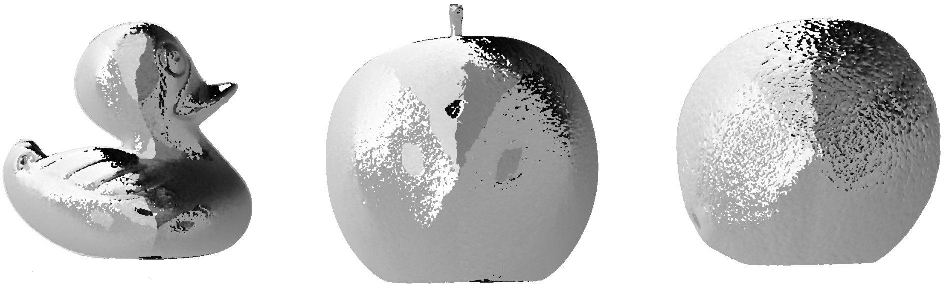

Figure 4.11: Azimuth angles disambiguated using the photometric stereo method for a slightly rough plastic duck, an apple and an orange.

an apple and an orange. The main effect of the added roughness is that the method is slightly more prone to error at small zenith angles. Note that, in terms of computation time (as opposed to practical flexibility or the number of views needed), this disambiguation method is by far the most efficient presented in this thesis.

4.3.2 Practical Considerations

The remainder of this section discusses the experimental arrangement needed for the technique. In particular, we consider how accurate the angles, θ_L and α_L, and distances to the light source, D_L, need to be for the method to work. The simple answer is that the values must only be approximately correct. Note that Miyazaki et al. [67] attempt to estimate the light source directions using specularities, a technique that could potentially be used to improve this technique.

The first point to note is that the technique is possible with only two light sources. These could be placed, for example, where sources 1 and 3 lie in Figure 4.8. The weakness of this arrangement is that for areas where $\alpha \approx 0$ or $180°$, the intensity from the light sources are almost equal and so the technique would be critically dependent on an accurate set-up. With three light sources, the technique is robust to both inaccurate positioning of light sources and to noise, as explained below. It is therefore unnecessary to include more than three light source positions.

In our non-retro-reflection experiments, we used a lightweight 30W filament lamp ($\sim$1cm diameter) with a condenser lens attached to focus the light onto the object. For the results presented here (Figures 4.10 and 4.11), we chose $\theta_L = 45°$ and $D_L = 1.5$m. For the light source azimuth angles, we used $\alpha_L^{(1)} = 270°$, $\alpha_L^{(2)} = 0$ and $\alpha_L^{(3)} = 90°$, as shown in Figure 4.9. We also found that using $\theta_L = 60°$ gave similar results. Although a more even distribution of the light sources is possible (i.e. separating the azimuth angles of the light sources by $120°$ instead of $90°$), it is not always practical to place light sources below the object. For this reason, we distribute the light sources as fully as possible, while keeping each one above, or level with, the object.

So far we have assumed that the light source angles, α_L and θ_L, and the distances, D_L, are fixed and known. In fact, these is room for uncertainty in these values, as we now show. It is impossible to exactly calculate the effects of source positioning errors because they depend upon the BRDF of the material under study. However, we can approximate the effects theoretically by assuming that the BRDF follows the Wolff diffuse reflectance model. Consider the cases where two of the light sources are fixed at the correct positions and an error is introduced to each of the three parameters (α_L, θ_L and D_L) of the third light source in turn.

First, consider the angle α_L. If any of the light sources shown in Figure 4.9 are rotated about the sphere centre by any angle up to $90°$, then Equations 4.15 to 4.17 remain valid. The technique is therefore highly robust to errors in α_L. Of course, if the error is exceptionally large, then the success of the disambiguation will be more dependent on the accuracy of θ_L and D_L.

Next, we examine the criticality of the distances D_L. We stress again that the effects of inaccurate light source distances depend upon the material BRDF, so we only consider the effects at the boundaries of region $\mathtt{A}$ in Figure 4.9 to act as a guide. Whether or not the method is successful for a given error in D_L depends partly on the surface zenith angle.

If L_s is the radiant light intensity leaving a point source, then the light intensity incident on the surface point (irradiance) is proportional to L_s/D_L^2 [35]. The Wolff model (Equation 4.3) predicts that

$$L_r\left(\theta_i, \theta_r, n, D_L\right) \propto \frac{\cos\theta_i}{{D_L}^2}(1-F(\theta_i, n))\left(1-F\left(\arcsin\left(\tfrac{\sin\theta_r}{n}\right), \tfrac{1}{n}\right)\right) \tag{4.18}$$

A similar expression applies to each light source, with the same constant of proportionality. In order for the required condition at the boundaries of region $\mathtt{A}$ to be met ($I^{(2)} > I^{(1)}$) we know that $L_r^{(2)} > L_r^{(1)}$. Substituting L_r from Equation 4.18 for each light source and rearranging, we get

$$\frac{D_L^{(1)}}{D_L^{(2)}} > \sqrt{\frac{\cos\theta_i^{(1)}}{\cos\theta_i^{(2)}} \quad \frac{1-F\left(\theta_i^{(1)}, n\right)}{1-F\left(\theta_i^{(2)}, n\right)}} \tag{4.19}$$

At the left and right boundaries to this region, the unit surface normal vectors for a given zenith angle, θ, are

$$\mathbf{n}^{\text{(left)}} = \begin{pmatrix} \sin\theta \\ 0 \\ \cos\theta \end{pmatrix} \qquad \mathbf{n}^{\text{(right)}} = \begin{pmatrix} \sin\theta/\sqrt{2} \\ \sin\theta/\sqrt{2} \\ \cos\theta \end{pmatrix} \tag{4.20}$$

If the source is not too close to the surface, we can state that the unit vectors from any surface point towards light sources 1 and 2, respectively, are

$$\mathbf{n}_L^{(1)} = \begin{pmatrix} 0 \\ -\sin\theta_L \\ \cos\theta_L \end{pmatrix} \qquad \mathbf{n}_L^{(2)} = \begin{pmatrix} \sin\theta_L \\ 0 \\ \cos\theta_L \end{pmatrix} \tag{4.21}$$

The angles between these vectors and the normal vectors at the region boundaries (given by Equation 4.20) are

$$\theta_i^{(\text{1,left})} = \arccos\left(\cos\theta\cos\theta_L\right) \tag{4.22}$$

$$\theta_i^{(\text{2,left})} = \arccos\left(\cos\theta\cos\theta_L + \sin\theta\sin\theta_L\right) \tag{4.23}$$

$$\theta_i^{(\text{1,right})} = \arccos\left(\cos\theta\cos\theta_L - \frac{\sin\theta\sin\theta_L}{\sqrt{2}}\right) \tag{4.24}$$

$$\theta_i^{(\text{2,right})} = \arccos\left(\cos\theta\cos\theta_L + \frac{\sin\theta\sin\theta_L}{\sqrt{2}}\right) \tag{4.25}$$

We can now substitute these equations into Equation 4.19 to obtain a condition for the ratio $D_L^{(1)}/D_L^{(2)}$ at the left boundary (Equations 4.22 and 4.23) and the right boundary (Equations 4.24 and 4.25). The method would successfully disambiguate an azimuth angle on the boundary provided that the condition is met. Figure 4.12 illustrates the different combinations of surface zenith angles and illumination directions that allow correct disambiguation for two different values of $D_L^{(1)}/D_L^{(2)}$.

Remembering that disambiguation is most important for large zenith angles (since the angular separation between the two possible surface normals is greatest in such cases), it would be reasonable to state the following qualitative result: For $\theta_L = 45^\circ$, as in Figure 4.10, the method gives good overall results provided that the distance between light sources does not vary by more than 10% and reasonable results at 20%.

Finally, consider the required accuracy of the angles, θ_L, along the left hand border of region `A`. This border is more critical than the borders of region `A'`, so gives a good overall picture. In the Wolff model, the incidence and exitance terms in the radiance function are separable. Since the exitance term is the same for both light sources 1 and 2 for a given point on the boundary, we use the fact that $I^{(2)} > I^{(1)}$ (Equation 4.15) to write the inequality

$$\cos\theta_i^{(2)}\left(1 - F\left(\theta_i^{(2)}, n\right)\right) > \cos\theta_i^{(1)}\left(1 - F\left(\theta_i^{(1)}, n\right)\right) \tag{4.26}$$

Each side of this inequality decreases monotonically with θ_i. For this to hold therefore, we can easily derive the condition that $\theta_i^{(2)} < \theta_i^{(1)}$. We can re-write this in terms of surface

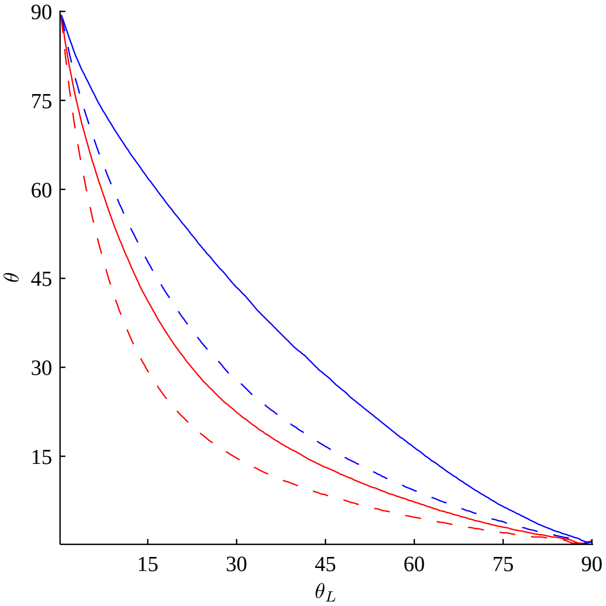

Figure 4.12: Combinations of surface zenith angles, θ, and light source angles, θ_L, that are recovered correctly for $D_L^{(1)}/D_L^{(2)} = 0.8$ (area above the top curve) and 0.9 (above the bottom curve). The solid curves relate to the left boundary, while the broken curves are for the right.

zenith angles and light source zenith angles using Equations 4.22 and 4.23:

$$\cos\theta_L^{(1)} > \cos\theta_L^{(2)} + \tan\theta\sin\theta_L^{(2)} \tag{4.27}$$

Manipulating this inequality numerically, the graph in Figure 4.13 is obtained. The figure shows the combinations of zenith angles, θ, and illumination angles, θ_L, that allow correct disambiguation for different uncertainties in $\theta_L^{(2)}$. A qualitative interpretation of Figure 4.13 is that, for most applications, an accuracy of about $15°$ is required.

In conclusion to this section, we make the following general points. The photometric stereo method discussed in this section provides a simple yet reliable method for azimuth angle disambiguation. It can be justifiably argued that, along with several other photometric stereo methods, the setup is highly specialised and potentially unwieldy. However, the above calculations show that the positions and angles in the rig do not need to be precise for the method to work well.

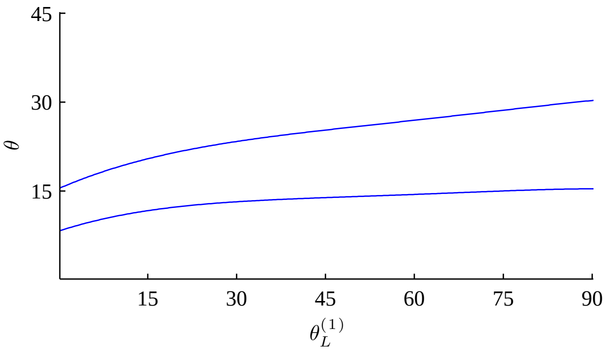

Figure 4.13: Minimum zenith angles at the left boundary of region `A` in Figure 4.9 that are disambiguable allowing for 30° error in $\theta_L^{(2)}$ (top curve) and 15° error (bottom curve).

4.4 Two-dimensional Reflectance Function Estimation

In this section, we extend the retro-reflectance function estimation technique, described in Section 4.2, to more general illumination directions. Using the photometric stereo method described in Section 4.3, we have a set of Cartesian triples $\mathcal{D} = \{(\theta_d, \alpha_d, I_d)\,; d = 1, 2...|\mathcal{D}|\}$. Our aim is to fit a two dimensional BRDF, $f\,(\theta, \alpha)$, to this data in order quantify the dependence of the image intensity on the surface zenith and azimuth angles. Technically, we are actually fitting a *radiance* function, $L_r\,(\theta, \alpha)$, to the data as before, which differs from the BRDF by a factor of $\cos\theta$ and a multiplicative constant. In doing so, we acquire a representation of a slice of the BRDF in the form of a discretely sampled radiance function, $\hat{L}_r = \left\{(\theta_k, \alpha_k, I_k)\,; k = 1, 2, ...|\hat{L}_r|\right\}$.

In our proposed method, we extend the 2D zenith-intensity histogram to the 3D zenith-azimuth-intensity histogram. We then fit a surface, representing the radiance function, to the histogram data using simulated annealing. The technique of simulated annealing [91] is a stochastic optimisation scheme. It attempts to minimise an energy functional by repeatedly forming trial solutions by making random changes to the current estimate. Each trial solution has an associated energy and can be accepted or rejected as the new current estimate according to a probability. This probability depends on the energy difference between the current estimate and the new trial solution.

Our proposed reflectance function estimation algorithm fits the surface by minimising an energy functional consisting of two terms. The first term is directly related to the histogram. The second is a form of elastic energy, which is based on the local curvature of the fitted surface and ensures a smooth BRDF. Due to the sparseness of data in the 3D histogram, compared to the 2D case, we do not attempt to find the outer envelope here since this may compromise robustness.

Simulated annealing is well suited to this task because, with sufficient care and attention to details such as the form of the energy functional and the cooling schedule, the algorithm should return a *globally* optimum solution. Also, as will become clear shortly, the energy functional used can be primarily based on a directly interpretable quantity (namely, the histogram frequencies). The downsides to simulated annealing are that inconsistent results are obtained for certain applications and convergence is often slow. The latter point however, may be addressed in future work using *mean field annealing* [76].

Before presenting the details of our method, we provide a brief overview of simulated annealing algorithms in general [91]. In Section 4.4.2, the three-dimensional histogram is formally defined, and it is shown how this leads to an energy field, which is to contain the fitted surface. Section 4.4.3 explains how the surface is to be represented within the energy space and how it is stochastically perturbed to form new trial states. The elastic energy is introduced in Section 4.4.4, which also describes how the total energy of a particular solution is calculated. Pseudocode for the algorithm is provided at the end of these sections. Finally, a set of experiments are described in Section 4.4.5 and fine tuning issues are discussed in Section 4.4.6.

4.4.1 A General Simulated Annealing Algorithm

Simulated annealing was originally inspired by the ability of molecules within a metal to tend towards their ground state as the sample is gradually cooled (annealed) from a high temperature towards a frozen state at a low temperature. In this original model [62] a thermodynamic system at a certain initial temperature in an arbitrary *microstate* is perturbed. In other words, a small change is made to the energy state of one or more molecules. This perturbation may result in a net increase or a net decrease in the total energy of the system. If the perturbation results in an energy decrease then the system moves into the new state. If the energy is increased by the perturbation, then the new state is accepted with a certain probability, which reduces with temperature. This process is repeated as the temperature is reduced, resulting in the system tending towards a global minimum energy state.

This scheme was generalised to combinatorial problems in 1980s by Kirkpatrick et al. [48] and Černý [10]. The technique can be stated generally as follows. Given a configuration space, a system state[2], f, and an objective function, $\varepsilon(f)$, simulated annealing finds the optimum state f_{opt} that minimises $\varepsilon(f)$. That is, the algorithm searches for $f_{\text{opt}} = \underset{f}{\operatorname{argmin}}\, \varepsilon(f)$. By analogy to the thermodynamic model, the current solution to the optimisation, f, relates to the current microstate of the sample, while the objective function, ε, takes the place of the system's internal energy.

Let the current state of the system be f and the so-called "neighbour state" (i.e. the state after the perturbation) be f'. The form of the perturbation depends upon the nature of the problem. The probability of accepting this new state is then given by the Boltzmann probability:

$$q_{\mathrm{B}} = \begin{cases} 1 & \Delta\varepsilon \le 0 \\ \exp\left(-\Delta\varepsilon/T\right) & \Delta\varepsilon > 0 \end{cases} \tag{4.28}$$

where $\Delta\varepsilon = \varepsilon(f') - \varepsilon(f)$. Alternative acceptance probabilities have also been suggested that are more appropriate for certain specific problems [91].

The reason for the success of simulated annealing is the fact that the energy of the system is allowed to increase on certain occasions. This prevents the system from becoming pinned to local minima. As the temperature is reduced, the probability of transition into higher energy states decreases, so that the system eventually settles. In fact, Aarts and Korst [1] were able to show that, if the temperature is reduced logarithmically and sufficiently slowly, then simulated annealing is guaranteed to find the global minimum.

---

[2] We adopt the symbol "f" since, for our case, the state is to represent a radiance function, which is intimately related to the BRDF.

Unfortunately, the amount of time needed for simulated annealing is often impractical and a compromise between accuracy and efficiency must be made. Indeed, the choice of the initial temperature and the cooling schedule can be of critical importance. Usually, the temperature is reduced by a fixed factor after each trial state, or after a fixed number of trials. If the factor is too small however (*quenching*), then the system may only converge to a *local* minimum. If it is too high on the other hand, then convergence will obviously take unacceptably long. The next few sections describe how the various ingredients of the simulated annealing algorithm are adapted to our needs in BRDF estimation.

4.4.2 Energy Field for Reflectance Function Calculation

The technique must fit a surface, representing a 2D slice of the BRDF, to 3D histogram data. The raw data for the histogram comes from the Cartesian triples, $\mathcal{D}$, acquired from the photometric stereo method discussed in Section 4.3. Simulated annealing is used by repeatedly perturbing the surface inside an "energy field" derived from the histogram. Clearly, regions of the histogram with high frequencies should have an associated low energy. We now detail the simple steps necessary to acquire this energy for each bin of the histogram.

It is possible to increase the data density in the histogram by noting that for isotropic surfaces

$$f\left(\theta, \alpha_L - \Delta\alpha\right) = f\left(\theta, \alpha_L + \Delta\alpha\right) \tag{4.29}$$

where $\Delta\alpha \in [0, 180^\circ)$ is the absolute difference between the light source azimuth angle and the surface azimuth angle. Figure 4.14 clarifies the definition of $\Delta\alpha$. We are essentially, exploiting the assumed isotropic nature of the surface reflection to increase the robustness of the surface fitting. With this in mind, we replace our α values in $\mathcal{D}$ with the reduced azimuth angles given by

$$\Delta\alpha = \min\left(|\alpha - \alpha_L|,\ 360^\circ - |\alpha - \alpha_L|\right) \tag{4.30}$$

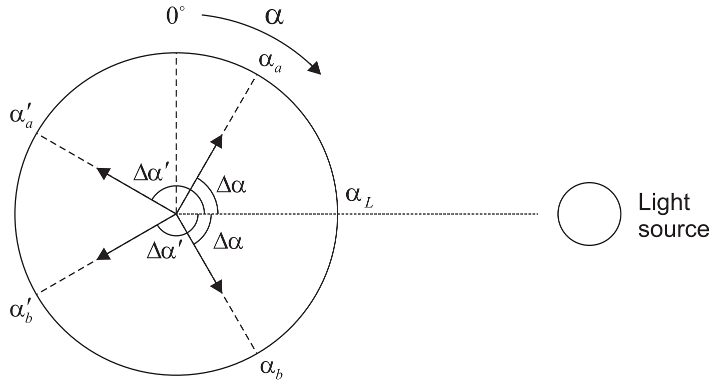

Figure 4.14: A light source illuminating a point. The observed intensity if the azimuth angle of the surface normal is α_a, should be the same as if the angle is α_b. Therefore, the angle $\Delta\alpha$ is sufficient to calculate the BRDF. Angles α'_a and α'_b illustrate another example.

Another advantage of this step is that $\Delta\alpha$ can be calculated at each pixel from all three views used in the photometric stereo stage of the technique and used to form the histogram. The density of the data in the histogram therefore increases sixfold. A twofold increase is due to the application of Equation 4.30, which reduces azimuth data from the interval $[0, 360°)$ to $[0, 180°)$. A further threefold increase is due to the use of three polarisation images.

Using a similar formulation to that used in Section 4.2, the histogram contents are given by $\mathcal{HC}_{a,b,c} = |\mathcal{BC}_{a,b,c}|$, where

$$\mathcal{BC}_{a,b,c} = \left\{ (\theta_d, \Delta\alpha_d, I_d)\,;\ \begin{array}{rcl} \theta_d & \in & \left[\theta_a^{(\text{bin})} - \frac{90°}{2k_{\text{max}}},\ \theta_a^{(\text{bin})} + \frac{90°}{2k_{\text{max}}}\right) \wedge \\ \Delta\alpha_d & \in & \left[\Delta\alpha_b^{(\text{bin})} - \frac{180°}{2k_{\text{max}}},\ \Delta\alpha_b^{(\text{bin})} + \frac{180°}{2k_{\text{max}}}\right) \wedge \\ I_d & \in & \left[I_c^{(\text{bin})} - \frac{255}{2k_{\text{max}}},\ I_c^{(\text{bin})} + \frac{255}{2k_{\text{max}}}\right) \end{array} \right\} \tag{4.31}$$

where $\{\theta_a^{(\text{bin})}; a = 1, 2, ...k_{\text{max}}\}$, $\{\Delta\alpha_b^{(\text{bin})}; b = 1, 2, ...k_{\text{max}}\}$ and $\{I_c^{(\text{bin})}; c = 1, 2, ...k_{\text{max}}\}$ are the equally spaced bin centres for zenith angles, reduced azimuth angles and intensities respectively. Since there are more bins in the 3D histogram than the 2D case for a given k_{max}, we use $k_{\text{max}} = 32$ for this section (recall that $k_{\text{max}} = 100$ was used in Section 4.2). This avoids having too many sparsely populated bins, and is another measure to maximise the robustness of the surface fitting. Figure 4.15a shows several slices of this histogram for the porcelain vase from Figure 3.10a. We use normalised units of bin lengths from

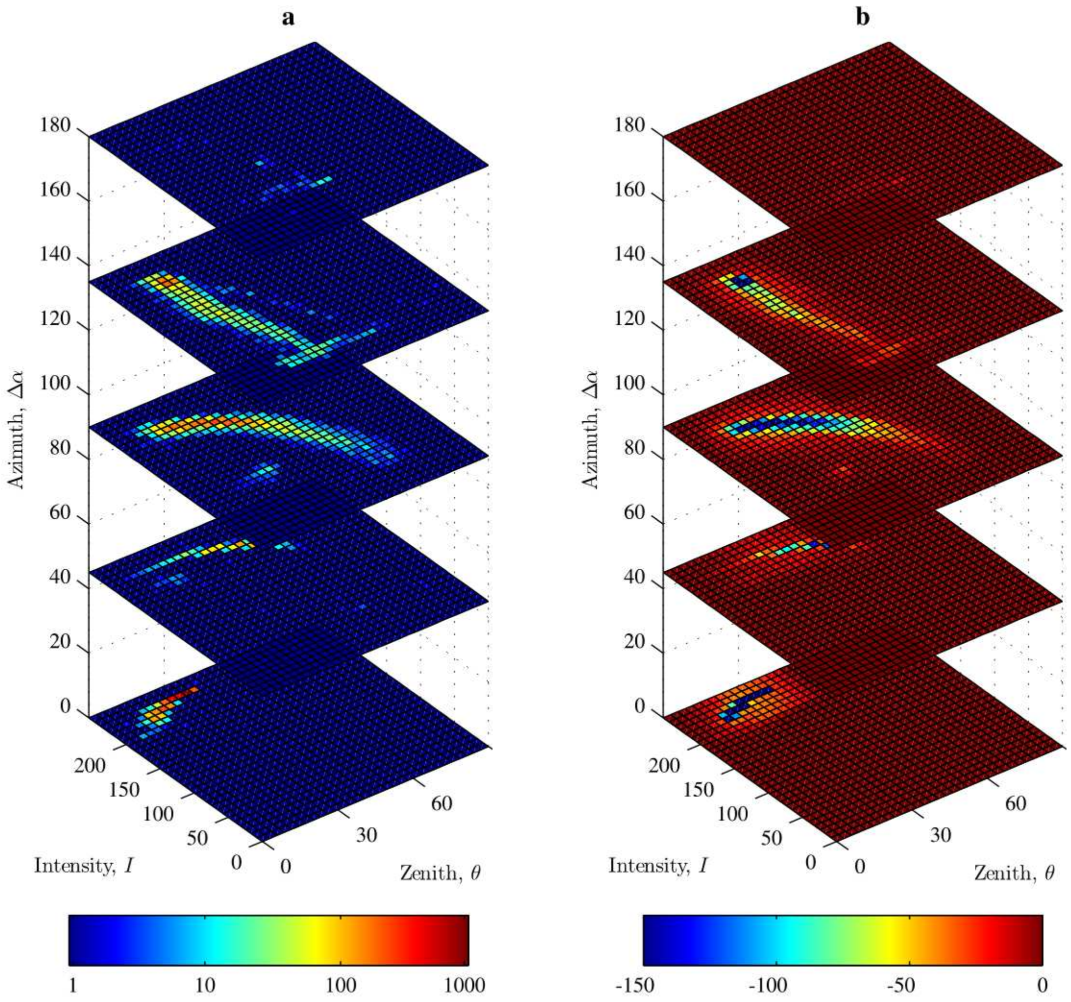

Figure 4.15: (a) A selection of slices of the 3D histogram of the porcelain vase using $\alpha_L = 0$ and $\theta_L = 45°$ (logarithmic scale). (b) Similar slices of the energy field (linear scale, with energies below -150 not shown for reasons of clarity).

this point on. That is, $\theta \leftarrow \theta k_{\text{max}}/90°$, $\Delta\alpha \leftarrow \Delta\alpha k_{\text{max}}/180°$ and $I \leftarrow I k_{\text{max}}/255$.

We wish to convert the histogram into an energy field using as simple a means as possible in order to preserve a direct a relation between the histogram frequency and the energy. The following equation is used for this task:

$$\varepsilon = -HC * \Xi - HC \tag{4.32}$$

where Ξ is a three-dimensional Gaussian kernel and $*$ denotes the convolution operation. The histogram contents, HC, are negated because higher frequencies should be associated with lower energies. This means that the maximum allowed value at a point in the discrete energy field is zero. Taking the reciprocal is another option, but this blurs the direct relationship between energy and the histogram. The Gaussian convolution is used to smooth the data. We use a high Gaussian standard deviation (≈ 5 bin lengths) so that a large portion of the energy field is non-zero. This will ultimately aid the simulated annealing algorithm in its search for a global minimum. Finally, the original histogram frequencies are subtracted from this smoothed version. This ensures that the energy field corresponding to non-zero histogram contents have significantly lower energy than surrounding areas. Figure 4.15b shows an example of the energy field generated in this way.

4.4.3 Initial State and Move Class

To fit a surface to the energy field described above using simulated annealing, we require an initial state for the surface and a means to perturb the surface (a move class). It was found that, for the completed algorithm, there was virtually no dependance on the initial estimate. We therefore simply used a plane as the initial state. In the normalised coordinates, the plane has corners at $(\theta, \Delta\alpha, I) = (0, 0, k_{\text{max}})$, $(k_{\text{max}}, 0, 0)$, $(0, k_{\text{max}}, k_{\text{max}})$ and $(k_{\text{max}}, k_{\text{max}}, 0)$. The plane is sampled at k_{max} equally spaced points for a given $\Delta\alpha$ and at k_{max} equally spaced values of $\Delta\alpha$, as shown in Figure 4.16a. This means that the number of points in the initial state is $k_{\text{max}}^2 = 32^2 = 1024$. This number is fixed throughout the annealing process.

Each perturbation is performed as follows. Let the Cartesian position of the discretely sampled points be $\hat{f} = \{(\theta_k, \Delta\alpha_k, I_k), k = 1, 2, ...k_{\text{max}}^2\}$ in the current state and $\hat{f}' = \{(\theta'_k, \Delta\alpha'_k, I'_k), k = 1, 2, ...k_{\text{max}}^2\}$ in the perturbed state. To select the state $\hat{f}'$ given $\hat{f}$, we wish to disturb the surface such that a random point is chosen as the epicentre of the perturbation. Surrounding points are then disturbed by an amount that decreases with the distance to the epicentre. The index of the random point is an integer chosen from a uniform distribution in the interval $[1, k_{\text{max}}^2]$. An example of a perturbation is shown in Figure 4.16b.

We choose the perturbation is to be Gaussian-like. That is, the point chosen as the

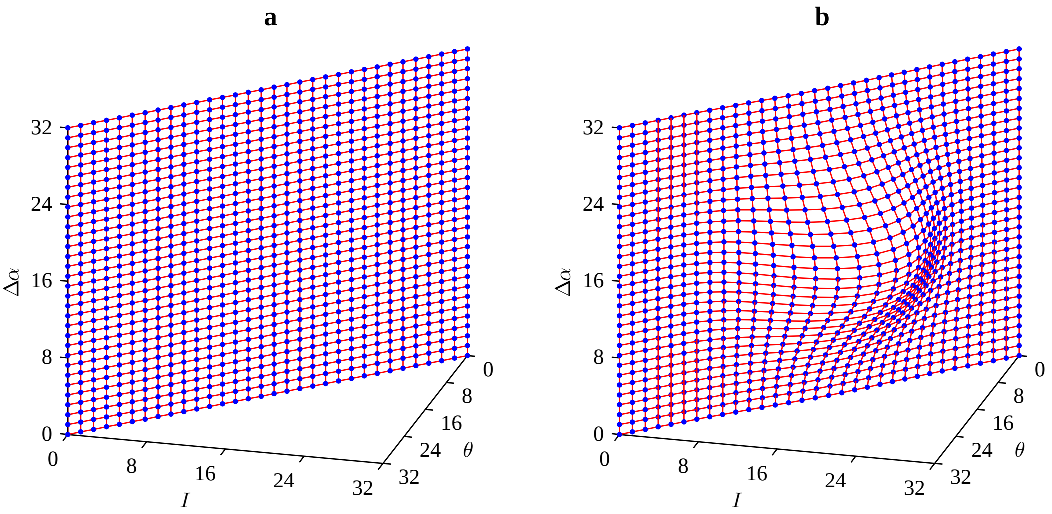

Figure 4.16: (a) Using a plane as the initial state. (b) State after a perturbation.

epicentre, $(\theta_{\text{epi}}, \Delta\alpha_{\text{epi}}, I_{\text{epi}})$, moves the greatest distance, Υ_{epi}, and surrounding points move by an amount that decays as a Gaussian. Consider the motion of the epicentral point. To preserve the stochastic nature required for simulated annealing, we let the magnitude of the perturbation, Υ_{epi}, also be chosen from a Gaussian distribution. We restrict the motion of any point, k, (at the epicentre or otherwise) to be along the straight line that connects the points $(\theta_k, \Delta\alpha_k, I_k)$ and $(0, \Delta\alpha_k, 0)$. In other words, points on the surface can only move *directly* towards the $\theta = I = 0$ line or directly away from it. This ensures that points remain evenly spread across the surface.

The Gaussian PDF used to select Υ_{epi} has zero mean. When $\Upsilon_{\text{epi}} < 0$, the perturbation is towards the $\theta = I = 0$ line, while $\Upsilon_{\text{epi}} > 0$ indicates the opposite. The standard deviation is chosen such that a sufficient number of large moves are possible to avoid convergence to a local minimum. Using any value between $k_{\text{max}}/20$ and $k_{\text{max}}/10$ seemed reasonable for all our experiments. If the epicentral point moves a distance Υ_{epi}, then the other points move by a distance that reduces as the Gaussian of the distance to the epicentre, which is given by

$$\Upsilon_k = \Upsilon_{\text{epi}} \exp\left(-\frac{(\theta_{\text{epi}} - \theta_k)^2 + (\Delta\alpha_{\text{epi}} - \Delta\alpha_k)^2 + (I_{\text{epi}} - I_k)^2}{2\sigma_{\text{area}}^2} \right) \tag{4.33}$$

where σ_{area} is the Gaussian standard deviation (which is given the symbol σ_{area} because the

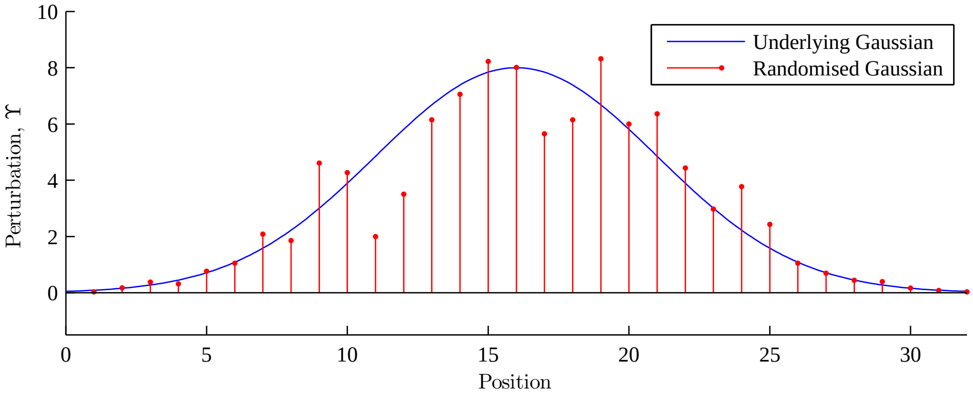

Figure 4.17: Comparison of our current perturbation method (where distances, Υ are taken directly from the underlying Gaussian) and the more randomised proposed method (where this gaussian is rescaled at discrete intervals).

area of the surface that is perturbed depends on it). We found that using $\sigma_{\text{area}} = k_{\text{max}}/10$ seemed to give best results. Using too small a value for σ_{area} sometimes results in a rough final surface. Although the exact values of these Gaussian standard deviations are not exactly critical, there is unfortunately some dependence on the results. However, we were able to use the same values for each experiment without loss of quality.

Note that, for a particular perturbation in our method, once the epicentre and magnitude have been randomly chosen, the perturbation is deterministic. Ideally, the perturbations should be as random as possible so all states are accessible. A better method would therefore use the Gaussian deterministic method in Equation 4.33 only as a basis for the perturbation. The distances Υ_k would then be rescaled by random amounts at discretely sampled points on an underlying Gaussian. Figure 4.17 gives an example of such a resampled non-deterministic Gaussian perturbation for a single slice of the initial state. We have not implemented the non-deterministic version to date however, since we feel that the advantages do not justify the extra computational expense.

4.4.4 Energy of State

We now have the configuration space and move class required for simulated annealing. The next ingredient is a means to calculate the energy of a particular state, $\varepsilon(\hat{f})$. As stated above, the energy functional that we use has two terms: the "bin contents energy," $\varepsilon_{\text{bc}}(\hat{f})$, derived from the histogram, and the "elastic energy," $\varepsilon_{\text{el}}(\hat{f})$, used to maintain a smooth surface. The former of these can be taken directly from the energy field discussed above: each point of the current state $\hat{f}_k$ has a bin contents energy taken as the energy associated with the histogram cell in which $\hat{f}_k$ is located. The sum of these energies for all points in the state is the total bin contents energy. Using this as the total energy state, without the introduction of the elastic term, we found that some areas of the fitted surfaces are invariably and unrealistically irregular. This did not prove too severe for areas of the histogram with high frequencies, but was much more problematic where data were more sparse.

In general, the elastic energy of a surface (or manifold) quantifies its bending energy and is based on curvature. Larger curvatures result in higher elastic energies, thus promoting smooth solutions to the problem being solved. The mathematics of elastic energy minimisation, or *elastica* [9], was originally developed and solved by Euler in 1744 [25]. The problem that he posed was to find plane curves, C, that minimise the integral

$$\varepsilon_{\text{el}}\left(C\right) = \int_C \left(a\kappa_1^2 + b\right) \mathrm{d}s \tag{4.34}$$

where a and b are constants, κ_1 is the curvature and $\mathrm{d}s$ is the differential unit of length. Horn [39] and Mumford [69] studied the mathematics of such curves for computer applications. One popular related application to computer vision was due to Kaas, Witkin and Terzopoulos [47], who add an additional "external" force term from an intensity image to generate "snakes". These snakes can be used to reconstruct edges that are partly occluded or broken.

The extension of Equation 4.34 to a surface, S, embedded in $\mathbb{R}^3$ is [42]

$$\varepsilon_{\text{el}}\left(S\right) = \int_S \left(a\kappa^2 + b\right) \mathrm{d}A \tag{4.35}$$

where κ is the mean curvature and $\mathrm{d}A$ is the differential unit of surface area. The second term in this integral results in a multiple of the area. We experimented with different values of b for our work but found that there was negligible overall impact. We therefore set $a = 1$ and $b = 0$, so our elastic energy becomes

$$\varepsilon_{\text{el}}\left(S\right) = \int_S \kappa^2 \mathrm{d}A \tag{4.36}$$

To calculate the integral in Equation 4.36, a method is needed to approximate the mean curvature at any surface point. Denote the location of point $\hat{f}_k$ by $\mathbf{f}_k$, a vector from the origin of the energy field. Given the discrete nature of our problem, the curvature, κ_k, at point k, is assumed to be constant over the area, A_k, of the barycell that contains the point. A barycell, as shown in Figure 4.18, is defined as to connect the midpoints of the "bonds" between $\mathbf{f}_k$ with its eight neighbours and the barycentres (centres of mass) of the triangles formed by these neighbours. Consider the point at location $\mathbf{f}_k = \left(\theta_k, \Delta\alpha_k, I_k\right)^T$. Let the point be bonded to eight neighbours, $\{\mathbf{f}_l(k); l \in \text{Nei}(k)\}$. The neighbours, Nei$(k)$, of a particular point, k, are fixed throughout the annealing process, and are defined as the eight closest points at the *initial state*. Fixing the bonds in this manner means that it is unlikely for any two points that are initially close together, to drift far apart at low temperatures.

It can be shown [63] that the mean curvature vector, $\mathbf{K}_k$, of a surface at point k can be calculated by

$$\mathbf{K}_k = \frac{1}{2A_k^{\text{(tot)}}} \sum_{l \in \text{Nei}(k)} \left(\cot\vartheta_{kl}^{-} + \cot\vartheta_{kl}^{+}\right)\left(\mathbf{f}_k - \mathbf{f}_l\right) \tag{4.37}$$

where $A_k^{\text{(tot)}}$ is the barycell area for point k and the angles ϑ_{kl}^{-} and ϑ_{kl}^{+} are defined in Figure 4.18. The mean curvature is then simply

$$\kappa_k = \frac{|\mathbf{K}_k|}{2} \tag{4.38}$$

Equation 4.37 assumes that the mean curvature vector is constant over a cell area. Meyer et al. [63] use the Voronoi cell for this purpose as it perfectly tessellates over a triangulated mesh. To make the algorithm more efficient however, we use the barycentric cell area as an approximation to this, as in Figure 4.18.

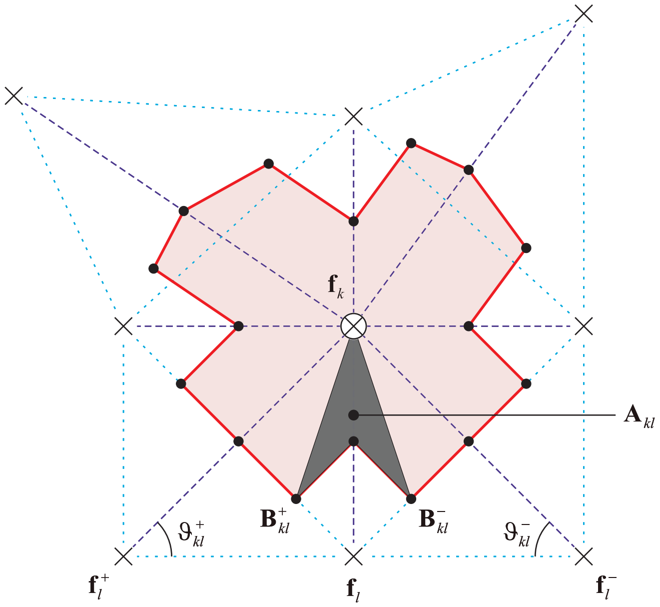

Figure 4.18: Example of a barycell (large shaded area) surrounding point $\mathbf{f}_k$. The eight neighbouring points are $\{\mathbf{f}_l\}$, which are here bonded to $\mathbf{f}_k$ by dark broken lines (the light broken lines are used to bond other points). The barycentres, B^-_{kl} and B^+_{kl}, chevron area, A_{kl} and angles, ϑ^-_{kl} and ϑ^+_{kl}, associated with point $\mathbf{f}_l$ are indicated.

The barycentric cell area is calculated as follows. For point $\mathbf{f}_k$, we have the set of neighbours, $\{\mathbf{f}_l\}$. The barycell can be divided into a set of chevron shapes for each l, one of which is highlighted in Figure 4.18. If the area of each chevron is A_{kl}, then the total cell area is clearly $A^{\text{(tot)}}_k = \sum\limits_{l \in \text{Nei}(k)} A_{kl}$.

The location of the barycentres of the triangles associated with point k are

$$\mathbf{B}^+_{kl} = \frac{\mathbf{f}_k + \mathbf{f}_l + \mathbf{f}^+_l}{3} \qquad \mathbf{B}^-_{kl} = \frac{\mathbf{f}_k + \mathbf{f}_l + \mathbf{f}^-_l}{3} \tag{4.39}$$

where $\mathbf{f}^+_l$ refers to the neighbouring point that appears next in a clockwise direction and $\mathbf{f}^-_l$ is in the anti-clockwise direction around $\mathbf{f}_k$. Using the vector product, we can write

$$A^{\text{(tot)}}_k = \frac{1}{4} \sum_{l \in \text{Nei}(k)} \left(\left| (\mathbf{f}_l - \mathbf{f}_k) \times \left(\mathbf{B}^-_{kl} - \mathbf{f}_k\right) \right| + \left| (\mathbf{f}_l - \mathbf{f}_k) \times \left(\mathbf{B}^+_{kl} - \mathbf{f}_l\right) \right| \right) \tag{4.40}$$

Collectively, Equations 4.37 to 4.40 allow for the calculation of a set of mean curvatures for each point in $\hat{f}$. The elastic energy of a state is then found using Equation 4.36 where the differential area is the barycell.

For the total energy of a state, we use

$$\varepsilon(\hat{f}) = \varepsilon_{\mathrm{bc}}(\hat{f}) \left(\varepsilon_{\mathrm{el}}(\hat{f}) + 1\right) \tag{4.41}$$

This form is adopted since $\varepsilon_{\mathrm{el}}$ has a minimum allowed value of zero, which corresponds to a plane. The elastic term can therefore be regarded as a multiplier that has no effect for the simplest surface (the plane) but increases with roughness.

The main ingredients of the simulated annealing algorithm are now in place. The configuration space consists of a 2-manifold in an $\mathbb{R}^3$ energy field formed from the zenith-azimuth-intensity histogram. The initial state is unimportant so is simply taken as a plane. The move class is a random Gaussian perturbation. Finally, the energy of a state $\varepsilon(\hat{f})$, is calculated using Equation 4.41.

One final point to note is that when the algorithm calculates $\Delta\varepsilon = \varepsilon(\hat{f}') - \varepsilon(\hat{f})$ using Equation 4.41 the total energy of state for the *entire surface* need not be calculated. Instead, the algorithm calculates the energy of the area of the trial surface within a distance of $2\sigma_{\mathrm{area}}$ to the epicentre of the perturbation. The energy of the corresponding region of the current surface is then subtracted from this energy.

Pseudocode

The algorithm described in this section is summarised by Pseudocode 2. The main equations needed are referred to where necessary. The function "*rand*" generates a random number between 0 and 1. Please note that a few of the lines have not yet been discussed, but appear later in Section 4.4.6.

4.4.5 Experiments

This section presents results from the various stages of the algorithm for the porcelain vase and bear using the illumination configuration shown in Figure 4.8. Some rendered images are then shown using the estimated BRDFs for various materials. In fact, an implementation of the above algorithm directly gave poor results, so it was necessary to

Pseudocode 2 Two-dimensional BRDF estimation.

```
Use photometric stereo to acquire $\mathcal{D} = \{(\theta_d, \alpha_d, I_d)\}$ (Section 4.3)
Replace $\alpha$ with $\Delta\alpha$ (Equation 4.30)
Calculate histogram (Equation 4.31) and energy field (Equation 4.32)
Generate initial state (Figure 4.16a)
while [solution not converged] do
    Perturb surface to obtain trial state (Equation 4.33)
    Calculate energy difference, $\Delta\varepsilon$, between states (Equation 4.41)
    Calculate transition probability, $q_{\mathrm{B}}$ (Equation 4.28)
    if [rand < $q_{\mathrm{B}}$] then
        Accept trial state
        Apply shadow constraint (Equation 4.43)
    Reduce temperature
Apply smoothness constraint (Equations 4.44 and 4.45)
```

make a few light assumptions about the BRDF. For example, we enforce the constraint that the BRDF is zero where the angle between the surface normal and the light source direction is greater than 90° (i.e. in shadow). To avoid confusing the reader, we discuss these matters after the results.

First consider Figure 4.19, which depicts a surface generated after several perturbations at high temperature. Points on the surface decorated with circles are those in areas with high mean curvature. Larger circles correspond to larger curvatures and, hence, higher elastic energies.

Figure 4.20 shows the state, $\hat{f}$, of the surface, at various temperatures during the annealing process when applied to the porcelain vase. The large indentation to the upper-right of the initial state is due to the shadow boundary constraint, which will be explained in Section 4.4.6. The temperature, T, was reduced by a factor of 0.999 after each trial state and the surface is shown for each order of magnitude of T. The algorithm took approximately three minutes to run.

The pie charts in Figure 4.20 show the distribution of transition probabilities between each decrement in the order of magnitude of temperature. In this case, the pie charts indicate a well-initialised execution, since initially about half of the transitions are accepted with unit probability, and the remainder probabilities given by the Boltzmann distribution. As the temperature is lowered and the surface begins to converge, the proportions

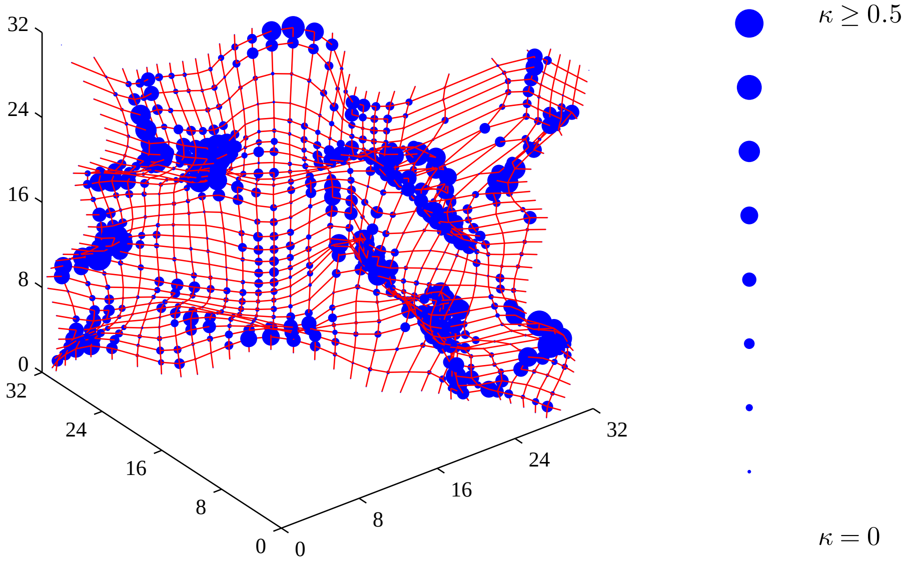

Figure 4.19: A rough arbitrary surface with the mean curvature indicated by the dot size at each point. Values were calculated using Equations 4.37 to 4.40.

change until, eventually, most of the probabilities are approximately zero. In fact, we used this as a convergence criterion: *when the fraction of transition probabilities equal to zero* (< 0.01) *exceeds 75% as the temperature drops by an order of magnitude, then the surface has converged.*

A simple monotonicity constraint (explained in Section 4.4.6 below) was then applied to the recovered radiance functions for the vase and bear to give the final estimate of the radiance function. Figure 4.21 shows some spheres that were rendered using the radiance functions calculated from the aforementioned objects at both $\theta_L = 45°$ and $\theta_L = 60°$. All rendered images have the light source placed at $\alpha_L = 0$. The images look realistic for the top half of the sphere (i.e. where $\Delta\alpha \leq 90°$), but tend to break down close to the shadow boundary. Extra constraints on the form of the BRDF are perhaps needed to address this issue. For comparison, Figure 4.22 shows spheres rendered under the same illumination conditions using the Wolff model.

Figure 4.23 shows a comparison of the intensities predicted by the estimated radiance functions, the Wolff model and ground truth (measured using a cylinder). The total diffuse

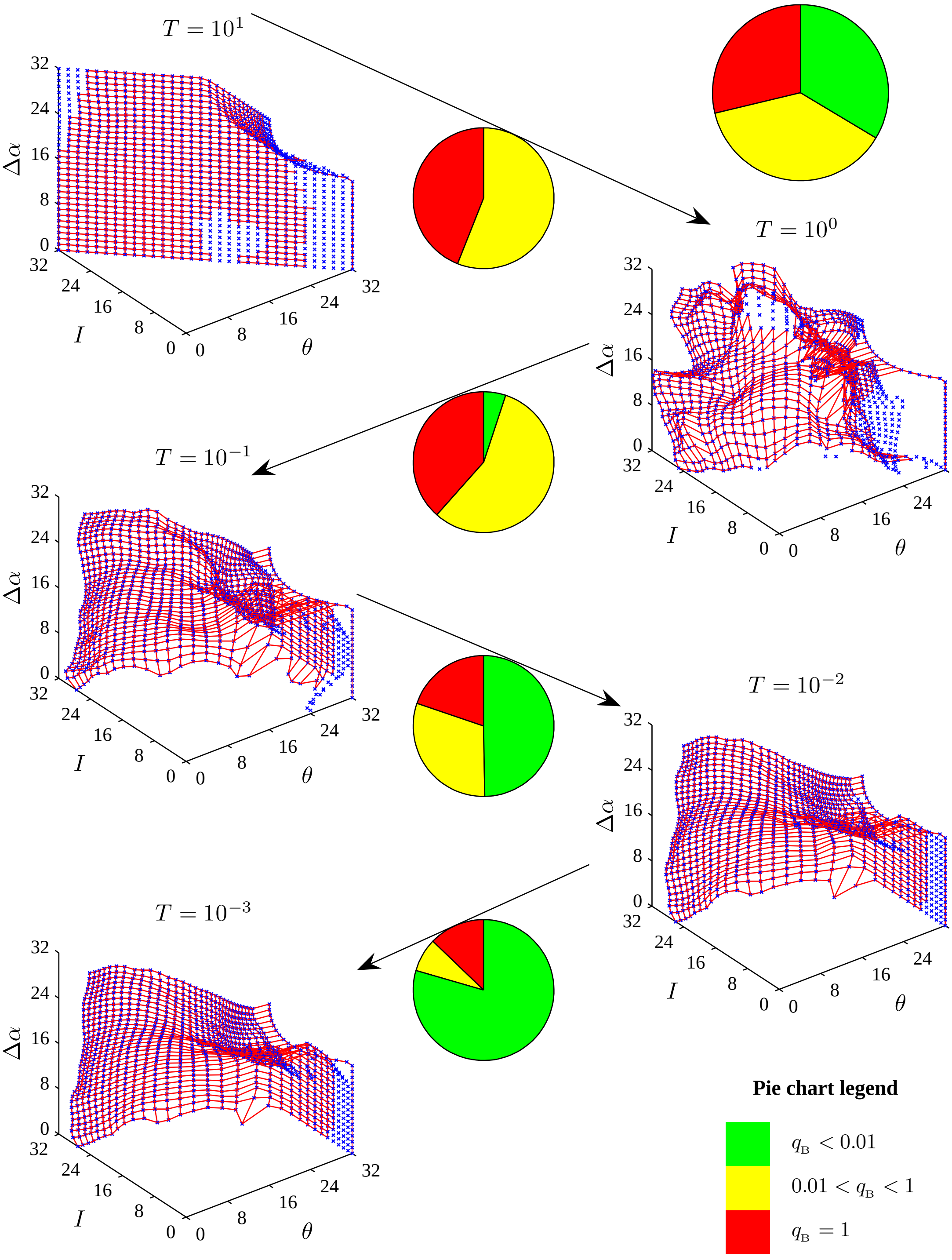

Figure 4.20: Annealing progression from an initial state at $T = 10$ (top) to a final state at $T = 0.001$ (bottom). Points not connected by lines are at locations with zero bin contents energy. The pie charts show the proportions of transition probabilities that are 1, (approximately) 0 and intermediate values for all transitions between each temperature shown. The overall proportions are shown in the top-right chart.

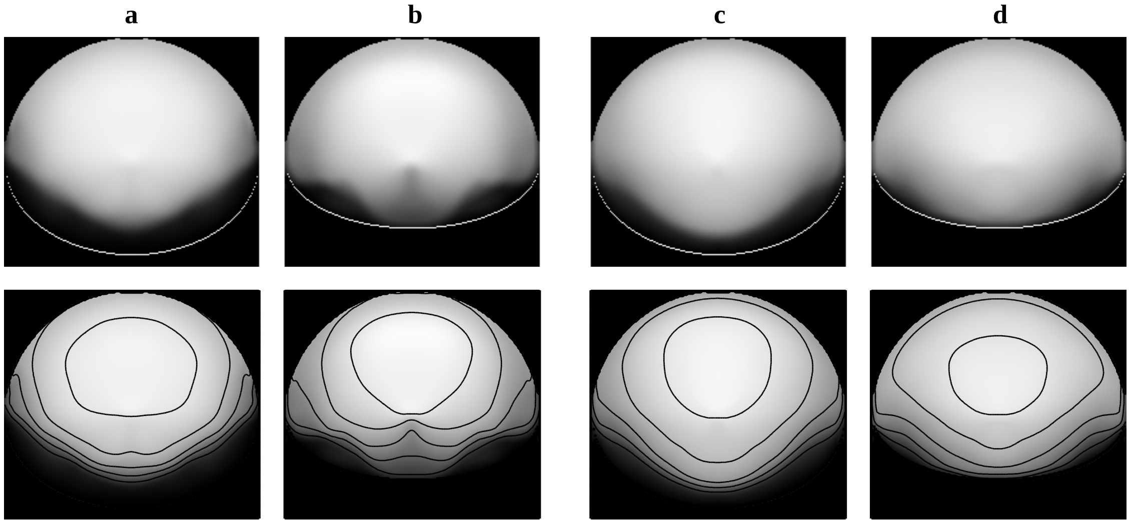

Figure 4.21: Top row: Spheres rendered using the BRDFs estimated from (a) the vase with $\theta_L = 45°$, (b) the vase with $\theta_L = 60°$, (c) the bear model with $\theta_L = 45°$, and (d) the bear model with $\theta_L = 60°$. The white lines indicate the calculated shadow boundary. Bottom row: same rendered spheres with black lines indicating intensity grey levels of 50, 90, 130, 170 and 210.

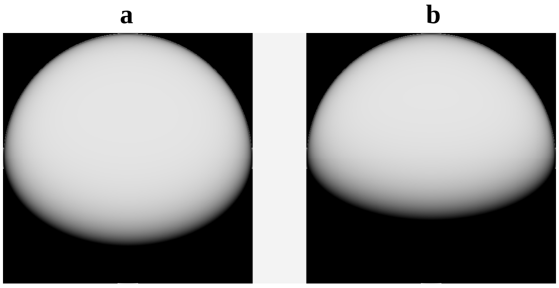

Figure 4.22: Spheres rendered using the Wolff model with (a) $\theta_L = 45°$ and (b) $\theta_L = 60°$, as in Figure 4.21.

albedo, ϱ, used in the Wolff model (Equation 4.3) was selected to give best agreement with the exact curves. The figure shows some good results and others that are reasonable, but less impressive. Combining data from both objects is a possibility to improve accuracy further.

The final experiments for this chapter involved using the 2D reflectance function estimation algorithm for colour images. This was done by applying the algorithm independently to each of the red, green and blue colour channels. One problem encountered with the *independent* estimations for each colour channel was that the colour balance present in the raw images did not always match that in the recovered BRDFs. A simple method we adopted to counter this issue was to linearly rescale the colour components of the BRDFs.

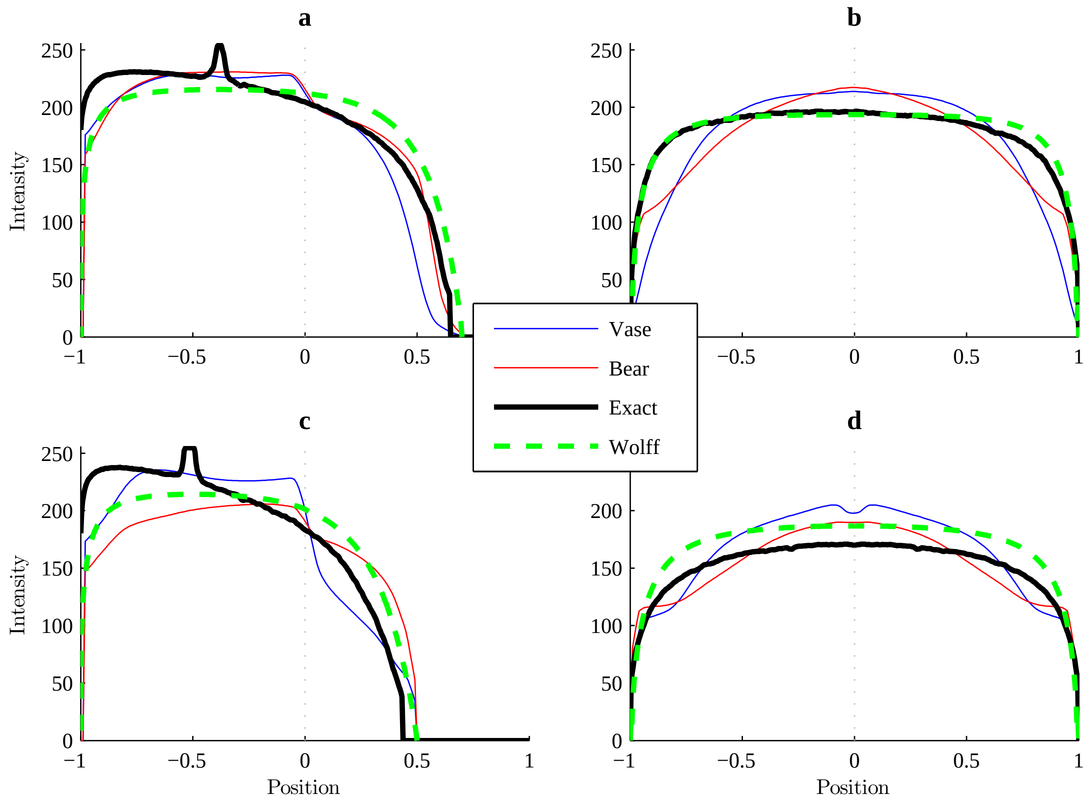

Figure 4.23: (a) Intensities as a function of the distance from sphere centre in the vertical direction (in relation to Figure 4.21) using $\theta_L = 45^\circ$ for estimated and theoretical radiance functions. (b) Same information for the horizontal direction. (c and d) Same graphs with $\theta_L = 60^\circ$.

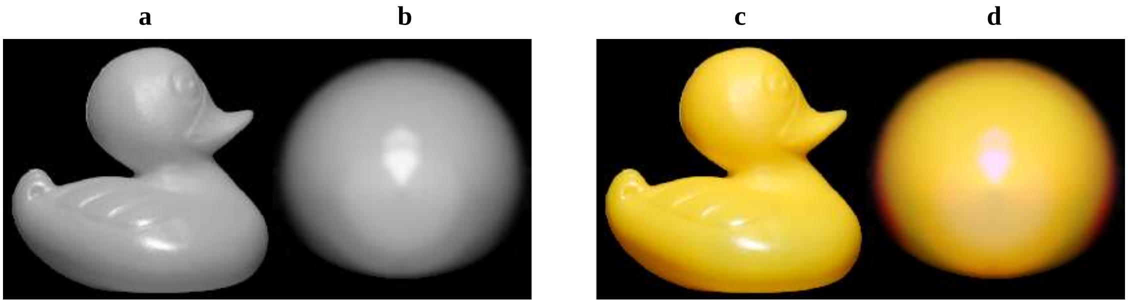

Figure 4.24: (a) Greyscale image of the slightly rough plastic duck. (b) Sphere rendered using the recovered BRDF. (c) Corresponding colour image and (d) sphere rendered using the three-colour BRDF.

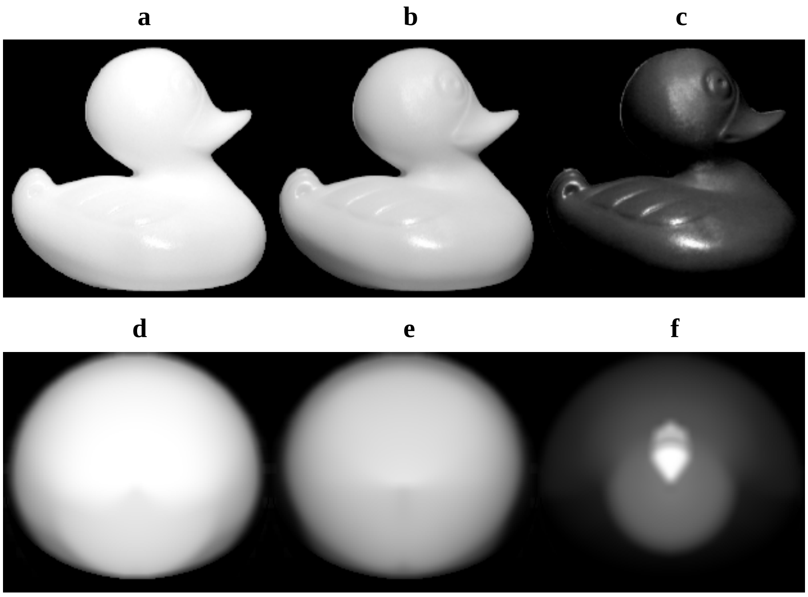

Figure 4.25: Raw images of the plastic duck in the (a) red channel, (b) green channel, and (c) blue channel. (d-f) rendered spheres for each channel.

This was done in such a way that the total colour ratios in the BRDFs matched the ratios in the raw image.

We have estimated the colour BRDFs of a few materials the under conditions where $\theta_L = 45^\circ$. Figure 4.24 shows the estimated BRDF of the plastic duck. Both the greyscale and colour versions of a rendered sphere are shown. For reference, Figure 4.25 shows the individual colour channels of the raw and rendered images. Both rendered images look realistic and appear to be made of the same material as the object. Since the object is slightly translucent, a small part of object at the base is visible where it would otherwise have been in shadow.

Figure 4.26: Greyscale and colour images of an apple and spheres rendered using the estimated BRDF.

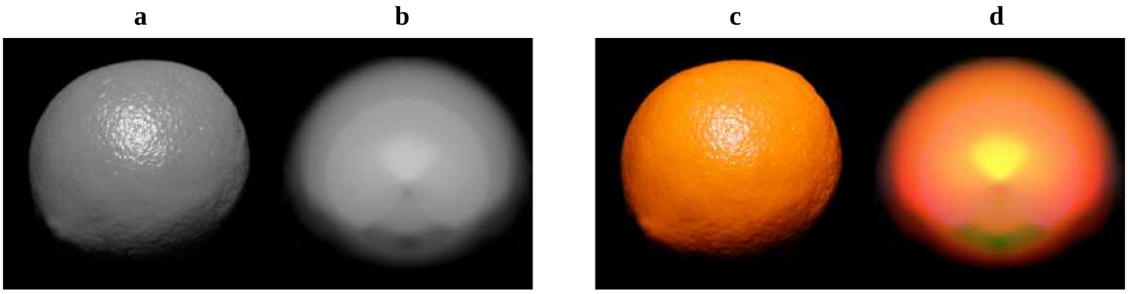

Figure 4.27: Greyscale and colour images of an orange and spheres rendered using the estimated BRDF.

Figures 4.26 and 4.27 show results for an apple and an orange. On these occasions the texture and roughness have caused greater problems for the algorithm and the results are less impressive. In particular, the rendered intensities tend to be too high, especially for regions of the apple with large $\Delta\alpha$. Also, the colour balance is still incorrect for a few small regions. Nevertheless, the results certainly show promise, and it is hoped that future work will yield more consistent results.

The intensities of cross-sections of the rendered spheres for each colour channel are plotted in Figure 4.28. The greyscale intensities are also shown. Note that some of the intensities are saturated for individual colour channels (which is also true for the raw images). This affects the mean (greyscale) intensity and therefore the original zenith angle estimates, although not necessarily critically. The graphs for the plastic duck are largely as one would expect, and show a broadened specularity where the blue component rapidly saturates. The other graphs are more complex than expected and may benefit from more sophisticated smoothing techniques. The figure also shows the intensities predicted by the Lambertian model. The greyscale BRDF predicted using our method closely matches this model for the horizontal cross-sections. This is not the case for the vertical cross-sections, which is not surprising considering the direction of illumination.

4.4.6 Fine Tuning

This section concludes with the discussion of a few minor points. These include extra constraints on the BRDF and the choice of parameters for the cooling schedule.

Shadow Boundary Constraint

It is possible to provide an extra constraint on the BRDF by consideration of the shadow boundary. Under the distant light source assumption, the unit vector in the direction of the light source, at any point, is

$$\mathbf{n}_L = \begin{pmatrix} \cos\alpha_L \sin\theta_L \\ \sin\alpha_L \sin\theta_L \\ \cos\theta_L \end{pmatrix} \tag{4.42}$$

Taking the scalar product of this and the normal vector, and noting that in shadowed areas the angle between these two vectors exceeds $90°$, we arrive at the following condition:

$$\cot\theta + \tan\theta_L \cos\left(\Delta\alpha - \alpha_L\right) \geq 0 \tag{4.43}$$

If this condition is not met, then that part of the surface must be in shadow.

The condition in Equation 4.43 is applied by detecting points in the current radiance function estimate that do not satisfy this condition and mapping them onto the nearest point in the energy field where the condition *is* satisfied. Given that the shadow boundary itself corresponds to the equality in Equation 4.43, we also force points in the current radiance function estimate at $I = 0$ to satisfy the equality. This process may be clarified by reference to Figure 4.20: the initial state is not entirely planar as points not satisfying the shadow constraint have been mapped onto allowed areas. Also, note that points at $I = 0$ are now constant through time as the equality version of Equation 4.43 is always enforced for these points.

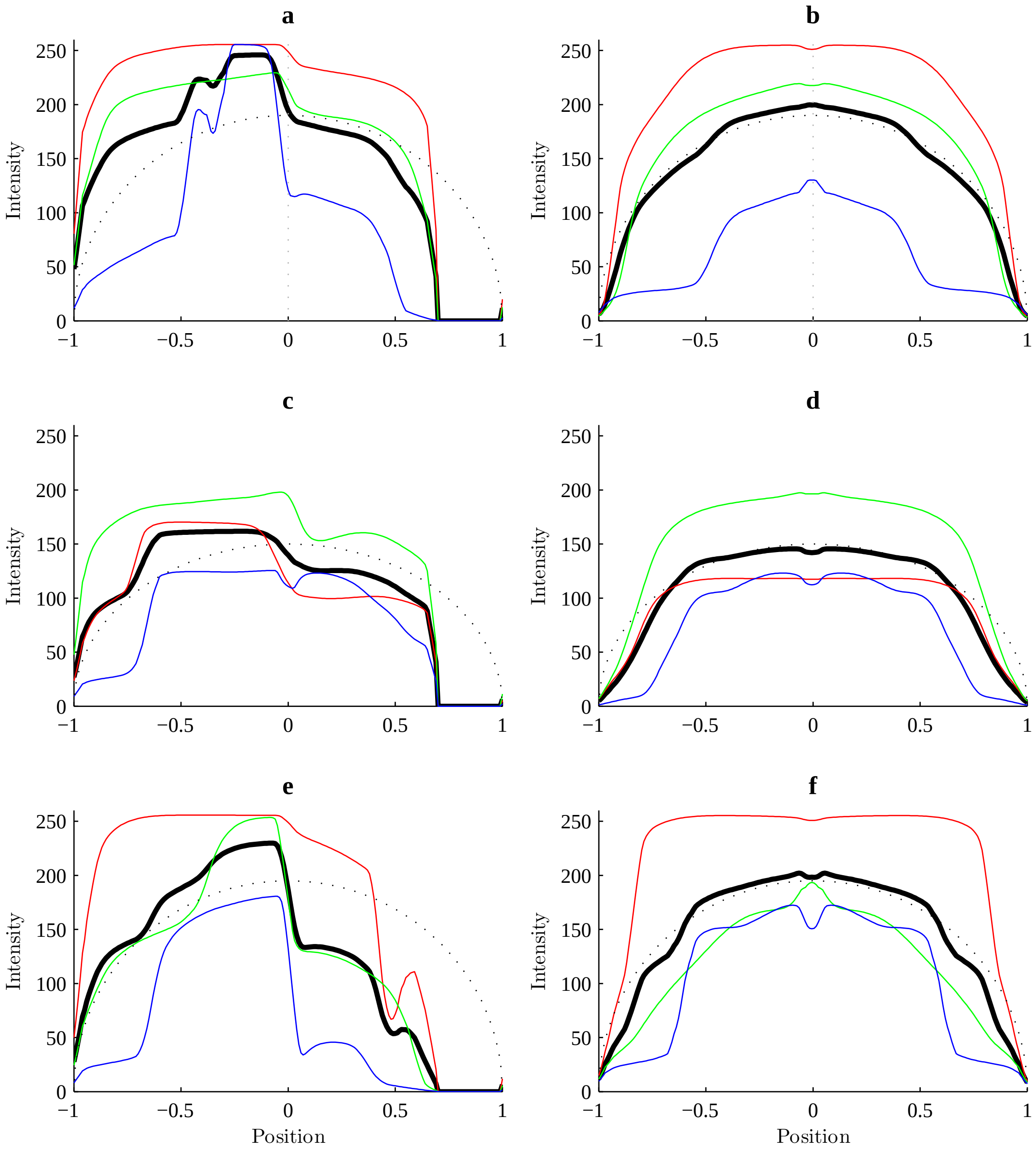

Figure 4.28: Intensities of the rendered spheres of the duck (top row), the apple (middle row) and the orange (bottom row). The left-hand column relates to vertical cross-sections of the spheres while the right-hand column is for horizontal cross-sections. The thin lines refer to the red, green and blue colour channel intensities. The thick lines relate to the greyscale (mean) intensities. The broken lines are for Lambertian reflectance.

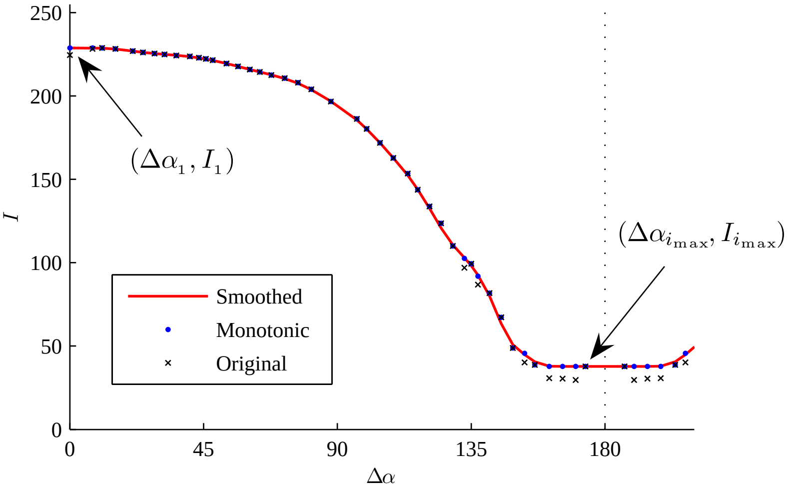

Figure 4.29: Portion of the BRDF corresponding to $31^\circ \leq \theta < 32^\circ$ and $\theta_L = 45^\circ$, showing how non-monotonicities are removed and how a small amount of smoothing is applied.

Monotonicity Constraint

One can intuitively state that most BRDFs are monotonic in $\Delta\alpha$ for a given zenith angle. That is, the following inequality can be enforced:

$$\frac{\partial f(\theta, \Delta\alpha)}{\partial(\Delta\alpha)} \leq 0 \qquad \forall\, \Delta\alpha, \theta \tag{4.44}$$

This is used as a final constraint, which is appended to the end of the algorithm. It turned out that the simulated annealing algorithm returned BRDFs that largely satisfy this condition anyway, so we enforce the constraint using as simple a method as possible. The following transformation was applied to achieve this task:

$$I_i^{(\theta)} \leftarrow \max\left\{I_j^{(\theta)}; j = i, i+1, \ldots i_{\max}\right\} \tag{4.45}$$

where $\left\{I_i^{(\theta)}; i = 1, 2, \ldots i_{\max}\right\}$ is the set of points on the BRDF with zenith angle θ (actually within a small range of θ). The indices i in $\left\{I_i^{(\theta)}\right\}$ run such that $\Delta\alpha_{i+1} > \Delta\alpha_i \,\forall\, i$. Figure 4.29 shows and example of this process. The algorithm also smooths the data by applying moving averages to the set $\left\{I_i^{(\theta)}\right\}$. This removes any sudden gradient variations.

Another possible constraint to ensure sensible BRDF estimation, but which we do not currently apply, is that:

$$\left.\frac{\partial f(\theta, \Delta\alpha)}{\partial(\Delta\alpha)}\right|_{\Delta\alpha=0,\,180^\circ} = 0 \qquad \forall\,\theta \tag{4.46}$$

This can be understood by regarding the azimuth interval in Figure 4.29 as the decreasing half of a periodic function, which can be extended indefinitely. The extension of the figure beyond 180° shows part of the increasing half of the period, which is formed by taking a mirror image of the first half. A similar extension is possible for negative azimuth angles.

Temperature Control

There are three aspects of temperature control that must be considered in simulated annealing algorithms: the initial temperature, the cooling schedule and the convergence criterion. The convergence rules were given in Section 4.4.5. To *guarantee* convergence to the global minimum, the initial temperature must be set high and the cooling coefficient (the factor by which the temperature is reduced after each trial) should be very low. All our results were obtained using an initial temperature of 10 and a cooling factor of 0.999.

In future work, the initial temperature could be set using a few random perturbations to estimate a small number of $\Delta\varepsilon$ values. These values could then be used with Equation 4.28, allowing to set an initial temperature that would give an even distribution of probabilities. Fortunately, the inclusion of an elastic energy already helps to prevent areas of the BRDF surface from converging to local minima. This means that the choice of the initial temperature is less critical here, than in some other simulated annealing applications. The choice of the cooling factor however, is more difficult. Its selection should ideally be determined by the specific application. If efficiency is important, then a lower value should be adopted compared to when this is not the case.

4.5 Conclusions

This chapter has accomplished the following:

1. It has presented the background theory and nomenclature about reflectance functions and the main results of the Wolff reflectance model.
2. Described a method to estimate one-dimensional BRDFs using zenith-intensity histograms and robust PDF fitting.
3. Described a technique to disambiguate azimuth angles using three light source directions.
4. Described a method to estimate two-dimensional BRDFs using zenith-azimuth-intensity histograms, simulated annealing and elastica.
5. Tested the above techniques on a range of objects and materials in both greyscale and colour.

The algorithm for 1D BRDF estimation is more reliable than the 2D version. This is due to the greater density of points in the histogram and the fewer degrees of freedom in the reflectance function. It is, of course, less general and complete than the 2D version. Results show that the 1D technique is reliable for all the materials used without the need for strong assumptions on the properties of the BRDF. The 2D case required several simplifying assumptions in order for the method to work well. Furthermore, the method shows some breakdown for the rougher, textured objects. Overall however, results are promising and allowed for good image rendering for porcelain and plastic. Using more light source directions would provide more data for the histograms and hence more robust BRDF estimates. The 1D algorithm is highly efficient, requiring only a few seconds of computation time. The 2D version requires a few minutes and depends on the choice of cooling parameters.

Finally, the photometric stereo technique is necessary to disambiguate the azimuth angles ready for the 2D BRDF algorithm. It can also be regarded as an alternative to the geometric stereo method, which is to be described in the next chapter. The method is more efficient than the geometric method, and often more accurate. However, the lighting configuration used in photometric stereo is not always practical and many applications require multi-view correspondence. Hence, our main motivation for Chapter 5.

Chapter 5

Multi-View Reconstruction

In this Chapter, we describe our method for shape estimation from two views. The goal is to establish and use dense stereo correspondence to resolve the azimuth angle ambiguities present in the surface normal estimates from a single view. The first step is to extract surface patches from the polarisation images. These patches are then aligned using an optimisation scheme to obtain correspondence indicators. The error criterion used for this task is similar to those developed by Cross and Hancock using the EM algorithm [17] and Chui and Rangarajan using softassign [12].

The input to the algorithm is the set of ambiguous azimuth angles estimated from the phase images, together with the zenith angles from the estimated radiance function and the raw intensity images. We use the set-up shown in Figure 3.4 with polarisation images taken from two views. For the second view, the turn-table is rotated by θ_{rot} about the x-axis (see Figure 3.4 for the co-ordinate axis convention). For this chapter, we use a known value of $\theta_{\text{rot}} = 20^\circ$. Note that both views give retro-reflection. This means that the 1D reflectance function algorithm from Section 4.2 is sufficient to estimate the zenith angles. Consideration of alternative illumination conditions is reserved for future work.

The set of surface patches are obtained by segmenting the two polarisation images into regions[1] according to the estimated surface normals. The segmentation is performed using a bi-threshold technique, where normals falling into an angular interval are selected

---

[1] We refer to each unconnected 2D image segment as a “region”, and use the term “patch” to refer to the 3D data (i.e. the surface normals that can be used to estimate surface geometry) for each of these segments.

from each image. The angular interval is shifted by θ_{rot} between images to account for the object rotation. Details of this procedure can be found in Section 5.1.

The correspondence problem is then reduced to matching patches from the two views. The solution is to align patch pairs for which a potential correspondence exists. We adopt a dual-step matching procedure in which both correspondence indicators and alignment parameters are sought. The cost function uses the correspondence indicators to gate contributions to a sum-of-squares alignment error for corresponding points. The alignment error measures the difference in surface normal components and a scale invariant measure of surface topography (the shape-index). The matching cost is calculated for all potentially corresponding patch pairs and the combination that gives the lowest total cost is taken as the set of patch correspondences. The derivation of the cost function that we adopt is given in Section 5.2.

Further details of the matching algorithm are provided in Section 5.3. It is first shown how the correspondences lead to unambiguous azimuth angle estimates. After this process, some areas of the images remain without detected correspondence. Monotonic interpolation is then used to estimate correspondences for such areas. Finally, the Frankot-Chellappa surface integration method is applied to recover a depth map from the field of unambiguous surface normals. This method was summarised earlier, in Section 3.2.2.

In the remainder of this section, we furnish details of the steps described above. The algorithm is described fully in Sections 5.1 to 5.3. Section 5.4 presents a series of experiments to reconstruct surfaces, and to render images under novel illumination conditions and using different reflectance functions.

Figure 5.1 shows a schematic of the overall algorithm structure and should be referred to throughout the chapter. The top part of the figure refers to the acquisition of the polarisation images from two views to obtain a set of intensities, degrees of polarisation and phase angles for each view. The reflectance function estimation described in the previous chapter is then applied to obtain a set of zenith angle estimates. The remainder of the figure refers to the processes discussed in this chapter.

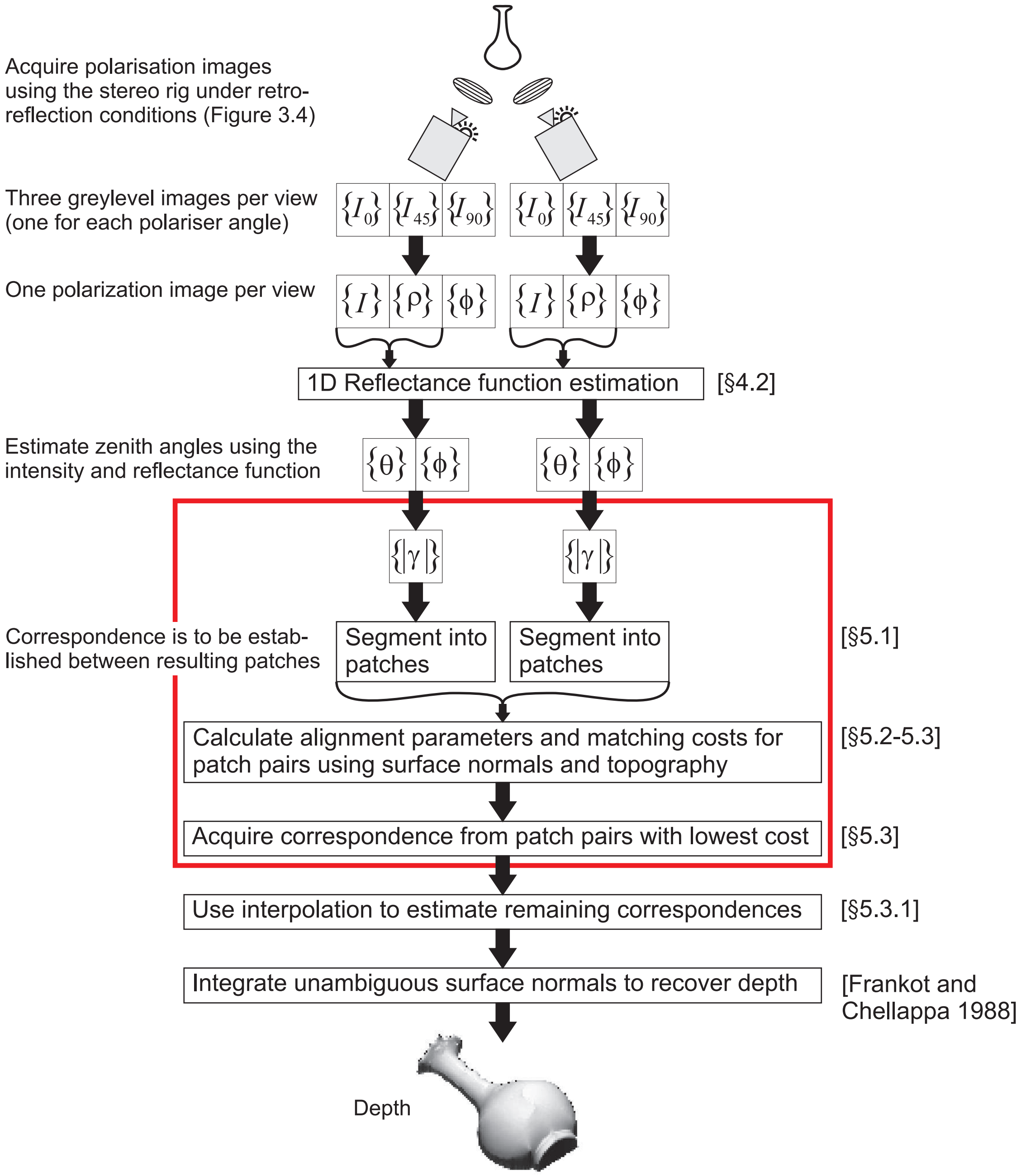

Figure 5.1: Flow diagram of the proposed shape recovery method. The stages of the algorithm in the large box are shown in more detail in Figure 5.2.

5.1 Segmentation

The purpose of this stage of the algorithm is to segment the two images into regions that are suitable for establishing correspondence. The segmentation is performed according to the angle defined by:

$$\gamma = \arctan\left(\sin\alpha\tan\theta\right) \tag{5.1}$$

This is the angle between the viewing direction and the projection of the surface normal onto the horizontal plane (see Figure 3.5). If the correct azimuth angle is greater than 180°, then γ is negative, i.e. the surface normal vector points to the left from the viewer's perspective. However, we do not know the sign of γ at this point and so can only calculate the modulus of the angle from the estimated phase, i.e.

$$|\gamma| = \arctan\left(\sin\phi\tan\theta\right) \tag{5.2}$$

The zenith angle, θ, is determined by the radiance function (Equation 4.2), which is estimated using the technique presented in Section 4.2. Let the minimum angle for reliable $|\gamma|$ estimates be γ_{lim}. Empirically, we found that this limit was approximately 20° for perfectly smooth objects. For rougher surfaces, a reasonable reconstruction was possible using this value, but results were better when $\gamma_{\text{lim}} = 30^\circ$ was used. We know that for regions where $\gamma < 0$ for both images, then where the pixels in the right-hand image (view 2 in Figure 3.4) satisfy the condition

$$\gamma_{\text{lim}} + \theta_{\text{rot}} < |\gamma| \leq 90^\circ \tag{5.3}$$

then the corresponding pixels in the left-hand image (view 1) must satisfy the condition

$$\gamma_{\text{lim}} < |\gamma| \leq 90^\circ - \theta_{\text{rot}} \tag{5.4}$$

Conversely, where $\gamma > 0$, then for regions of the right-hand image that satisfy the condition in Equation 5.4, then pixels in the corresponding left-hand image regions must satisfy

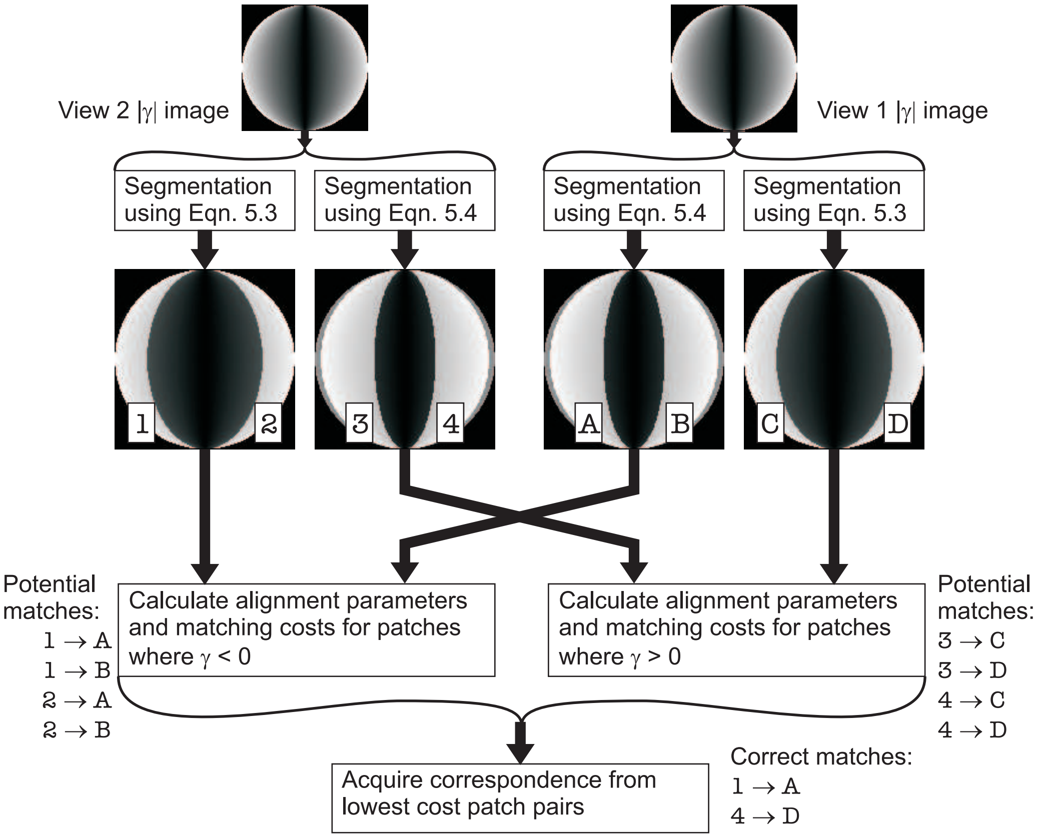

Figure 5.2: Flow diagram of the segmentation and correspondence part of the algorithm, using a synthetic sphere as an example. This figure corresponds to the large box in Figure 5.1.

the condition in Equation 5.3.

Our region segmentation algorithm makes two image passes for each view. On the first pass, the right-hand image is thresholded using the angle condition in Equation 5.3, while the left-hand image is thresholded using the condition in Equation 5.4. On the second pass, the angle conditions are interchanged between the left and right hand views. This two pass thresholding gives two sets of regions per view. Figure 5.2 illustrates the possible correspondence combinations between the region sets. In this way we impose angle interval constraints on the allowable region correspondences. Figure 5.3 shows a real-world example of the result of the segmentation process.

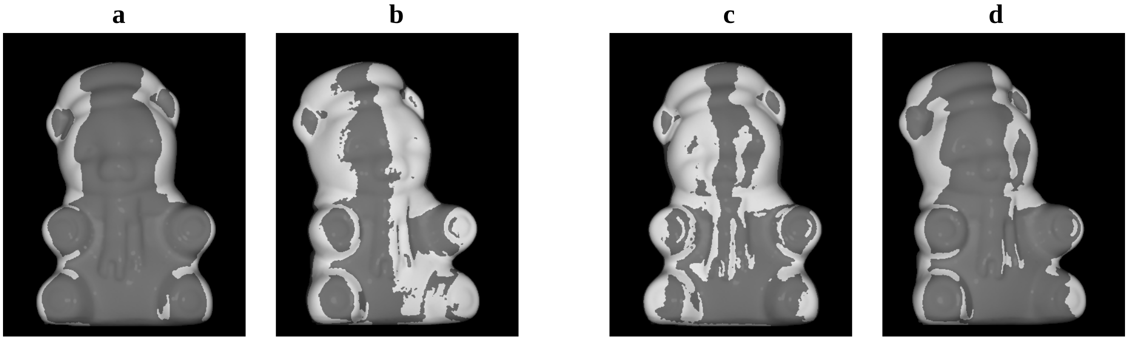

Figure 5.3: Segmentation of a real world image of a porcelain bear model. In (a) and (d), highlighted regions obey Equation 5.3, while in (b) and (c) they obey Equation 5.4.

Figure 5.4: Patch reconstructions of the leftmost regions of Figure 5.3a and b.

5.2 Cost Functions

As already mentioned, the algorithm seeks patch correspondences so as to minimise a cost (or energy) function. In this section we first describe a rudimentary cost function based purely on local patch reconstructions. We then introduce a more complex cost function based directly on the estimated surface normals and the scale invariant topographic quantity known as the shape index.

5.2.1 A Rudimentary Cost Function

This cost function is based on the similarity between local patch reconstructions obtained by applying the Frankot and Chellappa method to recover surface depth, h, from needle maps [27]. Figure 5.4 shows patch reconstructions from two of the larger regions from Figure 5.3a and b, using $\gamma < 0$.

Consider the alignment of two patches, $\mathcal{U}$ and $\mathcal{V}$, defined by: $\mathcal{U} = \{\mathbf{u}_a; a = 1, 2...|\mathcal{U}|\}$

and $\mathcal{V} = \{\mathbf{v}_b; b = 1, 2...|\mathcal{V}|\}$, where $\mathbf{u}_a = (x_a^{(\mathrm{u})}, y_a^{(\mathrm{u})}, z_a^{(\mathrm{u})})^T$ and $\mathbf{v}_b = (x_b^{(\mathrm{v})}, y_b^{(\mathrm{v})}, z_b^{(\mathrm{v})})^T$. The alignment is performed by applying a transformation function J to patch $\mathcal{U}$ to obtain the set of transformed positions: $J(\mathcal{U}) = \hat{\mathcal{U}} = \{\hat{\mathbf{u}}_a\}$, where $\hat{\mathbf{u}}_a = (\hat{x}_a^{(\mathrm{u})}, \hat{y}_a^{(\mathrm{u})}, \hat{z}_a^{(\mathrm{u})})^T$.

For our case where the target object has undergone a known rotation about the vertical (x) axis, the transformation contains a y-translation of Δy, a z-translation of Δz and a rotation about the x-axis of θ_{rot}. For our particular transformation therefore, $\hat{x}_a^{(\mathrm{u})} = x_a^{(\mathrm{u})} \forall a \in \{1, 2...|\mathcal{U}|\}$. We represent the transformation parameters as

$$\overline{\mathbf{\Theta}} = (\Delta y, \Delta z, \theta_{\mathrm{rot}})^T \tag{5.5}$$

where θ_{rot} is known. Our task then, becomes that of estimating the vector of parameter values $\overline{\mathbf{\Theta}}$ for each potentially corresponding patch pair that minimise an energy functional.

The rudimentary energy functional for aligning patches $\mathcal{U}$ and $\mathcal{V}$ is

$$\varepsilon_{\mathcal{U}\mathcal{V}}\left(\overline{\mathbf{\Theta}}\right) = \frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m_{ab}\left(\overline{\mathbf{\Theta}}\right) \left|\mathbf{v}_b - \hat{\mathbf{u}}_a\left(\overline{\mathbf{\Theta}}\right)\right|^2}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m_{ab}\left(\overline{\mathbf{\Theta}}\right)} \tag{5.6}$$

where $\left\{m_{ab}\left(\overline{\mathbf{\Theta}}\right)\right\}$ are the elements of the correspondence indicator matrix $M\left(\overline{\mathbf{\Theta}}\right)$, defined below. This is a modified version of the least squares alignment parameter estimation problem. The correspondence indicators exclude the contributions to the squared alignment error from non-matching points. Similar cost functions are obtained by Cross and Hancock [17] and Gold et al. [29]. In this paper, the elements of the correspondence indicator matrix are assigned as follows

$$m_{ab}\left(\overline{\mathbf{\Theta}}\right) = \begin{cases} 1 \quad \text{if} \begin{cases} a = \underset{a' \in \mathcal{A}(b)}{\operatorname{argmin}} \left|\hat{y}_{a'}^{(\mathrm{u})}\left(\overline{\mathbf{\Theta}}\right) - y_b^{(\mathrm{v})}\right| & \textbf{or} \\ b = \underset{b' \in \mathcal{B}(a)}{\operatorname{argmin}} \left|\hat{y}_a^{(\mathrm{u})}\left(\overline{\mathbf{\Theta}}\right) - y_{b'}^{(\mathrm{v})}\right| \end{cases} \\ 0 \quad \text{otherwise} \end{cases} \tag{5.7}$$

where $\mathcal{A}(b) = \{a \in 1, 2...|\mathcal{U}| \mid x_a^{(\mathrm{u})} = x_b^{(\mathrm{v})}\}$, is the set of labels for points in $\mathcal{U}$ that lie in

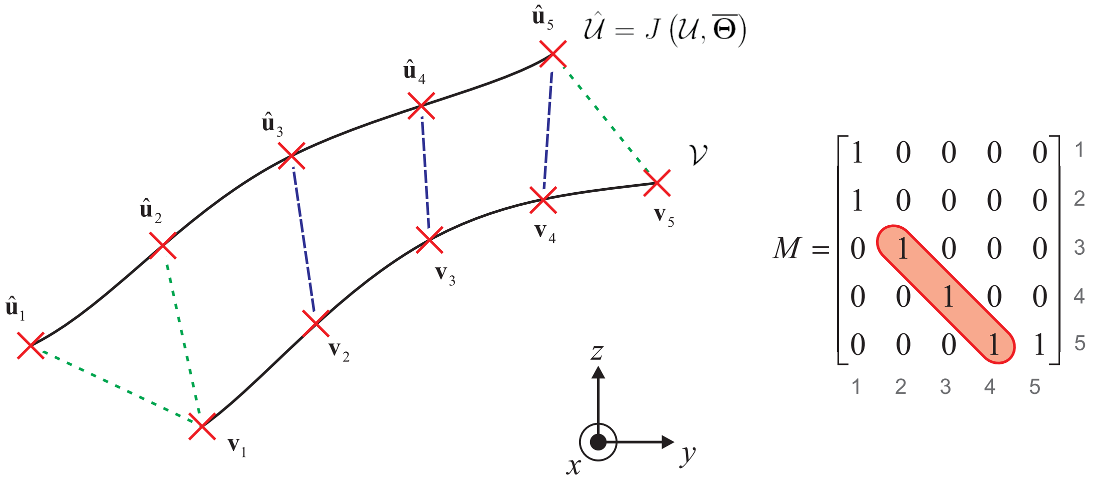

Figure 5.5: Illustration of how matrices $M(\overline{\boldsymbol{\Theta}})$ and $M'(\overline{\boldsymbol{\Theta}})$ are constructed. The points shown are from a horizontal cross section of two patches. All broken lines correspond to elements of $M(\overline{\boldsymbol{\Theta}})$ set to unity. For $M'(\overline{\boldsymbol{\Theta}})$ only the dark broken lines correspond such elements. The elements of $M(\overline{\boldsymbol{\Theta}})$ are shown to the right. For the reduced correspondence matrix, all elements are zero except for those highlighted.

the same horizontal plane as point $\mathbf{v}_b$. Similarly, $\mathcal{B}(a) = \{b \in 1, 2...|\mathcal{V}| \mid x_b^{(v)} = x_a^{(u)}\}$ are the labels for points in $\mathcal{V}$ in the same plane as point $\mathbf{u}_a$. In other words, $m_{ab}(\overline{\boldsymbol{\Theta}}) = 1$ if $\mathbf{v}_b$ is the closest point in $\mathcal{V}$ to $\hat{\mathbf{u}}_a(\overline{\boldsymbol{\Theta}})$ in the y-direction within the same horizontal cross-section. Similarly, $m_{ab}(\overline{\boldsymbol{\Theta}}) = 1$ also, if $\hat{\mathbf{u}}_a(\overline{\boldsymbol{\Theta}})$ is the closest point in $\hat{\mathcal{U}}(\overline{\boldsymbol{\Theta}})$ to $\mathbf{v}_b$. The construction of the matrix, $M(\overline{\boldsymbol{\Theta}})$, is illustrated in Figure 5.5. Note that Equation 5.7 does not include a term with the difference in z-position. This is because in the final algorithm, the energy functional is given in terms of the surface normals, so the height reconstruction (difference in z-position) is not needed to establish correspondence.

The matrix $M(\overline{\boldsymbol{\Theta}})$ is hence defined in a similar fashion to the correspondence matrix used by Chui and Rangarajan [12]. Note that Chui and Rangarajan assume that the correspondence is one-to-one, meaning that each row and each column of the final correspondence matrix has exactly one element set to unity. We do not enforce this constraint here.

Using the correspondence indicator matrix defined by Equation 5.7 means that any points in non-overlapping areas of the patches are inappropriately included in the energy calculation. We therefore use a reduced correspondence matrix, $M'(\overline{\boldsymbol{\Theta}})$, with elements

m'_{ab}, which is defined by

$$m'_{ab}\left(\overline{\boldsymbol{\Theta}}\right) = \begin{cases} m_{ab}\left(\overline{\boldsymbol{\Theta}}\right) & \text{if} \begin{cases} \min\limits_{a' \in \mathcal{A}(b)} \left\{\hat{y}^{(\mathrm{u})}_{a'}\left(\overline{\boldsymbol{\Theta}}\right)\right\} < y^{(\mathrm{v})}_{b} < \max\limits_{a' \in \mathcal{A}(b)} \left\{\hat{y}^{(\mathrm{u})}_{a'}\left(\overline{\boldsymbol{\Theta}}\right)\right\} \ \textbf{and} \\ \min\limits_{b' \in \mathcal{B}(a)} \left\{y^{(\mathrm{v})}_{b'}\right\} < \hat{y}^{(\mathrm{u})}_{a}\left(\overline{\boldsymbol{\Theta}}\right) < \max\limits_{b' \in \mathcal{B}(a)} \left\{y^{(\mathrm{v})}_{b'}\right\} \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{5.8}$$

This reduced correspondence indicator matrix therefore discards points in the patches where there is no overlap, as shown in Figure 5.5.

Note that this method is an alternative to the iterated closest point (ICP) algorithm [5]. We do not use ICP because that method would restrict our cost function to be based on the location of points in 3D space. Use of Equation 5.6 specifically, has no advantage over ICP in this respect. However, we would ideally like to base our cost primarily on the surface normals rather than patch reconstructions, as this avoids error propagation.

5.2.2 An Improved Cost Function

We aim here to refine the cost function to include surface normal information instead of the reconstructed height. However, if *only* the surface normals are used, then we encounter a problem: it is possible that two surface points on different patches could have identical surface normal directions but very different local shapes. We therefore introduce the shape index into the energy functional to account for local surface topography.

The shape index, s, is a view-invariant quantity that describes the local topography of a surface, and falls in the interval $[-1, +1]$. Where $s = -1$, the local surface shape is a concave spherical cup, where $s = 0$ a saddle point is found and where $s = +1$ the surface takes the form of a spherical dome. Between these values the surface takes intermediate topographic forms. The shape index is undefined for planar surfaces.

The starting point for the derivation of the shape index is the Hessian matrix, defined by:

$$H = \begin{bmatrix} \partial^2 h/\partial x^2 & \partial^2 h/\partial x \partial y \\ \partial^2 h/\partial x \partial y & \partial^2 h/\partial y^2 \end{bmatrix} = \begin{bmatrix} \partial p'_x/\partial x & \partial p'_x/\partial y \\ \partial p'_y/\partial x & \partial p'_y/\partial y \end{bmatrix} \tag{5.9}$$

where $h(x,y)$ is the surface height. p'_x and p'_y are the gradients of the surface and are related to the Cartesian components of the unit surface normal in Equation 3.16 by $p'_x = -p_x/p_z$ and $p'_y = -p_y/p_z$.

The Hessian matrix can be calculated from the normals estimated using polarisation data and the radiance function. Since we are performing a differentiation, the estimate of H is not entirely robust to noise. However, since the Hessian will not form the main part of our cost function (the normals themselves do), this problem is not severe.

Using Equation 5.9 directly leads to a view biased estimate of the Hessian, due to a foreshortening effect. We therefore introduce the viewpoint compensated Hessian derived by Woodham [122]:

$$C = \frac{1}{\sqrt{\left(1 + {p_x}^2 + {p_y}^2\right)^3}} \begin{bmatrix} p_y^2 + 1 & -p_x p_y \\ -p_x p_y & p_x^2 + 1 \end{bmatrix} H \tag{5.10}$$

The eigenvalues of this matrix are

$$\lambda_{\max} = -\frac{1}{2}\left(c_{11} + c_{22} - S\right) \tag{5.11}$$

$$\lambda_{\min} = -\frac{1}{2}\left(c_{11} + c_{22} + S\right) \tag{5.12}$$

where $S = \sqrt{\left(c_{11} - c_{22}\right)^2 + 4\left(c_{21} c_{12}\right)}$ and c_{ij} denotes elements of C. $\lambda_{\max}$ and $\lambda_{\min}$ are the principle curvature directions at the surface point and are the sole variables in the definition of the shape index:

$$s = \frac{2}{\pi} \arctan\left(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}\right) \tag{5.13}$$

In our energy functional, we motivate the use of the difference in shape index between surface points as follows. Consider two surface points with identical normal directions with one lying on a spherical dome ($s = +1$) and the other on a spherical cup ($s = -1$). The difference in shape index is then $+2$ which increases the matching cost. Based on this observation we propose the following form, which incorporates both the shape indices and the surface normals:

$$\varepsilon_{\mathcal{UV}}\left(\overline{\boldsymbol{\Theta}}\right) = \frac{\sum_{a=1}^{|\mathcal{U}|}\sum_{b=1}^{|\mathcal{V}|} m'_{ab}\left(\overline{\boldsymbol{\Theta}}\right) P_{ab}\left(\overline{\boldsymbol{\Theta}}\right) \Delta s_{ab}^2}{\sum_{a=1}^{|\mathcal{U}|}\sum_{b=1}^{|\mathcal{V}|} m'_{ab}\left(\overline{\boldsymbol{\Theta}}\right)} \tag{5.14}$$

where $P_{ab}\left(\overline{\boldsymbol{\Theta}}\right) = \left(p_{x_b} - \hat{p}_{x_a}\left(\overline{\boldsymbol{\Theta}}\right)\right)^2 + \left(p_{y_b} - \hat{p}_{y_a}\left(\overline{\boldsymbol{\Theta}}\right)\right)^2 + \left(p_{z_b} - \hat{p}_{z_a}\left(\overline{\boldsymbol{\Theta}}\right)\right)^2$ and $\Delta s_{ab} = s_b - s_a$. The “hat” notation is used to represent quantities having undergone transformation J, as before. Note that the shape index does not depend on the transformation parameters since it is rotation invariant.

Matrix $M'\left(\overline{\boldsymbol{\Theta}}\right)$ is calculated in the same way as before, using the patch reconstructions from the Frankot-Chellappa method to determine nearest neighbours in the y-direction. Note that this reconstruction is effectively used only as a means to account for foreshortening (i.e. so that patch $\mathcal{U}$ can be rotated by θ_{rot}).

5.3 Final Algorithm

The optimum alignment parameters for patches $\mathcal{U}$ and $\mathcal{V}$ are given by

$$\overline{\boldsymbol{\Theta}}_{\mathcal{UV}} = \underset{\overline{\boldsymbol{\Theta}}}{\text{argmin}}\, \varepsilon_{\mathcal{UV}}\left(\overline{\boldsymbol{\Theta}}\right) \tag{5.15}$$

For each pair of corresponding patches, $\mathcal{W}$ and $\mathcal{W}'$ say, the correspondence matrix is then taken as $M\left(\overline{\boldsymbol{\Theta}}_{\mathcal{WW}'}\right)$. The matching cost $\varepsilon_{\mathcal{UV}}\left(\overline{\boldsymbol{\Theta}}_{\mathcal{UV}}\right)$ is calculated for all potential matches. The algorithm sorts the list of costs into ascending order (i.e. the most confident match first) and confirms the matches sequentially, but rejects any that do not satisfy the following consistency constraint:

> If a point at y-position $y_a^{(\text{u})}$ corresponds to a point at $y_b^{(\text{v})}$, then two other corresponding points at $y_{a'}^{(\text{u})}$ and $y_{b'}^{(\text{v})}$ must satisfy:
>
> $$\text{sgn}\left(y_{a'}^{(\text{u})} - y_a^{(\text{u})}\right) = \text{sgn}\left(y_{b'}^{(\text{v})} - y_b^{(\text{v})}\right) \tag{5.16}$$

Due to occlusion, not all patches have a correspondence at all and so we ideally require a cost threshold. Any matches with a cost above the threshold would then be discarded. At present however, we are yet to develop an adaptive means of setting this threshold and intend to consider this issue in future work.

The algorithm was first tested using Equation 5.14 as the energy functional. One problem encountered when using this form is that an area of a patch may be occluded in one view but not in the other. Another issue is that, even though Equation 5.14 only needs the Frankot-Chellappa method as a convenient way to account for foreshortening, it still introduces a distortion in the patch shape and, therefore, a non-optimal correspondence matrix. We reduce the impact of these problems by replacing Equation 5.14 with the following functional, which takes the median energy of the separate costs for each horizontal cross-section of the patches:

$$\varepsilon_{\mathcal{UV}}\left(\overline{\boldsymbol{\Theta}}\right) = \frac{1}{K\left(\overline{\boldsymbol{\Theta}}\right)} \operatorname*{median}_{x\in\ell} \left(\frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m''_{ab}\left(\overline{\boldsymbol{\Theta}}, x\right) P_{ab}\left(\overline{\boldsymbol{\Theta}}\right) \Delta s_{ab}^{2}}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m''_{ab}\left(\overline{\boldsymbol{\Theta}}, x\right)} \right) \tag{5.17}$$

where $\ell = \{x_u\} \bigcap \{x_v\}$ is the set of unique x-positions covered by both patches and

$$m''_{ab}\left(\overline{\boldsymbol{\Theta}}, x\right) = \begin{cases} m'_{ab}\left(\overline{\boldsymbol{\Theta}}\right) & \text{if} \quad x_a^{(\mathrm{u})} = x_b^{(\mathrm{v})} = x \\ 0 & \text{otherwise} \end{cases} \tag{5.18}$$

We have also introduced the quantity $K\left(\overline{\boldsymbol{\Theta}}\right)$, which rewards patch overlap:

$$K\left(\overline{\boldsymbol{\Theta}}\right) = \frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m'_{ab}\left(\overline{\boldsymbol{\Theta}}\right)}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m_{ab}\left(\overline{\boldsymbol{\Theta}}\right)} \tag{5.19}$$

It is not necessary for the algorithm to calculate the cost for every conceivable pair of patches. This is because we know that the two pixels in any given correspondence pair lie on the same horizontal scanline for our particular experimental arrangement. Instead, it only calculates the cost of patch pairs where there is some overlap in the vertical co-ordinates (i.e. $\ell \neq \emptyset$). For such pairs, $\varepsilon_{\mathcal{UV}}\left(\overline{\boldsymbol{\Theta}}\right)$ is minimised using the Nelder-Mead method [56].

Figure 5.6: (a) Azimuth angles of disambiguated regions of the bear model. (b) Final estimates of the azimuth angles for the entire object.

We have also experimented with introducing an affine skew parameter into the vector $\overline{\Theta}$. This is to further reduce the impact of patch distortions arising from the local reconstruction process. The skew is particularly useful for long and thin patches such as those in Figure 5.4 since imperfect surface normal estimates have a greater detrimental effect in such cases. It is straightforward to add additional parameters to the transformation vector, $\overline{\Theta}$, using our method to increase the generality of the technique.

After the patch matching is complete, we can disambiguate the azimuth angles for the regions where correspondence was established. We do this using the knowledge that the patches reconstructed using $\gamma > 0$ have $\alpha < 180°$ (i.e. $\alpha = \alpha_1$ in Figure 3.5) and patches where $\gamma < 0$ have $\alpha > 180°$ ($\alpha = \alpha_2$). Figure 5.6a shows these unambiguous regions for the bear model. We still have two estimates of these unambiguous surface normals, one from each view. Our final value is taken from the view where the zenith angle was greatest, since the raw polarisation data would have contained less noise for this estimate.

5.3.1 Correspondence for Remaining Areas

Large areas of the images remain without correspondence, and therefore ambiguous, after the above matching procedure. To assign a value for the azimuth angle for these areas, we use *piecewise cubic Hermite interpolating polynomials* (PCHIP) [28]. This form of interpolation is similar to the spline technique in that a continuous curve is generated that passes exactly through the data points. The difference is that spline maintains continu-

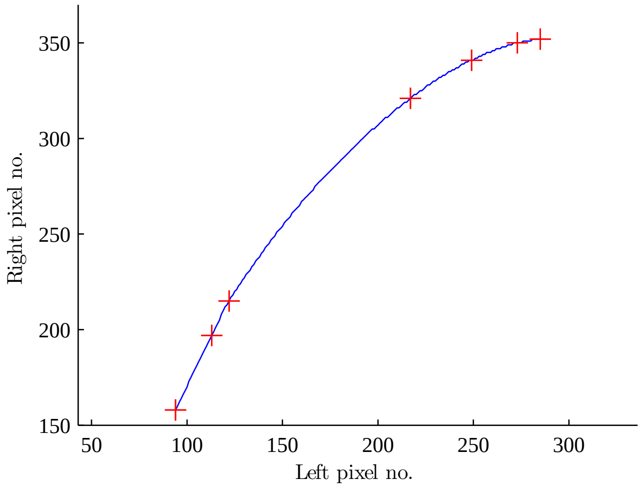

Figure 5.7: Example of using PCHIP to establish dense correspondence. The "+" markers indicate corresponding points on a particular scanline obtained from patch matching (not all are shown for reasons of clarity). The curve is the PCHIP interpolation.

ous first *and* second derivatives, while PCHIP allows discontinuities in the latter. The advantage of this from our point of view is that PCHIP ensures that a monotonic curve is generated, thus ensuring that the point correspondences satisfy the condition in Equation 5.16.

The algorithm effectively plots the y-positions of corresponding points from each view against each other and applies PCHIP to interpolate between correspondences. This method only provides relatively crude *correspondence* for areas not in the vicinity of the regions disambiguated previously. This is one reason for us currently relying on integration methods to recover depth instead of triangulation. However, it is sufficient as a means of selecting the most reliable azimuth angle (α_1 or α_2) for most image pixels. Figure 5.7 shows an example of using PCHIP for this task.

We do not use the results of PCHIP for areas of the image where $|\gamma| < \theta_{\text{rot}}$. Instead, we propagate from surrounding areas and align surface normals to the mean of local 3×3 windows. This is advantageous since these areas are the most susceptible to noise due to lower degrees of polarisation. Figure 5.6b shows the unambiguous azimuth angles for the full image for each view of the bear model.

To demonstrate the effects of disambiguation errors, and to further motivate our choice of the Frankot-Chellappa algorithm for shape reconstruction from surface normals, con-

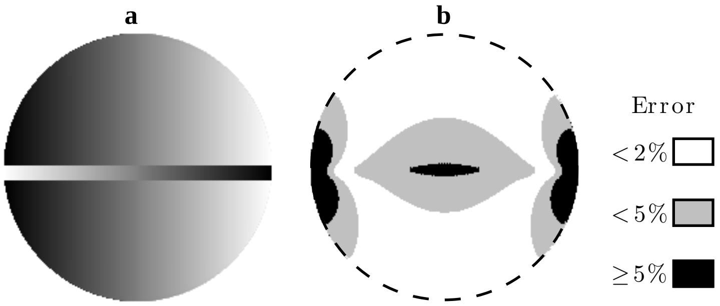

Figure 5.8: Illustration of the effects of an incorrect disambiguation on recovered height. (a) Image of p_y for a synthetic sphere with a disambiguation error simulated. (b) Magnitude of the error in the recovered height expressed as a percentage of the sphere radius.

sider Figure 5.8. For this figure, a synthetic sphere was created and the surface normals were calculated at each point. For a slice of the sphere, the signs of the p_x and p_y values were reversed to simulate the effects of a correspondence error. The surface normals were then integrated using the Frankot-Chellappa method to obtain a depth map. The difference between the original sphere and the recovered depth is shown in the figure and is clearly relatively small, given the magnitude of the error introduced. Had triangulation been used with such an error, then the reconstruction would be more severely deformed.

5.4 Results

This section presents the experimental results of the technique described above. The reconstruction algorithm has been applied to a range of objects and the strengths and weaknesses are clearly demonstrated by the results. Several ground truth comparisons are also shown. Finally, the section presents a few synthetic images of the reconstructed surfaces, rendered using the reflectance functions estimated in Chapter 4.

5.4.1 Surface Reconstructions

Figure 5.9 shows the recovered depth maps of some of the objects used in Chapters 3 and 4. A good reconstruction is clearly obtained for each of the six cases. There are two poorly recovered object areas. Firstly, the paws of the bear model do not protrude as they should, due to inter-reflections. Although using the estimated radiance function

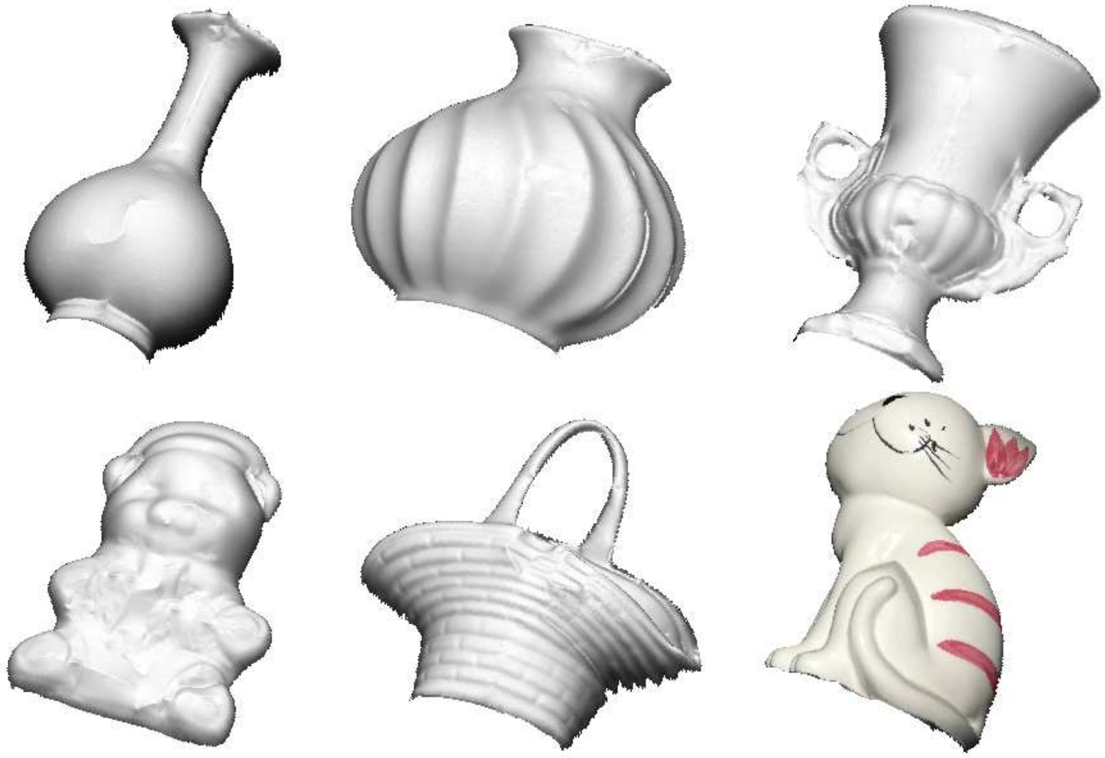

Figure 5.9: Recovered 3D geometry of some of the objects in Figures 3.10, 4.5 and 4.7. The texture on the cat model has been mapped back onto the surface.

helped to accurately recover the zenith angles here, the azimuth angles were incorrect for these areas. Secondly, the handle of the basket does not arch from one side of the object to the other, as Figure 5.10 shows. This is a result of the needle map integration method that we used and is the type of error that may be reduced if triangulation methods were incorporated into the final stage of the algorithm. Results for three other objects of different materials are shown in Figure 5.11. Here, the original RGB images have been used to reapply the texture onto the reconstruction.

Figure 5.12a shows a Lambertian rendering of the porcelain vase using ground truth height data. The ground truth was calculated using the volume of revolution of the object contour. In Figure 5.12b, three ground truth cross sections are compared to the recovered height. The comparison proves that the method works very well for the simplest case, namely that of a smooth object with basic geometry.

The difference image between the ground truth and the recovered height is shown in Figure 5.13a. Note that any symmetries in the true object geometry are not always

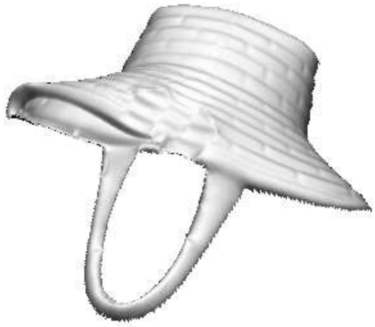

Figure 5.10: Different view of the porcelain basket from Figure 5.9, showing an error in the reconstruction of the handle.

Figure 5.11: Recovered shapes of objects of different materials with the original textures mapped onto the surface. The original images can be found in Figures 3.10 and 4.7.

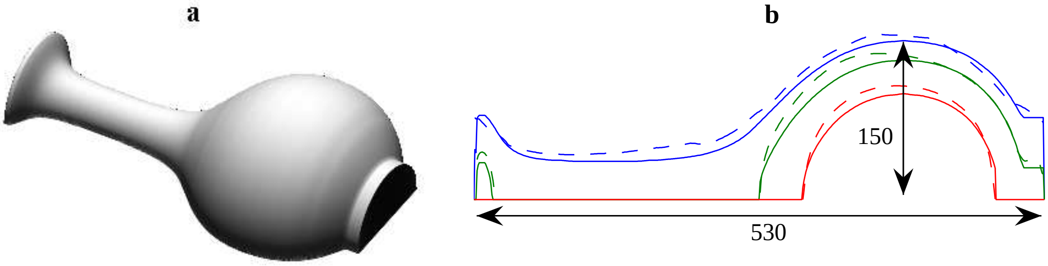

Figure 5.12: (a) Lambertian rendering of the porcelain vase using ground truth data. (b) Three ground truth cross-sections of the vase (solid lines) compared to the recovered height (broken lines). The distances indicated are such that one unit is equal to one pixel length.

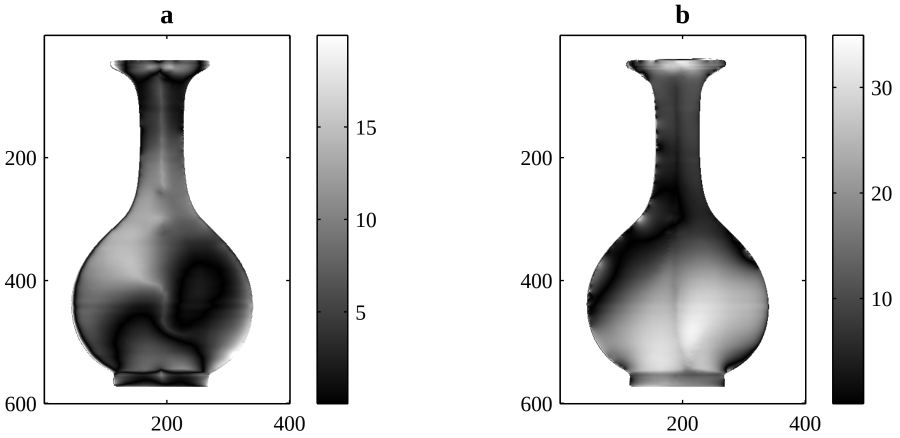

Figure 5.13: Difference between the recovered height and ground truth when (a) the radiance function is used to refine the zenith angles and (b) when the radiance function is not used. The scales to the right of the images indicate the absolute height error and are in the units of pixel lengths. Note that the scale is different for each image.

preserved in the error. This is a result of small environmental inter-reflections and is possibly accentuated by imperfect polariser calibration. For comparison, Figure 5.13b shows the difference image when the raw zenith angle estimates were used (i.e. those calculated directly from Equation 3.21 without estimating the radiance function). The root-mean-square height errors were eight units when the radiance function was used and eighteen units when it was not used. Note that the depth of the vase is 150 units. The effect of using the radiance function was therefore to reduce the average error of the reconstructed height from 12% to 5%.

Figure 5.14 shows a comparison of the ground truth profiles of the porcelain urn and the orange. The urn profile shows that, although the general shape is recovered, the height

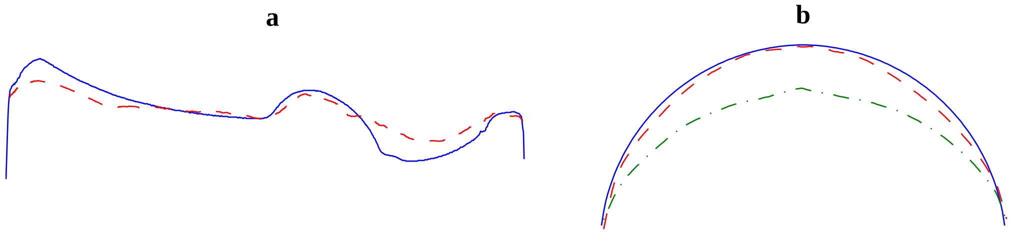

Figure 5.14: Ground truth profiles of (a) the porcelain urn and (b) the orange (solid lines). The broken lines show estimates of the profile. The bottom curve in (b) was estimated without using the radiance function.

variations are less pronounced than they should be for objects with complicated geometry. The method performed better for the orange. Figure 5.14b also shows the profile estimate of the orange without using the radiance function. Here, the radiance function was clearly essential for accurate shape estimation.

5.4.2 Image Rendering

With the shape reconstructions and radiance functions both to hand, it is possible to render images of the objects under novel illumination directions. This was done for Figure 5.15 for two of the porcelain objects. The reflectance functions used were those obtained using the 2D reflectance function estimation algorithm described in Section 4.4. The same radiance functions were used to render the spheres in Figure 4.21c and d.

The rendering process was simply a case of assigning an intensity to each pixel of the image based on the estimates of the surface orientations and the radiance function. For each point, the radiance function was interpolated to the surface normal angles estimated for that point. A ray tracer was used to determine which areas of the object should be in shadow. These areas were then set to zero intensity. The images look realistic apart from a few blemishes in the reconstructions.

In Figure 5.16, more images have been rendered, but this time using radiance functions of different materials. Figure 5.16a shows an orange rendered as though it were made of porcelain. Figure 5.16b shows a porcelain vase rendered using the three-channel radiance function of the plastic duck.

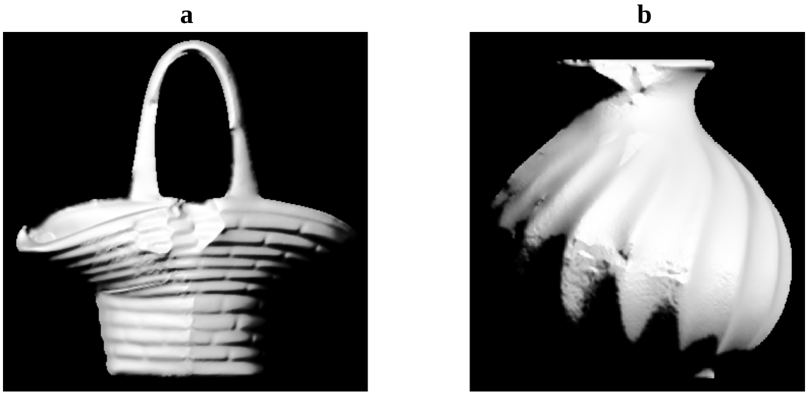

Figure 5.15: Using the reconstructions and estimated radiance functions to render images of porcelain objects under different illumination directions. (a) $\theta_L = 45^\circ$, $\alpha_L = 315^\circ$. (b) $\theta_L = 60^\circ$, $\alpha_L = 45^\circ$.

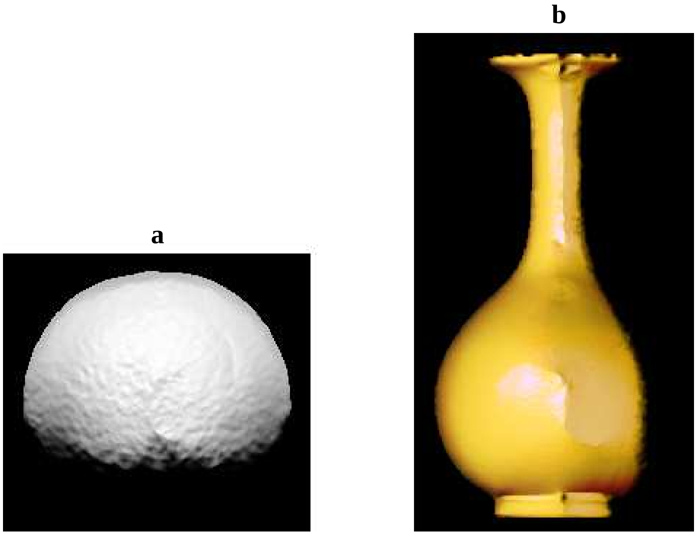

Figure 5.16: Rendering the reconstructed objects using different material reflectance functions and illumination directions. (a) Orange rendered as porcelain, $\theta_L = 60^\circ$, $\alpha_L = 0^\circ$. (b) Porcelain vase rendered as plastic, $\theta_L = 45^\circ$, $\alpha_L = 270^\circ$.

5.5 Conclusions

This chapter has accomplished the following:

1. It has shown how a pair of polarisation images can be segmented to reduce the stereo correspondence problem to a patch matching problem.
2. Described potential cost functions suitable for the matching process.
3. Described an algorithm to perform the matching.
4. Described a means to establish dense correspondence from the set of matched patches.
5. Conducted a series of shape reconstruction experiments to demonstrate the successes and limitations of the technique.
6. Compared experimental results to ground truth profiles.
7. Used the reconstructions together with the reflectance functions calculated in Chapter 4 for image rendering.

The first four points collectively form the stereo reconstruction algorithm, which is summarised by Figure 5.1. The next two points are the main experimental results for our method and are clearly superior to those from the basic method discussed in Section 3.2 (compare, for example, Figure 3.11 to Figure 5.9). The final point shows how the appearance of the reconstructed objects can be predicted after changes to the illumination conditions.

The development of the 3D reconstruction algorithm was the main goal of the thesis. Use of the 1D reflectance function estimation algorithm described in Section 4.2 to enhance zenith angle estimates is optional, but greatly improves results at low computational expense. Perhaps the main weakness of the technique is the special experimental arrangement that is currently required. In future work we hope to acquire the two polarisation images simultaneously using two cameras and fixed object and light source positions.

This is likely to involve use of the 2D reflectance function estimation algorithm described in Section 4.4 to enhance surface normal estimates under the more general illumination conditions that would be present. More details of plans for future work will be described in Chapter 6.

Chapter 6

Conclusion and Outlook

This conclusion commences with a summary of the contributions of Chapters 3 to 5. This will be followed by a discussion of the strengths and weaknesses of the computer vision techniques presented. Finally, directions for future work will be discussed.

6.1 Summary of Contributions

This thesis has advanced shape and reflectance analysis techniques in the field of computer vision, by use of polarisation information and Fresnel theory. The most significant contributions are twofold. First, was the new idea of using shading information to complement the polarisation data. This was achieved by the development of the BRDF estimation algorithms. These algorithms used the histograms of surface orientations and pixel intensities to find an optimum reflectance curve/surface. The second of the main contributions was the idea that polarisation vision can be used with several viewpoints. A novel patch matching algorithm was developed for this purpose, which minimised an energy function based on surface normal estimates and local topography.

Figure 6.1 shows a flow diagram to demonstrate the development of the main shape recovery algorithm in the thesis. We shall consider this algorithm first, and summarise the other contributions afterwards. The left-hand column of the figure formed the starting point for our method and was taken from existing theory. This foundation consisted of the

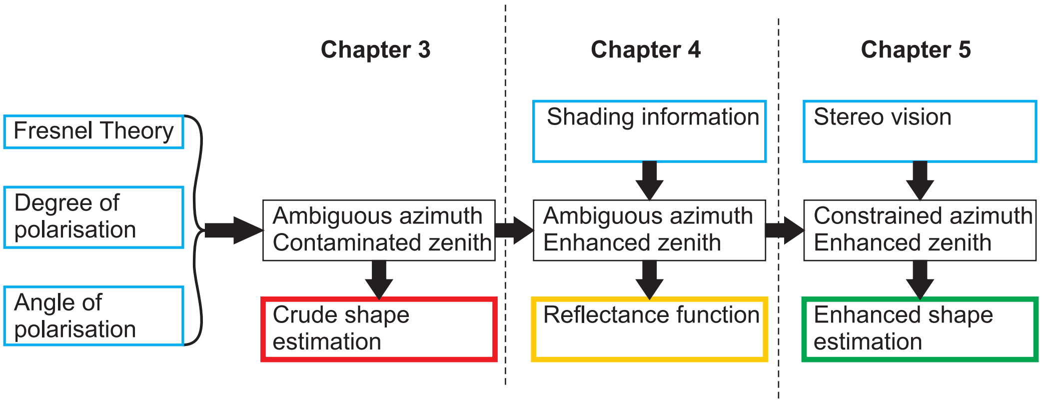

Figure 6.1: Development of the main shape recovery technique in the thesis.

long-established Fresnel theory and a simple method for estimating the degree and angle of polarisation from image data.

6.1.1 Development of the Shape Recovery Technique

The novel contributions of this thesis mostly involved processing polarisation images. Due to its robust nature, ease of implementation, and computational efficiency, we chose to rely on the Frankot-Chellappa method for converting needle maps into depth. The novel shape estimation techniques can therefore be regarded as methods for surface *normal* estimation from polarisation images. In developing the main technique, we also devised new reflectance function estimation and stereo correspondence methods.

As shown in Figure 6.1, the raw polarisation images provide only crude estimates of the normals. This is due to factors such as heavy camera noise, inter-reflections and roughness. In addition, the azimuth angles have a two way ambiguity. Chapter 3 showed the extent to which shapes could be estimated using minimal data processing and a single view. The result was a highly efficient method that did not attempt to refine the crude zenith angle estimates and relied on occluding boundaries for disambiguation. Although the technique was unable to recover details and was strongly affected by roughness, the method recovered the general shapes reasonably well and may be useful in future work if efficiency is an important factor.

An important novel concept used in the thesis is that shading information can be combined with polarisation to enhance the surface normal estimates. Chapter 4 used this idea to aid the development of an image-based reflectance function estimation algorithm, which was later used to improve shape recovery. To date, we have only applied the *one*-dimensional reflectance function estimation to shape recovery. We showed how consideration of the intensity-zenith histogram could be processed to deduce a 1D slice of the BRDF. This allowed for more accurate zenith angle estimates because the problems caused by inter-reflections, roughness and noise were significantly reduced. As Figure 6.1 shows, the incorporation of shading information led to accurate, but still ambiguous estimates of the surface normals.

Chapter 5 showed how two polarisation images can be used to obtain a set of patch correspondences between views. A novel patch alignment and matching algorithm was developed for this task. This involved minimising an energy functional based on surface normals (estimated from polarisation data and refined using the BRDF) and curvature. The multi-view reconstruction allowed for complete disambiguation and, therefore, an accurate field of surface normals. The results in Chapter 5 showed accurate shape reconstructions for a variety of object geometries and materials and are superior to many existing shape recovery techniques.

6.1.2 Further Contributions

The other novel contributions of the thesis are as follows. Chapter 3 contained a detailed analysis of the accuracy of raw polarisation measurements for various materials and describes a simple device for estimating refractive indices. In Chapter 4, the BRDF estimation technique was extended to two dimensions using the zenith-azimuth-intensity histogram. The BRDF was extracted from the histogram using simulated annealing and elastica. This technique was important because it removed the strong restrictions on the illumination conditions that were necessary for the 1D version. Chapter 4 also presents a photometric stereo method for azimuth angle disambiguation. Finally, Chapters 4 and 5 both demonstrate the use of the estimated reflectance functions for image rendering.

In the latter case, the reflectance functions were used in conjunction with the estimated shapes to render the objects under different illumination conditions and as if they were made from different materials.

6.2 Strengths and Weaknesses of the Thesis

Consider the main method presented in this thesis, illustrated by Figure 6.1. The most important advantage over the majority of previous stereo algorithms is that featureless surfaces can be used to establish correspondence. This means that the method complements many existing feature matching techniques.

In this thesis, we have focussed entirely on single object reconstruction. An advantage of the technique is that it should be possible to extend it to scenes with several objects. Indeed, it is hoped that the multi-view algorithm would be able to match whole objects between views, in addition to matching individual patches. The formulation used for the matching algorithm makes it easy to extend the method to more general transformations. This is simply done by adding extra parameters to the transformation vector (Equation 5.5).

Using shading information to enhance the surface normal estimates is another important strength of the thesis. This avoids the need for hefty geometric assumptions about the target object (unlike the work of Miyazaki et al. [67], for example), whilst making full use of each component of the polarisation image (unlike Rahmann and Canterakis [85, 84]). Despite this, the normal estimates are still less accurate than those possible in the presence of purely *specular* reflection. As mentioned in Section 3.1.5 however, the specular case presents its own difficulties.

Processing histograms to estimate reflectance functions, as described in Chapter 4, is a novel way to determine the relationship between the observed radiance and the surface orientation. Very few image-based methods have been proposed in the literature to estimate BRDFs. The advantage of image-based methods is that they are non-intrusive and do not require specialised equipment. Our 1D BRDF algorithm also has the strength of being highly computationally efficient, although this is not currently true for the 2D case. This latter weakness however, will hopefully be overcome in future work using the deterministic mean field annealing [76]. The photometric stereo method described in Chapter 4 is reliable, simple to implement algorithmically, highly efficient, and robust to inaccurate light source positioning. It does however, suffer from the common weakness of photometric stereo methods that it is not always practical to position the light sources as required.

One weakness of most polarisation vision techniques, including those presented in this thesis, is the extra time needed to acquire images using the rotating polariser. This restricts the techniques to stationary objects and scenes. As mentioned in Section 2.1.1, polarisation cameras partially overcome this weakness, at the expense of higher noise levels. It is hoped, but not yet tested, that the reflectance estimation algorithm would be robust to this extra noise. If so, then our technique could be applied to slow moving objects.

The depolarising effects of surface roughness is another common weakness for polarisation methods in general. The probability density function fitting for the reflectance function estimation helped to minimise effects here but certainly did not eradicate the problem entirely. Likewise, the dependence on the refractive index is another limitation. Perhaps a more sophisticated use of shading information would be needed to overcome these weaknesses simultaneously. The degree of polarisation would then only be used as an indicator of the *relative* size of the zenith angles for a particular object in the image. This would mean that the precise relationship between the degree of polarisation and the zenith angle would not be needed (i.e. Equation 3.21 would become obsolete). In addition, this would eliminate the need for the refractive index and mean that the only effect of roughness would be the introduction of more noise.

The final weaknesses relate to the special lighting configuration used for the stereo algorithms and the fact that direct specular reflections are not considered. We believe that these issues can be largely overcome using the techniques presented in this thesis. Specifically, it is thought that rather than rotating the object under fixed illumination and viewing, several cameras could be used simultaneously with a fixed object and light source. The 2D BRDF estimation algorithm could then be applied to refine zenith angle estimates. Since the specularities would not occur at the same surface points in the images, the specularity issue would be resolved. The method would also be less restricted by the special configuration needed by our current technique.

6.3 Outlook

In addressing some of the current weaknesses in the ways suggested above, it is thought that the generality of the technique can be extended. This can be done without the need for major new ideas or principles. One of the most important suggestions made in the previous section was to consider specular reflection more carefully. It may prove useful to separate specular and diffuse reflections from the images [109] and recover the shape from each component separately. Specularities also provide strong clues to the illumination distribution [67]. A major goal for future work would be to estimate the light source direction(s) from several views. This should allow for a more complete (3D) BRDF to be estimated, which could include specular spikes. If this algorithm is successful, then the technique may prove valid under non-controlled illumination conditions, or even outdoor scenes.

A further means to increase the generality of the technique, is to incorporate texture cues. The current method breaks down in the presence of significant texture. Use of single view shape-from-texture techniques [105] is one possibility. However, we believe that a form of feature matching (probably SIFT [60]) would be more useful. The main reason is that by matching features between images, we are better able to incorporate the extra information from the various viewpoints.

This brings us to the next point. Matching features between views provides a means to calculate depth directly using triangulation. Assuming that an algorithm is reliable in establishing correspondence, then triangulation is more accurate than needle map integration. The correspondence method described in Chapter 5 proved to be reliable for regions of large zenith angles and resulted in a good overall disambiguation of azimuth angles. For these regions therefore, the use of triangulation, instead of integration, is likely to result in a more accurate reconstruction. It is not yet clear however, whether a direct transition from integration to triangulation would be adequate for surface regions with small zenith angles. Perhaps triangulation should be used where surface features are detected using SIFT and for high confidence patch correspondences only. The gaps could then by filled in using integration methods. If such a combined triangulation/integration method were to be developed, it would not only be of use for polarisation methods, but also to combine feature matching with shape-from-shading and other methods.

The focus of this thesis was primarily on the development of novel software for computer vision. We end the thesis with a brief consideration of the hardware necessary to implement a complete and fully operational sensor. The sensor would use the techniques presented in this thesis for non-intrusive object detection and reconstruction.

Our proposal consists of two polarisation cameras, one of which may be mounted with a flash gun to provide the source of illumination. If the flash is used, then the 1D reflectance function estimation algorithm would be used to refine normal estimates for that camera and the 2D version for the other camera. If the flash is not used, then the 2D version must be used for both views. In this thesis, we relied on photometric stereo as part of the 2D algorithm (for disambiguation). There should however, be enough information to rely on geometric stereo for disambiguation instead.

The two cameras would communicate with each other to trigger simultaneous polarisation image acquisition. They would also communicate with a computer to provide information on their relative positions and orientations. This is so that θ_{rot} and any other required transformation parameters for Equation 5.5 are known (not all the parameters, of course, need to be known, as they can be estimated by the algorithm). The multi-view method can then be used as described in Chapter 5 to establish stereo correspondence from the refined normal estimates from each view. Finally, integration and/or triangulation can be used to estimate the depth.

A more ambitious goal would be to allow the system to learn. Polarisation cameras are able to acquire images rapidly, at above five frames per second [116, 66]. This would allow the relative positions of the cameras and object(s) to change. As more views become available, larger and more accurate portions of the BRDF could then be determined. The system would then rapidly improve the accuracy of its depth estimates with time as it learnt more about the material reflectance properties from different viewpoints. It may also be able to manage several materials within a scene by using the different BRDFs to segment the images. One certainty emerging from this thesis is that polarisation offers a great deal of information for vision systems and yields many possibilities for the future.

Bibliography

[1] E. Aarts and J. Korst. *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*. Wiley, 1988.

[2] A. Agrawal, R. Chellappa, and R. Raskar. An algebraic approach to surface reconstruction from gradient fields. In *Proc. ICCV*, pages 174–181, 2005.

[3] P. Beckmann and A. Spizzichino. *The scattering of electromagnetic waves from rough surfaces*. Pergamon, New York, 1963.

[4] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *Intl. J. Comp. Vis.*, 35:33–44, 1999.

[5] P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans. Patt. Anal. Mach. Intell.*, 14:239–256, 1992.

[6] M. Born and E. Wolf. *Principles of Optics. Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Pergamon, London, 1959.

[7] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In *Proc. Intl. Joint Conf. Artificial Intelligence*, pages 935–936, 1985.

[8] M. Z. Brown and G. D. Hager. Advances in computational stereo. *IEEE Trans. Patt. Anal. Mach. Intell.*, 25:993–1008, 2003.

[9] A. M. Bruckstein, R. J. Holt, and A. N. Netravali. Discrete elastica. *Applicable Analysis*, 78:453–485, 2001.

[10] V. Černý. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *J. Opt. Theory Appl.*, 45:41–51, 1985.

[11] H. Chen and L. B. Wolff. Polarization phase-based method for material classification in computer vision. *Intl. J. Comp. Vis.*, 28:73–83, 1998.

[12] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89:114–141, 2003.

[13] D. Clarke and J. F. Grainger. *Polarised Light and Optical Measurement.* New York: Pergamon, 1971.

[14] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Trans. Graphics*, 1:7–24, 1982.

[15] T. W. Cronin. Stomatopods. *Curr. Biol.*, 16:R235–236, 2006.

[16] T. W. Cronin, N. Shashar, R. L. Caldwell, J. Marshall, A. G. Cheroske, and T-H. Chiou. Polarization and its role in biological signalling. *Integr. comp. Biol.*, 43:549–558, 2003.

[17] A. D. J. Cross and E. R. Hancock. Graph matching with a dual-step EM algorithm. *IEEE Trans. Patt. Anal. Mach. Intell.*, 20:1236–1253, 1998.

[18] J. E. Cryer, P. S. Tsai, and M. Shah. Integration of shape from shading and stereo. *Pattern Recognition*, 28:1033–1043, 1995.

[19] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Trans. Graphics*, 18:1–34, 1999.

[20] O. Drbohlav and R. Šára. Using polarization to determine intrinsic surface properties. In *Proc. SPIE Conf. Polarization and Color Techniques in Industrial Inspection*, pages 253–263, 1999.

[21] O. Drbohlav and R. Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *Proc. ICCV*, pages 581–586, 2001.

[22] O. Drbohlav and R. Šára. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. ECCV*, pages 46–62, 2002.

[23] P. Dupuis and J. Oliensis. Direct method for reconstructing shape from shading. In *Proc. CVPR*, pages 453–458, 1992.

[24] J. D. Durou, M. Falcone, and M. Sagona. A survey of numerical methods for shape from shading. Technical report, IRIT, Université Paul Sabatier, Toulouse, 2004.

[25] L. Euler. Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes. Lausanne, 1744.

[26] D. A. Forsyth and J. Ponce. *Computer Vision, A Modern Approach*. Prentice Hall, Upper Saddle River, NJ, 2003.

[27] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Trans. Patt. Anal. Mach. Intell.*, 10:439–451, 1988.

[28] F. N. Fritsch and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM J. Numerical Analysis*, 17:238–246, 1980.

[29] S. Gold, A. Rangarajan, C. P. Lu, S. Pappu, and E. Mjolsness. New algorithms for 2D and 3D point matching: Pose estimation and correspondence. *Pattern Recognition*, 31:1019–1031, 1998.

[30] Y. Y. Han and K. Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. *ACM Trans. Graphics*, 36:741–748, 2003.

[31] P. Hanrahan and W. Krueger. Reflection from layered surfaces due to subsurface scattering. In *Proc. SIGGRAPH*, pages 165–174, 1993.

[32] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[33] X. D. He, K. E. Torrance, F. X. Sillion, and D. P. Greenberg. A comprehensive physical model for light reflection. *Proc. SIGGRAPH*, 25:175–186, 1991.

[34] G. H. Healey and T. O. Binford. Local shape from specularity. In *Proc. ICCV*, pages 151–160, 1987.

[35] E. Hecht. *Optics*. Addison Wesley Longman, third edition, 1998.

[36] C. Hernández, F. Schmitt, and R. Cipolla. Silhouette coherence for camera calibration under circular motion. *IEEE Trans. Patt. Anal. Mach. Intell.*, 19:343–349, 2007.

[37] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying BRDFs. *IEEE Trans. Patt. Anal. Mach. Intell.*, 27:1254–1264, 2005.

[38] B. K. P. Horn. *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View*. PhD thesis, MIT, 1970.

[39] B. K. P. Horn. The curve of least energy. *Mathematical Software*, 9:441 – 460, 1983.

[40] B. K. P. Horn and M. J. Brooks. *Shape from Shading*. Cambridge, Mass.: MIT Press, 1989.

[41] G. Horváth and D. Varjú. Underwater polarization patterns of skylight perceived by aquatic animals through Snell's window on the flat water surface. *Vision Res.*, 35:1651–1666, 1995.

[42] L. Hsu, R. Kusner, and J. Sullivan. Minimising the squared mean curvature integral for surfaces in space forms. *Experimental Mathematics*, 1:191–207, 1992.

[43] T. Hwang, J. J. Clarke, and A. L. Yuille. A depth recovery algorithm using defocus information. In *Proc. CVPR*, pages 476–481, 1989.

[44] K. Ikeuchi and B. K. P. Horn. Numerical shape from shading and occluding contours. *Artificial Intelligence*, 17:141–184, 1981.

[45] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *Proc. SIGGRAPH*, pages 511–519, 2001.

[46] H. Jin, A. Yezzi, and S. Soatto. Stereoscopic shading: Integrating shape cues in a variational framework. In *Proc. CVPR*, pages 169–176, 2000.

[47] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Intl. J. Comp. Vis.*, 1:321–331, 1988.

[48] S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[49] G. P. Können. *Polarized Light in Nature*. Cambridge University Press, 1985.

[50] J. J. Koenderink. What does the occluding contour tell us about solid shape? *Perception*, 13:321–330, 1984.

[51] K. Koshikawa. A polarimetric approach to shape understanding. In *Proc. IJCAI*, pages 493–495, 1979.

[52] S. Kotz and S. Nadarajah. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, 2000.

[53] K. J. Krakauer. *Computer Analysis of Visual Properties of Curved Objects*. PhD thesis, MIT, 1971.

[54] D. J. Kriegman and P. N. Belhumeur. What shadows reveal about object structure. *J. Opt. Soc. Am. A*, 18:1804–1813, 2001.

[55] E. P. F. Lafortune, S. C. Foo, K. E. Torrance, and D. P. Greenberg. Non-linear approximation of reflectance functions. In *Proc. SIGGRAPH*, pages 177–126, 1997.

[56] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optim.*, 9:112–147, 1998.

[57] J. H. Lambert. *Photometria sive de Mensura et Gratibus Luminus, Colorum et Umbrae*. Eberhard Klett, Augsburg, 1760.

[58] A. Laurentini. How far 3D shapes can be understood from 2D silhouettes. *IEEE Trans. Patt. Anal. Mach. Intell.*, 17:188–195, 1995.

[59] H. Li, S. C. Foo, K. E. Torrance, and S. H. Westin. Automated three-axis gonioreflectometer for computer graphics applications. *Optical Engineering*, 45:043605, 2006.

[60] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Intl. J. Comp. Vis.*, 60:91–110, 2004.

[61] S. R. Marschner, S. H. Westin, E. P. F. Lafortune, and K. E. Torrance. Image-based bidirectional reflectance distribution function measurement. *Applied Optics*, 39:2592–2600, 2000.

[62] N. Metropolis, W. A. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.

[63] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In *Proc. VisMath*, pages 35–57, 2002.

[64] D. Miyazaki, M. Kagesawa, and K. Ikeuchi. Transparent surface modelling from a pair of polarization images. *IEEE Trans. Patt. Anal. Mach. Intell.*, 26:73–82, 2004.

[65] D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi. Determining surface orientations of transparent objects based on polarisation degrees in visible and infrared wavelengths. *J. Opt. Soc. Am. A*, 19:687–694, 2002.

[66] D. Miyazaki, N. Takashima, A. Yoshida, E. Harashima, and K. Ikeuchi. Polarization-based shape estimation of transparent objects by using raytracing and PLZT camera. In *Proc. SPIE*, volume 5888, pages 1–14, 2005.

[67] D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi. Polarization-based inverse rendering from a single view. In *Proc. ICCV*, volume 2, pages 982–987, 2003.

[68] V. Müller. Polarization-based orientation in a natural environment. In *Intl. Workshop on Biologically Motivated Computer Vision*, pages 98–107, 2000.

[69] D. Mumford. *Elastica and Computer Vision*, pages 491–506. Algebraic Geometry and its Applications. Springer-Verlag, New York, 1994.

[70] S. K. Nayar, X. Fang, and T. Boult. Separation of reflectance components using colour and polarization. *Intl. J. Comp. Vis.*, 21:163–186, 1997.

[71] S. K. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. *Intl. J. Comp. Vis.*, 6:173–195, 1991.

[72] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13:611–634, 1991.

[73] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. Geometric considerations and nomenclature for reflectance. Monograph 161, National Bureau of Standards (US), October, 1977.

[74] J. A. Ogilvy. *Theory of Wave Scattering from Random Rough Surfaces*. Adam Hilger, Bristol, 1991.

[75] M. Oren and S. K. Nayar. Generalization of the Lambertian model and implications for machine vision. *Intl. J. Comp. Vis.*, 14:227–251, 1995.

[76] H. Orland. Mean field theory for optimization problems. *J. de Phys. Lett*, 46:763–770, 1985.

[77] A. P. Pentland. Finding illuminant direction. *J. Opt. Soc. Am.*, pages 448–455, 1982.

[78] B. T. Phong. Illumination for computer generated pictures. *Comm. ACM*, 18:449–455, 1975.

[79] E. Prados, F. Camilli, and O. Faugeras. A unifying and rigorous shape from shading method adapted to realistic data and applications. *J. Mathematical Imaging and Vision*, 25:307–328, 2006.

[80] E. Prados and O. Faugeras. Shape from shading: a well-posed problem? In *Proc, CVPR*, pages 870–877, 2005.

[81] H. Ragheb and E. R. Hancock. Surface radiance correction for shape from shading. *Pattern Recognition*, 38:1574–1595, 2005.

[82] H. Ragheb and E. R. Hancock. The modified Beckmann-Kirchhoff scattering theory for rough surface analysis. *Pattern Recognition*, 40:2004–2020, 2007.

[83] S. Rahmann. Polarization images: A geometric interpretation for shape analysis. In *Proc. ICPR*, volume 3, pages 542–546, 2000.

[84] S. Rahmann. Reconstruction of quadrics from two polarization views. In *Proc. IbPRIA*, pages 810–820, 2003.

[85] S. Rahmann and N. Canterakis. Reconstruction of specular surfaces using polarization imaging. In *Proc. CVPR*, pages 149–155, 2001.

[86] A. Rangarajan, H. Chui, and F. Bookstein. The softassign procrustes matching algorithm. In *Information Processing in Medical Imaging*, pages 29–42, 1997.

[87] T. Rindfleisch. Photometric method for lunar topography. *Photogram. Eng.*, 32:262–277, 1966.

[88] A. Robles-Kelly and E. R. Hancock. A graph-spectral method for surface height recovery from needle-maps. In *Proc. CVPR*, pages 141–148, 2001.

[89] A. Robles-Kelly and E. R. Hancock. Estimating the surface radiance function from single images. *Graphical Models*, 67:518–548, 2005.

[90] M. Saito, Y. Sato, K. Ikeuchi, and H. Kashiwagi. Measurement of surface orientations of transparent objects using polarization in highlight. In *Proc. CVPR*, volume 1, pages 381–386, 1999.

[91] P. Salamon, P. Sibani, and R. Frost. *Facts, Conjectures and Improvements for Simulated Annealing*. Society for Industrial and Applied Mathematics, 2002.

[92] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Intl. J. Comp. Vis.*, 47:7–42, 2002.

[93] Y. Y. Schechner and N. Karpel. Clear underwater vision. In *Proc. CVPR*, pages 536–543, 2004.

[94] Y. Y. Schechner, S. G. Narashimhan, and S. K. Nayar. Polarization-based vision through haze. *Applied Optics*, 42:511–525, 2003.

[95] C. Schlick. An inexpensive BRDF model for physically-based rendering. In *Proc. Eurographics*, volume 13, no. 3, pages 233–246, 1994.

[96] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. CVPR*, pages 519–526, 2006.

[97] R. Sekuler and R. Blake. *Perception*. McGraw-Hill, fifth edition, 2005.

[98] S. Shafer. Using color to separate reflection components. *Colour Research Applications*, 10:210–218, 1985.

[99] P. E. Shames, P. C. Sun, and Y. Fainman. Modelling of scattering and depolarizing electro-optic devices. i. characterization of lanthanum-modified lead zirconate titanate. *Applied Optics*, 37:3717–3725, 1998.

[100] N. Shashar, T. W. Cronin, L. B. Wolff, and M. A. Condon. The polarization of light in a tropical rain forest. *Biotropica*, 30:275–285, 1998.

[101] T. Shibata, T. Takahashi, D. Miyazaki, Y. Sato, and K. Ikeuchi. Creating photorealistic virtual model with polarization based vision system. In *Proc. SPIE*, volume 5888, pages 25–35, 2005.

[102] I. Shimishoni, Y. Moses, and M. Lindenbaum. Shape reconstruction of 3D bilaterally symmetric surfaces. *Intl. J. Comp. Vis.*, 39:97–110, 2000.

[103] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving Poisson equations in computer vision problems. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12:435–446, 1990.

[104] S. Soatto. 3D structure from visual motion: Modeling, representation and observability. *Automatica*, 33:1287–1312, 1997.

[105] B. J. Super and A. C. Bovik. Planar surface orientation from texture spatial frequencies. *Pattern Recognition*, 28:729–743, 1995.

[106] S. Tominaga and N. Tanaka. Refractive index estimation and color image rendering. *Pattern Recognition Letters*, 24:1703–1713, 2003.

[107] K. Torrance and M. Sparrow. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.*, 57:1105–1114, 1967.

[108] A. Treuille, A. Hertzmann, and S. Seitz. Example-based stereo with general BRDFs. In *Proc. ECCV*, pages 457–469, 2004.

[109] S. Umeyama. Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *IEEE Trans. Patt. Anal. Mach. Intell.*, 26:639–647, 2004.

[110] C. L. Vernold and J. E. Harvey. A modified Beckmann-Kirchoff scattering theory for non-paraxial angles. In *Proc. SPIE Conf. Scattering and Surface Roughness II*, volume 3426, pages 51–56, 1998.

[111] A. M. Wallace, B. Liang, J. Clark, and E. Trucco. Improving depth image acquisition using polarized light. *Intl. J. Comp. Vis.*, 32:87–109, 1999.

[112] G. J. Ward. Measuring and modeling anisotropic reflection. In *Proc. SIGGRAPH*, pages 265–272, 1992.

[113] R. White and D. A. Forsyth. Combining cues: Shape from shading and texture. In *Proc. CVPR*, pages 1809–1816, 2006.

[114] L. B. Wolff. Surface orientation from two camera stereo with polarizers. In *Proc. SPIE conf. Optics, Illumination, Image Sensing for Machine Vision IV*, volume 1194, pages 287–297, 1989.

[115] L. B. Wolff. Diffuse-reflectance model for smooth dielectric surfaces. *J. Opt. Soc. Am. A*, 11:2956–2968, 1994.

[116] L. B. Wolff. Polarization vision: A new sensory approach to image understanding. *Im. Vis. Comp.*, 15:81–93, 1997.

[117] L. B. Wolff and T. E. Boult. Constraining object features using a polarisation reflectance model. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13:635–657, 1991.

[118] L. B. Wolff, S. K. Nayar, and M. Oren. Improved diffuse reflection models for computer vision. *Intl. J. Comp. Vis.*, 30:55–71, 1998.

[119] K. Y. K. Wong and R. Cipolla. Reconstruction of sculpture from its profiles with unknown camera positions. *IEEE Trans. Im. Proc.*, 13:381–389, 2004.

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

[121] R. J. Woodham. Analyzing images of curved surfaces. *Artificial Intelligence*, 17:117–140, 1981.

[122] R. J. Woodham. Gradient and curvature from photometric stereo including local confidence estimation. *J. Opt. Soc. Am. A*, 11:3050–3068, 1994.

[123] P. L. Worthington. Re-illuminating single images using albedo estimation. *Pattern Recognition*, 38:1261–1274, 2005.

[124] P. L. Worthington and E. R. Hancock. New constraints on data-closeness and needle map consistency for shape-from-shading. *IEEE Trans. Patt. Anal. Mach. Intell.*, 21:1240–1267, 1999.

[125] T. P. Wu and C. K. Tang. Separating specular, diffuse, and subsurface scattering reflectances from photometric images. In *Proc. ECCV*, volume 2, pages 419–433, 2004.

[126] T. P. Wu, K. L. Tang, C. K. Tang, and T. T. Wong. Dense photometric stereo: a markov random field approach. *IEEE Trans. Patt. Anal. Mach. Intell.*, 28:1830–1846, 2006.

[127] R. Zhang, P. S. Tsai, J. E. Cryer, and M. Shah. Shape from shading: A survey. *IEEE Trans. Patt. Anal. Mach. Intell.*, 21:690–706, 1999.

[128] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13:680–702, 1991.

[129] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *Intl. J. Comp. Vis.*, 49:215–227, 2002.

[130] T. Zickler, R. Ramammoorthi, S. Enrique, and P. Belhumeur. Reflectance sharing: Predicting the appearance from a sparse set of images of known shape. *IEEE Trans. Patt. Anal. Mach. Intell.*, 28:1287–1302, 2006.