Shape from Shading under Relaxed Assumptions

Rui Huang

Submitted for the degree of Doctor of Philosophy

University of York

Department of Computer Science

Feburary 2012

Abstract

Shape-from-shading is a classical problem in computer vision. The aim is to recover 3D surface shape from a single image of an object, based on a photometric analysis of the pattern of shading. Since the amount of light reflected by a surface is a function of the direction of the incident light and the viewer, relative to the local surface normal, image intensity conveys information about surface orientation and hence, indirectly, surface height. This thesis aims to relax some of the assumptions typically made in the shape-from-shading literature. Our aim is to move shape-from-shading away from the overly simplistic assumptions which have limited its applicability to images captured in lab conditions. In particular, we focus on assumptions of surface smoothness, point source illumination and reconstruction in the surface normal domain.

In Chapter 3, we relax the assumption of a smooth surface. We exploit a psychology-inspired heuristic that pixels need only have a similar surface orientation if they are both in close proximity and have a similar intensity. This leads to an adaptive smoothing process which is able to preserve fine surface structure. We adapt a geometric shape-from-shading framework to overcome the problem of normals "flipping" between solutions which alternately satisfy data-closeness and smoothness terms. Under the classical assumption of point source illumination, we show that our method significantly outperforms a number of previously reported methods.

In Chapter 4, we relax the assumption of illumination being provided by a single point light source. Specifically, we consider environment illumination in which lighting is represented by a spherical function which describes the incident radiance from all directions in the scene. We use the well known result that Lambertian reflectance acts like a low pass filter and hence the convolution of environment lighting and surface reflectance can be efficiently represented using a low order spherical harmonic. With an order-1 approximation, we show how the image irradiance equation can be solved as a quadratically-constrained linear least squares optimisation. The global optimum is found using the method of Lagrange multipliers. The order-2 case is non-convex and prone to converge on local minima if solved using local optimisation. We reformulate the problem as a bilinear system of equations which leads to an efficient and robust solution method. In both

cases, we incorporate a structure-preserving smoothness constraint based on ideas from Chapter 3 to regularise the problem.

In Chapter 5, we continue with the relaxed illumination assumption (i.e. we model environment illumination), but we develop algorithms which operate in the domain of surface height rather than surface normals. This has the advantage of reducing the dimensionality of the problem at the expense of increased complexity. Moreover, integrability is implicitly enforced via the problem formulation. We describe two contributions. The first is a linear method for recovering surface height directly from images formed by taking ratios between colour channels. In this case, the nonlinear normalisation term is factored out. This allows us to form a linear system of equations relating image intensity and surface height via a finite difference approximation to the surface gradient. Finally, we relax the assumption that the object must be globally convex (i.e. contains no self occlusions). We show that self occluded intensity can be related to unoccluded intensity via a quadratic inequality constraint. This is too weak a constraint to be used for shape-from-shading on its own. However, we use it to develop an occlusion-sensitive surface integration algorithm. We show that the problem can be formulated as a convex optimisation and solved using semidefinite programming.

Contents

List of Figures

List of Tables

List of Symbols

$I(x)$ the image irradiance at the point x

ρ albedo

$\mathbf{n}(x)$ surface normal at point x

$R(\mathbf{n}(x))$ radiance function at point x

$\vec{\omega}$ a light source in vector representation

$L(\vec{\omega})$ the intensity of light source in direction $\vec{\omega}$

$V(x, \vec{\omega})$ visibility function at point x with the light source direction $\vec{\omega}$

$u(x)$ surface height function

p, q the gradients of the surface height in x and y directions

Θ rotation matrix

Y_l^m spherical harmonic of degree l and order m

ϕ azimuth angle

θ elevation angle

$c_{l,m}$ projection coefficient of degree l and order m for illumination function

$d_{l,m}$ projection coefficient of degree l and order m for reflectance function

Acknowledgements

First, I would like to acknowledge my parents for their support and encouragement during my PhD study.

I would like to gratefully acknowledge the enthusiastic supervision of Dr. William A.P. Smith for his great help whenever I have something excited or confused in mind. I would like to say, he is the best supervisor I have seen around, not only the method, but the way of thinking he gives me is great. I would also like to gratefully acknowledge Prof. Edwin Hancock, Prof. Richard Wilson, and Dr. Andrian Bors for their advice during the last 4 years.

Finally, I would like to thank my friends for their help and support.

Declaration

I declare that the research described in this thesis is original work, which I undertook at the University of York during 2008 - 2012. Except where stated, all of the work contained within this thesis represents the original contribution of the author.

I declare that all the work in this thesis is solely my own, except where attributed and cited to other authors. Most of the material in this thesis has been previously published or will be published. Below is a complete list of publications on which Chapters 2 to 4 are based.

1. R. Huang and W.A.P. Smith, Shape-from-shading under complex natural illumination. In Proc. ICIP (Special Session on Shape-from-shading), 2011.

2. R. Huang and W.A.P. Smith, Photometric Stereo under Low Frequency Environment Illumination. In Proc. ISVC, 2010.

3. R. Huang and W.A.P. Smith, A shape-from-shading framework for satisfying data-closeness and structure-preserving smoothness constraints. In Proc. BMVC, 2009

4. R. Huang and W.A.P. Smith, Structure-preserving Regularisation Constraints for Shape-from-Shading. In Proc. CAIP, pages 865-872, 2009

Chapter 1

Introduction

1.1 Introduction

Whenever a shape is illuminated by a light source, shading exists. In art, shading is used as a cue to shape by applying media more densely or with a darker shade for areas which are shadowed or angled away from the illumination, and less densely or with a lighter shade for areas which are brightly lit. In computer graphics, shading (or rendering) refers to the process of adjusting a color based on the angles of incident lights, viewing angles and the distance from the illumination to create a photo realistic effect. Computing shading patterns from descriptions of the shape of an object and information about the illumination is well understood. There is a vast literature of computer graphics techniques available to realistically render a wide range of surface types. The question we address in this dissertation is whether it is possible to recover shape from the shading information observed in one image alone. Figure 1.1 gives an overview of the problem.

Given a monocular image, light source and some auxiliary information, shape from shading is described as finding the shape of a smooth object from the image intensities [23]. This method was proposed to be a complementary to binocular stereo which relies on matching distinct features between images and fails on smooth objects. It is extended by photometric stereo where the viewing direction remains fixed but illumination varies. This research had a quite practical application background: to recover the surface of the moon at that time. It may also be used for face recognition where only a single training image is available.

However, shape from shading problem itself is an ill-posed problem. Even with naive lighting, reflectance and surface structure assumptions, ambiguities still exist and some questions about the uniqueness of solutions are still open. Moreover, a lack of methods which make realistic assumptions have meant the method has rarely been applied to real practical applications.
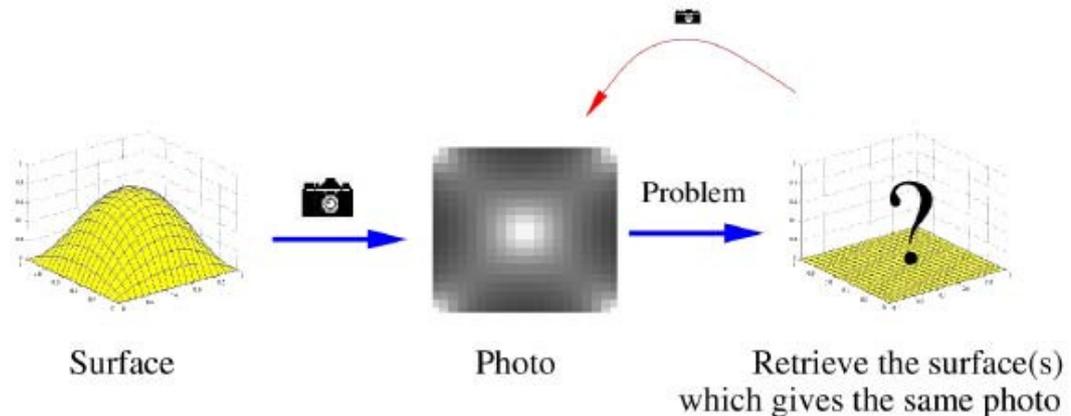
Figure 1.1: The "Shape From Shading" problem (from [44]).

Since the classic study by Horn and Brooks [25], a large volume of work on shape from shading has been undertaken. Some of this addresses the classic ambiguities, meanwhile, most paper focus on introducing more realistic assumptions or applying start-of-the-art mathematical tools to computing a result. We categorise these methods into three groups: methods solving partial differential equations; methods using minimization, and methods which attempt to relax assumptions. They will be discussed in detail in Chapter 2.

In this chapter, we will first formulate the shape from shading problem, some topics such as illumination estimation, shadowing and albedo will be discussed. Then, the classic ambiguities of the shape from shading problem will be listed. We will also show the motivation and contribution of each of the chapters containing our technical contributions.

1.2 Image Formation

Shape from shading effectively aims to invert the image formation process. By understanding the physical processes involved in moving from a description of a scene to an image of its appearance, we can derive a model that relates what we observe (an image of an object) to what we wish to estimate (the object shape). We divide our consideration into three parts: 1. reflectance models, 2. albedo and light source direction estimation, 3. visibility or cast shadows.

1.2.1 Reflectance Models

The reflectance of light by a surface is a key concept in understanding the shape from shading problem. It is the reflectance properties of a surface that determine how its orientation will relate to
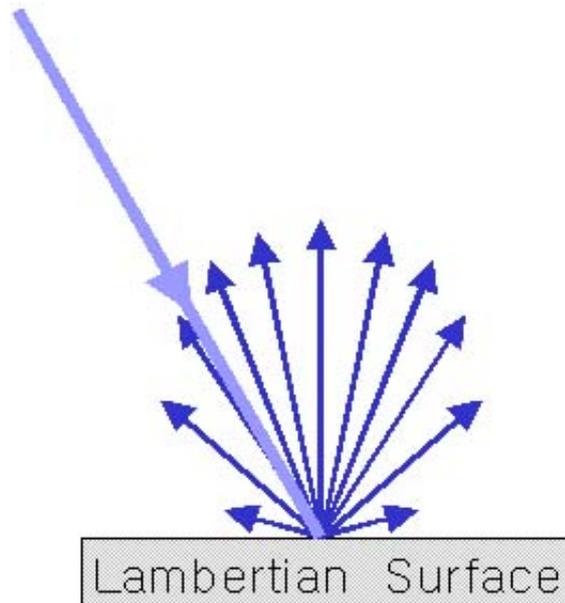
Figure 1.2: The Lambertian Surface. Light is scattered equally in all directions.

the amount of light reflected towards the viewer. Early measures of brightness used the candela, the only unit still defined with reference to human sensations. A more useful measure was introduced by the U.S. National Bureau of Standards(NBS): the *bidirectional reflectance distribution function* (BRDF) and a number of derived quantities [37]. The BRDF is determined by the microstructure of the object. The shape from shading problem is greatly simplified if the BRDF is known in advance, along with the light source and the viewer geometry.

The simplest model of reflectance is the ideal diffuser (or Lambertian) surface in which light is scattered equally in all directions in Figure 1.2. The brightness of such a surface can be shown to be proportional to the cosine of the incident angle, the angle between the surface normal and the incident ray. The general solution of the shape from shading problem revolves around the so-called *image radiance equation* relating image radiance to scene radiance:

$$I(x) = \rho(x)R(\mathbf{n}(x)) \tag{1.1}$$

where $I(x)$ is the image irradiance at the point x, while $R(\mathbf{n}(x))$ is the radiance of a surface patch with unit normal $\mathbf{n}(x)$ at the corresponding image point. Actually, image radiance is not equal to scene radiance, only proportional to it. The $\rho(x)$ is the refection coefficient, or albedo. It is defined as the ratio of reflected radiation from the surface to incident radiation upon it. We assume the constant and unit albedo here, which means the surface should be a textureless surface and do not absorb any reflected radiation. Objects with and without texture are shown in Figure 1.3.

In reality, the reflectance is not only decided by the surface normal and the light source direc-

3

(a)   (b)

Figure 1.3: Objects with texture in (a) and without texture in (b). We are focus on the object without the texture here

tion. In addition, the direction of the viewer is important when considering specular reflectance. Two popular models for simulating such reflectance, one ad hoc, the other physically-based, are the Phong model [43] and the Torrance and Sparrow model [58].

The Phong model is an empirical model, which is not based on physics but physical observation. The reflected intensity caused by the specular highlight is assumed to be a function of $(cos(a))^n$ with $n >= 200$ for a shiny surface and n small for a dull surface from the observation that for very shiny surfaces the specular highlight is small and the intensity fell off rapidly, while for duller surfaces it was larger and fell off more slowly, where a is defined as the angle between the view direction and the direction of mirror reflection. The complete illumination intensity model for reflection includes diffuse reflection from ambient light and a point light source, together with specular refection.

The Torrance and Sparrow model is a physically based model which captures important effects that the empirical Phong model cannot capture. Note that the specular lobes of most real life surfaces can exhibit off-specular behaviour. Phong model can only explain the perfect specular reflection in the specular lobe, but fails to describe the reason for off-specular behaviour. TS model assumes that reflectance consists of two components: the part which is associated with bulk material effects related to a Lambertian lobe at a particular position on the surface and the part which is purely related to surface scatter. The weakness that fails to capture the off-specular forward scattering in Phong model is overcome. By assuming the specular part is got from a zero-

mean Gaussian surface, the specular reflection is not a pure specular lobe towards one direction, but a collection of lobes satisfy an Gaussian distribution.

Most existing work deals with the shape from shading problem only under Lambertian reflectance. There have been few attempts to recover surface shape from the specular component of reflectance, the exception being Ragheb and Hancock [48]. Zickler et al. [70] propose a way to remove the specular reflectance from color images using color subspaces. By exploiting color information in images of dichromatic surfaces, a class of invariants is derived from illuminant-dependent 'subspaces' of RGB color space. In this work, the specular removal source-dependent in that is dependent on knowing (or estimating) the color of the light source. They project the color vector in the SUV color space to a new space using the light source vector as one of its axis and then project the Lambertian part back to RGB space. After removing the specular part, a shape from shading method is applied to recover the shape.

More realistic models of surface reflectance may also include effects such as inter-reflection and macroscopic shadowing. The problem with using such models in the context of shape from shading is that it becomes infeasible to find a simple relationship between intensity and shape. In practice, models such as Phong or Torrance and Sparrow offer enough realism for high accuracy.

In this thesis, we focus only on Lambertian reflectance. However, the models in Chapter 4 would extend naturally to more complex reflectance models as it has been shown previously that spherical harmonics can be used to efficiently describe glossy as well as matte surfaces. Instead, our principal focus is on objects immersed in a spherical illumination environment. In that case, image irradiance for a Lambertian surface is given by an integral over the upper hemisphere:

$$I(x) = \rho(x) \int_{\Omega(\mathbf{n}(x))} L(\vec{\omega})V(x,\vec{\omega})(\mathbf{n}(x) \cdot \vec{\omega})d\vec{\omega}, \qquad (1.2)$$

where $I(x)$ is the image irradiance at point x, $\rho(x)$ is the diffuse albedo, $L(\vec{\omega})$ is the illumination function with incident radiance from direction $\vec{\omega}$, $V(x,\vec{\omega})$ is a visibility function with the value 0 if x is occluded in direction $\vec{\omega}$, 1 the otherwise. The result of this equation is actually a calibrated value of the irradiance measure.

The vast majority of existing shape-from-shading work assumes a point light source (i.e. $L(\vec{\omega})$ is a Dirac delta function), constant and known albedo and no occlusion (i.e. $V(x,\vec{\omega}) = 1$ for all light source directions). It is only very recently, that there has been a small amount of work which takes environment illumination into account. This is described in the next chapter.

1.2.2 Albedo and Light Source Direction Estimation

The albedo is the intrinsic reflectivity of a surface, i.e. the proportion of light which is not absorbed by the surface. In the shape from shading literature, unit albedo is almost universally assumed. However, an accurate estimate of the albedo is useful if shape from shading algorithms are to be applied in realistic scenarios. Albedo may be determined experimentally if a calibrated camera and light source is used. However, for real images with no calibration computing the spatially varying albedo map is more difficult.

The most well known albedo estimation method was proposed by Blake [9]. The *log* image is a summation of albedo contribution and the intrinsic image contribution(image represents the intrinsic characteristics such as illumination, shape and the reflectance). The albedo information is acquired by thresholding to remove the latter part. The problem for this method is that the threshold used to decide the change of the albedo is still missing which is essential to the final result. Because the information of the threshold exists in that of the shape, which leads to a paradox that which comes first, the chicken or the egg. The state-of-the-art paper on albedo estimation is present by Biswas and his colleagues [8]. The input are image and average surface normal. A non-stationary stochastic filter framework is present to refine the albedo in iterations.

It is worth mentioning that there are few shape from shading paper working on both recovering shape and albedo information. A model based method proposed by Smith and Hancock [54, 55] focusses on the recovery of facial surfaces. A statistical model of variation in surface normal direction over the face surface is trained and used to act as a constraint on the surface. To overcome problems with applying linear statistical methods to non-linear unit surface normal vectors, the surface normals are projected to the tangent plane to the local average normal direction using the azimuthal equidistant projection. The average face is used as an initialisation. With the statistical model to hand, the algorithm proceeds by alternately finding the best fit of the model to the data and locally imposing the brightness constraint by rotating the normals back onto their irradiance cones.

Information about the light source is also important for the shape from shading problem, because all reflectance models use the intensity and direction of the light source to compute the amount of light reflected towards the viewer. In an experimental environment, the light source properties may be measured empirically, but for the general shape from shading problem, all we have is an image of the object. The direction of the light source also effects shadowing, which is caused by the cut of the ray through the boundary, a cue for shape recovery. In general, the shadow needs to be removed, which will be discussed in the next subsection.
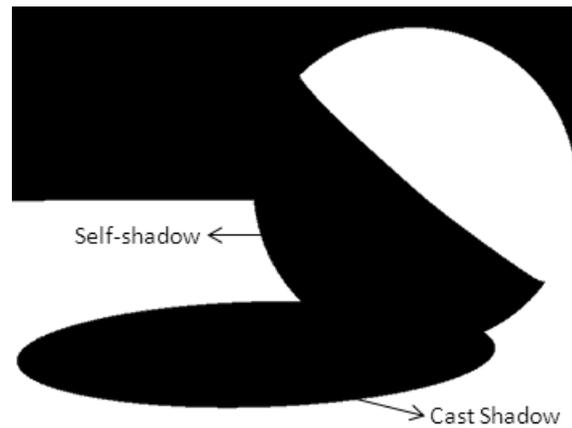
Figure 1.4: The self-shadow and cast shadow.

The pioneering method is proposed by Pentland in 1982 [40] and was extended by Chojnacki and Brooks [11]. The assumption of sampled points having the same distribution as a spherical object is the fundamental assumption of their method.

A simple estimation of illumination direction is introduced by Zheng and Chellappa [69]. In their paper, the illumination is estimated on smooth patches. The gradient map of those smooth patches are established at first, the difference of the average of the gradient map is related to the direction of the light source, and it is derived by solving a linear system of equations.

1.2.3 Visibility

Shadows are caused by the ray from the light source cutting by the boundary of surface which contribute 0 values in the integral of illumination function. Since the value of the visibility function is a combination of surface shape, light source and view point, the best way to solve it is removing it from the first step. There are two kinds of shadows: self-shadow and cast shadow. Cast shadow is caused by the light source being intercepted by another object, which provides little information for shape recovery. Self-shadow means the shadow is caused by part of the object, which is always believed to be the boundary information of the shape. (See Figure 1.4) In most situation, the shadow should be removed before the shape recovery process. It is every easy for human to point out the shadow from the surface because the inner structure maintains consistent and even the part which is saturated (the area where the mirror reflectance happens) can be recovered. Finlayson et al. [16] extract the intrinsic images using entropy minimization. The key idea in this paper is that the correct projection is that which minimizes entropy in the resulting invariant image. The intensity information is treated as the magnitude of the vector defined in the colour
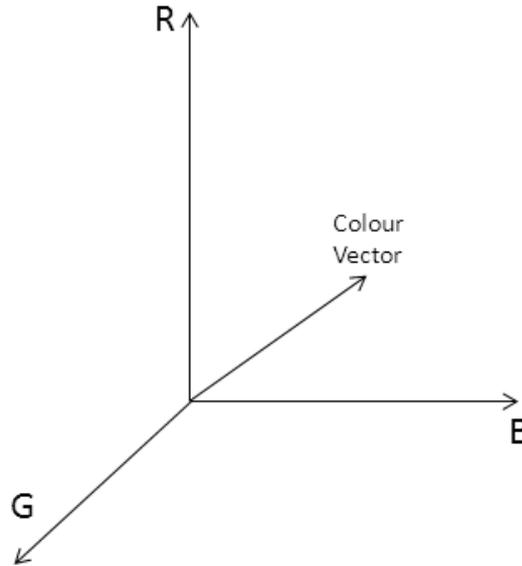
Figure 1.5: The Colour Vector in RGB space. Intensity is defined as the magnitude of the vector.

channel. Removing the shadow is a minimisation of the whole image defined within an entropy function in Figure 1.5. The algorithm follows the way of removing the shadow by human, but the problem is if the shadow is overwhelming, it may render all the image in the shadow. And it also fails when saturation occurs.

1.3   Classical Ambiguity

With a point light source, known constant albedo information, shape from shading problem is still an ill-posed problem. From each single intensity measurement, we aim to estimate the two components of surface orientation. Two classical demonstrations of the difficulties of shape from shading are the concave/convex and bas-relief ambiguity.

1.3.1   Concave/Convex Ambiguity

The concave/convex ambiguity has been noted in psychology. It is caused by the human assuming the light source to be at the top of the picture rather than at the bottom.
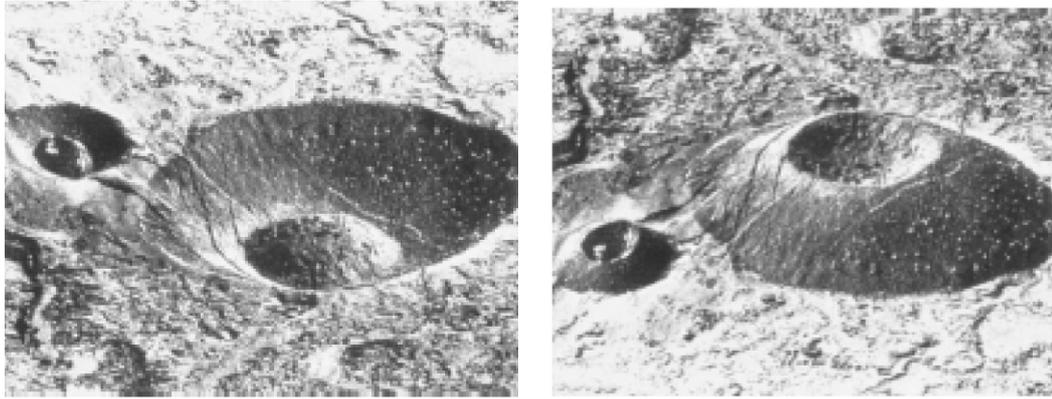
<center>(a)   (b)</center>

Figure 1.6: Examples of Concave/Convex ambiguity: The same image can produce two different surfaces which is a crater in (a) and volcano in (b).

Figure1.6 is a great example of such ambiguity. It leads to two possible solutions for shape from shading problem. There are few attempts to resolve this ambiguity in the literature. Perhaps the only examples are to use mutual illumination and interreflection [17, 18] or introducing a light source attenuation factor [46].

1.3.2   Bas-relief Ambiguity

The bas-relief originated from sculpture art. It is also called low-relief, and is a sculpture technique in which figures and/or other design elements are just barely more prominent than the (overall flat) background. In the shape from shading domain, the bas-relief ambiguity is a problem related to the unknown parameter of light source. In [6], Belhumeur and his colleagues analyze the results of the images under different shape, light source and albedo, and conclude that the light source and albedo are critical in determining the interpreted shape. In their paper, they formulate a bas-relief function $\bar{f}(x,y) = \lambda f(x,y) + \mu x + \nu y$ where $\bar{f}(x,y)$ is the recovered height function at point (x,y), and $f(x,y)$ is the ground truth height function. The coefficients λ, μ and ν are decided by the albedo and illumination direction.

x

(a)   (b)
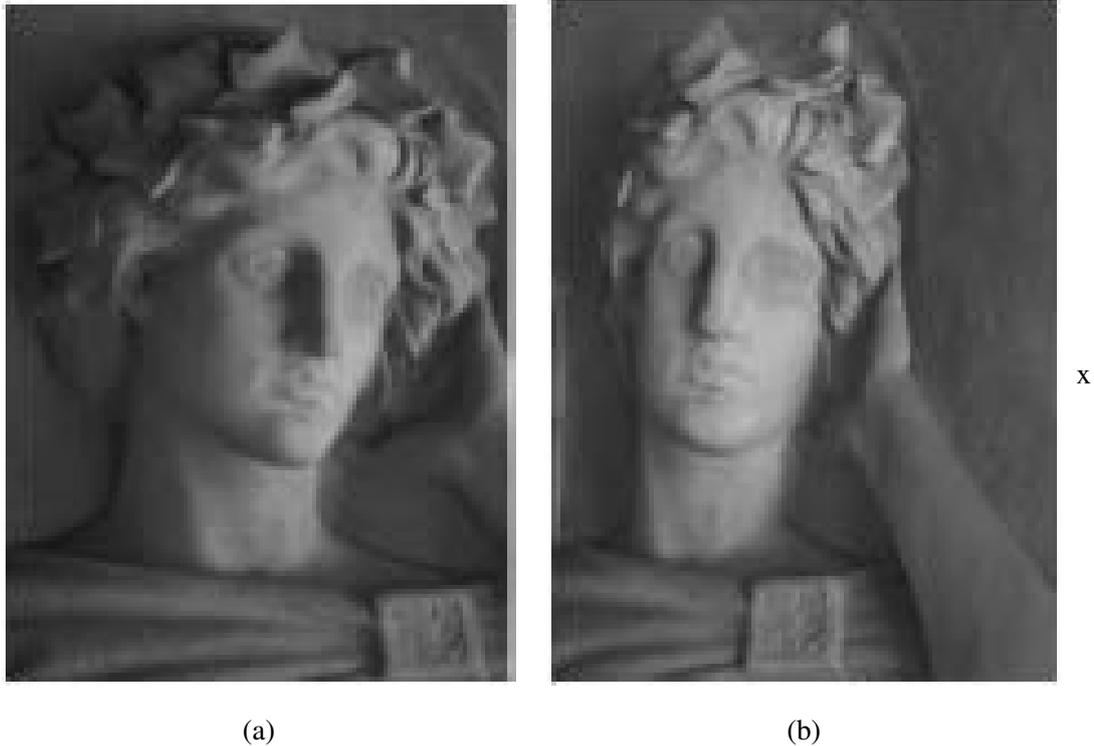
Figure 1.7: Examples of Bas-relief ambiguity: (a-b) Bas-relief Ambiguity: Frontal and side views of a marble sculpture. The different view point of the same object can produce two different surfaces which is normal in (a) and flatter in (b).

Resolving the bas-relief ambiguity requires additional information. Most previous work simply solves for a solution up the bas-relief ambiguity [5].

A key observation of this thesis is that complex environment illumination helps to overcome these ambiguities. Under complex illumination, the set of surface normal directions which can give rise to a particular colour are likely to be better localised than under point source illumination, where a cone of possible directions exists.

1.4   Modern Shape from Shading

Photometric estimation of shape has almost exclusively made the strong assumption that the illumination in a scene can be approximated by a single point light source [13]. This gives rise to simple expressions for image irradiance and, in turn, simple constraints on possible normal direction. For example, a Lambertian surface illuminated by a point light source defines a cone of possible normal directions, the opening angle of which is determined by the image intensity [64].

Aside from the unrealistic nature of this assumption, recent work [27] suggests that the com-

plexity of natural illumination may be a help rather than hindrance to the shape-from-shading problem. This is particularly so for the case of colour images where the illumination environment in each colour channel is independent. For example, in a typical indoor scene there may be natural light of a certain colour, artificial light of another colour and light reflected from any number of coloured objects. Points oriented towards each of these sources will predominantly reflect the colour of the source, providing a strong cue as to the orientation of the surface.

There have been two notable attempts to incorporate complex illumination into photometric shape recovery. The first assumes that illumination is perfectly ambient, such as may be experienced outdoors on a cloudy day. Most recently, Prados et al. [47] showed that this assumption leads to a a strongly non-local and non-linear Integro-Partial Differential Equation. The second uses the same spherical harmonic illumination model as in this paper. Basri et al. [5] assume that a spectral analysis of the photometric stereo data matrix will yield eigenvectors which coincide with the components of the surface normals. With both unknown lighting and shape, this factorisation is ambiguous and other constraints must be exploited.

To conclude, environment illumination is both a better approximation to reality and potentially provides a stronger constraint than point source illumination. Hence, in Chapters 4 and 5 we focus on such illumination models.

1.5   Thesis Structure

The thesis is structured as follows:

- Chapter 2 gives an overview of the shape from shading literature. PDE-based methods, minimization methods and methods approximating the irradiance function are discussed in detail. Some interesting method which do not fall into this classification or which address the resolution of ambiguities are also discussed.

- Chapter 3 introduces a new framework for a structure preserving smoothness constraint. We describe a classical Lambertian shape-from-shading algorithm (i.e. point source assumption) which requires a surface to be only piecewise smooth. The method is very simple yet outperforms a number of well known shape-from-shading algorithms. We provide results on objects with complex finescale detail and show that our method is able to preserve this detail.

- Chapter 4 begins by giving some background on environment illumination and its representation using spherical harmonics. Then we address the problem of minimising an objective

comprised of both irradiance and smoothness terms under a quadratic equality constraint. The methods for solving both first order illumination function and second order form are described and results provided on both synthetic and realworld objects under complex illumination.

- Chapter 5 focusses on the problem of surface height recovery. We first present a linear method for height from shading based on ratios taken between colour channels of an object under complex illumination. Second, we present an occlusion-sensitive surface integration method which takes an image and an estimated normal map as input. This problem is convex and the solution is found using semi-definite programming.

- Chapter 6 provides some conclusions and points out the direction for future research.

Chapter 2

Literature Review

2.1 Introduction

Shape-from-shading has attracted research attention for over 30 years. Hence, the literature is large and broad. Two well known surveys on the subject have been conducted: Zhang et al. [68] and Duroua et al. [13]. We classify existing work into three main categories: methods of solution of partial differential equations, methods using minimization or variational methods, and methods approximating the image irradiance equation. Work which does not fit into this categorisation is considered separately at the end of the chapter.

We will look at the simplest assumption for shape from shading problem, which works on intrinsic images (i.e. unit albedo is assumed and point source illumination is known) [3]. The shape from shading problem can be modelled by an irradiance function

$$I(x) = R(\mathbf{n}(x)) \tag{2.1}$$

where $I(x)$ is the image irradiance at the point x. While $R(\mathbf{n}(x))$ is the radiance of a surface patch with unit normal $\mathbf{n}(x)$ which corresponds to the point $(x, u(x))$. The height function u, which is unknown, has to be reconstructed on a compact domain $\Omega \subset \mathbf{R}^2$. Assuming a known point light source $\vec{\omega} = (\omega_x, \omega_y, \omega_z) \subset (R)^3$, then

$$R(\mathbf{n}(x)) = \vec{\omega} \cdot \mathbf{n}(x). \tag{2.2}$$

2.2 Partial Differential Equation Methods

A surface normal can also be represented as the first order partial derivative of the height:

$$\mathbf{n}(x) = \frac{1}{\sqrt{1 + p(x)^2 + q(x)^2}}(-p(x), -q(x), 1) \tag{2.3}$$

where $p = \partial u / \partial x$ and $q = \partial u / \partial y$, so that $\nabla u(x) = (p(x), q(x))$ Taking the Lambertian model into consideration, the shape from shading problem can be seen as a non-linear first order partial equation. The result of function u is that of shape:

$$I(x)\sqrt{1 + |\nabla u(x)|^2} + (\omega_x, \omega_y) \cdot \nabla u(x) - \omega_z = 0, \ x \subset \Omega \tag{2.4}$$

This a first order non-linear partial differential equation (PDE) of the Hamilton-Jacobi form. Points with maximal intensity have a single situation where the surface normal $\mathbf{n}(x)$ and light source direction $\vec{\omega}$ point at the same direction. These points are usually called the singular or critical points.

We consider the equation that appears most commonly: frontal illumination situated at infinity, so $\vec{\omega} = (0, 0, 1)$. Then the equation become a Eikonal Equation [10]:

$$|\nabla u(x)| = f(x), \ x \subset \Omega \tag{2.5}$$

where:

$$f(x) = \sqrt{\frac{1}{I(x)^2} - 1}. \tag{2.6}$$

In the past few years, new models have appeared. The main modifications of these models are changing the classical assumption in order to adapt to the real-life application. Among all the methods, the major modification is replacing the conventional orthographic projection with a more realistic perspective projection. The new model problem is called perspective shape from shading. Recent work described the non-linear PDE that arises under perspective projection [45, 46], the light source is no longer to be located at infinity, but at the centre of the projection.

Solving a partial differential equation usually requires the knowledge of boundary information to select a unique solution. There are two types of boundary information: the boundary value that imposes a value to the solution (Dirichlet type boundary condition) or the boundary value on the normal derivative(Neumann type boundary condition). Objects with occluding boundaries can naturally be used as boundary conditions $\partial \Omega$ of the reconstruction domain Ω. In Figure 2.1, the part of the image representing the object in grey levels is Ω (Dirichlet type), then $\partial \Omega$ coincides with the occluding boundary (Neumann type).

The solution of this partial differential equation for shape from shading problem was first proposed in Horn's PhD thesis in 1970 [23]. In order to solve this equation, he introduced the characteristic strip expansion method. The partial differential equation is converted into five ordinary differential equation, also called characteristic equation. Under the following two conditions, the equation can be solved. (i) the function u has to be a class C^2 (i.e. smooth); (ii) the (x, y, u, p, q)
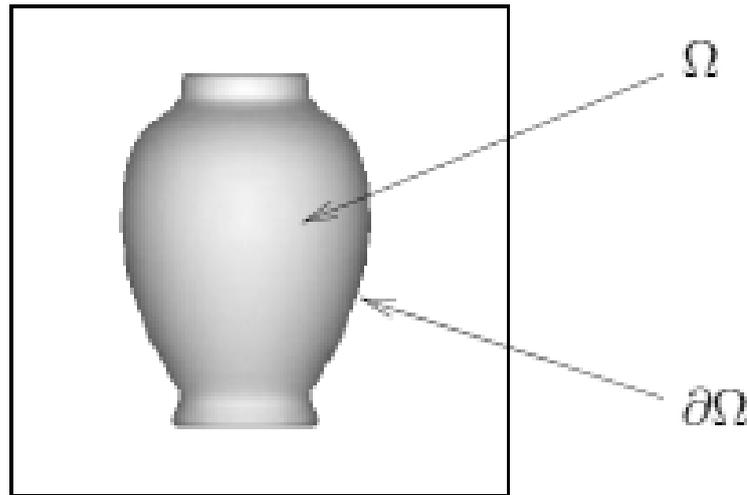
Figure 2.1: Two Boundary Conditions

in these ordinary differential equation must be known at every point of the curve called initial curve. This means two boundary information are both required in their method. This method is subject to the noise of brightness measure for the deviation of the characteristic strip along the image. In addition, the initial boundary data are impossible to obtain reliably in most situations, and the accumulation of errors in integration causes further difficulties.

A consistent numerical scheme based on Sethian's [29] fast marching method can be used to compute the reconstructed surface. Fast marching method is a useful method to solve the partial differential equation. In each iteration, the updated height of shape at specific point is calculated using neighbouring height information together with a compensation function. This compensation function is derived by the difference between the intensity from the predicted height and the shading information. The updating procedure always flows from small to large values of the solution u. The algorithm has been extended to the non-frontal illumination situation by rotating to the light source coordinates. But the accuracy and complexity are dependent on the structure of the surface. The point with the minimal height is required, and it is difficult to identify in most situations.

Recently, most papers focussing on PDE methods use the features of Hamilton Jacobi type equations. One of the weaknesses of PDE-based methods is that the surface must be differentiable. This is hard to achieve in reality. The edges and discontinuities prevail in the surface of the object. Thus, the viscosity solution is proposed to solve this problem. The theory of viscosity solution was first used in shape from shading problem by Rouy and Tourin [51]. The viscosity solutions

are PDE solutions in a weak sense. The differentiable of the surface is not a must and it allows for sharp edges. Their work is based on the continuous surface, which causes the problem when the boundary data are incorrect. A discontinuous scheme is also proposed by Prados to find a discontinuous result [45].

Another weakness of the PDE method is that the boundary information is needed to find a unique answer. To overcome this obstacle, Dupuis and Oliensis [12] consider C^1 solutions that only need to specify the value at the critical points (the points have the greatest intensity) which are locally minimal but no values at the boundary.

Most papers deal with the shape from shading problem under the orthogonal projection, which assumes all the light is parallel to each other. More recently, perspective projection has been considered [45], which is a more realistic model in many cases. It takes the geometry into consideration. Perspective shape from shading (PSFS) leads to an equation that is still a non-linear PDE but the perspective projection provides different boundary conditions from the conventional ones. In 2005, Prados and coworkers incorporated the perspective projection into the partial differential equation, and also introduced an attenuation term $\frac{1}{r^2}$, where r is the distance of the surface to the light source. Solving the Hamilton equations for the PDE, the shape from shading problem is proved to be a well-posed problem under their framework.

The latest paper on PDE method is proposed by Ecker and Jepson [14]. In fact, their equation is quite different from existing PDE method. They describe the problem as a polynomial system.

$$(1 + p^2 + q^2)I^2 - (-ap - bq + c)^2 = 0 \tag{2.7}$$

under the constraint that

$$-ap - bq + c \geq 0 \tag{2.8}$$

A linear search method is used to obtain the solution. This can obtain an exact solution in the absence of noise. To handle noise they propose a convex relaxation which they solve using a homotopy solver.

2.3   Minimization Methods

Minimization methods are one of the most general approaches to solving the shape from shading problem. It is straightforward to formulate an error function that measures the difference between observed and predicted intensity. The idea is then to minimize this value according to certain constraints that make sense in the context of shape from shading.
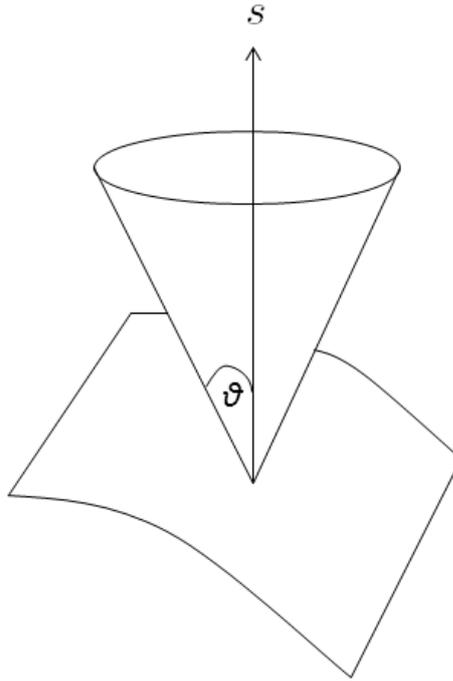
Figure 2.2: The cosine cone of the normal space

Reviewing the irradiance equation, this equation is satisfied in a continue space in which the inner product of surface normal $\mathbf{n}(x)$ and the light source $\vec{\omega}$ has the same value. It consists of a right circular cone with the light source direction as its axis and a half angle of ϑ. This angle is defined as the opening angle. See Figure2.2 All the normals that are ϑ degrees away from the point light source S have the same result. Illumination equation alone makes it an ill-posed question. Additional assumptions are made to make shape from shading a well-posed problem. Most commonly used assumptions are smoothness and integrability constraints.

Variational calculus is the mathematical foundation of these methods. Suppose we seek a smooth surface tat satisfies all the constraints over some domain. It is useful to define a non-negative objective function that measures the distance of a particular result to a satisfactory solution. The result here is the surface we seek, and the searching for the best result is minimizing the objective function. The search for a function that minimises an integral expression is the major concern of the calculus of variations [23]. The value of the extreme objective function must satisfy an associated Euler equation. The surface recovery problem here has been transformed to one of solving a system of partial differential equations.

A surface that we are seeking should be a global minimum of the objective function. However, a local minimum result also satisfies the Euler equation. This result is difficult to avoid given that we shall be dealing with objective functions involving a reflectance map whose analytic form

may not be known in advance. Even if we obtain a surface that generates a global minimum of the objective function, the result is still not the exact surface for the problem. For example, in the presence of noise, there may not be a smooth surface that satisfies the image irradiance equation exactly. We are solving for a surface that minimises the integral of the square of the difference between $I(x)$ and $R(\mathbf{n}(x))$. The surface recovered will depend on the formulation of the smoothness constraint.

The way of deriving iterative schemes for recovering surface shape is as follows

$$E(u(x,y)) = \iint_\Omega F(x,y,u(x,y)),...) + \lambda S(x,y,u(x,y)),...) + \gamma Int(x,y,u(x,y)),...) + ...dxdy \tag{2.9}$$

where $F(x,y,u(x,y)),...)$ is defined as the non-negative error function of the radiance function, $S(x,y,u(x,y)),...)$ and $Int(x,y,u(x,y)),...)$ are the smoothness and integrability terms with the coefficients λ and γ. The framework of this scheme was first introduced by Horn and Brooks in 1986 [25]. The objective function is the difference between the intensity calculated under current surface normal and that measured from the image $(I(x) - R(p(x),q(x)))^2$. An additional smoothness term and integrability term are added to govern the choice of surface normal with appropriate coefficient. An iterative scheme is utilized to update the results. In general, the regularizing terms used to guide the iteration are smoothness, integrability, gradient consistency and the unit normal constraint. Several problems arise when the optimization method is applied.

Strat [56] arrived at his method by application of the standard calculus to discrete domain. Integrability constraint is the regularisation term used to narrow down the search space. Since

$$\delta u = p\delta x + q\delta y \tag{2.10}$$

Integration along a closed path C, the integral should be zero. The basic form of integrability is

$$\oint_C (p(x,y)dx + q(x,y)dy) = 0 \tag{2.11}$$

Take the partial derivative of the objective function respect to x and y, the iterative process is updated with new p and q. However, there his scheme cannot incorporate the occluding boundary.

Ikeuchi and Horn [26] were the first to apply the calculus of variations to the shape from shading problem. To use the occluding boundary information, they mapped the surface function from pg space to stereographic fg space, where

$$f = \frac{2p}{1 + \sqrt{1+p^2+q^2}} \quad \text{and} \quad g = \frac{2q}{1 + \sqrt{1+p^2+q^2}} \tag{2.12}$$

They are using the smoothness constraint instead which has the form

$$\iint_\Omega (f_x^2 + f_y^2 + g_x^2 + g_y^2)dxdy \tag{2.13}$$

18

Smith [53] uses a similar method, the only difference being that they use n_x and n_y instead of p and q.

The problem is the data closeness, which is always traded off for the smoothness of the surface and integrability. The smoothness term is assigned with equal weight and tries to adjust the surface towards a minimal surface (often a plane, depending on the exact smoothness formulation used). The integrability is another problem for variational approach, and the computational expense is usually very prohibitive. Satisfaction of the data closeness constraint was not guaranteed until Worthington and Hancock [64] proposed a geometric update procedure that allows the image irradiance equation to be satisfied as a hard constraint. In addition, they explored various types of needle-map consistency constraint. Lambertian reflectance constitutes a cone for each normal which follows the model. The normal will be rotated back to the nearest point on the cone after each iteration. However, the data-closeness means it is very sensitive to noise.

Although optimization methods can give a robust result for the shape from shading problem, it is always computationally expensive for its nature of global minimization. The challenge is how to accelerate the calculation speed. Horn devised a new algorithm that simultaneously computes both the height field z and gradient field *(p,q)* from shaded images [24]. The regularizing terms used are the brightness constraint, the gradient smoothness constraint, and the integrability constraint. A local linear expansion of the reflectance function to stabilize the algorithm and accelerate its convergence is also used.

Multi-scale approaches are often used to increase speed and robustness in many computer vision problems. The well known pyramid algorithm is widely used in computer graphics for speeding up the calculation. A Multi-resolution shape from shading is discussed in Peleg and Ron's paper [38]. It concludes that the reduced image resolution and reduced shape resolution do not commute. But it fails to provide the correct form for reducing the image that corresponds to reducing the shape resolution.

Zheng and Chellappa [69] applied the intensity gradient constraint rather than a smoothness constraint for the minimization. They simplified the Euler equations by taking the first-order Taylor series of the reflectance map and representing the depth, gradient and their derivatives in discrete form. Depth and gradients are updated simultaneously. A hierarchical structure (pyramid) is also used to speed up the computation.

A hierarchical basis conjugate gradient descent method was developed to overcome the huge computational cost of conventional iterative methods [57]. Instead of using the calculus of variations to derive the Euler equations for the solution, they discretize the function using finite dif-

19

ferences and then minimize the new energy function directly using conjugate gradient descent. Unlike the hierarchy of discretization method, this approach does not require a hierarchy of problems at different resolutions. However, the improvement in the calculation part is still minute compared with the total computation cost.

2.4 Irradiance Approximation

Besides the above two classes, a number of papers also attempt to approximate the irradiance function. In general, they can be classified into two sub-classes according to the illumination condition: Point Light Source methods and Environment illumination methods.

2.4.1 Point Light Source methods

Point light source methods work on the assumption that the light source is a point light source located at an infinite distance. Some of them solve the problem locally. The computation of the normal at each point in the image is independent. They always need strong assumptions. Pentland [41] assumes that the surface is locally spherical. There are also some linear methods to approximate the reflectance function. A global linear approximation means that the same approximation is used for all points in the image. Even though it is proved in [42] that the greater the angle between the observer's direction and the lighting direction the better the approximation, this approach is still limited.

2.4.2 Natural illumination methods

There have been some limited attempts to incorporate more realistic models of illumination into shape-from-shading algorithms. The sort of illumination encountered in natural scenes is significantly more complex than a single point light source. For shape-from-shading to be applicable in the real world, this seems like a natural direction for future research.

In recent work, Johnson and Adelson [27] demonstrate that natural illumination actually narrows down the search space, in essence lying somewhere between well-posed photometric stereo and traditional point source shape from shading. See Figure 2.3.

(a)Point Light Source   (b)RGB photometric stereo   (c)Natural Illumination
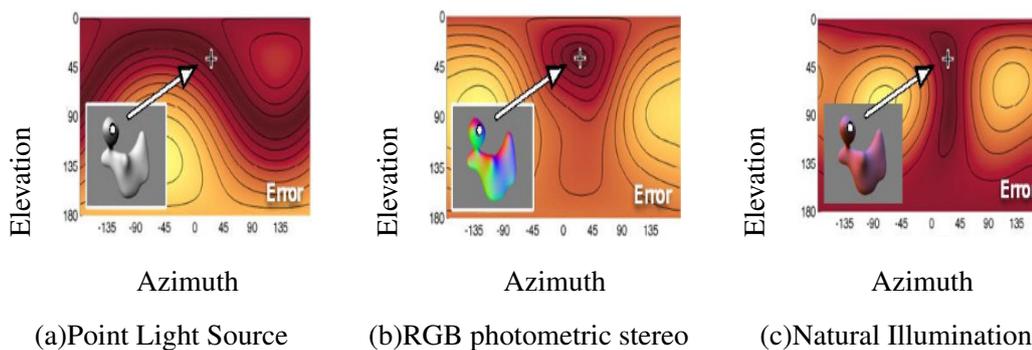
Figure 2.3: Illumination complexity reduces shape-from-shading ambiguity. Three errors maps are shown for a single observation under: (a) point source illumination, (b) three source photometric stereo (3 images mapped to red, green and blue colour channels), (c) natural illumination. Dark red represents low error. It is clear that 3 source photometric stereo has a well-defined, single minima corresponding to the true normal direction. Single source shape-from-shading is ambiguous with a distributed minima. Natural illumination lies somewhere in between with a more highly constrained minima but without a single well defined solution. (From [27]).

Previously, some work has also considered more complex illumination environments than single point sources. Koenderink et al. [30] allow an additional ambient component on top of a point source. Yuille et al. [65] considered a similar scenario in which an object is illuminated by a single point source and a diffuse component that remains constant over all the images. In shape-from-shading, Langer and Zucker [32] have considered the case of completely ambient illumination provided by skylight. In this case, shading is a function of ambient occlusion and is hence related to the global geometry of the surface. Prados et al. [47] have recently reformulated this problem as the solution of an integro-partial differential equation. Vogiatzis et al. [60] solve a nonlinear optimisation that includes estimation of the lighting environment within a multiview photometric stereo framework. This problem is geometrically constrained by the use of frontier points.

Spherical harmonics have been used previously in photometric analysis. Basri et al. [5] use the projection of the Lambertian reflectance function onto a spherical basis to derive harmonic images. These can be used to approximate an image of an object under arbitrary illumination. This allows shape recovery up to a bas-relief ambiguity using an SVD of the observed image intensities. Recently, Kemelmacher and Basri [28] used spherical harmonic images in the context of facial images. A single reference face shape is used to estimate the lighting coefficients and albedo information first. Then, the depth of the facial information is derived. The use of the spherical harmonic lighting in this paper is different to that of previous work. We use separate

21

spherical harmonic projections for reflectance and illumination in a manner that is popular in the graphics literature [21]. This renders shape recovery as a linear problem.

2.5 Other methods

While neither minimisation methods nor PDE methods offer reliable shape recovery from real world imagery, they can additionally be criticised on a number of counts. Minimization methods are critically dependent on the choice of additional constraints and are slow and computationally demanding due to their iterative nature. PDE methods require initial boundary conditions to help initialise the surface. For this reason, a number of other approaches have been explored, in particular, model-based methods, interactive approaches and those utilizing novel constraints.

Model based methods are widely used in computer graphics in applications such as deformable objects and grid animations. The most difficult part for the model-based method is the mapping strategy that involves warping and matching the input data to the model. A model based method proposed by Smith and Hancock [54, 55] focusses on the recovery of facial surfaces. A statistical model of variation in surface normal direction over the face surface is trained and used to act as a constraint on the surface. To overcome problems with applying linear statistical methods to non-linear unit surface normal vectors, the surface normals are projected to the tangent plane to the local average normal direction using the azimuthal equidistant projection. The average face is used as an initialisation. With the statistical model to hand, the algorithm proceeds by alternately finding the best fit of the model to the data and locally imposing the brightness constraint by rotating the normals back onto their irradiance cones. The problem for model based methods is that they are limited in their application. The database on which the model is trained determines the class of objects to which the algorithm can be applied. In addition, collecting the data and building the model is both time consuming and complex.

Interactive methods aim to overcome the ambiguities inherent in shape from shading by taking advantage of human expertise. User-guided shape from shading was proposed by Meyer and Pulido [36] to overcome the basic shape from shading ambiguities: bas-relief ambiguity and concave convex problem. Since these two problems are physical rather than analytical, an interactive scheme that requires the human to decide the structure is suggested. The method first determines the surface's flat areas in the image. Then, it assigns every flat area a weighted height. The ill-posed problem is circumvented by having the user set the relative heights of some areas. The fast marching method is applied to accelerate finding the heights of the rest of the surface. A similar

approach was also proposed by Zeng and Matsushita [66]. The difference here is that they use a shortest path searching method to find the relative height. Apart from the loss of automation, a potential weakness of interactive methods is that the recovered surface may be dependent on the perceptions of the user, especially for objects that are novel for the user. In other words, the process becomes subjective.

Drawing on the lack of robust constraints, some work has focussed on developing new constraints for shape from shading. A re-illumination-driven shape from shading, which is taken to be the generation of plausible, photo-realistic images under different lighting conditions from the original image, is suggested by Worthington [62, 63]. This scheme applies the smoothness constraint in a re-illuminated image computed by the estimated shape and new illumination situation, in combination with the hard data closeness constraint at each iteration. It suffers from problems inherent in the hard constraint, namely inability to deal with albedo changes and shadows that appear in real images.

2.6 Summary

In this chapter, we have summarised some of the most important historical and current work in shape from shading. We classify the shape from shading methods into three categories: partial differential equation method, minimization method and irradiance approximation methods. The PDE methods perform best in terms of efficiency, but perform badly when noise is introduced. The need for boundary information during the solving process is another limit. Minimization methods are flexible, which allow new constraints to be easily added. As we have discussed in this chapter, the result from minimization method are not guaranteed to be globally optimal. Their iterative nature is also time consuming, usually taking many hours to converge. Notice that the new constraints that are added into the scheme need a coefficient to adjust the weight, the selection of the value of the coefficient is also a notorious problem. Approximating the irradiance function by other representations is a trend within shape from shading. Better approximation of the irradiance model under environment illumination extends the potential application of shape from shading. Although the domain of the problem has been changed from using a monocular image to colour image, the result is convincing and some ambiguities have been avoided. There are also a number of methods that provide reasonable results for shape from shading, especially those using a statistical model. The only limit is the area of application.

In the next chapter, we propose a minimization method. We introduce a smoothness constraint

that is better able to preserve surface structure and a new framework for its application.

Chapter 3

A Structure-Preserving Smoothness Constraint

3.1   Introduction

The aim of computational shape-from-shading is to make estimates of surface shape from the intensity measurements in a single image. Since the amount of light reflected by a point on a surface is related to the surface orientation at that point, in general the shape is estimated in the form of a field of surface normals (a needle-map). The point light source illumination is considered in this chapter. Assuming a normalised and linear camera response, the image intensity predicted by the simplest Lambertian reflectance model is given by

$$I(x, y) = \rho(x, y)\mathbf{n}(x, y) \cdot \vec{\omega}, \tag{3.1}$$

where $\mathbf{n}(x, y)$ is the local surface normal, $\vec{\omega}$ is a vector in the light source direction and $\rho(x, y)$ is the diffuse albedo which describes the intrinsic reflectivity of the surface.

For an image in which the viewer and light source directions are fixed, the radiance function reduces to a function of the surface normal. The problem is an ill-posed problem because the surface normal has two independent variables. Some additional constraint are needed as the regularizing term. Two popular regularizing terms are smoothness constraint and integrability constraint. Solving the shape from shading problem is converted into a minimization problem which tries to minimize

$$E(u(x, y)) = \iint_{\Omega} F(x, y, u(x, y)), ...) + \lambda S(x, y, u(x, y)), ...) + \gamma Int(x, y, u(x, y)), ...) + ... dxdy \tag{3.2}$$

where $F(x, y, u(x, y)), ...)$ is defined as the non-negative error function of the radiance function, $S(x, y, u(x, y)), ...)$ and $Int(x, y, u(x, y)), ...)$ are the smoothness and integrability terms with the coefficients λ and γ.

The minimization of the above functional was done through variational calculus. However, the major problem for the above method is the smoothness error will not be zero even we provide the ground truth normal as initial input. An approach which overcomes these deficiencies was proposed by Worthington and Hancock [64]. Their idea was to choose a solution which strictly satisfies the brightness constraint at every pixel but uses the regularisation constraint to help choose a solution from within this reduced solution space. If we make the assumption that the reflectance properties are homogenous across the surface (i.e. constant unit albedo), we obtain a simple relationship between observed intensity and the angle of incidence, $\theta = \arccos(\mathbf{n} \cdot \vec{\omega})$, ($\vec{\omega}$ is the direction of light source)between the light source and surface normal:

$$I(x, y) = \mathbf{n}(x, y) \cdot \vec{\omega} = \cos \theta. \tag{3.3}$$

Geometrically, this means that the surface normal must lie on a right circular cone whose axis is the light source direction and whose opening angle is $\theta = \arccos(I)$. By constraining the surface normal to lie on the cone, we satisfy the image irradiance equation and hence ensure the fullest possible use of the input image.

To solve this minimisation, Worthington and Hancock use a two step iterative procedure which decouples application of the regularization constraint and projection onto the closest solution with zero brightness error:

1. $\mathbf{n}'_k = f_{\text{Reg}}(\mathbf{n}_k)$

2. $\mathbf{n}_{k+1} = \underset{\mathcal{E}_{\text{Bright}}(\mathbf{n})=0}{\arg \min} \; dis(\mathbf{n}, \mathbf{n}'_k),$

where $dis(., .)$ is the arc distance between two unit vectors and $f_{\text{Reg}}(\mathbf{n}_k)$ enforces a robust regularizing constraint. The second step of this process is implemented using

$$\mathbf{n}_{k+1} = \Theta(\mathbf{a}, \alpha)\mathbf{n}'_k, \tag{3.4}$$

where Θ is a rotation matrix which rotates a unit vector about an axis $\mathbf{a}$ by an angle α. To restore a normal to the cone we set $\mathbf{a} = \mathbf{n}'_k \times \omega$ and $\alpha = \theta - \arccos(\mathbf{n}'_k \cdot \omega)$. The result is the closest direction that satisfies $\theta = \arccos(I)$.

The performance of this approach is critically determined by the choice of regularisation term. Worthington and Hancock [64] proposed a number of different "robust regularisers". The idea

26

was that the update to a surface normal should be robust to extreme values in the current estimate of the local field of normals. This has the effect of preserving edges which are present at the previous iteration. However, the method relies on a correct piecewise segmentation of the surface at initialisation. Any surface structures which are not reflected in the normal map at initialisation will not be recovered as the method iterates. The result is that the method is sensitive to noise and initialisation and the final result may not capture the important structures of the surface.

3.2   Structure Perserving Smoothness Constraint

Individual pixel intensities provide a partial constraint on the local surface normal direction (namely the opening angle of the cone). However, the change in intensity across an image conveys information about the structure of the surface. Our aim is to exploit this information in our regularisation term.

The accuracy of the variational approach is entirely dependent on the choice of regularisation constraint. Typical smoothness constraints depend on the second derivative of surface height which is described as

$$\int \int (p_x^2 + p_y^2 + q_x^2 + q_y^2)dxdy, \tag{3.5}$$

where p and q are the first order partial derivatives of the surface height. The smoothness constraint is an isotropic constraint which assumes the surface changes are identical in every direction. Although the intensity gradient constraint was introduced to solve this problem [69], which requires that the intensity gradient of the reconstructed image be close to the intensity gradient of the input image in both the x and y directions, there is no existing framework for fully exploiting structural information from the local neighbourhood in every direction.

We propose an alternative regularisation constraint which employs information about the intensity gradient in all directions over a local neighbourhood. For a pixel (x, y), we define the local neighbourhood as $\Omega(x, y) = \{(x+1, y), (x-1, y), (x, y+1), (x, y-1)\}$. We precompute the change in incident angle between all pairs of neighbouring pixels:

$$S((x_1, y_1), (x_2, y_2)) = \frac{\arccos(I(x_1, y_1)) - \arccos(I(x_2, y_2))}{\Delta S_{max}}, \tag{3.6}$$

where ΔS_{max} is the largest change in incident angle over the image. We define a weight between adjacent pixels based on the magnitude of the change in incident angle: $W((x_1, y_1), (x_2, y_2)) = e^{KS((x_1, y_1), (x_2, y_2))}$, where the constant K determines the behaviour of the constraint (we use $K = 10$ as initial input). For small values, the constraint reduces to local smoothness, for large values more structure is preserved at the cost of increased sensitivity to noise.

The total of the weights between a pixel and its neighbour is given by:

$$Z(x,y) = \sum_{(i,j)\in\Omega(x,y)} W((x,y),(i,j)). \tag{3.7}$$

The surface normal at pixel (x,y) at iteration $t+1$ is given by the weighted average of its neighbouring normals at iteration t:

$$\mathbf{n}^{(t+1)}(x,y) = \frac{\mu^{(t+1)}(x,y)}{\|\mu^{(t+1)}(x,y)\|}, \tag{3.8}$$

where

$$\mu^{(t+1)}(x,y) = \sum_{(i,j)\in\Omega(x,y)} \mathbf{n}^{(t)}(i,j)\frac{W((x,y),(i,j))}{Z(x,y)}. \tag{3.9}$$

If we choose the function $S((x_1,y_1),(x_2,y_2))$ to be a constant equal to zero, the iterative process will simplify to the conventional smoothness constraint. If the change of the intensity in x and y directions alone are taken into consideration, the structural constraint here is equivalent to the intensity gradient constraint.

The regularisation constraint described above could be incorporated into the standard variational approach. However, this method would still suffer from poor data-closeness due to the treatment of the image irradiance equation as a soft constraint. Instead, we note that the application of this constraint retains surface structure and so iteratively applying the constraint until convergence does not result in an oversmoothed surface.

3.3   New Framework for the constraint

The variational method results in oversmooth surface estimates since the minimisation procedure does not strictly minimise the brightness constraint. In addition, the choice of coefficients for each regularisation term dramatically effect the result and in fact must be tuned for each input image. Worthington and Hancock [64] treat the brightness error as a hard constraint and rotate the normal such that Lambert's law is strictly satisfied at each iteration. The problem here is that the regularisation constraint is never satisfied since only a single application of the regularisation update is applied before the normal is rotated back to the cone. The effect is that surface normals oscillate between two positions: an on-cone position and a position representing a step towards reducing the error due to the regularisation constraint (See Figure3.1). The result does not improve significantly beyond the first iteration and decisions regarding the gross structure of the surface made at the stage of initialisation are simply reinforced.
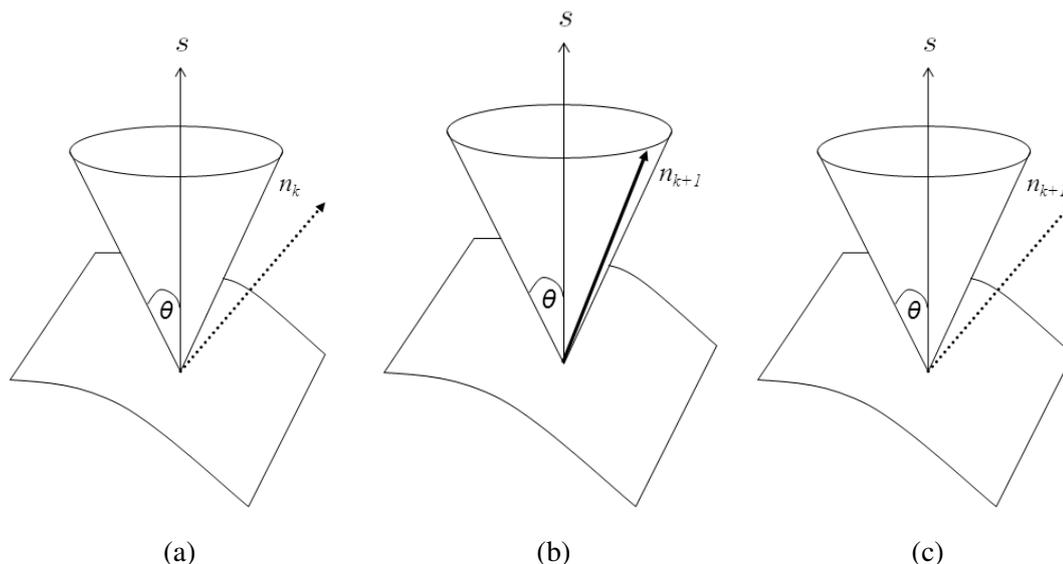
Figure 3.1: The oscillation phenomena in method [64]. (a) The smoothed normal after kth iteration. (b) Enforce the integrability constraint at the $k + 1$th iteration, the surface normal will go back to the cone which has the same open angle. (c)Apply the smoothness constraint, the normal go back to the same position which located at the kth iteration.

Our idea is to allow the regularisation update to be run to convergence before the surface normal is restored to the cone. This is only possible because our regularisation constraint seeks to preserve surface structure. Hence, even when run to convergence, the result still captures the gross structure of the surface. Our framework iteratively interleaves this process with rotating the normals back to their closest on-cone position. Since normals are allowed to move many times before being restored to the cone, changes to the gross structure of the surface are possible and information about surface structure is diffused over a wider region.

A summary of our algorithm is as follows:

1. Obtain $\mathbf{n}^{(0)}(x, y)$ using negative gradient initialisation [64]

2. Repeatedly apply (3.8) until convergence

3. Rotate normals back to cone: $\mathbf{n}(x, y) = \Theta \mathbf{n}^{(\text{final})}(x, y)$

4. Stop if converged, otherwise iterate to step 2

To obtain surface height estimates, we integrate the field of surface normals using the algorithm of Frankot and Chellappa [19].

Figure 4.20 shows the converged shapes described in step 2. From left to right, top to down are the converged shape in subsequent statues before rotating back to the cone. The shape preserves
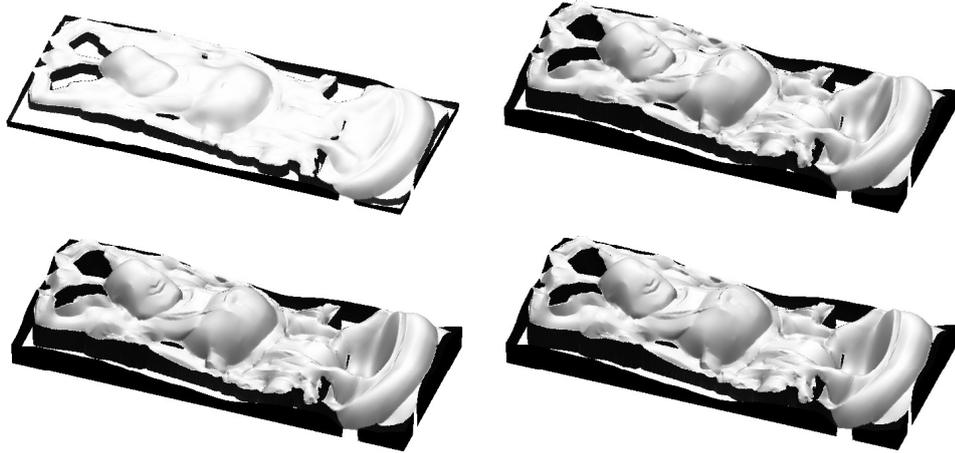
Figure 3.2: Converged shape in each step 2.

the structure features, such as the the belly and the head, and notice that from the second recovered shape, the result is still reasonable before back to cone. The advantage of this framework will be seen in the experiment. However, this framework only works for our constraint. The surface will downgrade to a plane surface if other smoothness constraints are used.

3.4 Experimental Results

We apply our method to both synthetic and real world images. We compare our result with two previous methods: Worthington and Hancock [64], and Haines and Wilson [22]. We also compare to the methods in the survey of Zhang et al. [68]. The unit albedo and frontal illumination is assumed for synthetic image.

Figure 3.3 and Figure 3.8 show the recovered surfaces on Stanford bunny and Buddha using the proposed algorithm, Worthington and Hancock [64] and Haines and Wilson [22]. The corresponding view of the ground truth surface is shown in the first panel. The input image is shown in the top left panel of Figure 3.4 The inaccuracy at the ear is because the ground truth data is pointed backwards and there is a discontinuity between the head and the ear. Note that the surface recovered by the proposed algorithm has better global structure whilst still containing much of the finescale surface detail. The advantage of our framework is also seen in this Figure Notice that the holes in Worthington and Hancock's method [64] are the results of ruining the smooth constraint by rotating back to the cone in each iteration.

Figure 3.4 and Figure 3.9 show the frontal view of meshes of two objects. The corresponding view of the ground truth mesh is shown in the first panel. The meshes from left to right are
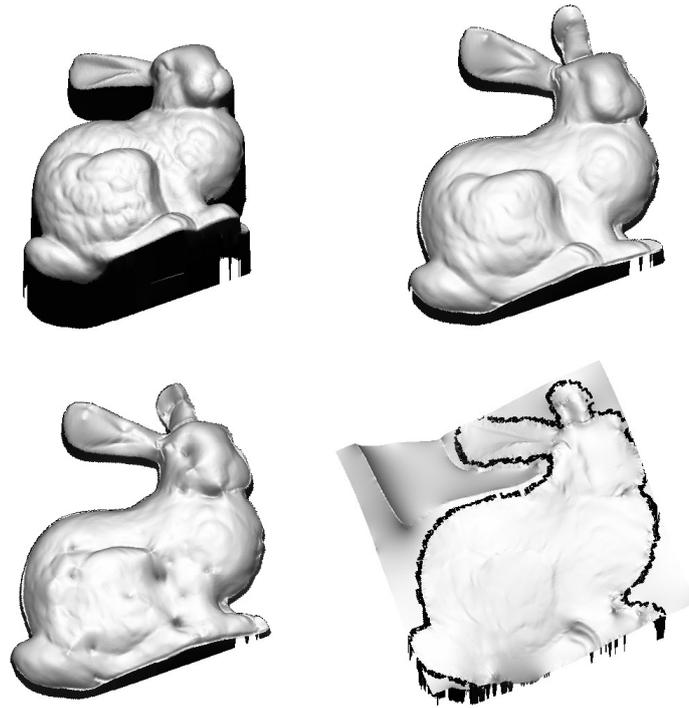
Figure 3.3: Surfaces recovered from the input image shown in the left panel of Figure 3.4 From left to right: ground truth, proposed algorithm, [64], [22].

proposed algorithm, Worthington and Hancock [64] and Haines and Wilson [22]. The surface recovered by the proposed algorithm is quite closed to the ground truth surface. The fine structure near the eyes and claws of the bunny and the belly hole of the Buddha are also preserved.
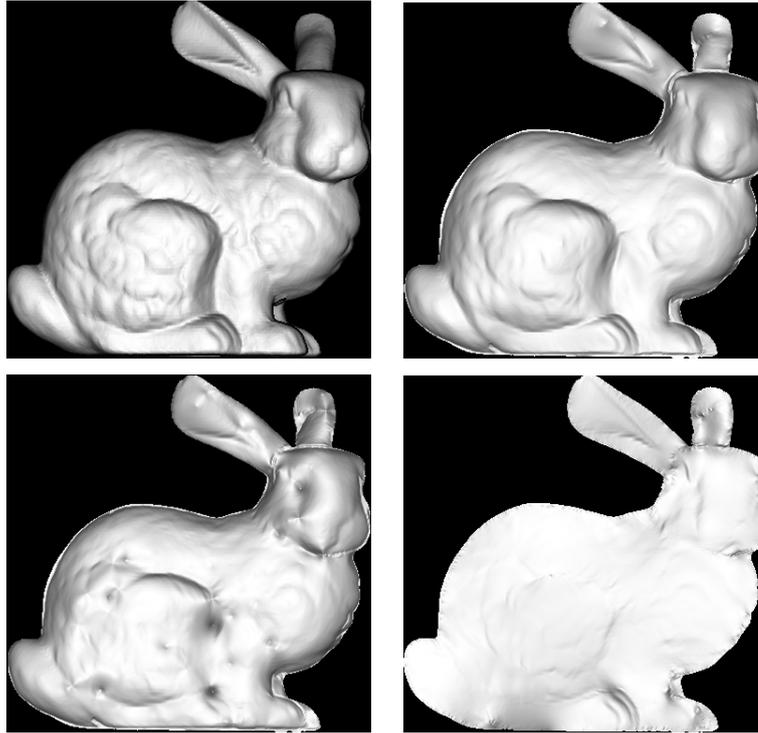
Figure 3.4: Frontal View of Mesh. From left to right: ground truth, proposed algorithm, [64], [22].

Figure 3.5 and Figure 3.10 compare the average angular error of normals in each iteration for both proposed method and Worthington and Hancock [64]. The proposed method performs better both at the error and converge speed. Note that from the use of data closeness constraint in the second time, the error increases rather than decreases. The reason for this is the concave/convex problem. The structure constraint smooths the surface to reduce the error while data closeness constraint increases the error.
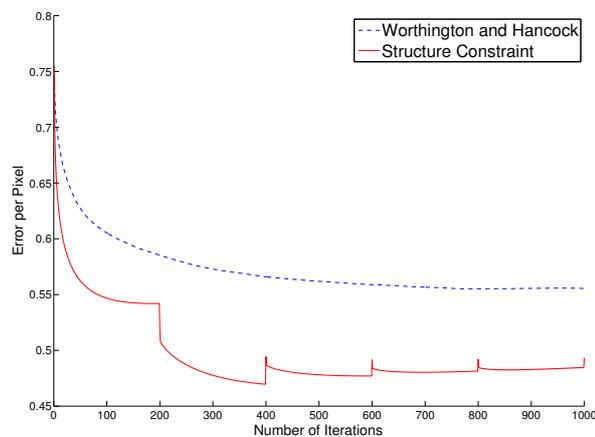
Figure 3.5: Mean Angular Error in each iteration.

Figure 3.6 and Figure 3.11 show the location of angular errors in each recovered needlemap. The angular error has been mapped into the scale of 0 to 1. The dark blue represents the error is zero and the red represents the biggest error. Note that the Worthington and Hancock's method [64] performs well at the bunny's ear at the concave and convex decision. The proposed method has a better performance in whole for the global images are dark blue. The overall body and head of the bunny is blue and the belly of the Buddha is also better than those of other methods. The results illustrate that under the correct convex and concave situation, the proposed method does well in the detail structure.
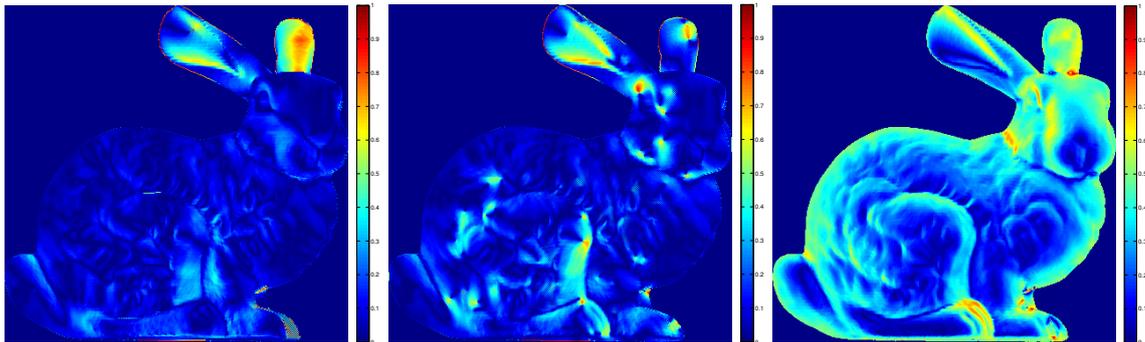
Figure 3.6: Error Map, From left to right: proposed algorithm, [64], [22].

Figure 3.7 and Figure 3.12 compare the histogram of the angular error. The propose method is much more accurate than those of other methods.
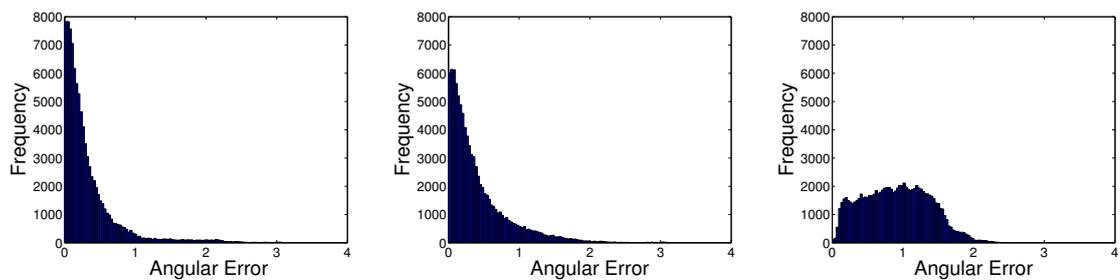
Figure 3.7: Histogram of Errors, From left to right: proposed algorithm, [64], [22].
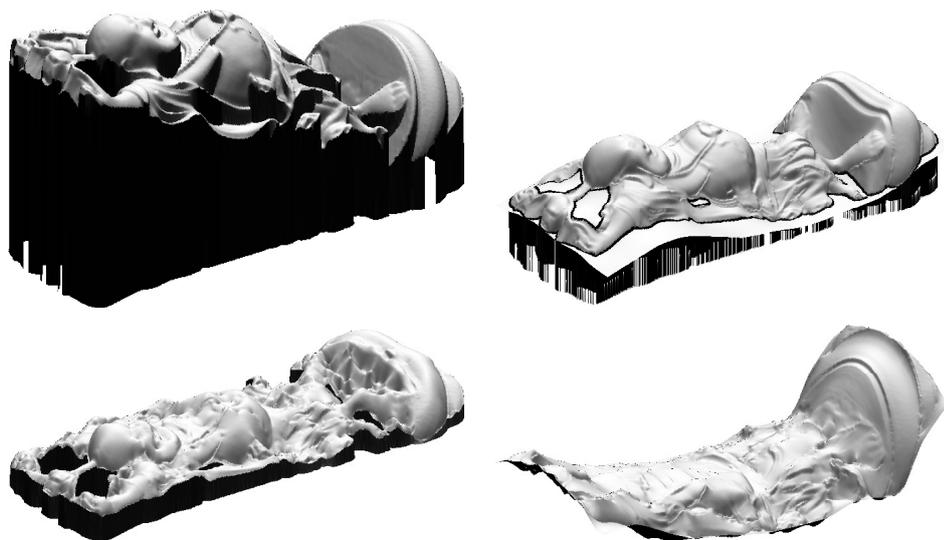
Figure 3.8: Surfaces recovered from the input image shown in the left panel of Figure 2. From left to right: ground truth, proposed algorithm, [64], [22].
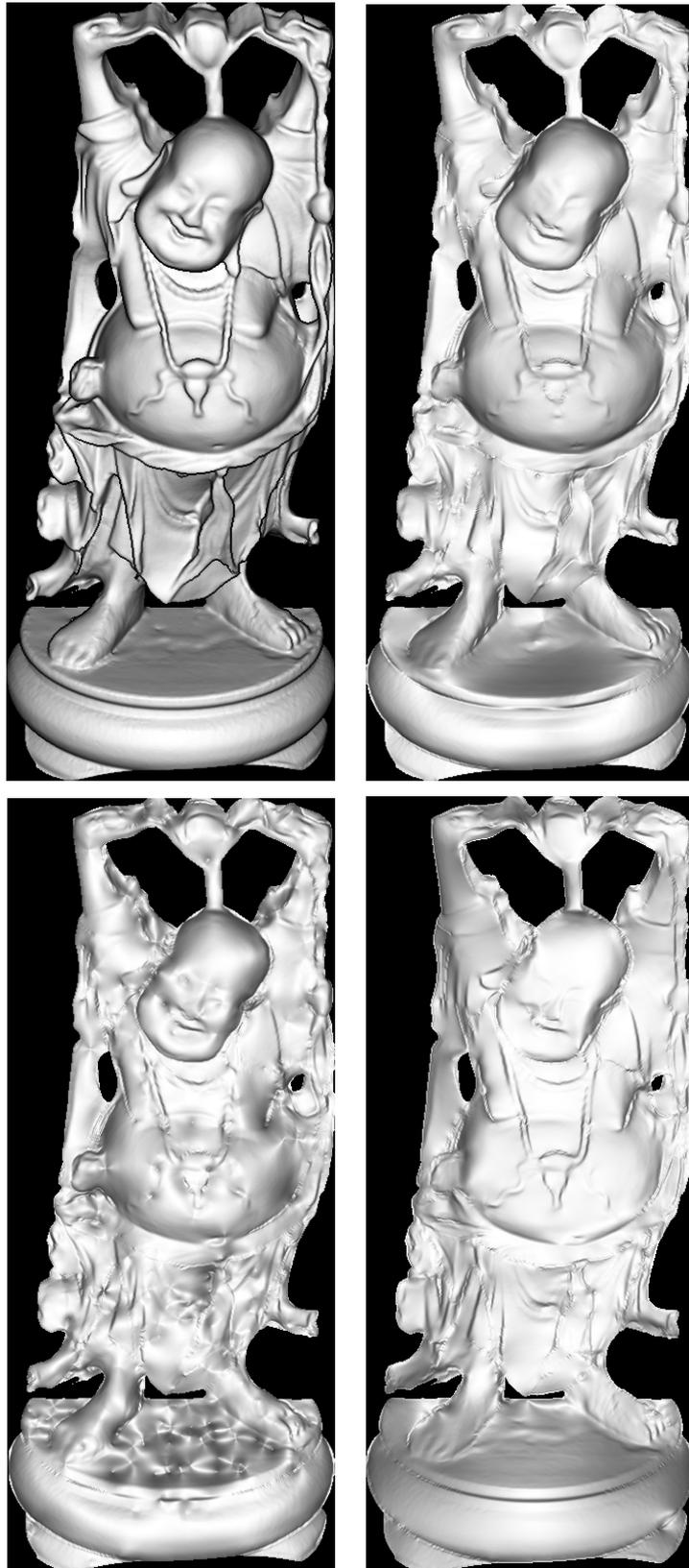
Figure 3.9: Frontal View of Mesh. From left to right: ground truth, proposed algorithm, [64], [22].
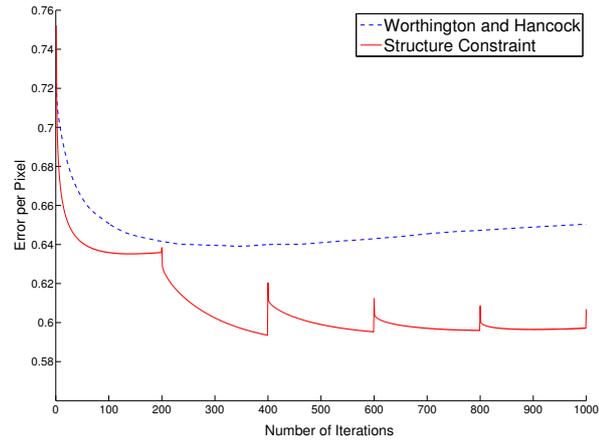
Figure 3.10: Mean angular error in each iteration.
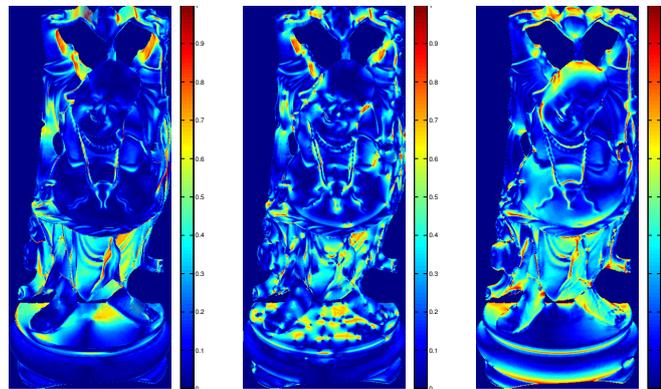
Figure 3.11: Error Map, From left to right: proposed algorithm, [64], [22].
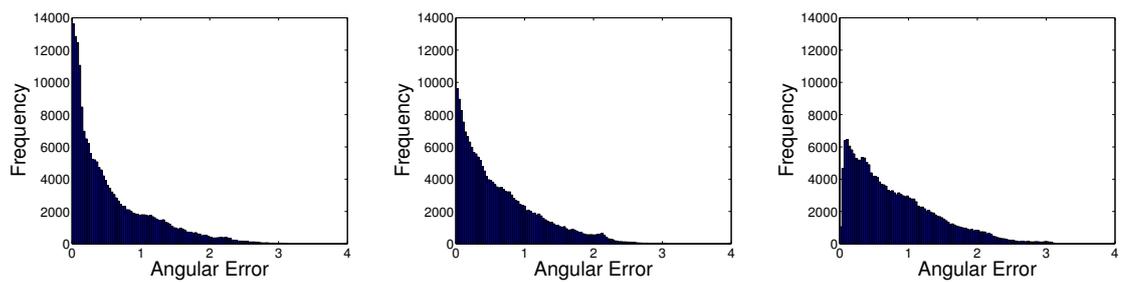
Figure 3.12: Histogram of Errors, From left to right: proposed algorithm, [64], [22].

Figure 3.13 and Figure3.14 show the recovered surfaces and surface normals along with reil-luminations of the surface normals under novel lighting. In column 1 we show the recovered surfaces rendered with frontal illumination. In column 2 we show the surface normals. The re-maining columns show renderings under novel illumination. The top row shows ground truth

images, the remaining rows show results from the proposed algorithm, Worthington and Hancock and Haines and Wilson respectively. We assume that the light source vector, L, is known. This is used in the initialisation and rotation back to the cone. Note that the quality of the estimated surface shape and reilluminations are considerably improved using the proposed algorithm.

Surface	Needle-map	Above	Below	Left	Right
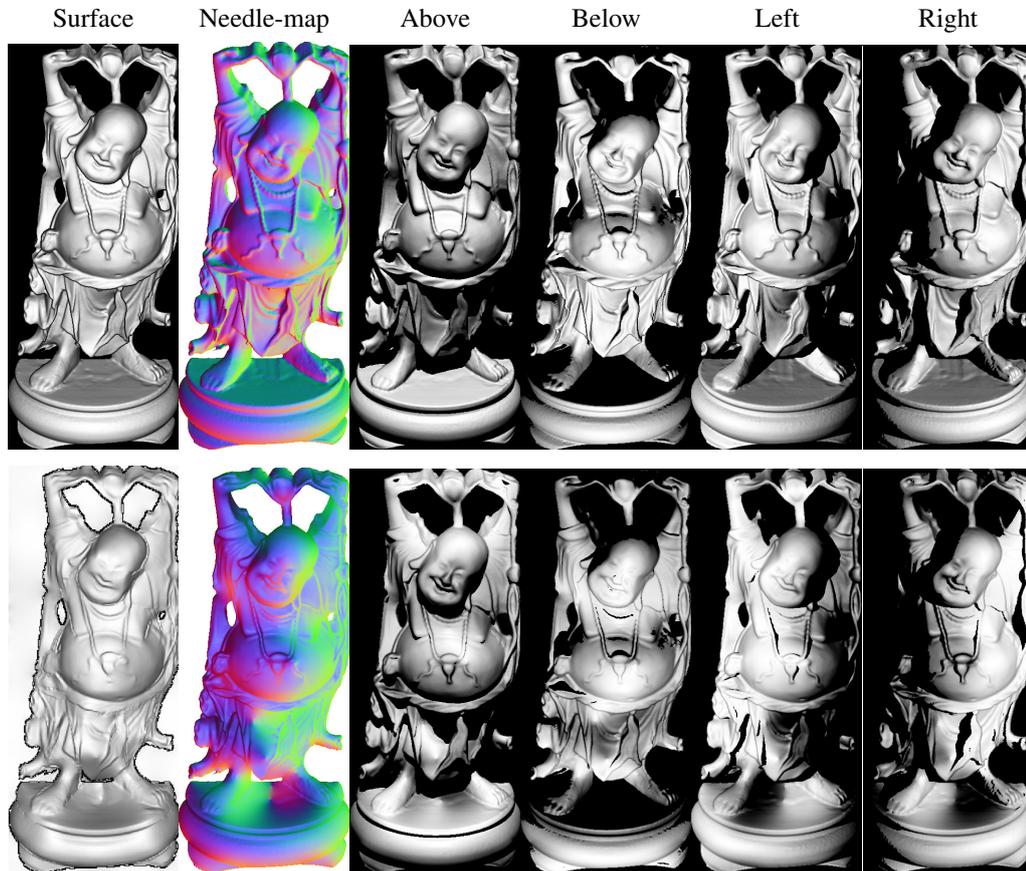
Figure 3.13: Recovered surfaces, surface normals and reilluminations. Ground truth in first column, remainder show: proposed algorithm.

| Surface | Needle-map | Above | Below | Left | Right |
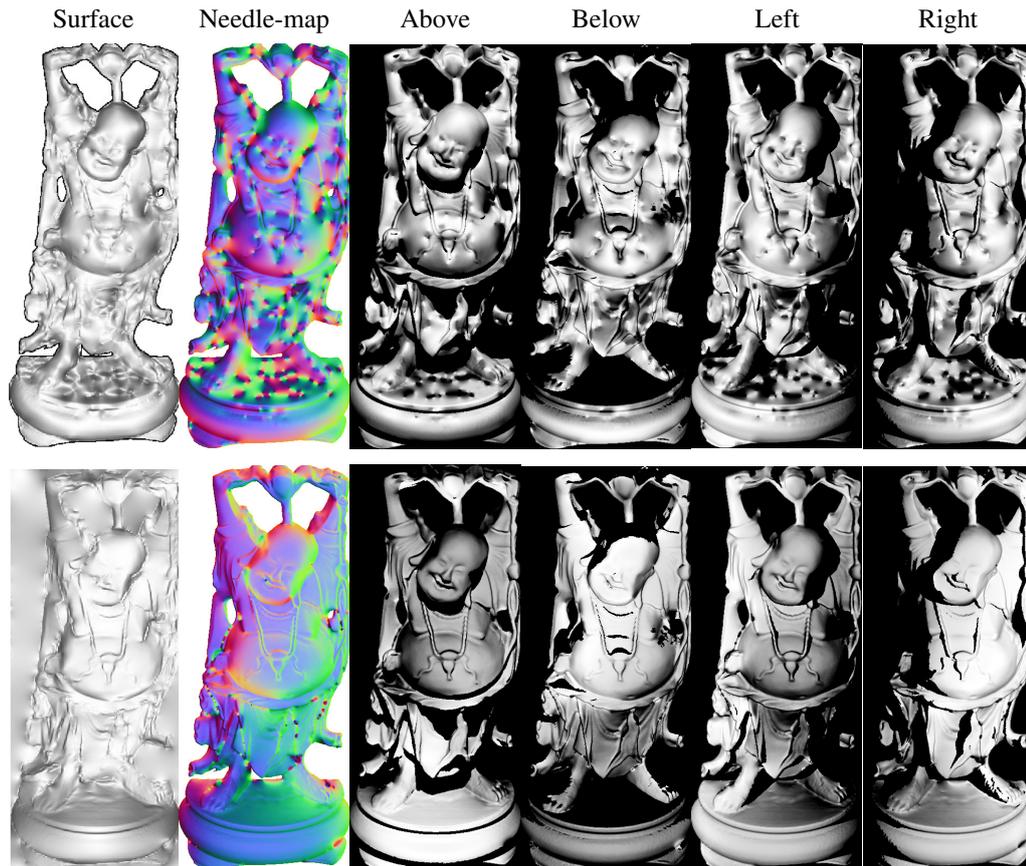
Figure 3.14: Recovered surfaces, surface normals and reilluminations. Ground truth in first column, remainder show: [64] and [22] respectively.
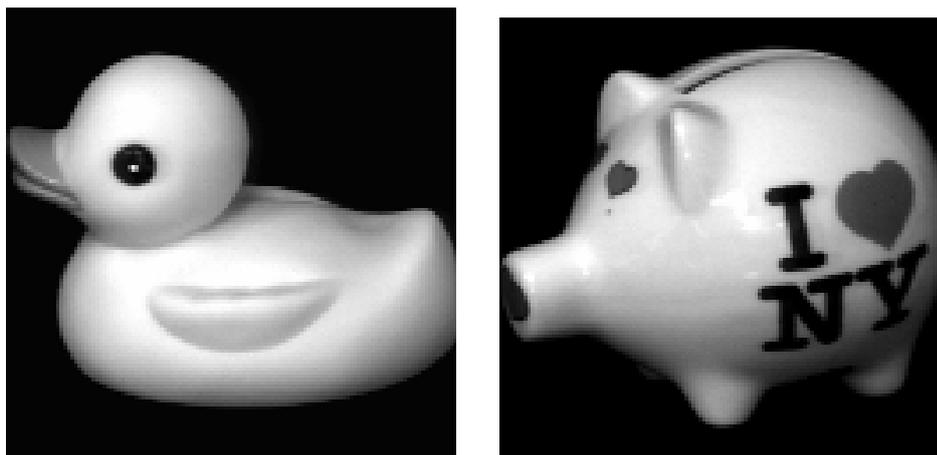
Figure 3.15: Real image in Coil database.
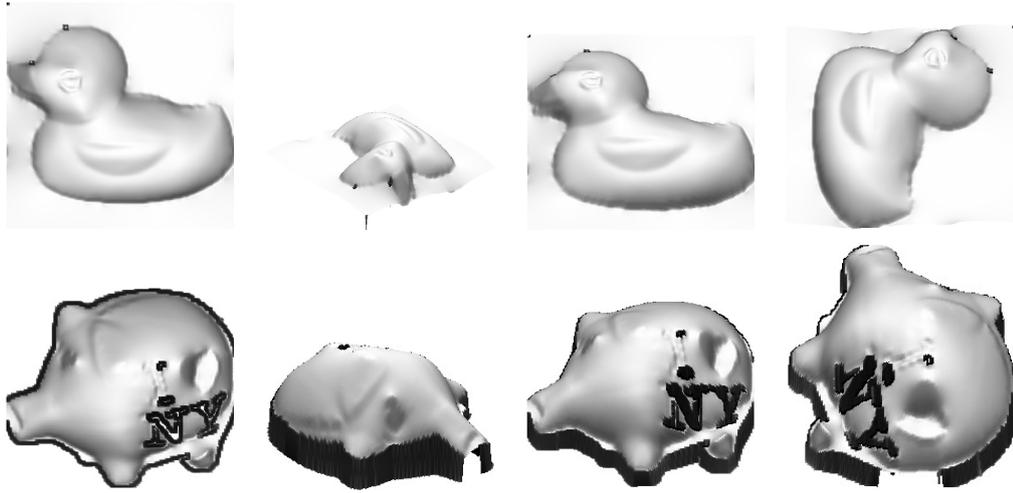
Figure 3.16: Surfaces recovered from real images.

In Figure 3.16 we show the result of applying our method to real images from Coil database. The input is shown in Figure3.15, each columns in Figure 3.16 show the surface from different viewpoints. Note that the surfaces in these images deviate from the Lambertian assumption and contain variations in albedo. Despite this, our algorithm recovers stable surface estimates which retain the fine surface detail (e.g. the wing of the duck).

We also apply our algorithm to a challenging real image captured of a snowy scene. The sky has been manually removed. The light source direction is considered to be from above. The proposed method also gives a plausible result under the non-frontal illumination case. Note that the parts with the cast shadow, the proposed method shows the potential in handling small change in albedo.

Figure 3.17: Snow mountain
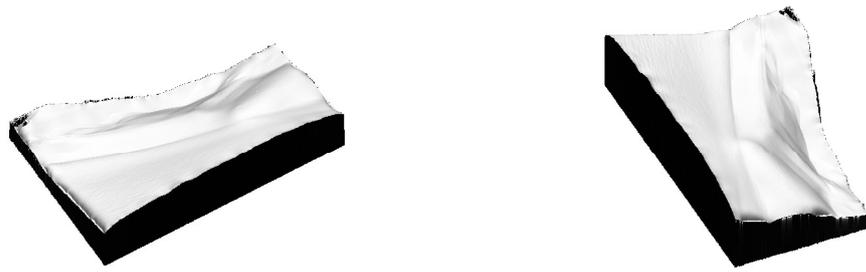
Figure 3.18: Realistic Image

We also show results with cast shadow for non-frontal illumination. The ground truth surfaces in Figure 3.19 are illuminated from an extreme angle resulting in much of the surface being in shadow. Note that our algorithm degrades gracefully, recovering unshadowed portions of the surface independently, whilst still retaining much of the global structure.
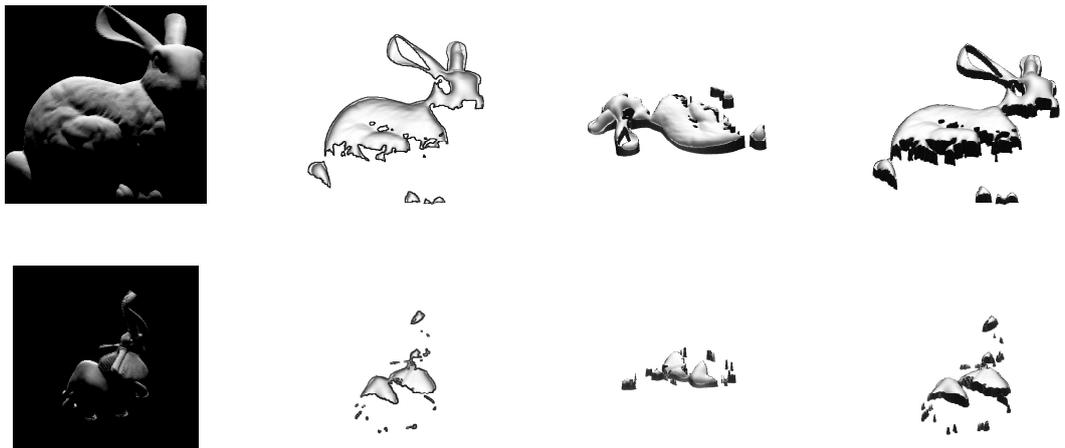
Figure 3.19: Surfaces recovered under extreme illumination.

Finally, we provide results for the classical shape from shading test objects: Mozart, Vase, and Penny. The surfaces recovered by our method show excellent reconstruction of both gross surface structure and finescale detail.

40
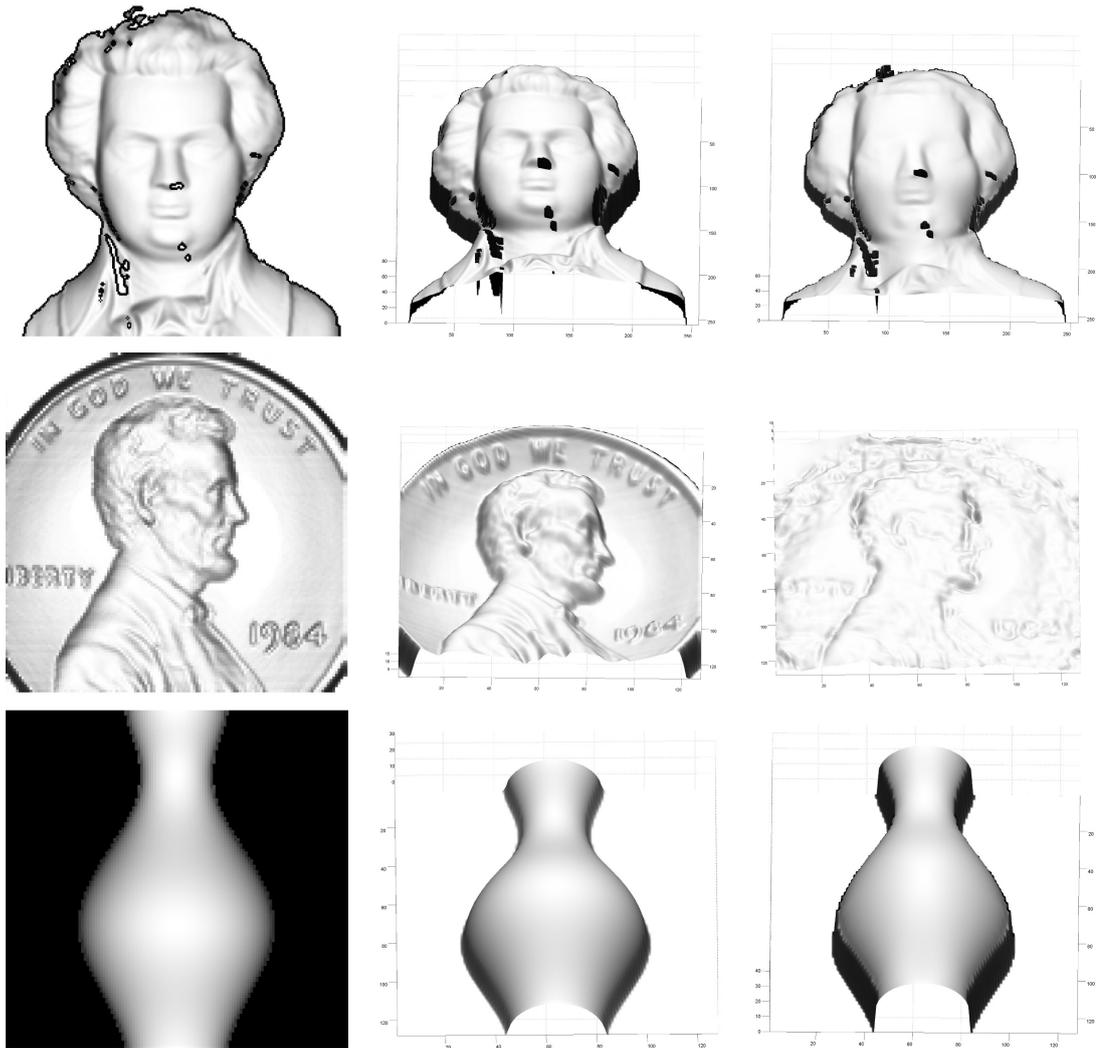
Figure 3.20: Three other often used objects for shape from shading.

The 3.1 gives the best value for coefficient K in all the synthetic image together with the related mean angular error.

Table 3.1: Synthetic objects with the best K and the Mean Angular Error

Object	K	ME
bunny	18.863	0.32644
buddha	1.0152	0.53059
vase	63.39	0.095695
Mozart	36.101	0.29148
Penny	997.3	0.59538

Notice the buddha and the penny have the largest error. The initialisation is the key to our method, and the discontinuity would also downgrade the accuracy.

We also compare the data of Mozart and Vase with those of listed in Zhang's survey paper [68]. Those methods are taken from the following papers [7, 33, 34, 42, 59, 69]

Table 3.2: Average Z error for synthetic images

Methods	Images			
	Vase		Mozart	
	S1	S2	S1	S2
Zheng & Chellappa	8.5	8.5	15.1	10.56
Lee & Kuo	10.0	7.9	13.7	9.77
Bichsel & Pentland	10.08	7.9	20.5	7.7
Lee & Rosenfeld	8.4	18.3	17.8	11.3
Pentland	11.2	9.0	15.7	19.7
Tsai & Shah	8.3	12.7	18.5	20.0
Our Method	1.99	4.99	3.84	5.45

Table 3.3: Standard deviation of Z error for synthetic images.

	Images			
	Vase		Mozart	
Methods	S1	S2	S1	S2
Zheng & Chellappa	11.1	13.9	18.4	15.9
Lee & Kuo	13.2	15.39	19,2	22.1
Bichsel & Pentland	13.8	16.9	37.4	14.6
Lee & Rosenfeld	14.6	22.3	33.0	30.3
Pentland	12.6	11.1	18.2	20.56
Tsai & Shah	15.0	19.7	33.3	30.5
Our Method	1.16	2.75	3.18	3.61

Table 3.4: Average p-q error for synthetic images.

	Images			
	Vase		Mozart	
Methods	S1	S2	S1	S2
Zheng & Chellappa	2.2	1.3	2.3	1.1
Lee & Kuo	1.6	0.9	1.7	0.6
Bichsel & Pentland	2.7	1.9	3.11	1.9
Lee & Rosenfeld	3.3	6.8	13.7	4.3
Pentland	1.8	1.2	1.3	1.3
Tsai & Shah	1.4	6.7	5.5	4.2
Our Method	0.29	0.37	0.43	0.45

Our method significantly outperforms all these previously published methods.

3.5 Summary

In this chapter, we present an iterative shape from shading method inspired by the original variational approach. We evaluate the existing smoothness constraint and develop a new structure preserving smoothness constraint. It is based on a simple heuristic that adjacent pixels with similar intensities should have similar surface orientation. The result is a balanced combination of

the local minimum of irradiance function and structure preserving smoothness function. The new framework which encourage the minimum to achieved for each part. The results outperforms a number of previous methods.

However, the main disadvantage of minimization method is its complexity. It takes many minutes to obtain a result. Also, the coefficient needs to be carefully chosen. It cannot guarantee the best result for different object even with same illumination. The result is also reliant on the initialisation. Finally, the assumption of a single point light source is unlikely to be sufficiently realistic for real world images.

Chapter 4

Shape-from-shading under Complex Illumination

4.1 Introduction

The shape-from-shading literature has almost exclusively modelled illumination as a single distant point light source. Such illumination is described by a unit vector representing the direction from which the light is incident on the surface and an intensity value representing the strength of the illumination. For a unit albedo, Lambertian surface, this leads to the very simple image irradiance equation used in the previous chapter:

$$I(x) = \mathbf{n}(x) \cdot \vec{\omega}. \tag{4.1}$$

where $\vec{\omega}$ is a unit vector in the light source direction.

In fact, if the angle between the light source and surface normal is greater than $90°$ then the surface is self-shadowed and no light is reflected. Extending the Lambertian model to account for self-shadowing requires clamping negative values to zero, i.e.

$$I(x) = \max(\mathbf{n}(x) \cdot \vec{\omega}, 0). \tag{4.2}$$

This is known as the half-cosine or clamped-cosine function. When considering environmental illumination (where light arrives from every direction), self-shadowing is very important as it excludes half of the environment from illuminating a point.

4.1.1 Background: Environmental Illumination

In the real world, a point on a surface is illuminated by light arriving from every direction in the local upper hemisphere. This light could be direct from emitting sources, which themselves

45

| (a) Environment Map | (b) Diffuse Convolution | (c) Specular Convolution |
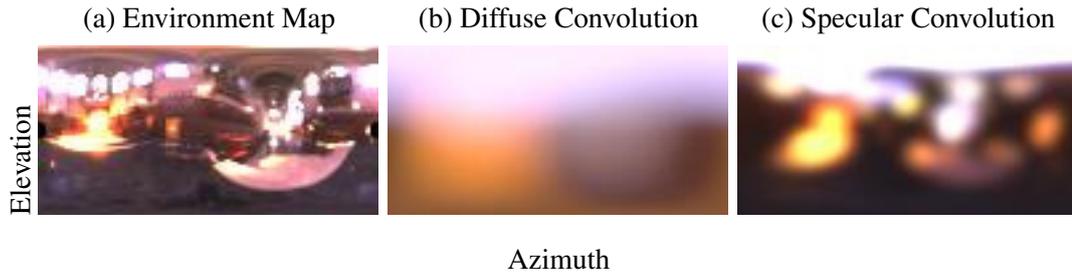
Azimuth

Figure 4.1: An environment map and its convolution with diffuse and specular reflectance functions. The horizontal coordinate represents the azimuth angle and the vertical coordinate represents the elevation angle.

could be directional (e.g. sunlight or artificial light) or ambient (e.g. skylight on a cloudy day). Alternatively, light could arrive at a surface having already been reflected by another object in the scene.

Environmental illumination refers to lighting which is: distant (i.e. directional dependence does not vary with position on the surface) and spherical (i.e. arrives from all direction). Such illumination is represented by a spherical function: $L(\vec{\omega})$, which describes the light source intensity (i.e. incident radiance) from direction $\vec{\omega}$. Point source illumination is a special case of environmental illumination in which L is a spherical Dirac delta function. In graphics, environmental illumination has been used for over a decade [61] and is now a standard rendering technique.

Image intensity under environmental illumination is calculated as an integral over the local incident hemisphere. For a Lambertian object with unit albedo this is given by:

$$I(x) = \int_{\Omega(\mathbf{n}(x))} L(\vec{\omega})V(x,\vec{\omega})(\mathbf{n}(x) \cdot \vec{\omega})d\vec{\omega}, \tag{4.3}$$

where $\Omega(\mathbf{n}(x))$ is the upper hemisphere relative to the local surface normal $\mathbf{n}(x)$. The *visibility function*, $V(x,\vec{\omega})$, is a binary function which describes occlusions by other parts of the surface. It is defined to be equal to zero if point x is occluded in direction $\vec{\omega}$, otherwise it is equal to one. For a convex object there are no occlusions and hence $\forall x, \vec{\omega} : V(x,\vec{\omega}) = 1$.

In practice, environmental illumination in a scene is measured by either placing a mirrored ball (known as a "light probe") at the point of interest in a scene or by using a fish-eye lens. The spherical environmental illumination function is represented as a discretised 2D image which is known as an environment map. Most commonly, an equirectangular projection is used to map the spherical function to 2D. Figure 4.1(a) shows an example of an environment map (in this case taken inside a cathedral - the bright sources are sunlight through windows). Row corresponds to

elevation angle (from $-90°$ to $+90°$) and column to azimuth angle (from $0°$ to $360°$). Hence, pixel rows correspond to lines of equal elevation (i.e. lines of latitude) and columns to lines of equal azimuth (i.e. lines of longitude). Note that this projection does not preserve area and the north and south poles are stretched to the whole of the top and bottom rows respectively. Image intensity calculated using an environment map transforms the continuous integral of Equation 4.3 into a discrete sum over the pixels of the environment map. In essence, each pixel of an environment map corresponds to a distant point light source whose direction is indicated by its position within the environment map image.

4.1.2 Background: Environmental Illumination Approximations

Real world illumination environments are, by their nature, complex (consider the space of all possible natural images). Their effect however is not. In particular, a diffuse surface acts like a low pass filter when it reflects environmental illumination towards the viewer. Hence, reflected intensity can be thought of as the convolution of the environment map with the surface reflectance function. Concretely, if occlusions are disregarded (i.e. the surface is assumed convex) then the image intensity reduces to purely a function of surface normal direction:

$$I(x) = \int_{\Omega(\mathbf{n}(x))} L(\vec{\omega})(\mathbf{n}(x) \cdot \vec{\omega})d\vec{\omega} = L * A(\mathbf{n}(x)), \qquad (4.4)$$

$A(\mathbf{n}(x))$ is the half-cosine function centred on the surface normal direction $\mathbf{n}(x)$. This observation applies to any reflectance function, not just Lambertian. In Figure 4.1 we show the convolution of the environment map with (b) diffuse and (c) specular reflectance functions. These convolution maps represent the intensity of the reflected light as a function of the surface normal direction under this particular environmental illumination. Hence, the mapping from surface normal direction to intensity can be described by a low frequency function.

This motivates the use of low frequency representations for object appearance. In particular, spherical harmonics have proved a particularly popular representation. Two spherical harmonic formulations have been used previously in vision and graphics. The first describes the illumination and reflectance functions separately using different sets of harmonic coefficients (we refer to this as "spherical harmonic lighting"). The second describes the convolution of the illumination environment with the reflectance function using a single set of coefficients (we refer to these as "spherical harmonic images"). The first representation is more flexible but may require many coefficients to adequately capture the high frequency variation present in a complex environment. The second representation is more efficient since convolution with a diffuse reflectance function guar-

antees that the resulting function will be accurately represented by a low frequency approximation. In this chapter, we explore the use of both representations.

4.1.3 Chapter Outline

We begin with an intuitive explanation of shape estimation under natural illumination and illustrate a simple preliminary result which was obtained in the course of this work. The convolution maps shown in Figure 4.1 associate an intensity (or colour) with each possible normal direction. If an object with the same reflectance properties as in the convolution images (or a combination of the two, e.g. the Phong model) is observed, points on the surface with a given normal direction will have the same intensity (or colour) as the corresponding normal in the convolution image. On a pixel by pixel basis, this gives us a means of computing an image radiance error (i.e. the difference between observed and predicted image radiance) which could be used in any existing optimisation-based shape from shading algorithm. Alternatively, if the surface is locally spherical, then patches around a point on the surface will exhibit the same image radiance (or colour) variation as those in the corresponding point in the convolution image. This may provide a means to estimate surface normal direction by matching patches around each pixel with patches in the convolution image. The reason such an approach is plausible is because complex illumination gives rise to complex shading patterns which may be sufficiently unique to allow patch-based matching.

This simple approach is demonstrated in Figure 4.2. A sphere and a complex object were rendered with the same glossy reflectance properties under environmental illumination. Template matching was used to find correspondences between patches on the object and the sphere. In other words, we made the assumption that the surface was locally spherical. The colour associated with each orientation is sufficiently distinct that this simple approach gives good global shape estimation, albeit with noise present in the reconstruction.

The remainder of this chapter is divided into two parts. We present methods for shape estimation using spherical harmonic representations. In the first part, we represent illumination and reflectance separately. In this case, reconstruction is only possible when lighting is low frequency. In the second part, we use spherical harmonics to represent the convolution of the illumination environment with Lambertian reflectance. In this case, we present methods for both first order and second order harmonic approximations.

Sphere Image	Input Image	Result
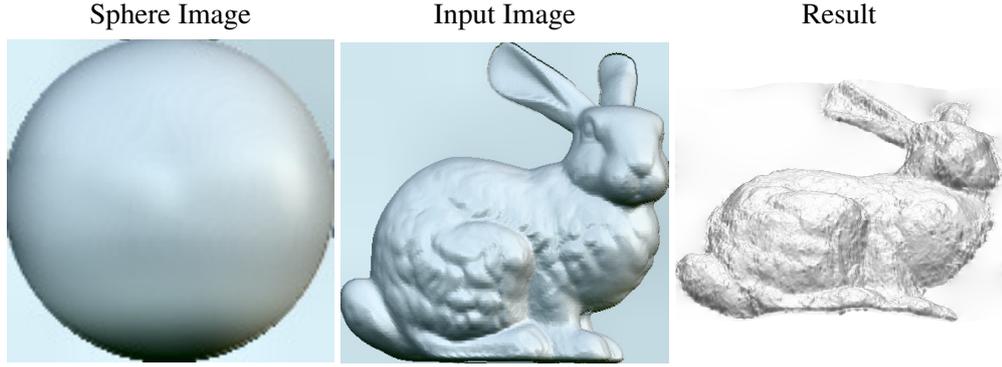
Figure 4.2: Proof of concept: a shape estimation result based on a preliminary method developed over the course of this thesis.

4.2 Shape Estimation using Spherical Harmonic Lighting

We begin by showing how to represent the illumination environment and surface reflectance functions in terms of a spherical harmonic basis. We use this representation to derive a shape-from-shading algorithm that is accurate when the illumination environment itself is low frequency and can be accurately described using a low order spherical harmonic approximation.

4.2.1 The Spherical Harmonics

A spherical harmonic of degree l and order m, $Y_l^m(\theta, \phi) : |m| \leq l \in N$ is defined on the unit sphere as:

$$Y_l^m(\theta, \phi) = k_{l,m} P_l^m(cos\theta) e^{im\phi}, \tag{4.5}$$

where $\theta \in [0, \pi], \phi \in [0, 2\pi]$, $k_{l,m}$ is a constant and P_l^m is the associated Legendre polynomial. The constant $k_{l,m}$ has the form

$$k_{l,m} = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} \tag{4.6}$$

The associated Legendre polynomials are defined recursively as

$$P_l^{-m} = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m \tag{4.7}$$

with base case:

$$P_0^0 = 1 \tag{4.8}$$

The indices l and m (which are integers) are referred to as the degree and order of the associated Legendre polynomial respectively. This equation has nonzero solutions that are nonsingular on

49

$[-1, 1]$ only if l and m are integers with $0 \leq m \leq l$, or with trivially equivalent negative values. When in addition m is even, the function is a polynomial. When m is zero and l integer, these functions are identical to the Legendre polynomials. In general, when l and m are integers, the regular solutions are sometimes called "associated Legendre polynomials", even though they are not polynomials when m is odd.

The real form of the spherical harmonics can be defined in terms of their complex analogues by setting

$$Y_l^m(\theta, \phi) = \begin{cases} \frac{1}{\sqrt{2}}(Y_l^m(\theta, \phi) + (-1)^m Y_l^{-m}(\theta, \phi)) = \sqrt{2}N_{(l,m)}P_l^m(\cos\theta)\cos m\phi & \text{if } m > 0 \\ Y_l^0(\theta, \phi) & \text{if } m = 0 \\ \frac{1}{i\sqrt{2}}(Y_l^{-m}(\theta, \phi) - (-1)^m Y_l^m(\theta, \phi)) = \sqrt{2}N_{(l,m)}P_l^{-m}(\cos\theta)\sin m\phi & \text{if } m < 0 \end{cases}$$
(4.9)

The spherical harmonics are orthogonal functions such that a spherical function can be expressed by a unique linear combination of spherical harmonic bases:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{l,m} Y_l^m(\theta, \phi),$$
(4.10)

where the coefficients $c_{l,m}$ are derived by projecting the spherical function onto the spherical harmonics basis:

$$c_{l,m} = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) Y_l^m(\theta, \phi) \sin(\theta) d\theta d\phi.$$
(4.11)

The quality of the approximation is determined by the degree of spherical harmonic used. Note that in practice, we limit the maximum degree, typically to a value of $l = 2$ or $l = 3$.

The first 4 terms of the spherical harmonic expansion of the Lambertian reflectance function with respect to Euclidian coordinates are [5]:

$$Y_0^0 = \sqrt{\frac{1}{4\pi}}$$

$$Y_1^{-1} = \sqrt{\frac{3}{4\pi}}x$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}z$$

$$Y_1^1 = \sqrt{\frac{3}{4\pi}}y$$
(4.12)

The first term is a constant and is unaffected by the surface normal direction. The real form of spherical harmonics from degree 0 to degree 2 are listed in Appendix A.

4.2.2 Spherical Harmonic Rendering

Spherical harmonics are an effective way to approximate a complex illumination environment. They have been used in computer graphics to speed up environment map rendering by replacing a numerical double integration with a matrix multiplication [21].

Rendering using spherical harmonics is achieved by approximating the environment map function and bidirectional reflectance distribution function (BRDF), respectively. The rendered image is then simply given by the inner product of the two coefficient vectors.

The illumination function $L(\vec{\omega})$ is stored in a discretised environment map in a latitude-longitude parameterisation: $L(\theta, \phi)$ (see for example Figure 4.3). Such environment maps can be acquired by photographing a mirrored ball (a 'light probe') or using a fish-eye lens. The relationship between $\vec{\omega} = [\omega_x \ \omega_y \ \omega_z]^T$ and θ, ϕ is as follows:

$$\omega_x = \sin \phi \sin \theta$$
$$\omega_y = \cos \phi \sin \theta \tag{4.13}$$
$$\omega_z = \cos \theta$$

We can approximate Equation 4.4 by decomposing the function into two spherical harmonic expansions: one for the environment map and one for the Lambertian reflectance function. The environment map is approximated by

$$L(\vec{\omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{l,m} Y_l^m(\vec{\omega}) \tag{4.14}$$

where

$$c_{l,m} = \sum_{\theta,\phi} L(\vec{\omega}) Y_l^m(\vec{\omega}) \tag{4.15}$$

and the reflectance function is approximated by

$$\max(0, \mathbf{n} \cdot \vec{\omega}^T) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} d_{l,m} Y_l^m(\vec{\omega}) \tag{4.16}$$

where

$$d_{l,m} = \sum_{\theta,\phi} \max(0, \mathbf{n} \cdot \vec{\omega}^T) Y_l^m(\vec{\omega}) \tag{4.17}$$

The summation is over the pixels of the environment map. For simplicity reason, it is also written in the following form

$$L(\vec{\omega}) = \sum_{i=0}^{\infty} c_i Y_i(\vec{\omega})$$
$$c_i = \sum_{\theta,\phi} L(\vec{\omega}) Y_i(\vec{\omega}) \tag{4.18}$$

and

$$\max(0, \mathbf{n} \cdot \vec{\omega}^T) = \sum_{i=0}^{\infty} d_i Y_i(\vec{\omega})$$
$$d_i = \sum_{\theta,\phi} \max(0, \mathbf{n} \cdot \vec{\omega}^T) Y_i(\vec{\omega}) \tag{4.19}$$

The relationship between i and l, m is $i = 2^l - l - 1 + m$. The coefficients c_i describe a specific illumination environment, $L(\vec{\omega})$. The quality of this approximation is dependent on the frequency composition of the environment map and the number of coefficients used. The coefficients d_i approximate the Lambertian reflectance function for a particular normal direction $\mathbf{n}$. An attractive property of the spherical harmonic functions is that they are not only orthogonal, but also orthonormal, which means the integral of the product of two spherical harmonic functions is one if they are the same, and zero if not. So, the reflectance function now is

$$I(x) = \rho(x) \int \sum_{i=0}^{num} c_i Y_i(\vec{\omega}) \sum_{i=0}^{num} d_i Y_i(\vec{\omega}) d\vec{\omega} \tag{4.20}$$

which is actually

$$I(x) \approx \rho(x) \sum_{i}^{n} c_i d_i \tag{4.21}$$

or equivalently:

$$I \approx \rho(x) \vec{c}^T \vec{d}. \tag{4.22}$$

The spherical harmonic approximation to the half cosine function (i.e. cosine function with negative values clamped to zero) is very accurate because of the low frequency nature of the function. This is discussed in more detail in Section 4.3. When it extends to the illumination environment, a higher order approximation is required. In graphics an approximation to order 25 is not unusual. Since all of these could be pre-calculated before rendering, it could achieve the real time purpose in rendering. Even cast shadows and interreflections can be approximated in a similar way. Figure 4.4 shows an example of using spherical harmonic lighting to render a Lambertian object under three different environment maps. These environment maps contain only low frequency illumination variations and are therefore well approximated by the spherical harmonics.
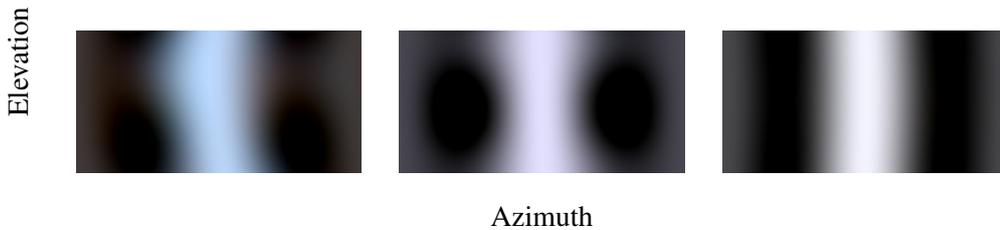
Figure 4.3: three examples of environment maps containing low frequency illumination
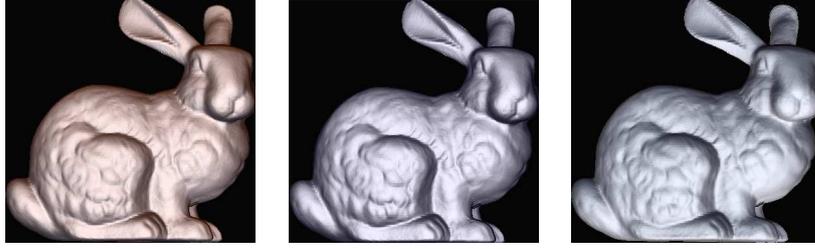
Figure 4.4: Examples of rendering using spherical harmonic lighting. An object rendered with the corresponding environment map.

Note that we work with colour images. In this case, we have a different set of coefficients for each colour channel of the environment map.

4.2.3 Shape Recovery from Multiple Images

We begin by considering the case of multiple images. Essentially, this is an extension of photometric stereo to the complex illumination case which uses a different formulation to that of Basri et al. [5]. For the case of multiple images, we derive a linear system of equations which is overdetermined. Consider the case of k colour images, giving $3k$ observations per pixel. The intensity at pixel x in the red channel for image i is denoted $I_r^i(x)$, similarly for the green and blue channels. The vector of coefficients obtained by a spherical harmonic projection of the red channel of the ith environment map is denoted $\vec{c_r^i}$. Hence, for each pixel x we have the following linear system of equations:

$$
\begin{aligned}
I_r^i(x) &= \rho(x)_r \vec{c_r^i}^T \vec{d}(x) \\
I_g^i(x) &= \rho(x)_g \vec{c_g^i}^T \vec{d}(x) \\
I_b^i(x) &= \rho(x)_b \vec{c_b^i}^T \vec{d}(x)
\end{aligned}
\tag{4.23}
$$

where $i = 1..k$. In other words, the observed intensities $\vec{I}(x) = [I_r^i(x)\ I_g^i(x)\ I_b^i(x)\ \dots\ I_b^k(x)]^T \in \mathbb{R}^{3k}$ can be written in terms of the matrix of environment map coefficients: $\mathbf{A}(x) \in \mathbb{R}^{3k \times n}$, where the first row of $\mathbf{A}(x)$ equals $\vec{c_r^1}^T$ etc, yielding the linear equation:

$$
\rho(x)\mathbf{A}(x)\vec{d}(x) = \vec{I}(x),
\tag{4.24}
$$

with $\vec{I}(x)$ being the unknown vector of coefficients which describes a Lambertian reflectance function for a point with normal direction $\mathbf{n}(x)$. The number of unknowns, num, in this system of equations (i.e. the elements of $\vec{d}(x)$) is determined by the number of degrees, l, used in the spherical harmonic expansion of the environment map and reflectance function: $num = (l+1)^2$.

Hence, for $l = 2$ degrees, the 9 observations given by 3 colour images are sufficient. In general, shape recovery using l degrees requires at least $\left\lceil \frac{(l+1)^2}{3} \right\rceil$ colour images. We also require that rank$(\mathbf{A}(x)) \geq num$. This requires that the colour channels of the environment map be linearly independent.

Note that the recovered vector of coefficients will be subject to a constant scaling, determined by the albedo. We therefore solve for $\rho(x)\vec{d}(x)$ for each pixel x using linear least squares. The remaining task is to recover the surface normal direction, $\mathbf{n}(x)$, which corresponds to the Lambertian reflectance function described by $\vec{d}(x)$. The most obvious way to accomplish this is to compute the spherical harmonic expansion of the coefficients $\vec{d}(x)$ using Equation 4.10 and find the maximum of the resulting function. The scaling by the albedo will apply a uniform scale factor but will not change the location of the maximum. However, such an approach is highly unstable as the exact location of the maximum is sensitive to noise. Instead, we optimise for the normal direction whose Lambertian spherical harmonic projection matches that of the recovered coefficients optimally in a least squares sense. This gives the recovered surface normal direction in global spherical coordinates. See Figure 4.5 for an example.

A more stable approach is instead to find the normal direction $\mathbf{n}^*(x)$, whose vector of coefficients $\vec{d}(\mathbf{n}^*(x))$ (computed according to Equation 4.17), minimises the angular difference with the vector of coefficients obtained by solving Equation 4.23:

$$\mathbf{n}^*(x) = \arg \max_{\|\mathbf{n}\|=1} \frac{\vec{d}(x) \cdot \vec{d}(\mathbf{n})}{\|\vec{d}(x)\| \|\vec{d}(\mathbf{n})\|}. \tag{4.25}$$

The advantage of this approach is that the albedo value is factored out by the normalisation of the lengths of the vectors. The albedo simply acts as a uniform constant scaling on $\vec{d}(x)$ and hence has no effect on the solution to Equation 4.25. The optimal normal direction can be obtained in a small number of iterations of any non-linear optimiser (we use a Matlab implementation of quasi-Newton optimisation).

4.2.4 Shape Recovery from a Single Image

A similar form of shape recovery is still possible given only a single colour image. The restriction is that only $l = 0, 1$ degrees can be used. Although this reduces the accuracy of the recovered shape it has been shown previously [20] that, assuming all lighting directions are equally likely, the accuracy of first degree spherical harmonic approximations is at least 87.5%. Although recovery using a first degree approximation would imply 4 unknowns per pixel with only 3 observations, we show that this problem can be made well-posed.

Elevation

Azimuth
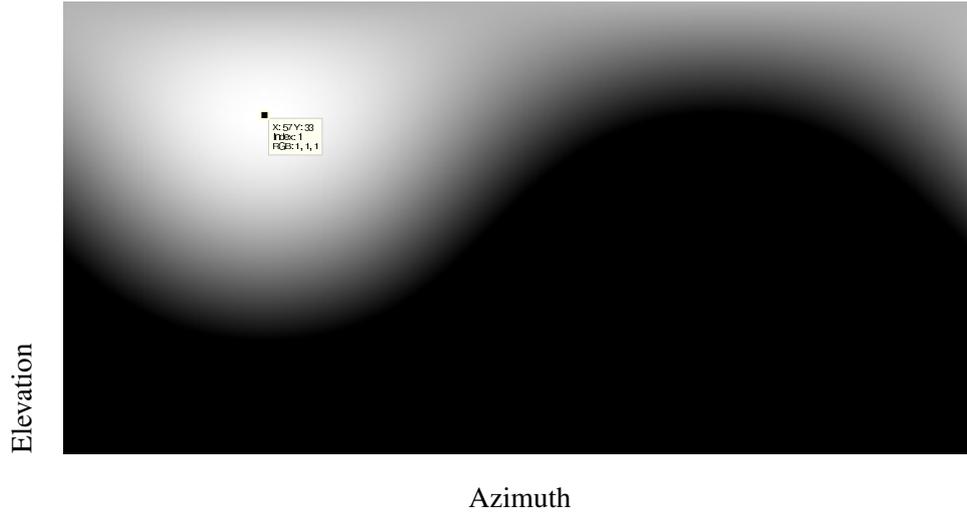
Figure 4.5: A spherical harmonic approximation to the Lambertian reflectance function for a particular surface normal direction. The function is displayed in the same equirectangular parameterisation as the environment maps. The maxima of this function indicates the surface normal direction (since Lambertian reflectance is maximal when the light source and normal are coincident). So, one way to recover the surface normal is to simply find the pixel with maximum intensity. This proves sensitive to noise, so instead we propose to search for the spherical harmonic coefficients that optimally match the observed reflectance function.

Draw from Green's method, an another interesting method is worth looking at which is known as rotational invariant. If a function g is a rotated copy of function f, then after the SH projection it is true that

$$\widetilde{g}(\vec{s}) = \widetilde{f}(\Theta \vec{s}) \tag{4.26}$$

where $\widetilde{g}$ and $\widetilde{f}$ are the approximation function of SH, Θ is the rotation matrix, and $\vec{s}$ is any input vector. In other words, it means SH projecting the rotated function g will give you exactly the same result as you have rotated the input to f before SH projecting. The approximation of the half cosine function can be seen as a rotated version of another half cosine function which the normal has the same direction of positive z axis or $[0\ 0\ 1]$. So

$$\vec{d}(\mathbf{n}) = D\vec{d}([0\ 0\ 1]) \tag{4.27}$$

It is proved that D can only change the coefficients within the same degree l. The matrix looks like a diagonal block with each block is a rotation matrix. For single image, we only use the matrix up

to degree 1. The matrix looks like

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta\sin\phi & -\cos\phi\sin\theta \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & \sin\theta & -\cos\theta\sin\phi & \cos\theta\cos\phi \end{bmatrix} \tag{4.28}$$

We follow the work of Ramamoorthi et al. [50], who show how to accelerate environment map rendering by performing a spherical harmonic projection of the environment map in a global coordinate system and of the Lambertian reflectance function under a local coordinate system. To perform rendering, a rotation matrix D is used to convert the global coordinates to local ones. Recall the irradiance function approximated by the spherical harmonics,

$$I(x) \approx \rho(x) \sum_{l}^{\infty} \sum_{m}^{l} c_{lm} d_{lm} \tag{4.29}$$

With the matrix, the image is derived by

$$I(x) = \vec{c}^T D \vec{d}([0\ 0\ 1]) \tag{4.30}$$

where $\vec{c}$ is the coefficient of the light source projected to the spherical harmonics, and $\vec{d}$ is the coefficient of the half cosine function projected to the spherical harmonics. In other words, the coefficients of the Lambertian reflectance function in a reference coordinate frame, $\vec{d}$, can be related to the coefficient for a specific normal direction by rotation: $\vec{d}(\mathbf{n}) = D\vec{d}([0\ 0\ 1])$.

We decompose the matrix D into two rotation matrix D_0 and D_1 which are related to the rotation of coefficients in different degrees. The rotation matrix for the degree 0 is an identity matrix with rank 1. The property of the diagonal block matrix makes the degree 0 term be eliminated while solving the linear equation.

$$\vec{c}^T D \vec{d}([0\ 0\ 1]^T) = c_0 D_0 d_0([0\ 0\ 1]) + \vec{c_1} D_1 \vec{d_1}([0\ 0\ 1]) \tag{4.31}$$

We still project the environment map under the global coordinates but the half-cosine function is represented under local coordinates. The linear equation we are solving is now:

$$\begin{aligned} I_r - c_r^0 d_0 &= \rho_r \vec{c_1}^T \vec{d_1}(\mathbf{n}) \\ I_g - c_g^0 d_0 &= \rho_g \vec{c_1}^T \vec{d_1}(\mathbf{n}) \\ I_b - c_b^0 d_0 &= \rho_b \vec{c_1}^T \vec{d_1}(\mathbf{n}) \end{aligned} \tag{4.32}$$

where the albedo ρ must be known or assumed to be unity. The rotation matrix which relates the reflectance function under local and global coordinates can be recovered by

$$D_1 = \vec{d_1}(\mathbf{n}) \vec{d_1}^{-1}([0\ 0\ 1]). \tag{4.33}$$

The components of the surface normal are embedded within the angular components of the rotation matrix D_1. To recover surface normal direction, we simply solve for the spherical coordinates which are consistent with the recovered rotation matrix.

4.3 Shape Estimation using Spherical Harmonic Images

We now present a completely different approach based on a spherical harmonic approximation to the convolution of the illumination function by the Lambertian reflectance function. This is the representation used by Basri et al. [4] which they refer to as spherical harmonic images. The integral of the Lambertian reflectance function over the upper hemisphere is equivalent to convolving the illumination function with a low pass filter. The low pass filter itself is a clamped-cosine function, which has the value of the cosine function at the upper hemisphere, and zero for rest of it. This method works under the assumption that no self-occlusion exists. The convolution function should be circularly symmetric. The spherical harmonics approximation provides a very efficient characterisation of the half cosine function, see Figure 4.6. Basri et al. [4] show that the first or-
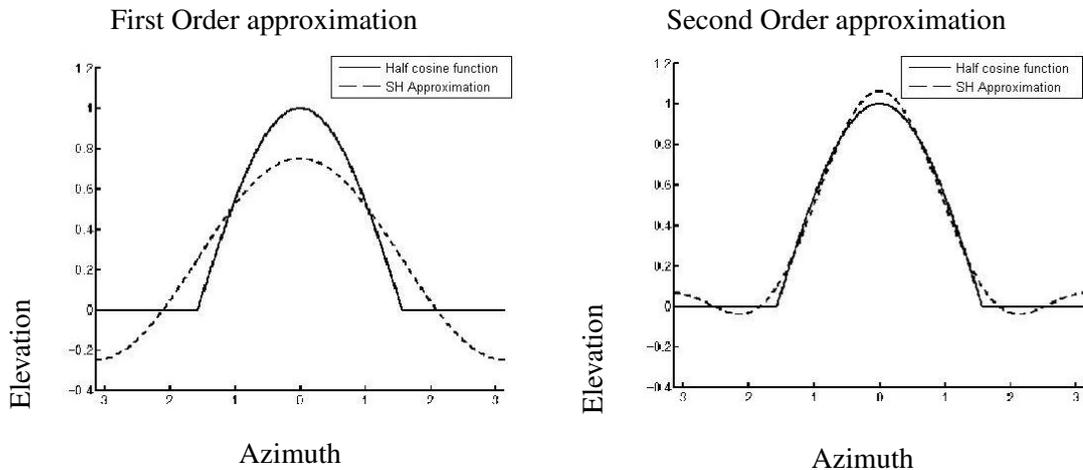
Figure 4.6: Approximation of spherical harmonics on half cosine function. The horizontal coordinate represents the input in radius where the vertical coordinate represents the value of function. The solid line represents the value of ideal half cosine function while the dash line represents the value from the approximated value

der spherical harmonics can approximate up to 85% of energy, while the second order spherical harmonics can approximate 98% of energy for any illumination condition.

Returning to the image formation model under environmental illumination (Equation 4.4), we

assume that there is no occlusion and constant unit albedo, so the visibility function is one for every value. We have

$$I(x) = \int_{\Omega(\mathbf{n}(x))} L(\vec{\omega})(\mathbf{n}(x) \cdot \vec{\omega})d\vec{\omega}, \tag{4.34}$$

The $\mathbf{n}(x) \cdot \vec{\omega}$ is the so called half cosine function. Thus, the 4.34 is analogous to a convolution, in which we centre a kernel (the half-cosine function), and integrate its product with a signal $L(\vec{\omega})$. Real spherical harmonics with up to second order is used here, as shown in Appendix A. According to the Funk-Hecke theorem, the reflectance function can be written as:

$$I(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (\alpha_l c_{lm}) Y_{lm} \tag{4.35}$$

with

$$\alpha_l = \sqrt{\frac{4\pi}{2l+1}} k_l \tag{4.36}$$

k_n can be pre-calculated.

$$
\begin{array}{ll}
k_0 = \frac{\sqrt{\pi}}{2} \approx 0.8862 & k_1 = \sqrt{\frac{\pi}{3}} \approx 1.0233 \\
k_2 = \frac{\sqrt{5\pi}}{8} \approx 0.4954 & k_4 = -\sqrt{\frac{\pi}{16}} \approx -0.1108 \\
k_6 = \frac{\sqrt{13\pi}}{128} \approx 0.0499 & k_8 = \sqrt{\frac{17\pi}{256}} \approx -0.0285
\end{array}
\tag{4.37}
$$

($k_3 = k_5 = k_7 = 0$).

Fig.4.7 shows Basri's method. Instead of approximating the illumination function, they apply the convolution on it directly.

The result is that the irradiance equation can be expressed in terms of spherical harmonics as:

$$I(x) = \sum_{l,m} a_l c_{lm} Y_{lm}(\mathbf{n}), \tag{4.38}$$

where c_{lm} are the illumination coefficients and a_l the Lambertian reflectance coefficients (which vary only with l since Lambertian reflectance has no azimuthal dependence). The first nine spherical harmonics (with $l \leq 2$) are simply constant ($l = 0$), linear ($l = 1$) and quadratic ($l = 2$) polynomials of the three components of the surface normal $\mathbf{n}$. A first order approximation captures 87.5% of the energy of Lambertian irradiance, whilst a second order approximation captures 99.22%. These approximations are sufficiently accurate to be of great utility in rendering [49].

4.3.1 First Order Solution

In an order-1 approximation, this leads to a linear image formation model which for RGB colour images can be written as: $I = [I_r \ I_g \ I_b]^T = \mathbf{A}\mathbf{n} + \vec{a}$, where $\mathbf{A} \in \mathbb{R}^{3 \times 3}$. The first row of $\mathbf{A}$
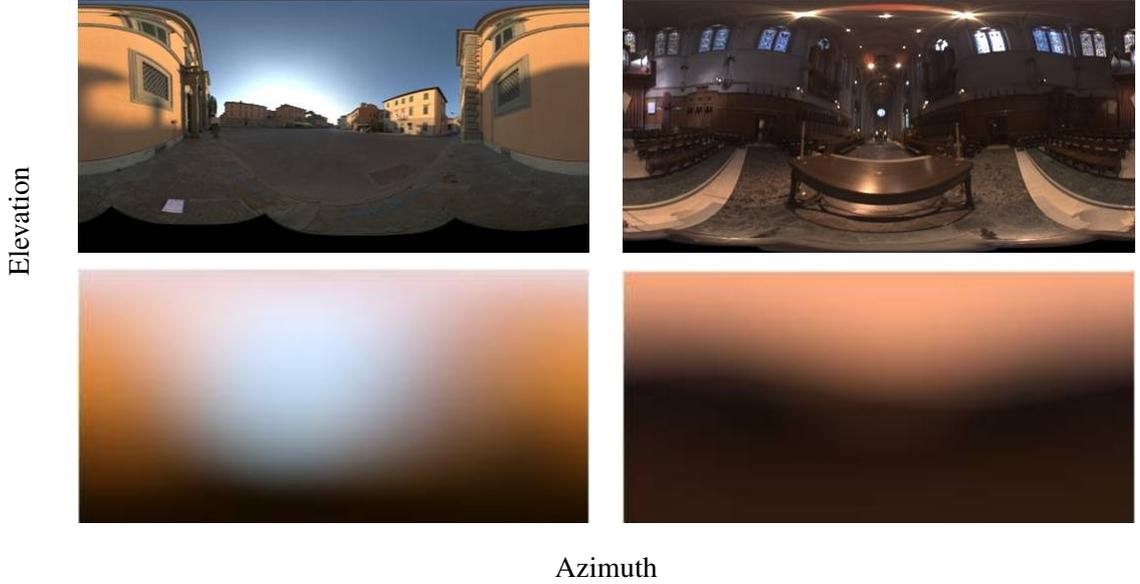
Figure 4.7: Irradiance function after applying the half cosine convolution filter to the illumination map. The images in the top row are the illumination function from HDR data. The images in the bottom row are the corresponding irradiance function. The coordinates are the longitude-latitude coordinates described in previous section.

contains the degree-1 spherical harmonic coefficients corresponding to the Lambertian reflectance function and the illumination environment in the red channel, i.e.

$$\mathbf{A}_1 = \rho_r \sqrt{\frac{3}{4\pi}} \begin{bmatrix} c_{11}^r & c_{1-1}^r & c_{10}^r \end{bmatrix}, \tag{4.39}$$

where c_{lm}^α is the degree l, order m illumination coefficient in colour channel α. We factor the albedo of the surface in the red channel, ρ_r, into the row. The second/third row are constructed similarly for the green/blue channels. The vector $\vec{a} \in \mathbb{R}^3$ contains the degree-0 coefficients for each colour channel: $\vec{a} = \frac{1}{\sqrt{4\pi}}[\rho_r c_0^r \quad \rho_g c_0^g \quad \rho_b c_0^b]^T$. c_{lm}^α is acquired by applying the spherical approximation on illumination map. For the first order

$$\begin{aligned} c_{1m}^\alpha &= \sum_{\vec{\omega}} L^\alpha(\vec{\omega}) Y_{1m}(\vec{\omega}) \\ &= \sqrt{\frac{3}{4\pi}} \begin{bmatrix} L_x^\alpha & L_y^\alpha & L_z^\alpha \end{bmatrix} \end{aligned} \tag{4.40}$$

4.3.1.1 Illumination Estimation

We begin by describing a simple method for estimating the illumination coefficients when the illumination environment is unknown.

We begin by using a crude estimate of the surface normals of the object inferred from the normals at the occluding boundary to make an estimate of $\mathbf{A}$ and $\vec{a}$. Our aim is to use a sample of

59

k estimated surface normals and their corresponding image intensities:

$$\mathbf{N} = \begin{bmatrix} \mathbf{n}_1^T & 1 \\ & \vdots \\ \mathbf{n}_k^T & 1 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} I_1^T \\ \vdots \\ I_k^T \end{bmatrix}, \tag{4.41}$$

to compute the illumination coefficients and surface albedos by solving the following linear system of equations:

$$\mathbf{I} = \mathbf{N} \begin{bmatrix} \mathbf{A}^T \\ \vec{a}^T \end{bmatrix}. \tag{4.42}$$

Note that the surface albedos are implicitly factored into the recovered coefficients.

In order to estimate these initial normals, we assume that the image has been segmented into a set of foreground, $\mathcal{F}$, and background, $\mathcal{B}$, pixels. Background pixels $x \in \mathcal{B}$ are excluded from the shape recovery. A surface normal $\mathbf{n}(x)$ at the occluding boundary, i.e. $x \in \mathcal{O}$, lies in the image plane (i.e. $n_z(x) = 0$) and is orthogonal to the tangent to the boundary. Hence, boundary normals can be computed by estimating the tangent to the foreground segmentation.

Unfortunately the boundary normals alone cannot be used to solve (4.42) since all entries in the third column of $\mathbf{N}$ would equal zero and hence: $\text{rank}(\mathbf{N}) = 3$. Instead, we use the boundary normals to obtain a very crude approximation to the surface by bicubically interpolating interior normals from the known boundary normals. We therefore solve (4.42) using all the estimated foreground normals. Although the normal estimates are very coarse, the system is highly overdetermined and in practice this yields good results. Note that an alternative is to instead use a weak perspective model. In this case, the z component of the boundary normals is related to the position in the image plane. However, the distance to the object and camera projection parameters must be known.

4.3.1.2 First Order Shape Estimation

With the illumination coefficients either estimated or known, we are ready to proceed to shape estimation.

We seek the unit vector $\mathbf{n}^*$ which minimises brightness error. This results in a quadratically constrained linear least squares problem:

$$\mathbf{n}^* = \arg\min_{\mathbf{n}} \|\mathbf{An} - \vec{b}\|^2, \text{ s.t. } \|\mathbf{n}\| = 1, \tag{4.43}$$

where $\vec{b} = \mathbf{I} - \vec{a}$. Contrast this to the classical photometric stereo equation in which albedos are factored into the unknowns and the unit norm constraint is therefore not required. Note that we

require colour channels of the illumination environment to be independent and hence $\text{rank}(\mathbf{A}) = 3$. This is similar to the photometric stereo requirement of point light source directions which are not co-linear.

In general, quadratically-constrained quadratic programs (QCQP) such as (4.43) are difficult to solve and may be non-convex. However, in this simple case the optimal solution can be found using the method of Lagrange multipliers. To do so, we reformulate our problem as finding the stationary points of:

$$\Lambda(\mathbf{n}, \lambda) = \|\mathbf{A}\mathbf{n} - \vec{b}\|^2 + \lambda(\|\mathbf{n}\|^2 - 1) \tag{4.44}$$

We take the singular value decomposition (SVD):

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{4.45}$$

and denote

$$\vec{c} = \mathbf{U}^T\vec{b} \ \ and \ \ \vec{y} = \mathbf{V}^T\mathbf{n} \tag{4.46}$$

One important property of singular value decomposition is the matrix $\mathbf{U}$ and $\mathbf{V}$ are real unitary matrix . So we have $\|\mathbf{n}\| = \|\vec{y}\|$. The diagonals of $\mathbf{\Sigma}$ are the singular values, denoted σ_i. The stationary points occur when all partial derivatives are zero

$$\partial_{y_i}\Lambda = \sigma_i^2 y_i - c_i\sigma_i + y_i\lambda = 0$$
$$\partial_\lambda\Lambda = \sum_i y_i - 1 = 0 \tag{4.47}$$

These are the *normal equations*, the last of which is the original quadratic constraint. By substitution we obtain a scalar equation in λ, the *secular equation*:

$$\sum_{i=1}^{3}\left(\frac{c_i\sigma_i}{\sigma_i^2 + \lambda}\right)^2 = 1. \tag{4.48}$$

When expanded, this gives an order-6 polynomial in λ with up to six real roots.

$$(\sigma_1 c_1(\sigma_2^2+\lambda)(\sigma_3^2+\lambda))^2+(\sigma_2 c_2(\sigma_1^2+\lambda)(\sigma_3^2+\lambda))^2+(\sigma_3 c_3(\sigma_2^2+\lambda)(\sigma_1^2+\lambda))^2 = (\sigma_1^2+\lambda)^2(\sigma_2^2+\lambda)^2(\sigma_3^2+\lambda)^2 \tag{4.49}$$

To solve this polynomial function, we compute the eigenvalues of the companion matrix of the coefficient as described in [15]. To simplify the description, we assume the function of λ with order 6 is represented as follows:

$$c_6\lambda^6 + c_5\lambda^5 + c_4\lambda^4 + c_3\lambda^3 + c_2\lambda^2 + c_1\lambda + c_0 = 0 \tag{4.50}$$

61

The companion matrix is constructed as

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & -c_0/c_6 \\
1 & 0 & 0 & 0 & 0 & -c_1/c_6 \\
0 & 1 & 0 & 0 & 0 & -c_2/c_6 \\
0 & 0 & 1 & 0 & 0 & -c_3/c_6 \\
0 & 0 & 0 & 1 & 0 & -c_4/c_6 \\
0 & 0 & 0 & 0 & 1 & -c_5/c_6
\end{bmatrix}
\tag{4.51}
$$

The eigenvalues of this matrix give the roots of the secular equation. The largest real root, λ^*, gives the smallest residual and hence the global optimum:

$$
\mathbf{n}^* = \mathbf{V}
\begin{bmatrix}
c_1\sigma_1/(\sigma_1^2 + \lambda^*) \\
c_2\sigma_2/(\sigma_2^2 + \lambda^*) \\
c_3\sigma_3/(\sigma_3^2 + \lambda^*)
\end{bmatrix} .
\tag{4.52}
$$

Note that there is an additional constraint on $\mathbf{n}^*$. Since the view vector is coincident with the positive z-axis, the z component of the surface normal must be negative, i.e. $\mathbf{n}^* \cdot [0\ 0\ 1]^T < 0$. In practice we find that this constraint is rarely breached and where it is, we evaluate the other roots of (4.48) to find the root with smallest error which has a positive z component. An alternative approach would be to add this inequality constraint to (4.43), which would require solution of the Karush-Kuhn-Tucker conditions.

4.3.1.3 Smoothness Constraint

The method above is purely local and per-pixel solutions are independent of each other. The surface normal is only decided by the intensity function but nothing else. This results in high sensitivity to noise and, depending on the nature of the illumination environment, may be under-constrained. However, it can be used to obtain a locally optimal initialisation which is refined by adding a regularisation term which encourages consistency between adjacent normals. A smoothness constraint can be incorporated without increasing the complexity of the optimisation (i.e. a global optimum can still be obtained).

We draw the idea from our work in Chapter 3. We formulate our smoothness constraint to require only piecewise smoothness (i.e. we do not smooth across discontinuities). We make the assumption that pixels in close proximity with a similar intensity have similar surface normal directions, i.e. we use the change in intensity as a measure of the change in surface normal direction. This is reasonable because the reflectance map as captured by the spherical harmonics

is smoothly varying. We define a region $\mathcal{R}(x)$ about a pixel x (where $\mathcal{R}(x) \subset \mathcal{F} \wedge \mathbf{n}(x) \notin \mathcal{R}(x)$). The surface normals over this region correspond to a distribution of points on the unit sphere. We therefore use the Karcher mean to compute the local average normal, by finding the point on the unit sphere which minimises angular distance to the sample points:

$$\bar{\mathbf{n}}(x) = \arg\min_{\|\mathbf{x}\|=1} \sum_{y \in \mathcal{R}(x)} \left[w_{x,y} \arccos(\mathbf{x} \cdot \mathbf{n}(y)) \right], \tag{4.53}$$

where the weights

$$w_{x,y} = \frac{\|\mathbf{I}_x - \mathbf{I}_y\|^{-2}}{\sum_{y \in \mathcal{R}(x)} \|\mathbf{I}_x - \mathbf{I}_y\|^{-2}} \tag{4.54}$$

are the reciprocal of the variance in pixel intensity. The weighted Karcher mean can be computed iteratively using gradient descent on the sphere [39]:

$$\bar{\mathbf{n}}(x)^{(t+1)} = \mathrm{Exp}_{\bar{\mathbf{n}}(x)^{(t)}} \left(\sum_{y \in \mathcal{R}(x)} w_{x,y} \mathrm{Log}_{\bar{\mathbf{n}}(x)^{(t)}}(\mathbf{n}(y)) \right), \tag{4.55}$$

where $\mathrm{Exp}_b(x)$ is the exponential map of x at base point b, similarly for the logarithmic map [39]. The local averages are then added to the least squares objective for each pixel and new normals estimated by resolving (4.43). The two steps of smoothing and solving local optimisations can be iterated until convergence. However, in the presented results we simply perform one smoothing iteration.

The conventional methods for adding smoothness constraint are either iterating the constraint in turn after data closeness constraint, or the method we describe earlier. They are both not applicable for the scheme here. We incorporate the smoothness constraint into the quadratic function. To do so, we add a term to the least squares objective which penalises departures from the local average direction:

$$\mathbf{A}_4 = w\bar{\mathbf{n}}^T \text{ and } b_4 = w \tag{4.56}$$

where w is a regularisation weight and $\bar{\mathbf{n}}$ the local average direction. $w = 0$ corresponds to the unregularised approach described above.

The units of $\|\mathbf{A}_4\mathbf{n} - b_4\|^2$ are squared distance, whereas the units of $\|\mathbf{A}_{1\ldots3}\mathbf{n} - \vec{b}_{1\ldots3}\|^2$ are squared intensity. To make the two errors commensurate, for a pixel x we set $w(x) = v\|\mathbf{I}(x)\|$, where v is a global parameter which controls the trade-off between satisfaction of the brightness and smoothness terms. With $v = 1$, a normal which is $90°$ from the local mean results in the same error as a brightness error equal to the brightness of the pixel. v is an empirical result, larger weight will give an over smoothed result, and smaller weight might not affect the result of local minimum. Experiment shows that the value of v should be in order of 100.

$\mathbf{A}$ is now a 4×3 matrix, the SVD of $\mathbf{A}$ is $\mathbf{U}$ 4×4 unitary matrix, $\boldsymbol{\Sigma}$ 4×3 matrix with three non-zero values, and $\mathbf{V}$ 3×3 unitary matrix. The only term that changes is $\vec{c} = \mathbf{U}^T \vec{b}$. Take the first three term of $\vec{c}$, the solving procedure is the same with the unregularised approach. Notice that no matter how many linear terms are added into $\mathbf{A}$, the equation solving procedure is the same. This advantage gives us another effective framework which can add all the constraints into it, only if the constraint can be formulated as a linear equation. During the experiment, we add an integrability constraint with proper weight. However, the integrability is difficult to describe as a linear constraint. An alternative way is calculating the non-integrable normal first, and then apply the smoothness constraint to obtain an integrable surface with correspondent normals. Integrable constraint now is formulated as

$$\mathbf{n} \cdot \mathbf{n_{int}} = 1 \tag{4.57}$$

where $\mathbf{n_{int}}$ is the integrable normal. The algorithm is described as follows:

1. Obtain $\mathbf{n}^{(0)}(x)$ from the equation with smoothness constraint

2. Add the integrable constraint from the result of $\mathbf{n}^{(0)}(x)$

3. Calculate $\mathbf{n}^{(1)}(x)$ with the integrability constraint is added

4. Stop after certain number of iteration(we use 20 here), otherwise iterate to step 2

Since the calculation for surface normal is trivial (a small eigendecomposition and matrix multiplications), even with dozens of iterations, the CPU time is still trivial.

4.3.2 Second Order Solution

The method described above for a first order approximation is both efficient and robust. However, it is well known that a second (e.g. 9 basis) spherical harmonic approximation gives a considerable increase in accuracy. The challenge here is that the image formation model becomes quadratic, leading to a higher order objective function.

Under a quadratic image formation model, the image intensity is given by:

$$\mathbf{I} = \begin{bmatrix} \mathbf{n}^T \mathbf{Q}_r \mathbf{n} + \vec{a}_r^T \mathbf{n} + d_r \\ \mathbf{n}^T \mathbf{Q}_g \mathbf{n} + \vec{a}_g^T \mathbf{n} + d_g \\ \mathbf{n}^T \mathbf{Q}_b \mathbf{n} + \vec{a}_b^T \mathbf{n} + d_b \end{bmatrix}, \tag{4.58}$$

where $\mathbf{Q}_c \in \mathbb{R}^{3 \times 3}$, $\vec{a}_c \in \mathbb{R}^3$ and $d_c \in \mathbb{R}$ are the quadratic, linear and constant coefficients respectively for colour channel $c \in \{r, g, b\}$. $\mathbf{Q}_c$ is symmetric and since $n_x^2 + n_y^2 + n_z^2 = 1$, it

only has 6 degrees of freedom. $\vec{a}_c$ and d_c are the same coefficients as in the linear model above. $\mathbf{Q}_c$ contains the order 2 spherical harmonic coefficients.

Our aim is to solve the following optimisation problem:

$$\mathbf{n}^* = \arg \min_{\mathbf{n}} \sum_{c \in \{r,g,b\}} \left(\mathbf{n}^T \mathbf{Q}_c \mathbf{n} + \vec{a}_c^T \mathbf{n} + d_c - I_c \right)^2, \tag{4.59}$$

$$\text{s.t. } \|\mathbf{n}\| = 1.$$

This involves minimisation of a quartic in three variables subject to a quadratic constraint. We are not aware of any globally optimal method for solving such a system. Forming the Lagrangian and setting partial derivatives to zero leads to a system of three cubic and one quadratic equation in four unknowns which cannot be solved in closed form.

The most obvious approach to solving Equation 4.59 would be to obtain a reasonable initialisation and then use a constrained optimisation to locally optimise the objective function subject to the unit norm constraint. This is exactly what was recently proposed by Johnson and Adelson [27]. However, in practice this is a highly unstable approach. The objective function is littered with local minima, which lead Johnson and Adelson [27] to have to engineer a highly complex optimisation procedure utilising coarse-to-fine search and boundary, smoothness and integrability constraints, all of which are controlled by numerous parameters which must be carefully selected.
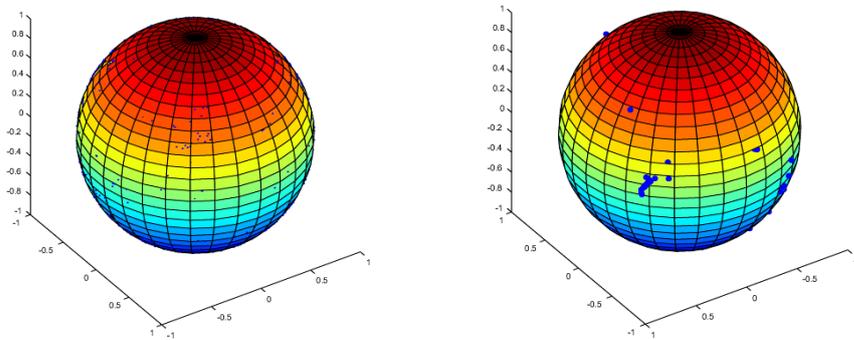
Figure 4.8: The left image shows 1,000 random initialisations for the local optimisation, and the right image shows the result after local optimisation has been applied (point shown in blue).

To illustrate this point, in Figure 4.8 we show an empirical demonstration of this phenomena. We use an order 2 approximation to the environment map in Figure 4.8 and compute the colour for a randomly selected surface normal. We then randomly generate 1,000 surface normal directions and apply local optimisation to minimise Equation 4.59. On the left we show the 1,000 initialisa-

tions and on the right we show the 1,000 converged solutions. It can be clearly seen that there are numerous local minima in which a local optimisation-based approach may become stuck.

We propose an alternative approach which performs very well in practice (i.e. is both efficient and almost always obtains the global minimum). We decompose the error function into two systems, each of which are linear in the unknowns and hence convex. This leads to a constrained bilinear system for which we can obtain the global minimum using alternating least squares.

Equation 4.58 can be rewritten as a linear system:

$$\mathbf{i} = \mathbf{B} \begin{bmatrix} 1 \\ \mathbf{n} \\ \phi(\mathbf{n}) \end{bmatrix} \tag{4.60}$$

where $\phi : \mathbb{R}^3 \mapsto \mathbb{R}^6$ is a feature map:

$$\phi(\mathbf{n}) = \begin{bmatrix} n_x^2 \ n_y^2 \ n_z^2 \ n_x n_y \ n_x n_z \ n_y n_z \end{bmatrix}^T, \tag{4.61}$$

and $\mathbf{B} \in \mathbb{R}^{3 \times 10}$ is formed by stacking the spherical harmonic coefficients as follows:

$$\mathbf{B} = [\mathbf{a} | \mathbf{A} | \mathbf{H}]. \tag{4.62}$$

$\mathbf{A}$ is the matrix of order-1 coefficients, as in Section 4.3.1. $\mathbf{H} \in \mathbb{R}^{3 \times 6}$ contains the order-2 coefficients. Each row contains the coefficients for one colour channel in the same order as in Equation 4.61. Equation 4.62 is the form that has been widely used in the literature, both for defining a linear subspace for recognition [67] and, in the method of Basri et al. [5], for shape estimation. As with the quadratic form, there is some redundancy caused by the unit norm constraint so $\mathbf{H}$ has 18 elements but only 15 degrees of freedom. The optimisation given in Equation 4.59 can be similarly rewritten:

$$\mathbf{n}^* = \arg \min_{\mathbf{n}} \left\| \mathbf{B} \begin{bmatrix} 1 \\ \mathbf{n} \\ \phi(\mathbf{n}) \end{bmatrix} - \mathbf{I} \right\|^2, \text{s.t.} \ \|\mathbf{n}\| = 1, \tag{4.63}$$

which can itself be transformed into an optimisation with 10 unknowns and 8 constraints:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{B}\mathbf{x} - \mathbf{i}\|^2, \tag{4.64}$$

$$\text{s.t.} \ x_1 = 1, \ x_5 = x_2^2, \ x_6 = x_3^2, \ x_7 = x_4^2,$$

$$x_8 = x_1 x_2, \ x_9 = x_1 x_3, \ x_{10} = x_2 x_3 \ \text{and}$$

$$x_5 + x_6 + x_7 = 1.$$

Note that there is no analytical gain in this transformation. Constructing the Lagrangian for this system, setting partial derivatives to zero, and simplifying and substituting leads to the same system of three cubic equations and one quadratic equation in four unknowns as for Equation 4.59. However, we note the following. Equation 4.64 is linear in each element of $\mathbf{x}$. Moreover, the constraints are linear in the second order terms (i.e. x_5 to x_{10}).

Hence, we attempt to solve Equation 4.64 by decomposing it into two linear systems which we solve using alternating least squares. The low order system is exactly that solved in the unregularised linear model. The high order system is a linear system with 5 unknowns.

We solve the high order system as follows. We wish to encourage consistency between the high and low order solutions. Note that we cannot enforce consistency as a hard constraint since this would completely determine the solution (i.e. solving the high order system would provide no extra information). Hence, when solving for the high order terms we strictly enforce the unit norm constraint (which becomes a linear equality constraint) and penalise errors between the high order terms and predictions of the high order terms from low order terms in a least squares sense. i.e.

$$
\mathbf{h}^* = \arg\min_{\mathbf{h}} \left\| \begin{bmatrix} \mathbf{H} \\ \mathbf{I}_6 \end{bmatrix} \mathbf{h} - \begin{bmatrix} \mathbf{I} - \mathbf{An} + \mathbf{a} \\ n_x^2 \\ n_y^2 \\ n_z^2 \\ n_x n_y \\ n_x n_z \\ n_y n_z \end{bmatrix} \right\|^2, \tag{4.65}
$$

$$
\text{s.t. } h_1 + h_2 + h_3 = 1
$$

$\mathbf{I}_6$ is an identity matrix, and n_x, n_y and n_z come from the solution of the low order system. Intuitively, the upper part of the vector difference encourages the high order terms to explain the residual difference between the observed intensity and the low order estimate. The lower part of the vector difference encourages the high order terms to be consistent with quadratic terms formed from the normal components estimated in the low order system.

Solving the high order system is simply a quadratic programming problem (linear least squares with a linear equality constraint). We iterate solution of the low and high order systems in an alternating least squares manner. Specifically, we initially compute the low order solution directly from the image intensities. We then solve for the high terms before subtracting the predicted high

67

order image intensity from the observations before resolving the low order system. This is iterated to convergence. Typically, this process converges within 30 iterations.

The low order system is indeed an minimization of linear equation under a quadratic constraint, which is the same with first order solution. We are solving for

$$\mathbf{n}^* = \arg \min_{\mathbf{n}} \|\mathbf{A}\mathbf{n} - \vec{d} - \vec{I} - \vec{I}_H\|^2, \text{ s.t. } \|\mathbf{n}\| = 1 \tag{4.66}$$

where $\vec{I}_H = \vec{H}h$ is the predicted high order contribution to the intensity.
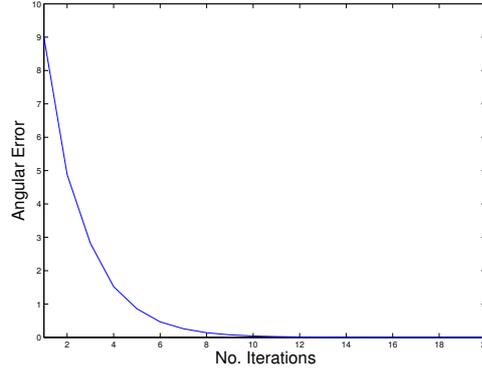
Figure 4.9: Convergence of our method

Figure 4.9 shows the convergence of our method on the same example as in Figure 4.8. Within 20 iterations our method obtains the globally optimal solution and shows smooth convergence properties. This behaviour was observed in all samples we generated.

4.3.2.1 Smoothness Constraint

As in the first order approach, the above solution to the image irradiance equation is independent for each pixel and, hence, highly susceptible to noise. For this reason it is necessary to add regularisation terms which encourage local consistency. Unlike the first order situation, the calculation in each iteration is only a penalise error for consistence, the smooth term need to be a weak constraint. In our method, we chose a isotropic smooth term. The linear smoothness term is similar to incorporating the integrability constraint in the first order solution. We do the smoothness with in a 3×3 window. For each normal in the window, we have 8 linear constraints

$$\mathbf{A}_{row} = w\mathbf{n}_{\mathbf{m}}^{\mathbf{T}} \text{ and } \mathbf{b_{row}} = \mathbf{w} \tag{4.67}$$

where $\mathbf{n}_{\mathbf{m}}^{\mathbf{T}}$ are the normals of the neighbouring pixel within the window from previous iteration, and for a pixel x we set $w(x) = v\|\mathbf{I}(x)\|$, where v is a global parameter which controls the trade-off between satisfaction of the brightness and smoothness terms.

Otherwise, the order-2 algorithm is identical to the order-1 algorithm and proceeds according to the same steps.

4.4 Experimental Results

We now provide experimental results for the methods described above. We divide this into three sections: 1. the method based on spherical harmonic lighting, 2. the first order harmonic image method, 3. the second order harmonic image method.

4.4.1 Spherical Harmonic Lighting Method

We now present experimental results of our method. We begin by showing results for synthetic imagery rendered under contrived low frequency environmental illumination. We select illumination environments which ensure our rank constraints are met. Figure 4.10 shows the three environment maps and three objects used in our experiment. The ground truth shape and the shape estimated by applying our method to three input images are shown in Figure 4.12. We integrate heightmaps from the estimated surface normals. The recovered surfaces retain much of the fine detail of the ground truth surfaces whilst also adhering to the global pattern of convexity and concavity.

Figure 4.10: The environment maps on the top with the corresponding input images.

Figure 4.11: Top row: ground truth surfaces, bottom row: estimated surfaces.

In Figure 4.12 we show an example of surface recovery from one input image. Although there is clearly an increase in the global distortion, much of the fine surface detail is still correctly recovered.
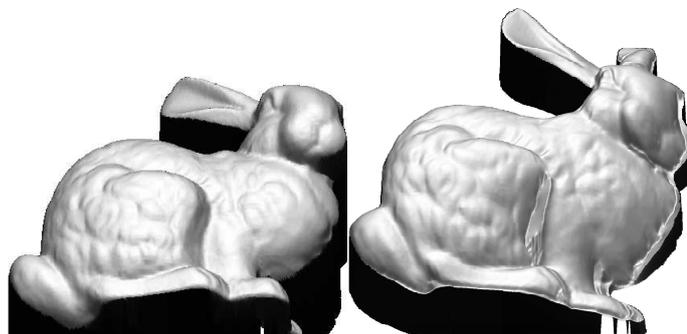
Figure 4.12: Heightmap recovered from a single image (right) versus ground truth (left).

Finally, we apply our method to some real world images. In order to approximate a low

frequency illumination environment, we photograph a white statue with approximately Lambertian reflectance inside a 'light stage' [35]. The illumination environments we approximate can be seen in the top row of Figure 4.13. Note that since the light stage only contains 42 LED lights, in practice the environment is a discrete approximation to those shown. Photographs of the object in the light stage under the corresponding illumination conditions are shown in the bottom row of Figure 4.13. We perform single image shape recovery for each of the three images. The object contains many non-convex regions and hence portions of the incident hemisphere are occluded. Despite this, the reconstructions shown in Figure 4.14 still capture the global and local shape of the face well. We show the result of reconstructing the statue using the more restrictive spherical gradient method of Ma et al. [35] on the left of the Figure (this method requires 4 images under a specific lighting setup).
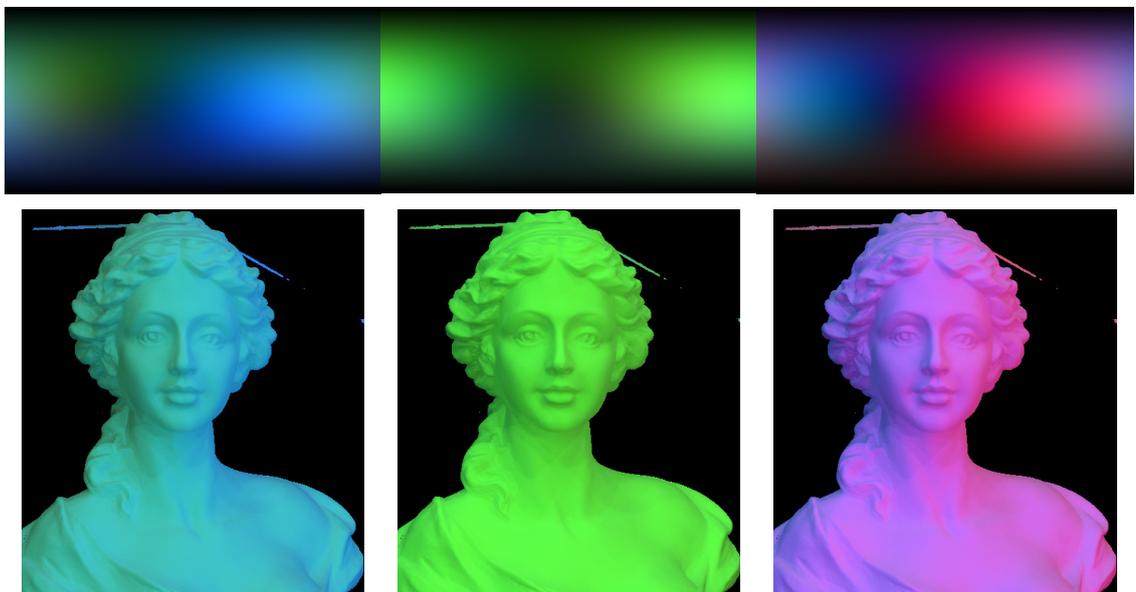
Figure 4.13: Real world input images (bottom) imaged in a light stage approximating the low frequency illumination shown in the top row.

Figure 4.14: Estimated height maps using the method of [35] followed by the reconstructions from the images shown in Figure 4.13.

4.4.2 First Order Solution

First, we show our method with synthetic data in Fig.4.16 with the natural illumination in Fig.4.15. The result greatly catch the convex concave information of the object and provide accurate data with details. Even without any additional constraint such as smoothness constraint, it is still plausible.
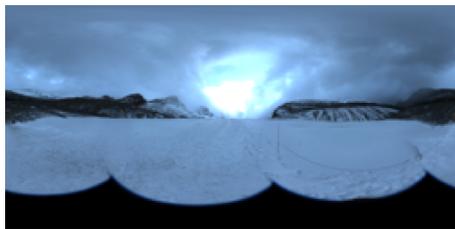
Figure 4.15: The Glacier illumination Environment.
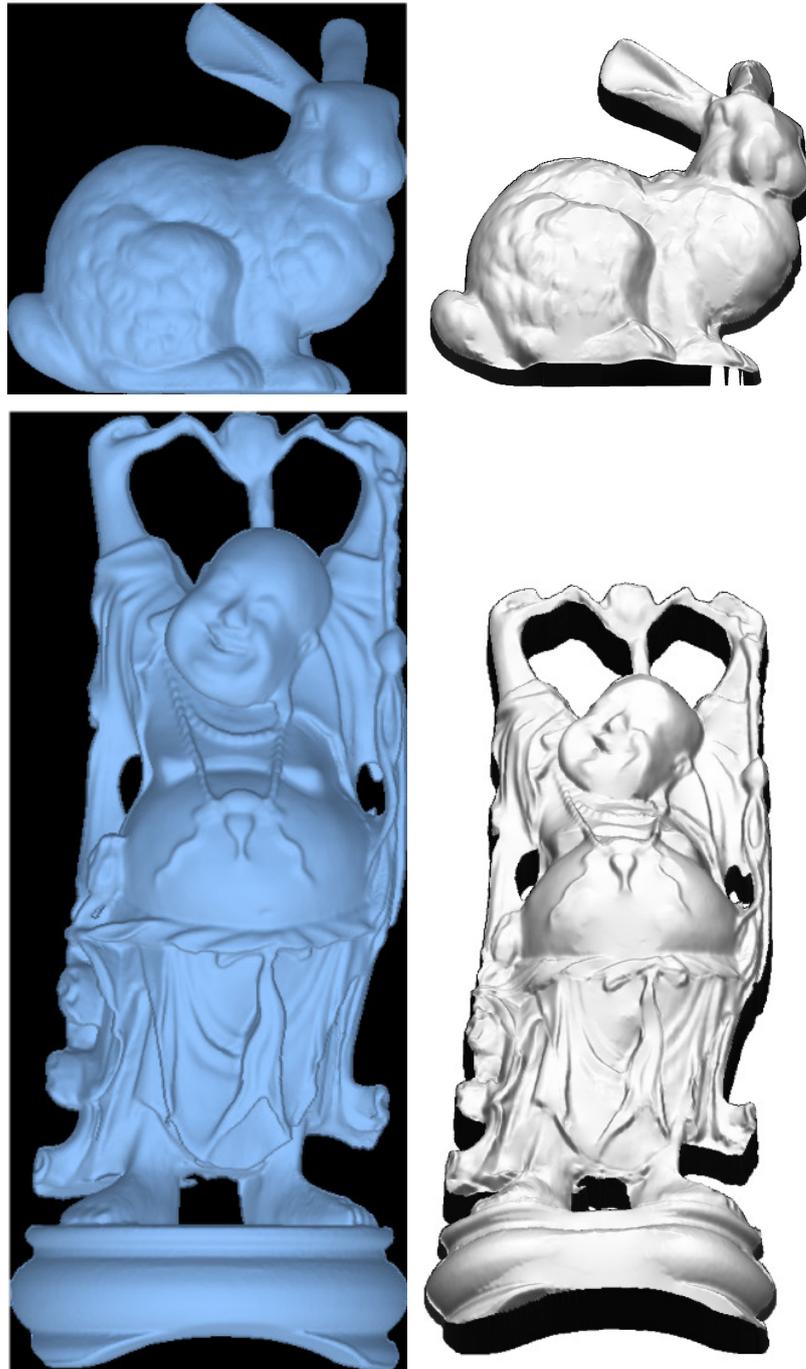
Figure 4.16: Left: Synthetic images rendered under Glacier environmental illumination. Right: Integrated surfaces recovered from the surface normals estimated by the order-1 method.

To prove the robustness of our method, we test it in two ways: 1)Estimated illumination 2)Cast shadow

74
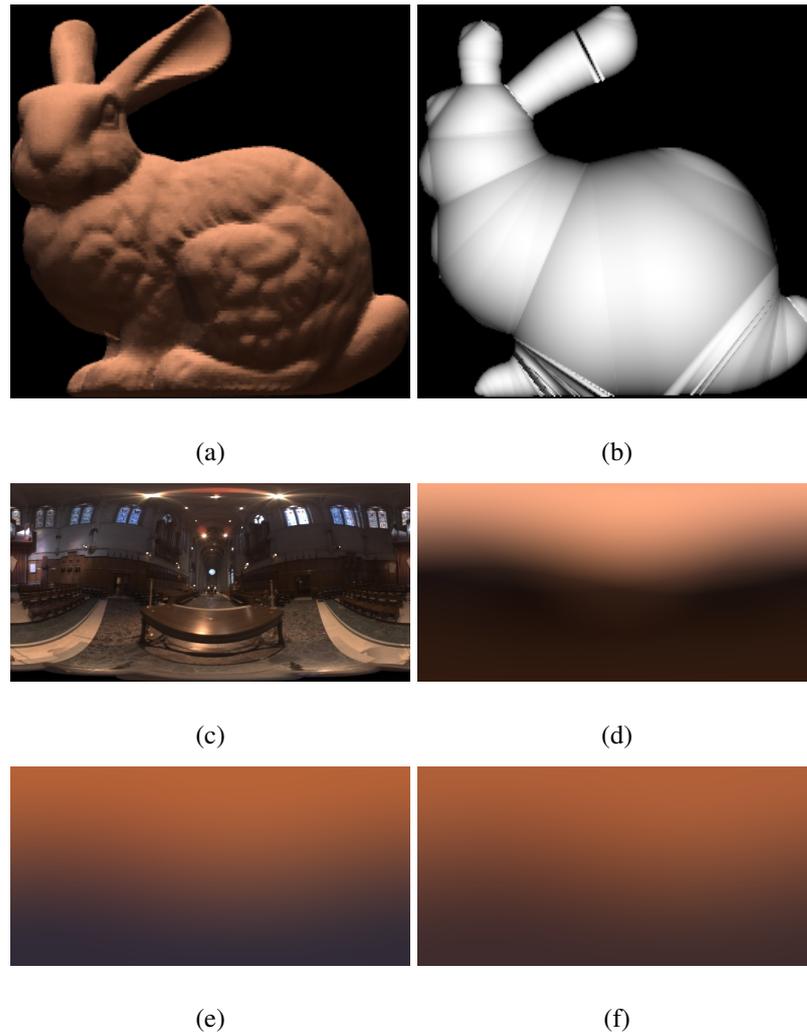
Figure 4.17: Illumination estimation example.

We show an example of illumination estimation in Fig. 4.17. From an input image (a), we compute a background segmentation which is used to compute boundary normals (b). The ground truth environment map (c) is convolved with the Lambertian reflectance function to yield the reflectance map shown in (d). The 4 basis approximation of this is shown in (e) whilst the 4 basis estimate using the normals in (b) is shown in (f). The relative error between (d) and (e) is typically less than 5%.

We show shape reconstructions for two objects under different environment maps in Fig. 4.18 and 4.20. The top row of Fig. 4.18 shows the estimated surfaces normals rendered with frontal illumination and with the surface normal components mapped to RGB. The integrated height map is shown below. The input image in the top left of Fig. 4.20 is rendered with the environment map in Fig. 4.19. Again we show estimated surface normals and integrated surface height.
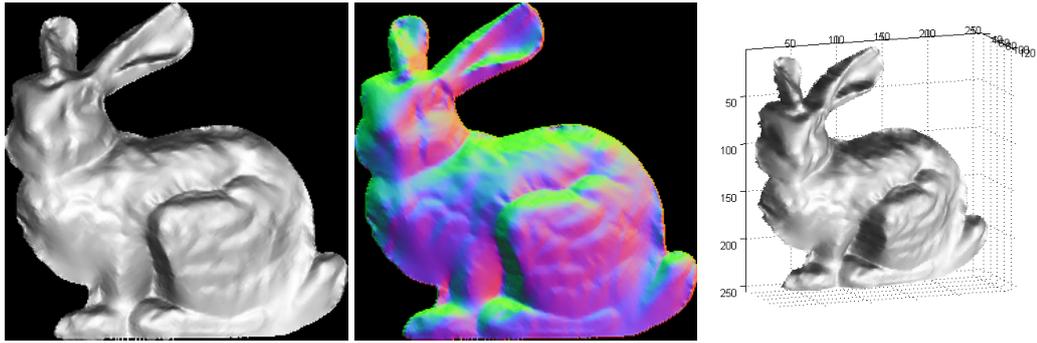
Figure 4.18: Shape estimated from input image in Fig. 4.17(a).

Figure 4.19: Environment map for image in Fig. 4.20.
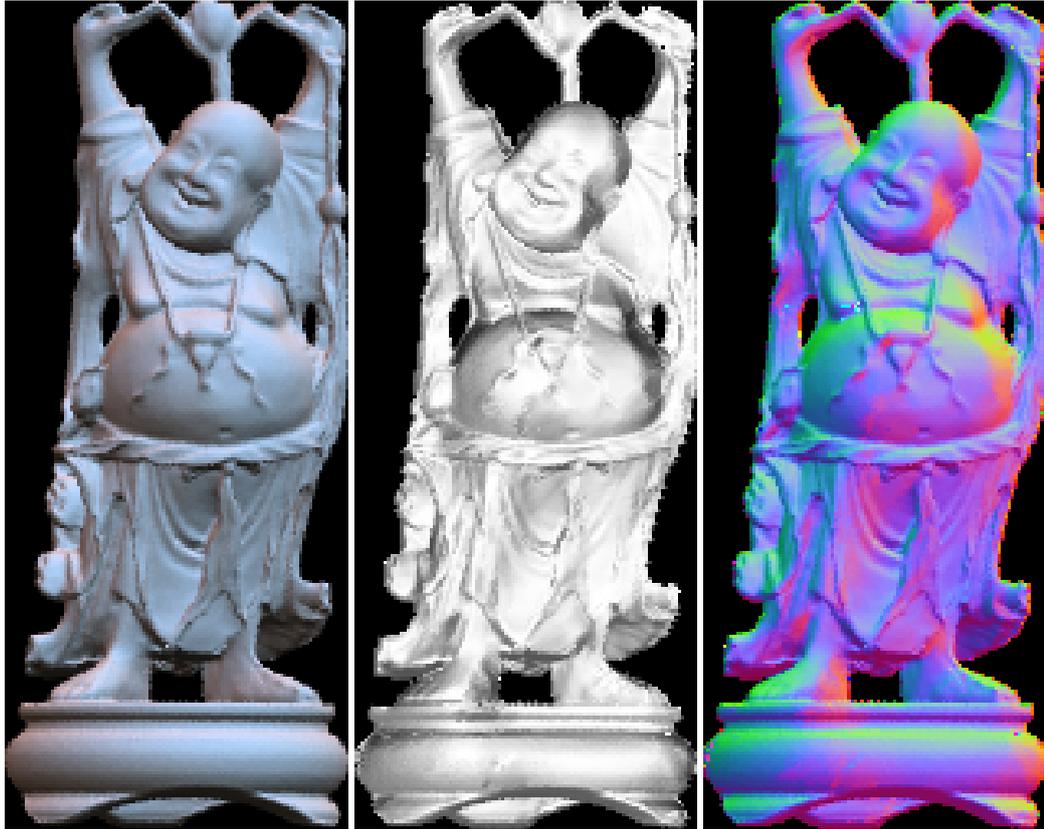
Figure 4.20: Shape estimate for a complex object.
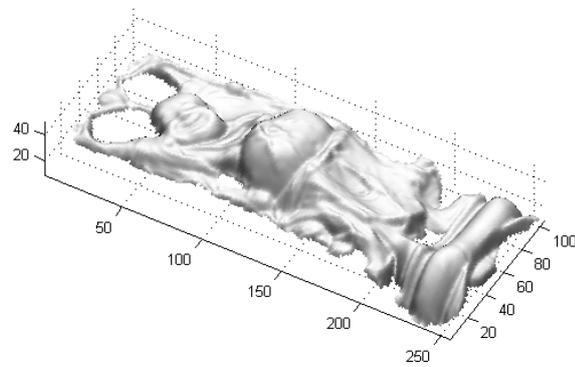
Figure 4.21: Integrated surface from the surface normals in Figure 4.20

Then, we also impose some cast shadow in a simple sine function with the same illumination map in Fig .4.15. The method fails at the area where the shadow dominant which shows in Fig .4.22
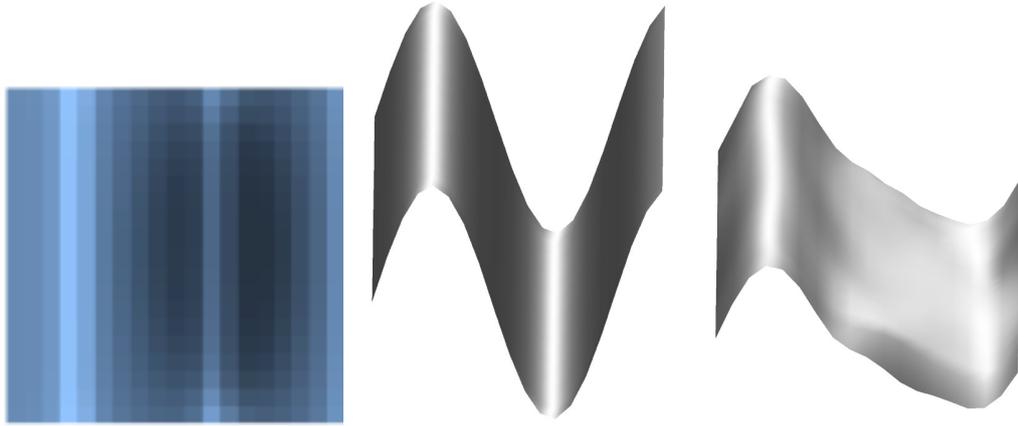
Figure 4.22: Shape estimate for object with cast shadow

4.4.3 Second Order Solution

Continue with the experiment in the linear order solution, we begin the results by showing the improvement of the second order solution with the first order solution. The complex illumination we used is in Figure4.23.

Figure 4.23: The Grace illumination environment.

Our compare tests begin with a simple object, a sphere. We also test it with the bunny and buddha we used for previous experiments. The improvement of the quadratic method over the linear method is not that obvious from the height map, because it improves at the points which have more high frequency parts, or details information. We compare it with the ground truth data. The improvement is about $3 - 5\%$ in average.
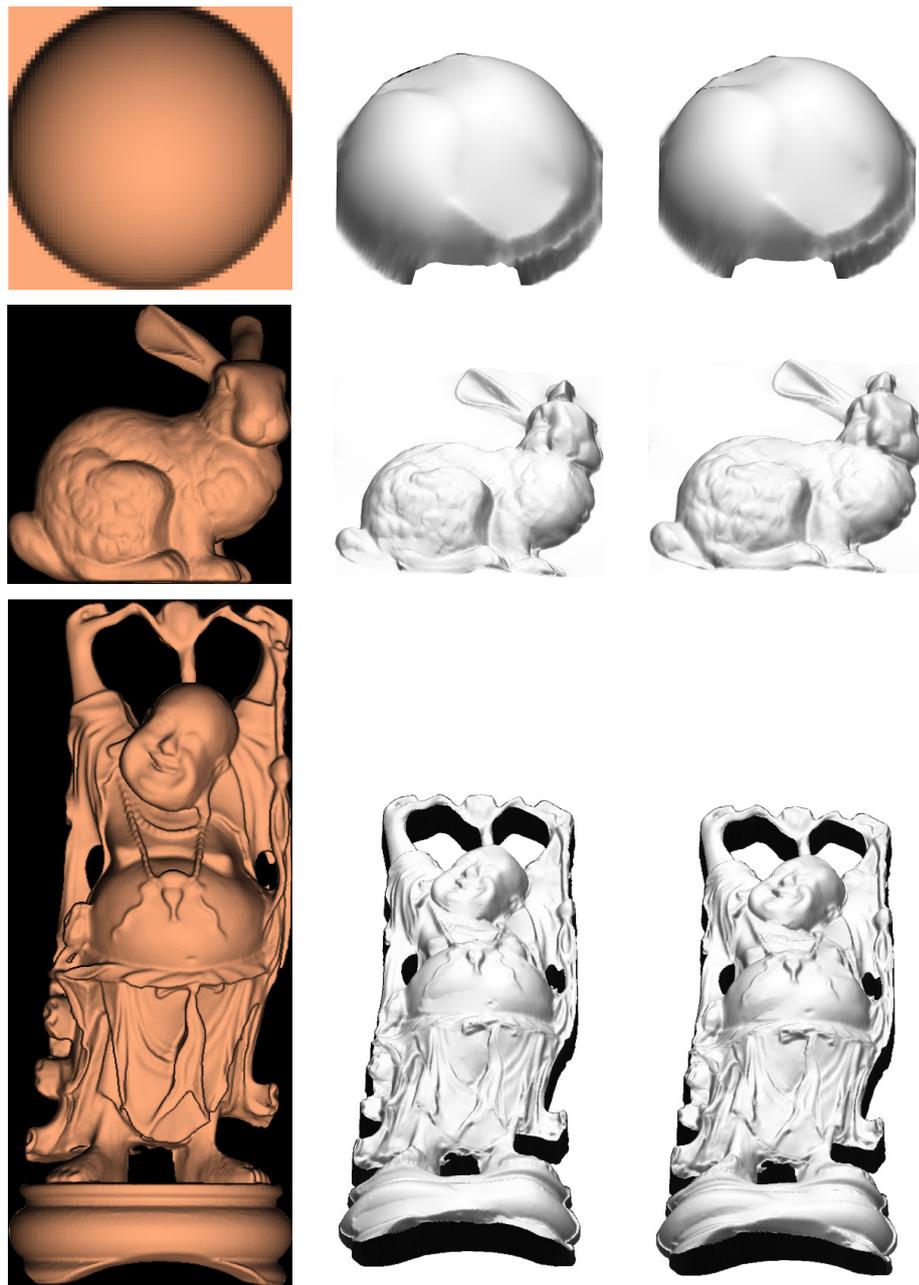
Image   Linear Method   Second Order

Figure 4.24: Surface recovered using both linear and second order methods

Table 4.1: Average Angular Error for synthetic images(in degree)

Objects	Linear	Second Order
Sphere	14.743	14.3125
Bunny	14.2036	13.7338
Buddha	12.4790	12.0149

We also show the recovered surface from the data set in MIT. [27] Figure4.27 are the recovered surfaces. The illumination function is measured by calibrating a sphere object under the same illumination condition. The object and the sphere have the same attribution for reflectance. We use two different sphere for illumination and they are listed in Figure 4.25, with corresponding recovered surface Figure 4.26.
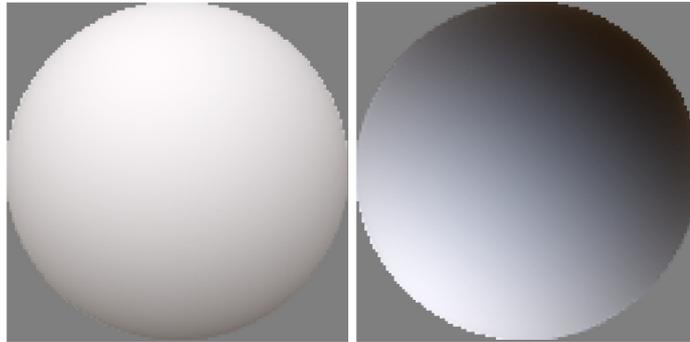
Figure 4.25: Sphere with illumination information

Image   Ground Truth Surface   Second Order
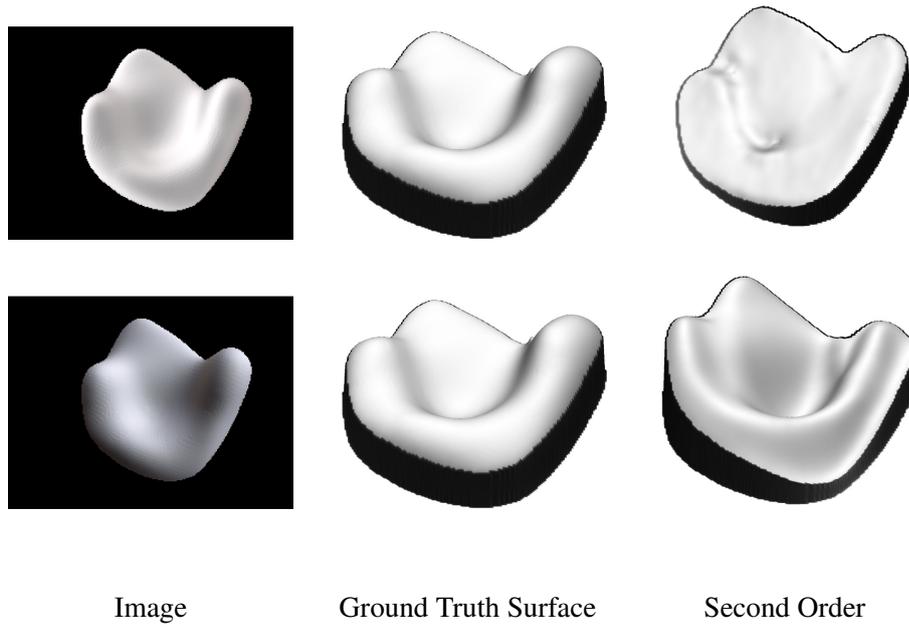
Figure 4.26: Surface recovered from different illumination conditions

Image   Ground Truth Surface   Second Order
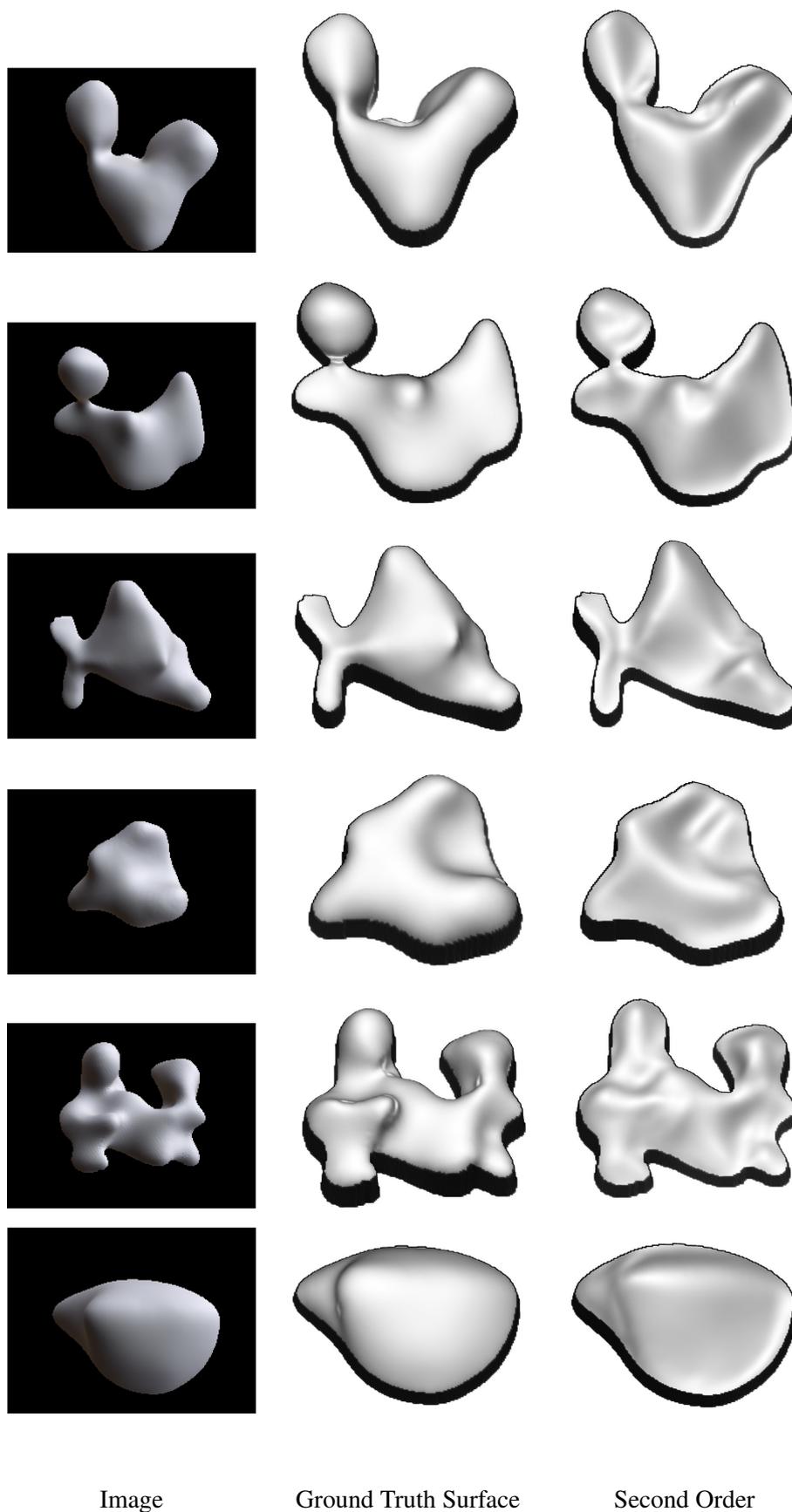
Figure 4.27: Surface recovered using second order methods

4.5 Conclusions

In section4.2.3, when an illumination environment can be accurately approximated by a spherical harmonic expansion, then photometric stereo under known environmental illumination can be effected by simply solving a system of linear equations. Shape recovery is even possible from a single image, though fewer coefficients can be recovered resulting in a loss of accuracy in the recovered shape. The quality of the recovered shape depends heavily on how well the spherical harmonic coefficients can approximate the environment map.

As this is the first work to consider photometric shape reconstruction under complex environmental illumination, there are many obvious limitations to our method. These provide several worthwhile avenues for future work. The first is to examine extending the method to more complex light source. Results in the graphics literature suggest that the spherical harmonic projection is still an efficient representation, though more coefficients are required to accurately represent highly specular reflectance. The second is to consider the effect of self occlusion, which may not only be a hindrance, but also potentially a cue to global geometry. Finally, we would like to explore whether any form of shape recovery is possible when the illumination environment is unknown.

In section 4.3.1, we have presented a shape-from-shading algorithm which is able to make shape estimates from objects illuminated by complex illumination. We believe we are the first to formulate the shape from shading problem as a minimization of a linear function under quadratic constraint and apply the direct mathematical solution to it. The results are encouraging and the method both time efficient and simple to implement. Finally, our local image formation model neglects occlusions which may form a useful cue to aid shape recovery [47].

In section 4.3.2, we extend the first order approximation to second order approximation. The extension from the lower order to higher order slightly increases the accuracy of the result. Our approach to optimising a quartic in three unknowns under a quadratic constraint is novel and may have applications outside shape from shading.

Although the complexity has increased by considering natural illumination, the complexity of shape from shading itself has actually been reduced. The classic ambiguities such as bas-relief and convex-concave ambiguity are avoided because the illumination environment provides direct information as to the set of possible surface orientations which give rise to a particular colour.

Chapter 5

Surface Recovery

5.1 Introduction

The goal of shape from shading is to reconstruct the 3-dimensional surface of an object from the image information. Existing methods can be divided into two categories depending on whether they output a "needle map" or height map.

(a)   (b)
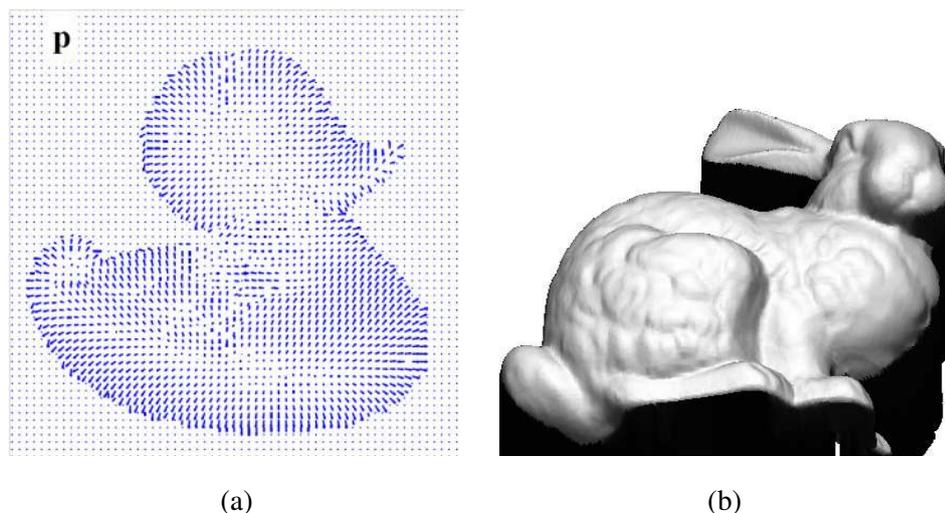
Figure 5.1: Two Possible results for shape from shading problem. (a) Needle map. (b) Height map

In most circumstances, the height map is the most desirable output as it is a true 3D representation of the surface. The integrability problem arises when the shape is represented as a field of surface normals or a "needle map". A field of surface normals are free to represent "impossible" surfaces, i.e. those that are not integrable. See Figure 5.2 for an example.

Integrability means that the surface obtained by integrating the surface normals is the same regardless of the path of integration taken. The gradient notation means that the surface is a scalar
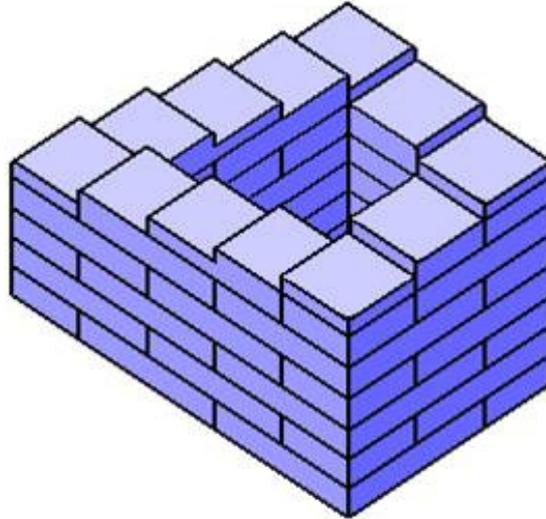
Figure 5.2: Impossible surfaces such as the constantly-descending staircase can be represented by a field of surface normals.

function and the second partial derivatives are independent of the order of differentiation, in other words they have zero curl. Shape from shading results from techniques based on the calculus of variations often fail to satisfy this constraint. Here, most results just incorporate a weighted term for the integrability constraint, but the iterative procedure cannot ensure that the constraint is strictly satisfied. For this reason, techniques for imposing the integrability constraint prove extremely useful for shape from shading.

Integrability constraint is the regularisation term used to narrow down the search space. Since

$$\delta u = p\delta x + q\delta y \tag{5.1}$$

Integration along a closed path C, the integral should be zero. The basic form of integrability is

$$\oint_C (p(x,y)dx + q(x,y)dy) = 0 \tag{5.2}$$

Commonly, a two-step approach is used whereby a field of surface normals is estimated from an image, after which a depth map is found that closely matches the estimated normals. This second step is known as surface integration and we describe some of the better known methods here.

Drawn from the idea that every continuous surface can be represented by a series of orthogonal basic functions with different frequency, Frankot and Chellappa [19] try to enforce the integrability on given gradient estimates by projecting the estimates onto the nearest integrable space of functions using the Fourier transform. The gradient of the reconstructed surface matches that of

the gradient map as closely as is possible with the basis functions. By choosing a different basis function (for example the discrete cosine transformation), a smooth integrable surface is generated. However, the nearest integrable function at some points may divert from the ground truth surface and it violates the Lambertian cone.

A Poisson solver algorithm was proposed by Simchony and Chellappa [52]. The least squared error estimate is used to solve the Poisson equation. It finds the solution that minimizes the norm of the correct gradient field. The solution is the gradient field that gives the global minimal error. But the result is still a non-integrable one because of the nature of least square error method. This work has recently been revisited by Chellappa and coworkers [1, 2].

A redundant non-orthogonal set of basis functions (shapelets) was introduced by Kovesi [31]. The relationship between surface and basis functions is equivalent to that of the gradients of the two. The continuity constraint is automatically performed together with the integration of the surface. Gaussian functions are chosen as the basis function. An interesting result of this approach is that it provides a scheme for integrating a surface from only the elevation angle alone.

In this chapter we consider two approaches that attempt to estimate surface height directly using intensity information. The ideas show promise but require refinement before they can be applied to real world problems.

The surface gradients can be approximated as linear functions of the surface height. In the simplest case by just taking single forward differences:

$$
\begin{cases}
p = z_{i+1,j} - z_{i,j} \\
q = z_{i,j+1} - z_{i,j}
\end{cases}
\tag{5.3}
$$

Since the surface normal is expressed by

$$
\begin{cases}
n_x = \dfrac{-p}{\sqrt{1+p^2+q^2}} \\
n_y = \dfrac{-q}{\sqrt{1+p^2+q^2}} \\
n_z = \dfrac{1}{\sqrt{1+p^2+q^2}}
\end{cases}
\tag{5.4}
$$

it is possible to relate image intensity directly to surface height. However, this increases the complexity of the problem substantially. Even under the simple point source Lambertian irradiance function, the image formation model becomes quadratic. Ecker and Jepson [14] proposed a Poisson solver to enumerate all the possible results and narrow down the result with additional constraints, although their method is extremely sensitive to noise.

86

5.2 Linear method for Direct Surface Recovery

The advantage of directly recovering surface height is obvious. Integrability is implicitly enforced via the problem formulation, and additional integration step is not required. Although solving the true surface height requires more information such as the distance from the object to the camera, and the calibration of the camera itself, we sidestep these issues by working in units of pixels. We present a very simple but potentially powerful approach for height from shading.

Recall the linear approximation of Lambertian irradiance function, and represent the surface normal in the form of surface gradients, we can have

$$i - i_0 = \frac{-\omega_x p - \omega_y q + \omega_z}{\sqrt{1 + p^2 + q^2}} \tag{5.5}$$

where i_0 is the 0th order contribution of illumination. Since we have the image in three channels, in fact, we have three similar equation. Note that the surface normal remains the same in three channels, we arrange the function

$$\frac{-\omega_x p - \omega_y q + \omega_z}{i - i_0} = \sqrt{1 + p^2 + q^2} \tag{5.6}$$

The right part is the same for all three functions and the value is a non-zero value. So we can set pairs of equations equal and factor out this part and get a system of two linear equations:

$$\begin{cases} i_1^r(-\omega_x^b p - \omega_y^b q + \omega_z^b) = i_1^b(-\omega_x^r p - \omega_y^r q + \omega_z^r) \\ i_1^r(-\omega_x^g p - \omega_y^g q + \omega_z^g) = i_1^g(-\omega_x^r p - \omega_y^r q + \omega_z^r) \end{cases} \tag{5.7}$$

where i_1^c is the intensity of first order in colour channel c.

Hence, every pixel gives rise to two linear equations which relate the surface gradients at a pixel to the observed intensity. Moreover, since the surface gradients can be approximated by linear functions of the surface height of the pixel and its neighbours, this leads to a large system of linear equations in the unknown surface heights.

Hence, we can write a sparse, overdetermined linear least squares optimisation of the form:

$$\mathbf{M} \begin{bmatrix} z_{0,1} \\ . \\ . \\ . \\ z_{i,j} \end{bmatrix} = \mathbf{b} \tag{5.8}$$

where $\mathbf{M}$ is the coefficient matrix. The least squares solution for $\mathbf{z}$ is given by the pseudo-inverse:

$$\mathbf{z} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{b} \tag{5.9}$$

However, since we are assuming an orthographic projection there is an unknown offset in the height (i.e. the surface can be shifted in the z direction without affecting the surface gradients). For this reason, we simply fix the bottom left pixel to zero height (the dimensionality of the unknown **z** reduces by one) and reformulate **M** accordingly. In practice, the dimensionality of **M** is very large (O(number of pixels $\times$ number of pixels)). However, since it is also very sparse it is perfectly feasible to solve the system of equations for large images using a sparse linear least squares solver.

As with the methods in the previous chapter, we assume that the illumination coefficients in the different colour channels are independent. As long as this assumption is met, this simple method provides encouraging results.

Again, we also assume that the illumination coefficients are known. However, it is also interesting to consider the potential for handling unknown illumination in the future. The matrix **M** is determined by the illumination coefficients and therefore has 12 degrees of freedom. Moreover, the illumination coefficients are constrained by the non-negativity of light (i.e. the spherical lighting function must be positive everywhere). Given an image, an estimate of the illumination coefficients yields a surface. This suggests a possible approach to joint illumination and surface estimation. The idea would be to define an objective for the surface (e.g. maximal smoothness or minimal depth range) and then optimise the illumination parameters to find a surface which optimises this objective.

5.3 Semidefinite Programming

Semidefinite programming (SDP) is a subfield of convex optimisation concerned with the optimisation of a linear objective function over the intersection of the cone of positive semidefinite matrices. The cone of positive semidefinite matrices can be seen as a nonlinear and non-smooth, but convex constraint. Semidefinite programming unifies several standard problems and finds many application in engineering and combinatorial optimisation.

5.3.1 Semidefinite programming

We consider the problem of minimising a linear function of a variable $\vec{x} \in \mathbf{R^m}$ subject to a matrix inequality:

$$
\begin{aligned}
&minimize && \vec{c}^T \vec{x} \\
&subject\ to && F(\vec{x}) \geqslant 0,
\end{aligned}
\tag{5.10}
$$

where

$$F(\vec{x}) \triangleq F_0 + \sum_{i=1}^{m} \vec{x}_i F_i$$

The problem data are the vector $\vec{c} \in \mathbf{R^m}$ and $m + 1$ symmetric matrices $F_0, \cdots, F_m \in \mathbf{R^{n \times n}}$. The inequality sign in $F(\vec{x}) \geq 0$ means that $F(\vec{x})$ is positive semidefinite, i.e., $\vec{z}^T F(\vec{x})\vec{z} \geq 0$ for all $\vec{z} \in \mathbf{R^n}$.

For a quadratically constrained quadratic programming problem, a convex quadratic constraint $(A\vec{x} + \vec{b})^T (A\vec{x} + \vec{b}) - \vec{c}^T \vec{x} - d \leq 0$, with $\vec{x} \in R^k$, can be written as

$$\begin{bmatrix} I & A\vec{x} + \vec{b} \\ (A\vec{x} + \vec{b})^T & \vec{c}^T \vec{x} + d \end{bmatrix} \geq 0 \tag{5.11}$$

The left-hand side depends affinity on the vector x, it can be expressed as

$$F(\vec{x}) = F_0 + \vec{x}_1 F_1 + \cdots + \vec{x}_k F_k \geq 0 \tag{5.12}$$

with

$$F_0 = \begin{bmatrix} I & \vec{b} \\ \vec{b}^T & d \end{bmatrix}, \quad F_i = \begin{bmatrix} 0 & \vec{a}_i \\ \vec{a}_i^T & \vec{c}_i \end{bmatrix}, i = 1, \cdots, k \tag{5.13}$$

where $A = [a_1 \cdots a_k]$. Therefore, a general quadratically constrained quadratic program (QCQP)

$$\begin{aligned} minimize & \quad f_0(\vec{x}) \\ subject\ to & \quad f_i(\vec{x}) \leq 0, i = 1, \cdots, L \end{aligned} \tag{5.14}$$

where each f_i is a convex quadratic function $f_i(\vec{x}) = (A_i\vec{x} + \vec{b})^T (A_i\vec{x} + \vec{b}) - \vec{c}_i^T \vec{x} - d_i$, can be written as

$$\begin{aligned} minimize & \quad f_0(\vec{x}) \\ subject\ to & \quad \begin{bmatrix} I & A_0\vec{x} + \vec{b}_0 \\ (A_0\vec{x} + \vec{b}_0)^T & \vec{c}_0^T \vec{x} + d_0 \end{bmatrix} \geq 0, \\ & \quad \begin{bmatrix} I & A_i\vec{x} + \vec{b}_i \\ (A_i\vec{x} + \vec{b}_i)^T & \vec{c}_i^T \vec{x} + d_i \end{bmatrix} \geq 0, i = 1, \cdots, L \end{aligned} \tag{5.15}$$

which is a semidefinite program with variables $\vec{x} \in \mathbf{R^k}$ and $t \in \mathbf{R}$.

5.3.2 SDP in shape from shading

We introduce the SDP framework for shape from shading as a first attempt to account for self-occlusion in the image formation process. However, we are only able to do so in the form of an inequality constraint. This is too weak to be used for shape from shading on its own. In practice,

we use it to constrain the surface integration process. In essence, providing an occlusion-sensitive surface integration approach.

Recall the full irradiance function with the visibility function:

$$I(x) = \rho(x) \int_{\Omega(\mathbf{n}(x))} L(\vec{\omega})V(x,\vec{\omega})(\mathbf{n}(x) \cdot \vec{\omega})d\vec{\omega}, \tag{5.16}$$

For any points that are partially occluded (i.e. some portion of the illumination environment over the upper hemisphere is blocked by another part of the surface), the intensity measured by the imaging device will be less than the predicted intensity with no occlusion.

Hence, using a model without occlusions, i.e. where the visibility function $V(x,\vec{\omega})$ is one for every direction of the semi-sphere, then

$$I(x) = \rho(x) \int_{\Omega(\mathbf{n}(x))} L(\vec{\omega})(\mathbf{n}(x) \cdot \vec{\omega})d\vec{\omega}, \tag{5.17}$$

We assume the unit know albedo here, and replace the surface normal with the gradient form, then we have

$$I(x) \leq \int_{\Omega} \frac{L(\vec{\omega})(-p(x)\omega_x - q(x)\omega_y + \omega_z)}{\sqrt{1 + p(x)^2 + q(x)^2}} d\vec{\omega}, \tag{5.18}$$

Take the discrete form of this equation and square both parts of the inequality sign, since both parts are positive values, we have

$$I(x)^2 \leq \frac{\sum_j^n L_x^j(-p(x))\omega_x + \sum_j^n L_y^j(-q(x))\omega_y + \sum_j^n L_z^j\omega_z}{1 + p(x)^2 + q(x)^2} \tag{5.19}$$

The $1 + p(x)^2 + q(x)^2$ is a positive value, so we can move it to the left part without changing the direction of the inequality sign. The summation is over a discretised environment map. The whole expression looks like Ecker and Jepson's [14] way of constructing the polynomial function, but the concept is different. The first difference is that we are using full environment illumination rather than the single light source. The second difference is that we take the self-occlusion into account.

The summation over the environment map must be restricted to the upper hemisphere (i.e. light source directions making an angle of greater than 90 degrees with the normal must be excluded). We achieve this by using an initial estimate of the surface normal (provided by the algorithms in Chapter 4) to determine which light source directions should be excluded.

The surface gradients are then replaced by the surface heights as in the previous section. So, for a single pixel with three different colour channels, we have three quadratic inequality constraints. Rather than solving the surface height on single pixel basis, we solve it on a small grid. The relationship of neighbouring pixels can increase the strength of the constraint. The reason we do
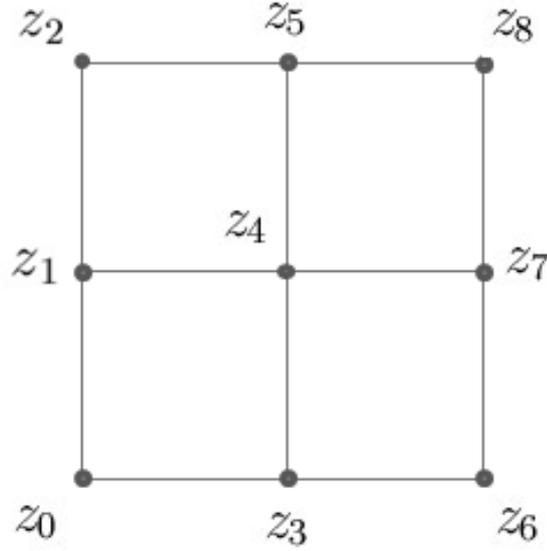
Figure 5.3: 3×3 grid of surface height with 4 pixels

not use a bigger grid is due to the limit of the SDP solver. For a matrix with large rank in quadratic constraint, the SDP solver has very poor performance. So, for an arbitary grid as shown in Fig.5.3 the gird consists of nine height values from z_0 to z_8, with the corresponding 4 pixels. However, the gradient expression can only represent the relationship of eight height values. The top right corner is excluded from our constraint. The bottom left height value is fixed to 0 as a reference height. For a 3×3 grid, we have 12 inequality quadratic constraint with 7 unknowns. Each quadratic inequality constraint has the similar form as

$$a \cdot z_{x,y}^2 + b \cdot z_{x,y+1}^2 + c \cdot z_{x+1,y}^2 + d \cdot z_{x,y} z_{x,y+1} + e \cdot z_{x,y} z_{x+1,y} + f \cdot z_{x+1,y} z_{x,y+1} + g \cdot z_{x,y} + l \cdot z_{x+1,y} + m \cdot z_{x,y+1} + n \leq 0$$
(5.20)

which is a standard quadratic constraint in semidefinite programming.

To write the equation within the form of semidefinite programming problem, we construct the Matrix Q_i with the quadratic term, Q_i must be a symmetric matrix. We also construct the vector $\vec{c}$ for the coefficients of the linear part. So the quadratic constraint functions are

$$\vec{z}^T Q_i \vec{z} + \vec{c}_i^T \vec{z} + d_i \leq 0$$
(5.21)

The intensity of pixel should also be positive, since $1 + p^2 + q^2$ is always bigger than 1, we have a group of linear inequality constraint as well. The objective function we used here is intended to

91

minimize the difference with initial surface. The whole SPD is described as follows

$$
\begin{aligned}
min \quad & \vec{z}^T Q_0 \vec{z} + \vec{c}_0^T \vec{z} + d_0 \\
& \vec{z}^T Q_i \vec{z} + \vec{c}_i^T \vec{z} + d_i \leq 0 \\
s.t. \quad & A\vec{z} + \vec{b} \geq \vec{0}
\end{aligned}
$$

Solving this quadratic constraint quadratic function needs the introducing of auxiliary variable t. Since the matrix Q_i is semidefinite, the matrix can be decomposed to $L_i^t L_i = Q_i$ using Cholesky decomposition method. This will transfer the problem to

$$
\begin{aligned}
min \quad\quad\quad & t \\
& \vec{z}^T L_0^T L_0 \vec{z} + \vec{c}_0^T \vec{z} + d_0 \leq t \\
s.t. \quad & \vec{z}^T L_i^T L_i \vec{z} + \vec{c}_i^T \vec{z} + d_i \leq 0 \\
& A\vec{z} + \vec{b} \geq \vec{0}
\end{aligned}
\tag{5.22}
$$

move the linear term to the right side of the inequality sign, finally, the SDP becomes a form as follows

$$
\begin{aligned}
min \quad\quad\quad\quad & t \\
& \left\| \begin{pmatrix} t - \vec{c}_0^T \vec{z} - 1 \\ 2L_0 \vec{z} \end{pmatrix} \right\| \leq t - \vec{c}_0^T \vec{z} + 1 \\
s.t. \quad & \left\| \begin{pmatrix} -d_i - \vec{c}_i^T \vec{z} - 1 \\ 2L_i \vec{z} \end{pmatrix} \right\| \leq d_i - \vec{c}_i^T \vec{z} + 1 \\
& A\vec{z} + \vec{b} \geq \vec{0}
\end{aligned}
\tag{5.23}
$$

This is the standard way of solving SPD. Notice that additional linear constraints can be added without increasing the complexity of the system. During the experiment, we just use the above 2 constraints. The whole algorithms is as follows

1. Obtain the initial surface from linear model described in previous chapter

2. Check the quadratic term and discard it if not semidefinite matrix

3. Apply the semidefinite solver

5.4   Experiment Result

We now present experimental results that demonstrate in principle that the two methods described can be used for shape estimation.

5.4.1 Linear Method

We begin by showing the performance of the linear method for direct surface estimation under ideal conditions. In this case, images are rendered using a linear illumination function, i.e. the model can describe the image exactly. The results are shown in Figure 5.4, where it can be clearly seen that the method recovers excellent surface estimates. Any deviation from ground truth is due to the numerical error in the approximation of surface gradients using finite differences.

We now repeat the same experiment but with quadratic illumination, i.e. there is a systematic error between the model approximation and the observed image. Results are shown in Figure 5.5. The reconstructions are remarkably stable given the inaccuracy of the illumination approximation. The reason for this is the strictness of the constraint afforded by solving for surface height directly. Only integrable surfaces are possible, which severely restricts the solution space.

In fact, the method is robust to Gaussian noise with a variance of up to 0.49 units of intensity. This is shown in Figure 5.6, where it can be seen that the surface height error is close to zero for all noise variances up to about 0.49.

Finally, we show an example on another form of systematic noise caused by occlusions. In Figure 5.7 we show a surface composed from a sine function. We render the surface with complex illumination and taking account of occlusions. This means that the valley is darker than would be predicted by the linear image formation model. As can be seen, the reconstruction fails in the occluded regions of the surface. This motivates the attempt to incorporate occlusion constraints in the SDP method.
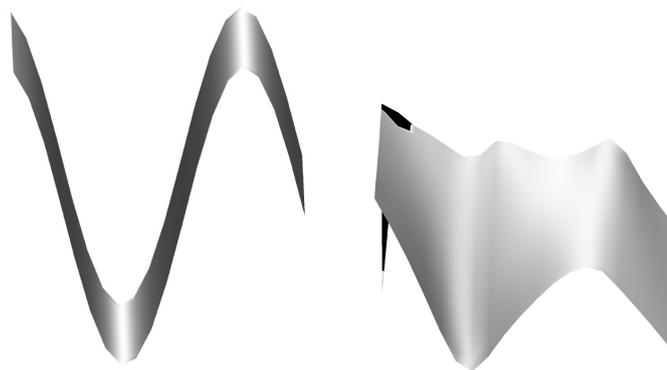
Figure 5.7: Cast Shadow Introduced for Linear Search

5.4.2 Semidefinite Programming Method

Finally, we show surface recovery results using the semidefinite programming method. In Fig.5.8 shows a surface composed from a sine function that is rendered with occlusions. The semidefinite programming method takes a gradient field as input, which forms a part of the objective function. The first row of Figure 5.9 shows performance when the SDP solver is given ground truth gradients as input. In this case, the solver is able to return the ground truth surface almost exactly as it satisfies all the inequality constraints. The second row shows an input from the order 1 algorithm of the previous chapter. Here, the normals are highly erroneous around the shadowed region. The SDP solution is able to reduce the surface height error by about 5%, although the effect is visually subtle.
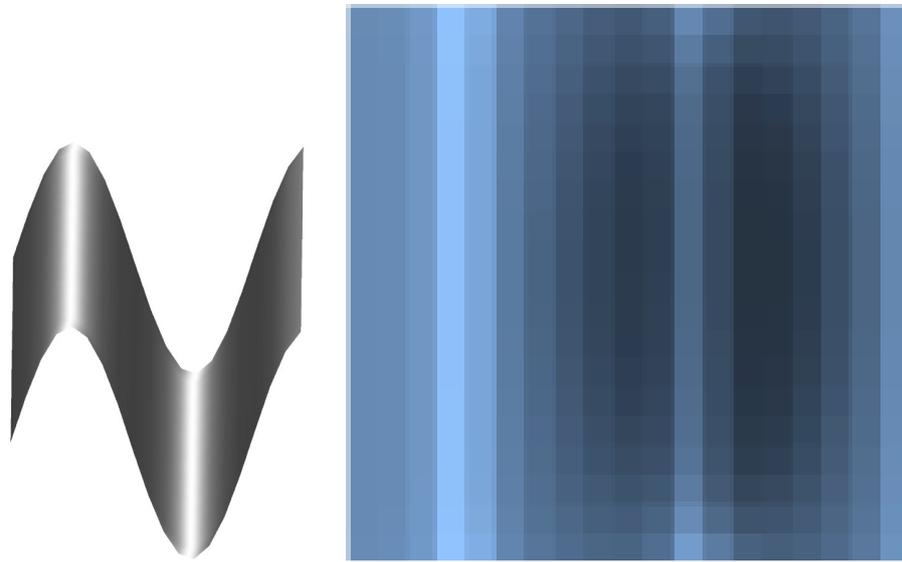
Figure 5.8: A non-convex surface (left) and its rendering with environment illumination and occlusions (right).
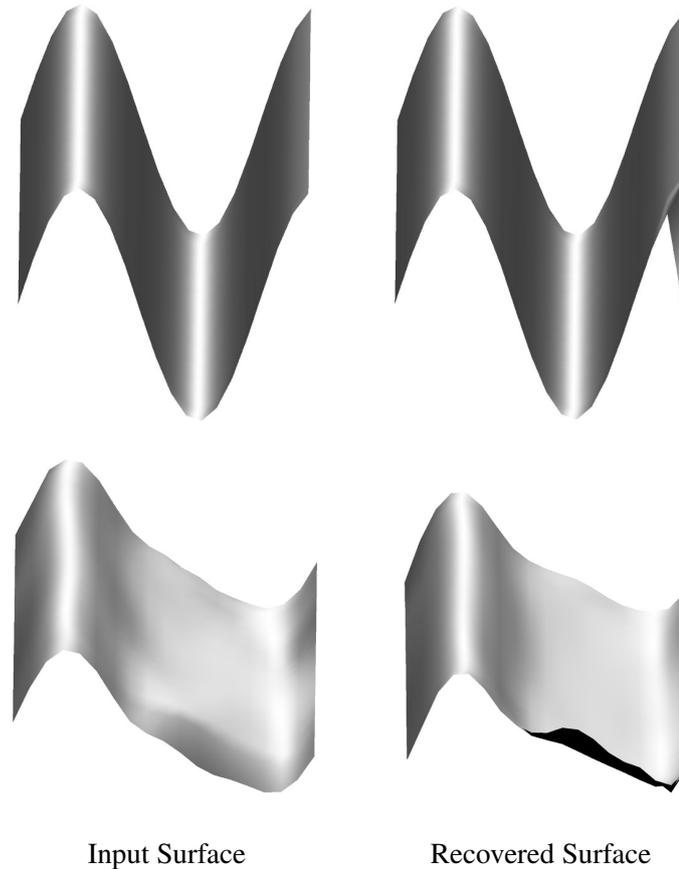
Figure 5.9: Left: the surfaces which are used to provide input gradient estimates for the SDP reconstruction. Right: the surfaces integrated from gradient estimates on the left with the occlusion sensitive inequality constraint.

5.5 Conclusions

In this chapter we have proposed two methods that represent a step toward direct surface height recovery in the context of shape from shading. This affords additional constraint by restricting the solution space to integrable surfaces only. The first approach is based on ratios between colour channels and requires only the solution of a system of linear equations. This method shows great promise although it fails to account for occlusions. The second approach attempts to account for occlusions in a semidefinite programming framework. Although we are able to construct meaningful constraints in the form of quadratic inequalities (and the resulting system is convex), in practice the constraint is too weak to allow surface recovery alone. Even when used as additional constraint in the surface integration process the improvement is extremely small. Moreover, existing solvers struggle to solve all but the most trivial of systems.

| Image | Ground Truth Surface | Linear Method |

Figure 5.4: Results of direct linear surface recovery for images rendered using a linear illumination function.

| | Image | Ground Truth Surface | Linear Method |

Figure 5.5: Results of direct linear surface recovery for images rendered using a linear illumination function.
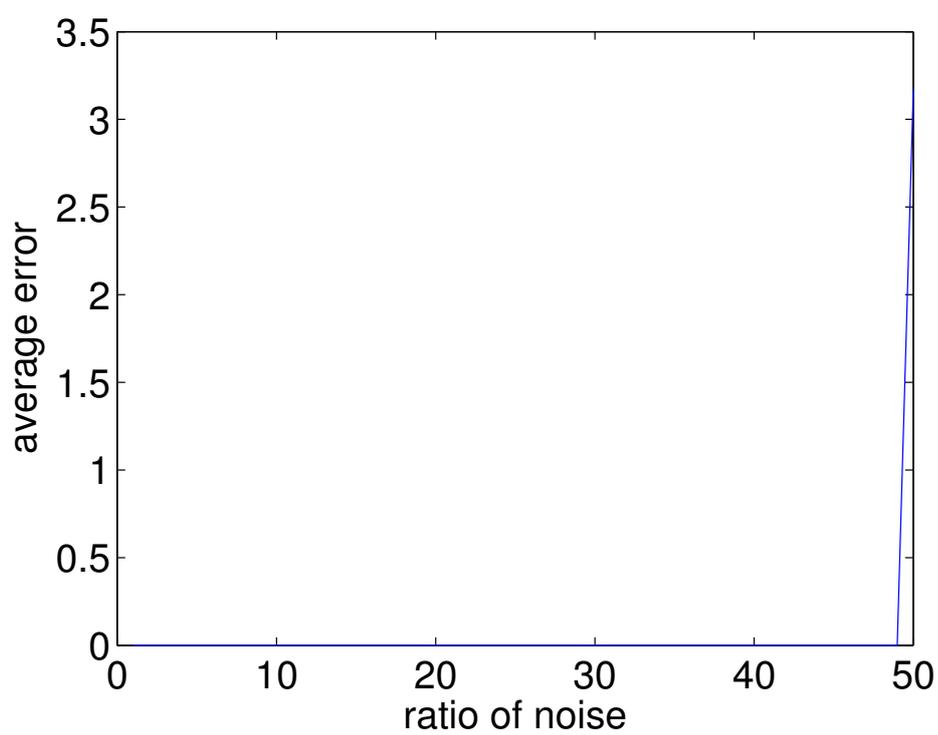
Figure 5.6: Average Z error as a function of the percentage of unit-variance Gaussian noise added to the image.

Chapter 6

Conclusions and Future Work

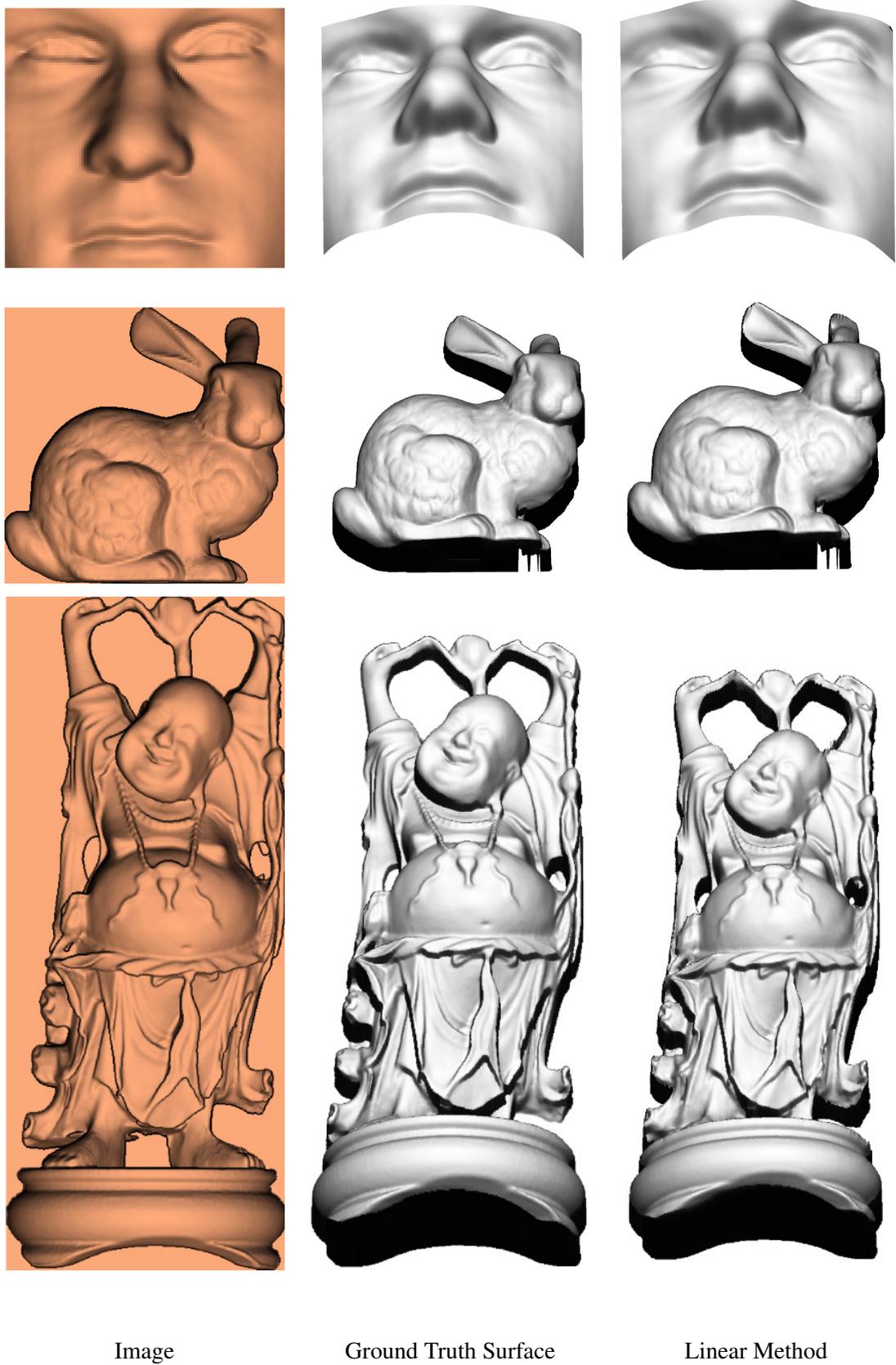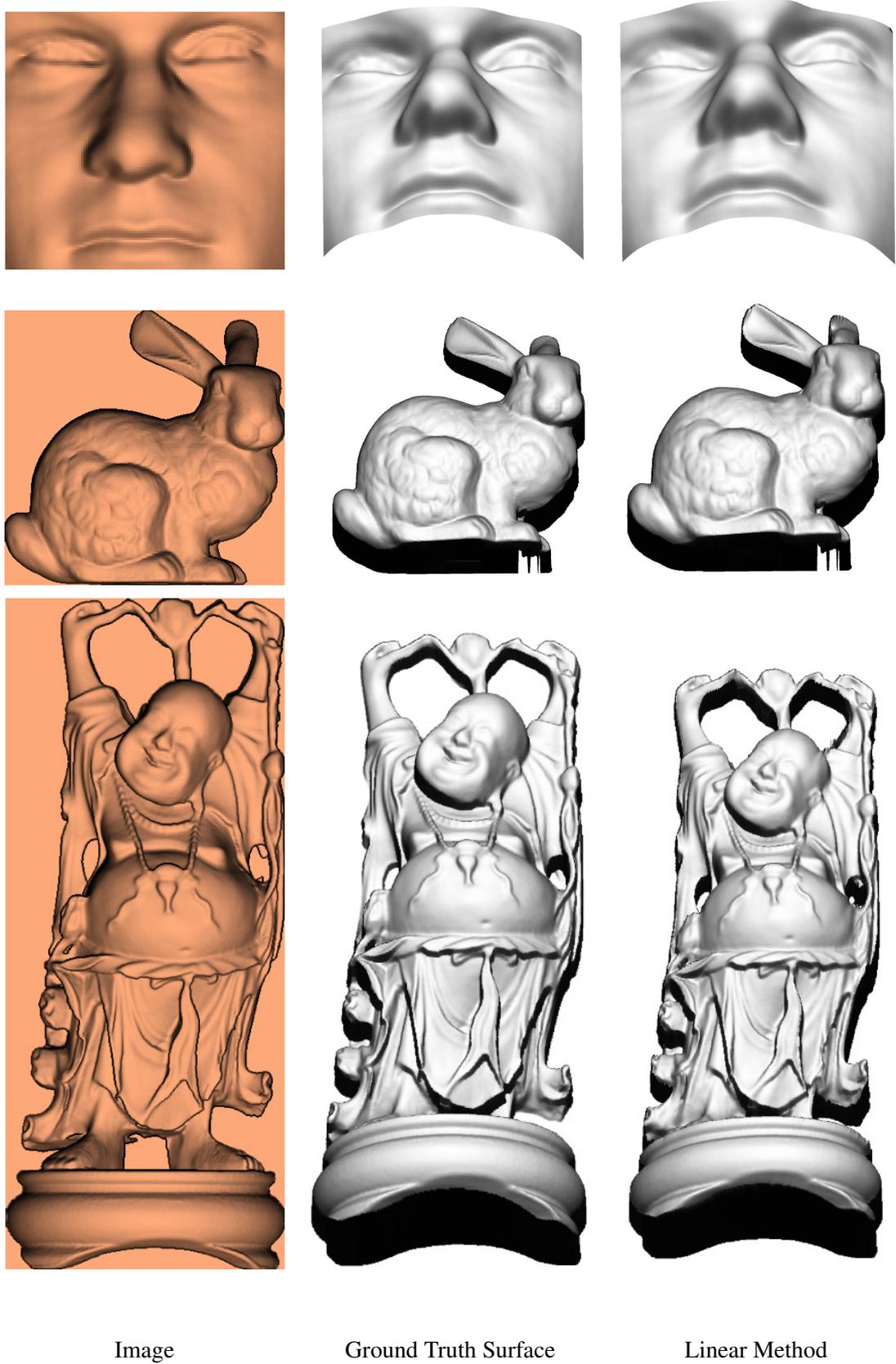In recent years, the shape from shading problem has again begun to attract attention. Since it was first proposed by Brooks and Horns nearly 30 years ago, a great effort has been made to provide reliable methods for real application. The problem itself has also developed from a single light source assumption, which is quite rare in the real world, to complex natural illumination. Many creative ideas has been introduced to this area. However, the problem is still unsolved in a general setting. Developing a practical method for real situations is still some way off. In this section we summarise the contributions made in this thesis, highlight some weaknesses of the work presented, and end by discussing directions for future work.

6.1   Summary of Contributions

In this thesis, we have proposed a number of novel methods for shape from shading, which range from conventional (surface normal estimation under point source illumination) to entirely new formulations (direct surface height estimation from images under complex illumination). The contributions made in this work are summarised in the following list:

1. We have shown a new smoothness constraint and iterative framework that improves the result of variational methods in Chapter 3. The constraint and framework together help more points to reach a global minima rather than the conventional methods.

2. We have presented a method for surface normal estimation under environment illumination in Chapter 4. The first order solution is highly efficient and simple, requiring only a small SVD decomposition and matrix multiplications to estimate the normal at each pixel. Our second order solution is based on a novel method for solving a particular form of quadrat-

ically constrained nonlinear optimisation. Again, the approach is efficient and obtains the correct global optimum on synthetic data.

3. We have shown two methods for direct surface height recovery in Chapter 5. The first is a shape-from-shading method that exploits ratios between colour channels to cancel for nonlinear normalisation terms. This leads to a remarkably simple linear algorithm. The second is a surface integration method that accounts for occlusions in its image formation model. This leads to a convex optimisation that can, in principle, be solved by semidefinite programming. In practice, present-day solvers struggle to obtain a solution for all but the most trivial examples.

6.2 Critical Analysis

There are some criticisms that may be levelled at the work proposed in this thesis. We discuss these weaknesses and then propose some possible directions for future work.

1. **Parameter Selection and Initial Estimation:** In Chapter 3, we showed that the accuracy of the recovered surface greatly depends on the result of the initialisation, or, more precisely, whether the concave-convex ambiguity is globally resolved correctly. Any errors in the initialisation that incorrectly resolve these ambiguities are unlikely to be fixed by the iterative refinement. In addition, the parameter controlling the behaviour of the smoothness constraint affects the final result. The optimum setting depends on the object under study.

2. **Noise Sensitivity:** In Chapter 4, we showed that although our first and second order algorithm work very well with clean, synthetic data, performance falls dramatically when noise is introduced. Even the smoothness constraint can only mitigate this effect to a certain extent. The reason for this is that small changes in pixel colour can lead to completely different global minima under environment illumination. In contrast, in classical point source shape from shading, small changes in intensity require only a small change in normal direction to minimise the irradiance error. In addition, when cast shadows are introduced, the result is degraded even further.

3. **Inefficient Solver and Constraint in SDP:** In Chapter 5, we showed a framework that could solve for surface height and obtain a global minimum for any set of constraints that could be written in linear form. We were able to incorporate occlusions into the image

formation model. However, the SDP solvers currently available are both slow and fail on large-scale problems.

6.3 Direction of Future Research

We conclude the thesis by providing directions for possible future research:

1. Concave-Convex Resolution One simple but effective way to improve our variational method is to provide a more accurate initialisation with respect to the resolution of concave-convex ambiguities. Although this is an ambiguity intrinsic to the shape from shading problem for the single light source case, some pragmatic methods could be applicable here. One approach would be to use recent advances in content-based retrieval to search a database to find similar objects, e.g. "shape by recognition". Alternatively, human interaction could exploit the ability of the human visual system to use top down processing resolve low level ambiguities.

2. Image Based Parameter Selection Correct selection of the parameter used in our variational method to balance the weight between intensity and smoothness constraints is critical. In order to make the method more widely applicable, it would be interesting to see whether it is possible to learn the optimal parameter setting from image information directly. The idea would be to apply shape from shading to a range of images of objects with known shape and find the optimal parameter setting for each object. Machine learning could then be used to related image statistics or bags of image features to the parameter value.

3. Cast Shadow Compatible Photometric Methods The model we used for photometric shape from shading under spherical harmonic model only works with objects that do not self-occlude, i.e. where intensity is a function of surface normal alone, not global geometry. However, most realworld objects do self-occlude. Moreover, there is information about the surface shape contained in the occlusion and shadowing information (e.g. Prados et al. [47] recently proposed a method to estimate shape using occlusion shading alone). An obvious avenue for future work would be to try and combine these two approaches so that occlusion information is explicitly accounted for and exploited in the shape recovery process.

4. Full Set of Constraints for Semidefinite Programming The semidefinite programming framework we propose has much promise for future work. However, the constraint we propose is too weak to be used alone for shape-from-shading. In effect, it just reduces the search space

for the surface integration problem. Future work could attempt to formulate more useful constraints within this framework (perhaps exploiting ratio images as in our linear method) to extend the method to a true shape from shading algorithm.

Appendix A

Real Form Spherical Harmonics

Real Spherical Harmonics with $l = 0$

$$Y_0^0 = \frac{1}{2}\sqrt{\frac{1}{\pi}}$$

Real Spherical Harmonics with $l = 1$

$$p_x = \sqrt{\frac{1}{2}}\left(Y_1^{-1} - Y_1^1\right) = \sqrt{\frac{3}{4\pi}} \cdot \frac{x}{r}$$

$$p_y = i\sqrt{\frac{1}{2}}\left(Y_1^{-1} + Y_1^1\right) = \sqrt{\frac{3}{4\pi}} \cdot \frac{y}{r}$$

$$p_z = Y_1^0 = \sqrt{\frac{3}{4\pi}} \cdot \frac{z}{r}$$

Real Spherical Harmonics with $l = 2$

$$d_{z^2} = Y_2^0 = \frac{1}{4}\sqrt{\frac{5}{\pi}} \cdot \frac{-x^2 - y^2 + 2z^2}{r^2}$$

$$d_{yz} = i\sqrt{\frac{1}{2}}\left(Y_2^{-1} + Y_2^1\right) = \frac{1}{2}\sqrt{\frac{15}{\pi}} \cdot \frac{yz}{r^2}$$

$$d_{xz} = \sqrt{\frac{1}{2}}\left(Y_2^{-1} - Y_2^1\right) = \frac{1}{2}\sqrt{\frac{15}{\pi}} \cdot \frac{zx}{r^2}$$

$$d_{xy} = i\sqrt{\frac{1}{2}}\left(Y_2^{-2} - Y_2^2\right) = \frac{1}{2}\sqrt{\frac{15}{\pi}} \cdot \frac{xy}{r^2}$$

$$d_{x^2-y^2} = \sqrt{\frac{1}{2}}\left(Y_2^{-2} + Y_2^2\right) = \frac{1}{4}\sqrt{\frac{15}{\pi}} \cdot \frac{x^2 - y^2}{r^2}$$

Glossary

Bidirectional Reflectance Distribution Function a four-dimensional function that defines how light is reflected at an opaque surface

Image irradiance equation a function that relates surface orientation to measured image irradiance

Phong model a phenomenological reflectance model for describing specular reflectance from shiny surfaces

Torrance and Sparrow model a microfacet-based reflectance model for describing specular reflectance from rough surfaces

Lambertian reflectance perfectly diffuse reflectance in which light is scattered equally in all directions

Albedo the ratio of reflected radiation from the surface to incident radiation upon it

Self-shadow shadowing that occurs when the surface is oriented more than $90°$ from the light source

Cast shadow shadowing that occurs when another part of the surface blocks light from reaching a point

Intrinsic image	an image which captures only intrinsic properties of an object or scene, extrinsic properties such as illumination and shadowing are removed
Point light source	a single identifiable localised source of light if the object is at a very great distance or the resolution of the lighting is too low to resolve its size
Convex relaxation	an approximation to a non-convex optimisation problem by a convex one
Homotopy solver	a tool for solving polynomial systems using the continuation method
Shape from shading	shape estimation using the shading information from a single, monocular image
Photometric stereo	shape estimation using the shading information in multiple images taken from a fixed viewpoint under varying illumination
Environment illumination	an approximation to realworld illumination conditions in which lighting is described by a spherical function representing the incident radiance from every direction

References

[1] Amit Agrawal, Rama Chellappa, and Ramesh Raskar. An algebraic approach to surface reconstruction from gradient fields. *International Conference on Computer Vision*, 1:174–181, 2005.

[2] Amit Agrawal, Ramesh Raskar, and Rama Chellappa. What is the range of surface reconstructions from gradient field. *European Conference on Computer Vision*, 3951:578–591, 2006.

[3] Harry J. Barrow and J. Martin Tenenbaum. Recovering intrinsic scene characteristics from images. *Computer Vision Systems*, 3:26, 1978.

[4] Ronen Basri and David Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:218–233, 2003.

[5] Ronen Basri and David Jacobs. Photometric stereo with general unknown lighting. *International Journal of Computer Vision*, 72:239–257, 2007.

[6] Peter N. Belhumeur, David J. Kriegman, and Alan L. Yuille. The bas–relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999.

[7] Michael Bichsel and Alex P. Pentland. A simple algorithm for shape from shading. In *Computer Vision and Pattern Recognition*, pages 459–465, 1992.

[8] Soma Biswas, Gaurav Aggarwal, and Rama Chellappa. Robust estimation of albedo for illumination-invariant matching and shape recovery. *IEEE Transaction on Pattern Analysis and Machine Intelligence.*, 31:884–899, 2009.

[9] Andrew Blake. Boundary conditions for lightness computation in mondrian world. *Computer Vision, Graphics and Image Processing*, 32:314–327, 1985.

[10] Anna R. Bruss. The eikonal equation: Some results applicable to computer vision. *Journal of Mathematical Physics*, 23(5):890–896, 1982.

[11] Wojciech Chojnacki and Michael J. Brooks. Revisiting pentland's estimator of light source direction. *Journal of Optical Society of America*, 11:118–124, 1994.

[12] Paul Dupuis and John Oliensis. An optimal control formulation and related numerical methods for a problem in shape reconstruction. *The Annals of Applied Probability*, 4(2):287–346, 1994.

[13] Jean-Denis Duroua, Maurizio Falcone, and Manuela Sagona. Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding*, 109(1):22–43, 2008.

[14] Ady Ecker and Allan D. Jepson. Polynomial shape from shading. In *Computer Vision and Pattern Recognition*, pages 145 –152, 2010.

[15] Alan Edelman and H. Murakami. Polynomial roots from companion matrix eigenvalues. *Mathematics Computation*, 64:763–776, 1995.

[16] Graham D. Finlayson, Mark S. Drew, and Cheng Lu. Intrinsic images by entropy minimization. *European Conference on Computer Vision*, pages 582–595, 2004.

[17] David Forsyth. Variable-source shading analysis. *International Journal of Computer Vision*, 91:280–302, 2011.

[18] David Forsyth and Andrew Zisserman. Reflections on shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):671 –679, 1991.

[19] Robert T. Frankot and Rama Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:439–451, 1988.

[20] Darya Frolova, Denis Simakov, and Ronen Basri. Accuracy of spherical harmonic approximations for images of lambertian objects under far and near lighting. *European Conference on Computer Vision*, pages 574–587, 2004.

[21] Robin Green. Spherical harmonic lighting: The gritty details. *Proceedings of the Game Developers Conference*, pages 1–47, 2003.

[22] Tom S. F. Haines and Richard C. Wilson. Belief propagation with directional statistics for solving the shape-from-shading problem. In *European Conference on Computer Vision*, pages 780–791, 2008.

[23] Berthold K. P. Horn. *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View*. PhD thesis, Massachusetts Institute of Technology, 1970.

[24] Berthold K. P Horn. Height and gradient from shading. *International Journal of Computer Vision*, 5(1):37–75, 1990.

[25] Berthold. K. P. Horn and Michael J. Brooks. The variational approach to shape from shading. *Computer Vision on Graphics and Image Processing*, 33(2):174–208, 1986.

[26] Katsushi Ikeuchi and Berthold K. P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 1981.

[27] Micah K. Johnson and Edward H. Adelson. Shape estimation in natural illumination. In *Computer Vision and Pattern Recognition*, pages 2553–2560, 2011.

[28] Ira Kemelmacher and Ronen Basri. 3d face reconstruction from a single image using a single reference face shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99:1–14, 2010.

[29] Ron Kimmel and James A. Sethian. Optimal algorithm for shape from shading and path planning. *Journal of Mathematical Imaging and Vision*, 14:237–244, 2001.

[30] Jan J. Koenderink and Andrea J. Van Doorn. The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*, 23:1573–1405, 1997.

[31] Peter Kovesi. Shapelets correlated with surface normals produce surfaces. In *International Conference on Computer Vision*, pages 994–1001, 2005.

[32] Michael S. Langer and Steven W. Zucker. Shape-from-shading on a cloudy day. *Journal of Optical Society of America*, 11(2):467–478, 1994.

[33] Chia-Hoang Lee and Azriel Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence.*, 26:125–143, 1985.

[34] Kyoung Mu Lee and C. C. Jay Kuo. Shape from shading with a linear triangular element surface model. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 15(8):815–822, 1993.

[35] Wan-Chun Ma, Tim Hawkins, and et al. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Eurographics Symposium on Rendering*, 2007.

[36] Alexndre Meyer, Hector Briceno Pulido, and Saida Bouakaz. User-guided shape from shading to reconstruct fine details from a single photograph. *Asian Conference on Computer Vision*, 1(4843):738–747, 2007.

[37] Fred E. Nicodemus, J.C. Richmond, and et al. *Geomertical Considerations and Nomenclature for Reflectance*. NBS Monograph 160, National Bureau of Standards, U.S. Department of Commerce, Washington, D.C, 1977.

[38] Shmuel Peleg and Gad Ron. Nonlinear multiresolution: A shape-from-shading example. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(12):1206–1210, 1990.

[39] Xavier Pennec and Inria Projet Epidaure. Probabilies and statistics on riemannian manifolds: Basic tools for geometric measurements. In *In IEEE Workshop on Nonlinear Signal and Image Processing*, pages 194–198, 1999.

[40] Alex P. Pentland. Finding the illuminant direction. *Journal of Optical Society of America*, 72:323–347, 1982.

[41] Alex P. Pentland. Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):170–187, 1984.

[42] Alex P. Pentland. Linear shape from shading. *International Journal of Computer Vision*, 4(2):153–162, 1990.

[43] Bui T. Phong. Illumination for computer generated pictures. *Communications of ACM*, 18:311–317, 1975.

[44] Emmanuel Prados. *Application of the theory of the viscosity solutions to the Shape From Shading problem*. PhD thesis, University of Nice Sophia-Antipolis, oct 2004.

[45] Emmanuel Prados and Olivier Faugeras. Perspective shape from shading and viscosity solutions. In *International Conference on Computer Vision*, volume 2, pages 826–831, 2003.

[46] Emmanuel Prados and Olivier Faugeras. Shape from shading: a well-posed problem. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Diego, California*, volume II, pages 870–877, jun 2005.

[47] Emmanuel Prados, Nitin Jindal, and Stefano Soatto. A non-local approach to shape from ambient shading. In *Proc. 2nd International Conference on Scale Space and Variational Methods in Computer Vision*, pages 696–708, 2009.

[48] Hossein Ragheb and Edwin R. Hancock. A probabilistic framework for specular shape–from–shading. *Pattern Recognition*, 36(2):407–427, 2003.

[49] Ravi Ramamoorthi and Pat Hanrahan. An efficient representation for irradiance environment maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 497–500, 2001.

[50] Ravi Ramamoorthi and Pat Hanrahan. On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *Journal of Optical Society of America*, pages 2448–2459, 2001.

[51] Elisabeth Rouy and Agnes Tourin. A viscosity solutions approach to shape-from-shading. *SIAM Journal of Numerical Analysis*, 29(3):867–884, 1992.

[52] Tal Simchony, Rama Cehllappa, and M. Shao. Direct analytical methods for solving poisson equations in computer vision problems. *IEEE Transaction on Pattern Analysis and Machine Intelligence.*, 12:435–446, 1990.

[53] Grahame Smith. The recovery of surface orientation from image irradiance. *DARPA Image Understanding Workshop, Palo Alto*, 1982.

[54] Willam A. P. Smith and Edwin R. Hancock. Recovering facial shape and albedo using a statistical model of surface normal direction. In *International Conference on Computer Vision*, pages 588–595, 2005.

[55] Willam A. P. Smith and Edwin R. Hancock. Recovering facial shape using a statistical surface normal model. In *The International Conference on Image Processing*, 2005.

[56] Thomas M. Strat. *A numerical method for shape-from-shading from a single image*. M.I.T., Department of Electrical Engineering and Computer Science, 1979.

[57] Richard Szeliski. Fast shape form shading. *Computer Vision, Graphics, Image Processing: Image Understanding*, 53:129–153, 1991.

[58] Kenneth E. Torrance and Ephraim M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal Of Optical Society Of America*, 57:1105–1114, 1967.

[59] Ping-Sing Tsai and Mubarak Shah. Shape from shading using linear approximation. *Image and Vision Computing*, 12(8):487–498, 1994.

[60] George Vogiatzis, Paolo Favaro, and Roberto Cipolla. Using frontier points to recover shape, reflectance and illumination. *The 10th IEEE International Conference on Computer Vision (ICCV'05)*, 1:228–235, 2005.

[61] Christine Waggoner, Son Chang, and Paul Debevec. Combining dynamic simulation, high dynamic range photography, and global illumination. *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 189–198, 1998.

[62] Philip L. Worthington. Reillumination-driven shape from shading. *Computer Vision and Image Understanding*, 98:326–344, 2004.

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

[64] Philip L. Worthington and Edwin R. Hancock. New constraints on data–closeness and needle map consistency for shape–from–shading. *IEEE Transaction on Pattern Analysis and Machine Intelligence.*, 21(12):1250–1267, 1999.

[65] Alan L. Yuille, Daniel Snow, Russell Epstein, and Peter N. Belhumeur. Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svdand integrability. *International Journal of Computer Vision*, 35:203–222, 1999.

[66] Gang Zeng, Long Quan, Yasuyuki Matsushita, and Heung-Yeung Shum. Interactive shape from shading. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.

[67] Lei Zhang and Dimitris Samaras. Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics. *IEEE Transaction on Pattern Analysis and Machine Intelligence.*, 28(3):351–363, 2006.

[68] Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Shape-from-shading: a survey. *IEEE Transaction on Pattern Analysis and Machine Intelligence.*, 21(8):690–706, 1999.

[69] Qinfen Zheng and Rama Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Transaction on Pattern Analysis and Machine Intelligence.*, 13(7):680–702, 1991.

[70] Todd Zickler, Satya P. Mallick, David J. Kriegman, and Peter N. Belhumeur. Color subspaces as photometric invariants. *International Journal of Computer Vision*, 2008.