ABSTRACT

SETHURAM, AMRUTHA SHREE. Reconstruction of Lambertian Surfaces from Photometric Stereo. (Under the direction of Dr. Wesley E. Snyder)

The objective of this thesis is to implement and compare two algorithms to reconstruct the shape of an object from photometric stereo. Photometric stereo is a practical technique for determining an object's shape and surface reflectance properties at a distance. The implementation proposes the use of three images of an object, recorded from the same viewpoint but with different illumination. The first algorithm employs the Fourier transform method to solve the minimization problem. The gradient data is obtained by incorporating photometric stereo method on image triplets. The Fourier transform of the unknown surface is then expressed as a function of the Fourier transform of the gradients. The relative depth values are then obtained by applying an inverse Fourier transform of the function. The second algorithm is based on iterative reconstruction which minimizes the cost function by gradient descent and annealing. Both these algorithms are implemented to reconstruct both real and synthetic surfaces and the results are compared. It is also shown that better reconstruction results are obtained by adopting the second algorithm in the presence of discontinuities in the image. Noise sensitivity of the frequency-domain method is also evaluated. An experimental setup to obtain real world images is also presented.

Reconstruction of Lambertian Surfaces from Photometric Stereo

By

Amrutha Shree Sethuram

A thesis submitted to the graduate faculty of
North Carolina State University
in partial fulfillment of the requirements for the degree of
Master of Science

Department of Electrical and Computer Engineering

Raleigh

2005

APPROVED BY

Chair of Advisory Committee

To my family

BIOGRAPHY

Amrutha Shree Sethuram, completed her Bachelors Degree in Electronics and Communication Engineering from R.V. College of Engineering in Bangalore, India. She has been a rank student throughout her schooling. After her under graduation, she worked as an Engineer in Picopeta Simputers Pvt. Ltd in Bangalore.

She then enrolled herself in the Masters Program in the Department of Electrical and Computer Engineering at North Carolina State University. She completed her thesis under the august direction of Dr. Wesley E. Snyder. Her research interests are in the area of image processing, pattern recognition and machine vision. She has also been trained in the fine arts of traditional Indian Carnatic music and dance.

ACKNOWLEDGEMENT

I would like to thank my advisor, Dr. Wesley E. Snyder for all his guidance and support throughout my research. He has always been a constant source of inspiration and motivation to excel in all my endeavors. I shall forever cherish my interactions with him and he shall remain my role model for his technical acumen and affable personality. I would also like to thank Dr. David Lalush and Dr. Edward Grant for having kindly agreed to be on my Masters Committee. I thank them for their time, patience and their valuable advice during the course of my thesis. I also thank the office of the Department of Electrical and Computer Engineering and my colleagues in the Image Analysis Laboratory Group for their assistance during my Masters.

I would like to thank my parents, Sethuram and Geetha, for their infallible love, encouragement and for having provided me with a conducive environment to achieve academic excellence. I also thank my brother, Narakesari, for his incessant support, love and for being my ideal sibling. I am grateful to my Father-In-Law, A.N.K.Murthy for instilling in me, a drive to achieve higher goals in life. The ever-lasting grace of my late Mother-In-Law, Nalini will be of perennial guidance in my life.

My deepest gratitude goes to my husband, Krishna Prasad, for having been bedrock of perpetual love, support and inspiration in life and at work. He has always been the light of my life and has been the torchbearer of my success. I offer my salutations to God for providing me with this opportunity to achieve my goals and I pray for His continued blessings.

TABLE OF CONTENTS

Page

LIST OF FIGURES vii
LIST OF TABLES x

1. INTRODUCTION 1

2. PROBLEM DESCRIPTION 3

2.1 Definitions 4
2.1.1 Radiometry 4
2.1.2 Lambertian Surface 6
2.1.3 Surface Orientation 8
2.1.4 The Reflectance Map 9
2.2 Problem Statement 12

3. LITERATURE REVIEW 16

3.1 Shape from Shading 16
3.2 Photometric Stereo 25

4. APPROACH 29

4.1 Projection Model 29
4.2 Range Map, Depth Map, Height Map, Gradient Map . 31
4.3 Backprojection 34
4.4 Visualization of Gradient Maps 34
4.4.1 Algorithm for needle map generation 35
4.5 Methodology 36
4.5.1 Step1: Generation of gradient maps using Photometric Stereo 37
4.5.2 Step 2: Generation of needle map 40
4.5.3 Step 3: Reconstruction from gradient maps . 40

5. ALGORITHM 1: FOURIER TRANSFORMS APPROACH 44

5.1 Theory and derivation 45
5.2 Algorithm 48
5.3 Implementation details 49

6. ALGORITHM 2: ITERATIVE RECONSTRUCTION 51

6.1 Using kernels to estimate derivatives 52
6.2 Function minimization 53
6.3 Algorithm 55
6.4 Implementation details 56

7. EXPERIMENTAL RESULTS 57

7.1 Introduction 57
7.2 Beethoven Image 57
7.2.1 Algorithm1: Results 59
7.2.1.1 Reconstruction Results 59
7.2.1.2 Noise Sensitivity 63
7.2.2 Algorithm 2: Results 69
7.3 Synthetic Images 73
7.3.1 Algorithm1: Results 74
7.3.1.1 Reconstruction Results 74
7.3.1.2 Noise Sensitivity 76
7.3.2 Algorithm 2: Results 80
7.3.3 More synthetic images 84
7.4 Real World Images 92
7.4.1 Experimental Setup 92
7.4.2 Experimental Images and Results 93

8. Conclusion and Future work 99

Bibliography 101

LIST OF FIGURES

Page

Figure 2.1 The direction of incident and emitted light rays are specified in a local coordinate system using q and j . . 5
Figure 2.2 Diagram to illustrate *BRDF* 6
Figure 2.3 An Ideal Lambertian surface 7
Figure 2.4 Representation of a surface point P on the Gaussian sphere through slant s and tilt q 8
Figure 2.5 Gnomic Projection 11
Figure 2.6 Stereographic Projection 11
Figure 2.7 Integration along a closed path 15
Figure 2.8 Object with discontinuities 15

Figure 4.1 Orthogonal parallel projection of a point P onto a point p . 30
Figure 4.2 A ray of projection for parallel projection model 31
Figure 4.3 Photometric stereo image triplet of Beethoven 37
Figure 4.4 (a) Gray value image depicting $p(x,y)$ (b) Gray value image depicting $q(x,y)$ 39
Figure 4.5 Needle map of Beethoven image triplet 40
Figure 4.6 Range images showing reconstruction of Beethoven triplet obtained by direct integration (a)Algorithm A (b)Algorithm B 42

Figure 7.1 Photometric stereo images of Beethoven 58
Figure 7.2 (a) Gray value image depicting $p(x,y)$ (b) Gray value image depicting $q(x,y)$ 58
Figure 7.3 Reconstruction of Beethoven image by Algorithm1 with $l=0$ and $m=0$ (a) Using hotmetal colormap (b)Grayscale representation 59
Figure 7.3 Profile of reconstructed Beethoven image by Algorithm1 (c)reconstructed image (d) Profile along VV axis (e) Profile along HH axis 60
Figure 7.4 Reconstruction of Beethoven image by Algorithm1 with $l=0.1$ and (a) $m=0.1$ (b) $m=0.3$ (c) $m=0.5$ 61
Figure 7.5 Reconstruction of Beethoven image by Algorithm1 with $m=0$ and (a) $l=0.2$ (b) $l=0.3$ (c) $l=0.5$ 62
Figure 7.6 Reconstruction of Beethoven image by Algorithm1 with $m=0.5$ and $l=0.5$ 63
Figure 7.7 Error images constructed by taking the difference between the reconstructed images of noisy gradients and the original reconstructed image of Beethoven by Algorithm1 with $l=0.1$ and (a) $m=0.1$ (b) $m=0.5$ (c) $m=10.0$ 64
Figure 7.8 Error images constructed by taking the difference between the reconstructed images of noisy gradients and the original reconstructed image of Beethoven by Algorithm1 with $m=0$ and (a) $l=0.1$ (b) $l=0.3$ (c) $l=0.5$ 65
Figure 7.9 Error as a function of variance of noise added to (a)P (b)Q

(Source: Beethoven Image) 67
Figure 7.10 Noise sensitivity as a function of (a) l (b) m
(Source: Beethoven Image) 68
Figure 7.11 (a)Reconstruction of Beethoven image by Algorithm2 (b)
(b)Reconstructed image showing profile axes (c) Profile
along VV axis (d) profile along HH axis 70
Figure 7.12 Reconstruction of Beethoven image by Algorithm2 with $l = 1.0$
and (a) $m = 1.0$ (b) $m = 5.0$ (c) $m = 10.0$ 71
Figure 7.13 Reconstruction of Beethoven image by Algorithm2 with $m = 5.0$
and (a) $l = 3.0$ (b) $l = 5.0$ (c) $l = 10.0$ 72
Figure 7.14 (a) Synthetic range image of a car
(b) Gray value image depicting $p(x, y)$ of (a)
(c) Gray value image depicting $q(x, y)$ of (a) 73
Figure 7.15 Reconstruction of the synthetic car image by Algorithm1
with $l = 0.1$ and $m = 0.1$ (a) Hotmetal Colormap (b) Grayscale 74
Figure 7.16 Profile of reconstructed synthetic car image by Algorithm1
(a) Reconstructed image (b) Profile along VV axis
(c) Profile along HH axis 75
Figure 7.17 Error as a function of variance of noise added to
(a)P (b)Q (Source: Synthetic Image) 78
Figure 7.18 Noise sensitivity as a function of (a) l (b) m
(Source: Synthetic Image) 79
Figure 7.19 (a)Reconstruction of synthetic image by Algorithm2
($l = 0.1$ and $m = 10.0$) (b) Profile axes on reconstructed
image (c) Profile along VV axis (d) Profile along HH axis 81
Figure 7.20 Graph of reconstruction error as a function of Temperature
$(l = 1.0 \text{ and } m = 10.0)$ 82
Figure 7.21 (a)Source Image (b)Gradient map $P(x, y)$
(c)Gradient map $Q(x, y)$
(d)Reconstruction by Algorthm1
(e)Reconstruction by Algorithm2 85
Figure 7.22 (a)Source Image (b)Gradient map $P(x, y)$
(c)Gradient map $Q(x, y)$
(d)Reconstruction by Algorthm1
(e)Reconstruction by Algorithm2 86
Figure 7.23 (a)Source Image (b)Gradient map $P(x, y)$
(c)Gradient map $Q(x, y)$
(d)Reconstruction by Algorthm1
(e)Reconstruction by Algorithm2 87
Figure 7.24 (a) Reconstructed image of frog by Algorithm1 showing
profile axis (b) Reconstructed image of frog by Algorithm2
showing profile axis (c) Profile along HH axis for
Algorithm1 (d) Profile along HH axis for Algorithm2 88
Figure 7.25 Graph of MSE as function of temperature(Source:banana.ifs) 89
Figure 7.26 Graph of MSE as function of temperature(Source:fighter.ifs) 90
Figure 7.27 Graph of MSE as function of temperature(Source:frog.ifs) . 90
Figure 7.28 (a), (b) and (c) photometric stereo images of a real
object, each obtained from a different light source
(d) Reconstructed image by Algorithm1
(e) Reconstructed image by Algorithm2 94

Figure 7.29 (a), (b) and (c) photometric stereo images of a real object, each obtained from a different light source (d) Reconstructed image by Algorithm1 (e) Reconstructed image by Algorithm2 95
Figure 7.30 (a), (b) and (c) photometric stereo images of a real object, each obtained from a different light source (d) Reconstructed image by Algorithm1 (e) Reconstructed image by Algorithm2 96
Figure 7.31 Reconstruction of images with shadows by (a1),(a2),(a3)Algorithm1 (b1),(b2),(b3) Algorithm2 98

LIST OF TABLES

Page

Table 4.1 Light source direction and strength of Beethoven image triplet 37
Table 6.1 Kernels to estimate first derivate 52
Table 6.2 Kernels to estimate second partial derivates 52
Table 7.1 Mean squared error as a function of l and $m = 0$ 66
Table 7.2 Mean squared error as a function of m 66
Table 7.3 Mean squared error as a function of l and $m = 10$ 77
Table 7.4 Mean squared error as a function of m 77
Table 7.5 Mean Squared Error as a function of l for Algorithm2 (Source:Auto.ifs) 82
Table 7.6 Mean Squared Error as a function of m for Algorithm2 (Source:Auto.ifs) 82
Table 7.7 MSE obtained for Algorithm1 and Algorithm2 (Source:Auto.ifs) 83
Table 7.8 MSE obtained for Algorithm1 and Algorithm2 in reconstruction of synthetic images 91

Chapter 1

Introduction

Smooth variations in the brightness or *shading* of objects in an image are an important cue in human vision for estimating the shape of objects. Researchers in human vision have attempted to understand and simulate the mechanisms by which the human eyes and brain use this shading information to recover three dimensional shapes. The problem of reconstruction of shape of an object from image shading has been of great interest to researchers in the field of computer vision. Approaches addressing the problem of reconstructing a surface given a brightness image are referred to as “Shape-From-Shading” (SFS) methods. SFS methods use a single image for shape recovery while “Photometric Stereo” techniques are shape recovery techniques which use more than one image. Various other methods used in reconstruction of surfaces include Shape-from-stereo, motion, texture, etc. The shading of an object depends on the shape of the object, reflectance properties of the surface and the distribution of light sources. The surfaces can be classified as Lambertian, specular and hybrid depending on their reflectance properties. Lambertian surfaces reflect light in all directions and have diffused reflectance while specular surfaces reflect all of incident light along a single direction. Hybrid surfaces are those with reflectance properties which are a combination of Lambertian and specular surfaces. In this thesis, we will restrict our study to Lambertian surfaces. The aim of this thesis is to examine the problem of reconstruction of Lambertian surfaces from photometric stereo. Two algorithms used to reconstruct the surface are implemented. Both these algorithms are based on reconstructing the surface by minimizing a cost function. This objective function

integrates surface derivatives, using the integrability constraints with the surface area and curvature constraints to solve the discontinuity problem.

The first algorithm employs the Fourier transform method to solve the minimization problem. The gradient data is obtained by incorporating a photometric stereo method on image triplets. The Fourier transform of the unknown surface is then expressed as a function of the Fourier transform of the gradients. The relative depth values are then obtained by applying an inverse Fourier transform of the function.

The second algorithm is based on iterative reconstruction and annealing. Both these algorithms are implemented to reconstruct both real and synthetic surfaces and the results are compared. Noise sensitivity of the frequency-domain method is also evaluated. An experimental setup to obtain real world images is also presented.

This thesis work is organized as follows. The problem description is presented in detail in Chapter 2. Chapter 3 is dedicated to literature survey. This is followed by a chapter which describes the approach. The details of the algorithms are presented in chapter 5 and chapter 6. The results obtained are presented in chapter 7. The conclusions and future work are presented in the last chapter.

Chapter 2

Problem Description

Photometric Stereo [1], introduced by R.J.Woodham is a novel technique used to recover shape of an object. The technique makes use of more than one image of the object, recorded from the same viewpoint but with different illumination. The central observation of this approach is that:

"Grey-levels recorded by an imaging system are *measurements* which result from real physical processes. Such a measurement may be viewed as a *constraint* on the possible interpretation of surface element that gave rise to it."

Photometric techniques may be applied when imaging variables, such as illumination and viewing geometry, can be controlled or at least known. It is probably the least computationally expensive method available for determining object shape. The investigation of this technique has been provided in [2]. Shape reconstruction techniques developed for photometric images can also be used to reconstruct the three-dimensional surface that corresponds to Scanning Electron Microscope (SEM) images [37]. SEM has been a principal tool for investigating materials in a wide range of fields from biological sciences to integrated circuit inspection and quality control. The algorithms that are implemented in this thesis can also be used to reconstruct the three-dimensional surface from surface orientations recovered from SEM images.

The main objective of this thesis is to examine the problem of reconstructing photometric stereo images. Two algorithms are implemented for the reconstruction of these photometric stereo images. The first algorithm is a regularization method which combines the integrability constraints, surface curvature, and area constraints into a single functional which is then minimized. The minimization is done by using Fourier transforms. This thesis also deals with the noise sensitivity of the Fourier transform method. The second algorithm is based on iterative reconstruction. This thesis shows that the algorithm based on iterative reconstruction procedure is more robust than the algorithm based on Fourier transform approach.

In the next section, some basic definitions from surface geometry are introduced.

2.1 Definitions

The image of a three-dimensional object depends on its shape, its reflectance properties, and the distribution of light sources. It also depends on the position of the object relative to the imaging system and on the object's attitude in space. A knowledge of radiometry is essential to understand how the brightness at a particular point in the image is determined as discussed in this section.

2.1.1 Radiometry

The amount of light falling on a surface is called the *irradiance*. It is the power per unit area incident on the surface. The amount of light radiated from a surface is called the *radiance*. It is the power per unit area per unit solid angle emitted from the surface. Brightness is determined by the amount of energy an imaging system receives per unit

apparent area. It is proved in [3] that the image irradiance E is proportional to the scene radiance L.

Scene radiance depends on the amount of light that falls on a surface and the fraction of the incident light that is reflected. It also depends on the geometry of light reflection. These directions are described in terms of a local coordinate system erected on the surface as shown in Figure 2.1

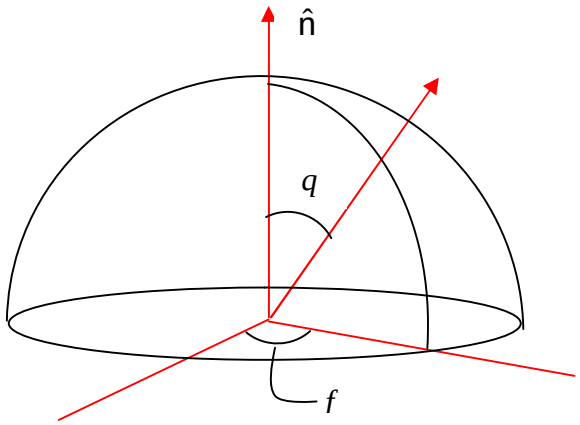


Figure 2.1 The direction of incident and emitted light rays are specified in a local coordinate system using q and f

The Figure 2.1 shows a perpendicular to the surface i.e. the normal $\hat{n}$ and an arbitrary reference line drawn on the surface. Directions are described by specifying the angle q between a ray and the normal and the angle f between a perpendicular projection of the ray onto the surface and the reference line on the surface. These angles are called the *polar angle* and *azimuth*, respectively. They are used to specify the direction (q_i, f_i) from which light falls on the surface and the direction (q_e, f_e) into which it is emitted towards the viewer as shown in Figure 2.2

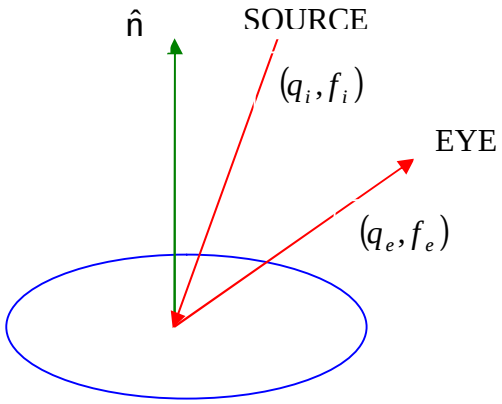


Figure 2.2 Diagram to illustrate *BRDF*

The *bidirectional reflectance distribution function* (BRDF) $f(q_i, f_i; q_e, f_e)$ is a measure of how bright a surface appears when viewed from one direction while light falls on it from another. If the light falling on the surface from the direction (q_i, f_i) i.e. the irradiance is $dE(q_i, f_i)$ and the brightness of the surface as seen from the direction (q_e, f_e) i.e. the radiance $dL(q_e, f_e)$, the BRDF is defined as the ratio of radiance to irradiance i.e.

$$f(q_i, f_i; q_e, f_e) = \frac{dL(q_e, f_e)}{dE(q_i, f_i)} \tag{2.1}$$

2.1.2 Lambertian Surface

An ideal Lambertian surface is one that appears equally bright from all viewing directions and reflects all incident light, absorbing none. From this definition, it can be deduced that the BRDF of an ideal Lambertian surface must be a constant. A Lambertian surface is one having perfect matte properties i.e. which obeys Lambert's cosine law.

Lambert's cosine law states that "The radiance in any direction from an element of a perfectly diffusing surface varies as the cosine of the angle between the direction of incidence and the normal vector of the surface at that element". Using the BRDF for a Lambertian surface, it can be shown that [3],

$$L = \frac{1}{p} E \cos q_i \quad \text{for } q_i \geq 0 \tag{2.2}$$

It can be seen that equation 2.2 is independent of the viewing direction i.e. the brightness of a Lambertian surface is a constant regardless of the direction from which it is viewed. An example of such a Lambertian surface is as shown in Figure 2.3

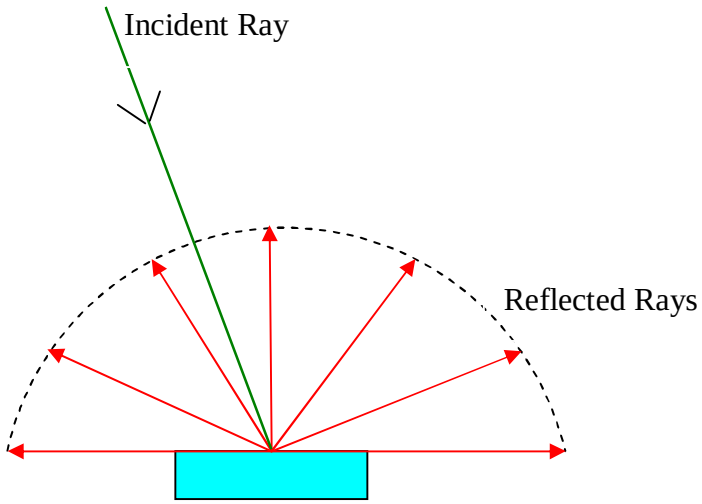


Figure 2.3 An Ideal Lambertian surface.

Surfaces covered with finely powdered transparent materials, such as barium sulphate or magnesium carbonate, come closest to obeying Lambert's law. It is a reasonable approximation for many other materials, such as paper, snow, and matte paint.

2.1.3 Surface Orientation

A reasonable notation for surface orientation has been developed in [3]. A smooth surface has a tangent plane at every point. The orientation of the tangent plane can represent the orientation of the surface at that point. The *surface normal*, a unit vector perpendicular to the tangent plane, is appropriate for specifying the orientation of this plane. The normal vector has two degrees of freedom, since it is a vector with three components and one constraint i.e. the sum of squares of the components must equal one. It can also be imagined that this vector is placed with its tail at the center of a unit sphere, called the *Gaussian sphere*.

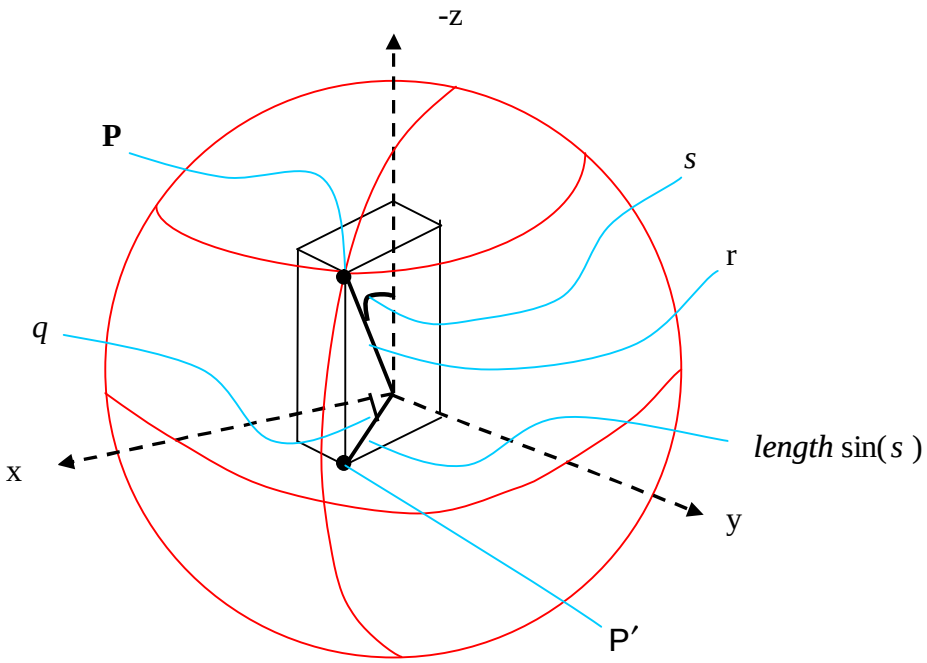


Figure 2.4: Representation of a surface point P on the Gaussian sphere through slant *s* and tilt *q*

The head of the vector touches the sphere at a particular point, which can be used to denote surface orientation as shown in Figure 2.4. The position of the point can be specified by two variables, slant and tilt.

A surface can be represented by a real valued function $z : X \times Y \to \Re$. This implies that a surface z will have a unique height $z(x, y) \in \Re$ at each point (x, y) in the domain

$$X \times Y \in \Re \times \Re . \tag{2.3}$$

The *gradient* of a differentiable surface z at a point (x, y) is defined by

$$\nabla z\Big|_{(x,y)} = \begin{bmatrix} p(x, y) & q(x, y) \end{bmatrix}^T \tag{2.4}$$

where

$$p(x, y) = \frac{\partial z}{\partial x}\Big|(x, y) \quad \text{and} \quad q(x, y) = \frac{\partial z}{\partial y}\Big|(x, y) \tag{2.5}$$

are the slopes of the surface in the $x-$and $y-$directions.

The gradient of a surface at a given point then corresponds to a point in the *gradient space* $P \times Q$. The upward unit surface normal at a point is given by

$$\hat{\mathsf{n}} = \frac{\mathsf{n}}{|\mathsf{n}|} = \frac{(-p, \quad -q, \quad 1)^T}{\sqrt{1 + p^2 + q^2}} \tag{2.6}$$

Both the surface gradient and the surface normal describe the orientation of a surface at a point (x, y).

2.1.4 The Reflectance Map

The reflectance map [4] encodes information about surface reflectance properties and light-source distributions. It is a representation tool used in developing methods for

recovering surface shape from images. It is a real-valued function defined on the gradient space which quantifies the brightness value to be observed at a point given the surface orientation at that point. It is well known [3, 5] that surface orientation affects the observed brightness. The brightness observed at a point depends on several factors: surface reflectivity, intensity of incident light, direction of incident light, and the viewing direction. Surface reflectivity is the ratio of power of reflected light from the surface to that incident on it. This is usually represented by a factor called albedo which indicates how much light the surface reflects relative to some ideal surface with the same geometry. Thus, if we measure the image irradiance $E(x, y)$, it is proportional to the radiance at the corresponding point on the surface imaged. If the surface gradient at that point is $(p(x, y), q(x, y))$, then the radiance at that point is $R(p(x, y), q(x, y))$. Thus by setting the constant of proportionality to one, we obtain

$$E(x, y) = R(p(x, y), q(x, y)). \qquad (2.7)$$

This image irradiance equation is fundamental to the method developed for recovering surface shape.

It is often convenient to plot the surface $R(p, q)$ as a function of the gradient (p, q). The $pq-$plane is called the gradient space, and every point in it corresponds to a particular surface orientation. As already explained in the previous section, a Gaussian sphere is used to represent the direction in which the surface normal is pointing. Sometimes, because of the curved surface of the Gaussian sphere, the points on the sphere are usually projected onto a plane to obtain the gradient space as shown in Figure 2.5. The gradient space is obtained my projecting the upper hemisphere of the Gaussian plane onto an infinite plane as shown in Figure 2.5 and is referred to as the *gnomic projection*. But in some cases, if the occluding

boundary is to be dealt with, another type of projection called the *stereographic projection* is used. In this type of projection, the whole sphere except for one point, is projected onto a plane, called the stereographic plane as shown in Figure 2.6

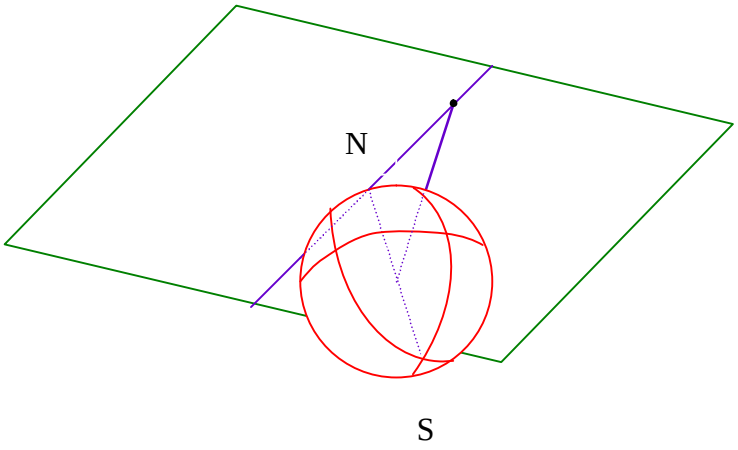


Figure 2.5 Gnomic Projection

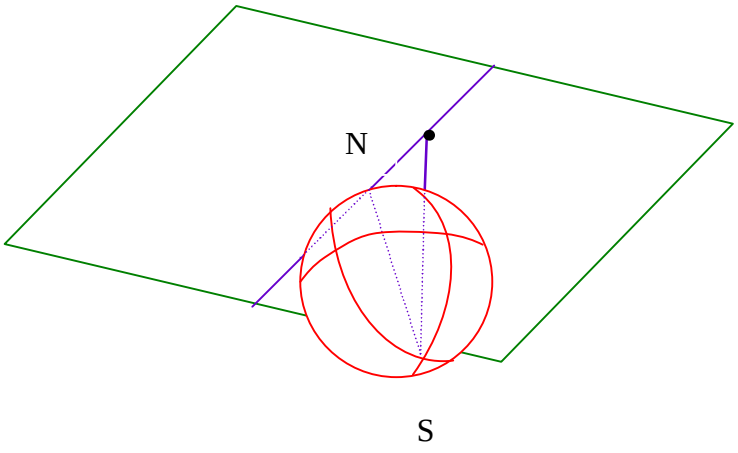


Figure 2.6 Stereographic Projection

2.2 Problem Statement

The objective of this thesis is to obtain photometric stereo images and to reconstruct the object from these images. Given three photometric stereo images, this work aims at initially obtaining the surface derivatives, p and q, and later reconstructing the surface from these gradients. Two algorithms are implemented to reconstruct the surfaces from the gradients. This thesis shows that the algorithm based on iterative reconstruction procedure is more robust than the algorithm based on Fourier transform approach.

The idea of photometric stereo [1] is to vary the direction of the incident illumination between successive views while holding the viewing direction constant. This thesis makes use of image triplets of an object whose surface is to be reconstructed. This is achieved by having a constant viewing direction and using different non-collinear light sources which are powered one at a time. In this section, it is shown as to how photometric stereo using three images can be employed to evaluate the surface normal and hence the derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

Suppose three images $I_1(x, y)$, $I_2(x, y)$ and $I_3(x, y)$ are obtained by varying the direction of incident illumination. Since there has been no change in imaging geometry, each picture element (x, y) in the three images correspond to the same object point and hence to the same gradient (p, q). The effect of varying the direction of incident illumination is to change the reflectance map $R(p, q)$ that characterizes the imaging situation. Thus we have three independent equations:

$$I_1(x,y) = R_1(p,q) \tag{2.8}$$

$$I_2(x,y) = R_2(p,q) \tag{2.9}$$

$$I_3(x,y) = R_3(p,q) \tag{2.10}$$

For reflectance characterized by $R_a(p,q)$ above, the three views uniquely determine both the surface orientation and the reflectance factor r at each image point. Let $\tilde{I} = [I_1, \quad I_2, \quad I_3]'$ be the column vector of intensity values recorded at a point (x,y) in each of the three views. Further, let $\tilde{n}_1 = [n_{11}, \quad n_{12}, \quad n_{13}]'$, $\tilde{n}_2 = [n_{21}, \quad n_{22}, \quad n_{23}]'$ and $\tilde{n}_3 = [n_{31}, \quad n_{32}, \quad n_{33}]'$ be unit column vectors which point in the direction of three directions of incident illumination. Let matrix N be defined by

$$N = \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & n_{32} & n_{33} \end{bmatrix} \tag{2.11}$$

Let $\tilde{n} = [n_1, \quad n_2, \quad n_3]'$ be the column vector corresponding to a unit surface normal at (x,y). Then

$$\tilde{I} = r \; N \; \tilde{n} \tag{2.12}$$

so that

$$r \; \tilde{n} = N^{-1} \; \tilde{I} \tag{2.13}$$

provided the inverse N^{-1} exists. The inverse exists if and only if the three vectors $\tilde{n}_1$, $\tilde{n}_2$ and $\tilde{n}_3$ do not lie in a plane. In this case, the reflectance factor and unit surface normal at (x,y) are given by:

$$r = \left| \; N^{-1} \; \tilde{I} \; \right| \tag{2.14}$$

$$\tilde{n} = \left(\frac{1}{r} \right) N^{-1} \tilde{I} \tag{2.15}$$

The above scheme can be generalized to any number of images k where $k >= 3$, with increased noise robustness due to the over determined system. However, when $k > 3$, the matrix N is rectangular and thus a pseudo-inverse of the matrix should be computed. Thus having obtained the surface normals at each point on the image, the problem at hand is now directed towards reconstructing the surface of the object from these normals. Let us suppose that we have determined surface orientation over the region Ω. The relative depth of surface points may be determined from the gradient (p, q) by means of the equality

$$dz = pdx + qdy$$

that relates infinitesimal changes in x, y and z. Integrating along a curve C from (x_0, y_0) to (x, y), we obtain

$$z(x, y) = z(x_0, y_0) + \int_C (pdx + qdy) \tag{2.16}$$

A depth value obtained at some point depends on the integration path that was taken to get there. If the surface orientation is known, elevation of the surface above some reference plane can be determined by integration along curves in the image. As shown in Figure 2.7, elevation of point (x, y) can be determined with reference to a point (x_0, y_0)

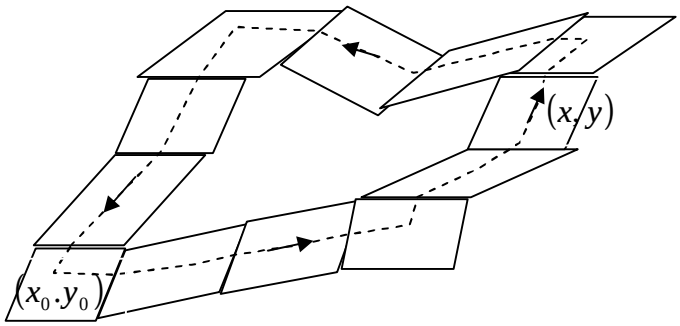


Figure 2.7 Integration along a closed path

However, in case of objects with discontinuities, the height of the surfaces can only be found up to an additive constant. In the figure 2.8 shown below, the difference in height between surfaces S1 and S2 i.e. h cannot be uniquely determined.

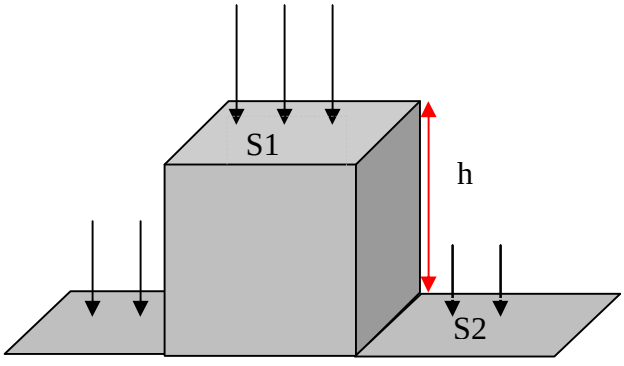


Figure 2.8 Object with discontinuities

Chapter 3

Literature Review

In the field of computer vision, the techniques used to reconstruct shape from brightness images are known as Shape-From-X techniques, where X can be one of shading, stereo, motion, texture, etc. Modern day researches in the field of human vision show that a human visual system makes use of all these cues while interpreting the shape of a three dimensional object. We shall however focus on Shape-From-Shading techniques since our goal is to relate image brightness at each pixel to the surface orientation at that pixel and recover shape from the recovered surface orientation.

3.1 Shape from Shading

"Shading appears to me to be of supreme importance in perspective, because, without it opaque and solid bodies will be ill-defined……" Leonardo Da Vinci, Notebooks

Human perception of objects in the real world strongly depends on shading of objects. Shading should not to be confused with shadowing. Shading provides cues all over the surface, not just along the special contours like shadowing. Artists have used this concept of shading for many centuries now, to convey the impression of a three-dimensional shape. Makeup artists exploit this effect to make a nose appear sharper by applying a thin line of light make-up along its ridge. Similarly, modern day computer graphics use shading to depict the illusion of the third dimension. It should be noted that given a surface and certain illumination conditions, it is fairly easy to determine the shading pattern. This is because the mapping from surface orientation to image brightness is unique [5]. However, the inverse

problem is generally not. The recovered shape can be expressed in several ways viz. depth $z(x, y)$, surface normal $[n_1, \quad n_2, \quad n_3]$, surface gradient (p, q), surface slant, f and tilt q. Depth can be considered as the relative distance between camera plane and surface points or the relative surface height above the $x - y$ plane. The surface normal is the orientation of a vector perpendicular to the tangent plane on the object surface. The surface gradient is the rate of change of depth in the x and y directions. The surface slant and tilt are related to the surface normal as $[n_1, \quad n_2, \quad n_3] = [l \sin f \cos q, \quad l \sin f \sin q, \quad l \cos f]$, where l is the magnitude of a vector in the direction of the surface normal.

The earliest work on the quantitative use of shading information was made in 1950-60 on recovering the shape of parts of the lunar surface in preparation for the human exploration of the moon [6]. In the field of computer vision, the problem was first formulated by Berthold K.P. Horn in 1970 in his PhD thesis [7]. Since then, a considerable amount of research has been undertaken in this field. In his work, Horn shows that the shape of an object can be obtained from shading if the reflectivity function and the position of light sources are known. He shows that a general solution of the shape-from-shading problem revolves around the image irradiance equation:

$$E(x, y) = R(p(x, y), q(x, y)) \tag{3.1}$$

where $E(x, y)$ is the image irradiance at the point (x, y), while $R(p(x, y), q(x, y))$ is the radiance of a surface patch.

The reflectivity and gradient of a surface are related by a non-linear first order partial differential equation in two unknowns. Horn solves this equation by setting up an equivalent set of five ordinary differential equations, three for the coordinates and two for the components of the gradient. These five equations are then integrated numerically along

certain curved paths. He argues that a component of the gradient can be found in one special direction since the gradient cannot be determined locally and repeats this process. This curve traced out of the object is called as ***characteristic curve.*** The projection of such a curve into the image is called a ***base characteristic***. The characteristic curve, along with the orientation defines a ***characteristic strip*** on the surface. However, to start the solution, an initial known curve is constructed near the singular points using spherical approximation. Singular points are those with maximum intensity. Thus the shape information is propagated simultaneously along the characteristic strips, assuming no crossover of adjacent strips. The direction of characteristic strips is identified as the direction of intensity gradients. Many closely spaced strips define the shape of the surface.

Oliensis [8] treats shape from shading as a partially well constrained problem. The author argues that the surface is uniquely determined over most of the image, but infinitely ambiguous in small regions bordering the image boundary, even though the image contains singular points. It is observed that the surface shape can be reconstructed from singular points instead of occluding boundary. It is also shown that the characteristic strips are curves of steepest ascent on the imaged surface.

The problem of surface reconstruction when viewed as an inverse problem on a non-linear first order differential equation in surface height can be solved through calculus of variations. The search for a function that minimizes an integral expression is the major concern of the calculus of variations [14]. Here, we find the valuable result that the extrema of functionals must satisfy an associated ***Euler equation***. A solution is constructed by seeking global optimum of a cost function that corresponds to the partial differential

equation under consideration, using variational calculus techniques and Euler equations. Since Euler equations nullify the partial derivatives of the cost function at all local extrema, any local minimum is treated as a global minimum. Thus a surface recovery problem can be transformed from one of minimizing a functional to one of solving one or more partial differential equations. Dupuis and Oliensis [9] constructed an algorithm based on calculus of variations which takes advantage of singular point constraints. Their algorithm does not use regularization and is capable of dealing with some orientation discontinuities. They have developed a solution using dynamic programming.

Based on the idea of [9], Bichsel and Pentland [10] developed an algorithm which is based on a minimum downhill principle which guarantees continuous surfaces and stable results. The authors believe that singular points play a key-role in shape from shading since each singular point introduces a three-fold ambiguity i.e. locally convex, concave or saddle shape to the solution and that a minimum downhill principle can remove this ambiguity. The authors also claim that the algorithm is applicable to a broad variety of objects and reflectance maps and converges in fewer than 10 iterations.

The main goal of a SFS problem is to obtain $p(x, y)$ and $q(x, y)$ such that the ***image irradiance equation*** is satisfied. Since each surface point has two unknowns for the surface gradient and each pixel in the image provides one gray value, we have an underconstrained problem. This causes the number of solutions to be infinite. To overcome this problem, additional constraints such as smoothness, integrability etc. should be introduced to limit the number of solutions to a few, preferably one. The general form of the optimization function to be minimized is given by:

$$W = \iint \left\{ \left(E(x, y) - R(p(x, y), q(x, y)) \right)^2 + l\, S(x, y) \right\} dxdy \qquad (3.2)$$

where $S(x, y)$ is a penalty term that penalizes deviation from some definition of regularity or smoothness.

Ikeuchi and Horn [11] were the first to use variational calculus in the derivation of iterative schemes for shape from shading. They employed the stereographic plane to express the orientations of surface patches, rather than the more commonly used gradient space. The use of the stereographic plane (shown in Figure 2.6) made it possible for them to incorporate occluding boundary information, but forced them to employ only a smoothness constraint. They solve the set of equations arising from the smoothness constraint and the image-irradiance equation iteratively, using occluding boundary information to supply boundary conditions. To ensure convergence, good initial values found at certain points were used. One disadvantage of the stereographic parameterization is that it is somewhat harder to express the condition of integrability. Thus the authors have chosen to ignore integrability, instead minimizing the integral of an unsmoothness penalty term. (The integrability condition does not allow the two partial derivatives of a function of two variables to be specified independently. Also, it requires that they must satisfy the condition that the two mixed derivatives of second order are the same. Mathematically, integrability condition can be written as $z_{yx} = z_{xy}$. This condition is referred to as ***integrability***, since one can recover the underlying surface $z(x, y)$ by evaluating line integrals of $(pdx + qdy)$ along arbitrary contours if the surface orientation information satisfies this constraint.)

Brooks and Horn [12], [15] expressed this minimization term in terms of surface normal. The authors point out that surface orientation has to obey an integrability constraint if it is to correspond to an underlying smooth surface and regularization methods do not guarantee

that the surface orientation recovered satisfies this constraint. They formulate a least-squares method for recovering height from estimates of the partial derivatives. They found that strictly enforcing the integrability condition does not lead to convergent iterative schemes. However, a convergent iterative scheme can be obtained by introducing a penalty term derived from the integrability constraint instead of enforcing integrability. The penalty term provided the iterative process with a sense of direction and the approach allowed one to recover surface gradients that are approximately integrable. The drawback however was that occluding boundary information could not be incorporated. This difficulty was overcome by expressing the integrability penalty term in terms of unit surface normal and its derivatives.

Frankot and Chellappa [16] enforced the integrability constraint in Brooks and Horn's algorithm in order to recover integrable surfaces, where integrability is defined by

$$z_{xy}(x,y) = z_{yx}(x,y) \forall (x,y) \tag{3.3}$$

They represent a nonintegrable estimate of surface slopes by a finite set of basis functions and enforce integrability by calculating the orthogonal projection onto a vector subspace spanning the set of integrable slopes i.e. they express surface orientations that correspond to integrable surfaces as a linear combination of a finite set of orthogonal Fourier basis functions and enforce integrability by projecting the nonintegrable surface slope estimates onto the nearest integrable surface slopes. Mapping the estimates to the nearest integrable field of surface orientations was done by finding the closest set of coefficients, in a least-squares sense, that also have the property of integrability. Their algorithm showed improvements both in accuracy and efficiency when compared to [12].

Later in his work, Horn [13] replaced the smoothness constraint by the integrability constraint of equation 3.3. Vega and Yang [19] developed a heuristic-based algorithm in which the heuristics were derived based on the geometrical interpretation of Brooks and Horn [12]. They performed their experiments using synthesized objects based on superquadric equations. The performance of their heuristic algorithm known as the "Shading Logic Algorithm" achieves a better performance than variational approaches and improves the stability of [12]. Additional techniques have been proposed to enforce the integrability constraint by several authors. Kim and Park [17] proposed an algorithm in which the 3D surface was approximated by Legendre polynomials and the relationships between the surface and its derivatives were represented in matrix forms using a polynomial coefficient vector. They have proposed shape from shading and photometric stereo algorithms in which the relative depth and its derivatives are iteratively updated.

Zheng and Chellappa [20] developed an iterative scheme using calculus of variations and a linear approximation of the reflectance map. They make use of a cost function that does not use the quadratic regularization term but instead, they implement the smoothness constraint by requiring the gradients of the reconstructed image to be close to the gradients of the input image.

It can be observed that a problem common to all the minimization techniques is their slow rate of convergence. To work around this problem, Szeliski [18] expressed the cost function purely from a discrete point of view using discrete approximation to the surface gradient on a regular grid, as opposed to previous techniques which developed iterative schemes on continuous coordinates using calculus of variations and discretized them later

for computations. The discrete formulation then enabled the use of functional optimization techniques such as the gradient descent that could provide faster convergence. On the basis of this theory, he employed a gradient descent method with Jacobi iteration and used hierarchical basis representations of height and gradient fields to further improve the rate of convergence.

Rather than using variational calculus techniques, Leclerc and Bobick [21] used a method that directly solved for surface height. By using a discrete formulation of the problem, they were able to achieve good convergence by employing numerical solution techniques. Further, they used a smoothness constraint to ensure convergence and used stereo processing to provide initial and boundary conditions.

Lee and Kuo [22] proposed to combine a triangular element surface model with a linearized reflectance map to formulate the shape-from-shading problem. They approximated a smooth surface by the union of triangular surface patches called “Triangular Elements” and expressed the approximating surface as a linear combination of a set of nodal basis functions. They determine the surface heights by minimizing a quadric cost functional corresponding to the squares of brightness errors which is solved by multigrid computation technique. Their algorithm does not require any integrability constraint or depth initialization.

Pentland [23], [24] recovered shape information from brightness patterns in small patches. His approach was to make an assumption about the surface shape instead of requiring prior information about the imaging environment. An assumption that the surface is locally spherical was made and a solution was sought based on the first and second partial derivatives of brightness. Based on the local intensity measurements, he estimated the

surface type to be one of planar, cylindrical, doubly curved or saddle and solved for surface orientations. He also could estimate the illumination direction based on intensity gradients.

Based on the same assumption, that the surface is spherical, Lee and Rosenfeld [25], [26], [27] formulated a method of estimating Lambertian surface shape with the help of a coordinate system having one axis in the assumed direction of the light source. They solved for estimates of slant and tilt angles in the coordinate system of the light source that are based on first-order derivatives of image brightness in contrast to the method adopted by Pentland [23] which used second order derivatives.

Pentland, later in his work in [28] used a linear approximation of the reflectance function in terms of surface gradient and applied the Fourier transform to the linear function to derive a closed form solution for the depth at each point. Further, the solution does not employ any assumptions about surface smoothness. Similarly, Tsai and Shah [29] used a linear approximation approach, the major difference being that they first used the discrete approximations for surface normal using finite differences and then linearized the reflectance function in depth instead of the surface normal. They employ a Jacobi iterative scheme for computing the depth map.

In general, shape from shading techniques can be divided into four groups:

i. Minimization Approaches: These approaches obtain solution by minimizing an energy function. [16], [15], [13], [11], [21], [22], [18], [19], [20] are minimization approaches.

ii. Propagation Approaches: These propagate the shape information from a set of surface points viz. singular points to the whole image. [10], [9], [7], [8] are propagation approaches.

iii. Local approaches: These derive shape based on the assumption of surface type. [23], [24], [25] are local approaches.

iv. Linear Approaches: These compute the solution based on the linearization of reflectance map. [28], [29] are linear approaches.

Zhang et al. [30] have conducted a survey on various shape-from-shading techniques and have implemented six algorithms to supplement their comparison and performance analysis. They found that the performance of each of the algorithms implemented were not very consistent across all the images. They have however inferred that minimization approaches are robust while the other approaches are faster.

3.2 Photometric Stereo

We have seen from the previous section that the problem of solving for surface orientation given a single brightness image is underconstrained. Shape-from-shading techniques as discussed in the previous section attempt to solve this underconstrained problem by introducing additional constraints such as smoothness constraint, integrability, a brightness constraint, etc. However, in situations where the imaging conditions can be controlled, a straightforward approach to constrain the problem can be developed by taking additional measurements under varying conditions. Thus we can convert this underconstrained problem into a constrained or overconstrained problem by taking additional images of the object under varying conditions. This is the fundamental theory of photometric stereo [1] as was explained in section 2.2.

In [31], Woodham assumes that three images of a surface with uniform reflectance properties are obtained from the same viewing angle but under different illumination

conditions. A fixed viewing direction eliminates the correspondence problem between image pixels. He then shows, using explicit formulations for various reflectance maps, that in a true image triplet not corrupted by noise or distortion, the three isophote curves intersect at only one point which is the true surface gradient (In this context, an isophote curve is a curve in gradient space (p, q) to which the possible solutions of equation 3.1 are limited. We thus have three isophote curves, one for each image). He points out that the images required for photometric stereo can be obtained either by explicitly moving a single light source, or by using multiple light sources calibrated with respect to each other or by rotating the object surface and imaging hardware together to simulate the effect of a moving light source. He also comments that the equivalent of photometric stereo can be achieved in a single view by using multiple illuminations which can be separated by color.

While using photometric stereo with three images, Zhang and Shah [32] opine that shape can only be recovered in the areas that are illuminated in all the three images. They consider a large sequence of images at any given time. Their process can be viewed as cascading shape from photometric stereo, which is formulated in the framework of a linear Kalman filter in order to iteratively recover and refine the shape and surface albedo. Their algorithm iteratively updates depth and albedo estimates calculated from subsequent photometric stereo triplets.

It can be noted that when the number of light sources is reduced to two, an ambiguity arises in the surface orientations. As clear from [31], the number of possible orientations in a two-image photometric stereo is two. [33] and [34] work around this problem by considering local integrability and convexity of surfaces respectively.

In all our discussion above, we have assumed uniform reflectance properties throughout the surface. This might not be the case in real surfaces. The non-uniformity can be attributed to albedo variations in the surface. It should be noted that we require four images to resolve albedo together with the set of surface orientations, except in special cases like Lambertian surfaces where only three images suffice both the recovery of surface orientation and albedo [3]. This technique of using four brightness images to estimate albedo together with surface normals is referred to as “Extended Photometric Stereo” [35].

Many of the shape-from-shading and photometric techniques work very well for continuous surfaces but fail when applied to discontinuous objects. Karacali and Snyder [36] describe an adaptive surface reconstruction method from a given gradient field that allows discontinuities in the solution. They formulate the problem in a discrete setting to relax the uniform integrability assumption to partial integrability to reconstruct discontinuous surfaces. Their method computes an optimal surface that corresponds to a given gradient field by projecting the gradient field onto a feasible subspace spanned by a set of orthonormal gradient fields. Further Karacali, in his PhD thesis [37] develops methods to reconstruct a three-dimensional surface together with a characterization of the surface composition given one or more images obtained from the same viewing direction. He applies his algorithm to scanning Electron Microscopy (SEM) images to extract specimen surface topography and material type information.

Wei and Klette [38] propose two improvements for reflectance-based shape recovery. Their algorithm is a more general version of Frankot-Chellappa-algorithm [16]. To improve the accuracy and robustness and to strengthen the relation between the depth map and

surface normals, they introduce two additional constraints into the cost function so that the partial derivatives of the unknown surface are as close as possible to the measured gradients.

$$z_{xx} = \frac{\partial^2 z}{\partial x^2} \tag{3.4}$$

and

$$z_{yy} = \frac{\partial^2 z}{\partial y^2} \tag{3.5}$$

The two new constraints model the behavior of the second-order derivatives such that the changes of depth maps will be more regular. They use Fourier transforms, as in [16], to solve for the unknown.

Wei and Klette further modified their algorithm in [39] and [40] by adding another surface curvature term into the cost functional to be minimized. Thus the new functional which they minimize is:

$$W = \iint_{\Omega} \left(|z_x - p|^2 + |z_y - q|^2 \right) dxdy + l \iint_{\Omega} \left(|z_x|^2 + |z_y|^2 \right) dxdy + m \iint_{\Omega} \left(|z_{xx}|^2 + 2|z_{xy}|^2 + |z_{yy}|^2 \right) dxdy \tag{3.6}$$

In the above cost function, the second term on the right hand side is a small deflection approximation of the surface area and the third term is a small deflection approximation of the surface curvature. The non-negative regularization parameters l and m establish a trade-off between the constraints. They express the Fourier transform of the unknown surface as a function of the given gradient's Fourier transforms. The relative depth values are then obtained by an inverse Fourier transform and by choosing associated weighting parameters.

Chapter 4

Approach

The two main steps involved in the reconstruction process are:

- Computation of surface gradients: This is done by employing the method of photometric stereo. As explained in section 2.2, three images of the object are obtained under different illumination conditions but with a fixed view. The surface normals and gradients are obtained at each point of the image using the equations shown in section 2.2
- Reconstruction from gradient vector fields.

Suppose the surface function $z(x, y)$ of an object is formed by an orthographic, i.e. parallel, projection of the surface into the $xy-$image plane, and defined in the image plane over a compact region Ω. The gradient values of this surface at discrete points $(x, y) \in \Omega$,

$$p(x, y) = \frac{\partial z(x, y)}{\partial x} = z_x \text{ and } q(x, y) = \frac{\partial z(x, y)}{\partial y} = z_y \qquad (4.1)$$

are computed from the first step. The algorithm used to find the gradient values is discussed in section 4.5

4.1 Projection Model

An orthographic projection i.e. orthogonal parallel projection is assumed in the implementation of the algorithms. For this kind of projection, every point $\mathsf{P} = (X, Y, Z)$ of R^3 is projected unambiguously onto a point $\mathsf{p} = (x, y)$ of the image plane (projection plane), with the projection equations

$$x = X \quad \text{and} \quad y = Y \tag{4.2}$$

As shown in Figure 4.1, the projection center is a point at infinity.

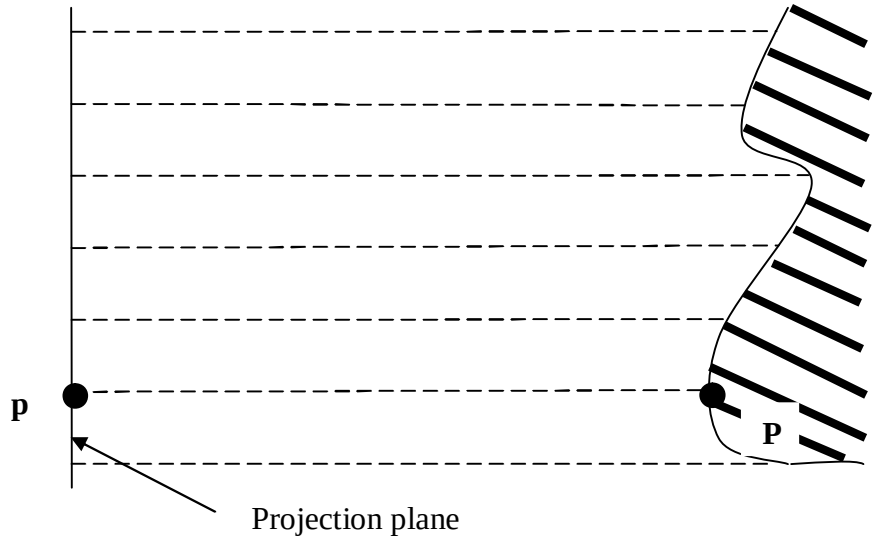


Figure 4.1 Orthogonal parallel projection of a point P onto a point p

Let f be the focal length of the camera. The real space R^3 i.e. $XYZ-$ space can be described by homogenous coordinates (t,u,v,w). The corresponding general parallel projection equation can be written as

$$(t', u', v', w')^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot (t, u, v, w)^T \tag{4.3}$$

In homogenous coordinates, a scene point

$$(X,Y,Z)=\left(\frac{t}{w},\frac{u}{w},\frac{v}{w}\right) \tag{4.4}$$

is projected into the image point (x, y, f), with

$$x=\frac{t'}{w'}=\frac{t}{w}=X\,, \qquad y=\frac{u'}{w'}=\frac{u}{w}=Y\,, \qquad z=\frac{v'}{w'}=\frac{f\cdot w}{w}=f \tag{4.5}$$

This use of orthogonal projection model is a valid assumption when objects are small compared to the camera constant or when objects are relatively far away from the camera.

4.2 Range Map, Depth Map, Height Map, Gradient Map

For parallel projection, the ray of projection into the scene space can be modeled as

$$g(t)=(x,y,t), \text{ with } t \geq f\,. \tag{4.6}$$

This ray passes through a certain image point $Q=(x, y, f)$ and is parallel to the optical axis. In this case the projection center is a point at infinity.

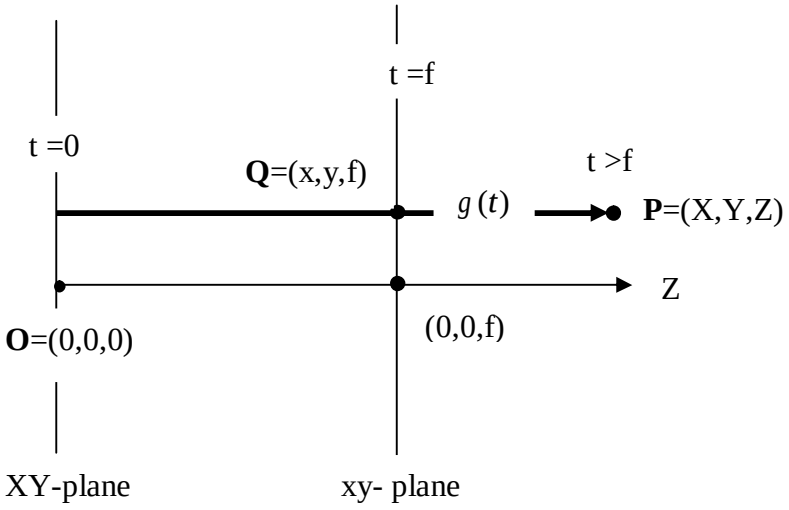


Figure 4.2 A ray of projection for parallel projection model

A surface point $\mathsf{P} = (X, Y, Z)$ is *projected into an image point* $\mathsf{p} = (x, y)$ of the image plane if a ray of projection exists which intersects the object surface at $\mathsf{P} = (X, Y, Z)$ for the first time (i.e. for the smallest $t \geq f$) after passing through the image point $\mathsf{Q} = (x, y, f)$. This definition of projection into the image plane specifies at most one surface point for each image point. This excludes the situation in the case of parallel projection in which all those object surface points with identical $XY-$coordinates are geometrically projected into the same point (x, y) (e.g. surface points on a face which is orthogonal to the image plane). In general, if several surface points are incident with the same ray of projection, then that point with the smallest $t-$coordinate value is selected for purposes of reconstruction. If a surface point is projected into an image point then this surface point is *visible* with respect to the model.

This definition allows synthetic objects to be dealt with as follows: given a surface function i.e. a synthetic model of a real-world object that is to be reconstructed, then the ray coordinates $g(t)$ can be substituted into the given surface equation in one variable t. The given equation in three variables X, Y, Z is transformed into an equation in one variable t. The solution(s) of this equation define the intersection point(s) of the ray of projection with the given surface. By specifying image coordinates (x, y), the relevant projected surface point(s) can be calculated. For such an approach, it can be assumed that the focal length f is given, and the generated images can be used to evaluate shape recovery techniques because the shape of the synthetic object i.e. the ground truth is known.

The *depth map* respectively the *height map*, and the *gradient map* are functions over the image grid $\{(x,y): 1 \le x \le M \wedge 1 \le y \le N\}$. The value distribution in depth maps or height maps can simple be visualized by gray value images. In the ideal case, at each point (x,y), a depth map $Z = Z(x,y)$ states the depth Z of that surface point $\mathsf{P} = (X,Y,Z)$ that is projected into the image point $\mathsf{p} = (x,y)$.

A *height map* $H(x,y)$ is defined relatively to an assumed background plane of height zero, which is parallel to the image plane. In the ideal case at each point (x,y) the value $H(x,y)$ is equal to the height of those surface points $\mathsf{P} = (X,Y,Z)$ that is projected into the image point $\mathsf{p} = (x,y)$. Height is measured with respect to the chosen background plane. Visualizations of the depth map and height map behave approximately like a positive and negative to each other.

For visualizing or representing $2\frac{1}{2}D-$ reconstructions the calculation of a range image is another option. However the Euclidean distance

$$d_2(\mathsf{P},\mathsf{Q}) = \sqrt{(X-x)^2 + (Y-y)^2 + (Z-f)^2} \tag{4.7}$$

between the visible point $\mathsf{P} = (X,Y,Z)$ and the projected point $\mathsf{Q} = (x,y,f)$ in the image plane are geometrically equivalent to depth Z. Sensors such as scanning lasers are often used to produce a range map. In this thesis, depth maps are used to present the final reconstruction of surfaces.

A *gradient map* is another representation of object surfaces. In the ideal case at each image point (x,y) a gradient map states the gradient $(p,q)^T$ of the object surface point $\mathsf{P} = (X,Y,Z)$ that is projected into this image point assuming that such a gradient is defined

at the projected point. Thus a gradient map can be viewed by two gray value images depicting the functions $p(x, y)$ and $q(x, y)$.

4.3 Backprojection

Surfaces of scene objects are projected into the image plane by projections as discussed above. The inverse mapping of projected object surfaces into the 3D space is generally called Backprojection. Depth maps (or height maps) and gradient maps provide important but differing information for such a back projection.

4.4 Visualization of Gradient Maps

Needle Diagrams are used to visualize gradient maps. Known gradient vectors are entered into the needle diagram as line segments starting at the image points where they are determined. For the representation of these gradient vectors as a needle diagram, it is assumed that values of calculated normals or gradients are given at image points (x, y) as input. The given vectors are either assumed as vectors of arbitrary length in $XYZ-$coordinates or as unit normals in coordinates of the Gaussian sphere (i.e. slant and tilt). A needle is defined to represent a normal as a line segment. A given normal $\mathsf{n} = (n_X, n_Y, n_Z)^T$ in the $XYZ-$coordinate system with a non-scaled length is represented by its unit normal $\mathsf{n}^0 = (n_1, n_2, n_3)^T$ since an un-normalized length would complicate the visual interpretation of the needle diagram. For the unit normal the representation of the Gaussian sphere is used as a basis for representation as shown in Figure 2.4. A needle diagram obtained for the Beethoven image is illustrated in Figure 4.5

The direction of the line segment is determined by q and the length of the line segment is determined by $\sin(s)$. Therefore for the coordinate differences that have to be calculated it holds that $\Delta x = n_1$ and $\Delta y = n_2$ with $0 \le \Delta x, \Delta y \le 1$.

Since these $\Delta -$values lie inside the interval $[0,1]$ a scaling factor s is used for the length of the line segment. Then the segment is drawn on the screen from the image point (x, y) that is the projective image of an object surface point P to an image point $(x + s \cdot \Delta x, y + s \cdot \Delta y)$. Furthermore, only the r^{th} line segment is drawn. It is also assumed that for every image point (x, y) at most one surface normal starting at (x, y) has to be represented.

4.4.1 Algorithm for needle map generation

The algorithm used to generate needle map of objects is shown below.

Algorithm: Needle map generation

1. **for** every image point $(x, y) = (r \cdot i, r \cdot j)$, in which a normal n was determined
2. **begin**
3. **if** (n determined in XYZ coordinates) **then**
4. **begin**
5. let $\mathsf{n} = (n_X, n_Y, n_Z)^T$;
6. $\|\mathsf{n}\| := \sqrt{n_X^2 + n_Y^2 + n_Z^2}$;
7. **if** $\|\mathsf{n}\| = 0$ **then** {error in normal determination}
8. **begin** $\Delta x := 0$; $\Delta y := 0$ **end**

9. **else**

10. **begin**

11. $\Delta x := n_X / \|\mathsf{n}\|; \quad \Delta y := n_Y / \|\mathsf{n}\|;$

12. **end**

13. **end**

14. **else** { n is a unit normal given in spherical coordinates s , q }

15. **begin**

16. $\Delta x := \cos(q) \cdot \sin(s);$

17. $\Delta y := \sin(q) \cdot \sin(s);$

18. **end**

19. $\Delta x := s \cdot \Delta x; \quad \Delta y := s \cdot \Delta y;$

20. draw a line from $(r \cdot i, r \cdot j)$ to $(r \cdot i + \Delta x, r \cdot j + \Delta y)$

21. **end**

4.5 Methodology

As explained in the beginning of this chapter, the first step in the process of surface reconstruction is the generation of gradient maps. The method of Photometric stereo is incorporated for the generation of these maps. The theory of Photometric stereo was explained in section 2.2. This section highlights the generation of gradient maps, needle map and a general reconstruction procedure for a set of three photometric stereo images [41] as shown in the Figure 4.3.

Beet1.jpg **Beet2.jpg** **Beet3.jpg**

Figure 4.3 Photometric stereo image triplet of Beethoven

Figure 4.3 shows three images of Beethoven from three different light sources which vary in their direction and strength. Table 4.1 shows these details of light source direction and strength for the images shown above.

Table 4.1 Light source direction and strength of Beethoven image triplet

Filename	Light Source Direction	Strength
Beet1.jpg	[-0.257723, -0.081351, -1]	1.0
Beet2.jpg	[0.017145, 0.202009, -1]	0.96688741
Beet3.jpg	[0.270128, -0.102021, -1]	1.03973

4.5.1 Step 1: Generation of gradient maps using Photometric Stereo

For the given set of images and their light source directions, the first step is to generate gradient maps. The algorithm used to generate gradient maps using Photometric stereo is explained below.

Algorithm: Generation of gradient maps using Photometric stereo

1. **Input** set of three photometric stereo images $I_1(x,y)$, $I_2(x,y)$ and $I_3(x,y)$

2. **Input** light source direction vectors

3. **Find** the corresponding unit vectors $\tilde{n}_1 = [n_{11}, \quad n_{12}, \quad n_{13}]'$, $\tilde{n}_2 = [n_{21}, \quad n_{22}, \quad n_{23}]'$ and $\tilde{n}_3 = [n_{31}, \quad n_{32}, \quad n_{33}]'$

4. **Form** matrix $N = \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & n_{32} & n_{33} \end{bmatrix}$

5. **Compute** N^{-1} i.e. the inverse of matrix N (Note: Inverse exists only if the three sources are non-collinear.

6. **for** each image point (x,y) **do**

7. **find** column vector $\tilde{I} = [I_1, \quad I_2, \quad I_3]'$ where I_1, I_2, I_3 are intensity values at point (x,y) in the three images

8. **compute** $N^{-1}\tilde{I}$ and $r = \left| N^{-1}\tilde{I} \right|$

9. **compute** $\tilde{n} = \left(\frac{1}{r} \right) N^{-1}\tilde{I}$ where $\tilde{n} = [n_x, n_y, n_z]'$

10. **compute** $p = n_x / n_z$ and $q = n_y / n_z$

11. **end**

12. Display gradient maps.

The gradient maps are generated by the above algorithm for the images shown in Figure 4.3. The resulting gradient maps i.e. the two gray value images depicting the functions $p(x, y)$ and $q(x, y)$ of Figure 4.3 are as shown in Figure 4.4 below.

(a)

(b)

Figure 4.4: (a) Gray value image depicting $p(x, y)$
(b) Gray value image depicting $q(x, y)$

4.5.2 Step 2: Generation of needle map

The next step is the generation of a needle map in accordance with the algorithm explained in section 4.4.1. Figure 4.5 shows the needle map generated for the Beethoven image triplet. A low sampling frequency (sampling interval =8) was used for the same. It should be noted that needles are dots for object faces that are oriented parallel to the image plane i.e. the normal is orthogonal to the image plane and $\sin(s) = 0$.Better results can be obtained with sophisticated graphics procedure to draw line segments.

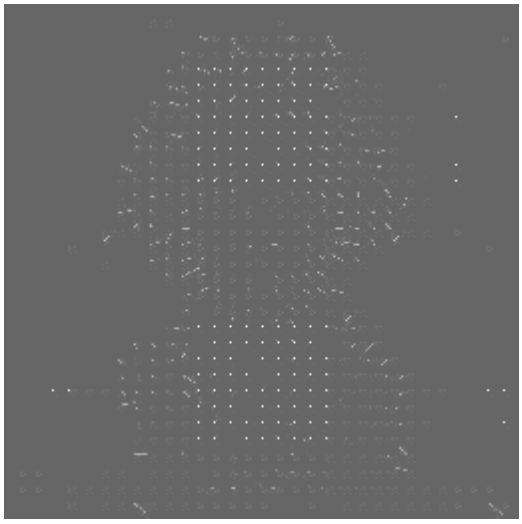


Figure 4.5 Needle Map of Beethoven image triplet

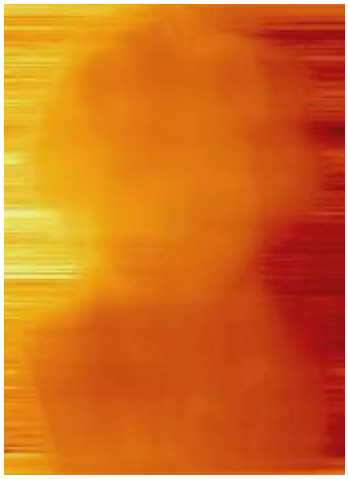
4.5.3 Step 3: Reconstruction from gradient maps

Having found the gradient maps from step 1, it would seem that an obvious approach to reconstruct the surface is by direct integration of p and q values at each point. This is

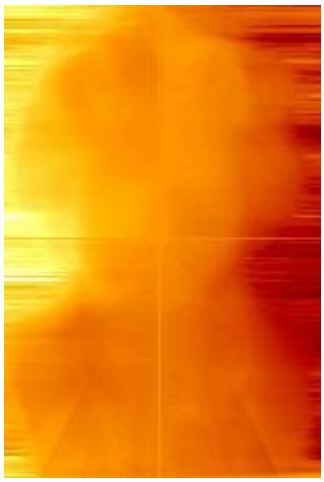
made clear by equation 4.1 which suggests that the surface z can be obtained by performing integration of both p and q values at each point on the surface. Such a technique is termed as a *local integration technique*. This method is conceptually simple and is based on the following curve integral:

$$z(x, y) = z(x_0, y_0) + \int_g p(x, y)dx + q(x, y)dy \tag{4.8}$$

where g is an arbitrarily specified integration path from (x_0, y_0) to (x, y). Starting with this initial height value, this technique propagates height values over the surface. The center of the image which lies within the object was chosen as an initial reference point. The height of every other point on the surface was found with respect to this reference point. Two different algorithms were implemented for the same. In the first algorithm, integration from the reference point to every other point on the surface was performed in two steps: first in direction of q gradient and next along the p gradient. In the second algorithm, integration was performed from the reference point to every point on the surface along the shortest path connecting the two. The reconstruction of the surface is depicted by range maps. The resulting surface obtained from both the algorithms is shown in Figure 4.6. A Hot Metal color map is used to display the same.



(a)



(b)

Figure 4.6 Range images showing reconstruction of Beethoven triplet obtained by direct integration
a) Algorithm A b) Algorithm B

It can be seen from Figure 4.6 that the direct integration method (local integration technique) resulted in poor reconstruction of the surface due to the fact that some of the assumptions made were not valid. Specifically,

- the surface is integrable at every point.
- the gradients exist at every point on the surface.
- all the points along the line of integration lie on the surface.

Also, in this approach, the locality of the computations propagates errors, i.e. this approach strongly depends on data accuracy. More often than not, the gradients obtained are corrupted by noise and hence the data available is not very accurate. To account for the drawbacks inherent in the above algorithm, two other algorithms were implemented. A global integration technique was adopted in these algorithms. This was done by minimizing a cost function and hence these algorithms can be classified as minimizing approaches. Regularization terms are introduced in this cost function which account for surface area and curvature. In chapter 5, an algorithm proposed by [40] is implemented. In this algorithm, Fourier transform of the unknown surface is expressed as a function of the Fourier transform of the gradients. The relative depth values are then obtained by an inverse Fourier transform. In chapter 6, unknown surface $z(x, y)$ is solved by minimizing the cost function by adopting iterative reconstruction.

Chapter 5

Algorithm 1: Fourier transforms approach

It was seen from the previous chapter that local integration methods do not perform well when the data is not accurate. Also, assuming that the gradient field is integrable meant that p and q could be added to get a depth map and the resulting surface would not be dependent on the choice of integration. It should be noted that integrability of surface orientations p and q means that the following equality holds

$$p_y(x,y) = q_x(x,y) \quad \forall(x,y) \tag{5.1}$$

This is also equivalent to having the equality

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

hold over all closed curves C in the surface domain. The method proposed by [16] enforced integrability by expressing surface orientations that correspond to integrable surfaces as a linear combination of Fourier coefficients.

Global integration techniques are based on minimizing the following cost function:

$$W = \iint_{\Omega_z} \left(\left| z_x - p \right|^2 + \left| z_y - q \right|^2 \right) dxdy \tag{5.3}$$

In the algorithm discussed below, the above functional is modified to reduce the error between the estimated surface and the original surface. The functional is modified as shown below.

$$W = \iint_{\Omega_z} \left(\left| z_x - p \right|^2 + \left| z_y - q \right|^2 \right) dxdy + l \iint_{\Omega_z} \left(\left| z_x \right|^2 + \left| z_y \right|^2 \right) dxdy + m \iint_{\Omega_z} \left(\left| z_{xx} \right|^2 + 2\left| z_{xy} \right|^2 + \left| z_{yy} \right|^2 \right) dxdy$$

(5.4)

where the subscripts indicate partial derivatives. In the above cost function, the second term measures small deflection approximation of surface area. The third term is related to surface curvature. l and m are regularization parameters to adjust the weighting of the constraints and are non-negative. The objective now is to solve for the unknown $z(x, y)$ using an optimization process which minimizes the above cost function W. Fourier transforms are employed to solve this problem.

5.1 Theory and derivation

Suppose that the Fourier transform of the surface function $z(x, y)$ is:

$$z_F(u, v) = \iint_{\Omega_z} z(x, y) e^{-j(ux+vy)} dxdy \tag{5.5}$$

and the inverse Fourier transform is

$$z(x, y) = \frac{1}{2p} \iint_{\Omega_F} z_F(u, v) e^{j(ux+vy)} dudv \tag{5.6}$$

where j is the imaginary unit. According to the differentiation properties of the Fourier transform, we have

$$z_x(x, y) \leftrightarrow j\, u\, z_F(u, v) \tag{5.7}$$

$$z_y(x,y) \leftrightarrow j\ v\ z_F(u,v) \tag{5.8}$$

$$z_{xx}(x,y) \leftrightarrow -u^2\ z_F(u,v) \tag{5.9}$$

$$z_{yy}(x,y) \leftrightarrow -v^2\ z_F(u,v) \tag{5.10}$$

$$z_{xy}(x,y) \leftrightarrow -u\ v\ z_F(u,v) \tag{5.11}$$

Let $P(u,v)$ and $Q(u,v)$ be the Fourier transforms of $p(x,y)$ and $q(x,y)$ respectively. Taking the Fourier transform of the cost function and using the above differentiation properties and using Parseval's formula,

$$\iint_{\Omega_Z} |z(x,y)|^2\ dx\ dy = \frac{1}{2p} \iint_{\Omega_F} |z_F(u,v)|^2\ du\ dv, \tag{5.12}$$

we obtain a new objective function to minimize:

$$\begin{aligned}
&\frac{1}{2p} \iint_{\Omega_F} \left[\left| j\ u\ z_F(u,v) - P(u,v) \right|^2 + \left| j\ v\ z_F(u,v) - Q(u,v) \right|^2 \right] du\ dv \\
&+ \frac{l}{2p} \iint_{\Omega_F} \left[\left| j\ u\ z_F(u,v) \right|^2 + \left| j\ v\ z_F(u,v) \right|^2 \right] du\ dv \\
&+ \frac{m}{2p} \iint_{\Omega_F} \left[\left| -u^2\ z_F(u,v) \right|^2 + 2 \left| -u\ v\ z_F(u,v) \right|^2 + \left| -v^2\ z_F(u,v) \right|^2 \right] du\ dv
\end{aligned} \tag{5.13}$$

The above expression can be expanded as

$$\begin{aligned}
&\frac{1}{2p} \iint_{\Omega_F} \left[u^2\ z_F\ z_F^* - j\ u\ z_F\ P^* + j\ u\ z_F^*\ P + P\ P^* + v^2\ z_F\ z_F^* - j\ v\ z_F\ Q^* + j\ v\ z_F^*\ Q + Q\ Q^* \right] du\ dv \\
&+ \frac{l}{2p} \iint_{\Omega_F} \left(u^2 + v^2 \right) z_F\ z_F^*\ du\ dv + \frac{m}{2p} \iint_{\Omega_F} \left(u^4 + 2\ u^2\ v^2 + v^4 \right) z_F\ z_F^*\ du\ dv
\end{aligned} \tag{5.14}$$

where $*$ denotes the conjugate. Differentiating the above expression with respect to z_F and z_F^*, we can deduce the following minimal conditions for the cost function:

$$C_{uv}\, z_F + j u P + j v Q = 0 \tag{5.15}$$

$$C_{uv}\, z_F^* - j u P^* - j v Q^* = 0 \tag{5.16}$$

where $C_{uv} = (1 + l)(u^2 + v^2) + m(u^2 + v^2)^2$. Adding the above two equations together, then subtracting the first one from the second one, results in the following equations:

$$C_{uv}(z_F + z_F^*) + j u(P - P^*) + j v(Q - Q^*) = 0 \tag{5.17}$$

and

$$C_{uv}(z_F - z_F^*) + j u(P + P^*) + j v(Q + Q^*) = 0 \tag{5.18}$$

solving the above equations except for $(u,v) \neq (0,0)$, we obtain

$$z_F(u,v) = \frac{-j u P(u,v) - j v Q(u,v)}{(1+l)(u^2+v^2) + m(u^2+v^2)^2} \tag{5.19}$$

where $(u,v) \neq (0,0)$. Therefore the Fourier transform of the surface is expressed as a function of the Fourier transforms of given gradients $p(x,y)$ and $q(x,y)$. Thus the result may be summarized as:

The cost function W ***is minimized by the surface*** $z(x,y)$ ***whose Fourier transform*** $z_F(u,v)$ ***satisfies*** $z_F(u,v) = \dfrac{-j u P(u,v) - j v Q(u,v)}{(1+l)(u^2+v^2) + m(u^2+v^2)^2}$.

5.2 Algorithm

The proposed method to calculate the depth from gradients uses the transformation as specified above. Final depth values are obtained by performing an inverse Fourier transform. These depth values are then displayed as a depth image. The step by step procedure in execution of this algorithm is presented below.

Algorithm: To compute height from gradient using Fourier transforms

1. **input** gradients $p(x,y)$, $q(x,y)$, l and m
2. **for** $0 \le x, y \le N-1$ do
3. **if** $\left(|\, p(x,y) | < \max_{pq} \& |\, q(x,y) | < \max_{pq} \right)$ **then**
4. Initialize the complex variable $P(x,y) = p(x,y)$;
5. Initialize the complex variable $Q(x,y) = q(x,y)$;
6. **else**
7. Initialize the complex variables to zero
8. **endif**
9. **end for**
10. Calculate the Fourier transforms of $P(x,y) : P(u,v)$
11. Calculate the Fourier transforms of $Q(x,y) : Q(u,v)$
12. **for** $0 \le u, v \le N-1$ **do**
13. **if** $\left(u \ne 0 \,\&\, v \ne 0 \right)$ **then**
14. $\Lambda = \left(1 + l \right) \left(u^2 + v^2 \right) + m \left(u^2 + v^2 \right)^2$;
15. $\Delta 1 = u \times \mathrm{Im}(P(u,v)) + v \times \mathrm{Im}(Q(u,v))$;
16. $\Delta 2 = -u \times \mathrm{Re}(P(u,v)) - v \times \mathrm{Re}(Q(u,v))$;

17. $H1(u,v) = \Delta 1 / \Lambda$;

18. $H2(u,v) = \Delta 2 / \Lambda$;

19. **else**

20. $H1(0,0) = d \quad ; \quad H2(0,0) = 0$;

21. **endif**

22. **end for**

23. Calculate inverse Fourier transforms of $H1(u,v)$ & $H2(u,v) : H1(x,y), H2(x,y)$;

24. **for** $0 \le x, y \le N-1$ **do**

25. $z(x,y) = H1(x,y) + Background\ value$;

26. **end for**

5.3 Implementation details

For any given set of photometric stereo images, the gradient values are computed as explained in section 4.5.1. These gradient values along with the parameters l and m serve as inputs to the algorithm. To avoid any kind of distortion in reconstructed surface, a maximum value of p and q is allowed in the algorithm. For the purpose of experimentation, performance of the algorithm was evaluated on a synthetic range image (obtained from [42]). When the input is a synthetic range image, the corresponding gradient values can be obtained by differentiating the image in x and y directions separately. The resulting gradient values were used as input to the algorithm. The resulting reconstructed surface is also a range image. It should be noted that the source image can be used as the ground truth in the

case of synthetic range images. The final reconstructed image can be compared with that of the ground truth and hence the error surface can be easily found.

The algorithm was tested on both the Beethoven image triplet and synthetic range image. The results obtained are shown in Chapter 7.

Chapter 6

Algorithm 2: Iterative Reconstruction

The previous chapter dealt with the reconstruction of the surface by minimizing the functional shown below by using Fourier transforms.

$$W = \iint_{\Omega} \left(|z_x - p|^2 + |z_y - q|^2 \right) dxdy + l \iint_{\Omega} \left(|z_x|^2 + |z_y|^2 \right) dxdy + m \iint_{\Omega} \left(|z_{xx}|^2 + 2|z_{xy}|^2 + |z_{yy}|^2 \right) dxdy \tag{6.1}$$

It was also seen that the functional had a form which made the application of Fourier transforms feasible. It should be noted that the surface which minimizes the second term on the right hand side will be one which has low values of the derivatives. This is because the term penalizes noise in a non-linear way. Thus the optimization process results in a solution with no sharp edges and thus the images appear blurred. This blurring can be avoided by changing the term to an exponential. In other words the second term on the right hand side can be changed to

$$-l \iint_{\Omega} \exp\left(-\left(z_x^2 + z_y^2\right)\right) dxdy \tag{6.2}$$

Similarly, the third term can be expressed in terms of the exponential as

$$-m \iint_{\Omega} \exp\left(-\left(z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2\right)\right) dxdy \tag{6.3}$$

Thus the new function to be minimized is

$$H = \iint_{\Omega} \left(|z_x - p|^2 + |z_y - q|^2 \right) dxdy - l \iint_{\Omega} \exp\left(-\left(z_x^2 + z_y^2\right)\right) dxdy - m \iint_{\Omega} \exp\left(-\left(z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2\right)\right) dxdy \tag{6.4}$$

Techniques like *Mean Field Annealing* (MFA) [43, 44, and 45] are used to minimize objective functions which result in a restored image that resembles the data (in the least squares sense) while at the same time preserve edges. The introduction of exponential terms makes it difficult for the functional to be represented by Fourier transforms. Instead, the above cost function can be minimized iteratively using gradient descent.

6.1 Using kernels to estimate derivatives

It can be seen from equation 6.4 that derivatives of the surface z should be computed in the process of minimization. This is usually done by using kernels to estimate derivatives. In this work, kernel operators are used to estimate both first and second order derivatives. The kernels used to differentiate in the $x-$ and $y-$ directions are as shown in table 6.1 and table 6.2 respectively.

Table 6.1 Kernels to estimate first derivate

-0.5	0	0.5

g_x **- in the** $x-$ **direction**

-0.5
0
0.5

g_y **- in the** $y-$ **direction**

Table 6.2 - Kernels to estimate second partial derivates

$\frac{1}{\sqrt{6}}$

0	0	0
1	-2	1
0	0	0

g_{xx}

$\frac{1}{\sqrt{6}}$

0	1	0
0	-2	0
0	1	0

g_{yy}

-0.5	0	0.5
0	0	0
0.5	0	-0.5

g_{xy}

As an alternative, 2 X 1 kernels which are more robust to noise can be used to obtain higher resolution.

Thus equation 6.4 in terms of kernel operators can now be written as

$$H = \iint_{\Omega} \left(\left|(z \otimes g_x) - p\right|^2 + \left|(z \otimes g_y) - q\right|^2 \right) dxdy - l \iint_{\Omega} \exp\left(-\left((z \otimes g_x)^2 + (z \otimes g_y)^2\right)\right) dxdy$$
$$- m \iint_{\Omega} \exp\left(-\left((z \otimes g_{xx})^2 + 2(z \otimes g_{xy})^2 + (z \otimes g_{yy})^2\right)\right) dxdy \tag{6.5}$$

6.2 Function minimization

Gradient descent approach can be employed to minimize the function shown in equation 6.5. In the simplest implementation of gradient descent, surface z is updated by

$$z_i^{k+1} = z_i^k - a \frac{\partial H}{\partial z_i} \tag{6.6}$$

where k denotes the iteration, i denotes the pixel on surface z and a is a constant.

While performing gradient descent, most local minima can be avoided by incorporating an *annealing* term, also referred to as the *temperature* in literature, into the functional. This term represented by t is initialized to a large value in the beginning of the process. As the gradient descent iterations are in progress, this annealing term t is slowly reduced. Annealing helps in faster convergence of the algorithm.

The function H containing kernel operators can be differentiated as explained in [43]. In the expression for the differential of H using kernel operators, the exponentials are just added instead of putting the summation in the argument of the exponential since

minimizing either of them results in a piecewise-linear image. Thus differential of H can be written as

$$\nabla_z H = \left(2(z \otimes g_x - p) \right) \otimes g_{xrev} + \left(2(z \otimes g_y - q) \right) \otimes g_{yrev}$$

$$-\frac{l}{t}\left\{ \left(-\frac{1}{t^2}(z \otimes g_x) \exp\left(-\frac{(z \otimes g_x)^2}{2t^2} \right) \right) \otimes g_{xrev} \right\}$$

$$-\frac{l}{t}\left\{ \left(-\frac{1}{t^2}(z \otimes g_y) \exp\left(-\frac{(z \otimes g_y)^2}{2t^2} \right) \right) \otimes g_{yrev} \right\}$$

$$-\frac{m}{t}\left\{ \left(-\frac{1}{t^2}(z \otimes g_{xx}) \exp\left(-\frac{(z \otimes g_{xx})^2}{2t^2} \right) \right) \otimes g_{xxrev} \right\}$$

$$-\frac{m}{t}\left\{ \left(-\frac{1}{t^2}(z \otimes g_{yy}) \exp\left(-\frac{(z \otimes g_{yy})^2}{2t^2} \right) \right) \otimes g_{yyrev} \right\}$$

$$-\frac{m}{t}\left\{ \left(-\frac{1}{t^2}(z \otimes g_{xy}) \exp\left(-\frac{(z \otimes g_{xy})^2}{2t^2} \right) \right) \otimes g_{xyrev} \right\}$$

(6.7)

where t in the above equation is the annealing term and $g_{xrev}, g_{yrev}, g_{xxrev}, g_{yyrev}, g_{xyrev}$ are the reverse kernels of g_x, g_y, g_{xx}, g_{yy} and g_{xy} respectively.

6.3 Algorithm

As explained in the previous section, this algorithm uses gradient descent to minimize the cost function. In each iteration, the surface values are updated. The steps involved in the implementation of this algorithm are as explained below.

Algorithm: To compute height from gradient using gradient descent

1. **Input** gradient images
2. **Initialize** the surface image to a pixel value of one
3. **Initialize** kernels and their respective reverse kernels
4. **Initialize** parameters a, l and m
5. **Set** the initial value of $t = t_{initial}$ (problem dependent parameter)
6. **while** $t > t_{final}$ **do**
7. implement equation number 6.7
8. solve $z^{k+1} = z^k - a \nabla_z H$ (gradient descent)
9. set annealing parameter $t = 0.99t$
10. **end**
11. **Display** the resulting reconstructed surface z

6.4 Implementation details

The above algorithm was implemented for reconstruction of Beethoven image from the existing photometric stereo images. The results obtained are shown in chapter 7. The algorithm was also implemented on some of the range images obtained from [42]. The performance of the algorithm depends heavily on the choice of different parameters viz. a , l , m and the initial temperature $t_{initial}$. These values are specific to the given image and hence these parameters can be considered as problem dependent. Having chosen an optimum $t_{initial}$ for the given image, iterating through the algorithm until t is reduced by two orders of magnitude usually gives good reconstruction. Since this is an iterative process, it was observed that it was much slower than Algorithm1.

Chapter 7

Experimental Results

7.1 Introduction

The description and implementation details of the algorithms were described in chapters 5 and 6. The results obtained for these surface reconstruction procedures are presented in this chapter. In section 7.2, results obtained for the reconstruction of the Beethoven image is presented for both the algorithms. This is followed by a section in which the results obtained for synthetic images are presented. In section 7.4, an experimental set-up used to obtain real world images is discussed and the results obtained in reconstructing these real world images are also presented.

In addition to experimental results, noise-sensitivity of the frequency-domain algorithm i.e. Algorithm 1 was evaluated. This was performed on both the Beethoven image and a synthetic image. Noise sensitivity curves obtained for these images are presented in sections 7.2.1 and 7.3.1 .

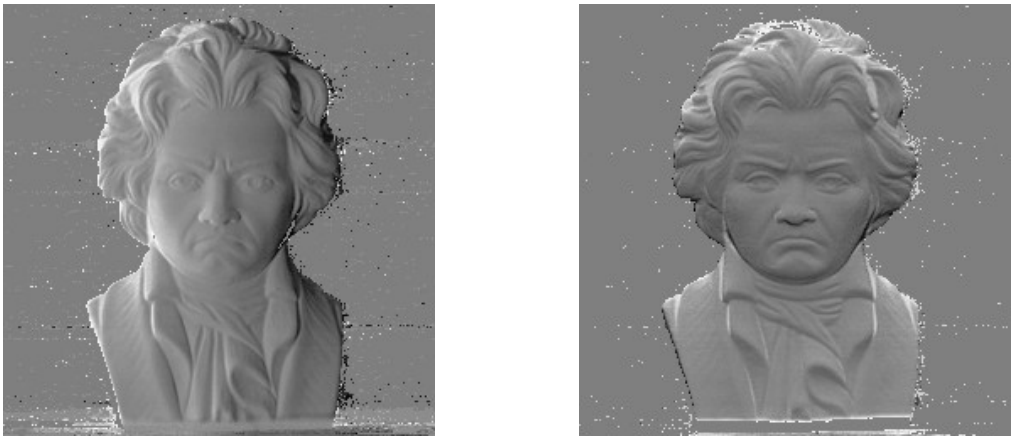
7.2 Beethoven Image

The photometric stereo images used for the algorithm are shown in Figure 7.1. These images were obtained from [41]. The results obtained for surface reconstruction of these images by direct integration method are shown in Figure 4.7.

Beet1.jpg **Beet2.jpg** **Beet3.jpg**

Figure 7.1 Photometric stereo images of Beethoven

The gradient maps obtained for the above photometric stereo images by adopting algorithm 4.5.1 are reproduced below in Figure 7.2.



(a) (b)

Figure 7.2: (a) Gray value image depicting $p(x, y)$
(b) Gray value image depicting $q(x, y)$

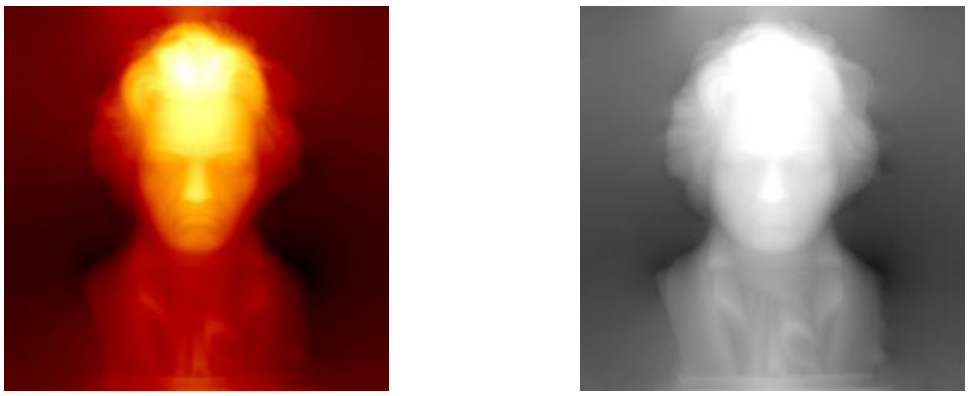
The gradient values obtained were used as input to the algorithms explained in chapters 5 and 6. The results obtained for surface reconstruction form these gradient maps are as shown below.

7.2.1 Algorithm 1: Results

Fourier transform methods were used to reconstruct surfaces in this algorithm. The theory, derivation and other details were explained in chapter 5. It is evident from the algorithm that for a given set of gradient images, the reconstruction process depends on the variables i.e. regularization parameters l and m.

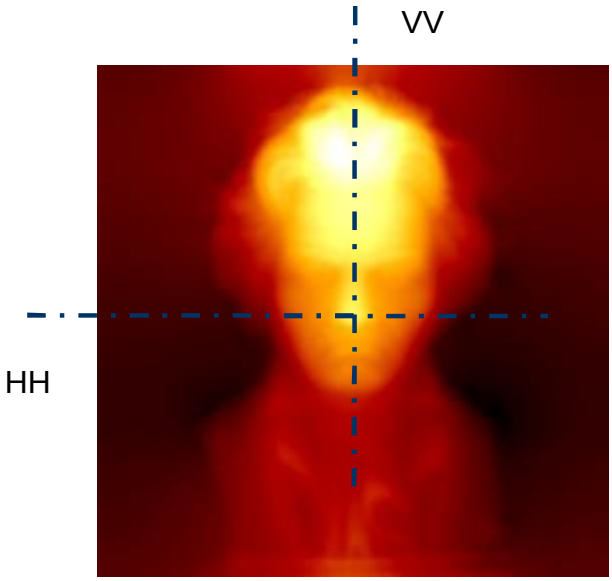
7.2.1.1 Reconstruction Results

The experiment was conducted for varying values of l and m. The reconstruction results are as shown in the figures below. Figure 7.3 shows the results when both l and m are equal to zero. Figure 7.4 shows the results for a fixed m and varying l while Figure 7.5 shows the results for a fixed l and varying m.

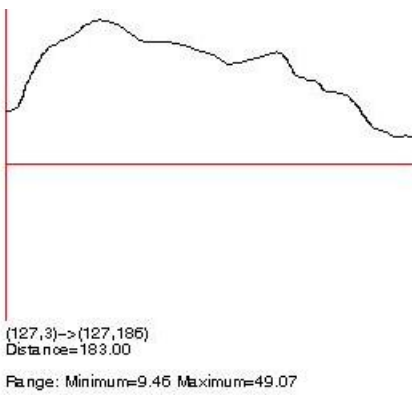


(a) (b)

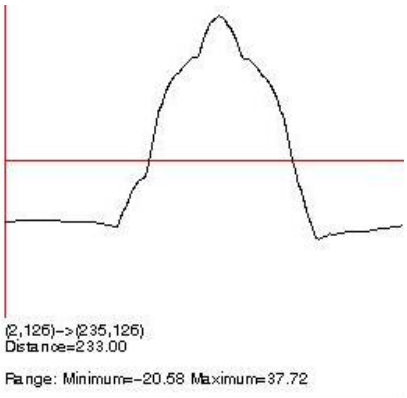
Figure 7.3: Reconstruction of Beethoven image by Algorithm1 with $l = 0$ and $m = 0$
(a) Using Hot-Metal colormap (b) Grayscale representation



(c)

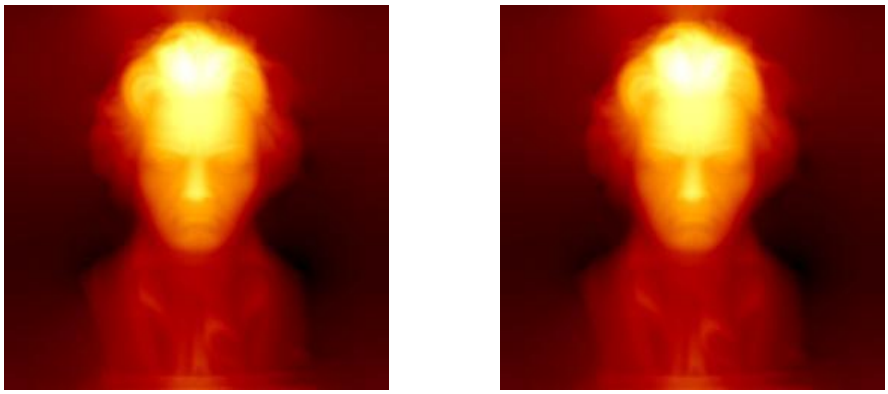


(d)

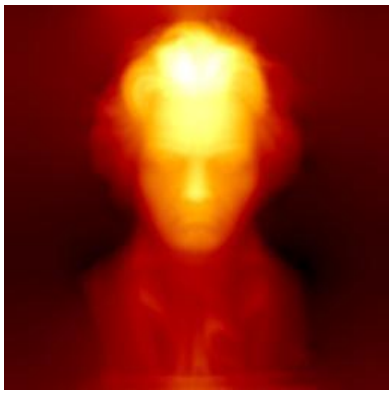


(e)

Figure 7.3: Profile of reconstructed Beethoven image by Algorithm1
(c)Reconstructed image (d) Profile along the VV axis
(e) Profile along the HH axis

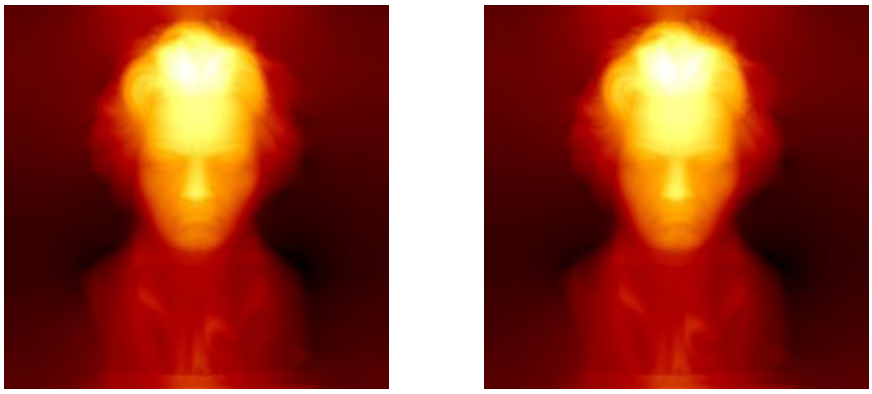


(a) (b)

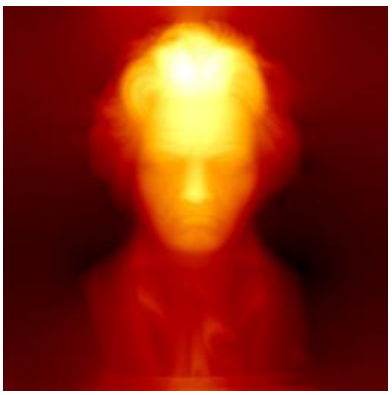


(c)

Figure 7.4: Reconstruction of Beethoven image by Algorithm1 with $l = 0.1$ and
(a) $m = 0.1$ **(b)** $m = 0.3$ **(c)** $m = 0.5$



(a) (b)



(c)

Figure 7.5: Reconstruction of Beethoven image by Algorithm1 with $m = 0$ and
(a) $l = 0.2$ **(b)** $l = 0.3$ **(c)** $l = 0.5$

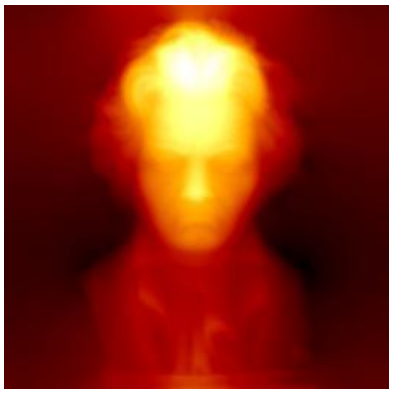


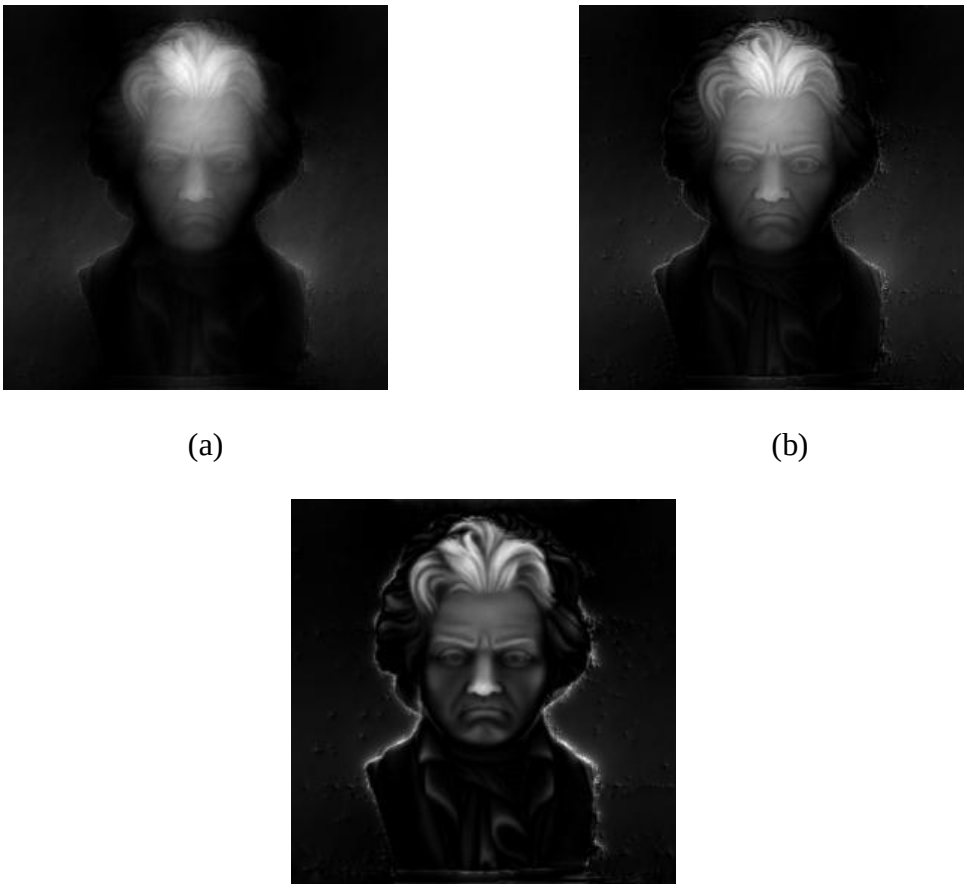
Figure 7.6: Reconstruction of Beethoven image by Algorithm1 with $m = 0.5$ and $l = 0.5$

It can be inferred from the above results that better results are obtained for a lower value of l . All the results shown above were reconstructed from gradient images without additional additive noise. In the next section, the effect of noise (present in the gradient images) on the reconstruction process is considered.

7.2.1.2 Noise Sensitivity

More often than not, the gradient values available to the algorithm for the purpose of reconstruction are corrupted by noise. To simulate this condition, zero-mean Gaussian noise at signal-to-noise (SNR) levels 20, 10 and 0 dB were added to the gradient images shown in Figure 7.2 prior to the process of reconstruction. Since the actual range image of Beethoven was not available for comparison purposes, the noise free reconstruction shown in Figure 7.3 was used as the ground truth. Mean Squared Error (MSE) was used as a measure to quantify the reconstruction error. Error images i.e. images showing the difference between the noisy

reconstruction and the original (ground truth) were constructed. Some of these error images (grayscale) are shown in Figure 7.7 and Figure 7.8



(a) (b)

(c)

Figure 7.7: Error images constructed by taking the difference between the reconstructed images of noisy gradients (SNR = 10dB) and the original reconstructed image of Beethoven by Algorithm1 with $l = 0.1$ and (a) $m = 0.1$ (b) $m = 0.5$ (c) $m = 10.0$

(a)

(b)

(c)

Figure 7.8: Error images constructed by taking the difference between the reconstructed images of noisy gradients (SNR=10dB) and the original reconstructed image of Beethoven by Algorithm1 with $m = 0$ and (a) $l = 0.1$ (b) $l = 0.3$ (c) $l = 0.5$

The mean squared error obtained in the reconstruction process for varying l is tabulated in Table 7.1 and that of varying m is tabulated in Table 7.2

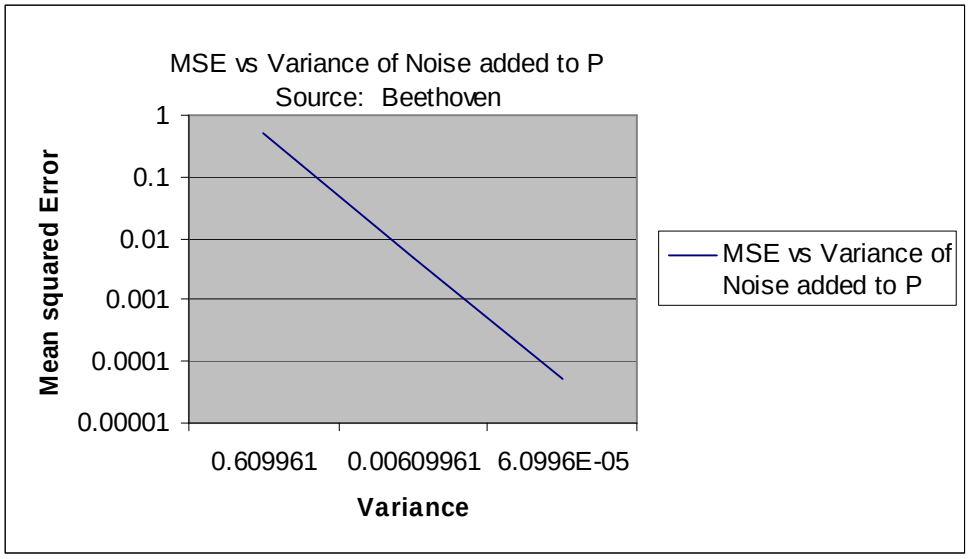
Table 7.1 Mean Squared Error as a function of l and $m = 0$

SNR	l =0	l =0.1	l =0.2	l =0.3	l =0.4	l =0.5
0 db	0.505391	3.613326	10.9498	20.528	31.196	42.2787
10 db	0.005054	3.12373	10.4743	20.0681	30.753	41.851
20 db	0.000051	3.111976	10.4582	20.049	30.732	41.829

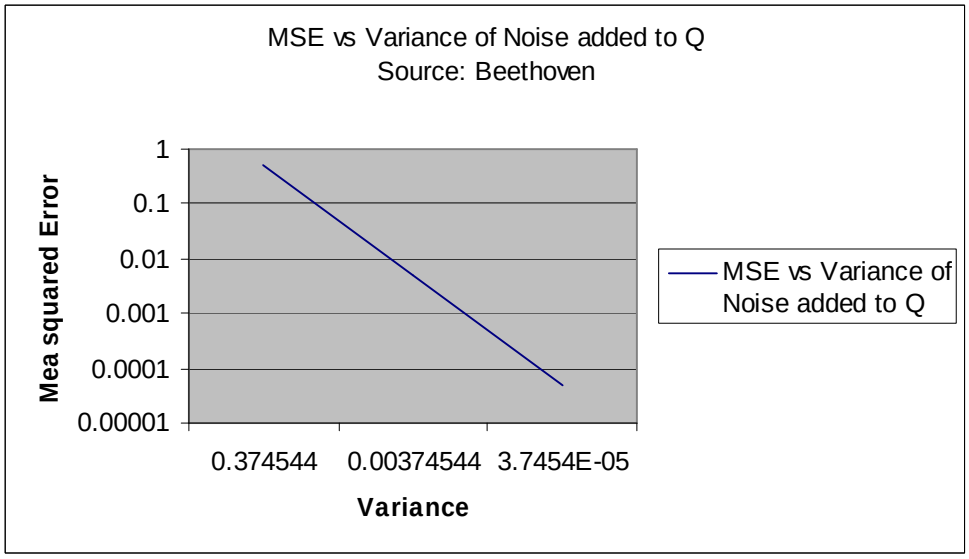
Table 7.2 Mean Squared Error as a function of m

SNR	l	m =0.1	m =0.3	m =0.5	m =2.0	m =10.0
0 db	0.1	3.58583	3.583367	3.5956	3.7545	4.714896
10 db	0.1	3.137	3.1653	3.1933	3.3963	4.409
20 db	0.1	3.1256	3.154	3.1824	3.3856	4.399

In Figure 7.9, graph of the mean squared error as a function of the variance of noise is shown.



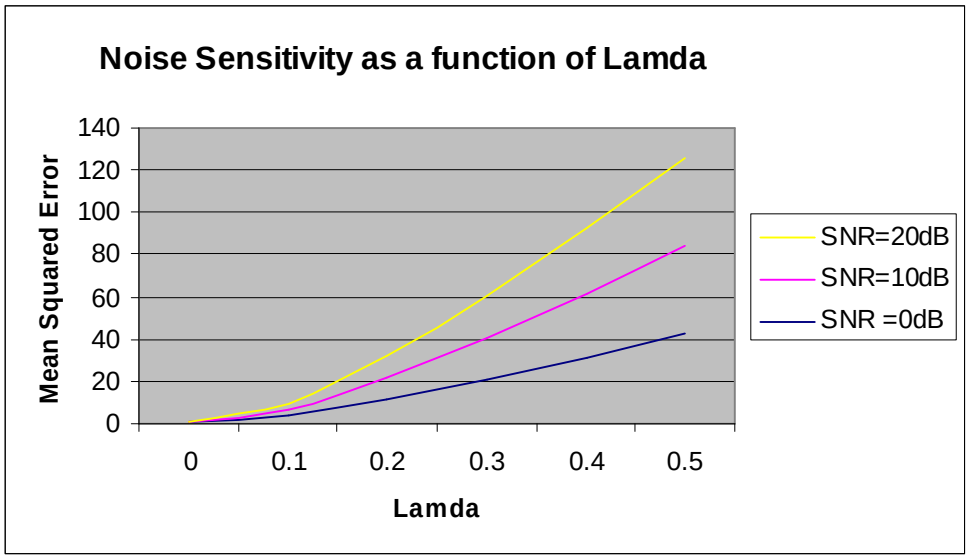
(a)



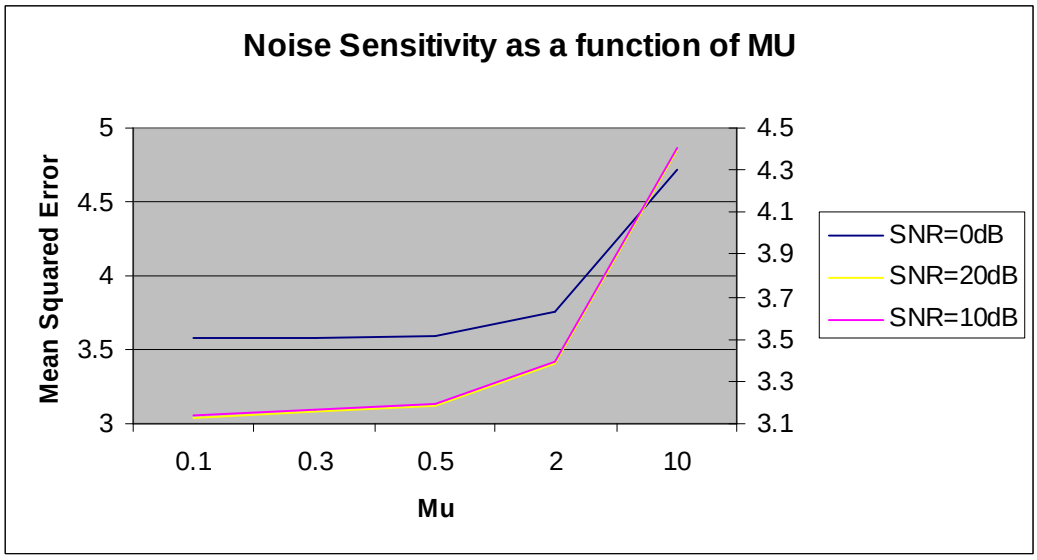
(b)

Figure 7.9 Error as a function of variance of noise added to (a) P (b) Q (Source: Beethoven Image)

In Figure 7.10, noise sensitivity curves are graphed as a function of l and m



(a)



(b)

Figure 7.10 Noise sensitivity as a function of (a) l (b) m (Source: Beethoven Image)

Inference:

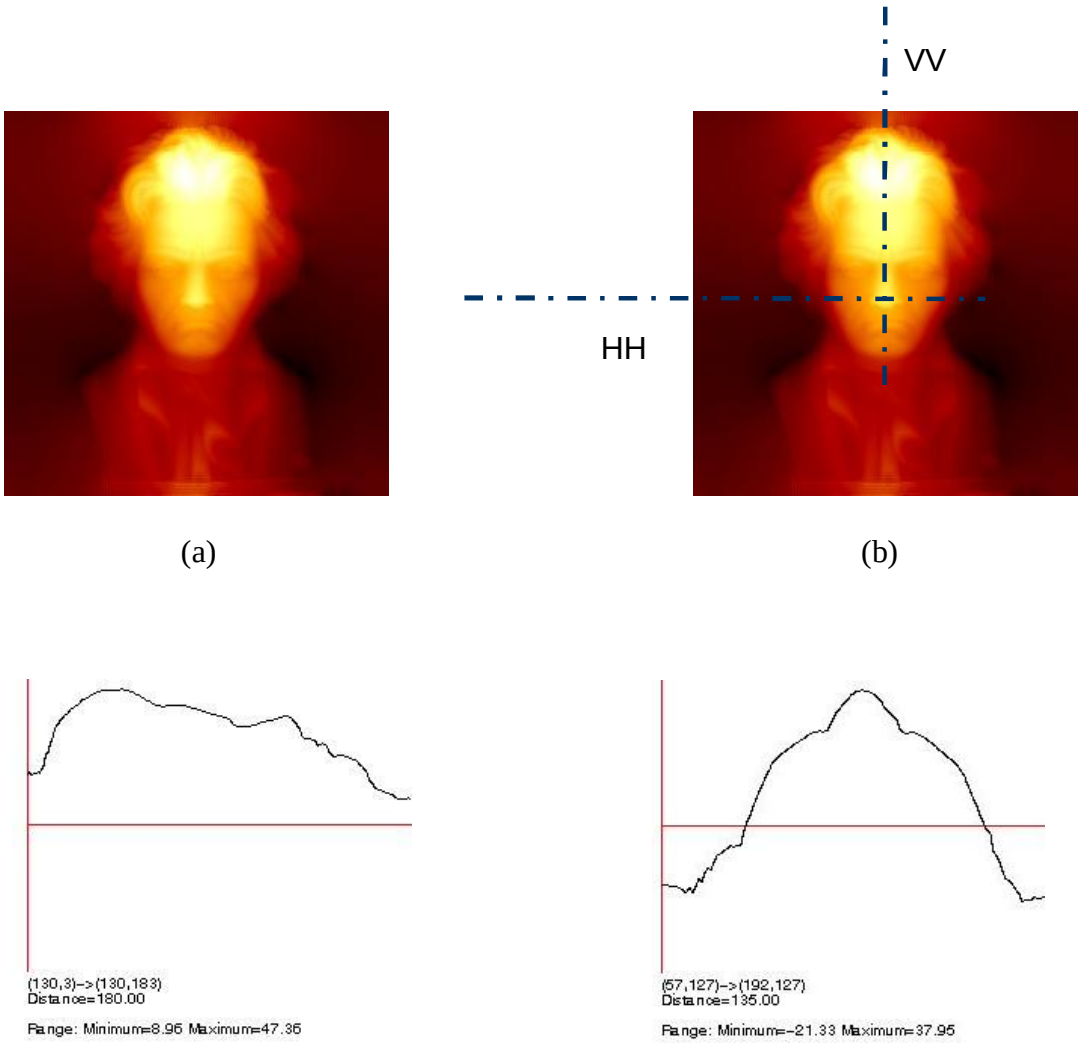
- It can be inferred from Figure 7.9 that the error increases with the amount of noise added to the gradient images. Large amount of noise added to the gradient images give rise to larger errors in the reconstructed images.
- It can be inferred from Table 7.1 and from Figure 7.10(a) that the reconstruction error increases with the increase in the value of l .
- It can be inferred from Table 7.2 and from Figure 7.10 (b) that for higher noise present in the image, the reconstruction error can be decreased by increasing the value of m to a certain extent (as is evident from Table 7.2 for 0dB SNR)

However it should be noted that, since the ground truth was not available for comparing the reconstruction error, the inference made from the Beethoven image only indicates that higher values of l lead to higher reconstruction errors. A more positive and definite conclusion can only be made by analyzing the results on a synthetic image where the ground truth is available for comparison purposes. This is shown in section 7.3.1.2.

7.2.2 Algorithm 2: Results

The methodology used in this algorithm was iterative reconstruction. The cost function was minimized adopting the gradient descent approach. The details of the algorithm itself were presented in chapter 6. In this section, results obtained for the reconstruction of the Beethoven image from the gradient images of Figure 7.2 by this iterative process are presented. Figure 7.11(a) shows one of the reconstructed images by Algorithm2. The profiles along the vertical and horizontal axes are presented in Figure 7.11 (c) and (d).

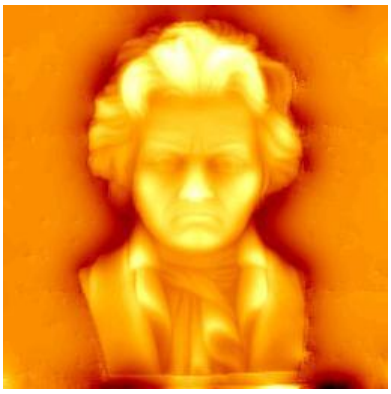
Figure 7.12 shows reconstructed images for varying m and reconstructed images for varying l are shown in Figure 7.13.



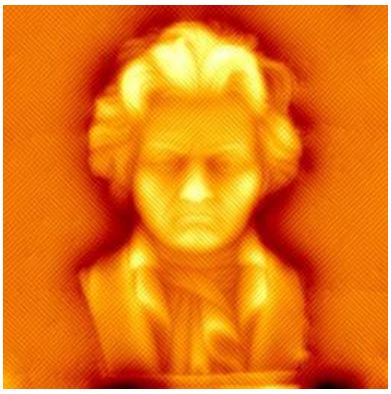
(a) (b)

(c) (d)

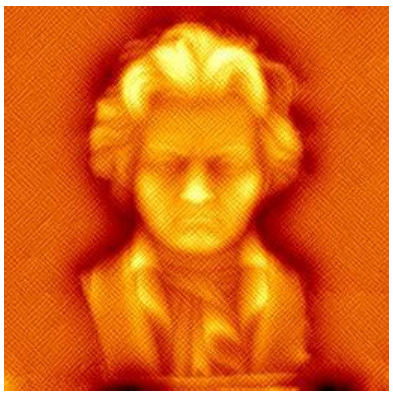
Figure 7.11 (a) Reconstruction of Beethoven image by Algorithm2
(b)Reconstructed image showing profile axes
(c) Profile along the VV axis
(d) Profile along the HH axis



(a)

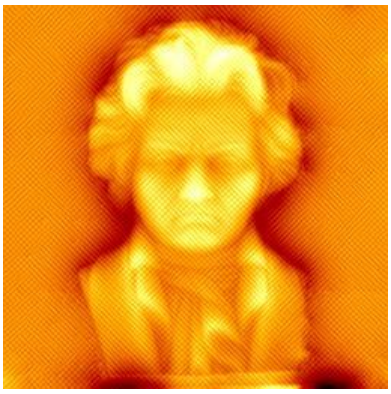


(b)

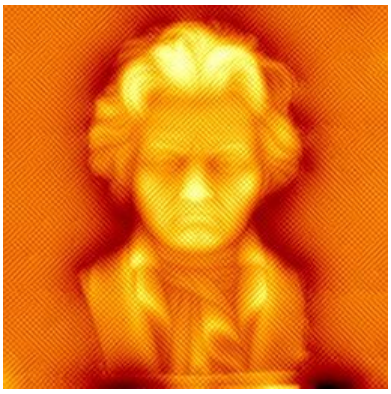


(c)

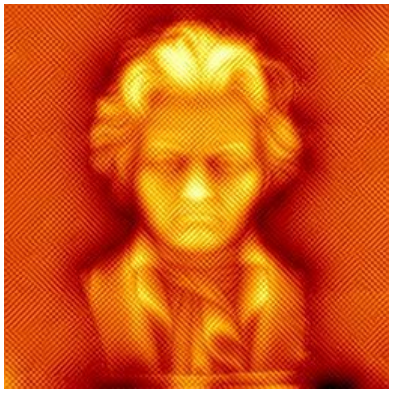
Figure 7.12: Reconstruction of Beethoven image by Algorithm2 with $l = 1.0$ and
(a) $m = 1.0$ **(b)** $m = 5.0$ **(c)** $m = 10.0$



(a)



(b)

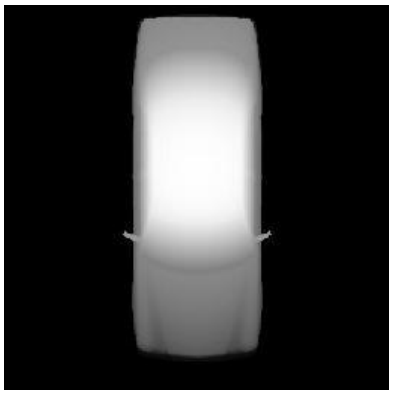


(c)

Figure 7.13: Reconstruction of Beethoven image by Algorithm2 with $m = 5.0$ and
(a) $l = 3.0$ **(b)** $l = 5.0$ **(c)** $l = 10.0$

7.3 Synthetic Images

Synthetic range images used for this part of the experimentation were obtained from [42]. The advantage of using synthetic images is that the ground truth is known apriori to evaluate reconstruction results. One such synthetic range image used to evaluate the noise sensitivity of the frequency domain reconstruction is shown in Figure 7.14 along with the corresponding gradient images constructed from finite differences. Three more synthetic range images were used for this experiment and the reconstruction results are produced in section 7.3.3



(a)

(b)

(c)

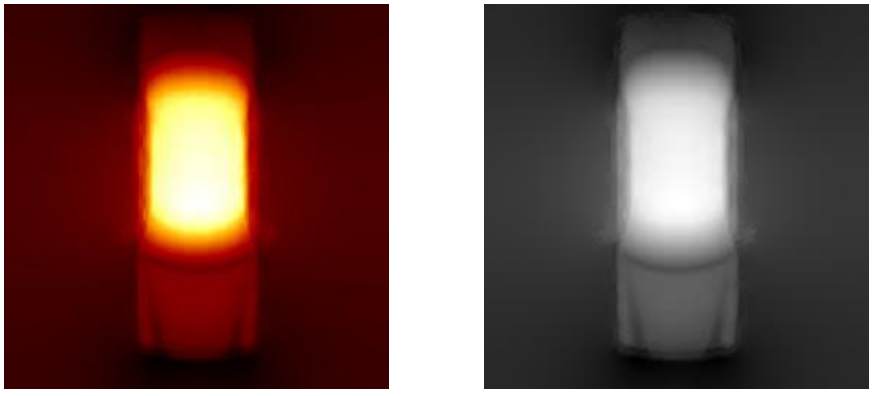
Figure 7.14: (a) Synthetic range image of a car
(b) Gray value image depicting $p(x,y)$ of a). (c) Gray value image depicting $q(x,y)$ of a)

7.3.1 Algorithm 1: Results

As explained in section 7.2.1, Fourier transform methods were used to reconstruct the synthetic range images from their respective gradient images. As in the case of Beethoven, various l and m values were tried. The results obtained were similar to those obtained in section 7.2.1 in that better results were obtained for a smaller value of l . Noise sensitivity was also evaluated for the reconstruction of synthetic range image of Figure 7.14 and the plots are shown in section 7.3.1.2

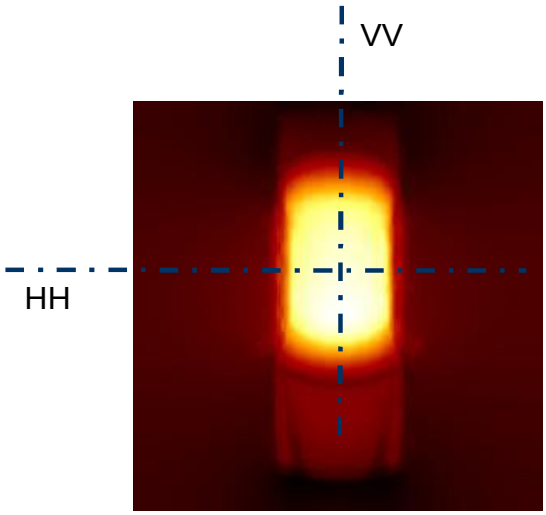
7.3.1.1 Reconstruction Results

The results were obtained for different values of l and m as in the case of Beethoven. The best reconstruction was obtained for $l = 0.1$ and $m = 0.1$. The reconstruction result is as shown in Figure 7.15

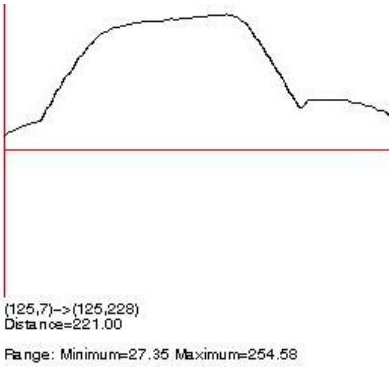


(a) (b)

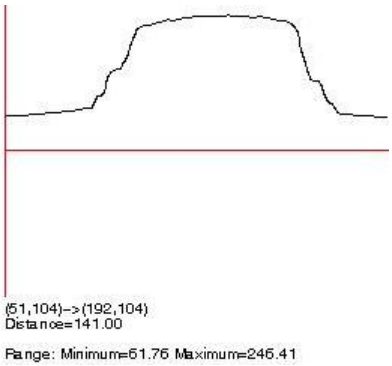
Figure 7.15: Reconstruction of the synthetic car image by Algorithm1 with $l = 0.1$ and $m = 0.1$
(a)HotMetal ColorMap (b) Grayscale



(a)



(b)



(c)

Figure 7.16: Profile of reconstructed synthetic car image by Algorithm1
(a)Reconstructed image (b) Profile along the VV axis
(c) Profile along the HH axis

7.3.1.2 Noise Sensitivity

The same procedure as explained in section 7.2.1.2 was followed to evaluate the noise sensitivity of the Fourier domain for synthetic images. The original range image shown in Figure 7.14 was used as ground truth. Zero mean Gaussian noise at SNR levels 20, 10 and 0 dB was added to the gradient maps of Figure 7.14 for evaluation purposes. As in section 7.2.1.2, mean squared error (MSE) was used as a measure to quantify the reconstruction error. Error images i.e. images showing the difference between the noisy reconstruction and the original (ground truth) were also constructed.

It should be noted that since the ground truth is available in the form of the original range image, the noise sensitivity plots obtained for the synthetic range image are more realistic. These noise sensitivity plots obtained for the synthetic image of Figure 7.14 are shown in Figure 7.17 and Figure 7.18. The MSE values obtained for different levels of SNR for varying l and m are tabulated in Table 7.3 and Table 7.4.

The mean squared error obtained in the reconstruction process of the synthetic range image of Figure 7.14 for varying l is tabulated in Table 7.3 and that of varying m is tabulated in Table 7.4

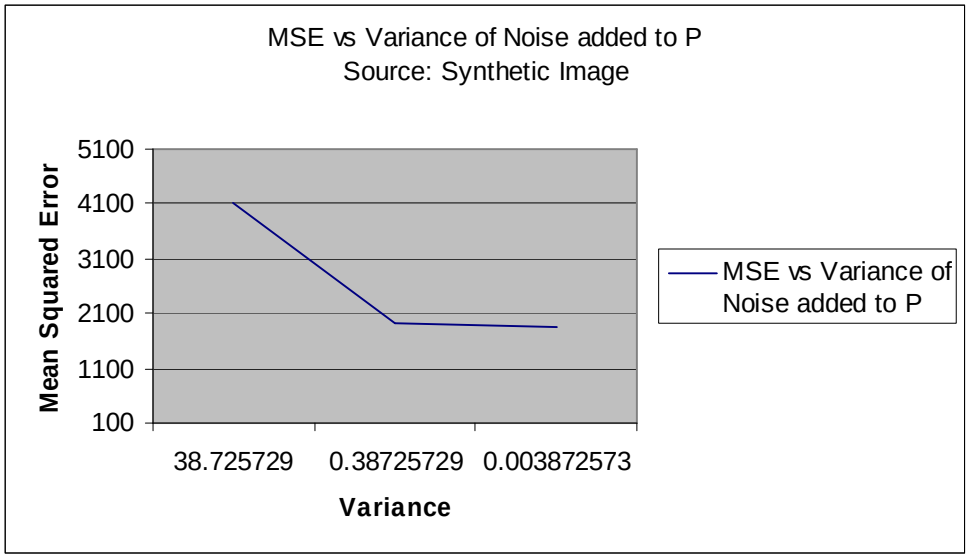
Table 7.3 Mean Squared Error as a function of l and $m = 10$

SNR	l =0	l =0.1	l =0.2	l =0.3	l =0.4	l =0.5
0 db	4102.66	4126.39	4148.04	4167.42	4184.78	4200.44
10 db	1912.25	1922.55	1931.83	1940.08	1947.47	1954.13
20 db	1854.77	1864.17	1872.49	1879.89	1886.5	1892.43

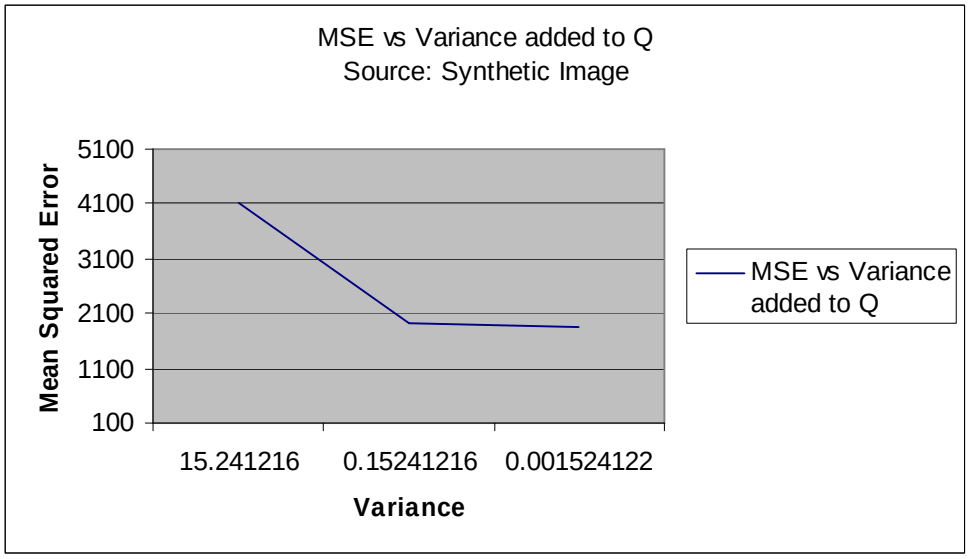
Table 7.4 Mean Squared Error as a function of m

SNR	l	m =0.1	m =0.3	m =0.5	m =2.0	m =10.0
0 db	0.1	4690.48	4589.89	4558.08	4421.97	4126.39
10 db	0.1	2194.84	2172.77	2156.62	2075.613	1922.55
20 db	0.1	2122.73	2103.06	2087.94	2010.92	1864.17

In Figure 7.17, graph of the mean squared error as a function of the variance of noise added to the synthetic image is shown.



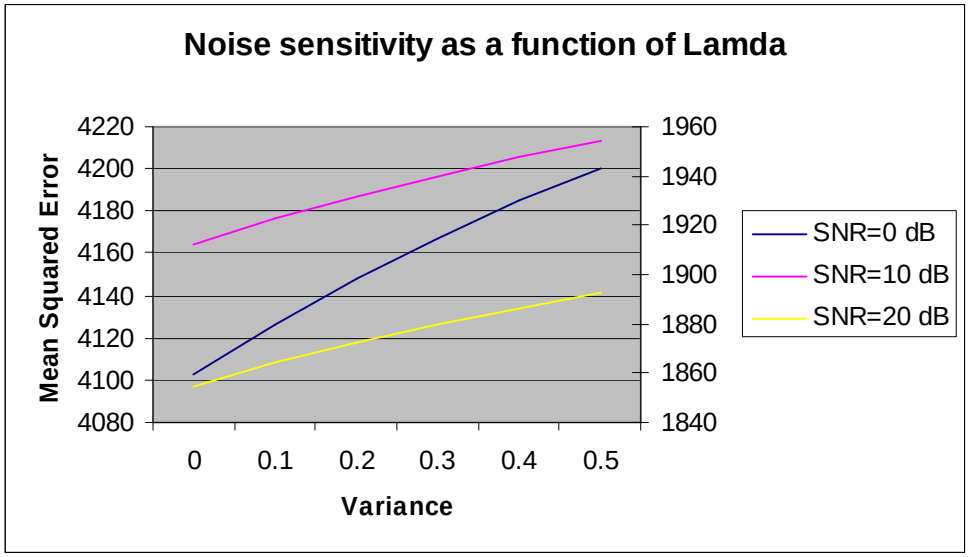
(a)



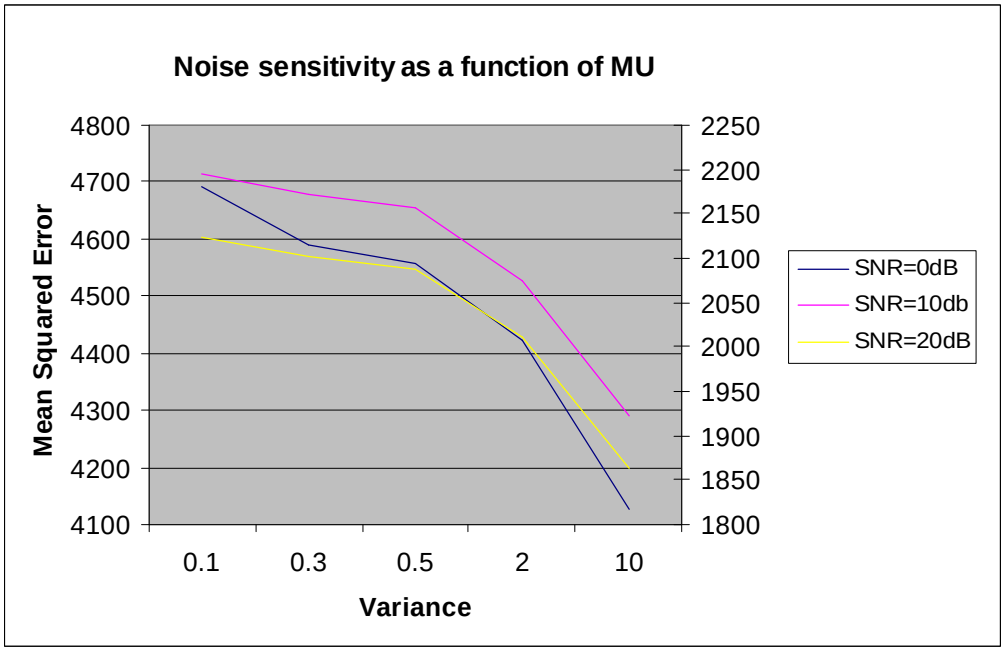
(b)

Figure 7.17 Error as a function of variance of noise added to (a) P (b) Q (Source: Synthetic Image)

In Figure 7.18, noise sensitivity curves are graphed as a function of l and m



(a)



(b)

Figure 7.18 Noise sensitivity as a function of
(a) l (b) m (Source: Synthetic Image)

Inference:

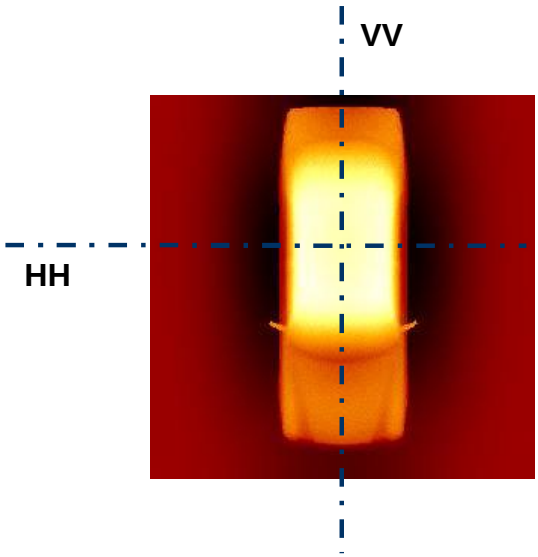
- It can be inferred from Figure 7.17 that the error increases with the amount of noise added to the gradient images as was inferred from the Beethoven images.
- From Figure 7.18 (a), it can be again inferred that the reconstruction error increases with the increase in the value of l . This can be attributed to the fact that higher values of l tend to blur the image and hence give rise to more inaccurate results.
- From Figure 7.18 (b), it can be inferred that higher values of m give rise to better reconstruction results and hence lower reconstruction errors.

7.3.2 Algorithm 2: Results

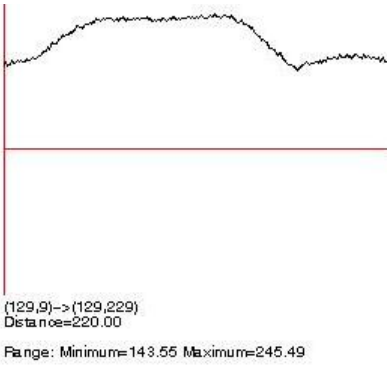
As explained in section 7.2.2, the methodology used in this algorithm was iterative reconstruction and the cost function was minimized by adopting the gradient descent approach. The details of the algorithm itself were presented in chapter 6. The algorithm was implemented for various values of l and m . The best reconstruction result obtained for this algorithm is shown in Figure 7.19.

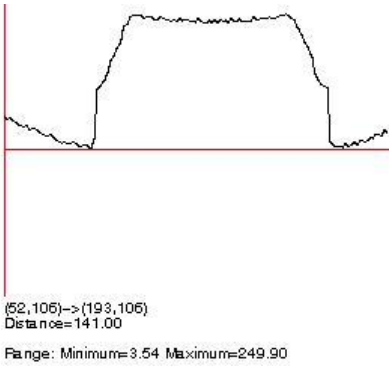
Mean Squared Error (MSE) between the original source image (Figure 7.14) and the reconstructed image (Figure 7.19) was computed. This reconstructed error was then graphed as a function of the parameters of the algorithm (viz. temperature). Figure 7.20 shows the graph of reconstruction error as a function of temperature. Table 7.5 and Table 7.6 show the values of reconstruction error as a function of l and m respectively.

(a)



(b)





(c) **(d)**

Figure 7.19 (a)Reconstruction of synthetic image by Algorithm2 ($l = 1.0$ and $m = 10.0$) (b)Profile Axes on reconstructed image (c)Profile along the VV axis (d) Profile along the HH axis

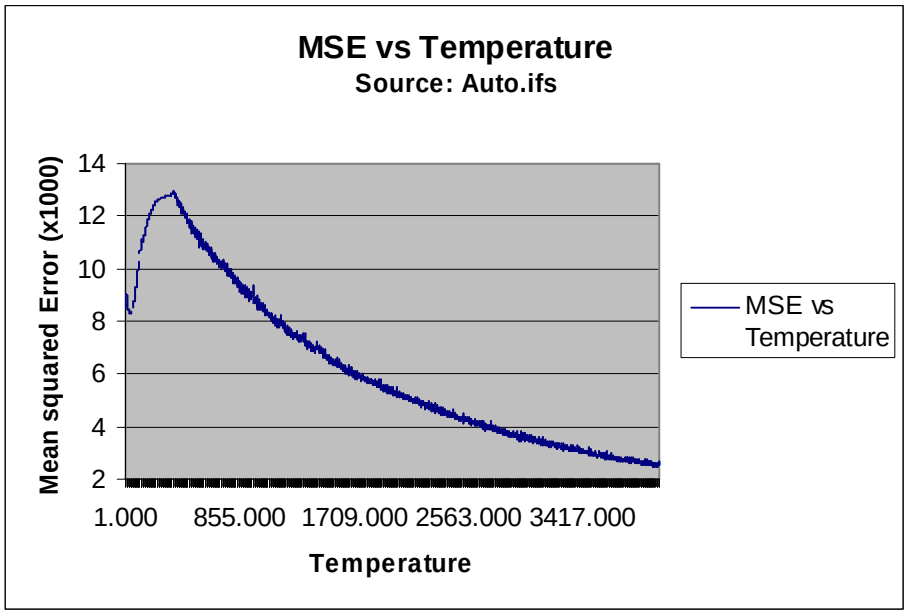


Figure 7.20 Graph of reconstruction error as a function of Temperature ($l = 1.0$ and $m = 10.0$) (source: Auto.ifs)

Table 7.5 Mean Squared Error as a function of l for Algorithm2 (Source: Auto.ifs)

$l \quad (m = 10.0)$	**MSE**
0.5	**420**
1.0	**301**
5.0	**2193**

Table 7.6 Mean Squared Error as a function of m for Algorithm2 (Source: Auto.ifs)

$m \quad (l = 1.0)$	**MSE**
10.0	**301**
20.0	**670.76**
50.0	**1580.91**

It is clear from Figure 7.20, as the iterative process proceeds, the temperature decreases and the reconstruction error decreases as the temperature decreases. It can be inferred from Table 7.5 and Table 7.6 that there is an optimum value for l and m for which the algorithm produces a minimum reconstruction error. Also, these values are problem dependent and need to be evaluated for every input image.

Comparison of Algorithm1 and Algorithm2:

The performance of both the algorithms can be evaluated by comparing the Mean Squared Error (MSE) obtained for each of these algorithms. The mean squared error obtained between the source image (Figure 7.14) and the images reconstructed by each of these algorithms, i.e. Algorithm1 (Figure 7.15) and Algorithm2 (Figure 7.19), were computed for different values of l and m. The results obtained are tabulated in Table 7.7 below.

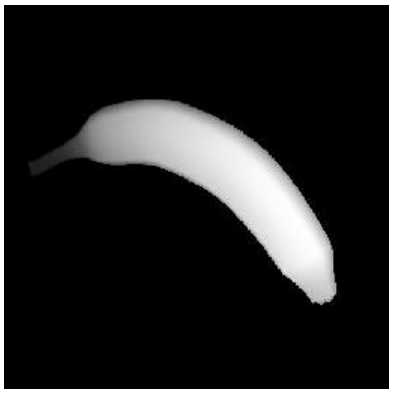
Table 7.7 MSE obtained for Algorithm1 and Algorithm2 (Source: Auto.ifs)

ALGORITHM1			**ALGORITHM2**		
l	m	**MSE**	l	m	**MSE**
0.1	**0.1**	**2189.21**	**0.5**	**10.0**	**420.3**
0.1	**10.0**	**1929.75**	**1.0**	**10.0**	**301.1**
0.1	**20.0**	**1870.09**	**5.0**	**10.0**	**2193**
0.5	**10.0**	**1958.07**	**1.0**	**20.0**	**670.76**
1.0	**10.0**	**1985.57**	**1.0**	**50.0**	**1580.91**

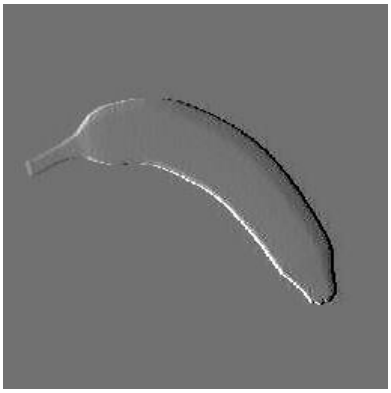
It can be seen from Table 7.7 that, the MSE obtained for the iterative method i.e. Algorithm2 is much lower than the MSE obtained for the Fourier transforms method i.e. Algorithm1. Thus it can be inferred that Algorithm2 produces more accurate reconstruction than Algorithm1.

7.3.3 More synthetic images

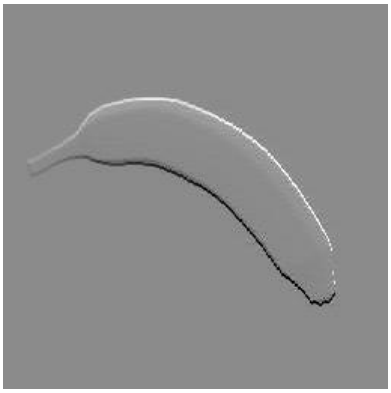
Algorithm1 and Algorithm2 were implemented on more synthetic range images obtained from [42]. The results of both the algorithms are shown in Figure 7.21 to Figure 7.23. However it should be noted that, if the object to be reconstructed has flat surfaces, the gradients of such surfaces are zero. Thus little information is available to the algorithms to reconstruct the surface. This leads to a poor reconstruction of the object as shown in Figure 7.22. It is evident in this figure that surfaces that make the body of the airplane are almost flat which means that the gradients at these surfaces are zero. Since not much information is available for the reconstruction process, the result obtained is rather poor.



(a)

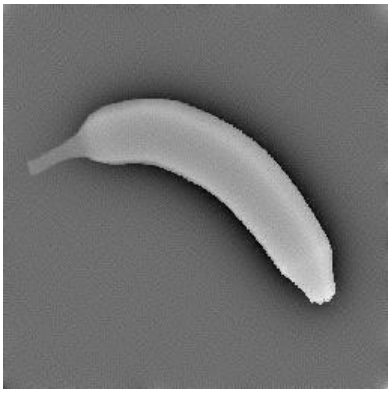


(b)



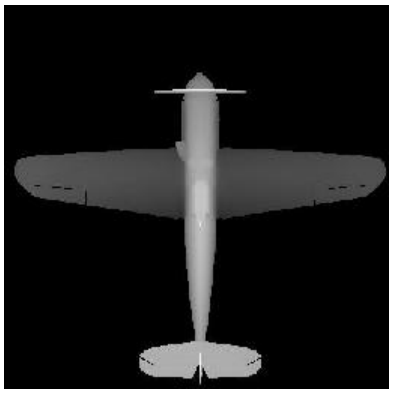
(c)

(d)

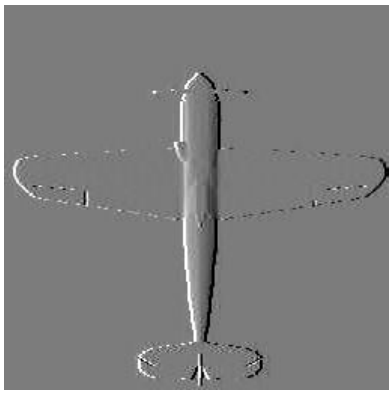


(e)

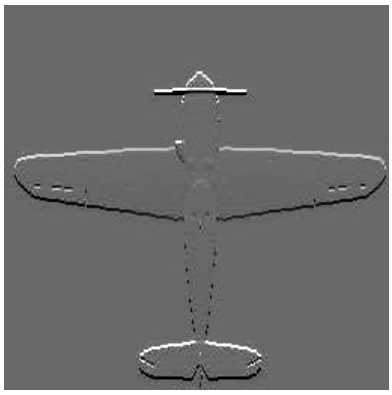
Figure 7.21 (a) Source Image (b) Gradient map $P(x, y)$
(c) Gradient Map $Q(x, y)$
(d) Reconstruction by Algorithm1
(e) Reconstruction by Algorithm2



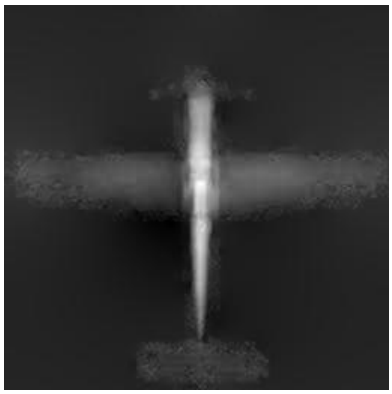
(a)



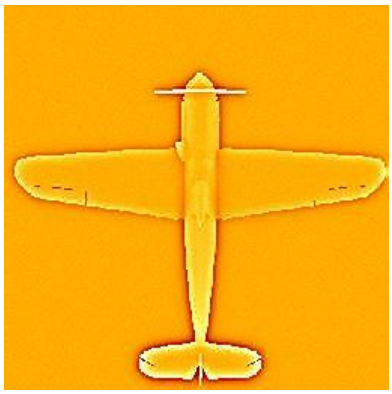
(b)



(c)

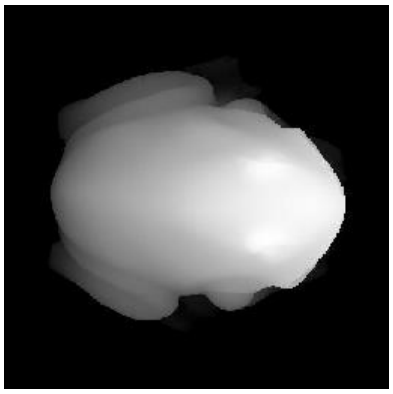


(d)



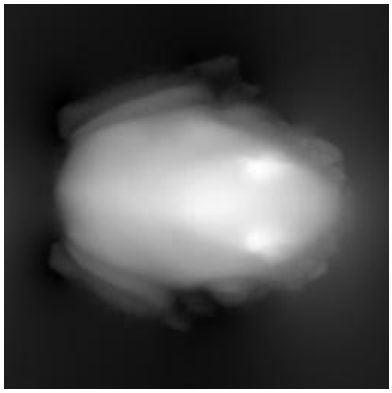
(e)

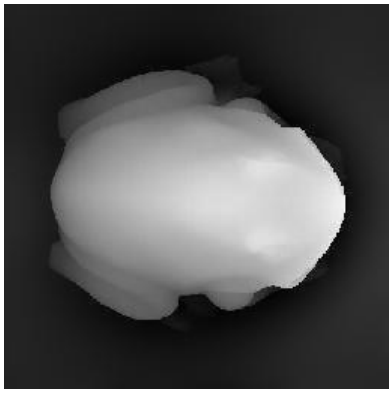
Figure 7.22 (a) Source Image (b) Gradient map $P(x, y)$
(c) Gradient Map $Q(x, y)$
(d) Reconstruction by Algorithm1
(e) Reconstruction by Algorithm2



(a)

(b) (c)

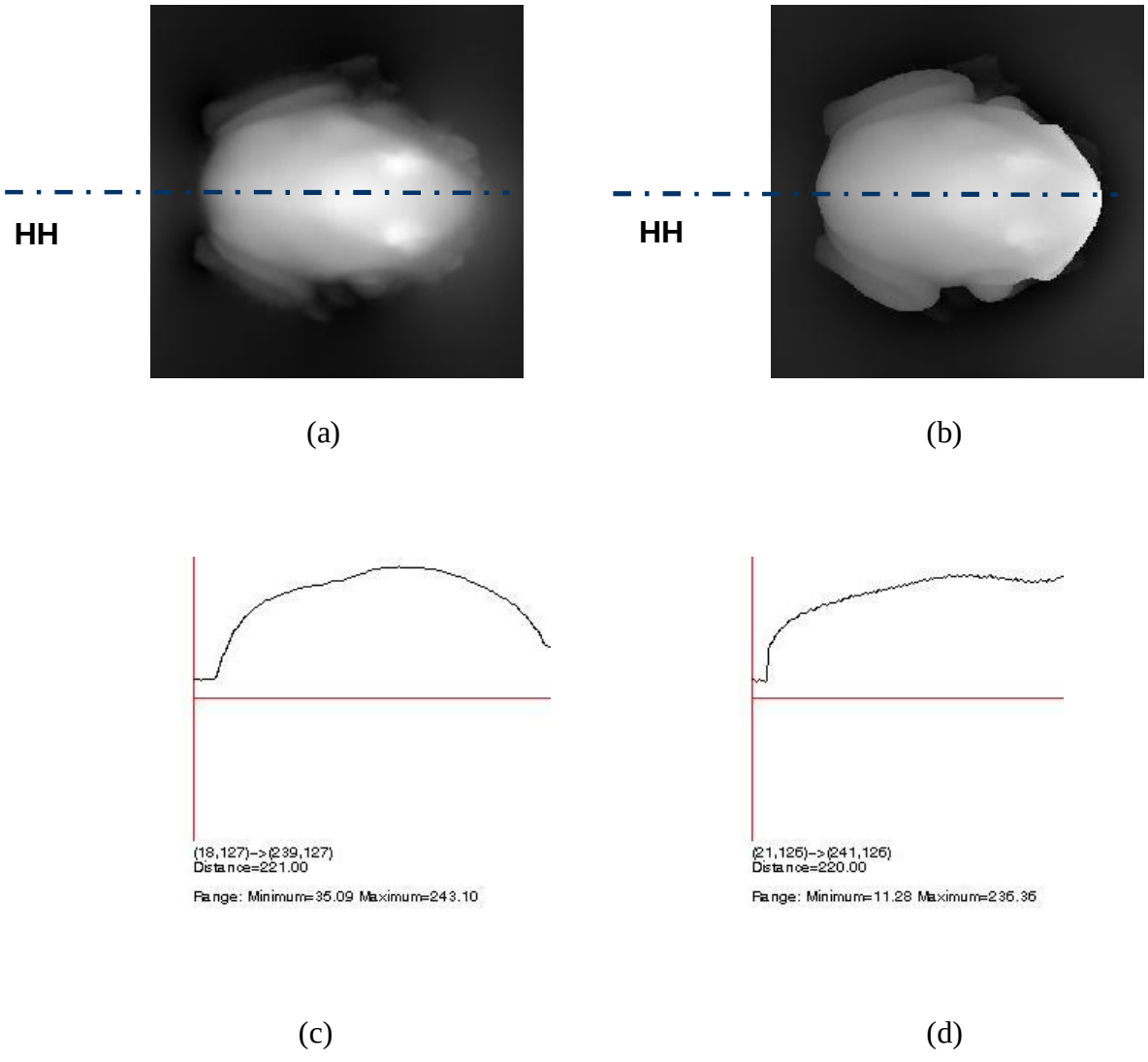




(d) (e)

Figure 7.23 (a) Source Image (b) Gradient map $P(x, y)$
(c) Gradient Map $Q(x, y)$
(d) Reconstruction by Algorithm1
(e) Reconstruction by Algorithm2

In Figure 7.24 below, the profile along the HH axis of the reconstructed synthetic frog image for both the algorithms is shown.



(a) (b)

(c) (d)

Figure 7.24 (a) Reconstructed image of frog by Algorithm1 showing profile axis.
(b) Reconstructed image of frog by Algorithm2 showing profile axis.
(c) Profile along HH axis for Algorithm1
(d) Profile along HH axis for Algorithm2

In the process of reconstruction, the reconstruction error i.e. MSE between the original source images and the reconstructed images was computed. For Algorithm2 (iterative method), this reconstructed error was then graphed as function of the temperature of the algorithm. The graphs obtained for reconstruction error as a function of temperature for each of the synthetic images of Figure 7.21-Figure 7.23 are shown in Figure 7.25-Figure 7.27.

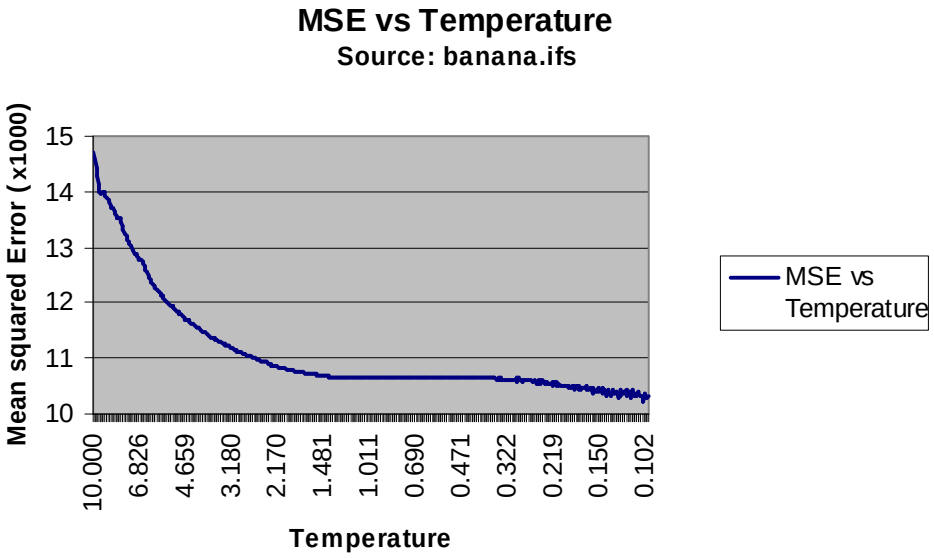


Figure 7.25 Graph of MSE as a function of temperature (Source: banana.ifs)

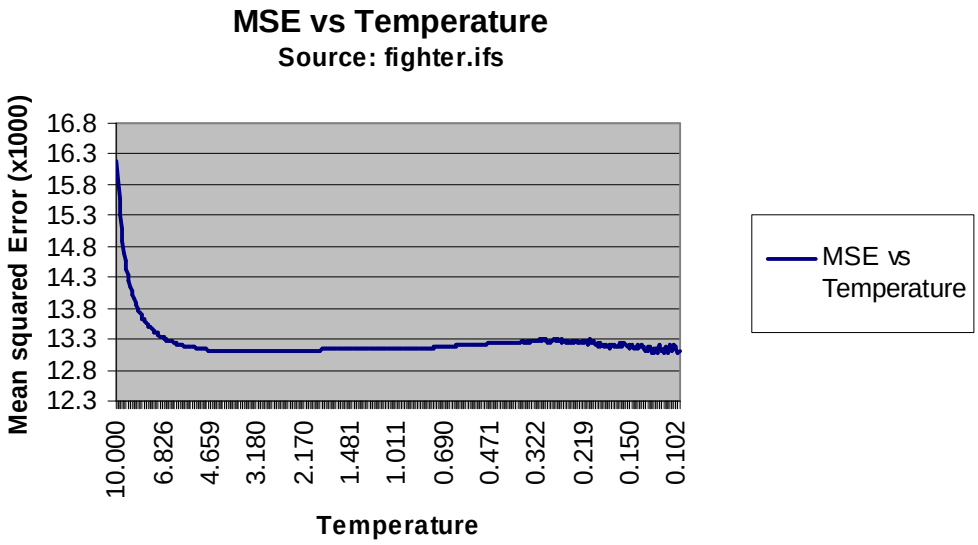


Figure 7.26 Graph of MSE as a function of temperature (Source: fighter.ifs)

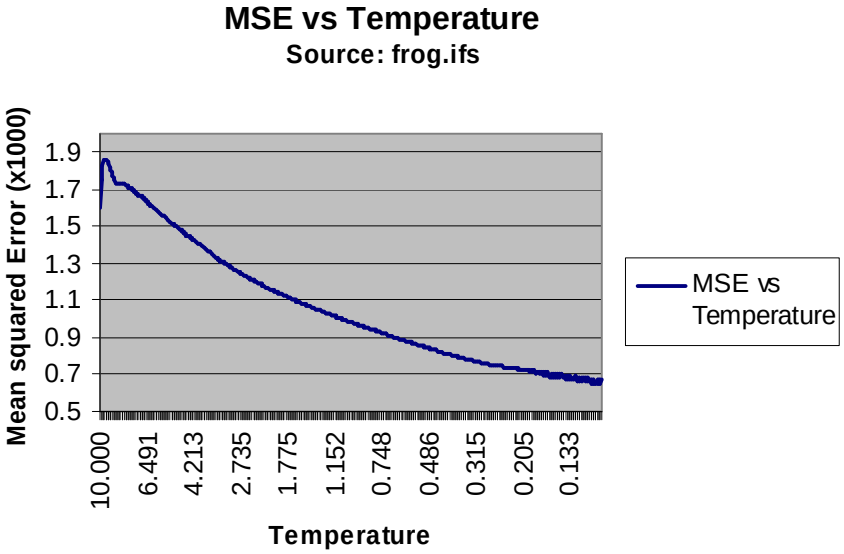


Figure 7.27 Graph of MSE as a function of temperature (Source: frog.ifs)

Comparison of Algorithm1 and Algorithm2:

It should be noted that when an image has a discontinuity, reconstruction cannot be performed by direct integration. Special techniques need to be developed to get perfect surface reconstruction of objects with discontinuities. One such technique is presented in [37] where the author relaxes the uniform integrability condition to partial integrability conditions. However, in the presence of discontinuities, the performance of the algorithms can be evaluated by comparing their MSE values. Among the algorithms described above, Algorithm2 gives better results when compared to Algorithm1 in the presence of discontinuities as is evident from Figure 7.22. The reconstruction error (MSE) obtained for both Algorithm1 and Algorithm2 in the process of reconstruction of synthetic images of Figure 7.21 – Figure 7.23 is tabulated in Table 7.8 below.

Table 7.8 MSE obtained for Algorithm1 and Algorithm2 in reconstruction of synthetic images

Source	MSE obtained for Algorithm1	MSE obtained for Algorithm2
Banana.ifs	2791.37	902.28
Fighter.ifs	1635.39	1592.55
Frog.ifs	1233.809	104.96

From the above table, it is clear that the reconstruction error obtained for Algorithm2 is lower than the reconstruction error obtained for Algorithm1. It can thus be inferred that Algorithm2 produces better reconstruction than Algorithm1 even in the presence of discontinuities.

7.4 Real World Images

The experiments were conducted on some real world images obtained from the imaging lab. As explained in chapter 2, the idea of photometric stereo is to vary the direction of the incident illumination between successive views while holding the viewing direction constant. This is achieved by having a constant viewing direction and using different non-collinear light sources which are powered one at a time. Since there has been no change in imaging geometry, each picture element (x, y) in the three images correspond to the same object point and hence to the same gradient (p, q). The effect of varying the direction of incident illumination is to change the reflectance map $R(p, q)$ that characterizes the imaging situation. Thus, it becomes necessary to capture three (or more) images of an object, each image corresponding to one light source. In section 7.4.1, an outline of the experimental setup is presented and in section 7.4.2, some real world images and the results obtained by reconstruction algorithms is presented.

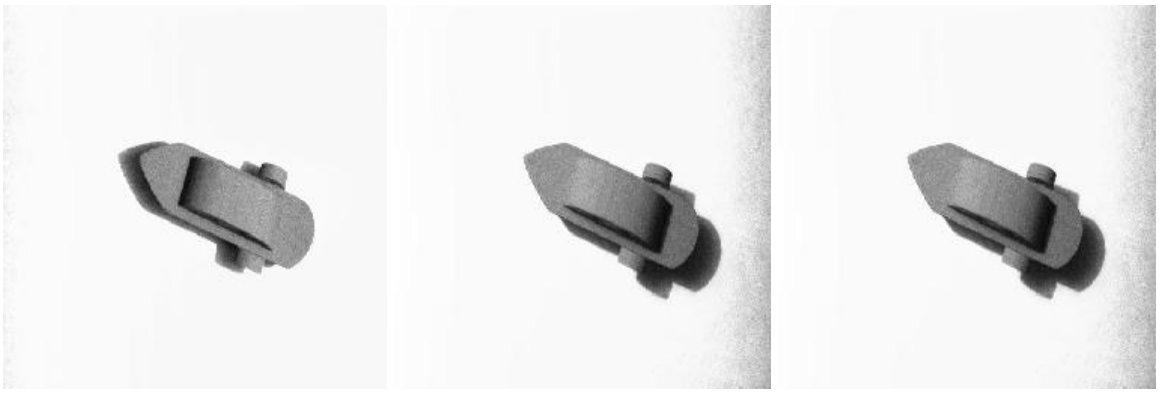
7.4.1 Experimental Setup

A system was setup in the imaging lab to obtain real world images to aid in this process of experimentation. A large wooden box was installed and was covered on the inside with black matte sheets to avoid any reflections on the object by the sides of the box. Three non collinear point light sources were mounted inside the box. A camera was mounted on the ceiling of the box with the lens pointing directly on the object. The objects chosen were such that their surface could be modeled as Lambertian reflectors since the algorithms are based on Lambert's cosine law.

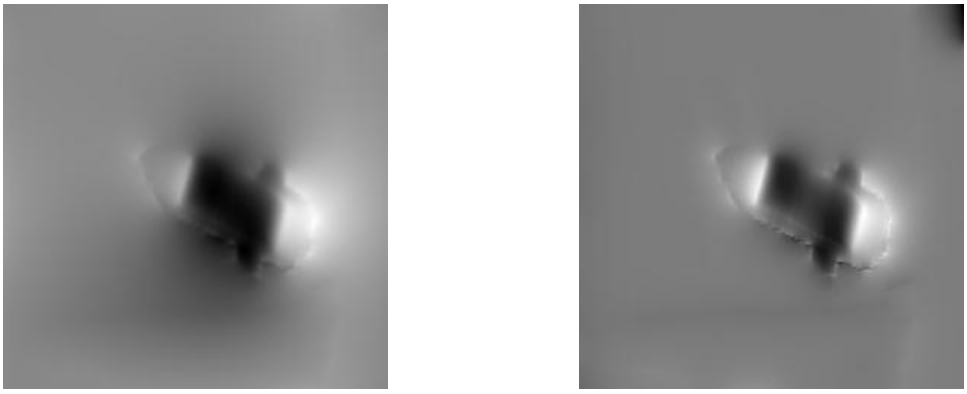
Since both the camera and the light sources were stationary, any error in measurement of the distances between the object and the camera and also the object and the light sources was thereby reduced. Care was taken to see that the camera axis and the z axis of the object coincided with each other. The distance between the object and the light source was such that every point on the object could be considered equidistant from a particular light source. Three images were captured – one image at a time corresponding to a particular light source. Distances of each light source from the object were measured. All this data was used as input to the algorithm 4.5.1 to generate gradient images by photometric stereo. These gradient images were used as input to the algorithms to reconstruct surfaces. Some real world images obtained from this setup are presented in the next section.

7.4.2 Experimental images and results

Some of the images obtained from the experimental setup and the reconstructed images are shown in Figure 7.28 to Figure 7.30

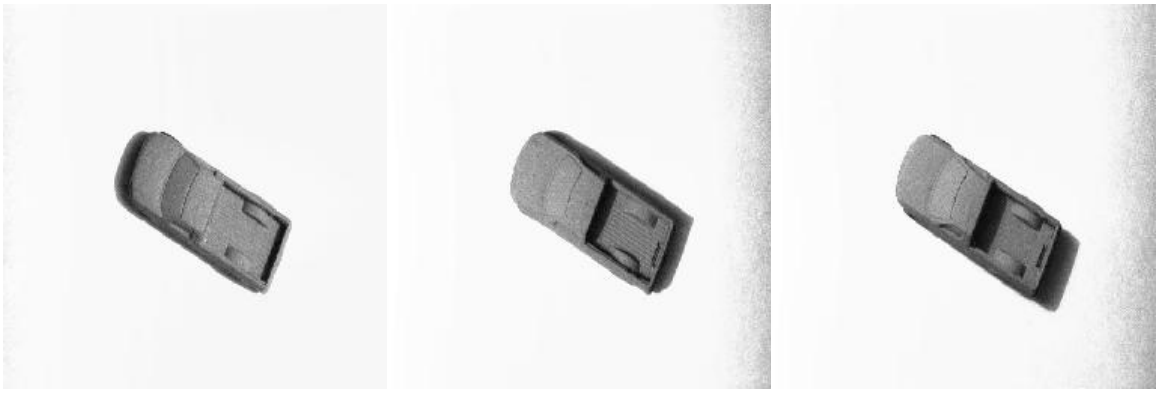


(a) (b) (c)

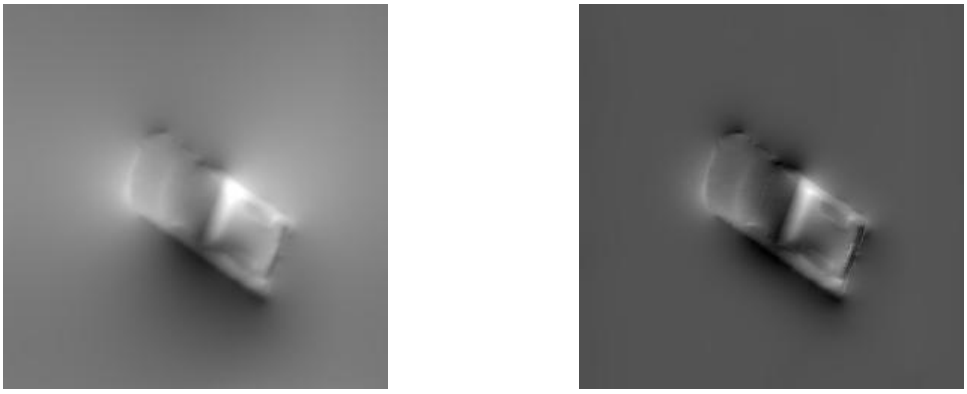


(d) (e)

Figure 7.28 (a), (b), and (c) Photometric stereo images of a real object, each obtained from a different light source
(d) Reconstructed image by Algorithm1
(e) Reconstructed image by Algorithm2

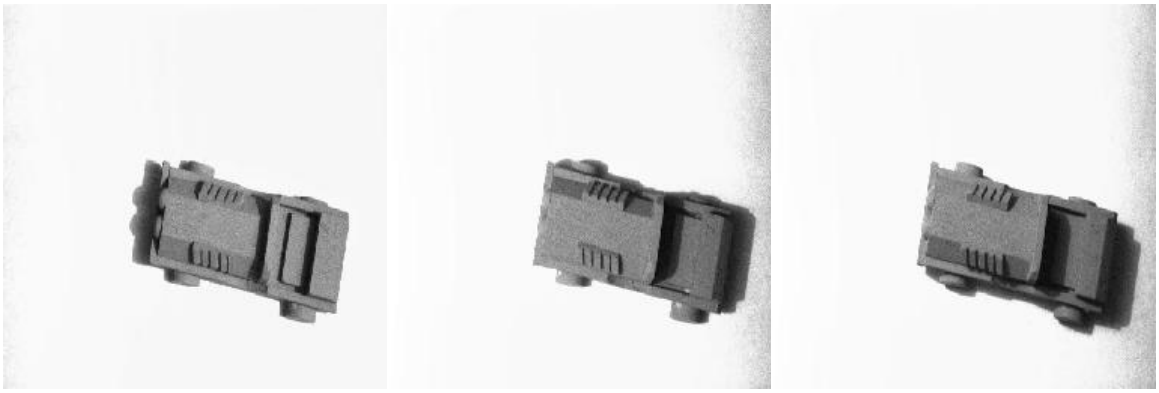


(a) (b) (c)

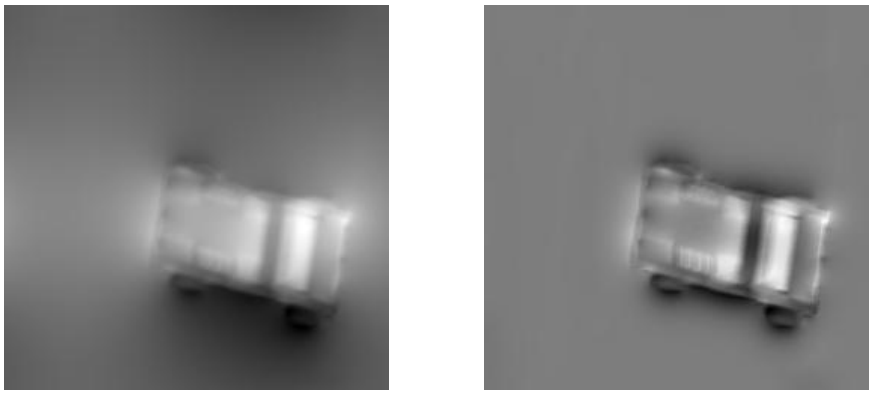


(d) (e)

Figure 7.29 (a), (b), and (c) Photometric stereo images of a real object, each obtained from a different light source
(d) Reconstructed image by Algorithm1
(e) Reconstructed image by Algorithm2



(a) (b) (c)



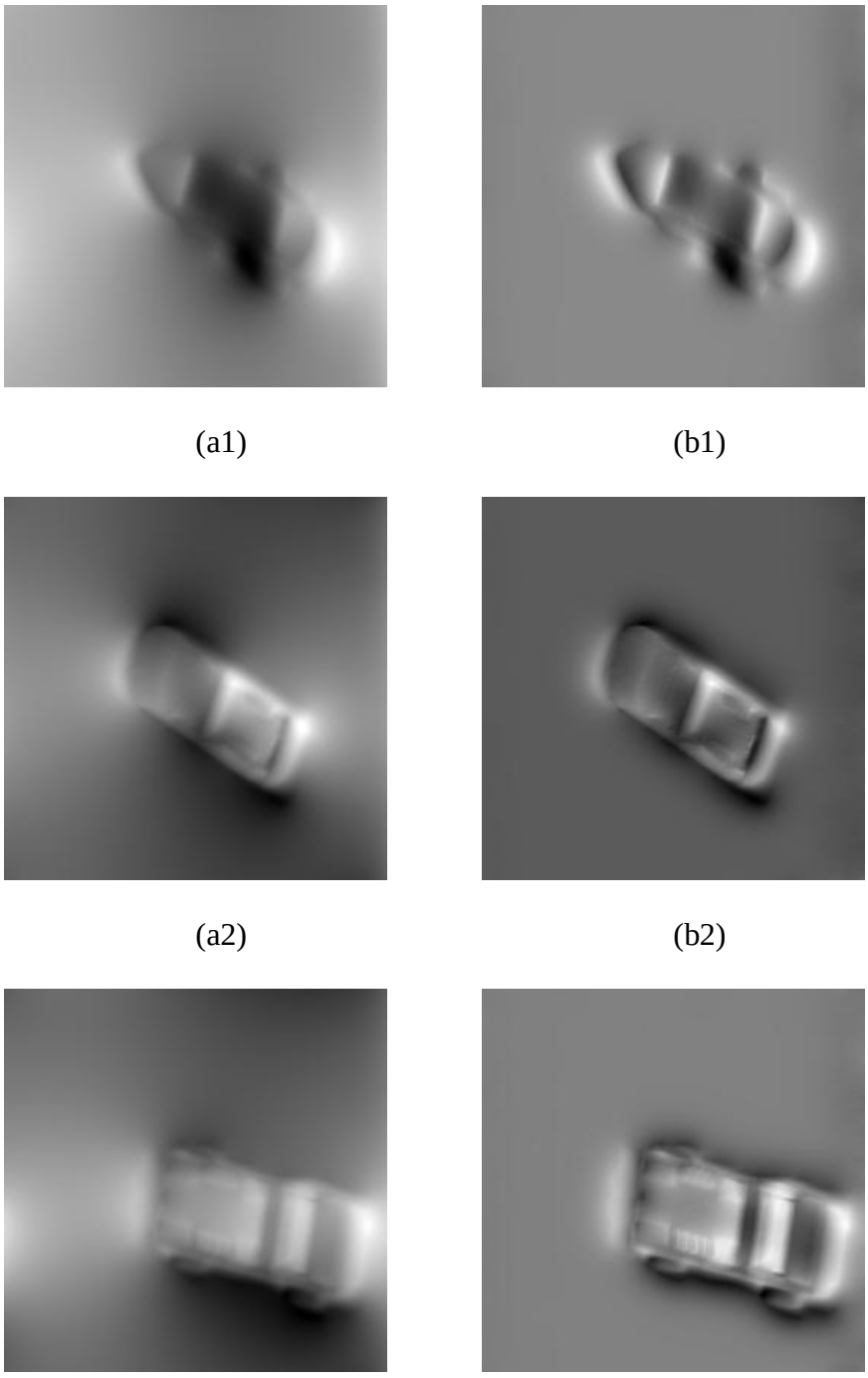
(d) (e)

Figure 7.30 (a), (b), and (c) Photometric stereo images of a real object, each obtained from a different light source
(d) Reconstructed image by Algorithm1
(e) Reconstructed image by Algorithm2

Discussion:

It can be seen from the real world images presented above that, the photometric stereo images obtained have shadows present in them. But in the analysis of the photometric

stereo method with three light sources, it was assumed that all the irradiances were positive i.e. the shadows were excluded from the analysis. Thus, the shadows present in the real images were manually removed prior to the generation of gradient images used in the reconstruction algorithms. The reconstructed images shown in the above figures were obtained from these preprocessed photometric stereo images. However, when the objects were reconstructed in the presence of shadows, the reconstruction results obtained for both the algorithms are as shown in Figure 7.31. In Figure 7.31, a1, a2 and a3 show the reconstructed images of the real world photometric stereo images with shadows, obtained from Algorithm1. The results obtained from Algorithm2 are shown in b1, b2 and b3.



(a1) (b1)

(a2) (b2)

(a3) (b3)

Figure 7.31 (a1), (a2) and (a3) Reconstruction of images with shadows by Algorithm1 (b1), (b2) and (b3) Reconstruction of images with shadows by Algorithm2

Chapter 8

Conclusion and Future work

In this thesis, the problem of reconstruction of Lambertian surfaces from photometric stereo was addressed. Gradient maps were obtained by the theory of photometric stereo. Two algorithms were implemented to reconstruct surfaces from photometric stereo images. The results of these reconstruction algorithms were presented in chapter 7. It can be summarized that,

- Algorithm1 was based on the Fourier transforms approach. The algorithm was implemented on the Beethoven, synthetic and real images. The results were shown in chapter 7.
- Algorithm2 was an iterative reconstruction approach based on minimizing a cost function by gradient descent. This algorithm was implemented on the Beethoven, synthetic and real images. The results obtained were shown in chapter 7.
- In addition, the noise sensitivity of Algorithm1 was tested on both the Beethoven and synthetic image. Noise sensitivity plots were obtained for both these images.
- Algorithm1 was found to have a lower computation time (10 seconds) than Algorithm2 (25-30 minutes) since it was a non-iterative process.
- The regularization term m was found to have a positive effect in the reconstruction process when the gradient images were corrupted by noise.
- Algorithm2 was found to be more robust than Algorithm1 i.e. it produced better reconstruction results than Algorithm1 even in the presence of discontinuities in the image.

Future work:

- The algorithms can be modified to make them more robust to the discontinuities present in the image.
- More than three light sources can be adopted to increase the efficiency of the algorithms on real images.
- Computer vision algorithms like erosion or other thresholding algorithms can be implemented to eliminate the shadows present in the real images.

Bibliography

[1] R. J. Woodham. Photometric stereo: a reflectance map technique for determining surface orientations from a single view. In *Proceedings of SPIE, Image Understanding Systems & Industrial Applications*, volume 155, pages 136-143, 1978.

[2] W.M.Silver. Determining Shape and Reflectance Using Multiple Images. *MS thesis, Massachusetts Institute of Technology, Boston, Massachusetts, 1980*

[3] B. K. P. Horn. *Robot Vision*. MIT Press, Cambridge, 1986

[4]B. K. P. Horn. Understanding image intensities. *Artificial Intelligence Journal*, 8(2):201-231, April 1977.

[5] B. K. P. Horn and M. J. Brooks, editors. *Shape from Shading*. The MIT Press, 1989

[6] J. Van Digellen. A photometric investigation of the slopes and heights of the ranges of hills in the maria of the moon. *Bulletin of the Astronomical Institute of the Netherlands*, 11(423), July 1951.

[7] B. 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, Boston, Massachusetts, 1970.

[8] J.Oliensis. *Shape from Shading as a Partially Well-Constrained Problem*. Computer Vision, Graphics, and Image Processing: Image Understanding, vol. 54, pp. 163-183, 1991

[9] P.Dupuis and J. Oliensis. Direct Method for Reconstructing Shape From Shading. *IEEE Proc. Computer Vision and Pattern Recognition, pp. 453-458, 1992*

[10] M.Bichsel and A.P.Pentland. A Simple Algorithm for Shape from Shading. *IEEE Proc. Computer Vision and Pattern Recognition, pp.459-465, 1992*

[11] K. Ikeuchi and B. K. P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17(1-3):141-184, August 1981

[12] B. K. P. Horn and M. J. Brooks. The variational approach to shape from shading. *Compute Vision, Graphics, and Image Processing*, 33:174-208, 1986.

[13] B. K. P. Horn. Height and gradient from shading. *International Journal of Computer Vision*, pages 37-75, 1989.

[14] R. Courant and D. Hilbert. *Methods of Mathematical Physics*. John Wiley & Sons, New York, 1962

[15] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 932-936, 1985.

[16] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 10(4):439-451, July 1988.

[17] B.-H. Kim and R.-H. Park. Shape from shading and photometric stereo using surface approximation by Legendre polynomials. *Computer Vision and Image Understanding*, 66(3):255-270, June 1997

[18] R. Szeliski. Fast shape from shading. *Computer Vision, Graphics, and Image Processing; Image Understanding*, 53(2):129-153, March 1991

[19] O. E. Vega and Y. H. Yang. Shading logic: A heuristic approach to recover shape from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):592-597, 1991

[20] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):680-702, 1991

[21] Y. Leclerc and A. Bobick. The direct computation of height from shading. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pages 552-558, 1991.

[22] K.M. Lee and C.C.J. Kuo. Shape from Shading with a Linear Triangular Element Surface Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):815-822, 1993

[23] A. P. Pentland. Local shading analysis. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 6(2):170-187, March 1984.

[24] A. P. Pentland. Local shading analysis. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*, chapter 15, pages 443-487. The MIT Press, 1989.

[25] C.-H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, 26(2):125-143, 1985.

[26] C.-H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*, chapter 11, pages 323-347. The MIT Press, 1989.

[27] C.-H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Machine Intelligence*, 13(7):680-702, 1991.

[28] A. Pentland. Shape Information From Shading: A Theory About Human Perception. *Proc. Of Int'l Conf. Computer Vision, pp 404-413, 1988.*

[29] P.S. Tsai and M. Shah. Shape From Shading Using Linear Approximation. *Image and Vision Computing Journal, 12(8): 487:498, 1994.*

[30] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):690-706, August 1999.

[31] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Shape from Shading*, chapter 17, pages 513-532. The MIT Press, 1989.

[32] R. Zhang and M. Shah. Iterative shape recovery from multiple images. *Image and Vision Computing*, 15(11):801-814, November 1997

[33] R. Onn and A. Bruckstein. Integrability disambiguates surface recovery in two-image photometric stereo. *International Journal of Computer Vision*, 5(1):105-113, 1990.

[34] J. Yang, N. Ohnishi, and N. Sugie. Two-image photometric stereo method. In *Proceedings of SPIE, Intelligent Robots and Computer Vision XI: Biological, Neural Net, and 3-D Methods*, volume 1826, pages 452-463, 1993.

[35] T. E. Kim, S. H. Lee, S. H. Ryu, and J. S. Choi. Shape reconstruction using extended photometric stereo. In *Proceedings of the International Conference on Information, Communications and Signal Processing, Trends in Information Systems Engineering and Wireless Multimedia Communications*, volume 2, pages 718-720, 1997.

[36] B.Karacali and W.Snyder. Reconstructing discontinuous surfaces from a given gradient field using partial integrability. *Computer Vision and Image Understanding 92 (2003) pp 78 -111*

[37] B. Karacali. *Vector Space Methods For Surface Reconstruction From One Or More Images Acquired From The Same View With Applications To Scanning Electron Microscopy Images*. PhD thesis, North Carolina State University, Raleigh, North Carolina, 2002.

[38] T. Wei and R.Klette. A New Algorithm for Gradient Field Integration. Image and Vision Computing New Zealand. (IVCNZ'2001), Dunedin, New Zealand. (2001)

[39] T.Wei and R.Klette. Depth Recovery from Noisy Gradient Vector Fields Using Regularization. *CAIP 2003, LNCS 2756, pp 116-123, 2003*

[40] T.Wei and R.Klette. Regularization method for Depth from Noisy Gradient Vector Fields. *http://www.citr.auckland.ac.nz/techreports/2002/CITR-TR-115.pdf*

[41] http://www.ece.ncsu.edu/imaging/Software/index.html

[42] http://range.informatik.uni-stuttgart.de/htdocs/html

[43] W.Snyder and H.Qi. *Machine Vision*. Cambridge University Press, 2004

[44] G.Bilbro and W.Snyder, "Mean Field annealing, an Application to Image Noise Removal," *Journal of Neural Network Computing, Fall, 1990*

[45] G.Bilbro, W.Snyder, S.Garnier, ad J.Gault, "Mean Field annealing:a Formalism for Constructing GNC-like Algorithms," *IEEE Transactions on Neural Networks, 3(1) pp. 131-138, 1992*