

**VIEWPOINT INVARIANT SURFACE RECONSTRUCTION FROM
GRADIENT DATA**

by

Yossarian Yggy King

B.Sc. (Computer Science) Dalhousie University, 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

IN THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF COMPUTER SCIENCE

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 1993

© Yossarian Yggy King, 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Computer Science

The University of British Columbia
Vancouver, Canada

Date 30/8/93

Abstract

Surface reconstruction recovers complete and coherent surfaces from scattered, noisy data. Viewpoint invariant reconstruction allows seamless combination of multiple views, recognition of objects in arbitrary poses, and consistent quantitative measurements.

Viewpoint invariant reconstruction from gradient data is studied in a variational framework. An invariant smoothing functional based on the bending energy of a thin plate is derived. The functional is nonconvex, and three convex approximations are presented. An invariant but nonconvex metric for measuring fit to gradient data is given, based on the angle between unit surface normals. A convex constraint metric is obtained from the second order Taylor series approximation. The chief goal of this thesis is to evaluate and compare the different functionals, with viewpoint invariance the prime evaluation criterion.

The approximately invariant regularizing functionals are discretized on a regular grid using a finite difference method. Reconstruction from gradient data proceeds in three stages: interpolation of the initial data; smoothing of the gradient map; and recovery of height from gradient.

Data were generated for multiple views of a synthetic surface. Each reconstruction method is applied to recover a surface for each view. The viewpoint invariance of the reconstruction is evaluated by measuring, qualitatively and quantitatively, how well the surfaces from different views match up.

Experiments are performed with real data from photometric stereo in the context of building 3D object models by fitting together surfaces reconstructed from multiple views. Viewpoint invariance is important in order for the different surfaces to fit together. Multiple views of a simple ellipsoid of known shape allows accurate evaluation of the results, while data gathered from a doll's face provides more of a challenge to the reconstruction techniques.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	v
List of Figures	vi
1 Introduction	1
2 Surfaces, Shapes, and Objects	5
2.1 Surface Reconstruction from Range Data	5
2.2 Shape from Shading	7
2.3 Acquisition of 3D Object Models	9
3 Variational Approach to Surface Reconstruction	12
3.1 Problem Statement	12
3.1.1 The Ill-Posed Problem	13
3.1.2 Well-Posed Variational Formulation	13
3.2 Viewpoint Invariance	15
3.2.1 Requirements for Invariance	15
3.2.2 An Invariant Constraint Metric	16
3.2.3 An Invariant Smoothing Functional	19
3.3 Convex Approximation	21
3.3.1 Approximating the Constraint Metric	21
3.3.2 Small Deflection Approximation of the Smoothing Functional	22
3.3.3 First Derivative Parameterization	23
3.3.4 Principal Directions Parameterization	23
3.4 Recovering Height from Gradient	27
4 Surface Reconstruction on a Regular Grid	29
4.1 Discretization	31
4.1.1 Regular Grid	32
4.1.2 Approximating Derivatives with Finite Differences	33
4.2 Initial Interpolation of the Gradient Map	35
4.3 Minimization of the Discretized Functional	37
4.4 Boundary Conditions	39

5	Investigations with Synthetic Data	41
5.1	Synthetic Data Sets	42
5.2	Failure of Invariant Constraint Metric	46
5.3	Use of Different Functionals for Height Recovery	48
5.4	Comparison of Smoothing Functionals	49
5.5	λ Normalization	57
5.6	The Data-Point Correspondence Problem	58
5.7	Effects of Noise	60
5.8	Effects of Varying Smoothing Parameters	64
6	Building Object Models from Real Data	69
6.1	Data Collection	69
6.1.1	Photometric Stereo	70
6.1.2	Imaging Setup and Calibration	71
6.2	A Simple Ellipsoid	75
6.2.1	Surface Recovery	75
6.2.2	Building Models	78
6.3	Facial Reconstruction	96
7	Conclusions and Future Work	105
7.1	Summary of Results	106
7.2	Issues Raised	107
7.3	Future Work	109
	Bibliography	112
	Appendices	116
A	Invariant Quadratic Constraint Metric Impossible	116
B	Interpolation by Iterated Local Averaging	120
C	Conjugate Gradient Descent	123

List of Tables

5.1	Distance between recovered surfaces	56
5.2	Distance between recovered surfaces for noisy data	64
6.3	Pixel dimensions at different distances	74

List of Figures

1.1	Model reconstruction from multiple views	4
3.1	An invariant measure of the distance between two vectors	17
4.1	Surface reconstruction from gradient data	30
4.2	Estimating derivatives over a rectangular patch	34
5.1	A synthetically generated surface	42
5.2	Synthetic z_x	43
5.3	Synthetic z_y	43
5.4	Cross section through three synthetic surfaces	45
5.5	Oversmoothing caused by angle-based constraint metrics	47
5.6	Differential surfaces for the two constraint metrics	47
5.7	Height recovered with different integrators	48
5.8	Surface reconstructed using Ω_1	50
5.9	Surface reconstructed using Ω_2	51
5.10	Surface reconstructed using Ω_3	52
5.11	Surface slices reconstructed using Ω_1	53
5.12	Surface slices reconstructed using Ω_2	54
5.13	Surface slices reconstructed using Ω_3	55
5.14	The data-point correspondence problem	59
5.15	Reconstruction with corresponding point data fit term	61
5.16	Surface reconstructed from noisy data without smoothing	62
5.17	Noisy, sparse p component of gradient map	63
5.18	Reconstruction from noisy data with Ω_1	65
5.19	Reconstruction from noisy data with Ω_2	66
5.20	Reconstruction with different λ values	67
5.21	Reconstruction with different ν values	68
6.1	The imaging setup	72
6.2	Three images of the ellipsoid	73
6.3	Optical bench calibration	74
6.4	Mathematical specification of the ellipsoid	76
6.5	p and q components of gradient map from photometric stereo	77
6.6	Surface height recovered without smoothing	78

6.7	Surface height recovered with Ω_1 smoothing	79
6.8	Surface height recovered with Ω_2 smoothing	79
6.9	Slice through reconstructed surfaces	80
6.10	Slices through smoothed gradient maps	81
6.11	Model built with no smoothing	86
6.12	Model built using Ω_1 smoothing	87
6.13	Model built using Ω_2 smoothing	88
6.14	Cross-section of registered surfaces with no smoothing	89
6.15	Cross-section of registered surfaces with Ω_1 smoothing	90
6.16	Cross-section of registered surfaces with Ω_2 smoothing	90
6.17	Cross-section of ground truth definition of ellipsoid	91
6.18	Cross-section of registered surfaces for trimmed data with no smoothing .	92
6.19	Cross-section of registered surfaces for trimmed data with Ω_1 smoothing .	93
6.20	Cross-section of registered surfaces for trimmed data with Ω_2 smoothing .	93
6.21	Cross-section of registered trimmed surfaces with no smoothing	94
6.22	Cross-section of registered trimmed surfaces with Ω_1 smoothing	95
6.23	Cross-section of registered trimmed surfaces with Ω_2 smoothing	95
6.24	Images of doll's face	97
6.25	Gradient data for the first view of the doll's head	98
6.26	Face reconstructed in first view without smoothing	100
6.27	Face reconstructed in first view with Ω_1 smoothing	101
6.28	Face reconstructed in first view with Ω_2 smoothing	102
6.29	Face surface alignment without smoothing	103
6.30	Face surface alignment with Ω_1 smoothing	103
6.31	Face surface alignment with Ω_2 smoothing	104
B.1	Sparse data on a grid	121
B.2	Surface interpolated from sparse data	122
C.1	Convergence criterion for conjugate gradient descent	125

Introduction

Visual perception of the three dimensional world begins in two dimensions. A flat image, together with knowledge of how that image is formed, provides information about the shape and structure of the objects and scenes in view. Stereo, optical flow analysis, shape from shading, and shape from texture are some of the techniques of computational vision used to derive 3D information from 2D images, typically providing estimates of distance or surface orientation at points in the visual field.

Surface reconstruction is the task of determining complete and coherent surfaces from the sparse, noisy, and possibly inconsistent local constraints provided by the various modules of early vision. The reconstruction of visual surfaces is an intermediate step leading towards full 3D scene analysis, object recognition, object model acquisition, and similar high-level tasks.

The form of objects in the world is independent of how they are viewed, and their perception should mirror this invariance—the shape of a reconstructed surface should not depend on the viewpoint from which it is seen. Viewpoint invariant surface reconstruction from range data has been attempted in previous work (Stevenson & Delp, 1992; Vemuri & Skofteland, 1992), and this thesis seeks to provide a similar investigation for reconstruction from gradient data.

A variational framework is used, and the task of surface reconstruction is posed as one of functional minimization, wherein viewpoint invariance can be achieved by using invariant functionals. Use of gradient data has typically received attention in the context

of shape from shading, which recovers surface orientation, and often surface height, from the shading of one or more 2D images. Shape from shading has been shown to be an area where a variational formulation is particularly appropriate (Horn & Brooks, 1986), but little emphasis has been placed on viewpoint invariance.

Automatic model acquisition seeks to integrate information from multiple views to form a complete 3D model of an object, such as might be required for object recognition (Goad, 1983), pose estimation (Li, 1993), or for inclusion in a computer graphics or CAD system. Viewpoint invariance is important for successfully merging information from multiple views, since the recovered shape of the object in overlapping areas must be the same in order for them to fit together. This thesis uses 3D model acquisition as a scenario for evaluating the invariance of the proposed reconstruction methods.

Chapter 2 provides an overview of some of the relevant literature in the area of surface reconstruction, shape from shading, and 3D model acquisition. Aspects of viewpoint invariance are emphasised, particularly as they relate to variational methods.

Chapter 3 presents a variational formulation that regularizes the ill-posed surface reconstruction problem. The requirements for viewpoint invariance are elucidated. An invariant data fit metric based on the angle between surface normals is derived. The energy density of a thin-plate under deformation is used as an invariant smoothing term. This derivation yields a non-convex functional which is difficult to minimize with standard techniques. Three quadratic approximations to the invariant smoothing functional are derived, following the development of (Stevenson & Delp, 1992). The first of these is the small deflection approximation, equivalent to the quadratic variation stabilizer that has previously been used in the surface reconstruction literature (Grimson, 1981; Terzopoulos, 1988). The second approximation is based on having estimates of the first derivatives of surface height at each point in the image, a technique that is particularly appropriate for application to gradient data. The final approximation is based on estimates of the

principal directions at each point on the visible surface.

Continuous-domain formulations are discretized in Chapter 4, using a finite difference method. The conjugate gradient descent method is described for solving the discrete minimization problem. Surface reconstruction over a regular grid is achieved through a three stage process that interpolates the initial gradient data, smooths the resulting gradient map by application of the invariant functional, and enforces integrability constraints to recover surface height.

Chapter 5 describes a number of experiments on synthetic data that were used to evaluate the reconstruction methods in a carefully controlled manner. Invariance of the three approximations is investigated, and the effects of noise and different smoothing parameters are explored. Various issues related to the regularizing process and the comparison of different regularizing functionals are also shown to be important.

Chapter 6 uses real gradient data derived from photometric stereo (Woodham, 1980) to reconstruct surfaces from multiple views of an object mounted on a rotating turntable. Data are gathered for four views of an ellipsoid of known shape. The different surfaces are matched together to produce a complete 3D object model. For direct matching of surfaces from adjacent views, viewpoint invariance is essential. Since gradient data provides no information about absolute depth, neighbouring surfaces need to be registered. Additional surface matching experiments were performed with data from two views of the more complicated surface of a pottery doll's face. Figure 1.1 shows the construction of surfaces from multiple views and illustrates the correspondence problem and the need for invariance.

Some success is achieved, though in practice effects of view variation in the data and enforcement of boundary conditions may be more important than the choice of smoothing functional. These and other results are summarized in Chapter 7, together with directions for future work.

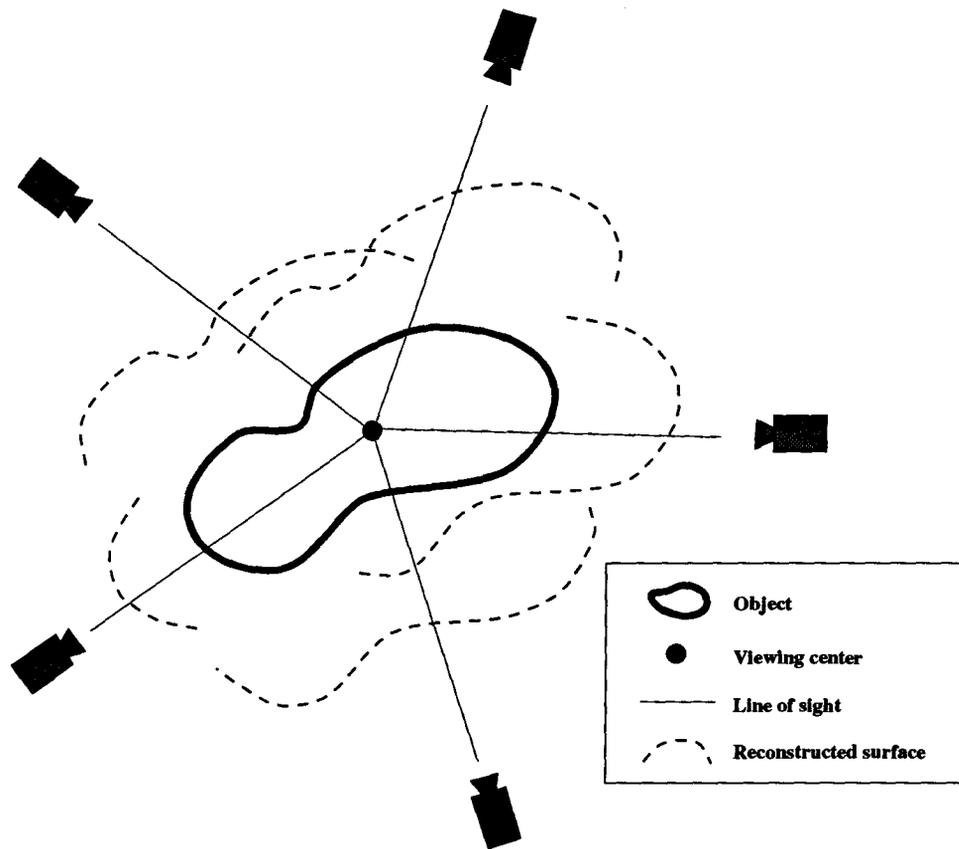


Figure 1.1: Model reconstruction from multiple views. Invariance is required so that adjacent surfaces fit together. Recovered surfaces need to be translated along their respective lines of sight in order to bring them into registration.

Surfaces, Shapes, and Objects

The literature abounds with work on surface reconstruction from range data and the determination of shape from shading. Both of these areas are very relevant to the problem of surface reconstruction from gradient data. The proposed invariant reconstruction techniques are to be tested in the context of fitting surfaces together to form 3D object models, another area that has seen much work. This chapter provides a review of some of the more relevant literature in these areas.

Past work on surface reconstruction is reviewed in Section 2.1. Related issues from shape from shading are examined in Section 2.2. Various approaches to 3D model acquisition are described in Section 2.3.

2.1 Surface Reconstruction from Range Data

Stereo vision, motion analysis, and passive and active rangefinders all provide estimates of the distance to the nearest surface at points in the visual field. Interpolation and smoothing of these data are required to combat noise and sparsity. Reconstruction of visual surfaces from range data has received attention in previous research.

Grimson (Grimson, 1981) provides an early variational formulation of the surface reconstruction problem. Reconstruction is performed on a grid, with the quadratic variation stabilizer used as a smoothing term. Under certain assumptions, this stabilizer is equivalent to the small deflection approximation of the bending energy of a thin plate, a

stabilizer which is examined in this thesis. The quadratic variation stabilizer is thus approximately invariant when the surface slope is small, although the property of invariance is not explicitly considered in Grimson's work.

The desire for invariance is expressed by Terzopoulos (Terzopoulos, 1988), but the formulation presented is, again, only invariant under the small deflection assumption. "Splines under tension" are proposed, providing a combination of thin-plate and membrane functionals that allows controlled continuity. An efficient multigrid method is implemented.

A recent implementation of the thin-plate/membrane spline approximation provides approximately invariant surface and motion estimation from stereo and rangefinder data (Vemuri & Skofteland, 1992). Kalman filtering techniques are used to integrate data from a stream of images and to maintain uncertainty estimates of the values recovered. The small-deflection approximation of the thin-plate energy density is again used, which seems adequate in practice.

Viewpoint invariance is one of the major evaluation criteria of the survey of 3D reconstruction techniques provided by (Bolle & Vemuri, 1991). Various methods are surveyed, using a broad categorization into methods for explicit and implicit surfaces. A specific scheme is proposed for the reconstruction of parametric surfaces. Provided the parameterization is viewpoint invariant, then so is the reconstruction. The notion of an invariant parameterization is an important one, as is discussed in Section 5.6.

The reconstruction methods described above all refer to invariance as a desirable property, but are generally more concerned with accurate recovery of the surface. Viewpoint invariance is the chief focus of (Stevenson & Delp, 1992), which derives an invariant variational formulation of the surface reconstruction problem based on the thin-plate energy density. The formulation is non-convex, and hence difficult to solve. Several convex approximations are derived and tested on synthetic and real range data.

This thesis is largely based on the work of Stevenson and Delp, and provides an invariant formulation for surface reconstruction from gradient data that is analogous to their formulation for range data. Stevenson and Delp suggest a constraint metric for use with gradient data based on the sine of the angle between surface normals, but do not test it. A metric very similar to this is shown to be inadequate in Chapter 5 of this thesis.

Attempts have previously been made to incorporate slope information into the surface reconstruction problem. Harris presents a method for surface reconstruction that integrates range, slope, and higher-order derivatives (Harris, 1987). The solution is formulated as the minimization of the heat energy in a multi-level analog resistor mesh. Discontinuities of any order are handled by “breaking” the mesh at the appropriate level. A parallel iterative scheme is presented for the solution. The formulation is not view invariant.

This coupled depth/slope model has been implemented and combined with a method to incorporate discontinuities based on the natural boundary conditions of the problem (Li, 1991). A number of experiments are performed, indicating that natural boundary conditions work well, except at self-occluding boundaries, where additional information is needed to achieve satisfactory results. This thesis does not directly address the issues of discontinuities and boundary conditions, but they are shown to be important to the general task of surface reconstruction.

2.2 Shape from Shading

Gradient information has been used for surface description and reconstruction primarily in the context of shape from shading. The observed shading of surfaces is obviously affected by the local orientation of the surface, and this fact has been exploited in vision systems to determine local shape. Given orientation, full recovery of the 3D surface is a

logical extension that is often incorporated.

The shape from shading problem is easily cast in a variational framework, and much of the early work on shape from shading can be understood in these terms (Horn & Brooks, 1986). Horn and Brooks develop a “systematic approach to the discovery of parallel iterative schemes for solving the shape from shading problem on a grid,” based on satisfying a discretized version of the Euler equations for the variational problem. The importance of enforcing integrability is emphasized, and a method for doing so is derived which recovers height and orientation simultaneously, using the height function as a check on the integrability of the gradient field. Integrability is not strictly enforced, but rather encouraged by the use of a penalty term.

Horn and Brooks are not alone in emphasizing integrability. Frankot and Chellappa present a method for projection of a non-integrable gradient field onto its “closest” integrable counterpart (Frankot & Chellappa, 1988). The solution is direct, but requires that the surface be represented as a linear combination of a finite set of basis functions. The technique can be retrofit to existing shape from shading methods to maintain integrability.

Horn provides a historical perspective on the shape from shading problem and its solution, and presents a method for the simultaneous recovery of height and gradient (Horn, 1990). It is emphasized that shape from shading is not ill-posed if suitable boundary conditions are known, though these boundary conditions can be difficult to incorporate with the gradient representations of surface orientation. The method presented seeks to encourage integrability and satisfy the image irradiance equation. An iterative scheme is derived based on the Euler equations of the problem.

A caution against overly sophisticated techniques is given by (Tsai & Shah, 1992), who present a shape from shading algorithm that is implemented in 20 lines of C code that requires only “one to two iterations” to converge. The first derivatives of the surface are

approximated with forward differences of the surface height function, and the reflectance function is linearized directly in terms of surface height rather than first derivatives. Convergence is assured by the linearity of the solution. The algorithm produces results that appear qualitatively correct.

Motivated by the observation that surface shape is perceived largely independently of illumination conditions, (Koenderink & van Doorn, 1980) show that a large class of local brightness extrema occur along parabolic lines of a surface—points at which one of the principal curvatures is zero (i.e., the surface is locally flat in one direction). These lines divide the surface into patches of positive and negative Gaussian curvature. Surface curvature is a viewpoint invariant surface property, and so a link is provided between the shape from shading problem and the quest for invariance. Reliable local estimates of curvature can be derived from shading information using the photometric stereo technique (Woodham, 1990).

Despite these results, little attention has been given to the issue of viewpoint invariance in solving the shape from shading problem. This thesis addresses one aspect of this problem by testing various approximately invariant gradient-based smoothing functionals.

2.3 Acquisition of 3D Object Models

A logical extension of the idea of surface reconstruction is to merge surfaces from multiple views to derive full 3D models of the object in view. Many proposed techniques for object recognition rely on having an internal model of the object to be recognized (Roberts, 1965; Goad, 1983; Lowe, 1987; Wang & Iyengar, 1992). Automatic generation of object models would facilitate these techniques.

Surfaces generated from different views can only be merged directly if their overlapping regions fit together (i.e., if they have been calculated in an invariant manner). This thesis proposes to test the invariance of surface reconstruction techniques in the context of piecing together 3D object models such as might be used in object recognition tasks.

The octree representation is a method of recursively subdividing 3D space to create an occupancy structure describing an object. Szeliski uses an octree representation for automatic recovery of object models (Szeliski, 1990a). The object is placed on a rotating turntable, with multiple silhouette views used to “carve out” the octree representation. Holes can be handled, but the use of the silhouette limits the forms of concavity that can be determined. Since most of the information in each image is being ignored, many images are required for an accurate model. Octrees are efficient to represent and manipulate (Chen & Huang, 1988), but suffer from an inability to represent local shape.

Szeliski presents an alternative method for automatic model building in which optical flow is used to calculate the range of points on an object as it rotates on a turntable (Szeliski, 1990b). Disparity measurements are projected to 3D points with associated uncertainties. Points from multiple views are brought into a common coordinate frame and nearby points merged. A method of smoothing and interpolating the resulting point set is suggested but not implemented, though later work explains how this might be achieved (Szeliski & Tonnesen, 1991). Results are presented for several simple polyhedra.

Viewpoint invariance is not emphasized in Szeliski’s work, but is indirectly important to Ferrie *et al*, who use estimates of the local Darboux frame (representing invariant surface properties) to segment a surface and fit volumetric primitives to the resulting patches (Ferrie *et al*, 1990). The process generates a “blob” model of the object in which fine surface detail is lost.

Earlier work attempts to fit a polyhedral object model to a set of 3D points (Faugeras *et al*, 1984). Range data is acquired from multiple views as an object rotates on a

turntable. Points from each view are brought into a common coordinate system, and an iterative process is used to determine a triangulation that lies “sufficiently close” to the observed data.

It has recently been proven that “there are no view invariant features for the general case of n 3D points under perspective [or orthographic] projection” (Burns *et al*, 1993), which presents difficulties for feature-based object recognition. If recognition is not to depend on viewpoint, then it is desirable that it be based on viewpoint invariant measures, be they discrete features or dense measures. The authors suggest an analysis of the view variation of different types of features to derive distribution functions to aid recognition. A more direct approach seems to be to augment object models with local shape information, thus providing more constraints on the matching process.

Chapter 6 describes an attempt to build 3D object models by registration of surfaces calculated from overlapping views of an object. Local surface orientation can be maintained at each stage of the process and used to augment the final model. If surface reconstruction is invariant, then the multiple surfaces should fit together well. The actual fit observed provides a test of invariance.

Variational Approach to Surface Reconstruction

The surface reconstruction task is to produce a dense map of depth estimates that is consistent with data sensed from the scene, possibly using prior knowledge and expectations about the scene domain. This chapter formalizes this problem. Section 3.1 states the problem in its ill-posed form, and presents a well-posed variational formulation. The conditions required for viewpoint invariance are elaborated and an invariant regularizing functional is derived in Section 3.2. The functional derived is non-quadratic and difficult to minimize. More tractable approximations to the invariant functional are discussed in Section 3.3. The presentation and content of this chapter somewhat follows the development of (Stevenson & Delp, 1992), which was the work that inspired much of this thesis.

3.1 Problem Statement

This section provides a mathematical statement of the surface reconstruction problem and a well-posed variational formulation.

3.1.1 The Ill-Posed Problem

A visual surface can be represented (barring transparency) as an explicit surface $z(x, y)$, giving the distance to the scene at each (x, y) point in the visual field¹. Low level visual processes provide estimates of properties of the scene related to shape, in particular estimates of the distance or surface normal at points in the scene. These estimates may be noisy, sparse, and inconsistent.

Given a set \mathcal{S} of estimates of the surface distance or surface normal at discrete points in the visual field V , a naive statement of the surface reconstruction problem is to find a real-valued function $\hat{z}(x, y)$, defined for all $(x, y) \in V$, which matches the point-wise estimates of distance and orientation given by \mathcal{S} .

A problem is well-posed, in the sense of Hadamard, if a unique solution exists that depends continuously on the initial data (Tikhonov & Arsenin, 1977). The surface reconstruction problem, as posed above, may violate any of these conditions—inconsistencies in the constraints may imply that there is no exact solution; sparsity of constraints allows many solutions; detection and exploitation of depth or orientation discontinuities may lead to a solution that does not depend continuously on the initial data.

3.1.2 Well-Posed Variational Formulation

Regularization theory can be applied to yield a well-posed problem by restricting the space of allowable solutions, and has been successfully applied in many areas of computer vision (Poggio *et al*, 1985). Incorporation of *a priori* expectations and knowledge about the scene domain places restrictions on the form of the estimated surface. The problem now becomes one of finding the surface estimate that best fits the sensed data while simultaneously meeting the desired smoothness and coherence criteria. The problem

¹Though there may be many different object and background surfaces in view, we can speak of a single “surface,” perhaps with discontinuities, covering the entire field of view.

as initially posed sought to find the surface height function from among all real-valued functions. Regularization generally limits the solution space to functions with continuous derivatives to some specified order.

The regularized problem can be mathematically stated as finding (for fixed \mathcal{S}) z^* which minimizes a measurement functional

$$M[z^*, \mathcal{S}] = \sum_{c \in \mathcal{S}} \rho_c(z^*, c) + \lambda \Omega[z^*] \quad (3.1)$$

where Ω is a stabilizing functional that measures the “smoothness” of the surface in some appropriate way; ρ_c measures the fit of the surface to the constraint c ; and λ is a regularizing parameter that controls the tradeoff between smoothness and fit to the data. Note that Ω serves a dual purpose—it incorporates expectations about the general form of the surface to interpolate between data points, and it helps compensate for the effects of noise in the data. Note also that the solution of the variational problem is not the same as the solution of the original problem. If the smoothness and data fit terms of the regularizing functional are chosen appropriately then the solution should be satisfactory, but it is a different problem that is being solved.

The solution to the regularized surface reconstruction problem is the surface \hat{z} satisfying

$$M[\hat{z}, \mathcal{S}] = \min_{z^*} M[z^*, \mathcal{S}] \quad (3.2)$$

A unique solution is guaranteed if and only if the measurement functional M has a unique global minimum².

²If there are no local minima other than the global minimum then the solution will be easier to find (using gradient descent methods).

3.2 Viewpoint Invariance

For seamless combination of multiple views or recognition of objects in arbitrary poses, it is desirable to reconstruct surfaces in a manner that is independent of viewpoint. This section describes the conditions that must be met for this to be achieved and derives a regularizing functional for invariant reconstruction.

3.2.1 Requirements for Invariance

Suppose that a surface estimate \hat{z} is derived from a set of sensed constraints \mathcal{S} . Applying a transformation T to the constraint set yields

$$\mathcal{S}' = T[\mathcal{S}] \quad (3.3)$$

Let \hat{z}' be the surface reconstructed from \mathcal{S}' . If the reconstruction process is invariant to the transformation T , then the surface reconstructed from the transformed constraints should be the same as that obtained by transforming the surface reconstructed from the original constraints, i.e.

$$\hat{z}' = T[\hat{z}] \quad (3.4)$$

This can be achieved if the regularizing functional is invariant to the transformation, that is

$$M[z^*, \mathcal{S}] = M[T[z^*], T[\mathcal{S}]] = M[z^{*'}, \mathcal{S}'] \quad (3.5)$$

That this is sufficient for invariance can be seen from the following

$$M[\hat{z}, \mathcal{S}] = \min_{z^*} M[z^*, \mathcal{S}] = \min_{z^{*'}} M[z^{*'}, \mathcal{S}'] = M[\hat{z}', \mathcal{S}'] \quad (3.6)$$

If M has a unique global minimum then \hat{z} and \hat{z}' are unique, and by Equation 3.5 \hat{z}' is the transformed version of \hat{z} .

For viewpoint invariance, the relevant transformations are rigid translations and rotations³. An invariant regularizing functional can be constructed from a constraint metric ρ and a stabilizing functional Ω based on surface characteristics that are unaffected by rigid motion.

3.2.2 An Invariant Constraint Metric

The fit between a postulated surface estimate and a set of data constraints should be an invariant calculation if the final surface estimate is to be independent of viewpoint. Viewpoint invariance requires that the constraint metric not depend on the coordinate system within which the surface is embedded.

For depth data, such as those from a range-finder or stereo module, a suitable invariant metric is the squared distance between each range point and the postulated surface. This thesis addresses the use of slope data, for which an invariant measure can be based on the angle between the normal of the estimated surface and the normal obtained from a slope data point.

Given a gradient estimate (p, q) , where p and q represent the surface slope in the x and y directions, the (un-normalized) surface normal corresponding to that data point is $\mathbf{n}_d = (-p, -q, 1)$. Similarly, the surface normal at a point on the estimated surface is $\mathbf{n}_s = (-z_x, -z_y, 1)$, where z_x and z_y are the first partial derivatives of surface height with respect to x and y . The angle θ between the two vectors is given by

$$\begin{aligned} \cos \theta &= \frac{\mathbf{n}_d \cdot \mathbf{n}_s}{\|\mathbf{n}_d\| \|\mathbf{n}_s\|} \\ &= \frac{pz_x + qz_y + 1}{\sqrt{(p^2 + q^2 + 1)(z_x^2 + z_y^2 + 1)}} \end{aligned} \quad (3.7)$$

³Note that preservation of visibility is assumed. Wherever the term “invariant” is used, “visible-invariant” should be understood.

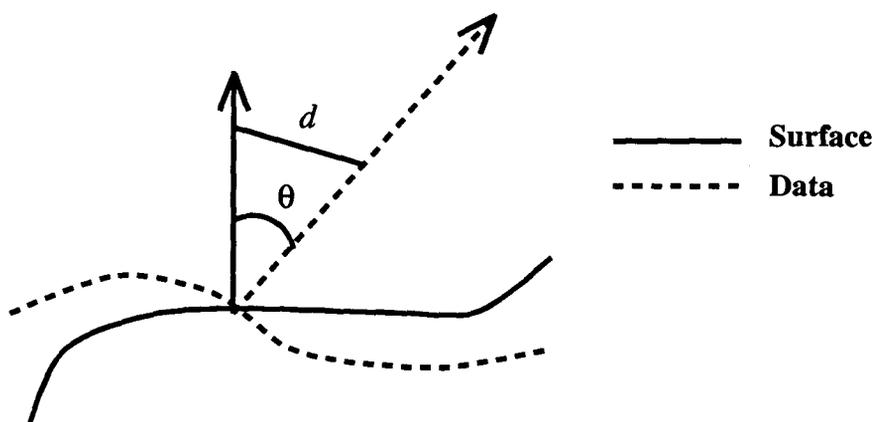


Figure 3.1: An invariant measure of the “difference” between two surface normal vectors.

An invariant constraint metric that increases with increasing angle (up to the maximum possible angular separation π) is the squared distance d^2 between the endpoints of the normalized vectors (see Figure 3.1), where

$$\frac{d}{2} = \sin \frac{\theta}{2} \quad (3.8)$$

Trigonometric manipulation leads to

$$d^2 = 2(1 - \cos \theta) \quad (3.9)$$

yielding the angle-based constraint metric

$$\rho_\theta(z^*, (p, q)) = 1 - \frac{pz_x + qz_y + 1}{\sqrt{(p^2 + q^2 + 1)(z_x^2 + z_y^2 + 1)}} \quad (3.10)$$

This metric is not quadratic with respect to p and q , which makes minimization somewhat harder (as explained in Section 3.3). Stevenson and Delp suggest a similar constraint metric, using $\sin^2 \theta$ (Stevenson & Delp, 1990), but apparently has not implemented it.

A quadratic constraint metric can be obtained from the squared distance between the endpoints of the non-normalized surface normal vectors:

$$\rho_{SD}(z^*, (p, q)) = \|\mathbf{n}_s - \mathbf{n}_d\|^2 = (z_x - p)^2 + (z_y - q)^2 \quad (3.11)$$

This constraint is coordinate system dependent and thus not invariant⁴. This is the metric used by (Horn, 1990) as an integrability constraint for recovering surface height from gradient.

The ideal constraint metric would be both viewpoint invariant, and quadratic in z_x and z_y . Unfortunately, such a metric does not exist (see Appendix A for proof). An alternative is to use a quadratic approximation, such as that derived in Section 3.3.1.

⁴For example, if $(p, q) = (1, 0)$ and $(z_x, z_y) = (0, 0)$, then a penalty value of 1 is obtained. Rotating these vectors by 45° about the y axis causes the penalty to blow up to ∞ , with all other penalty values in $[1, \infty)$ assumed for smaller rotations.

3.2.3 An Invariant Smoothing Functional

Ideally, the stabilizing functional Ω of Equation 3.1 would be viewpoint invariant and would assume greater values as the surface departs from consistency. Since most surfaces commonly encountered in the visual world are (piecewise) smooth and tend not to vary rapidly, a stabilizing functional can be based on surface curvature, an invariant property.

Reconstructing the surface by fitting a thin plate model to the available data achieves this end. The energy density of a thin plate deformed from a planar rest state is given as (Courant & Hilbert, 1989)

$$A \left(\frac{\kappa_1^2 + \kappa_2^2}{2} \right) + B\kappa_1\kappa_2 \quad (3.12)$$

where κ_1 and κ_2 are the principal curvatures of the plate. The parameters A and B are properties of the material, where $A = E$ is Young's modulus, a measure of the stiffness of the material⁵ and $\frac{B}{A} = \nu$ is Poisson's ratio, which measures the lateral contraction of the material per unit elongation (Volterra & Gaines, 1971). For most metals and glass, Poisson's ratio is approximately $\nu \approx 0.3$, while for some rubbers and plastics it can approach the volume-conserving theoretical limit of 0.5 (Jastrzebski, 1976). Equation 3.12 can be rewritten in terms of these more familiar constants as

$$E \left\{ \left(\frac{\kappa_1^2 + \kappa_2^2}{2} \right) + \nu\kappa_1\kappa_2 \right\} \quad (3.13)$$

An alternative measure of curvature is provided by the Gaussian and mean curvatures

$$K = \kappa_1\kappa_2 \quad (3.14)$$

$$H = \frac{1}{2}(\kappa_1 + \kappa_2) \quad (3.15)$$

Which yields an equivalent expression for the energy density

$$E(2H^2 - (1 - \nu)K) \quad (3.16)$$

⁵Analogous to the spring constant of an ideal spring that obeys Hooke's law.

The stabilizing functional to be used is the total deformation energy of the plate, obtained by integrating the energy density over the surface area of the plate

$$\Omega[z^*] = \int_U E(2H^2 - (1 - \nu)K) dA \quad (3.17)$$

For an explicit surface $z(x, y)$, the Gaussian and mean curvatures are (Pogorelov, undated)

$$K = \frac{z_{xx}z_{yy} - z_{xy}^2}{(z_x^2 + z_y^2 + 1)^2} \quad (3.18)$$

$$H = \frac{(z_x^2 + 1)z_{yy} - 2z_xz_yz_{xy} + (z_y^2 + 1)z_{xx}}{2(z_x^2 + z_y^2 + 1)^{3/2}} \quad (3.19)$$

while the area of an infinitesimal planar patch is

$$dA = \sqrt{z_x^2 + z_y^2 + 1} \quad (3.20)$$

Combining these gives an expression of the stabilizing functional for explicit surfaces

$$\begin{aligned} \Omega[z(x, y)] = \iint & \frac{[(z_x^2 + 1)z_{yy} - 2z_xz_yz_{xy} + (z_y^2 + 1)z_{xx}]^2}{2(z_x^2 + z_y^2 + 1)^{5/2}} \\ & - (1 - \nu) \frac{z_{xx}z_{yy} - z_{xy}^2}{(z_x^2 + z_y^2 + 1)^{3/2}} dx dy \end{aligned} \quad (3.21)$$

where the constant factor E due to the material has been dropped.

In short, the surface reconstruction problem is stated as a functional minimization problem. The invariant functional to be minimized acts to fit a thin plate approximation to the given slope data. Combining 3.10 and 3.21, the complete invariant functional is

$$\begin{aligned} M[z, \mathcal{S}] = \sum_{(p,q) \in \mathcal{S}} & \left(1 - \frac{pz_x + qz_y + 1}{\sqrt{(p^2 + q^2 + 1)(z_x^2 + z_y^2 + 1)}} \right) \\ & + \lambda \iint \frac{[(z_x^2 + 1)z_{yy} - 2z_xz_yz_{xy} + (z_y^2 + 1)z_{xx}]^2}{(z_x^2 + z_y^2 + 1)^{5/2}} \\ & - 2(1 - \nu) \frac{z_{xx}z_{yy} - z_{xy}^2}{(z_x^2 + z_y^2 + 1)^{3/2}} dx dy \end{aligned} \quad (3.22)$$

where constant factors have been combined with the regularizing parameter λ .

This formulation is in terms of the surface $z(x, y)$, but it is to be understood that the application will be in the gradient domain, with second derivatives of surface height becoming first derivatives of gradient.

This is a non-convex functional which will be difficult to minimize without making some approximations. The following sections present several possible approximations, which will be tested for accuracy and viewpoint invariance.

3.3 Convex Approximation

The reconstruction process has been stated as the minimization of the functional given by Equation 3.22. In order for a unique solution to exist, the functional must have a unique global minimum. A unique minimum guaranteed to exist and is much easier to find if the functional is convex. A functional $F[x]$ is convex over a solution space V if and only if (Maddox, 1988)

$$\forall x, y \in V, \beta \in [0, 1] : F[\beta x + (1 - \beta)y] \leq \beta F[x] + (1 - \beta)F[y] \quad (3.23)$$

Any quadratic function is necessarily convex.

Neither the smoothing functional nor data fit term of Equation 3.22 are convex. Several quadratic approximations are possible, and are detailed in the following sections.

3.3.1 Approximating the Constraint Metric

While the data fit term ρ_θ is not convex, it does have continuous derivatives and a single global minimum. This makes minimization by gradient descent methods feasible, but a quadratic functional like the integrability constraint ρ_{SD} of Equation 3.11 allows

greater efficiency (since the conjugate gradient descent method is intended for quadratic or approximately quadratic functionals—see Section 4.3.)

As stated previously, no invariant quadratic constraint metric exists (see Appendix A). We can instead use the second order Taylor series approximation of ρ_θ , expanded in terms of z_x and z_y about the point (p, q) :

$$\rho_\theta = \rho_{\text{Tay}} + O(z_x - p)^3 + O(z_y - q)^3 \quad (3.24)$$

where

$$\rho_{\text{Tay}}(z^*, (p, q)) = \frac{(z_x - p)^2 + (z_y - q)^2 + (z_x q - z_y p)^2}{(1 + p^2 + q^2)^2} \quad (3.25)$$

The Taylor series is not convergent (unless $(z_x - p) < 1$ and $(z_y - q) < 1$), and the expression ρ_{Tay} is not invariant, but it is quadratic and can be hoped to be an improvement over ρ_{SD} . (Note that where $p = q = 0$, ρ_{Tay} and ρ_{SD} are equivalent.)

For the experiments, all three constraint metrics— ρ_θ , ρ_{SD} , and ρ_{Tay} —were tried and compared.

3.3.2 Small Deflection Approximation of the Smoothing Functional

A common approximation to Equation 3.21 comes from the assumption that z_x and z_y are small, yielding (Courant & Hilbert, 1989; Volterra & Gaines, 1971)

$$\Omega_1[z(x, y)] = \iint (z_{xx} + z_{yy})^2 - 2(1 - \nu)(z_{xx}z_{yy} - z_{xy}^2) dx dy \quad (3.26)$$

Setting $\nu = 0$ yields the quadratic variation stabilizer, which has often been used for surface reconstruction (Grimson, 1981; Terzopoulos, 1988).

This approximation goes back to Lagrange (Volterra & Gaines, 1971) and is frequently used in the mechanical engineering field, where the bending of plates is analyzed in a

coordinate system based on the rest state of the plate and the amount of bending is usually small. For surface reconstruction, where the coordinate system is arbitrarily imposed by the viewing direction and surfaces may have many folds and wrinkles, the approximation seems less appropriate.

3.3.3 First Derivative Parameterization

A more accurate approximation can be derived if estimates of the first derivatives z_x and z_y are available. Given values of p and q as estimates of z_x and z_y , Equation 3.21 can be written as

$$\Omega_2[z(x, y)] = \iint (Az_{xx} + Bz_{xy} + Cz_{yy})^2 - 2(1 - \nu)D(z_{xx}z_{yy} - z_{xy}^2) dx dy \quad (3.27)$$

where A - D are constant at each point on the surface, and are defined as

$$A = \frac{1 + p^2}{(1 + p^2 + q^2)^{\frac{5}{4}}} \quad (3.28)$$

$$B = -\frac{pq}{(1 + p^2 + q^2)^{\frac{5}{4}}} \quad (3.29)$$

$$C = \frac{1 + q^2}{(1 + p^2 + q^2)^{\frac{5}{4}}} \quad (3.30)$$

$$D = \frac{1}{(1 + p^2 + q^2)^{\frac{3}{2}}} \quad (3.31)$$

(Note that A - D are all less than 1 in absolute value.)

This is the first attempt at an invariant parameterization made by (Stevenson & Delp, 1992)⁶.

3.3.4 Principal Directions Parameterization

An alternative parameterization can be based on the first and second fundamental forms of the surface (Pogorelov, undated), which serve as a unique local characterization of the

⁶In their paper, A and B are erroneously squared—this is a typographic error (Stevenson, personal correspondence).

surface. The fundamental forms can be used to derive the principal curvatures of the surface, from which the smoothing functional naturally follows.

First fundamental form. An explicit surface can be represented as the set of points in \mathbb{R}^3 of the form $\mathbf{z} = (x, y, z(x, y))$. The first fundamental form at such a point is

$$\begin{aligned} I &= d\mathbf{z}^2 \\ &= (dx, dy, z_x dx + z_y dy)^2 \\ &= dx^2 + dy^2 + z_x^2 dx^2 + 2z_x z_y dx dy + z_y^2 dy^2 \end{aligned} \quad (3.32)$$

The first fundamental form is often expressed in terms of the parameters E , F , and G as

$$\begin{aligned} I &= [dx \ dy] \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \\ &= Edx^2 + 2Fdx dy + Gdy^2 \end{aligned} \quad (3.33)$$

where

$$\begin{aligned} E &= \mathbf{z}_x^2 = (1 + z_x^2) \\ F &= \mathbf{z}_x \cdot \mathbf{z}_y = z_x z_y \\ G &= \mathbf{z}_y^2 = (1 + z_y^2) \end{aligned} \quad (3.34)$$

Second fundamental form. The second fundamental form is defined similarly as

$$II = -d\mathbf{z} \cdot d\mathbf{n} \quad (3.35)$$

where \mathbf{n} is the unit surface normal at the point, which can be expressed as

$$\begin{aligned} \mathbf{n} &= \frac{\mathbf{z}_x \times \mathbf{z}_y}{\|\mathbf{z}_x \times \mathbf{z}_y\|} \\ &= \frac{(-z_x, -z_y, 1)}{\sqrt{z_x^2 + z_y^2 + 1}} \end{aligned} \quad (3.36)$$

The second fundamental form is generally expressed in terms of the parameters L , M , and N as

$$\begin{aligned} II &= [dx \ dy] \begin{bmatrix} L & M \\ M & N \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \\ &= Ldx^2 + 2Mdx \, dy + Ndy^2 \end{aligned} \quad (3.37)$$

where

$$\begin{aligned} L &= -\mathbf{z}_x \cdot \mathbf{n}_x = \mathbf{z}_{xx} \cdot \mathbf{n} = \frac{z_{xx}}{\sqrt{z_x^2 + z_y^2 + 1}} \\ M &= -(\mathbf{z}_x \cdot \mathbf{n}_y + \mathbf{z}_y \cdot \mathbf{n}_x) = \mathbf{z}_{xy} \cdot \mathbf{n} = \frac{z_{xy}}{\sqrt{z_x^2 + z_y^2 + 1}} \\ N &= -\mathbf{z}_y \cdot \mathbf{n}_y = \mathbf{z}_{yy} \cdot \mathbf{n} = \frac{z_{yy}}{\sqrt{z_x^2 + z_y^2 + 1}} \end{aligned} \quad (3.38)$$

The forms I and II can be thought of as expressing properties of curves on the surface $z(x, y)$. Suppose that γ is the curve formed by the intersection of the surface with a plane perpendicular to the tangent plane at a point P . If the direction of the curve at P is given as (dx, dy) , then the curvature of γ at P is (Pogorelov, undated)

$$k_{dx \, dy} = \frac{L \, dx^2 + 2M \, dx \, dy + N \, dy^2}{E \, dx^2 + 2F \, dx \, dy + G \, dy^2} = \frac{II}{I} \quad (3.39)$$

The quantity $k_{dx \, dy}$ is known as the *normal curvature* of the surface in the direction $(dx : dy)$ ⁷. This expression can be used to calculate the smoothing functional of Equation 3.12, since the principal curvatures κ_1 and κ_2 are just the normal curvatures in the principal directions.

Given the principal directions $\{(dx_1 : dy_1), (dx_2 : dy_2)\}$, the principal curvatures can be written as

$$\kappa_1 = \frac{1}{\sqrt{z_x^2 + z_y^2 + 1}} (\omega_1 z_{xx} + \omega_2 z_{xy} + \omega_3 z_{yy})$$

⁷If the direction is specified as an angle θ measured from the tangent in the x -direction, then the normal curvature can be calculated from $dx = \cos \theta, dy = \sin \theta$.

$$\kappa_2 = \frac{1}{\sqrt{z_x^2 + z_y^2 + 1}} (\omega_4 z_{xx} + \omega_5 z_{xy} + \omega_6 z_{yy}) \quad (3.40)$$

where

$$\begin{aligned} \omega_1 &= \frac{dx_1^2}{E dx_1^2 + 2F dx_1 dy_1 + G dy_1^2} \\ \omega_2 &= \frac{2 dx_1 dy_1}{E dx_1^2 + 2F dx_1 dy_1 + G dy_1^2} \\ \omega_3 &= \frac{dy_1^2}{E dx_1^2 + 2F dx_1 dy_1 + G dy_1^2} \\ \omega_4 &= \frac{dx_2^2}{E dx_2^2 + 2F dx_2 dy_2 + G dy_2^2} \\ \omega_5 &= \frac{2 dx_2 dy_2}{E dx_2^2 + 2F dx_2 dy_2 + G dy_2^2} \\ \omega_6 &= \frac{dy_2^2}{E dx_2^2 + 2F dx_2 dy_2 + G dy_2^2} \end{aligned} \quad (3.41)$$

The principal directions can be calculated from the Weingarten mapping

$$\beta = \begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{bmatrix} L & M \\ M & N \end{bmatrix} \quad (3.42)$$

The principal directions are simply the eigenvectors of β : $\mathbf{e}_i = (dx_i, dy_i)$.

An approximation to the principal curvatures of Equation 3.40 can be derived if estimates p and q of z_x and z_y are available, from which the Weingarten mapping can be calculated, and thus the principal directions can be derived. This yields estimates of the parameters ω_1 – ω_6 at each point in the image⁸.

The smoothing functional (see Equation 3.13) now has the convex approximation (again, neglecting a constant factor)

$$\Omega_3[z(x, y)] = \int_U (\hat{\kappa}_1^2 + \hat{\kappa}_2^2) + 2\nu \hat{\kappa}_1 \hat{\kappa}_2 dA$$

⁸To estimate the second fundamental form it is necessary to numerically differentiate p and q in order to estimate second derivatives of surface height. This is the only real difference between Ω_3 and Ω_2 . This method of arriving at Ω_3 is not terribly good (as explained in the conclusions section).

$$\begin{aligned}
&= \iint \frac{1}{\sqrt{z_x^2 + z_y^2 + 1}} [(\hat{\omega}_1 z_{xx} + \hat{\omega}_2 z_{xy} + \hat{\omega}_3 z_{yy})^2 + (\hat{\omega}_4 z_{xx} + \hat{\omega}_5 z_{xy} + \hat{\omega}_6 z_{yy})^2 \\
&\quad + 2\nu(\hat{\omega}_1 z_{xx} + \hat{\omega}_2 z_{xy} + \hat{\omega}_3 z_{yy})(\hat{\omega}_4 z_{xx} + \hat{\omega}_5 z_{xy} + \hat{\omega}_6 z_{yy})] dx dy \quad (3.43)
\end{aligned}$$

where $\hat{\omega}_1$ - $\hat{\omega}_6$ are the estimated parameters derived from the Weingarten mapping. The parameters are derived from estimates of the principal directions—at singularities in the principal direction field (umbilic points), any arbitrary orthogonal directions can be chosen, since the curvatures are the same in all directions. The factor $\sqrt{z_x^2 + z_y^2 + 1}$ in the denominator can be approximated by $\sqrt{p^2 + q^2 + 1}$.

This is the approximation used by (Stevenson & Delp, 1992), with $\nu = 0^9$.

Any of the smoothing functionals Ω_1 - Ω_3 can be combined with any of the data fit terms ρ_θ , ρ_{SD} or ρ_{Txy} (Equations 3.10, 3.11, and 3.25) to yield a regularizing functional. The minimization of the functional yields a complete, smoothed gradient map. The invariance of the reconstruction will depend on how accurately the approximation chosen reflects the original invariant functional. This thesis seeks to evaluate and compare the different approximations on this invariance criterion.

3.4 Recovering Height from Gradient

Application of the smoothing functional is designed to produce a complete and coherent gradient map, which leaves the task of recovering the actual surface height. This is an integration process, which seeks to find the surface estimate which comes closest to being the integral of the gradient map. (Since integrability has not been enforced, direct integration is not possible.) Separation of processing for gradient smoothing and height recovery was chosen as a means of simplifying the computation, but is rather an artificial split (see Chapter 7).

⁹The factor $(z_x^2 + z_y^2 + 1)^{-\frac{1}{2}}$ is eliminated in Stevenson and Delp's formulation by virtue of the minimization scheme they have chosen (Stevenson, personal correspondence).

This can easily be done by minimizing a functional analagous to the integrability constraint of Equation 3.11, or any of the other data fit functionals used while smoothing the gradient map. The difference now is that the penalty term is applied for every point in the smoothed gradient map, rather than just at the points corresponding to the original data. In short, we find the surface which minimizes

$$R[z] = \iint \rho((z_x, z_y), (p, q)) dx dy \quad (3.44)$$

The squared differences constraint, being quadratic, is easier to minimize, while ρ_θ meets our desired invariance criteria, but is non-convex and hence harder to minimize. Alternatively, $\rho_{T_{xy}}$ is quadratic and approximately invariant.

Note that during the stages of gradient map smoothing and surface height reconstruction, only first order derivatives ever need to be numerically estimated, a benefit of working with slope data rather than range data.

4

Surface Reconstruction on a Regular Grid

Surface reconstruction has been formalized as a problem of minimizing a functional that imposes smoothness and fit to the data constraints on the surface (see Equation 3.22). The previous chapter derived an invariant regularizing functional and several quadratic approximations, all in the continuous domain. This chapter presents a discretization of the problem and an algorithm for its solution.

The task considered is to derive robust estimates of surface orientation and relative distance at each point on a regular grid, given noisy estimates of surface orientation at a possibly sparse set of points on the grid. The grid corresponds to a digital image, and it is assumed that the imaging process can be modeled by orthographic projection. The problem of dealing with depth and orientation discontinuities, while important in the general case, is only addressed superficially here (see Section 4.4).

The problem is solved in three stages (see Figure 4.1):

1. initial interpolation of the gradient data;
2. smoothing of the gradient map; and
3. integrating gradient to recover surface height.

The continuous formulations presented in Chapter 3 are recast in the discrete domain in Section 4.1, using a finite difference method. A method for interpolating the initial gradient data is explained in Section 4.2. Section 4.3 describes gradient descent methods

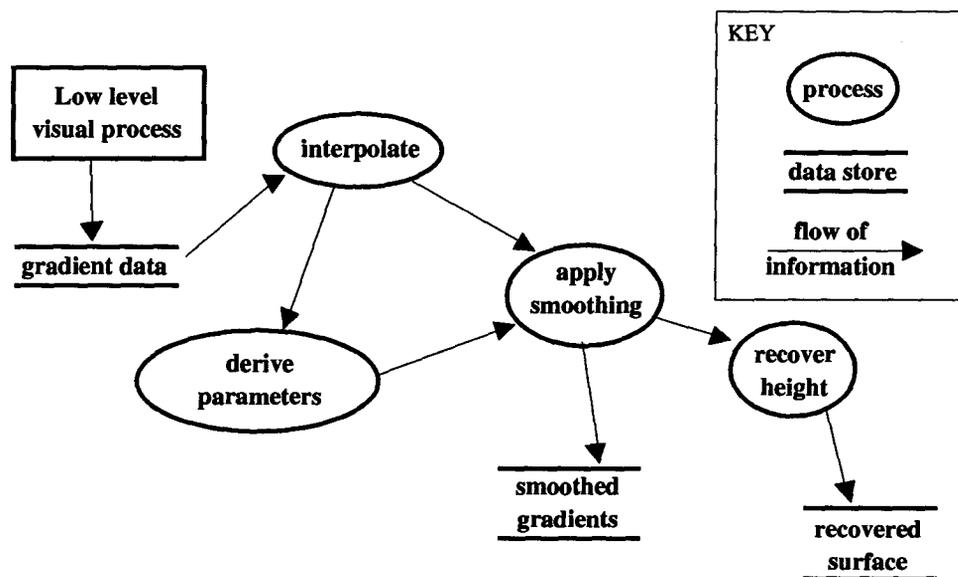


Figure 4.1: Surface reconstruction from gradient data.

for minimization of the discretized problem. Boundary conditions apply at the edges of the reconstruction, and are discussed in Section 4.4.

The algorithm was implemented in Lucid Common Lisp on a Sparc workstation.

4.1 Discretization

In Chapter 3, the surface reconstruction problem was stated as the variational problem of minimizing a functional defined over a continuous surface height function $z(x, y)$. For a discrete surface height function defined at points in a digital image, the problem becomes one of minimizing a function $M(\mathbf{z})$, where \mathbf{z} is an array of surface height values corresponding to points in the image.

The function M has the form of Equation 3.22, but in the discretized problem, the double integral is replaced by a double summation over the rows and columns of the image, and the values of derivatives are replaced with their numerical estimates. For an $M \times N$ image, the discrete versions of the various smoothing functionals are thus (see Equations 3.26, 3.27, and 3.43)

$$\Omega_1[z(x, y)] = \sum_{i=1}^M \sum_{j=1}^N (\hat{z}_{xx} + \hat{z}_{yy})^2 - 2(1 - \nu)(\hat{z}_{xx}\hat{z}_{yy} - \hat{z}_{xy}^2) \quad (4.1)$$

$$\Omega_2[z(x, y)] = \sum_{i=1}^M \sum_{j=1}^N (\hat{A}\hat{z}_{xx} + \hat{B}\hat{z}_{xy} + \hat{C}\hat{z}_{yy})^2 - 2(1 - \nu)\hat{D}(\hat{z}_{xx}\hat{z}_{yy} - \hat{z}_{xy}^2) \quad (4.2)$$

$$\begin{aligned} \Omega_3[z(x, y)] = \sum_{i=1}^M \sum_{j=1}^N \frac{1}{\sqrt{z_x^2 + z_y^2 + 1}} & [(\hat{\omega}_1\hat{z}_{xx} + \hat{\omega}_2\hat{z}_{xy} + \hat{\omega}_3\hat{z}_{yy})^2 + (\hat{\omega}_4\hat{z}_{xx} + \hat{\omega}_5\hat{z}_{xy} + \hat{\omega}_6\hat{z}_{yy})^2 \\ & + 2\nu(\hat{\omega}_1\hat{z}_{xx} + \hat{\omega}_2\hat{z}_{xy} + \hat{\omega}_3\hat{z}_{yy})(\hat{\omega}_4\hat{z}_{xx} + \hat{\omega}_5\hat{z}_{xy} + \hat{\omega}_6\hat{z}_{yy})] \end{aligned} \quad (4.3)$$

where z_x and z_y are the variables to be solved for during the smoothing process, and the second derivative estimates \hat{z}_{xx} , \hat{z}_{xy} , and \hat{z}_{yy} are formed as described in Section 4.1.2 (Equation 4.11). For each term in the summation, the values are taken from the corresponding point (i, j) in the image.

Similarly, the functional for recovery of surface height from the smoothed gradient map (Equation 3.44) is discretized as:

$$R(\mathbf{z}) = \sum_{i=1}^M \sum_{j=1}^N \rho((\hat{z}_x, \hat{z}_y), (p, q)) \quad (4.4)$$

with first derivatives estimated by Equation 4.12. Using the different constraint metrics derived previously yields three possible functionals for surface height recovery

$$R_{\text{SD}}(\mathbf{z}) = \sum_{i=1}^M \sum_{j=1}^N \rho_{\text{SD}}((\hat{z}_x, \hat{z}_y), (p, q)) \quad (4.5)$$

$$R_{\theta}(\mathbf{z}) = \sum_{i=1}^M \sum_{j=1}^N \rho_{\theta}((\hat{z}_x, \hat{z}_y), (p, q)) \quad (4.6)$$

$$R_{\text{Tay}}(\mathbf{z}) = \sum_{i=1}^M \sum_{j=1}^N \rho_{\text{Tay}}((\hat{z}_x, \hat{z}_y), (p, q)) \quad (4.7)$$

Rather than directly discretizing the continuous functionals, an alternative approach is to derive the Euler-Lagrange equations for the continuous problem statement, which are then discretized to form the basis of an iterative solution. (See (Courant & Hilbert, 1989) for an explanation of the Euler-Lagrange equations of a variational problem, and (Horn & Brooks, 1986) for examples of their application to shape from shading.) Discretization of the original functional was chosen as the more direct way to tackle the problem.

4.1.1 Regular Grid

Surface reconstruction will be performed over an $M \times N$ grid. The grid points are h units apart in the x direction and v units apart in the y direction. Each intersection corresponds to a pixel in a video image, with h and v indicating distances in the world corresponding to the width and height of a single pixel. This assumes that the imaging process can be modelled as orthographic projection, with

$$x = \frac{X}{h} \quad y = \frac{Y}{v} \quad (4.8)$$

where (X, Y, Z) is a point in the world corresponding to position (x, y) in the image grid¹. The parameters h and v , determined during camera calibration, encode information about the average distance to the scene, aspect ratio, focal length, analog to digital conversion, etc. Data values are available at a subset of the pixels in the image.

4.1.2 Approximating Derivatives with Finite Differences

Given a small $h \times v$ rectangular patch in the discretized gradient map, with estimates of the gradient (p, q) at each of the four corners, a quadratic surface height function can be defined over the patch

$$z(x, y) = ax^2 + by^2 + cxy + dx + ey + f \quad (4.9)$$

The first derivatives of this height function are

$$\begin{aligned} z_x(x, y) &= 2ax + cy + d \\ z_y(x, y) &= 2by + cx + e \end{aligned} \quad (4.10)$$

which will be equal to the known values of p and q at the corners of the patch. Note that these patches are nonconformal, since adjacent patches will not in general coincide along their edges (except, by construction, at the corners).

Plugging in the known values of p, q, x , and y at the corners of the rectangular patch yields a system of eight linear equations in the five unknowns a – e (note that gradient information alone places no constraint on the constant term f). A least squares solution to this system provides estimates of the second derivatives of $z(x, y)$, corresponding to first derivatives of the gradient map

$$\hat{z}_{xx} = 2a = \frac{p_{10} - p_{00} + p_{11} - p_{01}}{2h} = p_x$$

¹This is sometimes referred to as a *weak perspective* model of projection, as opposed to true orthographic projection where $x = X, y = Y$.

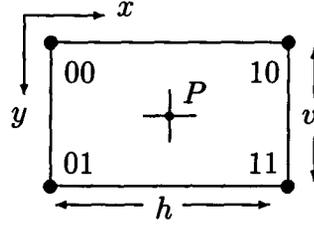


Figure 4.2: A rectangular patch. The derivative calculation yields first derivative estimates valid at point P in the center of the patch.

$$\begin{aligned}
 \hat{z}_{yy} &= 2b = \frac{q_{01} - q_{00} + q_{11} - q_{10}}{2v} = q_y \\
 \hat{z}_{xy} &= c = \frac{v(p_{01} - p_{00} + p_{11} - p_{10}) + h(q_{10} - q_{00} + q_{11} - p_{01})}{2(h^2 + v^2)} \\
 &= \frac{v^2 p_y + h^2 q_x}{h^2 + v^2}
 \end{aligned} \tag{4.11}$$

where the numerical subscripts of p and q indicate different corners of the rectangular patch (see Figure 4.2). One might expect the estimate of z_{xy} to be $\frac{1}{2}(p_y + q_x)$ —observe that this is only the case when $h = v$.

For reconstruction of height from gradient, the first derivatives of surface height need to be computed. This can be done in an exactly analogous fashion

$$\begin{aligned}
 \hat{z}_x &= \frac{z_{10} - z_{00} + z_{11} - z_{01}}{2h} \\
 \hat{z}_y &= \frac{z_{01} - z_{00} + z_{11} - z_{10}}{2v}
 \end{aligned} \tag{4.12}$$

These calculations can be performed by convolution with the simple masks

$$\frac{1}{2h} \begin{array}{|c|c|c|} \hline -1 & & +1 \\ \hline -1 & & +1 \\ \hline \end{array} \quad \text{and} \quad \frac{1}{2v} \begin{array}{|c|c|} \hline -1 & -1 \\ \hline +1 & +1 \\ \hline \end{array} \tag{4.13}$$

Note that this method of calculating derivatives has the desirable property that the estimates of z_x and z_y are valid at the same point, unlike a scheme such as $\hat{z}_x = z_{10} - z_{00}$, $\hat{z}_y = z_{01} - z_{00}$.

Calculating derivatives in this way yields estimates that are valid at points offset by half a pixel in x and y from the original grid. This has implications for the iterative reconstruction process, since data values and estimated parameters are valid at slightly different points on the surface than the derivative estimates. In practice, it has been found that the only substantial effect of the staggered grids is that the reconstructed surface is offset by half a pixel from the data. If necessary, this can easily be compensated for by resampling.

4.2 Initial Interpolation of the Gradient Map

The initial input to the surface reconstruction process is a set of possibly noisy gradient constraints measured at discrete points on the surface. These constraints need to be interpolated to fill in missing values and smoothed to reduce the effects of noise. Rather than immediately applying the full smoothing functional, it is useful to first interpolate the data, in order to get closer to the minimum before incurring the computational overhead of more complex later stages.

The initial interpolation is done by component-wise interpolation of the surface normals given by the gradient data. This estimate passes directly through the data points and does nothing to minimize noise, which will be compensated for by later smoothing stages. The interpolation proceeds as follows.

Consider the abstract case of a sparse data set $\{u_i(x_i, y_i) | i = 1, \dots, N\}$ defined at discrete points in the (x, y) plane. We can fit a piecewise planar surface \hat{u} to the data by minimizing the small deflection approximation of the membrane energy functional (Courant & Hilbert, 1989)

$$\iint (\hat{u}_x^2 + \hat{u}_y^2) dx dy \quad (4.14)$$

subject to the constraints $\hat{u}(x_i, y_i) = u(x_i, y_i)$. Terzopoulos (Terzopoulos, 1988) uses

this penalty term as part of a controlled continuity stabilizer, as it allows orientation discontinuities.

This interpolation scheme could be applied separately to the p and q components of a set of gradient constraints², equivalent to minimizing

$$\iint (p_x^2 + p_y^2 + q_x^2 + q_y^2) dx dy \quad (4.15)$$

which is just the quadratic variation of the underlying surface. This is essentially equivalent to the smoothing functional used by (Grimson, 1981) for surface reconstruction from depth constraints.

However, this functional is not viewpoint invariant. Another approach is to calculate the unit surface normals corresponding to the known values of p and q from

$$\mathbf{n}(p, q) = (n_x, n_y, n_z) = (-p, -q, 1) / \sqrt{p^2 + q^2 + 1} \quad (4.16)$$

Each component n_x, n_y, n_z of the surface normal “map,” can be interpolated as described above. This process is viewpoint invariant, and (provided the angle between nearby normals is less than about $\frac{\pi}{2}$) is roughly equivalent to linear interpolation of the angular separation between adjacent normals. Following interpolation, the normals are converted back to gradient values by

$$p(\mathbf{n}) = -\frac{n_x}{n_z} \quad (4.17)$$

$$q(\mathbf{n}) = -\frac{n_y}{n_z} \quad (4.18)$$

The interpolation is carried out using the simple algorithm of Appendix B. It is this interpolated gradient map that forms the input to the second stage of the reconstruction process. The interpolated gradient map is also used to calculate the parameters required to form Ω_2 and Ω_3 .

²Note that C^0 continuous interpolation of gradients yields C^1 continuity of the corresponding surface height.

4.3 Minimization of the Discretized Functional

Discretization of the variational problem yields a problem of function minimization. Consider a functional $M[f]$, where f is a function. We can think of M as a function of infinitely many variables, corresponding to the values $z = f(\mathbf{x})$ at each point \mathbf{x} in the domain of f . Discretizing the domain of f approximates the functional $M[\cdot]$ with a function $M(\cdot)$ of a finite (though large) number of variables—the values of f at each node in the discretization.

The discrete problem is to find \mathbf{z} such that

$$M(\mathbf{z}) = \min_{\mathbf{z}^*} M(\mathbf{z}^*) \quad (4.19)$$

The typically large number of variables foils attempts at an analytic solution, but various iterative numerical methods are available. Many of these methods are based on the idea of gradient descent—starting at some point in the parameter space of the function, keep moving “downhill” along the surface of the function until the minimum is reached (Faddeev & Faddeeva, 1963).

Steepest descent is the simplest of these methods. Given an initial point \mathbf{z}_0 , the gradient of the function is computed at that point, corresponding to the locally steepest direction on the function surface. The function is minimized along this direction³, yielding a new point \mathbf{z}_1 . More precisely

$$\begin{aligned} \mathbf{g}_k &= \nabla M(\mathbf{z}_k) \\ \alpha_k &= \arg \min_{\gamma \in \mathbb{R}} M(\mathbf{z}_k - \gamma \mathbf{g}_k) \\ \mathbf{z}_{k+1} &= \mathbf{z}_k - \alpha_k \mathbf{g}_k \end{aligned} \quad (4.20)$$

³This is an optimistic form of steepest descent. A more cautious approach is to take small steps of a predetermined size in the direction of steepest descent, perhaps updating the step size via some heuristic.

where this process is repeated until a suitable stopping criterion is met⁴. Note that the line minimization to find α_k may itself be an iterative approximation process, depending on the form of $M()$.

The trouble with steepest descent is that while each new direction is perpendicular to the previous direction (minimizing the function in a particular direction means that the component of the gradient in that direction is zero), it may be close to parallel to some prior direction. This results in the algorithm taking many more iterations than should be necessary. For example, if the minimum is in a “valley” of the function, we would like the algorithm to move to the floor of the valley in one step and to the minimum in the next. Steepest descent will not do this, but instead will jump back and forth across the valley, taking many small steps before it finally reaches the minimum.

The method of conjugate gradient descent provides faster convergence by seeking to optimize the direction in which to move at each step. Furthermore, for a quadratic function, the conjugate gradient method guarantees convergence after at most n steps, where n is the number of unknowns in \mathbf{z} (Faddeev & Faddeeva, 1963).

The basic idea is to keep moving in a direction that minimizes the function over the subspace spanned by the previously chosen direction vectors. Appendix C presents an overview of the algorithm; see (Grimson, 1981) or (Press *et al*, 1988) for further details.

All such gradient descent methods are vulnerable to getting stuck in local minima. Various stochastic techniques exist which attempt to avoid this problem (Blake, 1989), but they tend to be very computationally intensive. The simplest alternative is to stick to convex functions, in particular quadratics, which is why so much effort was taken to provide quadratic approximations to the invariant functional.

Once the gradient map has been found to solve the minimization problem, we can recover height from gradient by solving another discrete minimization problem. The

⁴e.g. the norm of the gradient, $\|\mathbf{g}_k\|$, is below some tolerance.

continuous solution is described in Section 3.4. We simply apply the gradient descent minimization techniques to the discretized versions of the continuous problem (see Equations 4.5, 4.6, and 4.7).

4.4 Boundary Conditions

The three stages of the surface reconstruction process—gradient interpolation, gradient smoothing, and recovery of height from gradient—all act to recover functions by imposing conditions on the function over some region of interest. Additional issues are raised in considering the boundaries of the region.

The Dirichlet boundary conditions are used when values can be prescribed at the boundary. If appropriate boundary conditions are not known, then rather than prescribing arbitrary values (say 0), the Neumann, or natural, boundary conditions can be used. For the gradient smoothing stage, the natural boundary conditions are:

$$\frac{\partial p}{\partial n} = \frac{\partial q}{\partial n} = 0 \quad \text{and} \quad \frac{\partial z}{\partial n} = (p, q) \cdot \mathbf{n} \quad (4.21)$$

where \mathbf{n} is the outward normal to the boundary (Ascher & Carter, 1993). That is, p and q are extended linearly beyond the boundary, while the changes in z across the boundary are obtained from the values of p and q . The natural boundary conditions can be derived from the Euler equations for the problem (Courant & Hilbert, 1989). Alternatively, if the functionals are discretized directly then the natural boundary conditions simply fall out of the discretization.

(Li, 1991) showed that natural boundary conditions work well for surface reconstruction from slope data, except at self-occluding object boundaries. At occluding boundaries, it is desirable to have prescribed values or to use Dirichlet conditions.

In this work, direct discretization of the functionals over the image ensures that the

natural boundary conditions are implemented at the boundaries of the image, but not at object boundaries *within* the image. This is a serious flaw if the goal is accurate surface reconstruction, but handling of such discontinuities is an issue that is not addressed. Provision has been made for inclusion of boundary locations—if the algorithm is provided with the location of discontinuities then derivatives will not be calculated across these boundaries and values for gradient or height can be explicitly set. In particular, the photometric stereo process yields an object silhouette image which can be used to find occluding boundaries⁵.

⁵In practice, use of this method of setting boundary conditions did not prove useful and was not used.

Investigations with Synthetic Data

A method has been described for reconstructing visual surfaces from gradient data. This chapter presents the results of a number of experiments with synthetic data devised to evaluate the method. Although reconstruction accuracy is important in general, this study focuses primarily on the view variation of the smoothing functionals. Smoothing and accuracy can conflict, but finding the optimal degree of smoothing is not considered.

Section 5.1 describes the synthetically generated data sets. Section 5.2 discusses problems with the invariant constraint metrics. Use of different functionals for determining height from gradient is discussed in Section 5.3. The results obtained using different smoothing functionals are presented in Section 5.4. A “ λ normalization” technique required for the comparison of results from different reconstruction functionals is described in Section 5.5. Functionals for surface reconstruction were developed with pointwise invariance in mind—the problem of invariance to reparameterization is considered in Section 5.6. The effects of noise in the data are investigated in Section 5.7. Section 5.8 demonstrates the results of varying the regularizing and smoothing parameters.

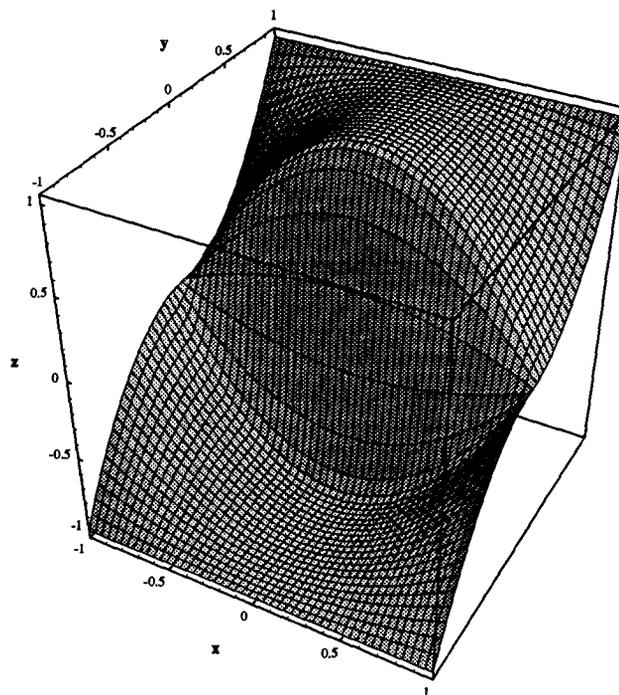


Figure 5.1: Synthetically generated surface height function.

5.1 Synthetic Data Sets

A smoothly varying synthetic surface was generated that contains a wide range of surface normal directions. The mathematical form of the surface is

$$z(x, y) = 2(1 + x^2(y^2 - 1))(CND_{0, \frac{1}{15}}(y) - 1/2) \quad (5.1)$$

where $CND_{\mu, \sigma}$ is the one-dimensional cumulative normal distribution function with mean μ and standard deviation σ . The surface was sampled at points on a 50×50 grid laid over the range $x, y \in [-1, 1]$ (see Figure 5.1). This sampling yields values for the grid spacing parameters of $h = v = 2/50 = 0.04$. The input to the reconstruction process is the gradient at each point (Figures 5.2 and 5.3).

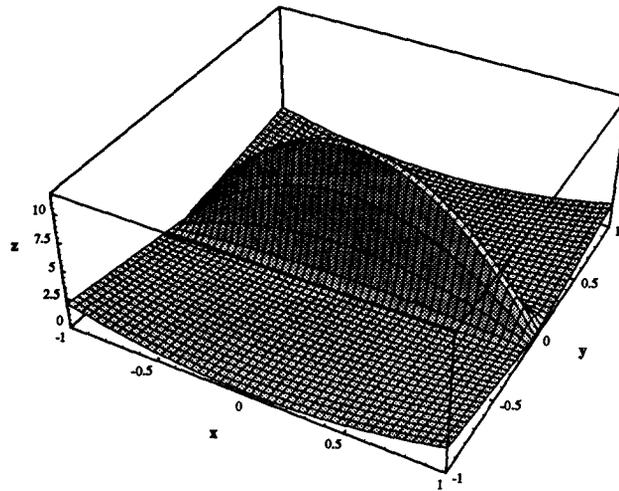


Figure 5.2: First component of the gradient map— z_x values for synthetic surface.

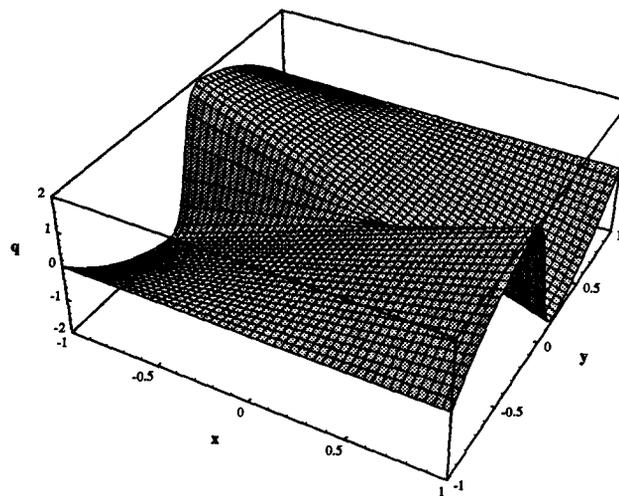


Figure 5.3: Second component of the gradient map— z_y values for synthetic surface.

Two more data sets were generated by rotating the analytical surface by 30° and 60° about the x axis and resampling over the range $x \in [-1, 1], y \in [-\frac{1+\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2}]$. This corresponds to the same “piece” of the surface as that for the unrotated data set. For the rotated data sets, the grid spacing parameters are $h = 0.04, v = (1 + \sqrt{3})/50 = 0.0546$. The unrotated surface and its two rotations will be denoted S_0, S_{30} and S_{60} . Figure 5.4 shows a cross-section of the three surfaces¹.

The basic method of evaluating the invariance of a surface reconstruction method will be to reconstruct the surface for each of the data sets, then rotate the resulting surfaces into a canonical position and see how well they match. Matches will be evaluated qualitatively by comparing surface plots and quantitatively by measuring the distance between different reconstructions.

The functional to be minimized (Equation 3.22 and its discrete approximations) contains the smoothing parameter λ , controlling the tradeoff between smoothness and fit to the data, and the parameter ν relating to the thin-plate material (see Section 3.2.3). For accurate surface reconstruction, many techniques exist for determining a value for λ that is appropriate for the available data (Wahba, 1990).

This thesis seeks to evaluate different smoothing functionals, and so a rather larger λ was used than might have been necessary, in order to emphasise the effects of smoothing. In addition, the values used in conjunction with different smoothing functionals must be equalised to allow meaningful comparison of the results. The value $\lambda_0 = 1.0$ was used throughout the experiments, with the “normalization” process described in Section 5.5. The choice of ν had little effect on the results, and $\nu = 0.0$ was used throughout. The effects of varying these parameters are demonstrated in Section 5.8.

¹Points in 2D plots and grid intersections in 3D plots correspond to actual data points. For the 3D plots, the data has been sub-sampled, so not all points in the data set are plotted.

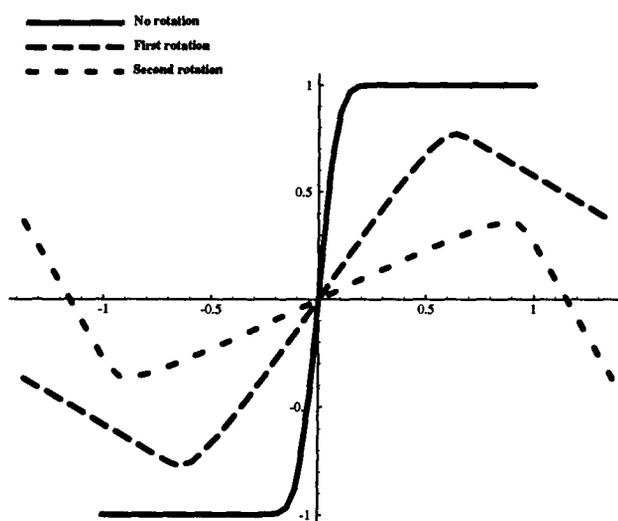


Figure 5.4: Cross section of row 25 of the three synthetic surfaces.

5.2 Failure of Invariant Constraint Metric

On attempting to use the invariant constraint metric ρ_θ its inadequacy was immediately apparent (see Figure 5.5). This failure is due to the small range of values taken on by the constraint metric (since $0 \leq (1 - \cos\theta) \leq 2$) compared to the large range for the smoothing functional (infinite at discontinuities). The data fit term is completely overwhelmed by the smoothing term at points of high curvature on the surface, meaning that these points are oversmoothed. The linear tradeoff provided by the regularizing parameter λ is not sufficient to combat this effect. ρ_{Tay} suffers from similar defects.

Tikhonov and Arsenin, in their treatment of the solution of ill-posed problems, assume that L_2 norms are to be used (Tikhonov & Arsenin, 1977). Of the constraint metrics studied here, only ρ_{SD} is an L_2 norm. The other constraint metrics may thus be inappropriate, but the mathematics has not been pursued.

Figure 5.6 plots a slice through the “derivative surface” for ρ_θ and ρ_{SD} . The plot shows the partial derivative of the constraint metric with respect to a particular element of the p component of the gradient map. Although both plots have the same sign, the magnitude of ρ_θ becomes very small across the jump in height, thereby failing to enforce data fit at those points and yielding to oversmoothing. ρ_{Tay} suffers from similar problems. (The calculation is performed using an intermediate gradient map estimate during smoothing of S_0 with Ω_2 and ρ_{SD} .)

The constraint metric ρ_{SD} based on squared differences (see Equation 3.11) provides a better match with the smoothing functionals used, and will be employed for the remaining examples.

On a purely practical note, the quadratic metric ρ_{SD} yields convergence in fewer and faster iterations than the doubly iterative solution using ρ_θ .

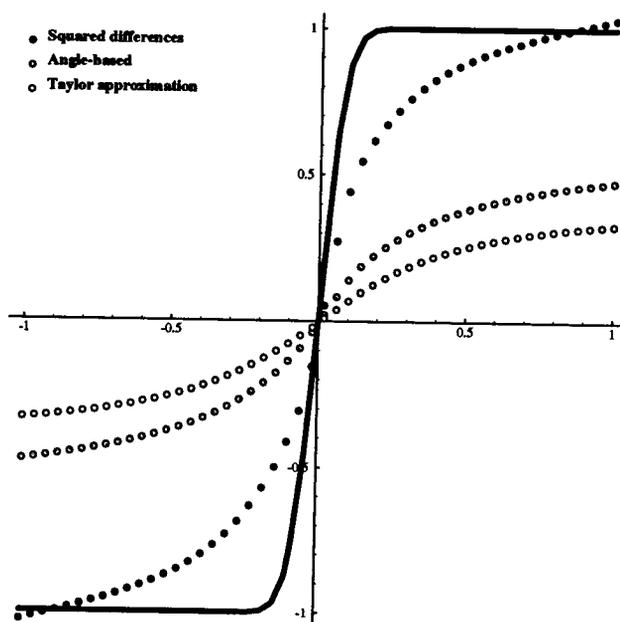


Figure 5.5: The inadequacy of the angle-based constraint metrics is demonstrated by their application to the unrotated surface. The smoother is Ω_2 in each case.

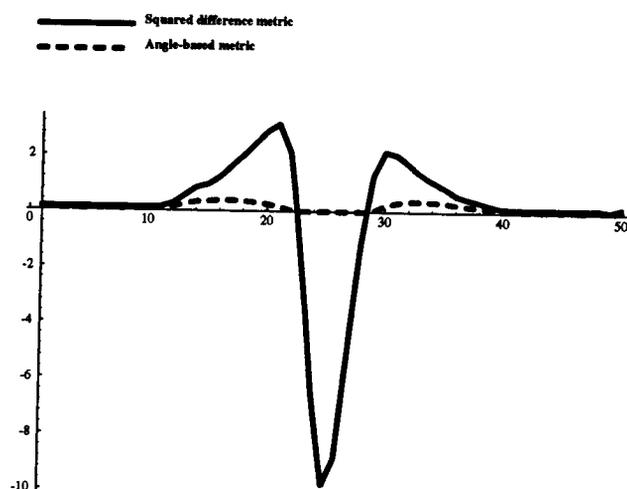


Figure 5.6: Differential surfaces for the two constraint metrics.

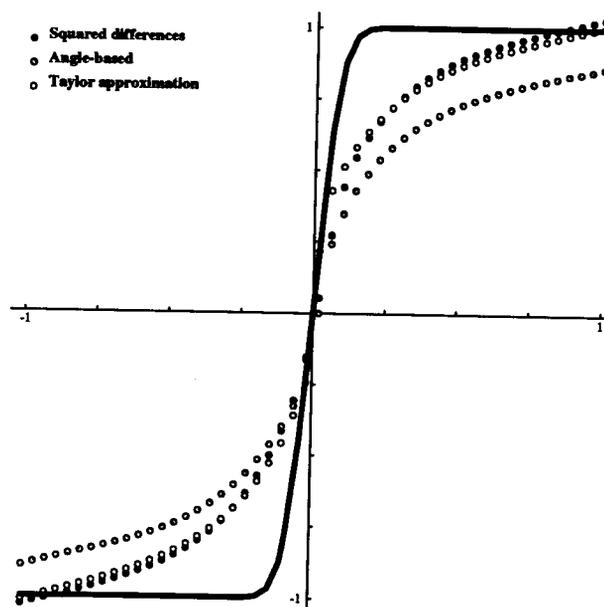


Figure 5.7: Height recovered with different integration functionals (after smoothing). The solid line is the corresponding slice through the original surface.

5.3 Use of Different Functionals for Height Recovery

Height can be recovered from a gradient map by minimizing any of the functions R_{SD} , R_{θ} , or R_{Tay} (see Equations 4.5, 4.6, and 4.7). Not all integrators are created equal though, as was evident in practice. R_{SD} converged most rapidly and produced the best results; R_{Tay} converged more slowly but gave roughly the same results; R_{θ} was the slowest and most erratic to converge, and gave poor results.

Figure 5.7 shows a slice through the surfaces produced by the different integrators when applied to the gradient map for S_0 smoothed with a combination of Ω_2 and ρ_{SD} .

The reasons for the differences in performance are not entirely clear, but it is believed

that R_θ is inheriting some of the shortcomings of ρ_θ . $R_{T,\mathbf{v}}$ is thought to produce slow convergence because of the interaction of p and q in the penalty term. Note that for an integrable surface, any reasonable functional that pushes the surface in the right direction should work (indeed, a direct solution should be possible). It is only when integrability has not been enforced that the choice of the height recovery functional becomes important. Unless otherwise stated, all following experiments use R_{SD} .

5.4 Comparison of Smoothing Functionals

The primary goal of this thesis is to evaluate the performance of the different smoothing functionals Ω_1 , Ω_2 , and Ω_3 that were derived in the preceding chapters. Each of the smoothers was paired with the constraint metric ρ_{SD} and used to smooth the gradient map corresponding to the surfaces S_0 , S_{30} , and S_{60} . Height was recovered from the smoothed gradient map using R_{SD} as an integrator (see Equation 4.5). Plots of the surfaces recovered for the original, unrotated data set S_0 are shown in Figures 5.8, 5.9, and 5.10.

Cross-sectional plots of the surfaces reconstructed for each of the three data sets are shown in Figures 5.11, 5.12, and 5.13 for Ω_1 , Ω_2 , and Ω_3 respectively. Each figure shows a slice of the surface for the original data and for each rotated data set, where the reconstructions for the rotated data sets have been rotated back into the position of the original surface. The mathematical surface definition is also shown to indicate the reference surface.

The effect of staggered grids (see Section 4.1.2) is apparent near the origin, where it is clear that the reconstructed surfaces are displaced slightly to the right from the original surface. Note that the reconstructions have not been rescaled—the reconstruction process

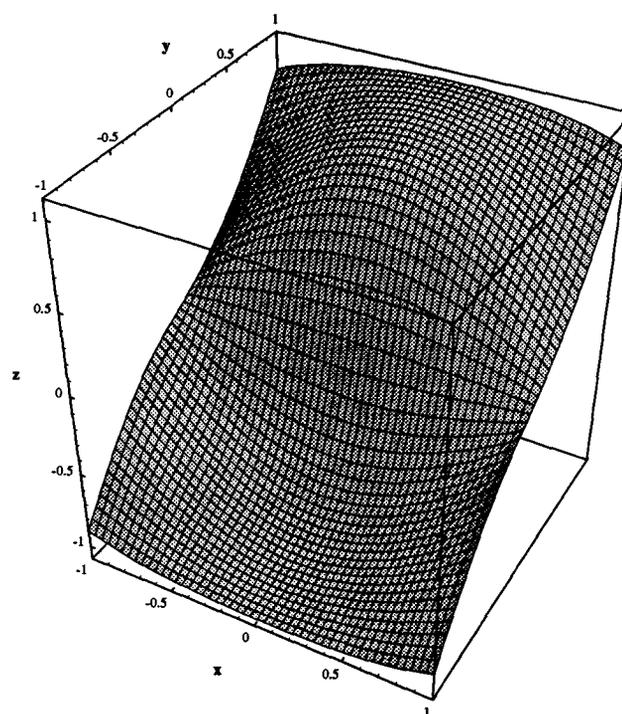


Figure 5.8: Surface reconstructed using Ω_1 .

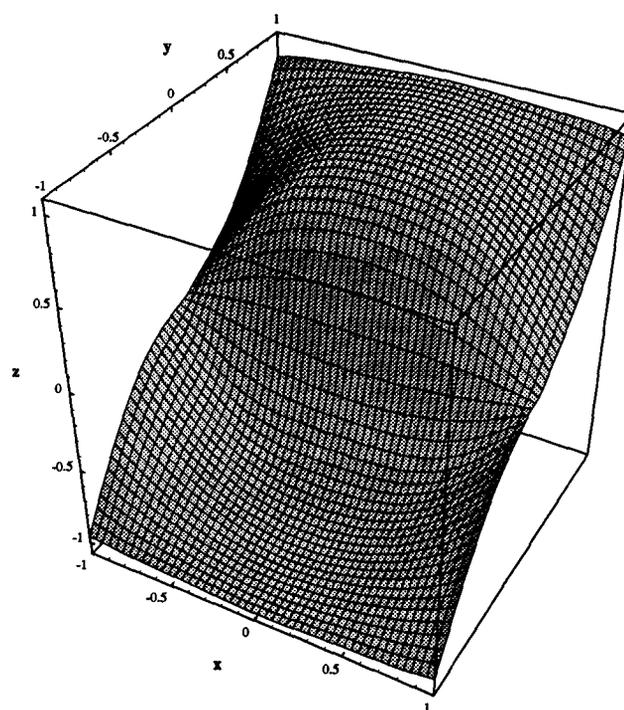


Figure 5.9: Surface reconstructed using Ω_2 .

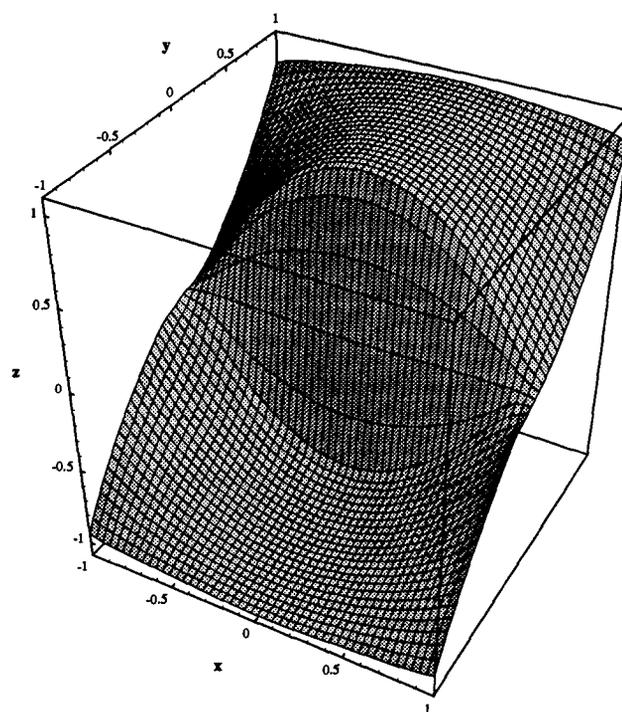


Figure 5.10: Surface reconstructed using Ω_3 .

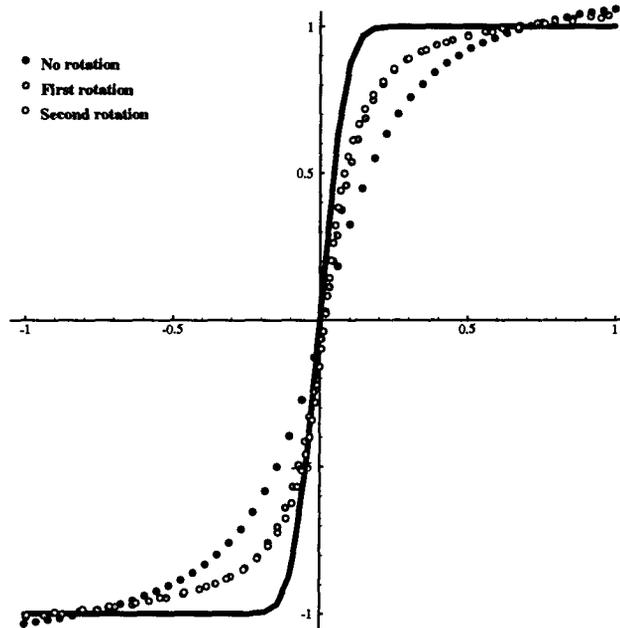


Figure 5.11: Surface slices reconstructed using Ω_1 . Each slice is through row 25 of the reconstructed surface.

accurately recovers the correct range of surface heights².

Ω_2 , the smoother based on first derivative parameterization, appears (at least qualitatively) to offer no significant benefits over the quadratic variation stabilizer Ω_1 . In both cases, the reconstructions for the two rotated data sets S_{30} and S_{60} line up very well, while the reconstruction for the original data set S_0 differs significantly. The trouble with the original data set is due to the high gradients encountered when y is close to 0. The surface is so steep at this point as to be almost discontinuous, which presents a problem for the smoother. This does not happen for the rotated surfaces, since the gradients remain much smaller.

For a quantitative comparison of the reconstructed surfaces, we can measure the

²The absolute position of the surface is also recovered correctly, which one would not generally expect from gradient data. This happens because of the symmetry of the surface.

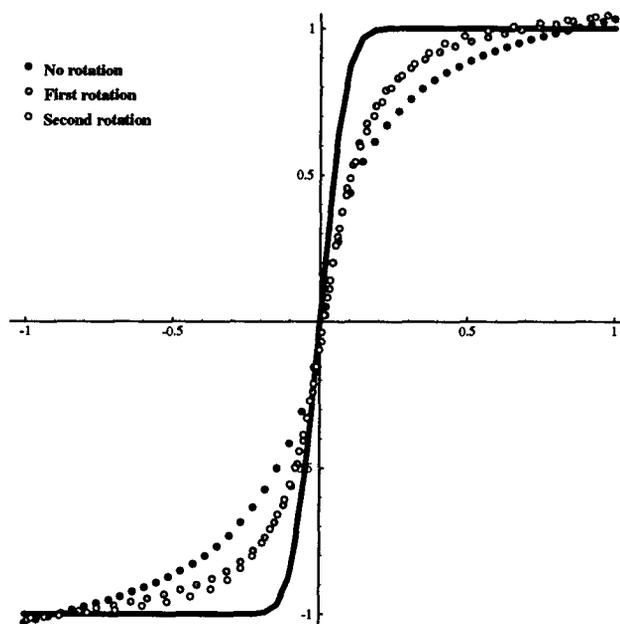


Figure 5.12: Surface slices reconstructed using Ω_2 . Each slice is through row 25 of the reconstructed surface.

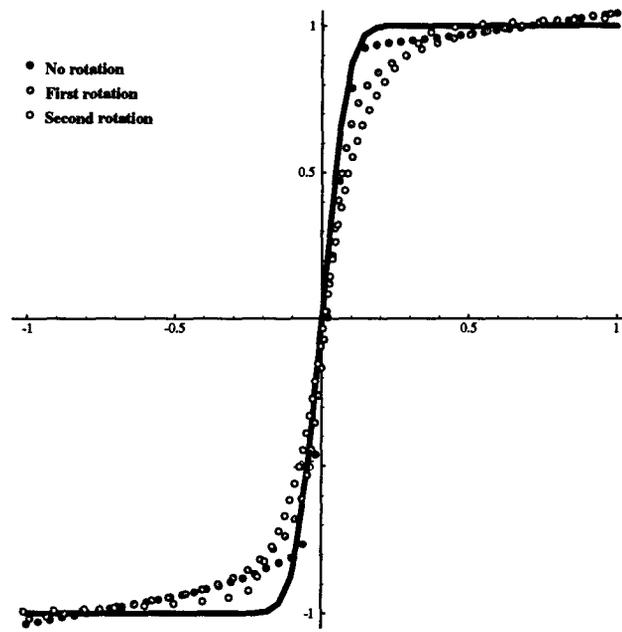


Figure 5.13: Surface slices reconstructed using Ω_3 . Each slice is through row 25 of the reconstructed surface.

	$S_{30} \rightarrow S_0$	$S_{30} \rightarrow S_{60}$
Ω_1	0.0542	0.0049
Ω_2	0.0355	0.0095
Ω_3	0.0302	0.0145

Table 5.1: Distance between recovered surfaces for different smoothing functionals.

“distance” between the reconstructions from different views. The distance metric used is the average vertical distance between the reconstructed surfaces, after rotation into the same position. S_{30} was rotated into the position of each of the other surfaces and this metric calculated. The results are shown in Table 5.1, which indicates that Ω_1 does best when the gradients are relatively small, while Ω_2 does perhaps marginally better overall. (This table provides a measure of reconstruction invariance, rather than closeness to original data.)

Ω_3 appears to be worse than both Ω_1 and Ω_2 . In retrospect, the method selected for implementation of Ω_3 was not a good choice. The parameterization is based on estimating the principal directions at each point on the surface. (Stevenson & Delp, 1992) report good results for this smoother in the context of range data, but their method is based on a more sophisticated method of estimating principal directions. In the current implementation, the only “new information” introduced to derive Ω_3 that was not used for Ω_2 was to numerically differentiate the gradients p and q to form estimates of the Weingarten mapping (see Section 3.3.4). This implies that when the estimates of p and q are poor, the parameterization of Ω_3 will be even worse. However, it was exactly when p and q were inaccurate that Ω_3 was meant to be an improvement over Ω_2 . Ω_3 was not considered in further tests, including those with real data.

5.5 λ Normalization

Selection of a regularizing parameter that gives the “best” results for the level of noise in the data is a difficult problem that has received much attention (Wahba, 1990). Setting $\lambda = 0$ will interpolate the data, while $\lambda = \infty$ will find the best-fit plane. Aside from reconstruction accuracy, an additional issue arises when the task is evaluation of different smoothing functionals.

In comparing the different smoothing functionals, it proved necessary to “normalize” the values used for the regularizing parameter λ . For a fixed value of λ , the tradeoff between smoothness and data fit will be different when different functionals are used. It is thus inaccurate to compare different smoothing functionals using the same regularizing parameter, as the results for one functional may be smoothed more than for another.

In order to achieve comparable degrees of smoothing with the different functionals, the notion of an *effective λ value* is introduced, denoted λ_0 . Given a desired value λ_0 , and smoothing and data fit terms Ω and ρ , the actual value of λ to be used is determined by scaling λ_0 by the “base response” of the smoothing and data fit terms to some representative gradient map data set $PQ_0 = \{(p, q)_{ij}\}$.

The base smoothing response is defined to be just the response of the smoothing term to PQ_0

$$\text{BSR} = \Omega(PQ_0) \quad (5.2)$$

To define a base data fit, the average gradient (\bar{p}, \bar{q}) is first calculated by averaging the surface normals present in PQ_0 . The base data fit is then defined as the sum of the response of the data fit term for each element of PQ_0 compared to the average gradient

$$\text{BDF} = \sum_{(p,q) \in PQ_0} \rho((p, q), (\bar{p}, \bar{q})) \quad (5.3)$$

The actual value of λ to use is then calculated as

$$\lambda = \lambda_0 \frac{\text{BDF}}{\text{BSR}} \quad (5.4)$$

Unless stated otherwise, the effective regularizing parameter used in these experiments is $\lambda_0 = 1.0$. The S_{30} data set was used for PQ_0 . This yielded actual λ values of 0.0403, 0.0810, and 0.0757 for Ω_1 , Ω_2 , and Ω_3 when paired with ρ_{SD} .

This is by no means the only way of solving this problem, but it seems to give reasonable results, as the surfaces in Figures 5.11–5.13 appear to be smoothed by comparable amounts. The problem of “normalizing” the regularizing parameter arises whenever different smoothing functionals are to be compared, but does not appear to have been addressed in the literature. (Stevenson & Delp, 1992), for example, state that $\lambda = 0.1$ was used throughout their experiments with different smoothing functionals.

5.6 The Data-Point Correspondence Problem

One problem with reconstruction on a regular grid, as proposed, is that the correspondences between data points and reconstructed surfaces are different in different views (see Figure 5.14). This means that while a smoothing functional such as Ω_2 may be viewpoint invariant on a point-by-point basis, it is not invariant to the reparameterization of the surface induced by sampling on a grid from a different direction. This is an important issue that has not been adequately addressed in this work.

In an attempt to demonstrate that this is indeed a problem, a new data fit term was derived that attempts to match corresponding points. The actual data fit calculation is identical to that of ρ_{SD} , but data points are matched against the closest point on the surface, rather than matching against the same grid point. The correspondences were determined by using the original data surface and the surface reconstructed using Ω_2

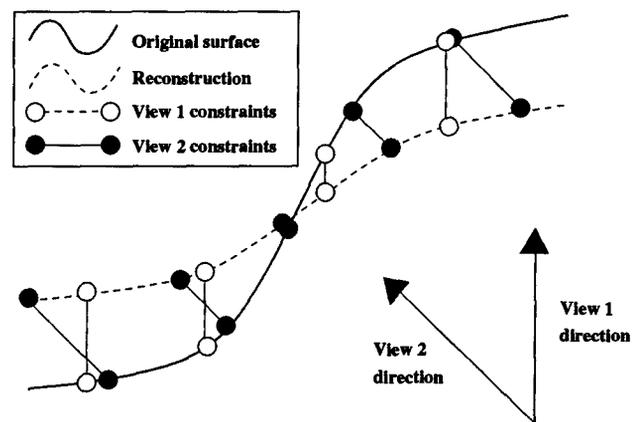


Figure 5.14: The data-point correspondence problem. Data points in different views are matched against different points on the postulated surface.

and ρ_{SD} to see which point on the surface was closest to each of the data points. These correspondences were then used throughout the reconstruction process³.

The results of the reconstruction (using Ω_2) for S_0 and S_{30} are shown in Figure 5.15. The results are, unfortunately, even worse than before, but the correspondence problem is still thought to be an issue that needs addressing. For reconstruction from range data, it is possible to derive correspondences by finding closest points, but the situation is more difficult for gradient data, since the absolute position of each data point is unknown. Under certain assumptions, such as that of convex surfaces, it might be possible to match directly at the level of p and q to determine the correspondences. This was not pursued.

5.7 Effects of Noise

To test the performance of these reconstruction methods in the presence of noise (before attempting to use real data), the gradient maps for the three surfaces were corrupted and the reconstruction re-applied. The addition of noise was achieved by treating each component of the gradient map as the tangent of an angle—the underlying angle was distorted by a random Gaussian with mean 0 and standard deviation 3° . In addition, each element of the gradient map was dropped from the data set with a probability 0.25. Figure 5.17 shows a slice through the p component of the gradient map for the original surface and the distorted p . Noise was added in this manner to each data set S_0 , S_{30} , and S_{60} independently to form S_0^n , S_{30}^n , and S_{60}^n .

Figure 5.16 shows the results of interpolating and integrating S_0^n without smoothing. Flaws in the surface are obvious.

³This is an artificial way of determining correspondences that could obviously not be used in practice (since the required surfaces are not available). It was done just to make a point.

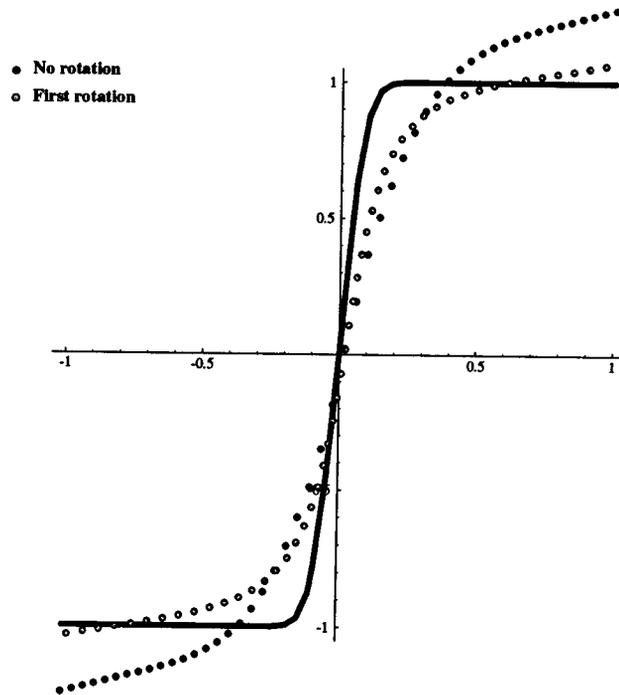


Figure 5.15: Reconstruction with corresponding point data fit term. Invariance is still not achieved.

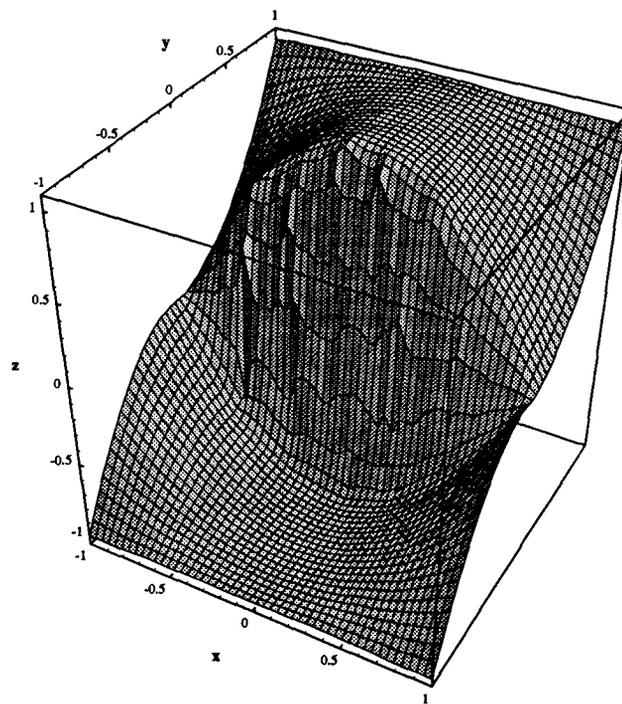


Figure 5.16: Surface reconstructed from noisy data without smoothing.

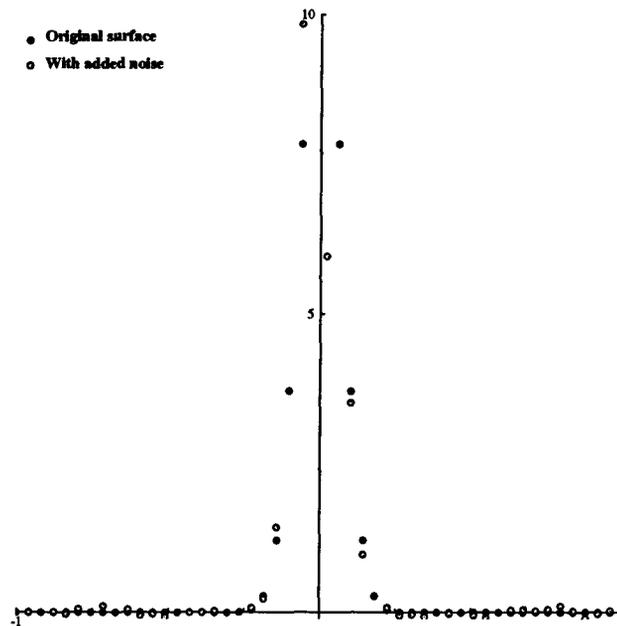


Figure 5.17: Noisy, sparse p component of gradient map compared to the uncorrupted gradient values.

	$S_{30} \rightarrow S_0$	$S_{30} \rightarrow S_{60}$
Ω_1	0.0583	0.0579
Ω_2	0.0554	0.0200

Table 5.2: Distance between recovered surfaces for different smoothing functionals applied to noisy data.

The results of the reconstruction are shown in Figure 5.18 and 5.19 for Ω_1 and Ω_2 , respectively. Comparison with Figures 5.11 and 5.12 indicates that this level of noise and sparsity is not a serious problem. It does appear that Ω_2 is better able to handle the noise, as the reconstructions for the rotated noisy data sets S_{30}^n , and S_{60}^n line up better than they do using Ω_1 (though the results for S_0^n still differs substantially).

The average pointwise distance between the reconstructed surfaces is again measured, and appears in Table 5.2. The quantitative results indicate that Ω_2 outperforms Ω_1 , though both struggle with the steep surface S_0 .

5.8 Effects of Varying Smoothing Parameters

While not central to the considerations of viewpoint invariance, it is instructive to examine the effects of varying the parameters that control the smoothing functional. These parameters are the regularizing parameter λ controlling the tradeoff between smoothness and fit to the data, and ν , a property of the material modelled by the variational formulation (see Section 3.2.3).

Figure 5.20 shows the effect of different degrees of smoothing on the surface S_{30} (values shown are for the *effective lambda*, as defined in Section 5.5). As expected, higher values of λ weight the smoothing term more heavily, causing the surface to be smoothed more. Higher values are appropriate when the input data is noisier. If the data is perfect, then

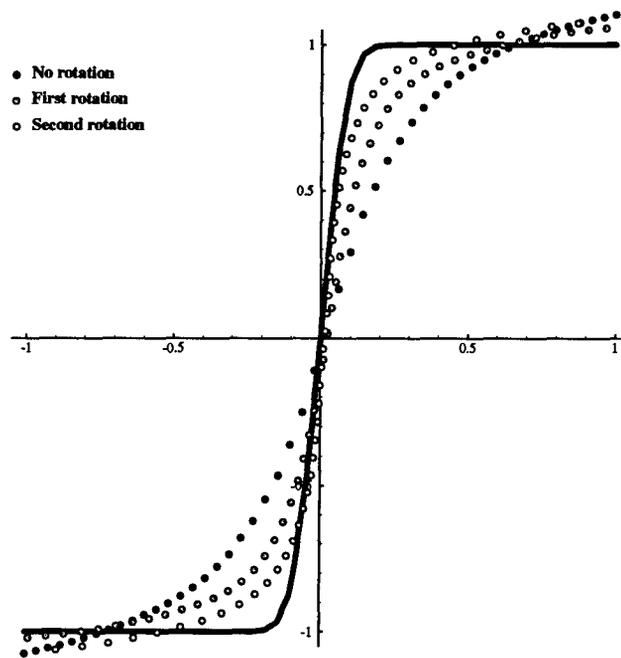


Figure 5.18: Reconstruction from noisy data with Ω_1 .

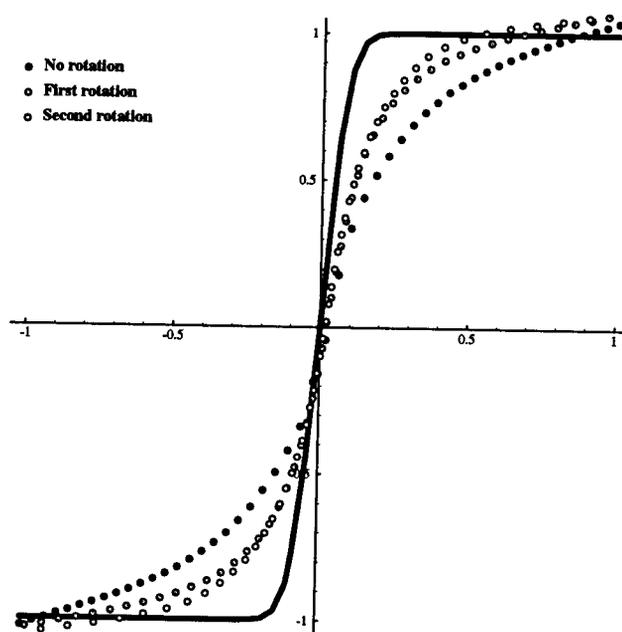


Figure 5.19: Reconstruction from noisy data with Ω_2 .

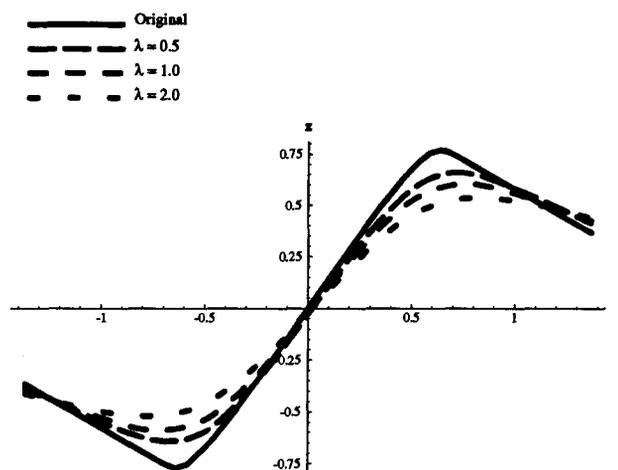


Figure 5.20: Reconstruction with different λ values. The values shown have been normalized, as described in the text.

λ can be set to 0, eliminating smoothing altogether.

Figure 5.21 shows the effects of different values of ν . Setting $\nu = 0.5$ is the theoretical volume-conserving limit if the formulation is to correspond to a physically realizable material. $\nu = 1$ was tried for the sake of comparison. As can be seen, variations in this parameter seem to have little effect on the reconstruction, as has been observed previously (Terzopoulos, 1988). $\nu = 0$ was used for all the preceding experiments.

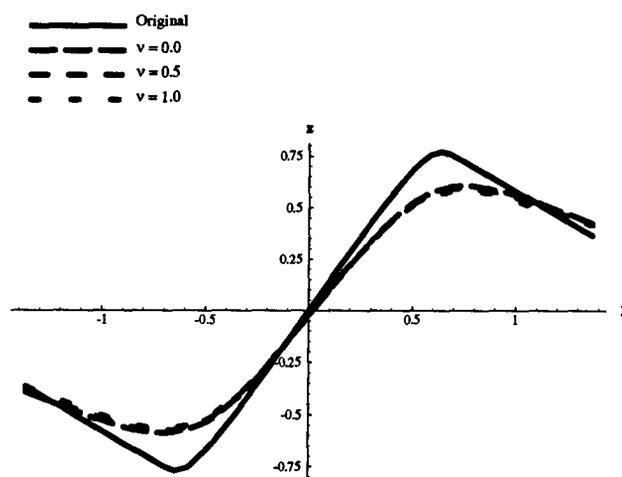


Figure 5.21: Reconstruction with different ν values. ν is a property of the material related to surface contraction per unit elongation.

6

Building Object Models from Real Data

The previous chapter has investigated various properties of the smoothing functionals using synthetically generated data. This chapter applies the reconstruction techniques to real data obtained with photometric stereo. Invariance of the results is explored in the context of fitting together surfaces from different views, a technique which can be used to build 3D object models.

Data collection is discussed in Section 6.1, including descriptions of the photometric stereo technique, the imaging setup used, and the calibration performed. Section 6.2 discusses surface reconstruction and model building for a simple ellipsoid object. Some samples of the surfaces that were recovered are shown in Section 6.2.1. Section 6.2.2 describes the surface alignment process used to build 3D models and shows the models recovered. Additional reconstruction and alignment experiments were performed with data collected for the more complex surface of a doll's face, as described in Section 6.3.

6.1 Data Collection

Image data was collected under controlled conditions and photometric stereo was used to recover the gradient at each point in the image. Section 6.1.1 describes the photometric stereo process. The imaging setup and its calibration are described in Section 6.1.2.

6.1.1 Photometric Stereo

This thesis investigates the use of gradient data for surface reconstruction. A robust and reliable source of gradient data is photometric stereo (Woodham, 1980). Various shape from shading techniques (Horn & Brooks, 1986) exploit the fact that the shading of a surface is related to its orientation by the image irradiance equation

$$E(x, y) = R(p, q) \quad (6.1)$$

which relates irradiance E at a point (x, y) in the image to surface orientation (p, q) at a corresponding point in the scene. The function R is the reflectance map, which is determined by the reflective properties of the surface, the illumination conditions, and the imaging geometry. R generally depends nonlinearly on the gradient.

Given measurements of the irradiance E at each point in the image, and knowledge of the reflectance map R , the image irradiance equation provides a constraint on the gradient. However, the equation is not invertible, since we have just one constraint on two unknowns. The basic idea of photometric stereo is to collect multiple images from the same viewing direction, but under different conditions of illumination, thus providing multiple constraints on the gradient and allowing direct local recovery of surface orientation.

Consider the case of three images gathered under three different illumination conditions. The constraints imposed on surface orientation can be written as

$$E_i(x, y) = R_i(p, q) \quad i = 1, 2, 3 \quad (6.2)$$

The implementation of photometric stereo can be thought of as a lookup table which matches brightness triples (E_1, E_2, E_3) to gradients (p, q) . The reflectance map functions R_i are not generally known *a priori*, but can be determined from a calibration object of known shape (e.g. a sphere).

For more details on the photometric stereo technique, see (Woodham, 1980) or Chapter 10 of (Horn, 1986).

6.1.2 Imaging Setup and Calibration

Image collection was done on a Newport calibrated optical bench. For the ellipsoid object, an ellipsoidal PVC¹ model was mounted on a rotational platform, allowing easy collection of multiple views of the object with known inter-view rotations. A PVC sphere was used for calibration. The imaging setup is shown in Figure 6.1. Data collection for the doll's head example was done similarly, though no ground truth was available.

A 3 CCD colour RGB camera was used in conjunction with three coloured light sources, allowing the RGB channels to be used for simultaneous gathering of the multiple images required for photometric stereo. Figure 6.2 shows the images obtained on the three colour channels for the first view of the ellipsoid used in the model building experiments.

The use of photometric stereo assumes that the projection can be modelled as orthographic (or, more precisely, weak perspective). This assumption is valid if the variation in range within the scene is small with respect to the distance to the camera. The projection equations for this simple model are

$$x = \frac{f_x X}{Z_0} \qquad y = \frac{f_y Y}{Z_0} \qquad (6.3)$$

where Z_0 is the average scene distance and f_x and f_y combine factors due to focal length, aspect ratio, etc. These equations map world points (X, Y, Z) to image points (x, y) . For the experiments performed here, the average scene distance was taken to be the distance from the image plane of the camera to the center of the rotational platform, $Z_0 = 86''$.

Calibration of the system was done to determine the parameters $h = \frac{Z_0}{f_x}$ and $v = \frac{Z_0}{f_y}$, which can be thought of as the height and width of a pixel in the units of the world

¹Poly-vinyl chloride.

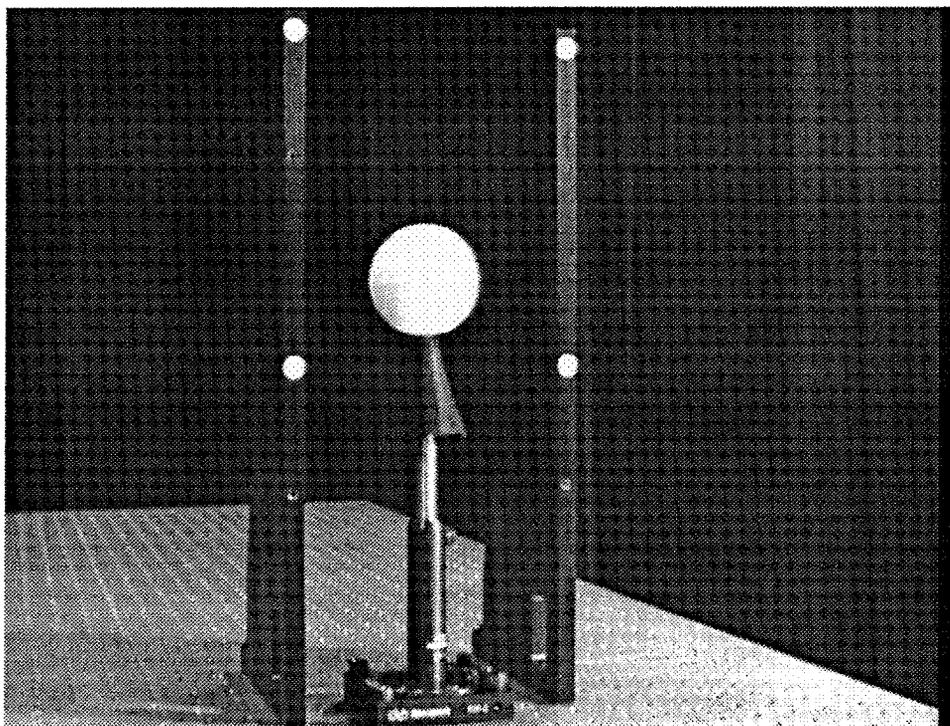


Figure 6.1: The imaging setup used for data collection. Objects are mounted on a rotating platform. The sphere was used for photometric stereo calibration. The four white dots were used to determine projection parameters.

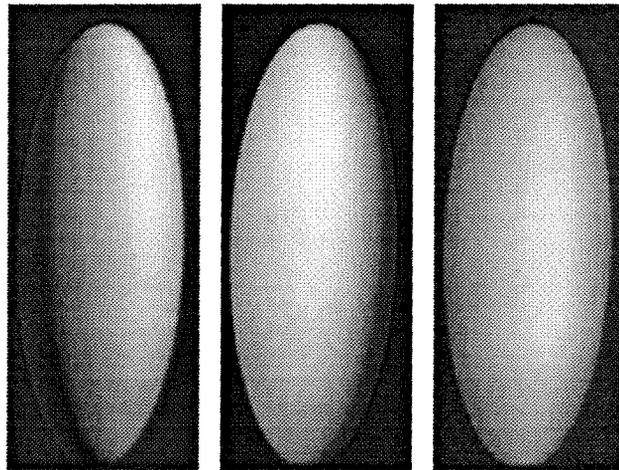


Figure 6.2: Three images of the ellipsoid gathered under different lighting conditions. Left is the red channel, center is the green, right is the blue.

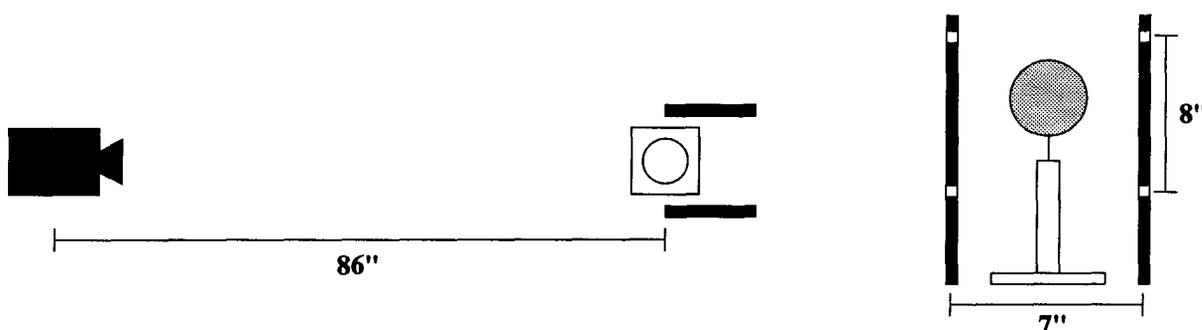


Figure 6.3: Optical bench calibration.

	h	v
$Z_0 - 1''$	0.02533	0.02012
Z_0	0.02563	0.02036
$Z_0 + 1''$	0.02587	0.02060

Table 6.3: Pixel dimensions at different distances.

coordinate system (see Section 4.1.1). Two vertical braces were mounted on either side of the rotational platform, with two circular white dots affixed to each (see Figure 6.3). The four white dots form a rectangle of known dimensions around the area of interest in the image. The centroids of the dots were calculated in a calibration image, giving the “pixel unit” dimensions of the rectangle, thus allowing easy determination of h and v . To ensure the robustness of this method, the calibration procedure was repeated with the side braces moved forward one inch, and again with them moved backward one inch. The results appear in Table 6.3. Differences between the three calibration positions agree well with that expected by a 1” variation at 86” ($\frac{1''}{86''} \times 100\% = 1.16\%$). These results were averaged to determine a camera aspect ratio of $\bar{v}/\bar{h} = 0.795$.

6.2 A Simple Ellipsoid

Experiments were performed with the data produced by photometric stereo for multiple views of the ellipsoidal model. The following subsections describe the reconstruction of surfaces for each view and the alignment of these surfaces to create a 3D model.

6.2.1 Surface Recovery

Once the photometric stereo method has been used to recover gradients from the raw image data, the surface reconstruction techniques described in previous chapters can be applied to recover the shape of the surfaces in view. This was done for multiple views of a PVC ellipsoid which were then meshed to form a full 3D model (Section 6.2.2). The ellipsoid was machined according to the mathematical specification

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 \quad (6.4)$$

where a , b , and c are the axes of the ellipsoid (see Figure 6.4).

Images of the ellipsoid were gathered for four different views, with a 90° rotation of the object between each view. The two components of the gradient map recovered for the first view of the ellipsoid are shown in Figures 6.5. The gradient map is recovered over a 280×130 image (which has been subsampled to plot the figures).

These data were integrated using R_{SD} to recover a surface with no smoothing. The result appears in Figure 6.6. The coherence of the surface indicates the regularity of the data obtained with photometric stereo. Smoothing was then applied, using Ω_1 and Ω_2 with ρ_{SD} , and height recovered from the smoothed gradient maps using R_{SD} . The effective λ value (see Section 5.5) in each case was $\lambda_0 = 1.0$. Results are shown in Figure 6.7 for Ω_1 and Figure 6.8 for Ω_2 . The fact that boundary conditions have not been applied is

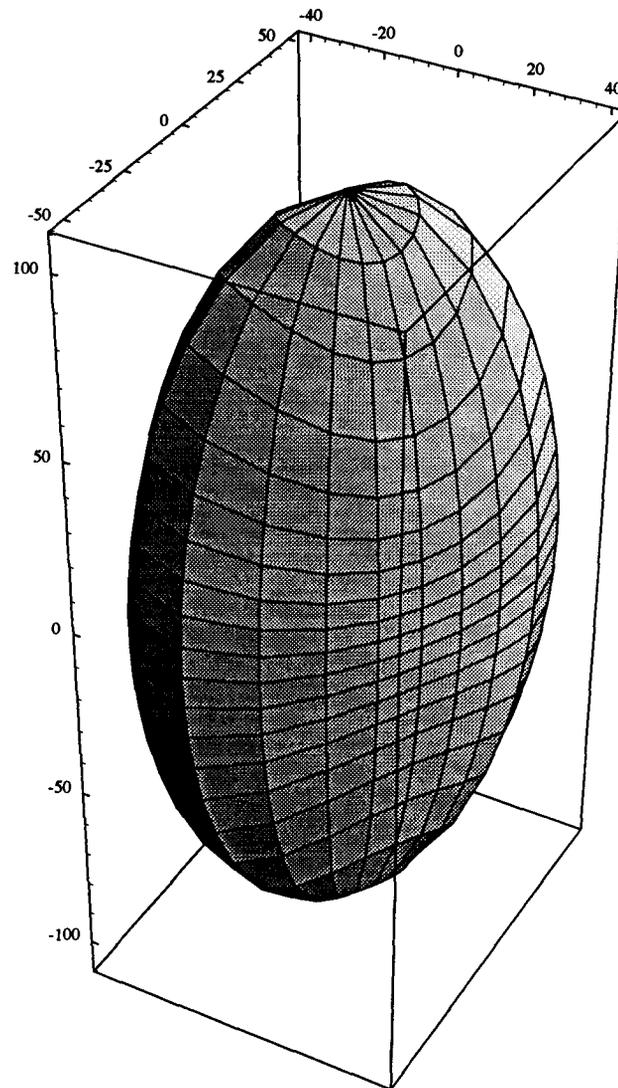


Figure 6.4: Mathematical specification of the ellipsoid. The three axes of the ellipsoid are 1.79", 2.99", and 5.39". The units shown in the figure are "pixel widths" (1 pixel = 0.02563").

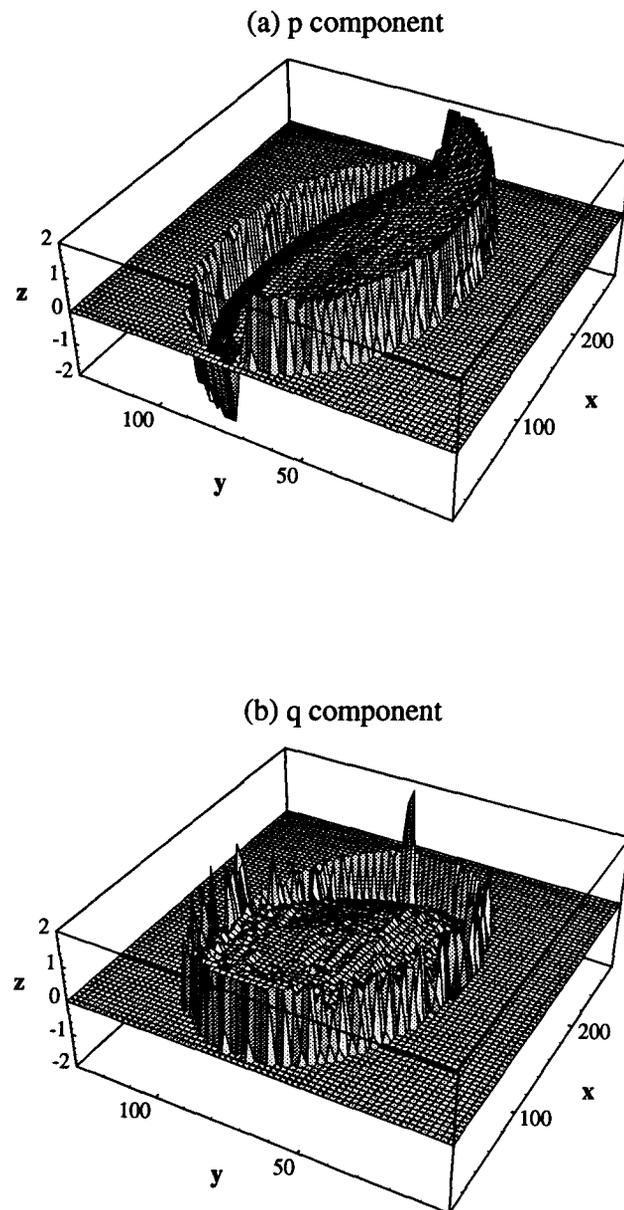


Figure 6.5: p and q components of the gradient map calculated from photometric stereo.

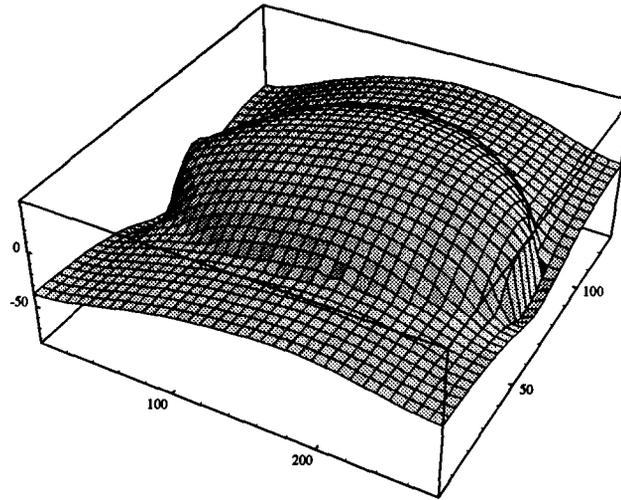


Figure 6.6: Surface height recovered without smoothing.

seen in the distortion of the surface at the edges of the ellipsoid². For the sake of easier comparison, slices through the 3D surfaces have been plotted in Figure 6.9. The surfaces obtained in each case are clearly very similar (an indication that perhaps smoothing is not really needed for this data). Figure 6.10 shows slices through the p component of the original gradient map and those obtained by smoothing with Ω_1 and Ω_2 .

6.2.2 Building Models

In order to test the invariance of the reconstruction methods, the surfaces reconstructed from multiple views of an object were meshed together to form a complete 3D model of the object. The more invariant the reconstruction technique, the better the different surfaces will fit together. Although the inter-view transformation is known, the use of gradient data means that the absolute distances of the recovered surfaces are unknown,

²The edge portions of these surfaces were trimmed with the aid of the silhouette image generated by the photometric stereo process before models were built.

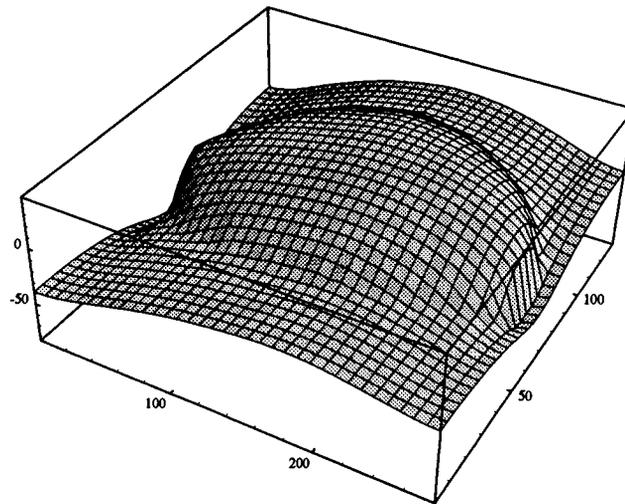


Figure 6.7: Surface height recovered with Ω_1 smoothing.

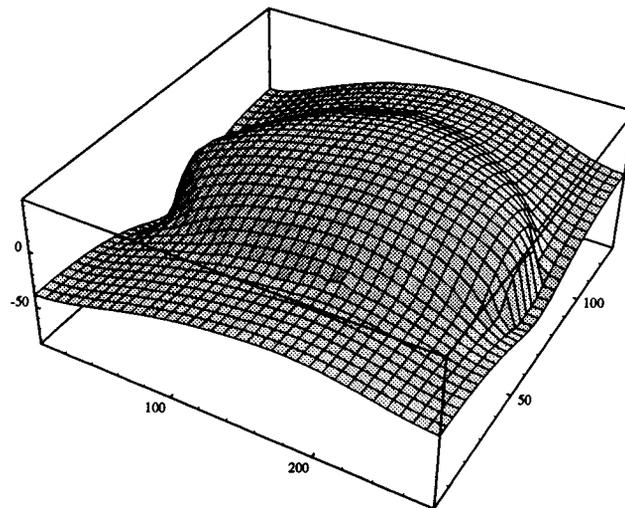


Figure 6.8: Surface height recovered with Ω_2 smoothing.

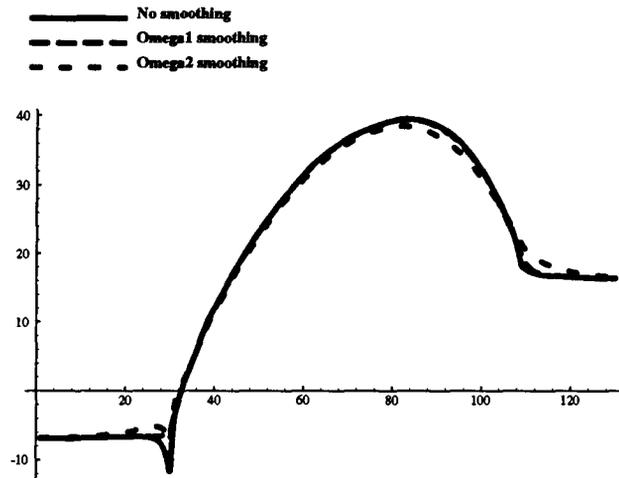


Figure 6.9: Slice through row 140 of surfaces reconstructed with no smoothing, Ω_1 smoothing, and Ω_2 smoothing. (Absolute surface height is arbitrarily determined by the reconstruction process.)

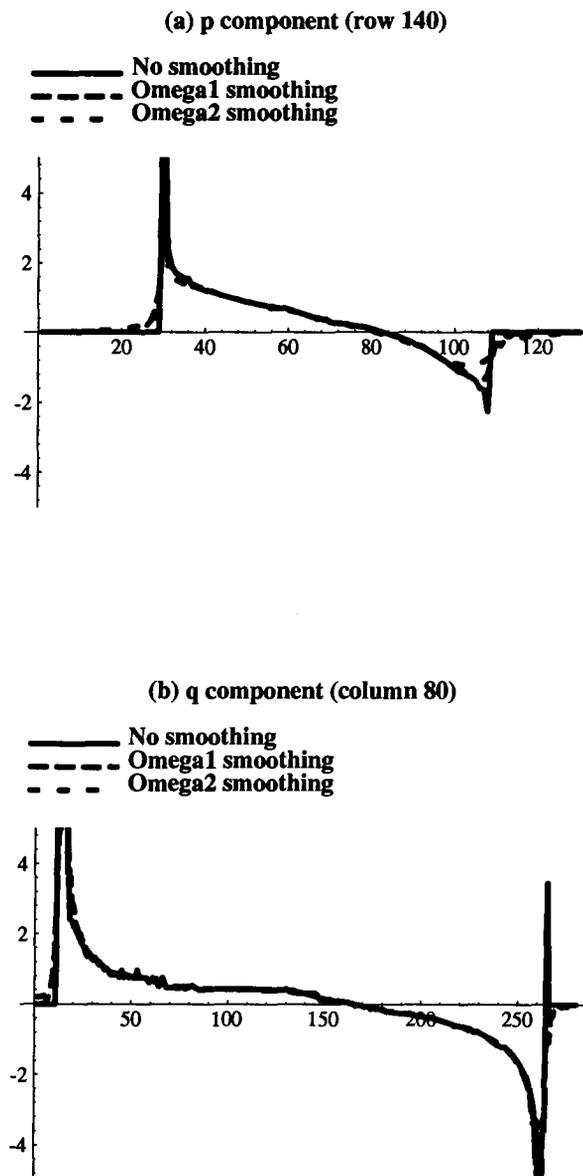


Figure 6.10: Slices through row 140 of p and column 80 of q component of smoothed gradient maps.

and so the alignment process must determine one degree of translational freedom per view.

Following sections describe the alignment problem and an algorithm to solve it, and present results from the model building process.

Registering Surfaces from Multiple Views

Multiple gradient views of an object have been obtained from known viewing directions and the object surface reconstructed in each view. A Cartesian coordinate system was used, with the positive z axis towards the viewer; the y axis vertical, increasing downwards; and the x axis horizontal, increasing to the right. If the intersection of the optical axis with the axis of rotation is taken as the origin, then the inter-view transformation is a known rotation about the y axis. In order to bring the multiple surfaces into alignment, the unknown absolute distance of each surface must be determined. If the axis of rotation is known, then there is one degree of translational freedom along the viewing axis associated with each surface (see Figure 1.1).

In practice, the origin was the upper left corner of the image, with the z position of the origin determined arbitrarily by the reconstruction algorithm. The y coordinates are the same in each view, leaving two degrees of freedom to be solved for to determine the x and z alignment of neighbouring surfaces.

Given an $M \times N$ surface $z(x, y)$ that has been recovered for a particular view, it is simple to transform the surface into a 3D point set in world coordinates by using the pixel dimensions determined during camera calibration

$$S = \{(X, Y, Z) = (hx, vy, z(x, y)) | 1 \leq x \leq M, 1 \leq y \leq N\} \quad (6.5)$$

where h and v are the height and width of a pixel in the units of the world coordinate system (determined during calibration). Given the point sets S and S' for two different

surfaces, it still remains to apply the rotation and determine the absolute depth of each surface.

If points $(x, y, z) \in S$ and $(x', y', z') \in S'$ correspond to the same physical point on the object, then the transformation between the two surfaces is

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z + \Delta z_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \Delta z_2 \end{bmatrix} \quad (6.6)$$

where Δz_1 and Δz_2 represent the unknown translation along the line of sight for the two surfaces and R is the rotation matrix for the known rotation about the y axis by angle θ

$$R = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (6.7)$$

Alternatively, and equivalently, Equation 6.6 can be expressed as

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ 0 \\ \Delta z \end{bmatrix} \quad (6.8)$$

indicating that the rotation can first be applied and then two translational parameters determined. The equivalence of the two formulations is seen from

$$\Delta x = \Delta z_1 \sin \theta \quad \Delta z = \Delta z_1 \cos \theta + \Delta z_2 \quad (6.9)$$

Formulating the problem in this way also obviates the need to know the position of the (vertical) axis of rotation and justifies the use of the upper left corner of the image as the origin.

Given two corresponding points, it is a trivial matter to determine the translation that aligns them. Unfortunately, the correspondences are unknown. This problem can

be overcome by using an iterative closest point algorithm similar to that of (Besl & McKay, 1992)³. Given two point sets A and B related by an unknown translation, the point correspondences are approximated by finding the closest point in B to each point in A . A translation vector is then calculated by averaging the difference vectors of each of these point matches. This translation is then applied to A and the process is iteratively repeated until A is “sufficiently close” to B . The success of this process hinges on the ability to use closest point matches in place of the (unknown) point-to-point correspondences.

In the context of surfaces recovered from gradient data, each point in the point set obtained by Equation 6.5 can be tagged with the gradient at that point. Given point sets from two adjacent, overlapping surfaces, the gradients at each point can be used to determine the mutually visible points—i.e. the area of overlap. Having determined the points in the overlapping regions, the iterative closest point algorithm can be applied to register the point sets.

Registering surfaces in this way results in a dense cloud of points that have all been brought into a common coordinate system. In order to provide a more succinct representation of the object model, the point cloud was aligned with the 3D origin and converted to a spherical coordinate representation (r, θ, ϕ) . The model is generated as an array $R(\theta, \phi)$ which stores an r value for each (θ, ϕ) . When multiple points map to the same (θ, ϕ) , their r values are averaged. This representation is suitable as long as the objects to be modelled are star-shaped (i.e. every ray emanating from the origin intersects the surface of the object exactly once).

The model building process may now be summarized as follows:

1. Reconstruct surfaces from multiple views.

³Besl and McKay solve for rotational as well as translational alignment.

2. Convert surfaces into point set representation and apply known inter-view rotation.
3. Determine overlapping region of adjacent surfaces by applying mutual visibility calculation to gradients.
4. Use the overlapping regions to register the point sets with the iterative closest point algorithm.
5. Merge the registered point sets into one point cloud representing the entire 3D surface of the object.
6. Align the object point cloud with the origin and convert to to the more compact spherical coordinate representation.

Results for Full Surfaces

The model building process described in the previous section was applied to the data gathered for four views of the ellipsoid separated by 90° ⁴. Surfaces were reconstructed for each view using (a) no smoothing, (b) smoothing with Ω_1 , and (c) smoothing with Ω_2 . The constraint metric used was ρ_{SD} in each of the smoothed cases. Height was recovered from gradient using R_{SD} .

The model recovered with no smoothing is shown in Figure 6.11. The (θ, ϕ) space of the spherical coordinate system has been sampled on a 36×36 grid. The models recovered after surface reconstruction with Ω_1 smoothing and Ω_2 smoothing are shown in Figures 6.12 and 6.13. As would be expected from the similarity of the surfaces recovered in each view, the models are also similar.

While the recovered models appear qualitatively similar to the ground truth shape of the object (Figure 6.4), there are a number of problems. The surfaces reconstructed from

⁴Nothing in the method described requires 90° rotations. The only requirement is that there be sufficient overlap between views to solve the registration problem.

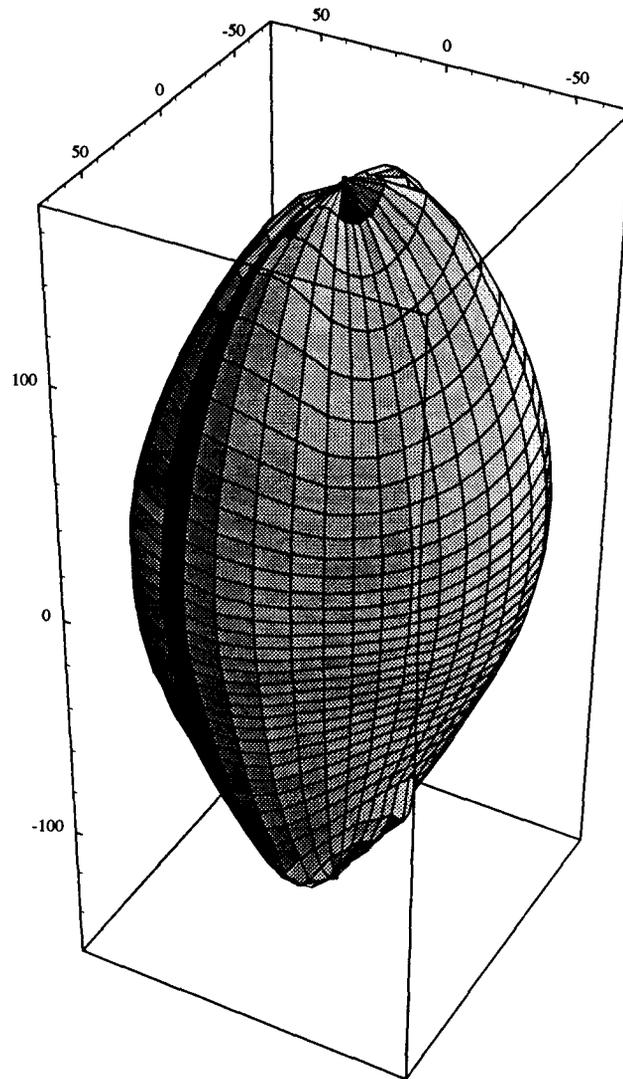


Figure 6.11: Model built with no smoothing.

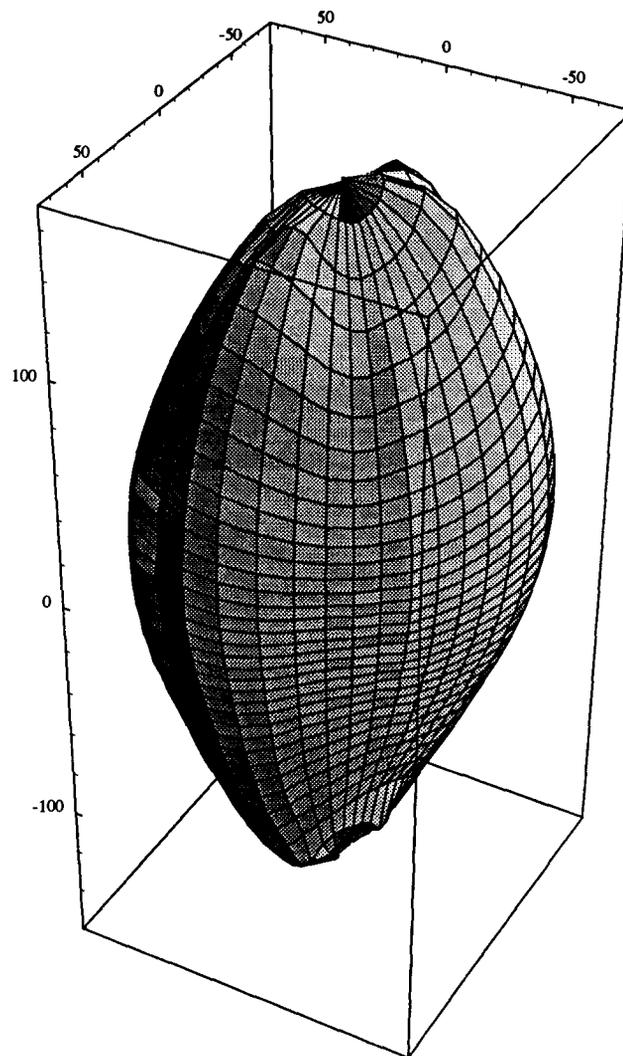


Figure 6.12: Model built using Ω_1 smoothing.

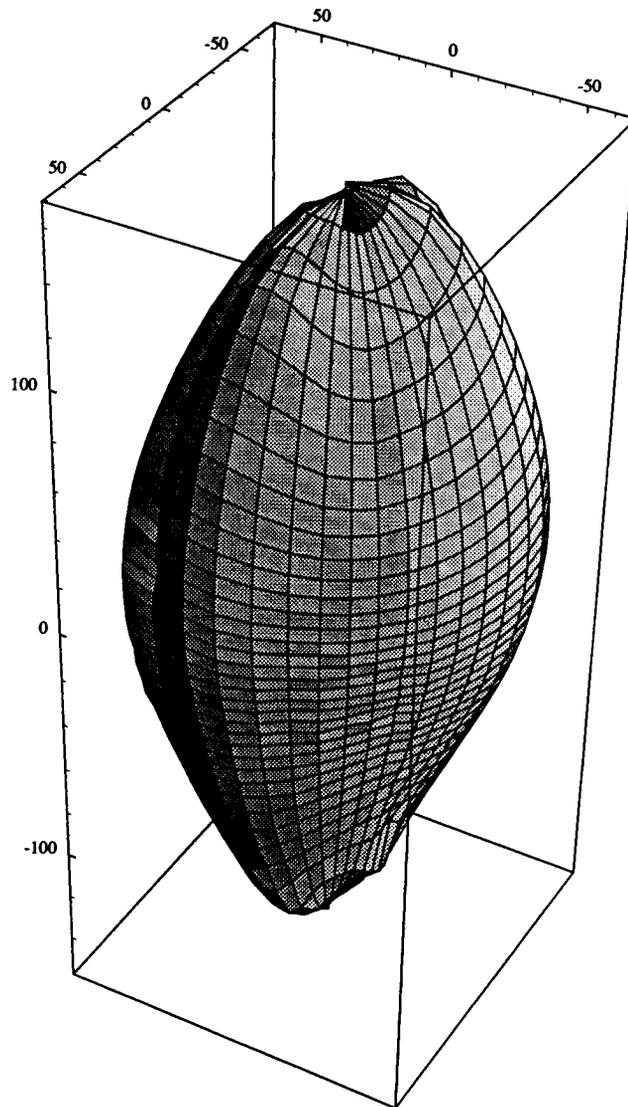


Figure 6.13: Model built using Ω_2 smoothing.

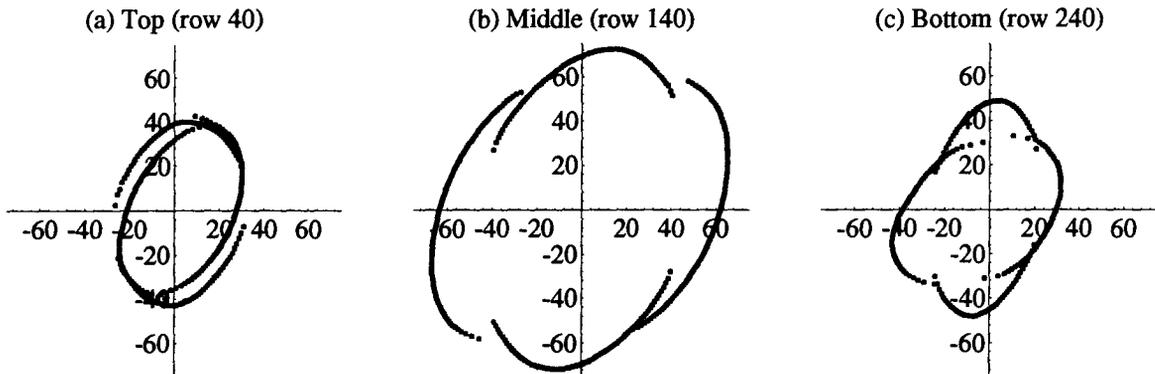


Figure 6.14: Cross-section of registered surfaces with no smoothing.

each view do not fit together particularly well, and the apparent coherence of the models is the result of extensive averaging—each r value in the spherical coordinate representation was calculated from as many as several hundred 3D points from the reconstructed surfaces.

A clearer picture of the results of the the alignment is obtained by taking slices through the registered surfaces, as is done in Figures 6.14, 6.15, and 6.16. As can be seen, the alignment between adjacent surfaces is poor, and correcting the alignment within one cross-section would make it worse in another. Figure 6.17 shows the corresponding slices through the mathematical definition of the ellipsoid, providing a ground truth reference.

Results for Trimmed Surfaces

The lack of boundary conditions at the occluding boundaries of the object was thought to be a reason for the poor alignment of the surfaces. Since smoothing is imposed across the boundaries, the sharp edges will be smoothed away and bends introduced into the surface.

In order to test this hypothesis, an experiment was conducted in which the original

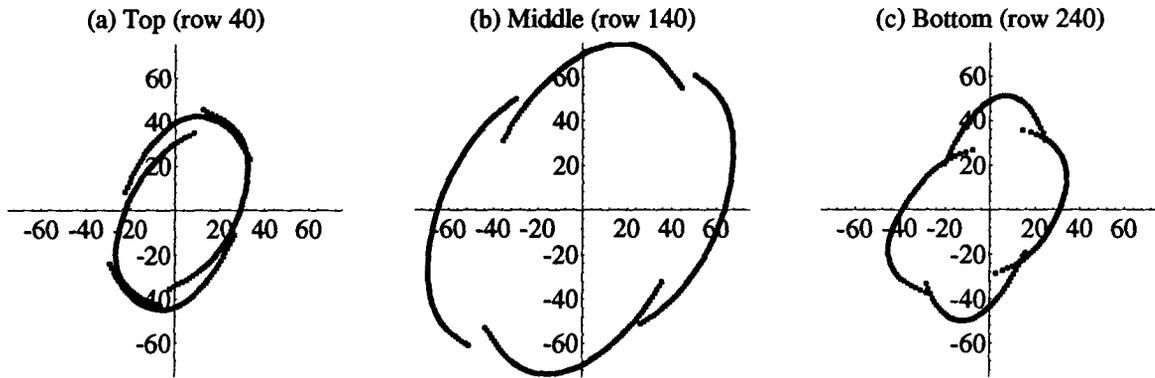


Figure 6.15: Cross-section of registered surfaces with Ω_1 smoothing.

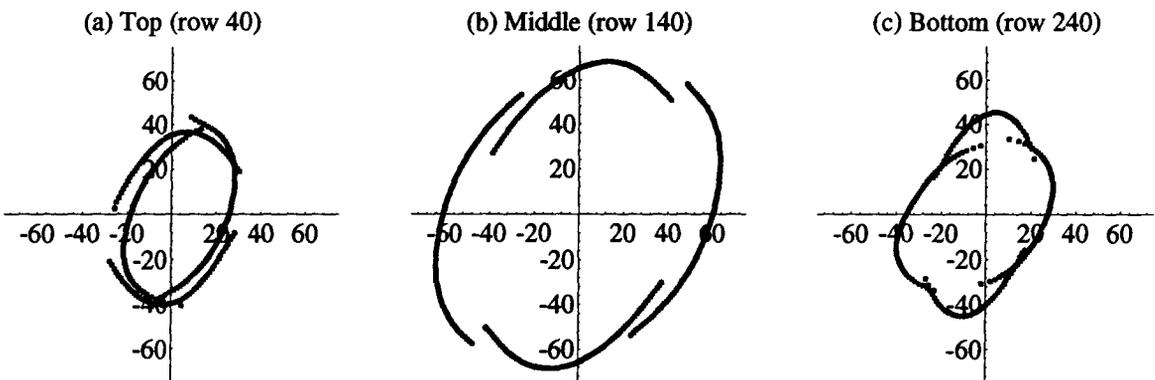


Figure 6.16: Cross-section of registered surfaces with Ω_2 smoothing.

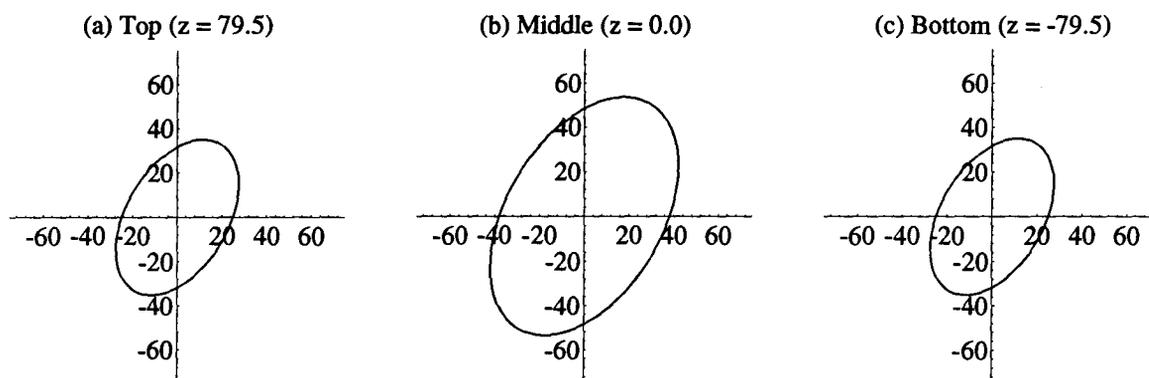


Figure 6.17: Cross-section of ground truth definition of ellipsoid.

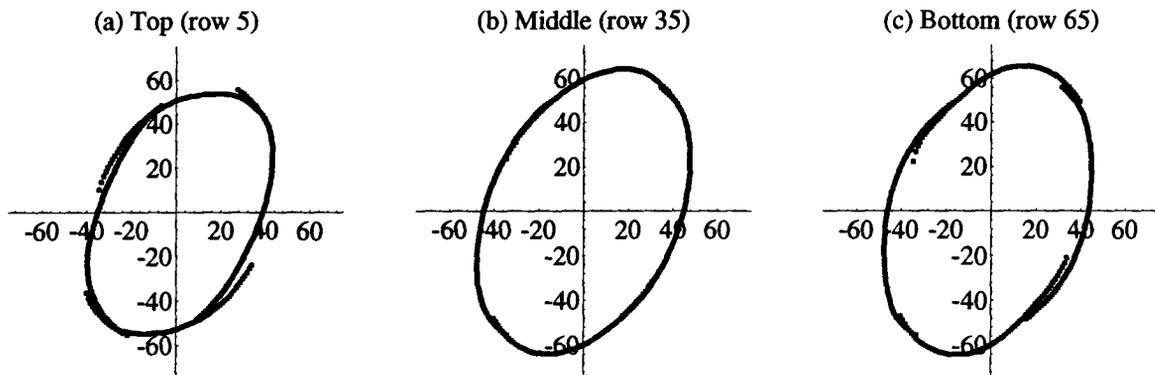


Figure 6.18: Cross-section of registered surfaces for trimmed data with no smoothing.

data for each view was trimmed to produce a sub-image with the gradient known at every point, so that no object boundaries were present in the image. The smoothing and reconstruction algorithms impose natural boundary conditions at the edge of the image, which could be expected to produce better results than the case in which object boundaries were smoothed over.

A more sophisticated alternative would be to incorporate natural boundary conditions along irregular boundaries within the image. Such boundaries could be obtained either from a silhouette image or by thresholding the gradient map (so as to avoid the extreme gradient values encountered at the occluding boundaries of an object). Handling of such boundaries and discontinuities was an issue consciously avoided from the outset. Trimming the data so that no object boundaries fall within the image was a pragmatic choice that allowed a hypothesis to be tested within the framework of the algorithms that had already been developed.

The results are indeed much improved, as is seen in Figures 6.18, 6.19, and 6.20, which show slices through the aligned surfaces for reconstruction with the trimmed data. The surfaces align well, even near the edges of the data.

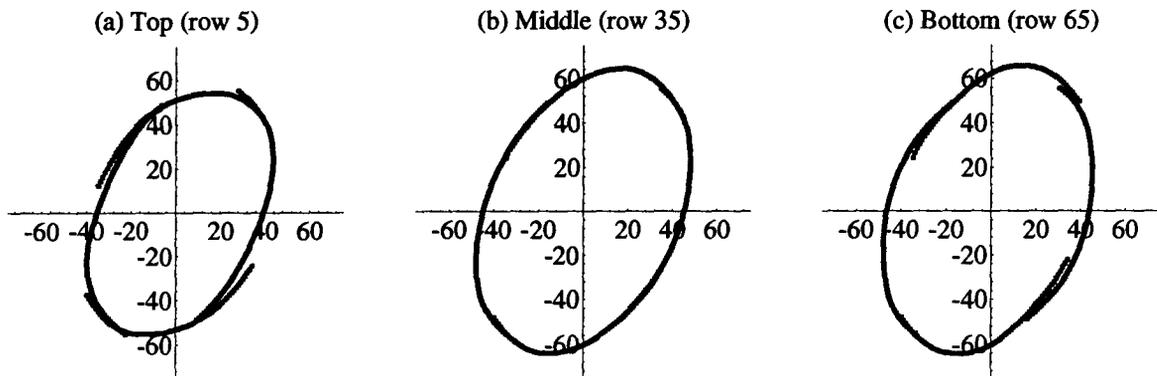


Figure 6.19: Cross-section of registered surfaces for trimmed data with Ω_1 smoothing.

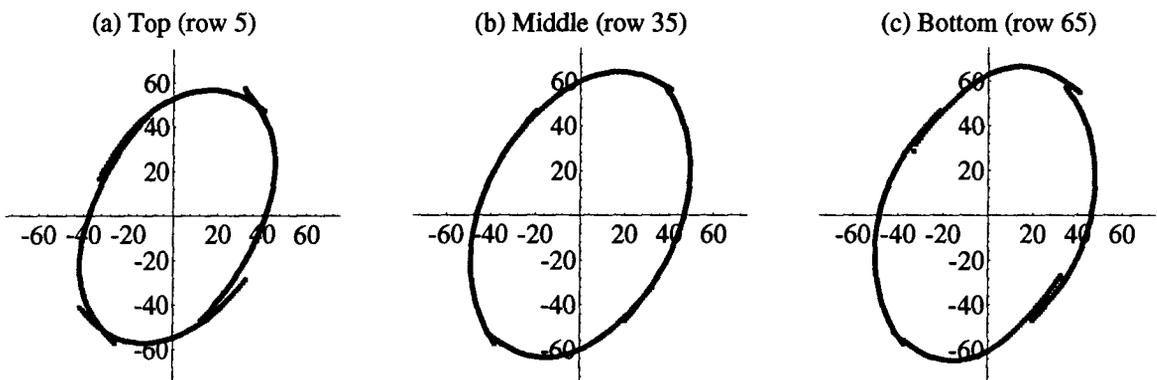


Figure 6.20: Cross-section of registered surfaces for trimmed data with Ω_2 smoothing.

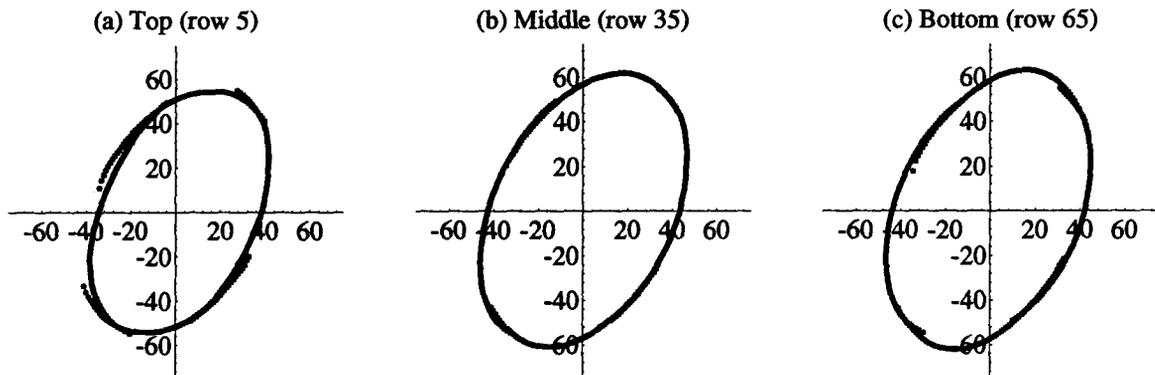


Figure 6.21: Cross-section of registered trimmed surfaces with no smoothing.

However, inadequate boundary conditions do not seem to be the root of the problem. Another experiment was performed in which the reconstructed surfaces, rather than the original data, were trimmed. Trimming was performed on the surfaces recovered from the full original data, with the regions extracted exactly corresponding to the trimmed regions in the previous experiment. These trimmed portions of the surfaces were then aligned. Results are shown in Figures 6.21, 6.22, and 6.23, and are seen to be as good as those obtained by trimming the data and reconstruction over a region interior to the object. This result would not be expected if errors induced at the object boundaries were propagating over the surface and making alignment impossible. The difficulties in aligning the surfaces thus seem due to the accumulation of errors across the reconstructed surface.

In each of these cases, the differences between results obtained with no smoothing and those obtained with Ω_1 or Ω_2 smoothing are minimal.

Viewpoint invariance seems to be possible when the extent of the regions to be aligned is not too great, as otherwise errors accumulate across the surface. Unlike range data, slope data do not provide any fixed 3D points to which the surface can be “anchored,”

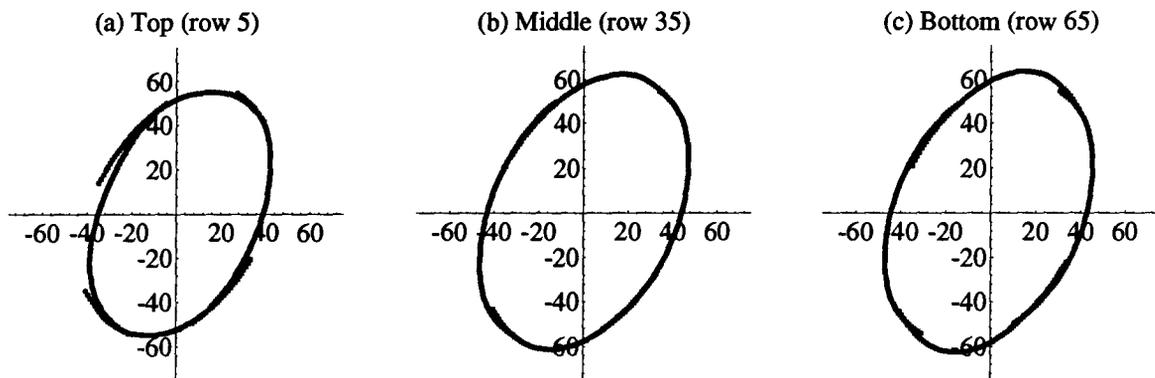


Figure 6.22: Cross-section of registered trimmed surfaces with Ω_1 smoothing.

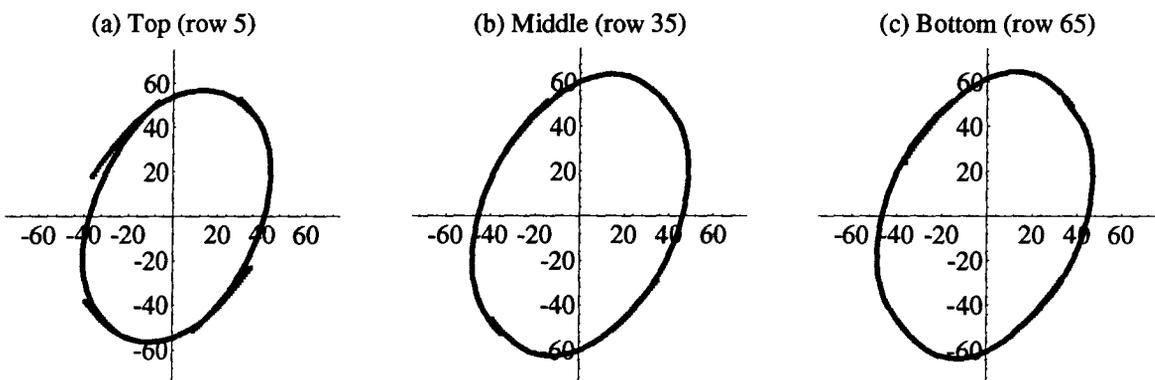


Figure 6.23: Cross-section of registered trimmed surfaces with Ω_2 smoothing.

meaning that errors can propagate unchecked. This seems to have more effect on the ability to align surfaces recovered from multiple views than does the choice of smoothing functional⁵.

An additional difficulty for the alignment algorithm is the nature of the example that was chosen. For a smooth object like the ellipsoid, cross-sectional segments that are being aligned are smooth, relatively featureless curves. This makes it difficult to perform accurate alignment, and it is not surprising that the cross-sections are “spread,” since there are no features to hold them together. This may be the reason the reconstructed models invariably have a wider cross section than the ground truth model (compare the various alignment results with Figure 6.17). The following section presents results for a more complex surface which provides features that aid the alignment process.

6.3 Facial Reconstruction

Experiments were performed using data gathered from a pottery doll’s head. Photometric stereo was used to gather gradient data for two views of the head separated by a 30° rotation about the vertical axis. Figure 6.24 shows the red channel for the two views of the doll’s head. The region actually used for reconstruction is delimited by a rectangle overlaid on each view. The raw p and q data for the first view are shown in Figure 6.25. (Note that this data set mostly avoids the extreme values of p and q found at occluding boundaries of objects, which were present in the ellipsoid data.)

Surfaces were reconstructed for the two views and aligned using the iterative closest point algorithm described above in Section 6.2.2⁶.

The reconstruction and alignment process were performed using no smoothing, Ω_1

⁵Though this might not be true of data obtained from a noisier source than photometric stereo.

⁶Because of the way the data has been trimmed, the mutual visibility condition cannot be used to determine the region of overlap, and so it was delimited manually.

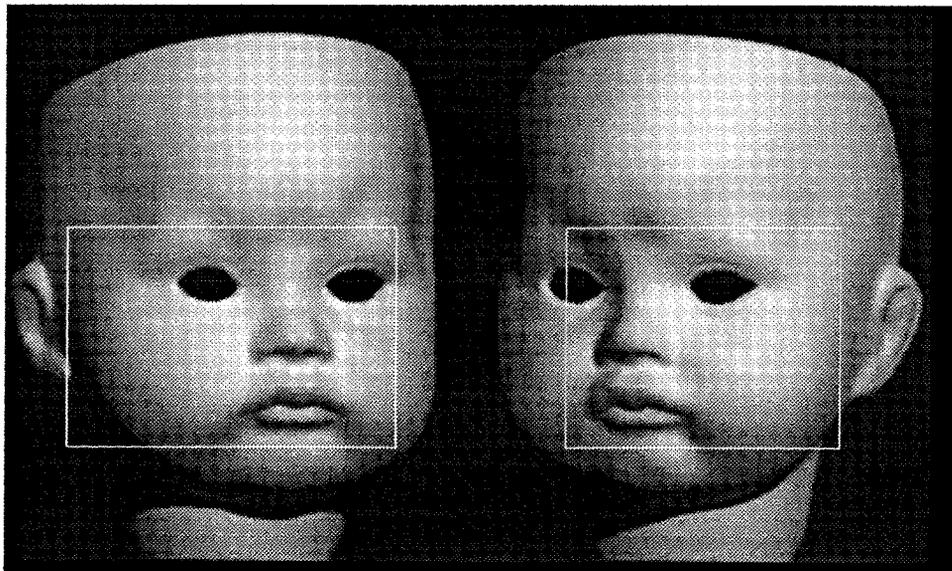


Figure 6.24: Images of doll's face showing regions from which data was selected.

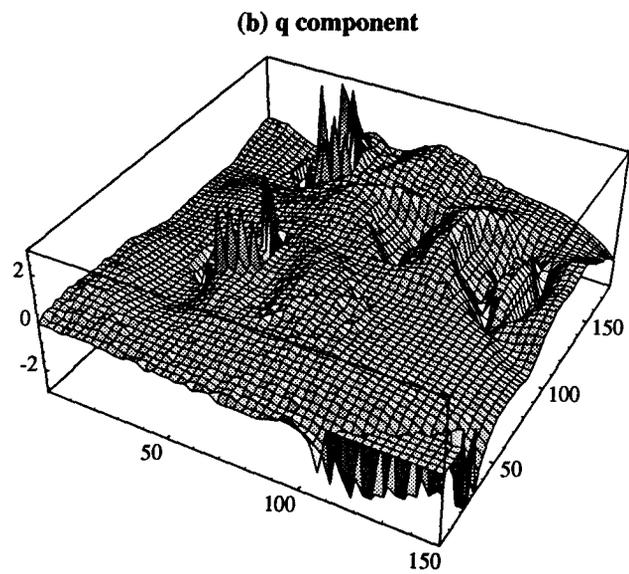
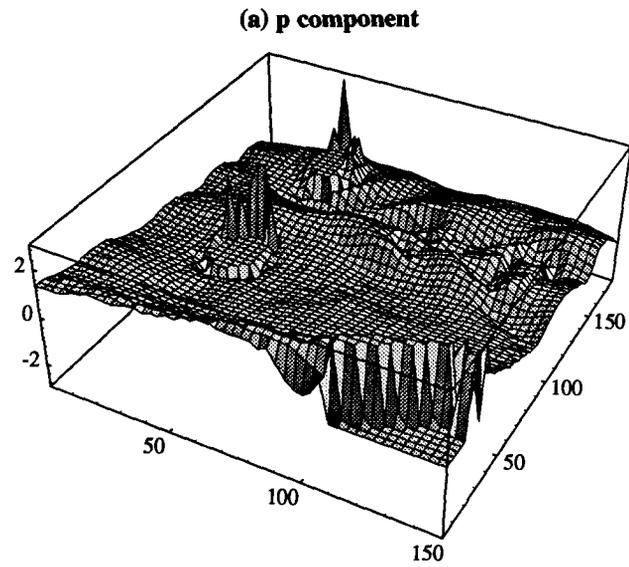


Figure 6.25: Gradient data for the first view of the doll's head.

smoothing, and Ω_2 smoothing. The surfaces recovered for the first view are shown in Figures 6.26, 6.27, and 6.28. The reconstruction is similar in the three cases, though the effects of smoothing are obvious, particularly in the Ω_2 case (this may be due in part to inadequacies in the λ normalization scheme).

The viewpoint invariance of the reconstructions can be seen in Figures 6.29, 6.30, and 6.31, which show quite successful alignment. Models were not constructed because the spherical coordinate representation does not lend itself to the non-convex shape of the doll's face.

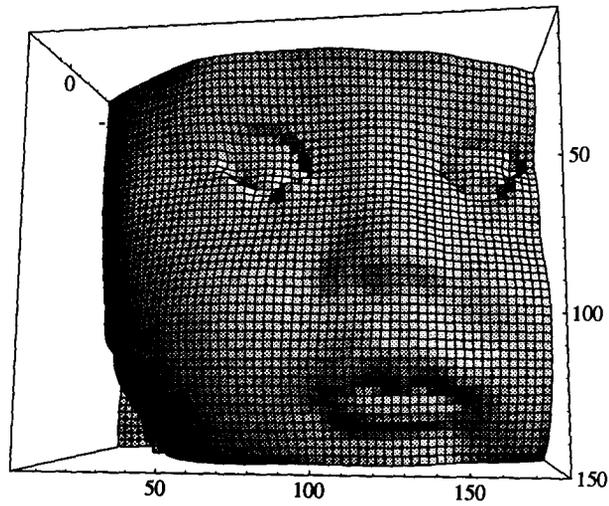
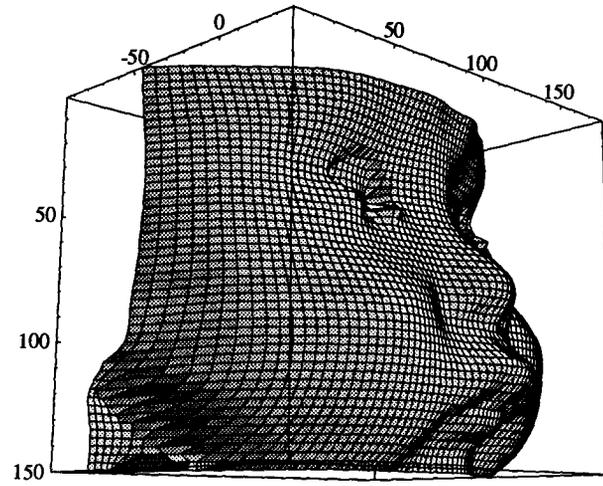


Figure 6.26: Two angles on the face reconstructed from the first view without smoothing.

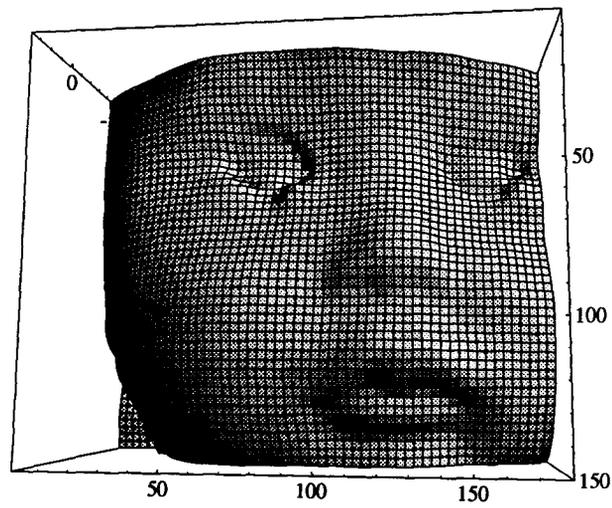
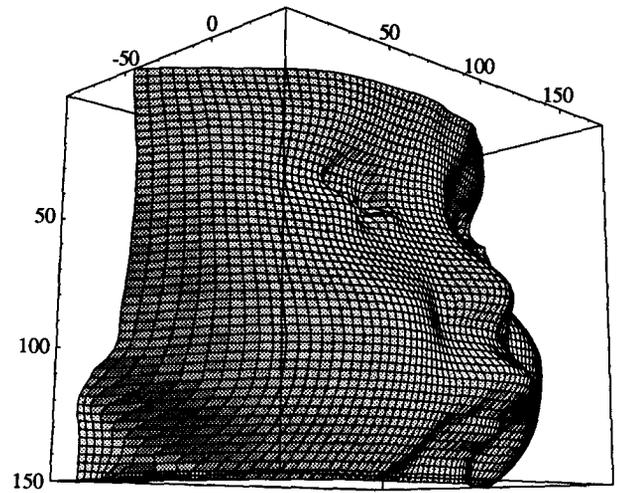


Figure 6.27: Two angles on the face reconstructed from the first view with Ω_1 smoothing.

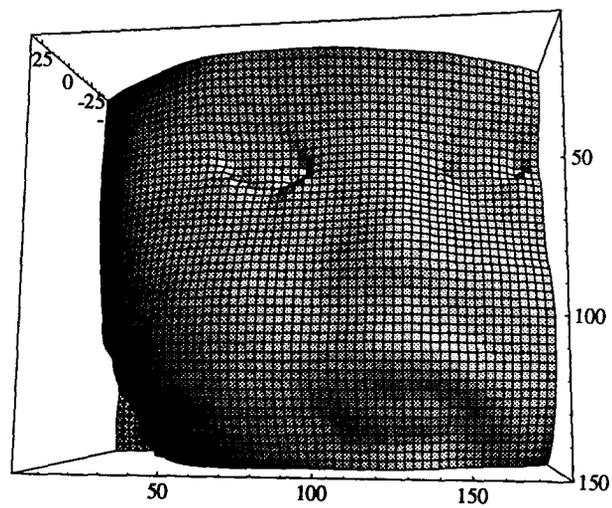
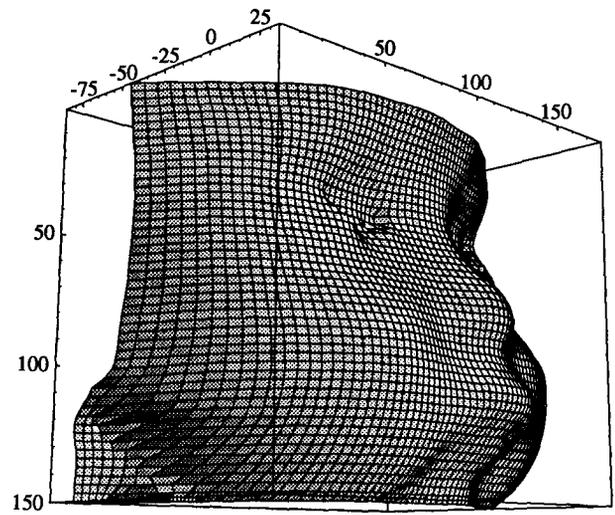


Figure 6.28: Two angles on the face reconstructed from the first view with Ω_2 smoothing.

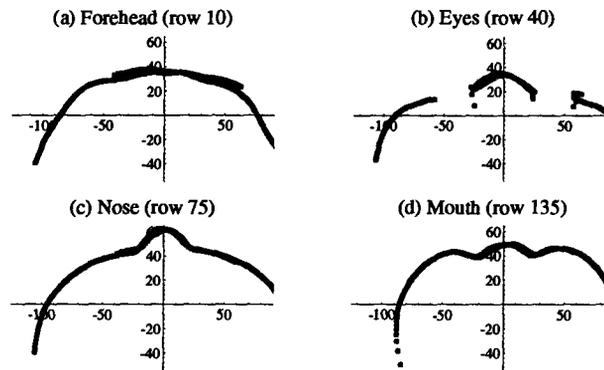


Figure 6.29: Face surface alignment without smoothing. Slices are plotted through representative features of the face.

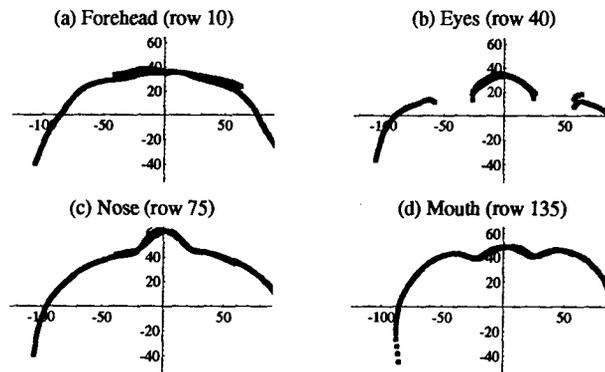


Figure 6.30: Face surface alignment with Ω_1 smoothing.

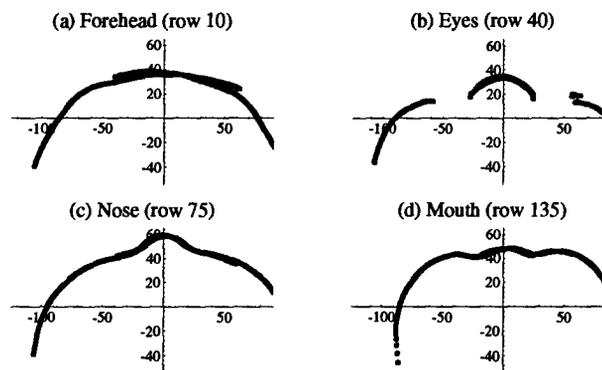


Figure 6.31: Face surface alignment with Ω_2 smoothing.

Conclusions and Future Work

Viewpoint invariant surface reconstruction from gradient data has been addressed in a variational framework. The property of invariance is important for the merging of information from multiple views, such as for automatic acquisition of 3D object models, for recognition of objects in arbitrary poses, and for accurate surface reconstruction and measurement generally. An invariant regularizing functional based on the energy of a deforming thin plate was presented. Three quadratic approximations to the smoothing component of the functional were derived— Ω_1 is a small deflection approximation; Ω_2 is an approximation based on first derivative estimates; and Ω_3 is based on estimates of the principal directions at each point of the surface. An invariant data fit metric based on the angle between surface normals was also derived.

The evaluation method employed was to obtain gradient data for a surface viewed from different directions, apply each reconstruction functional to each view, and match the surfaces from different views to see if the same shape was recovered. Results were presented for synthetic and real data gathered using photometric stereo. Real data experiments were performed with both a simple ellipsoid of known shape and a more complex doll's face (without ground truth).

A brief synopsis of the results is presented in Section 7.1. Various issues were raised during the course of experimentation, and are discussed in Section 7.2. Ideas for extensions and refinements to the work are outlined in Section 7.3.

7.1 Summary of Results

Some degree of invariance was achieved, though the choice of smoothing functional proved to be less important than anticipated. The invariant data fit constraint metric proved unsuitable as part of the regularizing functional, and was abandoned. It was proved that there is no invariant, quadratic data fit term, and a non-invariant metric was adopted, thereby somewhat compromising the intent of the thesis to perform invariant reconstruction.

Ω_1 is equivalent to the quadratic variation stabilizer used in previous work by Grimson (Grimson, 1981) and Terzopoulos (Terzopoulos, 1988), and produced fairly invariant results when gradient values were relatively small. Ω_2 performs similarly, but seemed more robust in the presence of (synthetic) noise. It was realized that the method of implementing Ω_3 was inadequate, so only a few experiments were run with this functional.

Real data from photometric stereo were introduced in the context of building full 3D models of objects by splicing together surfaces recovered from multiple views. Invariance is needed for the different views to be meshed together. Results were somewhat equivocal, as other factors such as boundary conditions and data collection seem to play a greater role than the choice of smoothing functional. Both Ω_1 and Ω_2 smoothing appeared to give similar results. Experiments were performed for four views of a simple ellipsoid of known shape and for two views of the more complex surface of a doll's face.

Models were generated for the ellipsoid with some success. The biggest problem is suggested to be the propagation of errors across the surface, as gradient data provides only relative constraints on local surface height. Knowledge of suitable boundary conditions can be used to mitigate this problem.

Experiments were also conducted with data from a more complicated doll's head object. The presence of distinctive features on the surface led to more successful alignment.

Again, little difference was observed between the two smoothing functionals. The coherence of the data obtained from photometric stereo meant that global smoothing was perhaps unnecessary for either the ellipsoid or the doll's face.

3D model acquisition by splicing surfaces obtained from multiple views appears feasible, but perhaps not optimal. An alternative approach is to align and integrate the gradient data from multiple views before attempting to reconstruct surface height. For some applications, orientation-based representations of shape that do not explicitly require surface height can be used (Li, 1993), and the need to construct surface height from gradient can perhaps be avoided altogether.

7.2 Issues Raised

A variety of issues have been raised by this work, many of which were not foreseen at the outset. These problems are primarily concerned with the form of the variational formulation used to solve the surface reconstruction problem.

Regularizing a problem generally involves choosing a smoothing term and a fit to the data term. Appropriate choices for these norms are imperative to a successful solution. This was made painfully obvious by the failure of the invariant data fit term that was proposed. This data fit term does not match well with the smoothing functionals and leads to excessive smoothing. The constraint metrics suggested are not L_2 norms, which may be the source of the difficulty (Tikhonov & Arsenin, 1977).

A major goal of this thesis was to compare results obtained from different regularizing functionals combining different smoothing and data fit terms. It was rapidly realized that in order to do sensible comparisons, it is necessary to “normalize” the regularizing parameter λ controlling the tradeoff between smoothness and data fit. If this is not done, then different functionals will produce different degrees of smoothing and it becomes hard

to make meaningful comparisons. This is an issue that has apparently not been dealt with in the literature. For example, Stevenson and Delp use a constant value for λ across all the smoothing functionals that they compared (Stevenson & Delp, 1992).

The smoothing functional chosen was the energy of an idealized thin plate under deformation. A parameter ν of the formulation is an elastic property of the material being modelled. Variation in this parameter was observed to have little effect on the reconstruction, as has been noted previously (Terzopoulos, 1988).

Smoothing was applied to the gradient data, and a separate stage of processing integrated the gradients to yield surface height. The choice of functional used for the integration is not considered to be important, particularly when the gradient map is close to integrable. Integrability was not enforced in the present work, but should have been.

Viewpoint invariance was the guiding light for the choice of smoothing functionals, but the emphasis was placed on deriving functionals that were invariant on a pointwise basis. An issue that was overlooked was the change in surface parameterization that results from resampling the surface on a grid imposed by a different viewing direction. When working with range data, distances in 3D space can be used to aid in solving this correspondence problem, but for gradient data this becomes more difficult. The surface could be reparameterized using coordinates intrinsic to the surface, rather than imposed by the viewpoint. Such a framework readily lends itself to invariant reconstruction. Such a reparameterization method has been developed for positional data (Bolle & Vemuri, 1991). The formulation relies on relative positions between adjacent points, and so could be readily extended to gradient data.

Comparison of different smoothing functionals was the primary focus of this work, but the choice of smoothing functional is not the only factor to consider in trying to

achieve viewpoint invariance. Other factors such as boundary conditions, tradeoffs between smoothness and data fit, and collection of data may overwhelm the effects of the smoothing functional. In particular, if noise affects the data in a viewpoint dependent manner, then it may be impossible to achieve invariant reconstruction. For the photometric stereo data that was used, the inherent smoothness of the data meant that the application of global smoothing may have been a flawed idea from the start.

7.3 Future Work

The techniques developed for surface reconstruction suffer from a number of shortcomings that are pointed out in the previous section. Many of these problems are soluble and various extensions to the work are also possible.

Integrability was not enforced during smoothing of the gradient map, an oversight that should be corrected. An easy way to do this is to smooth the gradient map and recover surface height simultaneously, so that numerical estimates of the derivative of the height estimate can be compared to the gradient map to derive a penalty term that encourages integrability. This method has been successfully employed in previous work on shape from shading (Horn, 1990; Ascher & Carter, 1993).

Another alternative is to project the gradient map onto the “closest” integrable map (Frankot & Chellappa, 1988), either after each smoothing iteration or after the smoothing is complete. Once an integrable gradient field is thus derived, surface height can be recovered by direct integration.

Testing different smoothing functionals was the original intent of the thesis, and the issue of discontinuities and boundary conditions was consciously ignored. Tests with real data showed this to be a shortcoming. The importance of boundary conditions for height reconstruction from gradient data has been demonstrated in (Li, 1991), and their

inclusion should be relatively straightforward.

A possibility that has not been explored is the use of alternative representations of slope to aid in the derivation of invariant functionals. Unit normals or stereographic coordinates (Horn, 1990) may prove easier to work with in an invariant manner.

Integration of slope and range data in the context of invariant surface reconstruction may prove fruitful. Slope and range data have previously been combined for reconstruction purposes (Harris, 1987), and invariant reconstruction from range data has received attention (Stevenson & Delp, 1992). Combining range and gradient data allows for integration of different visual modules and thus greater robustness in surface recovery. In addition, tagging gradient points with a range value may aid in deriving functionals that are invariant to the reparameterization of the surface induced by a new viewing direction. (Alternatively, the surface could be reparameterized in object-centered coordinates (Bolle & Vemuri, 1991).)

Another source of data to be integrated is curvature data, such as can be determined from photometric stereo (Woodham, 1990). Curvature information could be used to independently determine principal directions on the surface and thus provide a better way of parameterizing Ω_3 . The expression for the energy density of a thin plate is based on changes in curvature (Pogorelov, 1988). Curvature data could be used to define a non-planar rest state for the plate that is fitted to the data, a novel idea that seems appropriate for the recovery of the complex surfaces one expects to find in typical scenes.

For 3D model acquisition, reconstructing explicit surfaces in several views and then splicing the surfaces together leads to inconsistencies and errors at the points of overlap. An alternative is to match gradient data (or, rather, surface normals) directly from multiple views and then reconstruct the full 3D surface from the merged data. For example, it is possible to reconstruct a convex polyhedron from its extended Gaussian image (Little, 1985). Note that in the work presented here, 3D model acquisition

was intended as a means to test the invariance of the reconstruction algorithms, rather than an end in itself.

Bibliography

- Anton, H. (1987). *Elementary Linear Algebra*, 5th edition. John Wiley & Sons, Toronto.
- Ascher, U. M. & Carter, P. M. (1993). A multigrid method for shape from shading. *SIAM Journal of Numerical Analysis*, 30(1), 102–115.
- Besl, P. J. & McKay, N. D. (1992). A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239–256.
- Blake, A. (1989). Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 2–12.
- Bolle, R. M. & Vemuri, B. C. (1991). On 3D surface reconstruction methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(1), 1–13.
- Burns, J. B., Weiss, R. S. & Riseman, E. M. (1993). View variation of point-set and line-segment features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(1).
- Chen, H. H. & Huang, T. S. (1988). A survey of the construction and manipulation of octrees. *Computer Vision, Graphics, and Image Processing*, 43, 409–431.
- Courant, R. & Hilbert, D. (1989). *Methods of Mathematical Physics*, volume 1. John Wiley & Sons, Toronto.
- Faddeev, D. K. & Faddeeva, V. N. (1963). *Computational Methods of Linear Algebra*. W. H. Freeman and Company, San Francisco, Translated from Russian by R. C. Williams.
- Faugeras, O. D., Hebert, M., Mussi, P. & Boissonnat, J. D. (1984). Polyhedral approximation of 3D objects without holes. *CVGIP*, 25, 169–183.
- Ferrie, F. P., Lagarde, J. & Whaite, P. (1990). Recovery of volumetric object descriptions from laser rangefinder images. In *European Conference on Computer Vision*, pages 387–396.

- Frankot, R. T. & Chellappa, R. (1988). A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10, 439–451.
- Goad, C. (1983). Special purpose automatic programming for 3D model-based vision. In *Proceedings of the ARPA Image Understanding Workshop*.
- Grimson, W. E. L. (1981). *From Images to Surfaces: A Computational Study of the Human Early Visual System*. The MIT Press Series in Artificial Intelligence. The MIT Press, Cambridge, Massachusetts.
- Harris, J. G. (1987). A new approach to surface reconstruction: the coupled depth/slope model. In *International Conference on Computer Vision*, pages 277–283. IEEE.
- Horn, B. K. P. & Brooks, M. J. (1986). The variational approach to shape from shading. *Computer Vision Graphics and Image Processing*, 33, 174–208.
- Horn, B. K. P. (1986). *Robot Vision*. MIT Press/McGraw-Hill, Cambridge, MA.
- Horn, B. K. P. (1990). Height and gradient from shading. *International Journal of Computer Vision*, 5, 37–75.
- Jastrzebski, Z. D. (1976). *The Nature and Properties of Engineering Materials*, 2nd edition. John Wiley & Sons, Toronto.
- Koenderink, J. J. & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Acta Optica*, 27, 918–996.
- Li, Y. (1991). Surface reconstruction by coupled depth/slope model with natural boundary conditions. Technical Report 91-24, University of British Columbia Department of Computer Science, Vancouver.
- Li, Y. (1993). *Orientation-based Representations of Shape and Attitude Determination*. PhD thesis, University of British Columbia, Vancouver, British Columbia.
- Little, J. J. (1985). *Recovering shape and determining attitude from extended Gaussian images*. PhD thesis, The University of British Columbia, Vancouver, BC, Ph.D., supervisor R.J. Woodham.
- Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31, 355–395.
- Maddox, I. J. (1988). *Elements of Functional Analysis*, 2nd edition. Cambridge University Press, Cambridge.

- Poggio, T., Torre, V. & Koch, C. (1985). Computational vision and regularization theory. *Nature*, 317, 314–319.
- Pogorelov, A. V. (1988). *Bending of Surfaces and Stability of Shells*. American Mathematical Society, Providence, Rhode Island, Translated from Russian by J. L. Schulenberger.
- Pogorelov, A. V. (undated). *Differential Geometry*. P. Noordhoff N. V., Groningen, The Netherlands, Translated from Russian by L. F. Boron.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1988). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge.
- Roberts, L. G. (1965). Machine perception of three dimensional solids. In Tippett, J. T., Berkowitz, D., Clapp, L., Koester, C. & Vanderburgh, A., editors, *Optical and Electro-optical Information Processing*, pages 159–197. MIT Press, Cambridge, MA.
- Stevenson, R. L. & Delp, E. J. (1990). Invariant reconstruction of curves and surfaces with discontinuities with applications in computer vision. Technical Report TR-EE 90-35, Purdue University.
- Stevenson, R. L. & Delp, E. J. (1992). Viewpoint invariant recovery of visual surfaces from sparse data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(9), 897–909.
- Szeliski, R. & Tonnesen, D. (1991). Surface modeling with oriented particle systems. Technical Report 91/14, Digital Equipment Corporation, Cambridge Research Lab.
- Szeliski, R. (1990). Real-time octree generation from rotating objects. Technical Report 90/12, Digital Equipment Corporation, Cambridge Research Lab.
- Szeliski, R. (1990). Shape from rotation. Technical Report 90/13, Digital Equipment Corporation, Cambridge Research Lab.
- Terzopoulos, D. (1988). The computation of visible surface representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4), 417–438.
- Tikhonov, A. N. & Arsenin, V. Y. (1977). *Solutions of ill-posed problems*. V.H. Winston & Sons, Washington, DC.
- Tsai, P-S. & Shah, M. (1992). A simple shape from shading algorithm. Technical Report CS-TR-92-24, University of Central Florida, Orlando, Florida.

- Vemuri, B. C. & Sifteland, G. (1992). Invariant surface and motion estimation from sparse range data. *Journal of Mathematical Imaging and Vision*, 1(1), 43–64.
- Volterra, E. & Gaines, J. H. (1971). *Advanced Strength of Materials*. Prentice-Hall, Inc., Englewood Cliffs, NJ.
- Wahba, G. (1990). *Spline Models for Observational Data*, volume 59 of *Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia.
- Wang, W. & Iyengar, S. S. (1992). Efficient data structures for model-based 3D object recognition and localization from range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10), 1035–1045.
- Woodham, R. J. (1980). Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19, 139–144.
- Woodham, R. J. (1990). Surface curvature from photometric stereo. Technical Report 90-29, University of British Columbia.

Invariant Quadratic Constraint Metric Impossible

A constraint metric for fit to gradient data is a function $\rho(z_x, z_y, p, q)$ that measures the “closeness” of the gradient (z_x, z_y) at a point on an estimated surface to the corresponding data point (p, q) . To achieve invariant surface reconstruction requires an invariant constraint metric. For ease of minimization, a constraint metric quadratic in z_x and z_y is desirable. Unfortunately, these two qualities are incompatible—no quadratic function of z_x and z_y is invariant to rotations.

Consider the surface normals $(-z_x, -z_y, 1)$ and $(-p, -q, 1)$ corresponding to surface and data gradients. A metric that depended only on the angle between these surface normals would be invariant to rotations of the data and surface. Intuitively, it seems reasonable that such a metric would not be quadratic—if the metric is independent of the length of the vectors, then some form of normalization must be involved, which is a non-quadratic procedure.

An arbitrary constraint metric that is quadratic in z_x and z_y can be written as

$$\rho(z_x, z_y, p, q) = k_1(p, q)z_x^2 + k_2(p, q)z_y^2 + k_3(p, q)z_x z_y + k_4(p, q)z_x + k_5(p, q)z_y + k_6(p, q) \quad (\text{A.1})$$

where the $k_i(p, q)$ are arbitrary functions of p and q . If ρ is invariant to rotations, then

$$\rho(z_x, z_y, p, q) = \rho(z'_x, z'_y, p', q') \quad (\text{A.2})$$

where z'_x, z'_y, p', q' are the rotated gradients.

The rotation of a surface normal vector can be calculated by multiplying by a rotation matrix

$$\begin{aligned} \mathbf{n}' &= R\mathbf{n} \\ (n'_x, n'_y, n'_z) &= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \end{aligned} \quad (\text{A.3})$$

The rotated gradients can thus be calculated as

$$z'_x = -\frac{n'_x}{n'_z} = -\frac{r_{11}z_x + r_{12}z_y - r_{13}}{r_{31}z_x + r_{32}z_y - r_{33}} \quad (\text{A.4})$$

$$z'_y = -\frac{n_y}{n_z} = -\frac{r_{21}z_x + r_{22}z_y - r_{23}}{r_{31}z_x + r_{32}z_y - r_{33}} \quad (\text{A.5})$$

Substituting these expressions for the rotated gradients into Equation A.1 and plugging the results into Equation A.2, cross-multiplying, and then equating coefficients of like powers of z_x and z_y yields 15 equations which must be satisfied by the coefficient functions k_i if ρ truly is invariant

$$(1) \quad -r_{13}^2 k_1(p', q') - r_{23}^2 k_2(p', q') - r_{13} r_{23} k_3(p', q') + r_{13} r_{33} k_4(p', q') + r_{23} r_{33} k_5(p', q') + r_{33}^2 k_6(p, q) - r_{33}^2 k_6(p', q') = 0$$

$$(2) \quad 2 r_{12} r_{13} k_1(p', q') + 2 r_{22} r_{23} k_2(p', q') + r_{13} r_{22} k_3(p', q') + r_{12} r_{23} k_3(p', q') - r_{13} r_{32} k_4(p', q') - r_{12} r_{33} k_4(p', q') + r_{33}^2 k_5(p, q) - r_{23} r_{32} k_5(p', q') - r_{22} r_{33} k_5(p', q') - 2 r_{32} r_{33} k_6(p, q) + 2 r_{32} r_{33} k_6(p', q') = 0$$

$$(3) \quad -r_{12}^2 k_1(p', q') + r_{33}^2 k_2(p, q) - r_{22}^2 k_2(p', q') - r_{12} r_{22} k_3(p', q') + r_{12} r_{32} k_4(p', q') - 2 r_{32} r_{33} k_5(p, q) + r_{22} r_{32} k_5(p', q') + r_{32}^2 k_6(p, q) - r_{32}^2 k_6(p', q') = 0$$

$$(4) \quad -2 r_{32} r_{33} k_2(p, q) + r_{32}^2 k_5(p, q) = 0$$

$$(5) \quad r_{32}^2 k_2(p, q) = 0$$

$$(6) \quad 2 r_{11} r_{13} k_1(p', q') + 2 r_{21} r_{23} k_2(p', q') + r_{13} r_{21} k_3(p', q') + r_{11} r_{23} k_3(p', q') + r_{33}^2 k_4(p, q) - r_{13} r_{31} k_4(p', q') - r_{11} r_{33} k_4(p', q') - r_{23} r_{31} k_5(p', q') - r_{21} r_{33} k_5(p', q') - 2 r_{31} r_{33} k_6(p, q) + 2 r_{31} r_{33} k_6(p', q') = 0$$

$$(7) \quad -2 r_{11} r_{12} k_1(p', q') - 2 r_{21} r_{22} k_2(p', q') + r_{33}^2 k_3(p, q) - r_{12} r_{21} k_3(p', q') - r_{11} r_{22} k_3(p', q') - 2 r_{32} r_{33} k_4(p, q) + r_{12} r_{31} k_4(p', q') + r_{11} r_{32} k_4(p', q') - 2 r_{31} r_{33} k_5(p, q) + r_{22} r_{31} k_5(p', q') + r_{21} r_{32} k_5(p', q') + 2 r_{31} r_{32} k_6(p, q) - 2 r_{31} r_{32} k_6(p', q') = 0$$

$$(8) \quad -2 r_{31} r_{33} k_2(p, q) - 2 r_{32} r_{33} k_3(p, q) + r_{32}^2 k_4(p, q) + 2 r_{31} r_{32} k_5(p, q) = 0$$

$$(9) \quad 2 r_{31} r_{32} k_2(p, q) + r_{32}^2 k_3(p, q) = 0$$

$$(10) \quad r_{33}^2 k_1(p, q) - r_{11}^2 k_1(p', q') - r_{21}^2 k_2(p', q') - r_{11} r_{21} k_3(p', q') - 2 r_{31} r_{33} k_4(p, q) + r_{11} r_{31} k_4(p', q') + r_{21} r_{31} k_5(p', q') + r_{31}^2 k_6(p, q) - r_{31}^2 k_6(p', q') = 0$$

$$(11) \quad -2 r_{32} r_{33} k_1(p, q) - 2 r_{31} r_{33} k_3(p, q) + 2 r_{31} r_{32} k_4(p, q) + r_{31}^2 k_5(p, q) = 0$$

$$(12) \quad r_{32}^2 k_1(p, q) + r_{31}^2 k_2(p, q) + 2 r_{31} r_{32} k_3(p, q) = 0$$

$$(13) \quad -2 r_{31} r_{33} k_1(p, q) + r_{31}^2 k_4(p, q) = 0$$

$$(14) \quad 2 r_{31} r_{32} k_1(p, q) + r_{31}^2 k_3(p, q) = 0$$

$$(15) \quad r_{31}^2 k_1(p, q) = 0$$

These equations place very harsh constraints on the coefficient functions:

- (i) Equation (15), together with the fact that, in general, the components of the rotation matrix R are non-zero, implies $k_1(p, q) = 0$.
- (ii) Similarly, (5) indicates $k_2(p, q) = 0$.
- (iii) Equation (9) and (ii) imply $k_3(p, q) = 0$.

(iv) Equation (13) and (i) imply $k_4(p, q) = 0$.

(v) Equation (4) and (ii) imply $k_5(p, q) = 0$.

So it must be the case that $\rho(z_x, z_y, p, q) = k_6(p, q)$, which is independent of the values of z_x and z_y and thus clearly useless as a constraint metric.

Interpolation by Iterated Local Averaging

Given a set of data points $\mathcal{D} = \{z_{ij}\}$ at a sparse set of points (i, j) on a regular $M \times N$ grid, we want to fill in the missing values in the grid by interpolating the data points. A simple way of doing this is iterated local averaging (ILA), which repeatedly replaces the value at a grid point (i, j) with the average of the values over the local neighbourhood. The data values act as “anchors” and are left unchanged.

The algorithm is not particularly efficient, taking more iterations than necessary, but each iteration is rapid. The algorithm can be stated as follows:

```

while not done
  for i = 1 to M
    for j = 1 to N
      if Z[i,j] not in data set
        sum = 0
        n = 0
        if i > 1 then sum += Z[i-1,j]; n += 1
        if i < M then sum += Z[i+1,j]; n += 1
        if j > 1 then sum += Z[i,j-1]; n += 1
        if i < N then sum += Z[i,j+1]; n += 1
        Z[i,j] = sum/n
      endif
    endfor
  endfor
endwhile

```

A suitable stopping criterion is when no value changes by more than some small value ϵ . As presented, the algorithm sweeps sequentially across the image—conceptually it should act in parallel, though (in the limit) parallel and serial implementations are equivalent.

The membrane energy functional measures the deformation energy of an idealized

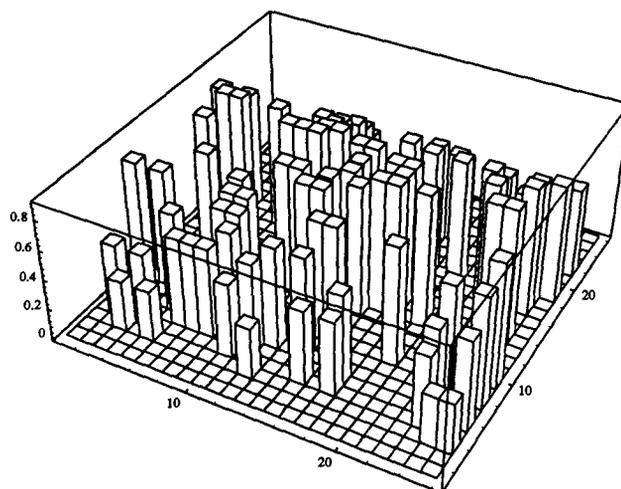


Figure B.1: Sparse data on a grid.

membranous surface. As an interpolation method, it fits a piecewise planar surface to a sparse set of data. The membrane functional is

$$\iint z_x^2 + z_y^2 dx dy$$

The ILA algorithm can be shown to be equivalent to steepest descent minimization of the membrane equation, using the forward difference discretization

$$\hat{z}_x = z_{i,j+1} - z_{i,j}$$

$$\hat{z}_y = z_{i+1,j} - z_{i,j}$$

Figure B.1 shows a sparse data set consistent with a cylinder aligned along the axis. Z values are defined at 101 of 625 points on a 25×25 grid. Figure B.2 shows the interpolated surface. The reconstruction took 107 iterations to reach a point where no surface height value changes by more than 0.000001 in a single iteration. Implemented in C, this procedure took about 2 seconds on a Sparc workstation.

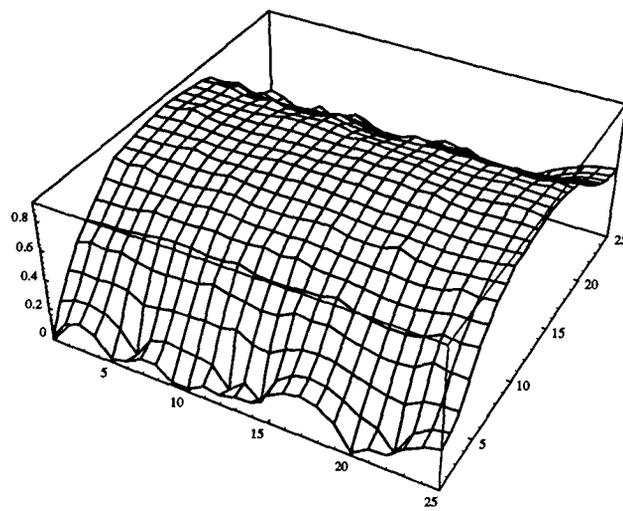


Figure B.2: Surface interpolated from sparse data.

Conjugate Gradient Descent

The problem is to find the solution vector \mathbf{z} which minimizes a given measurement function $M(\mathbf{z}^*)$. Suppose M can be approximated by the quadratic form

$$M(\mathbf{z}) = \frac{1}{2}\mathbf{z}^T Q \mathbf{z} - \mathbf{b}^T \mathbf{z} \quad (\text{C.1})$$

where Q is an $n \times n$ symmetric positive definite matrix, and any constant term has been dropped as irrelevant to the minimization problem. The elements of Q correspond to the coefficients of quadratic combinations of the elements of \mathbf{z}

$$q_{ij} = \begin{cases} \text{the coefficient of } z_i z_j \text{ if } i \neq j \\ \text{twice the coefficient of } z_i^2 \text{ if } i = j \end{cases} \quad (\text{C.2})$$

while linear terms are of the form $b_i z_i$.

Starting at a point $\mathbf{z}_0 \in V$, the task is to generate a sequence $\{\mathbf{z}_k\}$ that converges to the function minimum. At each step in the process, let \mathbf{g}_k be the gradient of the function

$$\mathbf{g}_k = \nabla M(\mathbf{z}_k) = Q \mathbf{z}_k - \mathbf{b} \quad (\text{C.3})$$

The chosen direction in which to move at each step will be denoted \mathbf{d}_k , with $\mathbf{d}_0 = -\mathbf{g}_0$. Finding the line minimum in the direction \mathbf{d}_k is equivalent to finding α_k to minimize the one-dimensional function

$$h_k(\alpha) = M(\mathbf{z}_k + \alpha \mathbf{d}_k) \quad (\text{C.4})$$

For the quadratic case, we can solve the one-dimensional minimization analytically by setting the derivative to zero and solving for α

$$\frac{dh_k}{d\alpha}(\alpha_k) = \alpha_k \mathbf{d}_k^T Q \mathbf{d}_k + \mathbf{z}_k^T Q \mathbf{d}_k - \mathbf{b}^T \mathbf{d}_k = 0 \quad (\text{C.5})$$

yielding

$$\alpha_k = -\frac{(\mathbf{z}_k^T Q - \mathbf{b}^T)\mathbf{d}_k}{\mathbf{d}_k^T Q \mathbf{d}_k} = -\frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{d}_k^T Q \mathbf{d}_k} \quad (\text{C.6})$$

We now move to the new point

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \alpha_k \mathbf{d}_k \quad (\text{C.7})$$

All that remains is to specify how to choose the new optimized direction

$$\mathbf{d}_{k+1} = \beta_k \mathbf{d}_k - \mathbf{g}_{k+1} \quad (\text{C.8})$$

where

$$\beta_k = \frac{\mathbf{g}_{k+1} \cdot \mathbf{g}_{k+1}}{\mathbf{g}_k \cdot \mathbf{g}_k} \quad (\text{C.9})$$

This is the method of Fletcher-Reeves. (Press *et al*, 1988) suggest that the Polak-Ribiere variation works better when the function being minimized is not quadratic. The only difference is in the calculation of β_k

$$\beta_k = \frac{(\mathbf{g}_{k+1} - \mathbf{g}_k) \cdot \mathbf{g}_{k+1}}{\mathbf{g}_k \cdot \mathbf{g}_k} \quad (\text{C.10})$$

The net effect of the conjugate gradient method is that all the \mathbf{d}_k are mutually conjugate and the \mathbf{g}_k are mutually orthogonal (i.e. $\mathbf{d}_i^T Q \mathbf{d}_j = 0$ and $\mathbf{g}_i^T \mathbf{g}_j = 0$, for all $i \neq j$). The process is analogous to the Gram-Schmidt method for finding an orthogonal set of basis vectors for an inner product space (Anton, 1987).

As with steepest descent, the iterative process is halted when suitable termination criteria are met. The gradient is zero at the solution, so the norm of the gradient can be used as a stopping criterion. Since the minimization problem is equivalent to the solution of the linear system $Q\mathbf{z} = \mathbf{b}$, we see that the gradient is just the residual $Q\hat{\mathbf{z}} - \mathbf{b}$. Alternatively, iteration can be halted when the solution remains relatively unchanged from one iteration to the next. This was the approach chosen, with a threshold being

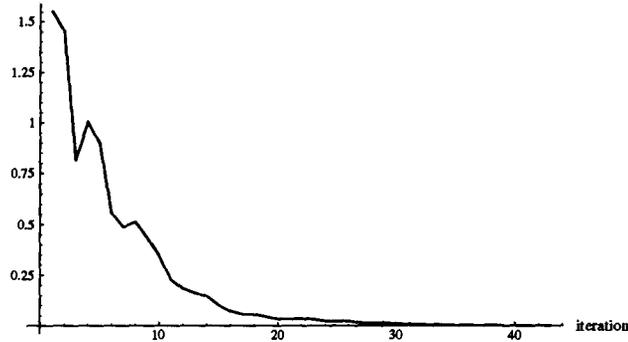


Figure C.1: Maximum change in any variable as a function of iteration number. This value was used as a convergence criterion during conjugate gradient descent.

applied to $\|\mathbf{z}_k - \mathbf{z}_{k-1}\|^\infty$ —i.e. the process was halted when no element of \mathbf{z} changed by more than some small constant ε in a given iteration. This seemed to give reasonable results.

Figure C.1 shows the value of this norm for a typical run. The example is for gradient map smoothing of the first synthetic surface (shown in Figures 5.2 and 5.3) using Ω_2 and ρ_{SD} , for which 43 iterations were sufficient for convergence to the 0.001 level.

Note that the notion of conjugate directions is really only meaningful for quadratic forms, as it depends on the matrix Q . Under the assumption that the function M is approximately quadratic, the calculation can, however, proceed without explicit knowledge of Q , provided that the gradient can be determined and the one-dimensional minimization required to determine α_k can be performed. In the system as implemented, the analytic solution for α_k is used when the function is quadratic, while Brent's minimization method (Press *et al*, 1988) is used in the absence of a quadratic form. In addition, the Fletcher-Reeves technique is used for quadratics, while the Polak-Ribiere variant is used for non-quadratics.