



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

RECOVERING SHAPE FROM SHADING AND CONTOUR

Pierre Breton

Department of Electrical Engineering
McGill University

November 1994

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

© PIERRE BRETON, 1994



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN
IRREVOCABLE NON-EXCLUSIVE
LICENCE ALLOWING THE NATIONAL
LIBRARY OF CANADA TO
REPRODUCE, LOAN, DISTRIBUTE OR
SELL COPIES OF HIS/HER THESIS BY
ANY MEANS AND IN ANY FORM OR
FORMAT, MAKING THIS THESIS
AVAILABLE TO INTERESTED
PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE
IRREVOCABLE ET NON EXCLUSIVE
PERMETTANT A LA BIBLIOTHEQUE
NATIONALE DU CANADA DE
REPRODUIRE, PRETER, DISTRIBUER
OU VENDRE DES COPIES DE SA
THESE DE QUELQUE MANIERE ET
SOUS QUELQUE FORME QUE CE SOIT
POUR METTRE DES EXEMPLAIRES DE
CETTE THESE A LA DISPOSITION DES
PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP
OF THE COPYRIGHT IN HIS/HER
THESIS. NEITHER THE THESIS NOR
SUBSTANTIAL EXTRACTS FROM IT
MAY BE PRINTED OR OTHERWISE
REPRODUCED WITHOUT HIS/HER
PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE
DU DROIT D'AUTEUR QUI PROTEGE
SA THESE. NI LA THESE NI DES
EXTRAITS SUBSTANTIELS DE CELLE-
CI NE DOIVENT ETRE IMPRIMES OU
AUTREMENT REPRODUITS SANS SON
AUTORISATION.

ISBN 0-612-05679-1

Canada

ABSTRACT

The classical approach to shape from shading problems is to find a numerical solution of the image irradiance partial differential equation. It is usually assumed that the parameters of this equation (the light source direction and surface albedo) can be estimated in advance. For images which contain shadows and occluding contours, this decoupling of problems is artificial and the classical approach fails. We develop a new approach to solving these equations using the image geometric structures instead of the image photometric structure. Our approach is based on modern differential geometry, and solves for light source and surface shape changes concurrently. Local scene elements (scenels) are estimated from the shading flow field and the tangent field, and smoothness, material, and light source compatibility conditions resolve them into consistent scene descriptions. Shadows and related difficulties for the classical approach are discussed.

RÉSUMÉ

L'approche classique du problème de l'inférence des formes à partir de la variation de l'intensité lumineuse ou en anglais "*Shape from Shading*" consiste à trouver une solution numérique aux équations différentielles partielles qui décrivent la réflexion de la lumière sur une surface mate. L'estimation des paramètres de ces équations (la direction de la source de lumière et le coefficient de réflexion de la surface) est toujours présumée préalablement résolue. En général, ce découplage du problème de l'inférence de forme et de l'estimation des paramètres est artificiel et l'approche classique donne des résultats erronés. Nous développons une nouvelle approche pour résoudre ces équations en se servant de structures géométriques extraites de l'image au lieu de la structure photométrique. Notre approche est fondée sur la géométrie différentielle moderne, et résout simultanément les changements d'illumination et de forme. Des descriptions locales de la scène sont estimées à partir de ces structures géométriques de l'image. Des conditions de compatibilité liées à la continuité des surfaces et des conditions d'éclairage permettent par la suite d'identifier une description cohérente. Les cas-problèmes de l'approche classique sont étudiés.

ACKNOWLEDGEMENTS

First, I would like to express my sincerest thanks to Professor Steven Zucker, my Ph.D. supervisor: his altruism created an atmosphere so conducive to research; his presence, his insightful feedback, his constant support made these years at McGill a very pleasant learning experience.

Throughout the years of my graduate studies at CIM, I had the privilege and the pleasure of interacting with and learning from many individuals. I would like to thank all of them:

Allan Dobbins, Benoit Dubuc, James Elder, Benjamin Kimia, Lee Iverson, David Lamb, Michael Langer, and the late Doug Miller, for all the discussion that took place around our office;

André Foisy, Thomas Jelonek, Frederic Leymarie, Anna Lin, and Peter Whaite, for their contributions to the lab's spirit;

Frank Ferrie and Vincent Hayward, for taking time to be on my committee and for their presence in the lab.

I would like to thank Martin Grant (who was co-supervisor at my arrival at McGill) for taking time to be on my committee and for his support.

I would like to extend special thanks to Lee Iverson and Peter Whaite, for endless little reasons, let's just say for all the time they gave me.

Special thanks also goes to Michael Langer, for reviewing a draft copy of this thesis, and for allowing me to use figures of joint papers; and to Michael Glaum, for his help and review of the more theoretical sections.

I am grateful to Professor André-Marie Tremblay, my M.Sc. supervisor, who motivated my interest in this field of research.

Finally, I would like to thank Karima, for everything that cannot be expressed in words. Un gros merci...

TABLE OF CONTENTS

ABSTRACT	ii
RÉSUMÉ	iii
ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	x
LIST OF TABLES	xiii

Part 1

ON THE SHAPE FROM SHADING PROBLEM

CHAPTER 1. Introduction	2
1. OVERVIEW	5
2. CLAIMS OF ORIGINALITY	8
CHAPTER 2. Overview of the Vision Problem	10
1. THE GENERAL VISION PROBLEM	10
1.1. Data and Knowledge	11
1.2. Representations of the Image and the Scene	13
2. ELEMENTS OF THE SOLUTION	14
2.1. Low-level and High-level Vision	14
2.2. Data-driven and Knowledge-based Approaches	15
CHAPTER 3. The Classical Shape from Shading Problem	17
1. BRIEF HISTORY	17
2. CLASSICAL DEFINITION OF THE SHAPE FROM SHADING PROBLEM	19

TABLE OF CONTENTS

3. ANALYSIS OF A "CLASSICAL" SHAPE FROM SHADING ALGORITHM	22
3.1. Overview of Bichsel and Pentland's Algorithm	22
3.2. Solution to the "Classical" Shape from Shading Problem	22
3.3. Limitations and Difficulties	26

Part 2

THE SHADING FLOW FIELD

CHAPTER 4. The General Shape from Shading Problem	31
1. THE OBJECTIVES	31
2. THE LINE OF THOUGHT	32
3. LOCAL VS. GLOBAL SCENE CONSTRAINTS	34
3.1. Generalizing the Scene Model	34
3.2. Introducing a Model with Discontinuities	35
3.3. Coupling Surface and Light Source Estimation	36
4. GEOMETRIC VS. PHOTOMETRIC IMAGE STRUCTURES	36
4.1. The Tangent Field	38
4.2. The Shading Flow Field	39
4.2.1. Other Motivations	
5. SUMMARY	43
6. THE GENERALIZED SHAPE FROM SHADING PROBLEM	44
CHAPTER 5. Computing the Shading Flow Field	46
1. INTRODUCTION	46
1.1. Fuzzy Derivatives	47
1.2. Limitation and Difficulties	48
2. PIECEWISE CONTINUOUS FLOW FIELD	50
2.1. Computing Intensity Discontinuities	50
2.2. Defining the Shading Flow Field Operator's Domain	52
2.2.1. Discrete Shading Flow Field Representation	
3. RELATION TO RELAXATION LABELLING	54
3.1. The General Relaxation Labelling Paradigm	55
3.2. The Two-Labels Relaxation Labelling Paradigm	57
3.3. Implementation Remarks	58

TABLE OF CONTENTS

4. COHERENT SHADING FLOW FIELD	58
4.1. Using Curvatures as Constraints	58
4.2. Using Differential Invariants as Constraints	59
4.3. Compatibilities between Local Shading Flow Descriptions	60
5. NUMERICAL RESULTS	61
5.1. Abrupt Albedo Change	61
5.2. Multiple Light Sources	63
5.3. The Car's Fender	63
CHAPTER 6. Interpretations of Discontinuities	66
1. VARIOUS CASES OF SCENE DISCONTINUITY	66
1.1. Shadows	67
1.1.1. The Cast Shadow Boundary	
1.1.2. The Attached Shadow Boundary	
1.2. Change in Reflectivity	69
1.3. Occlusions and Other Surface Discontinuities	70
2. SUMMARY OF RESULTS	70

Part 3

THE SCENEL BUNDLE APPROACH

CHAPTER 7. The Scenel Bundle Approach	75
1. THE OUTLINE OF THE SCENEL BUNDLE APPROACH	75
1.1. Informal Description of Scenel Bundles	80
2. THE LOCAL REPRESENTATION OF THE SCENE	81
2.1. The Illumination Model	81
2.2. The Surface Model	82
2.3. Fibre Bundles	87
2.4. The Scenel Bundle	88
2.5. Topology	89
2.6. Geometry	90
CHAPTER 8. The Scene Element	92
1. THE DISCRETE SCENEL SPACE	92
1.1. The Quantization of the Scenel Attributes	92
1.2. The Set of Scenels	94

TABLE OF CONTENTS

2. THE INFERENCE OF THE SCENEL FIBRE	95
2.1. The Local Image Structures generated by the Scenel	96
2.1.1. The Expected Shading Flow Field	
2.1.2. The Expected Edge Map	
2.2. The Labelling	100
3. THE CONSISTENT SCENEL LABELLING	102
3.1. The Solution: a Cross-Section of the Scenel Bundle	102
3.2. Compatibility between Neighbouring Scene Elements	104
CHAPTER 9. Implementation	109
1. NUMERICAL RESULTS	109
1.1. Abrupt Albedo Change	109
1.2. Multiple Light Sources	111
1.3. The Car's Fender	113
CHAPTER 10. Discussion	116
1. FUTURE DIRECTIONS	117
1.1. Implementation and Experimentation	117
1.2. Consequences of the Shading Flow Field	117
1.3. Integration of other Cues	118
1.3.1. Visual Texture	
1.3.2. Highlights and Shadows	
1.4. Improving the Scenel Model	119
1.4.1. Nearby and Extended Light Sources	
1.4.2. Light Interaction with the Ambient Medium	
Appendices	
APPENDIX A. Flows and Scalar Field: Evidence for Multiple Pathways	123
1. INTRODUCTION	123
1.1. The Allman/Zucker Hypothesis	124
1.2. Two Classes of Textures	125
2. PRELIMINARY OBSERVATIONS	127
3. EXPERIMENTAL PARADIGM	128
4. EXPERIMENTAL PROCEDURE	130
	viii

TABLE OF CONTENTS

5. RESULTS	132
6. DISCUSSION	133
APPENDIX B. Higher Order Constraints	136
1. DEFINITIONS OF THE COORDINATE SYSTEMS	136
2. SHADING AND SPIN FORMS	137
2.1. The Partial Derivative of a Surface Patch Normal	137
2.2. The Partial Derivatives of the Image Irradiance	140
2.2.1. Viewer's Coordinates and Surface Patch Local Coordinates	
2.2.2. The Image Irradiance	
2.2.3. The Shading Flow Field	
APPENDIX C. Surface Shape and Occluding Contours	144
1. DIFFERENTIAL PROPERTIES OF A PLANAR CURVE	144
2. SELF OCCLUDING EDGES	145
APPENDIX D. Shading from Shape	147
1. PREVIOUS WORK	147
2. LIGHT RAY MANIFOLD	150
2.1. Local Coordinates: Continuous Case	150
2.2. Local Coordinates: Discrete Case	151
3. COORDINATE EVOLUTION	152
4. PARALLEL IMPLEMENTATION	155
4.1. Memory Costs	155
4.2. Time Costs	155
5. SCATTERING AND ABSORPTION IN FREE SPACE	157
6. SUMMARY	158
APPENDIX E. Labelling	161
REFERENCES	163

LIST OF FIGURES

1.1	An Ideal Image for Classical Shape from Shading	2
1.2	An Example of Broken Assumption	3
1.3	An Image of a Scene illuminated by Two Light Sources	4
1.4	An Image of a Scene with Shadows	5
1.5	Example of Geometrical Structures within the Image	6
1.6	Photometry	6
1.7	Photometric class of solution	7
1.8	Topology	7
1.9	Constraints	8
2.10	An image of a car assembly line	11
3.11	The Classical Setting	20
3.12	Example of a "successful" shape recovery	23
3.13	Inaccuracy due to poor singularity identification: Discretization	24
3.14	Inaccuracy due to poor singularity identification: Noise	25
3.15	Shape recovery with two light sources	27
3.16	Shape Recovery of a Multiple Albedo Scene	28
4.17	The Image Structures	33
4.18	The Line Drawing	38
4.19	The Shading	40
4.20	The Shading Alone	40
5.21	Shading Flow Field from Fuzzy Derivatives	47
5.22	Fuzzy Derivative and Intensity Discontinuity	48
5.23	Fuzzy Derivative and Shading Flow Discontinuity	49
5.24	Gradient with Intensity Discontinuity	51

LIST OF FIGURES

5.25 Neighbouring Data Points	53
5.26 Graphical Interpretation of the Shading Flow Field	60
5.27 Example of a shading flow field recovery	62
5.28 Example of a shading flow field recovery	64
5.29 Fender of a car	65
7.30 The Scene Element	76
7.31 The Scenel Fibre	77
7.32 The Scenel Bundle	78
7.33 Compatibility	79
7.34 Frame, Angles and Coordinates System	83
7.35 Representation of the Normal of the Surface Patch	84
7.36 Representation of the Shape of a Surface Patch	86
8.37 The Voronoi Cells	93
8.38 Variation of the Shading Flow within a Scenel	101
8.39 Variation of the Shading Flow for Neighbouring Scenel	102
8.40 Samples used to Illustrate the Shading Flow Variations	103
8.41 Compatibility	105
8.42 Compatibilities wrt. Voronoi Cells	106
9.43 Example of a shape recovery	110
9.44 Example of a shape recovery	112
9.45 Fender of a car	114
10.46 Geometry of Specular Reflection	119
10.47 Penumbra caused by an Extended Light Source	120
10.48 Mutual Illumination	120
A.49 Two Types of Texture Patterns	126
A.50 The Kolers paradigm for apparent motion	129
A.51 Illustration of the Texture Filling-in Experiments	131
A.52 Psychophysics Results: Variation in Luminance	133
A.53 Psychophysics Results: Variation of Color	134
A.54 Psychophysics Results: Variation in Texture Fields Densities	135
D.55 Radiance of a Ray	149

LIST OF FIGURES

D.56 Graphic representation of the Radiosity Equation	149
D.57 Parameterization of a Light Ray	151
D.58 Discretization of Light Ray Directions	152
D.59 The Light Ray Manifold	153
D.60 Local Transformation	154
D.61 A Ring in a Box	156
D.62 Balls floating in Fog	159

LIST OF TABLES

4.1	Classical vs. Generalized Shape from Shading	31
4.2	Classical vs. Generalized Shape from Shading	35
6.3	Interpretation of Discontinuities	71
6.4	Discontinuities	73
E.5	Labelling of the Surface Normals	161
E.6	Labelling of the Surface Curveness	162
E.7	Labelling of the Surface Shape Vector	162

Part 1

ON THE SHAPE FROM SHADING PROBLEM

CHAPTER 1

Introduction

In this thesis, we address a well known problem in computer vision: shape from shading. The problem consists of recovering the shape of surfaces of a three-dimensional scene from a single static two-dimensional intensity image. It is one of the classical problems, and it has already received a great deal of attention.

Previous research on shape from shading concentrated on images such as Fig. 1.1, i.e., images of a smooth surface of constant albedo illuminated by a single distant light source under Lambert's reflectance model. For such images, it is tempting to decouple the surface estimation problem from the light source estimation problem and the surface albedo estimation problem.

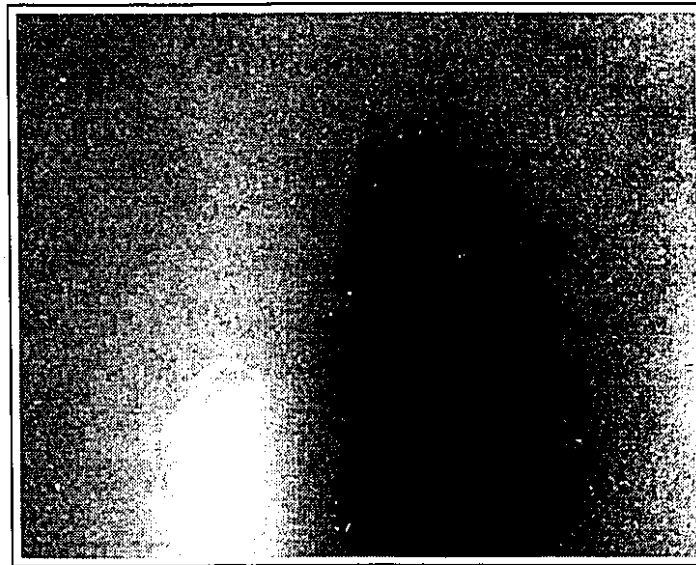


FIGURE 1.1. This image provides an example in which the scene is composed of a single smooth surface of constant albedo which is illuminated from a single distant point source.

However, this is a major drawback, because such a decoupling of problems yields algorithms whose domain of application is very limited. To illustrate, such algorithms could not deal with:

- images of scenes with shadows, i.e. parts of the scene that are illuminated differently;
- images of scenes in which surfaces have different albedos (or reflectance coefficients), i.e. the scene is not entirely the same colour;
- images of scenes with surface geometric discontinuities, i.e. in which the scene contains multiple surfaces either abutting or not.

In fact, it is quite common to encounter such images, and it is quite rare to encounter the classical ones. Apart from contrived settings such as those found in a research laboratory, we seldom come across uniformly coloured, uniformly illuminated scenes consisting of a single smooth surface.



FIGURE 1.2. This image provides an example in which the scene is not illuminated from a single point source at infinity. There are singularities (maximally bright points) on the front of the nose and on the back of the neck. The classical shape from shading setting which allows for a unique light source direction, thus infers that the nose and the back of the neck are facing the same direction — an obvious fault.

The classical assumptions encounter real difficulties when evaluated against natural images. To illustrate the nature of one of these difficulties, consider the image in Fig. 1.2. One of the most common assumptions in shape from shading is that the illumination is constant over the entire scene, which generally implies that there is only one effective light source. This is not the case for the scene in our figure. Notice that the front and the back of the head are the brightest areas in the image. The classical assumptions of a unique light

source and a Lambertian model of reflectance imply that the brightest areas are facing the light source. For this example, it would mean that both the front and the back of the head are facing the same direction. This is an obvious error.

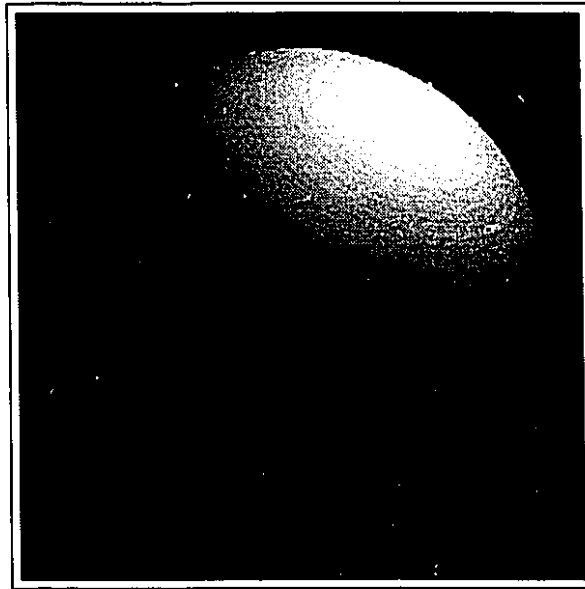


FIGURE 1.3. This image provides another example in which the scene is illuminated from multiple distant point sources.

The scene does not need to be very complex before such difficulties arise. Figure 1.3 shows a very simple example that captures the essence of the difficulty. The viewer can only see half of the sphere's surface from a given viewpoint. Similarly, any given point light source can only illuminate half of the sphere's surface. Unless the sphere, the viewer and the light source are aligned, some visible part of the sphere will not be illuminated by the given light source.

Although the sphere seems like an extremely simple example, this discussion illustrates how it is sufficiently complex to demonstrate many subtleties in shape from shading. We shall use the sphere example throughout this thesis.

Figure 1.4 shows an example of an image of a scene in which shadows are observed. In the image, the shadow is simply a black zone. It is not possible to infer the shape of the surface in that zone. Unless this fact is acknowledged, the algorithm will yield erroneous information.



FIGURE 1.4. This image provides another example in which the scene is illuminated from a single distant point source. Here, there are parts of the surface which are visible but are in a shadow.

1. Overview

In this thesis, we define a generalized shape from shading problem which takes discontinuities into account — lighting discontinuities, reflectance discontinuities, and surface discontinuities. The added complexity leads to a new approach that relies on geometric structures inferred from the image, rather than on the image itself. This observation marks a definite break with previous work based on the photometric properties of the image.

Our shape inference process comprises three steps:

- (i) The first step is to extract local geometrical structures from the image. We rely on the tangent field (information about contours) and the shading flow field (information about smooth scenes) as input for the shape inference (Fig. 1.5).
- (ii) The second step is to find local scene models that can account for the local geometric structures within the image (Fig. 1.6). Once this is done, at any given pixel location, there exist several candidate local scene models (Fig. 1.7).
- (iii) We then consider how the different models at different locations fit together (Fig. 1.8), and study local constraints to find consistent surfaces. Such local constraints seek pairs of local scene models that can describe the same smooth surface (Fig. 1.9). The third and last step is thus to find among these possible local scene models, global structures for which the local scene models are locally consistent.

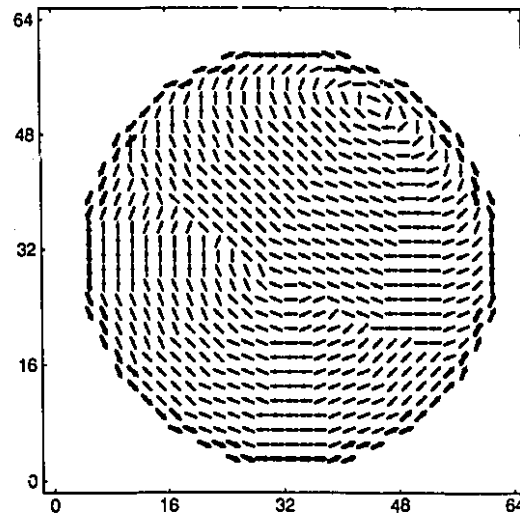


FIGURE 1.5. The shading flow field is depicted by thin arrows and the tangent field, by thick arrows. These geometrical structures correspond to the image shown in Fig. 1.3.

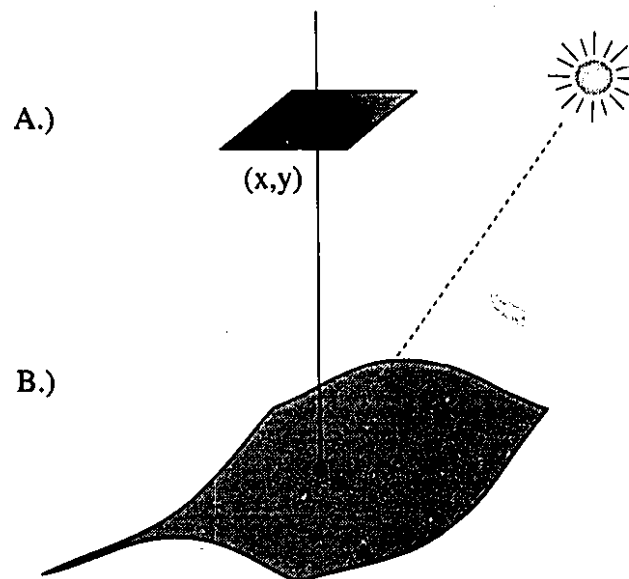


FIGURE 1.6. The local scene model (b) has to account for the local image geometric structures (a).

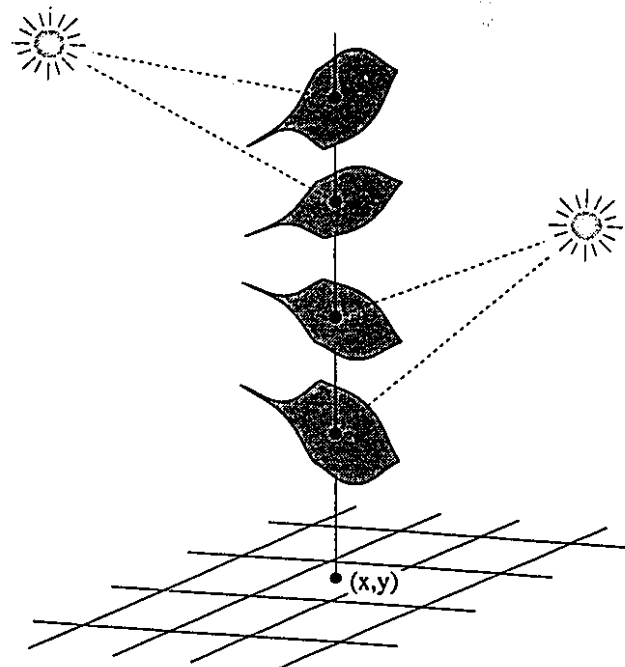


FIGURE 1.7. For a given image position, a collection of possible scene models can account for the observed local image geometric structures.

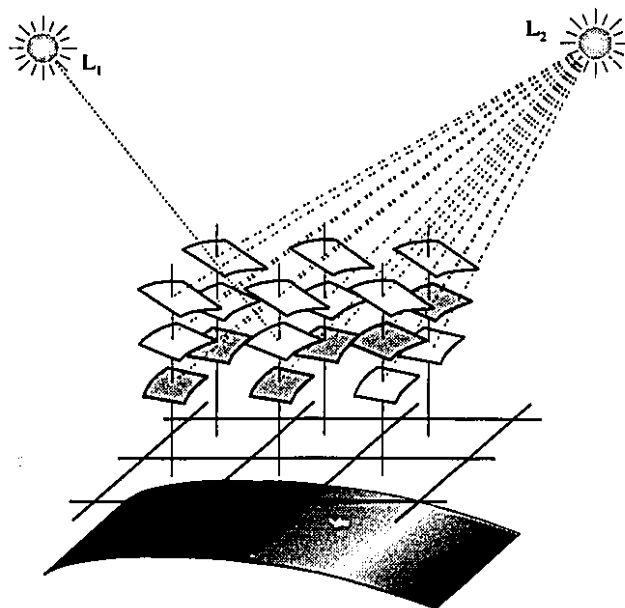


FIGURE 1.8. A local scene model also has to be consistent with the neighbouring local scene model.

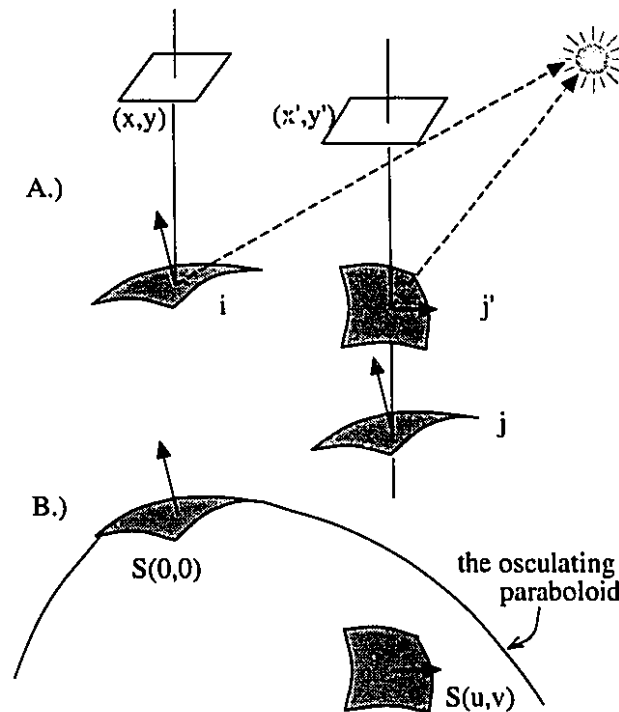


FIGURE 1.9. The lighting and surface continuity properties introduce constraints on the local scene models.

2. Claims of Originality

In this thesis, we define a new generalized shape from shading problem, and study its solution.

- We show that, to infer shape from shading under the condition that albedo and lighting are variable, *photometric* structures of the image cannot be relied upon and we thus exploit the *geometric* structures of the images — the shading flow field and the image curves.
- We have built a new local representation of the surface orientation and shape for which *continuous* transformations of the surface are mapped to *continuous* curves in a five-dimensional space. This representation, along with an illumination model, allows us to build a local scene model which is rich enough to treat the generalized shape from shading problem.
- We have established the relationship between local descriptors of the scene and local geometric properties of the images i.e. the shading flow field and the edge map. Local differential properties of the image are used since the image irradiance values are not

sufficient to resolve the ambiguity created by the possibility of multiple illumination sources and albedos.

In this thesis, we propose a new computational framework to infer the shape of the surfaces of these more complicated scenes. In the process of developing this framework, several results are worth mentioning as they are either an integral part of the computational framework or serve to justify choices.

- We have established a compatibility relationship between neighbouring local scene models based on differential geometric properties of surfaces and continuity properties of light sources.
- We have defined the shading flow field as the dual of the gradient field, and present a new approach to extract these flow fields from an intensity image.
- We have studied topological and geometrical properties of the shading flow field and of the scene bundle in order to provide a theoretical rationale for our framework.
- Since we consider shape from shading as an entirely geometrically driven process (we consider that shape from shading has to rely on shading flows instead of image irradiance intensities), we provide psychophysical evidence in support of the idea that the primate brain does not process scalar (image) and vector properties (geometric structures) in the same way.
- We show the relationship between different types of discontinuities in the image and different types of discontinuities in the scene as an asset to interpret the image.
- We also present a computer graphics implementation of the radiosity equation that takes interaction with the ambient media into account. The novelty here lies on the massively parallel nature of the implementation.

CHAPTER 2

Overview of the Vision Problem

This chapter only contains background material intended for a naive reader. It is meant to provide an overview of the vision problem and to show where the topic of this thesis fits into “the” big picture. Anyone familiar with this research field should move on to Chapt. 3.

1. The General Vision Problem

The goal of a vision system is to provide useful information about the environment from photosensitive devices. Evolution provided animals with such visual systems for their specific needs. There is a wide range of different visual systems which cater to a wide range of different behaviours. Some are very primitive, such as the barnacle’s. It only contains a few photosensitive cells which cause it to retreat in its shell when a sudden illumination decrease occurs. Some visual systems are very developed, such as the human visual system. It comprises: two eyes facing in the same direction, thus allowing stereopsis; different photo-receptors, allowing both colour vision and low intensity vision; a fovea, allowing greater resolution where attention is needed; and roughly a third of the brain’s cells, allowing for a lot of processing power. Between these extremes, we can note that some animals cannot discriminate colors (e.g. owl monkeys), some animals have eyes on opposite sides of the head giving them a large field of view but little or no stereopsis (e.g. pigeons). Some animals have even developed special purpose visual devices such as the fly’s ultra-violet sensors, presumably for orientation.

Since vision is often the most important source of information for humans, it is not surprising that we want to give that faculty to machines. Automation of many tasks in industry requires information about the environment in which the machine works. An ambulatory robot must sense its environment to be autonomous, especially when the environment is

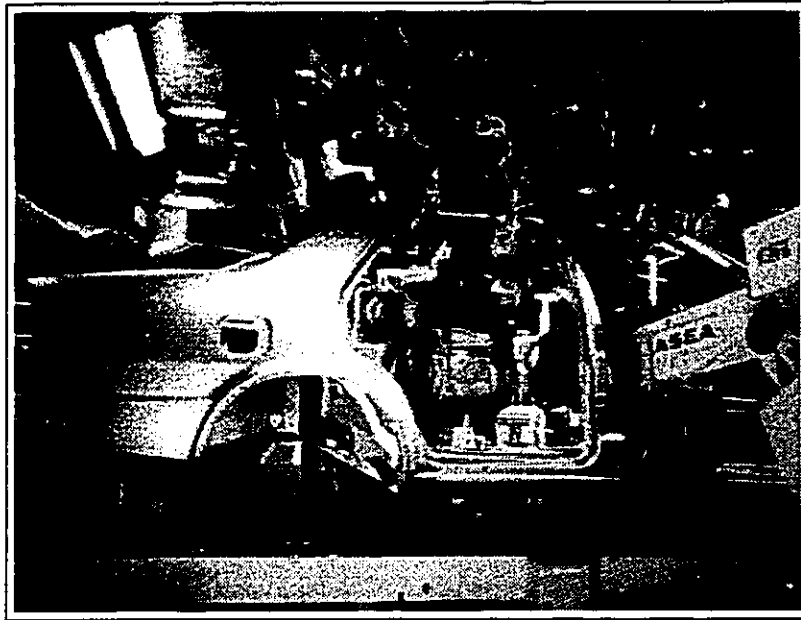


FIGURE 2.10. This image provides an example of the data that the machine would acquire from a sensor. This image is composed of 471×358 pixels with the intensity of each pixel coded with one byte.

hazardous. Biomedical imaging and remote sensing creates such a huge quantity of data that the automation of the interpretation process also becomes desirable.

The problem setting can be described generally as follows. Light interacts with the environment and is captured by sensors. Three-dimensional structures in the environment project onto two-dimensional structures in the image. The problem of vision is to infer useful information about the environment from the output of these sensors. Somehow, the two-dimensional image structures must support inferences about the three-dimensional structures in the environment. Because this inference process is under-determined, some prior knowledge of the image formation process and of the interaction between the light and the environment is needed to interpret the data.

1.1. Data and Knowledge. To solve the vision problem, there are two types of information: data and knowledge.

Data is the information provided by the sensors. For humans, the sensors are the eyes. For machines, the sensors are usually cameras. The information is usually represented by images or two-dimensional arrays of numbers. The image represents the activity of photosensitive cells in the retina or the response of CCD-elements of the camera — scalar quantities. It also provides a neighbourhood structure that is related to the neighbouring relationship of the incoming light rays.

Knowledge is the information about the environment that the system has acquired prior to data acquisition. This knowledge can comprise detailed descriptions of various objects of interest — e.g. the components of an automated production chain. It can include details of the relationship between objects and their context (i.e., that telephones are often found on tables) and so on.

It also provides the tools needed to begin an understanding of image intensities. This includes knowledge of physics. For example:

- spatially, the environment is in a three-dimensional space;
- light travels in straight lines in the void;
- light interacts with matter...

Knowledge can also be specific to the vision system used. For example, the following knowledge base could describe a particular vision system:

- the vision system incorporates two sensors;
- each sensor comprises an optical system and a photosensitive surface;
- the properties of the optics in front of the photosensitive surfaces provide the information about how the various rays coming from different points in the scene will project onto the photosensitive surfaces;
- these surfaces are composed of photosensitive cells arranged in some irregular foveated tessellation;
- the image intensity values correspond to the response of the respective cells;
- the response of each cell depends on its particular photometric and spectral response functions...

Knowledge can also depend on the expected scenery. The algorithm presented in Chapt. 3 is based on a set of assumptions. Some of these assumptions are:

- the visible surface is continuous;
- the visible surface has a uniform matte colour, and is uniformly illuminated;
- all visible surface patches are illuminated from the same direction, and this direction is known;
- the brightest spot in the shading corresponds to a surface patch directly facing the light source...

When this knowledge is rich but specific, the vision problem can be easier to solve, but the solution may have only a limited interest. On the other hand, when this knowledge is general but limited, the solution to the vision problem will be more widely applicable, but it may also be much more difficult to reach.

While we recognize that stereo-vision is a powerful cue for depth perception, we ask the question: what happens when we close one eye?

In this thesis, we will consider a vision system with only a single CCD-camera and we will only consider static images¹. Since humans are able to extract a lot of information from such an image, we will explore the ways that a machine could do it. We will assume that the photosensitive surface of our sensor is composed of regular square cells all having the same response function. We will also assume that the image is obtained by an orthographic projection. Such an image corresponds to the set of parallel incoming rays from the scene that are perpendicular to the image plane.

1.2. Representations of the Image and the Scene. The solution of the vision problem can take various forms. We have said that the problem of vision is to infer some useful information about the environment from the output of the sensors. The notion of usefulness depends on the task considered. A machine whose task is to turn on the artificial light on a highway at dusk will find the simplest representation useful, whereas an autonomous vehicle on Mars will need a much richer description. There is in fact a hierarchy of such representations of the useful information implicit in a scene.

The lowest representation of the scene information is the intensity image. It is provided directly by the sensors. The intensity image gives us information on the amount of light that emanates from given directions. It can also contain some information about the spectral composition of the light.

Since we are considering only single static images, the next representation could be the local structures that can be found in the intensity image. There are two types of features that have been identified based on the dimensionality of their image support [136].

- Type I. Image curves: edges and lines. These features are locally characterized by a contrast, an orientation and a curvature.
- Type II. Vector and direction fields: texture flow fields and shading flow fields. These features are locally characterized by an orientation and higher order descriptors. For texture flow fields, these higher order descriptors could be two curvatures [66]: the curvature of the flow field itself and the curvature of its dual. For shading flow fields, these higher order descriptors could also include the curl of the vector field,

¹We briefly mention the works of other researchers interested in shading analysis but on different types of vision systems. Some have explored the possibilities of combining stereopsis and shading [10, 18, 61, 129], motion and shading [2, 108], and multiple images and shading [82, 130]. This list is not meant to be exhaustive but it should provide some pointers to anyone who wishes to pursue these directions.

its deformation and the axis of deformation. Shading flow fields are studied in Chapt. 4 of this thesis.

The next representation could be a local description of the scene. The scene is described with respect to the image coordinate system, and might comprise two models:

- A local surface model. It has two components:
 - first, a geometric component, described by a position and the surface patch's orientation and shape;
 - second, an appearance component, described by a colour (intensity, hue, saturation), a visual texture pattern, and perhaps a reflectivity function.
- A local lighting model. The lighting model is usually composed of the intensity and direction of the virtual illuminant seen by the surface patch. A richer model could include the incoming intensity of the rays from every direction.

The next representation might be a global description. The objects in the scene are segmented and described globally in term of outlines (snakes), surfaces (thin sheets), or volumetric primitives (spheres, cylinders, super-quadrics, etc.).

Finally, the highest representation could be the identity of objects in the scene, along with a description of how they relate to each other. This level is comparable to the description of a scene that a person might make in order to communicate with another person.

2. Elements of the Solution

The existence of a solution to the vision problem is provided by biology since evolution has succeeded in implementing one for human needs. From an engineering point of view, it can be advantageous to examine such working examples. Psychologists and neurophysiologists have been studying the visual system of different animals for a long time. Hence throughout our attempts to find a solution to the vision problem, we were inspired by some of their findings.

In the next section, we will mention one early work in this field — Helmholtz's distinction between low-level and high-level vision. Then, in the following section, we will look at the approaches that were used in the attempt to solve the vision problem.

2.1. Low-level and High-level Vision. Heinrich von Helmholtz (1821 – 1894) [125] distinguished two levels of visual processing: the low-level and the high-level. These are sometimes also referred to as early processing and later processing. Early processing refers to what is going on from the retina to the cortex, whereas later processing refers to what is going on in the later stages of cortical processing.

DEFINITION 2.1. *Low-level vision is dominated by physics and physical models of the environment. The problem is one of finding general constraints on special-purpose hardware. Thus, if something is understood about structure of the image or the scene, then something can be inferred about functions in the visual system [137].*

DEFINITION 2.2. *High-level vision is dominated by reasoning. The problem is one of finding specific constraints on general-purpose hardware. Thus, if something is understood about function, then something can be hypothesized about structures in the scene or in the image [137].*

When Roberts [115] described what is probably the first artificial vision system, he too decomposed it into two stages: a low-level stage in which image curves were obtained and a high-level stage in which the image curves were matched against a model knowledge-base of polyhedra — the blocks world. Even though Roberts severely restricted the world of possible objects, matching remained a difficult problem.

As we attempt to create artificial vision systems to deal with a much richer world, it is necessary to infer more intermediate representations before matching the low-level results with a model base. The classical problem of “shape from shading”, which we will discuss in Chapt. 3, is an example of an attempt to enrich the representation at an early processing stage. It provides a clear example of how physics and physical models of the environment can be used as general constraints to solve a low-level vision problem.

2.2. Data-driven and Knowledge-based Approaches. We said earlier that there are two types of information useful for solving the vision problem: data and knowledge. Two approaches have been attempted, each basically driven by one type of information.

The data-driven approach builds on the data. It infers more structured representations from the data. It thus increases the database. Typically, one incorporates general constraints following the principle of least commitment that Marr imported to computer vision [93]. Nevertheless, these constraints are assumptions that will require careful analysis before they can be accepted as a fact consistent within the resulting representations.

The knowledge-based approach seeks answers to questions. Detailed hypotheses are inferred from the knowledge, and conclusions are confirmed by the database. One strategy consists in orienting the search through the model base by using contextual information. Ultimately, the approach seeks to match a model with the data or the representations inferred from it.

These two approaches have the potential to complement each other. As we just said, the knowledge-based approach seeks to match models to the data. This matching is a very difficult step when the model base is rich and the data limited to intensity images. For some applications, it is possible to limit the model base enough to succeed in solving a particular problem (see the pattern recognition literature). Unfortunately, in most potential applications, it is not that simple. An alternative is to enrich the data with more structured representations. This is precisely what a data-driven approach provides. Each structured representation comes with a set of assumptions. Whereas these assumptions should hold most of the time for the lower levels of a structured representation, at the higher levels, the assumptions can be more context dependent and hence, less reliable. The problem becomes one of assessing the reliability of the results. This is the type of problem that is well-suited for a knowledge-based approach.

In this thesis, we scrutinize the "shape from shading" problem. It is a problem that is now generally considered to be an early vision problem. It is dominated by the physics of the interaction of light with matter and by physical models of the reflecting surface. General constraints on the interaction of light with matter and on the reflecting surface are used to guide the inference of a local representation of the surface shape.

In Chapt. 3 and in Parts 2 and 3, different data-driven approaches to deal with the shape from shading problem are presented. These approaches differ significantly in their underlying assumptions and thus lead to very different algorithms. It is our belief that the new approach, based on shading flow fields and scene bundles (described in Parts 2 and 3) is a significant advance over the classical formulation (presented next, for comparison and introduction).

CHAPTER 3

The Classical Shape from Shading Problem

Shape from shading is a classical problem in computer vision research. It is usually associated with a particular setting: a single distant point light source uniformly illuminates a smooth surface of constant albedo according to Lambert's reflection law. Throughout this thesis, we refer to this setting as *classical shape from shading*.

In this chapter, a brief history of the shape from shading problem is presented. The classical shape from shading problem is described, together with the various attempts to solve it. The classical setting introduces constraints that allow fast solutions to the shape from shading problem. But we show that it also severely limits the domain of application. This chapter is intended to provide the background for our new approach.

1. Brief History

The quest to determine the shape of the surface features in the maria of the moon [26, 114] is at the origin of recent work on shading analysis [50, 54]. In those days, the only available images of the moon were obtained from telescopes on earth. Since the moon is always more or less presenting the same face to earth observers, motion or stereo vision algorithms could not provide precise estimates of its topography. Even if images were taken from distant points on the earth, these algorithms could not provide precise estimates of features on the moon — the ratio of the radius of the earth to the earth-moon distance is too small. Thus, other cues received more attention. Photometric stereo algorithms [61, 129, 130] rely on images of a scene taken under identical viewing conditions but different lighting conditions. Many shape from shading algorithms [51] rely on a single image but assume that the reflectance function is known.

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

The shape from shading problem is classical in vision; Ernst Mach (1866) was perhaps the first to establish a formal relationship between image and scene domains, and to capture their inter-relationships in a partial differential equation [91,112].

Consider a sufficiently smooth surface¹ characterized by a depth map $z(x, y)$ where (x, y) are the image coordinates. Let

$$(p, q) \equiv \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right)$$

denote the gradient of the surface, so that the surface normal field

$$\mathbf{N}(x, y) = \frac{(p, q, -1)}{\sqrt{p^2 + q^2 + 1}}.$$

For a matte surface (modelled by Lambert's reflectance function) which is illuminated by a distant point light source (e.g. the sun) from direction $\mathbf{L}$, the image irradiance $E(x, y)$ can be modelled as

$$(3.1) \quad E(x, y) = \mathbf{N}(x, y) \cdot \mathbf{L}.$$

Mach assumed that the surface could be obtained by integrating Eq. 3.1. It took quite a while before someone picked up that challenging problem.

Horn set the modern approach by focusing on the solution of these partial differential equations by classical and numerical techniques [51, 52, 56, 62]. In his early attempts to solve the shape from shading problem, Horn [51] introduced the *image illumination equation*:

$$E(x, y) = F(x, y, z, p, q).$$

This equation is often referred to as the *image irradiance equation* (see for example [15]). This general formulation states that the image irradiance for a given position (x, y) depends on the position in space (x, y, z) of the corresponding surface point and on the surface gradient (p, q) . The explicit dependency on surface position allows for a dependency on material properties as well. Furthermore, it allows the modelling of a nearby light source for which the distance between the source and the surface patch is crucial.

Since the general form of the image irradiance equation did not yield practical numerical solutions, the dependence of $F(x, y, z, p, q)$ on position was dropped [53] thereby reducing the equation to the more simple form:

$$E(x, y) = F(p, q).$$

¹By smooth here, we mean the function is differentiable as many times as needed.

The basic assumption here is that all variations in the image irradiance are due to variations in the surface gradient, thus to variations in surface orientation.

Horn's early attempts to solve this problem yielded promising results and others have built upon it. In the following sections, we provide the classical definition of the problem and then we discuss some of the solutions proposed.

2. Classical Definition of the Shape from Shading Problem

We take the classical setting in computer vision for shape from shading to be the following: a point light source at infinity uniformly illuminates a smooth matte surface of constant albedo whose image is formed by orthographic projection.

PROBLEM 3.1. CLASSICAL SHAPE FROM SHADING

Assuming that

- (i) *the scene has the same reflectivity everywhere;*
- (ii) *the scene is illuminated by a single distant point light source;*
- (iii) *the scene is composed of a single smooth surface;*
- (iv) *the surface is matte;*
- (v) *the image is formed by orthographic projection.*

Given

- (i) *the albedo ρ ;*
- (ii) *the illumination λ ;*
- (iii) *the direction of the light source $\mathbf{L}$;*
- (iv) *the image irradiance $E(x, y)$.*

Recover

- (i) *the surface shape.*

The matte surface is traditionally modeled with Lambert's reflectance function [78] so the image irradiance equation becomes

$$(3.2) \quad I(x, y) = \rho \lambda \mathbf{L} \cdot \mathbf{N}(x, y)$$

where $I(x, y)$ is the image intensity² at a point (x, y) ; ρ is the albedo of the surface, i.e. the fraction of the shining light which is reflected; λ is the illumination, i.e. the amount of shining light; $\mathbf{L}$ is the light source direction; and $\mathbf{N}(x, y)$ is the normal at the surface point

²The image irradiance $E(x, y)$ is often replaced by the image intensity $I(x, y)$. One assumes that the relation between the image irradiance $E(x, y)$ and the image intensity $I(x, y)$ is linear. One further assumes that the highest image intensity value corresponds to a surface facing the light source so that one can use the image intensity normalized by the highest value, and take ρ and $\mathbf{L}$ being given the unit value.

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

corresponding to an image point (x, y) . The physical model leading to this equation has recently been reviewed by Nayar et al. [98].

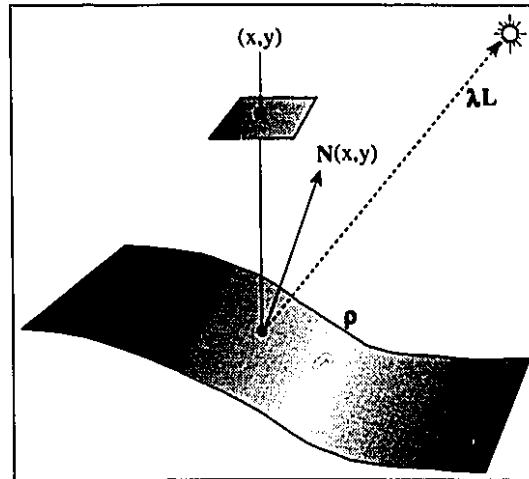


FIGURE 3.11. A continuous surface, described locally by its orientation N , is illuminated from a direction L . The image of this scene is obtained by an orthographic projection. We use a coordinate system in which the image plane defines the x - y plane and depth is the distance from the image plane along the z axis. Two scalar quantities are important for the understanding of matte reflection: the albedo ρ and the illumination λ .

Fundamental to the classical shape from shading problem is the following implicit basic assumption:

ASSUMPTION 3.1 (HORN). *Variations in image intensity are entirely due to variations in surface orientation.*

Researchers followed two basic approaches in their attempts to solve the problem of recovering shape from shading. The first approach assumes that the model is correct and proceeds directly to compute the exact depth map $z(x, y)$. The second approach assumes that there are errors in the model and therefore introduces additional constraints to recover shape.

Horn's early attempts to solve this problem lie at the heart of the first approach. He realized that along characteristic strips [41], the image irradiance equation could be

rewritten in terms of five ordinary differential equations [51, 52]:

$$\begin{aligned}\frac{dx}{ds} &= \frac{\partial F}{\partial p} & \frac{dy}{ds} &= \frac{\partial F}{\partial q} & \frac{dz}{ds} &= p \frac{\partial F}{\partial p} + q \frac{\partial F}{\partial q} \\ \frac{dp}{ds} &= \frac{\partial E}{\partial x} & \frac{dq}{ds} &= \frac{\partial E}{\partial y}\end{aligned}$$

where s is a parameter which varies with distance along a characteristic strip. When the surface normal is known at a pixel (x_0, y_0) , then one can integrate the differential equations to obtain a curve on the surface. In order to get an initial estimation of the normal, Horn used the singular points. At the maxima of intensity, the normal vector $\mathbf{N}$ and the vector pointing toward the light source $\mathbf{L}$ are equal.

Several authors studied the conditions under which these known points uniquely constrain the solution and showed that recovering shape from shading is not always ill-posed [15, 16, 99, 100, 118]. Elegant solutions have since been recently proposed [6, 86, 101]. Unfortunately, the characteristic strip method remains numerically unstable.

The second approach is precisely concerned with issues such as numerical stability. The susceptibility to noise is at the root of the need for some additional constraint. There are two main classes of algorithms: local and global algorithms.

Global algorithms, or variational approaches, were developed to extract estimates of surface orientation with an additional assumption about surface shape — the surface is smooth in some sense. The smoothness assumption is used to relate adjoining points [13]. It enables spatially isolated information about absolute surface orientation to be iteratively propagated across the surface. Such iterative methods were implemented on a regular grid, the integral surface being defined by either a C^2 function [14, 56, 121] or a C^1 function [14, 56, 62, 88]. Multi-grid implementations were also proposed [122, 123] to alleviate the computational burden.

The issue of integrability has also received attention [37, 55, 56] — integrability can be enforced when both the surface height and the gradient of surface height function are represented. It provides a powerful constraint on the surface function.

A class of local algorithms were also developed [31, 68, 87, 106] at about the same time. In order to extract estimates of surface orientation, again an assumption about surface shape was made — the local surface curvature constrains the surface. This local curvature assumption relates information about absolute surface orientation within a small image neighbourhood. Whereas Frankot and Chellappa's algorithm enforced integrability, in a similar spirit, Ferrie and Lagarde [30] considered how to enforce curvature consistency.

From their comparative study of shading techniques, they concluded that it stabilizes and improves local estimates of shape from shading.

Other types of algorithms have also been considered. Among others, we briefly mention the existence of Pentland's algorithm [107] using an assumption about surface reflectance.

It is clear that the "classical" shape from shading problem has received a lot of attention. In the next section, we will examine the nature of the solution and how it behaves in situations for which the basic model is not appropriate.

3. Analysis of a "Classical" Shape from Shading Algorithm

Recently, ways to obtain the solution of the "classical" shape from shading problem were published [6, 86, 99-101, 135]. It is reported that these algorithms provide solutions that are reasonably good when the usual assumptions are met. We have implemented Bichsel and Pentland's algorithm [6] to illustrate the solution to the classical shape from shading problem³.

3.1. Overview of Bichsel and Pentland's Algorithm. Bichsel and Pentland's (B&P's) algorithm takes gray level images as input. The singular points of an image also need to be identified, as they play a key role by introducing a three-fold ambiguity (locally convex, concave, or saddle) to the solution. Since these points correspond to the points of maximum brightness, Bichsel and Pentland assume that this identification task should be straightforward [6].

B&P's algorithm uses a "minimum downhill principle" to remove this ambiguity. The downhill principle consists in passing surface information only to pixels which are more distant from the light source. For the minimum downhill principle, one chooses among different possible paths, the path that leads the least away from the light source. Bichsel and Pentland claim that this "minimum downhill principle" guarantees the convergence of their algorithm.

The solution of the shape from shading problem is a description of the shape of the scene that generated the image. For this algorithm, the description consists of a depth map. It is obtained by an iterative scheme in which the local surface patch depth is updated according to the minimum downhill principle.

3.2. Solution to the "Classical" Shape from Shading Problem. Figure 3.12(a) shows a test image used to analyse B&P's algorithm. The scene consists of a sphere in front

³The two key subroutines were copied directly from their conference paper. However, it was necessary to write code for the input/output as well as the calls for these subroutines.

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

of a plane, both matte surfaces. The scene is illuminated from a distant point which lies directly on the line of sight, in mathematical terms: $L = (0, 0, 1)$. The shading is generated artificially using Lambert's reflectance function. In Fig. 3.12(b), we show the corresponding depth map obtained by the algorithm. All the algorithm's underlying assumptions are satisfied for this example, and the scene obtained by this algorithm should resemble the original data when illuminated from the given direction.

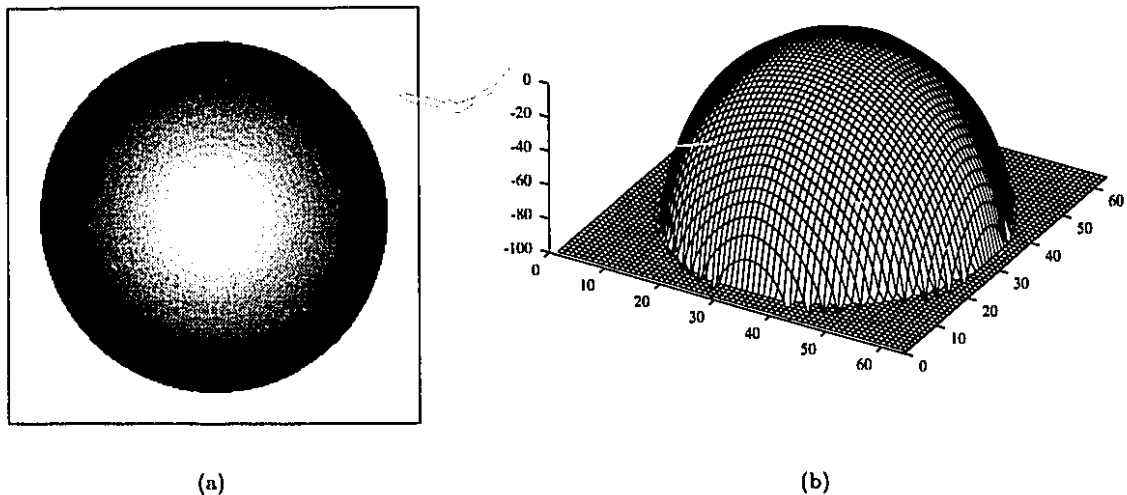


FIGURE 3.12. This very simple test image provides an example of a "successful" shape recovery. We implemented an algorithm using the classical shape from shading assumptions. An ideal intensity image of a constant albedo matte scene illuminated by a single distant point light source (a) and the corresponding recovered depth map (b) are shown in this figure. The algorithm is from Bichsel and Pentland [6]. The code of the two main subroutines came directly from their conference paper; only the input/output and the call functions inside an iteration loop had to be re-coded.

OBSERVATION 3.1. *The depth map from B&P's algorithm captures accurately the shape of the sphere and the plane. However, it is important to note that although the algorithm returned a "good" solution, other "good" solutions are also possible. In the original scene, the depth of the plane is largely arbitrary. Furthermore, the same image could have been generated with the plane in other orientations if the sphere and the plane were not constrained to be of the same reflectivity.*

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

Identifying the singular points is not a straightforward task when we deal with noisy discrete data. Singular points may not show as global maximum in the image domain. Most local maxima in the image domain do not correspond to a singular points; they are simply the result of noise.

In Fig. 3.13(a), we present another test image. The scene is composed of three identical damped anisotropic waves slightly interfering. Again the scene is illuminated from a distant point which lies directly on the line of sight and the shading is generated artificially using Lambert's reflectance function. Because of the discretization of the image, a surface point directly facing the light may not be represented in the image by a global maximum. The depth map obtained with B&P's algorithm is shown in Fig. 3.13(b). This example reveals that B&P's algorithm is quite dependent on the accuracy of the singularity detection.

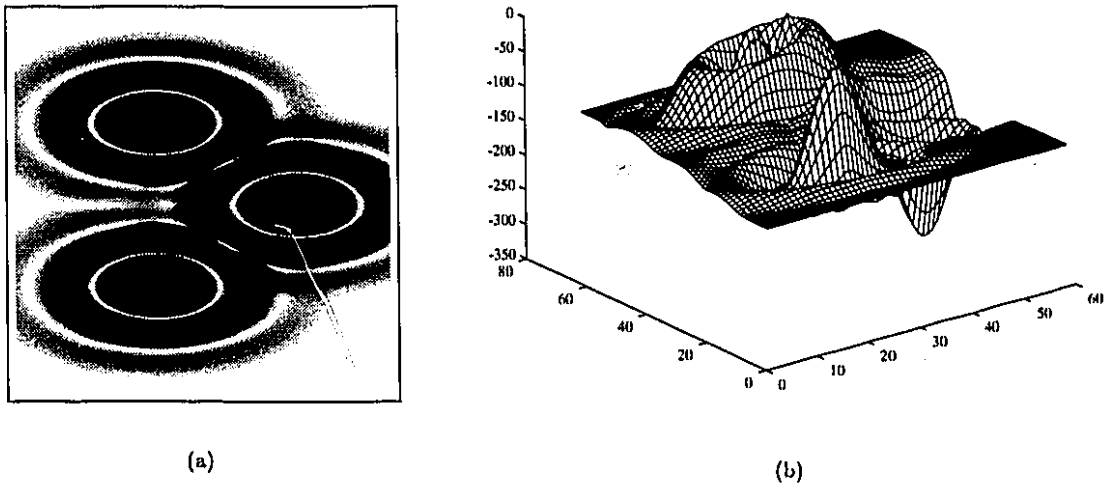


FIGURE 3.13. The image of an undulated constant albedo matte surface illuminated by a single distant point light source (a) and the corresponding recovered depth map (b) are shown in this figure. This depth map is the output of B&P algorithm using the brightest pixels in the image as singularities. The surface should correspond to the interference pattern of three identical damped anisotropic waves. The recovered surface does not fit this description. This test image provides an example of an inaccurate shape recovery that occurs when the initial data are not reliable. Here, the singularities are not all well identified. Since the image is a discrete function, singularities for the continuous image irradiance function do not necessarily appear as the brightest point in the image.

In Fig. 3.14(a), we present a third test image. The scene consists of a mask in front of a flat background. The scene was originally captured by a laser range finder and the shading

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

was obtained by considering a light source at a distant point lying directly on the line of sight. In this image, since noise introduces local maxima, it is even more difficult to detect the singularities that are key for the shape recovery. The depth map obtained with B&P's algorithm is shown in Fig. 3.14(b). Our best guess leads to a reasonable result⁴.

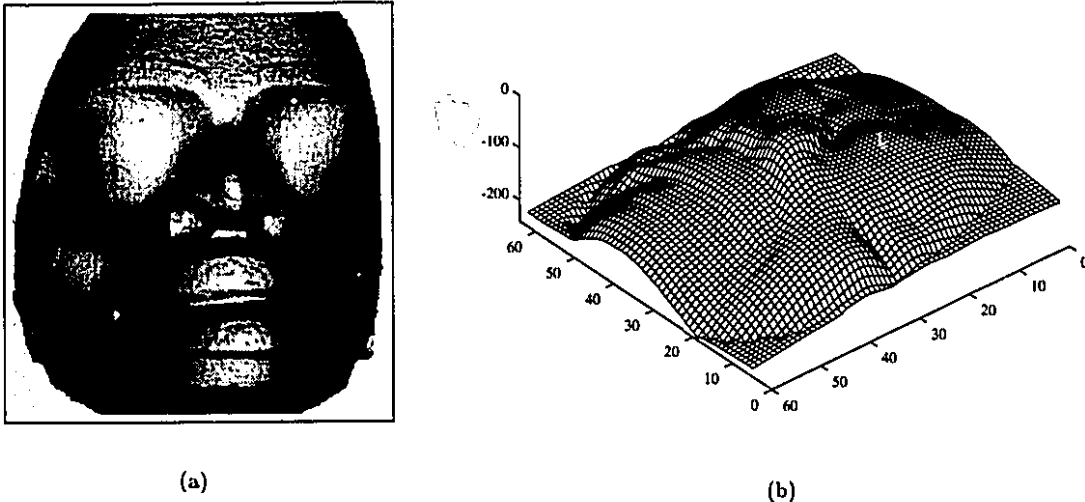


FIGURE 3.14. In this figure, we show an intensity image (a) that was made by artificially shading the range data image of a mask, and the corresponding recovered depth map (b). We can note here that B&P algorithm did not accurately recover every detail of the shape of the mask. The shape recovery was based on singularities that were perceptually salient, and these were identified manually. The emphasis in this figure is not on the performance of the algorithm, but on the nature of the initial data. Here, finding the singularities is not straightforward because of the noise. Using only global maxima likely misses important singularities, whereas using local maxima introduces a lot of insignificant singularities.

OBSERVATION 3.2. *In order to recover the shape of surfaces in the scene, the initial data must be reliably available.*

From these three examples, we can at least conclude that the use of B&P's algorithm is interesting for situations over which we have complete control. Thus, we will use this algorithm in the following section to show the weaknesses of the classical shape from shading assumptions. In their very recent comparative study of the various shape from shading techniques, Zhang et al. [135] have also implemented B&P's algorithm. They have run the

⁴For this example, the singularities are not determined by the brightest pixels, but are instead chosen based on our visual perception.

algorithm on images for which the classical setting is usually appropriate and have also acknowledged the accuracy and drawbacks of the method. Given the correct singularity information, B&P's algorithm provides typical results for the "classical" shape from shading setting.

3.3. Limitations and Difficulties. The literature cited in this chapter describes the various attempts to solve the classical shape from shading problem from first principles. We emphasize, however, that, to make these approaches tractable, certain parameters are assumed known. Typically the surface albedo ρ , the illumination λ and the direction from which it comes L are presumed to be constant and known. Operationally this assumption decouples problems; e.g., it decouples the shape from shading problem from light source estimation problems [105]. Although this decoupling simplifies significantly the shape from shading problem, it also limits severely the range of its applicability. The result of any algorithm based on such assumptions could be misleading.

In Fig. 3.15(a), we show an image of a sphere illuminated from two different directions. Parts of the sphere are illuminated by only one source and the central part is illuminated by both light sources. As previously, the shading is generated artificially using Lambert's reflectance function. The depth map obtained with B&P's algorithm is shown in Fig. 3.15(b). We can observe a deformation of the sphere.

OBSERVATION 3.3. *An algorithm designed to solve the "classical" shape from shading problem fails to accurately recover the shape of a scene if shadows are present.*

The error is due to the fact that the "classical" shape from shading algorithm assumes that the surface is illuminated from the wrong direction on parts of the image. Objects create shadows and an object which lies in the shadow is illuminated differently. Since different surface patches are often illuminated differently, this type of error is likely to occur in general.

In Fig. 3.16(a), we show an image of the sphere in front of a plane but part of the plane and sphere are in a darker colour than the rest. Again the scene is illuminated from a distant point which lies directly on the line of sight and the shading is generated artificially using Lambert's reflectance function. The depth map obtained with B&P's algorithm is shown in Fig. 3.16(b). We can observe a deformation of the sphere.

OBSERVATION 3.4. *An algorithm designed to solve the "classical" shape from shading problem fails to accurately recover the shape of a scene if it does not have a constant albedo.*

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

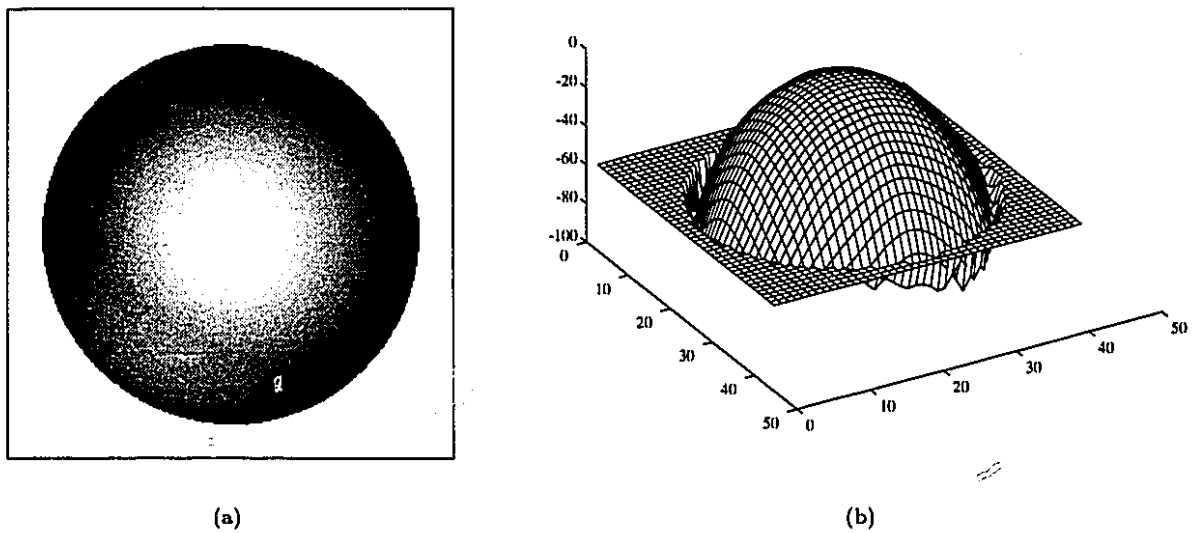


FIGURE 3.15. To analyse the limitation of a shape from shading algorithm based on the classical assumptions, this test image provides an example of shape recovery for a simple scene in which two light sources are present. Parts of the sphere are illuminated with only one light source — the upper right and the lower left parts — and the central part of the sphere and the back plane are illuminated by both light sources. The intensity image (a) and the corresponding recovered depth map from B&P's algorithm (b) are shown in this figure. Note the "valley" introduced by the algorithm to account for the shading variability around the edge.

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

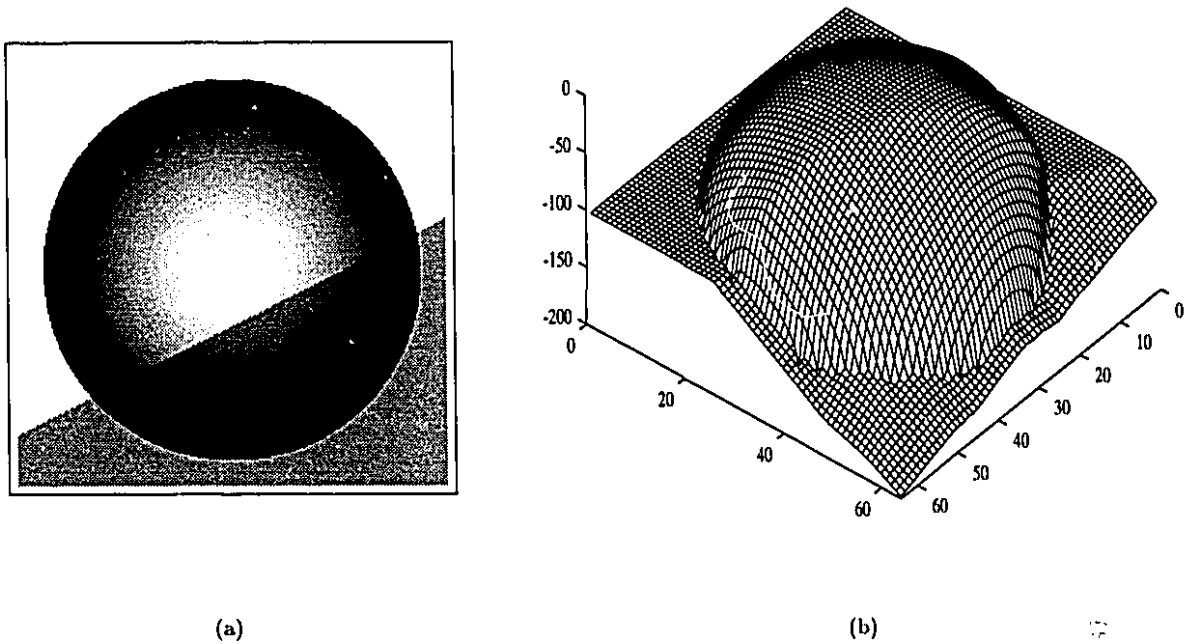


FIGURE 3.16. This test image provides an example of shape recovery attempt with B&P's algorithm for a simple scene in which the surface albedo changes abruptly — the scene is not entirely the same colour. The intensity image (a) and the corresponding recovered depth map (b) are shown in this figure. This result shows clearly that the algorithm based on the classical shape from shading assumptions fails to accurately recover a multiple albedo scene.

3. THE CLASSICAL SHAPE FROM SHADING PROBLEM

The error is due to the fact that, on a certain part of the image, the algorithm wrongly assumes the surface reflectance coefficient. Since the world is not entirely the same colour, this type of error is likely to occur in general.

Although promising results for the classical shape from shading problem have been obtained by numerous researchers, problems remain which are not naturally treated in the classical sense, especially those related to discontinuities and shadows. We emphasize again that in order to make the shape from shading problem tractable, certain parameters have been assumed known. This is tantamount to a decoupling of problems, for example, the shape from shading problem and the light estimation problem. We submit that such decoupling, while appropriate for certain highly engineered situations, is not always necessary; moreover, it can make shading analysis impotent precisely when it should be useful.

Part 2

THE SHADING FLOW FIELD

CHAPTER 4

The General Shape from Shading Problem

A human observer confronted with a static, monocular view of a scene will succeed in obtaining some estimate of the shapes of the surfaces within it, even when some of the classical setting's constraints are relaxed. The presence of a shadow, a diffuse light source, or even a patterned surface does not necessarily interfere with our ability to recover shape from shading. Thus the classical constraints can be relaxed in principle; but how far?

In this chapter, we redefine the shape from shading problem and we describe how the classical assumptions need to be relaxed to deal the problems that we have identified.

1. The Objectives

Scenes often contain multiple surfaces. It is common that these surfaces partly occlude one another and even that a surface self-occludes. It is also common for different surfaces to have different reflectance coefficients. Furthermore, occlusions are common along lines between a light source and points on a surface; this latter instance causes lighting conditions to abruptly change and creates shadows.

Classical Shape from Shading	Generalized Shape from Shading
Single smooth surface	Multiple smooth surfaces
Single light source No shadow	Multiple light sources Shadows
Single albedo scene	Multiple albedo scene

TABLE 4.1. The classical shape from shading problem and the generalized shape from shading problem differ in the scenes they can deal with.

We have observed in Chapt. 3 that these scenes cause the classical shape from shading algorithms to fail. These failures are not due to bad algorithm design, rather the cause is intrinsic to the problem's definition. The assumptions underlying the classical shape from

shading problem are too restrictive. To handle the situations in Chapt. 3, the classical setting needs to be generalized as in Table 4.1.

We want to be able to also recover shape from shading for these problematic scenes simply because they are so common. Therefore we have to redefine the shape from shading problem and attempt to address all the problems raised by our observations.

2. The Line of Thought

The scenes mentioned in the previous section do not comply with the constraints inherent in the assumptions of the classical shape from shading problem. Thus, first and foremost, we have to relax these constraints on the scene in order to correct this situation. But how much can we relax them?

Can we get rid of them altogether? Consider the general form of the image irradiance equation:

$$I(x, y) = \Phi(x, y) \Xi(x, y) ,$$

where the functions

$$\begin{aligned} \Phi(x, y) &= \rho(x, y) \lambda(x, y) , \\ \Xi(x, y) &= \mathbf{L}(x, y) \cdot \mathbf{N}(x, y) . \end{aligned}$$

How can we recover $\Phi(x, y)$ and $\Xi(x, y)$ given only their product? This is obviously impossible without any assumption.

So we redefine the shape from shading problem in a way such that we retain only the essence of the basic assumption implicit in the classical shape from shading problem, that is:

ASSUMPTION 4.1. *For the generalized shape from shading problem, smooth variations in intensity are entirely due to smooth variations in surface orientation.*

To make the distinction clear, we emphasize the fact that this assumption is only concerned with variations that are *smooth*. Abrupt variations of intensity can be due to abrupt variations in surface orientation, but they can also be due to abrupt variations in the surface albedo or abrupt variations in the lighting conditions. We examine such abrupt changes later (see Chapt. 6).

The basis for the distinction between Assumptions 3.1 and 4.1 lies in the observation that intensity changes occur at different scales [79, 80]. Small changes in image intensity

characterize smooth change in scene features (i.e. smooth surface¹) although the quantization of space and intensity introduces some noise. Large changes in image intensity characterize changing scene features, i.e. wherever the surface reflectance, the lighting condition, or the surface orientation change abruptly. This notion of scale was fundamental to Land's Retinex theory [79,80], and will be important to our generalization of shape from shading.

The changes of intensity occurring at various scales reveal that there is more to an image than just the photometric aspect. The various structures are shown in Fig. 4.17. Along with the intensity, we can distinguish the shading flow field (small changes) and the image discontinuity curves (large changes).

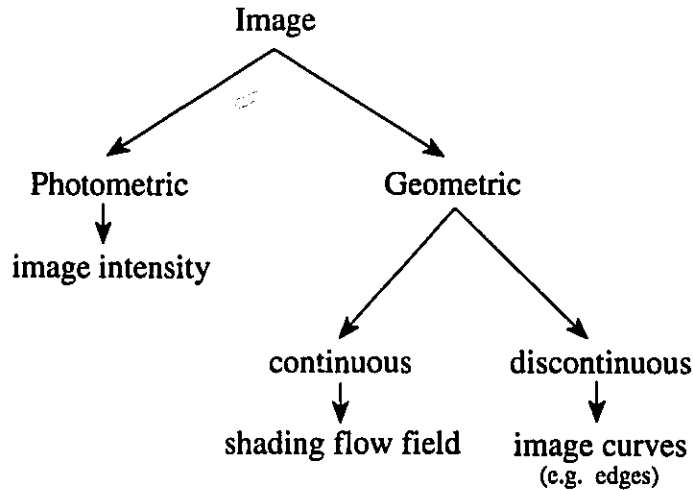


FIGURE 4.17. An image comprises a set of pixel intensity values. These capture the photometric aspect of the image. An image also implicitly defines geometric structures. These emerge from the neighbourhood relationships of the image pixels. These structures can describe either continuous properties of the image or discontinuous ones.

The significance of Assumption 4.1 for the reflectance function of a matte surface can be expressed by the following equation:

$$(4.1) \quad \nabla I(x, y) = \rho(x, y) \lambda(x, y) \mathbf{L}(x, y) \cdot \nabla \mathbf{N}(x, y) .$$

At every point (x, y) where the image intensity gradient $\nabla I(x, y)$ is defined, this equation implies that the surface albedo $\rho(x, y)$, the illumination $\lambda(x, y)$ and the direction of the illuminant $\mathbf{L}(x, y)$ are *locally* constant. It also implies the existence of a differentiable surface normal $\nabla \mathbf{N}(x, y)$. Note that the other terms in the total derivative ∇I are assumed

¹In Chapt. 10, we look at alternative assumptions.

nonexistent according to the above comments about scale; i.e., $\nabla L = 0$, $\nabla \rho = 0$, and $\nabla \lambda = 0$.

There are two key conceptual differences between our approach to the generalized shape from shading problem and the approaches to the classical shape from shading problem described in Chapt. 3:

- (i) in order to generalize the scene model, the scene constraints are considered to be local and not necessarily global;
- (ii) since the surface reflectance coefficient and illumination cannot be taken as given, surface shape recovery is based on the geometric structures of the image rather than on the photometric structures.

We elaborate on the importance of these differences in the following sections. Alternative options for relaxing the basic assumption are discussed in Chapt. 10.

3. Local vs. Global Scene Constraints

Using local scene constraints instead of global scene constraints has several noteworthy consequences:

- It addresses the problematical issues in the scene model that caused the classical shape from shading algorithm to fail.
- It introduces discontinuities into the scene model. The projection of these discontinuities onto the image plane provides another powerful cue to aid surface shape estimation.
- It does not allow for a global description of scene descriptors, such as the albedo ρ , the illumination λ or the illuminant direction L , and thus these cannot be taken as given. Rather, it stresses the need to evaluate these concurrently with the surface shape.

3.1. Generalizing the Scene Model. We have demonstrated in Chapt. 3 that the classical global constraints are too restrictive to deal with several features that commonly occur. The key to our approach is to reconsider these same constraints but with a reduced range; e.g. whereas for the classical shape from shading problem, the albedo was constant over the entire scene, for the generalized shape from shading problem, it is assumed constant only over a neighbourhood. Table 4.2 summarizes the constraints used by the different scene models.

The reflectance properties of the scene are only considered *locally* constant; the scene is not presumed to be described with a single albedo. The illumination of the scene is also

Classical Shape from Shading	Generalized Shape from Shading
Global constraints	Local constraints
Smooth surface	Piecewise smooth surface
Constant illumination	Piecewise constant illumination
Constant albedo	Piecewise constant albedo

TABLE 4.2. The classical shape from shading problem and the generalized shape from shading problem differs on the surface, the illumination, and the reflectance constraints and the resulting model.

considered to be *locally* constant; the scene is not presumed to be everywhere illuminated in the same way. The differentiability (a local property) of the surface normal function implies that this function is continuous (another local property). Where the image intensity gradient is not defined, the surface normal function can be discontinuous; the scene is not presumed to be described by a single continuous surface.

3.2. Introducing a Model with Discontinuities. A model with piecewise constant albedo implies that the projection onto the image plane will be marked by discontinuities in albedo. A model with piecewise constant lighting also implies that the projection onto the image plane will be marked by discontinuities in illumination and illuminant direction. Similarly, a model with piecewise smooth surfaces implies that the projection onto the image plane will be marked by discontinuities in surface depth, orientation and shape. All of these discontinuities form contours.

OBSERVATION 4.1. *The discontinuities in the scene are marked by curves of discontinuity in the image plane.*

The relations between the curves of image discontinuities and the curves of scene discontinuities provide an important insight to distinguish the various scene discontinuities (see Chapt. 6).

Since geometric discontinuities (in curvature, orientation, and depth) are unavoidable and their projection onto the image has a widely recognized importance [8, 69, 73], we allow for discontinuities. We therefore assume that the scene is composed of piecewise smooth surfaces².

²Since the reflectance function depends on the existence of a differentiable surface normal, allowing surfaces that are nowhere smooth is clearly inappropriate.

OBSERVATION 4.2. *The projection onto the image of geometric discontinuities provides an additional cue to help shape recovery.*

3.3. Coupling Surface and Light Source Estimation. The classical shape from shading assumptions allowed the scene to be modelled by a single albedo, a single illumination and a single illuminant direction. The albedo and the illumination always enter the equations as a product, so, for the entire image, a single vector $\rho\lambda\mathbf{L}$ is needed. When the classical assumptions hold, one can expect to extract this single vector based on statistical considerations [105].

Since we are considering local constraints for the generalized shape from shading problem, the vector $\rho\lambda\mathbf{L}$ is constant only within some neighbourhood. One should typically expect this vector to take multiple values for any given scene, i.e. according to our assumptions, it will take a given value for points corresponding to a certain patch of the image, and it will take a different value for points corresponding to a different patch. Several patches of various sizes and forms have to be delimited and modelled by different vector values. The problem is more complex and one cannot expect to extract the vector $\rho\lambda\mathbf{L}$ based on the statistical considerations used for the classical problem. The problems of shape from shading and light source estimation can no longer be decoupled. Light sources and surface properties must be handled concurrently; neither problem must be solved “before” the other.

4. Geometric vs. Photometric Image Structures

To emphasize the dichotomy between the photometric and geometric structures, we rewrite Eq. 4.1 as follows:

$$\nabla I(x, y) = \Phi(x, y) \tilde{\Gamma}(x, y)$$

where the photometric and geometric aspects of the scene are respectively:

$$\Phi(x, y) = \rho(x, y) \lambda(x, y)$$

$$\tilde{\Gamma}(x, y) = \mathbf{L}(x, y) \cdot \nabla \mathbf{N}(x, y) .$$

The photometric structure of the image, the image intensity, is directly related to the photometric aspect of the scene, the product of the illumination and the albedo of the surface $\Phi(x, y)$ as shown in Eq. 3.2. Since we are letting both the illumination and the albedo be functions of position, the image intensity is a poor choice for initial data of the shape from shading problem. The geometric information about the scene would be confounded with the photometric.

OBSERVATION 4.3. *The relaxation of the scene constraints leads to the rejection of the photometric structure of the image as initial data for the shape from shading problem.*

The geometric aspect of the image reveals information of a different nature. We consider two types of structures [136]:

- type I. Contours. These are one-dimensional structures in a two-dimensional space. Of particular importance are those curves that correspond to discontinuities.
- type II. Flow patterns. These are two-dimensional structures in a two-dimensional space. Vector fields and direction fields are examples; we shall consider a specific one called the shading flow field.

To elaborate, viewing the intensity image as a scalar field, we derive two geometric structures from it. One of these corresponds to the regions depicting smooth intensity changes, or regions in which the intensity gradient is well-behaved. This will become the shading flow field, and it is developed later in this section. Separating these regions of smooth intensity variation are Jordan curves of discontinuities — these correspond to so-called “edges” in images, and they depict the locus of positions along which surface and lighting properties change abruptly. Note that the curves and fields are complementary, but that they both provide information about shape.

Singularities of the flow field constitute another class of geometric structures.

- Point singularities. These are zero-dimensional structures in a two-dimensional space. A specularity on a sphere is an example of a situation that generates such a structure. Such a singularity corresponds to a point where the shading flow field is undefined — the singularity being a sink (or a source) of the gradient field. Point singularities can be characterized by their index.
- Line singularities. These are one-dimensional structures in a two-dimensional space. A highlight on a cylinder is an example of a situation that generates such a structure. These structures in the shading flow field are characteristic of parabolic surface patches (or line [72]) as the surface curves in only one direction.
- Undefined regions. These are two-dimensional structures in a two-dimensional space. The back plane in Fig. 3.12 is an example of a situation that generates such a structure. These structures in the shading flow field are characteristic of planar surface patches as the surface does not curve in any direction, hence the intensity remains constant.

We continue next with a discussion of image curves, represented locally as a tangent field. First, however, note that in a static image, we can extract a shading flow field and

a texture flow field [136]. Although shading may overlap a texture flow, it is difficult to recover when the changes occur at the same scale. Since we consider only the smallest scale, the shading flow field [12] and the texture flow field can be viewed as exclusive and complementary. Both provide information about shape³ [109].

4.1. The Tangent Field. The tangent field [136] comprises the position, orientation and curvature information about image curves: $\Upsilon^t = \{x, y, \theta^{(t)}, \kappa^{(t)}\}$. There are several types of image curves, e.g. bright lines, dark lines, edges. The term “edge” refers to the image curve along which an abrupt change in the image intensity occurs.

Artists have drawn sketches in terms of line drawings for several centuries. It has been acknowledged as a powerful cue for the perception of shape. When humans look at line drawings, they can often extract a qualitative description of what they see [43]. When viewing the line drawing in Fig. 4.18, most people immediately recognize the humanoid figure.



FIGURE 4.18. The line drawing provides sufficient information to humans to allow them to recognize some scenes. This line drawing was extracted from the gray-level image with Iverson's logical-linear operators for edges and lines [66].

Often, line drawings contain sufficient data to drive the recognition process. Occluding contours have probably received the most attention. Richards et al. [113] attempted to

³Note that the scope of this current work includes contours and shading flows. The integration of texture information into this framework is left for the future.

uncover how unique three-dimensional interpretations could be made from two-dimensional silhouettes. At the heart of their work were two rules:

- (i) Do not propose undulations of the 3-D surface without evidence for such.
- (ii) Pick the most general position 3-D interpretation, namely, that 3-D shape that preserves the signs of the curvature of the silhouette over the widest range of view-points.

and three constraints:

- (i) The sign of the Gaussian curvature of points on the 3-D surface that project onto the silhouette is the same as the sign of curvature of those projections [71]. This qualitative constraint on the surface shape is independent of viewing distance [69].
- (ii) For generic surfaces, the flexional (parabolic) lines are closed and non-intersecting.
- (iii) A region of negative curvature on a silhouette is always interpreted in three dimensions as a furrow (or neck), never as a dent.

In a companion paper, Beusmann et al. [5] proposed a method of representing the shape of complex objects as convex parts. They showed that the part boundaries, hyperbolic regions, could be inferred from the occluding contours. Koenderink and van Doorn [73] described how this type of representation is common in fine arts.

OBSERVATION 4.4. Occluding contours only depend on the geometric properties of the surface with respect to the viewer. The relationship between the scene and the orientation and curvature of an edge element is therefore independent of the illumination model and the surface reflectance model.

At those positions where the image intensity gradient is not defined, image curves can provide precious information.

4.2. The Shading Flow Field. If we provide the shading information along with the line drawing, we can notice details that we have missed previously. The shading provides evidence of undulations that the line-drawing cannot capture.

In Fig. 4.19, the hollow between the shoulder blades is an example. Here the shading provides evidence of an undulation of the 3-D surface that was not available from the line drawing. Neither the lighting condition nor the reflectance properties of the surface change abruptly, and since the viewing position is facing the hollow, there is no self-occlusion. This is typical of situations in which the line drawing extracted from an image (Fig. 4.18) is insufficient to capture the (smooth) variations in shape.

4. THE GENERAL SHAPE FROM SHADING PROBLEM



FIGURE 4.19. The shading provides additional information to humans. It allows us to notice some shape detail on the surface.

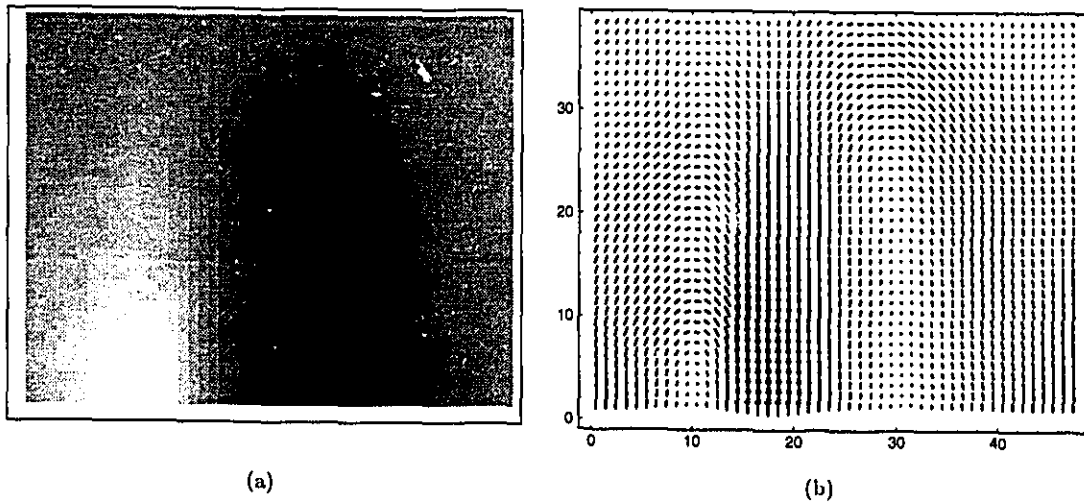


FIGURE 4.20. An image (a) of the shoulder blade region in Fig. 4.19 and its shading flow field (b). The shading is not sufficient to characterize unambiguously the shape.

An image without discontinuities is sometimes quite confusing. In Fig. 4.20, we have cropped the image and we have only taken the region between the shoulder blades. Although we can still get a qualitative description of the shape, its orientation is not reliable [75]. We note here that a similar phenomenon occurs with texture when viewed through an aperture (the surfaces always seem fronto-parallel).

OBSERVATION 4.5. *The line drawings and the shading provide complementary information about the shape of surfaces.*

Richards et al.'s first rule consists of not proposing undulations of the 3-D surface without evidence for such. The shading flow field provides precisely this evidence.

If we look at the first order directional derivatives of the intensity function,

$$I_x = \rho\lambda \mathbf{L} \cdot \frac{\partial}{\partial x} \mathbf{N} ,$$

$$I_y = \rho\lambda \mathbf{L} \cdot \frac{\partial}{\partial y} \mathbf{N} ,$$

or the second order directional derivatives of the intensity function,

$$I_{xx} = \rho\lambda \mathbf{L} \cdot \frac{\partial^2}{\partial x^2} \mathbf{N}$$

$$I_{xy} = \rho\lambda \mathbf{L} \cdot \frac{\partial^2}{\partial x \partial y} \mathbf{N}$$

$$I_{yy} = \rho\lambda \mathbf{L} \cdot \frac{\partial^2}{\partial y^2} \mathbf{N}$$

we note that all these partial derivatives depend on the variable illumination and albedo. Again, as in the image intensities, these measures confound the scene's geometric information with the scene's photometric properties. Therefore, they are not appropriate to solve the generalized shape from shading problem.

The directional derivatives' information can nevertheless be combined in ways that are independent from the variable illumination and albedo. One such way is the *shading flow field*.

DEFINITION 4.1. *The shading flow field is the vector field which indicates the direction in which the image intensity remains constant. It is therefore perpendicular to the image intensity gradient field.*

Although the following calculations are elementary, their implications are deep for our model.

THEOREM 4.1. *For the generalized shape from shading problem, the orientation of the intensity gradient field (where it is defined) is independent of the variable illumination (λ) and albedo (ρ). Thus the orientation of the intensity gradient field only depends on the geometric properties of the surface $\bar{\Gamma}(x, y)$ and the lighting with respect to the viewer.*

Proof: The orientation of the gradient field θ is related to a ratio of directional derivatives:

$$\tan \theta = \frac{I_y}{I_x}.$$

Since both directional derivatives are directly proportional to the product of the illumination and the albedo,

$$\tan \theta = \frac{\mathbf{L} \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L} \cdot \frac{\partial}{\partial x} \mathbf{N}} = \frac{\Gamma_y}{\Gamma_x}.$$

The orientation of the intensity gradient field is independent from the variable illumination (λ) and albedo (ρ). $\square$

COROLLARY 4.1. *The shading flow field (where it is defined) is independent of the variable illumination and albedo, and thus only depends on the geometric properties of the surface and the lighting with respect to the viewer.*

We acquire two vector fields based on this orientation information. We can consider the normalized gradient field and the shading flow field as unit vector fields (direction fields).

COROLLARY 4.2. *The curvatures of both the shading flow field and the normalized gradient field are also independent of the variable illumination and albedo, and thus only depend on the geometric properties of the surface and the lighting with respect to the viewer.*

The curvature of shading flow field κ_s and the curvature of the normalized gradient field κ_g are respectively:

$$\kappa_s = \frac{2I_x I_y I_{xy} - I_x^2 I_{yy} - I_y^2 I_{xx}}{(I_x^2 + I_y^2)^{\frac{3}{2}}};$$

$$\kappa_g = \frac{I_x I_y (I_{xx} - I_{yy}) - I_{xy} (I_x^2 - I_y^2)}{(I_x^2 + I_y^2)^{\frac{3}{2}}}.$$

Since the first and second directional derivatives are directly proportional to the product of the illumination and the albedo, in both instances, the numerator and the dominator would be proportional the cube of this product and cancel out.

The shading flow field can exhibit singularities. Singularities of index one would correspond to circulations, either clockwise or counterclockwise. Of particular importance for

Bichsel and Pentland's algorithm were those points that correspond to maxima of intensity. These are counterclockwise circulations. Note that the shading flow field does not give rise to sink or source because of the way it is defined with respect the image intensity — just recall that the curl of the gradient of a scalar field is identically null. Singularities of index minus one would correspond to saddles. We obtain the index of an isolated point singularity by summing the angle differences between vectors (divided by 2π) as we follow a closed path around the singularity in a counterclockwise fashion. Thus, from the shading flow field, we have a straightforward way to locate singularities.

4.2.1. *Other Motivations.* Observe that a sensitivity issue arises in the shape from shading problem; spatial quantization of the image induces a quantization of the scene domain. Analogously to the manner in which integer solutions are not always possible for algebraic equations, we begin with “quantized” initial data as well. In particular, we derive our initial estimates from *the shading flow field* instead of directly from the intensity image. This field is the first order differential structure of the intensity image expressed as the isoluminance direction and augmented with the gradient magnitude; we supplement it with the intensity “edge” image. We suggest that dealing with uncertainties at the level of the shading flow field will expose more of the natural spatial consistency of the intensity variation, and will thus lead to more robust⁴ processing than the raw intensities. The shading flow field ideas are related to Koenderink's isophotes [72] and Wolff photometric flow fields [128].

Our motivation for starting from the shading flow field is also biological. We take shading analysis to be an inherently geometric process, and hence handled within the same cortical systems that provide orientation selection and texture flow analysis (see App. A). Shading flow is simply a natural extension.

5. Summary

We re-examine the problems uncovered in Chapt. 3 for the classical formulation of the shape from shading. The discussion of the “local vs. global” scene constraints and the “geometric vs. photometric” image structures yields the following proposals.

PROPOSAL 4.1 (OBSERVATION 3.1 REVISITED). *The first difficulty observed was that the computed shape of a background surface is constrained by an occluding object. By assuming that the scene can comprise several smooth surfaces, we allow for surface discontinuities.*

⁴Throughout this thesis, the term “robust” is used as meaning “not sensitive to noise”.

Thus, the shape of the background surface is no longer necessarily constrained by occluding objects.

PROPOSAL 4.2 (OBSERVATION 3.2 REVISITED). *The second difficulty observed was that shape recovery depends on an estimation of the initial data. The reliability of the estimate is crucial to an accurate recovery of shape. By taking the abrupt changes in the image intensity and the shading flow field, we have chosen as initial data only geometric structures that can be more reliably extracted from the image. Unlike maxima or minima, each of these data can be refined according to their own geometry.*

In his Ph.D. thesis, Lee Iverson [66] describes a reliable way to extract the abrupt changes in the image intensity. In Chapt. 5, we propose a reliable way to extract the shading flow field.

PROPOSAL 4.3 (OBSERVATION 3.3 REVISITED). *The third difficulty observed was the failure to deal with shadows. Shadows are zones where the lighting conditions are different, as one light source is somehow occluded. By assuming that several distant light sources illuminate the scene, we allow for the changing lighting condition that creates shadows. The lighting conditions are not taken as given.*

PROPOSAL 4.4 (OBSERVATION 3.4 REVISITED). *The fourth difficulty observed was the failure to deal with scenes that are not entirely the same colour. By assuming that the scene is composed of surfaces for which the reflectance coefficients are piecewise constant, we have the possibility to deal with scene which are not entirely the same colour. The reflectance coefficients are not taken as given.*

6. The Generalized Shape from Shading Problem

These proposals result in a new definition of the shape from shading problem — even though it still deals with a single monochromatic image, it allows more general scenes.

PROBLEM 4.1. GENERALIZED SHAPE FROM SHADING

Assuming that

- (i) *colours can be present in the scene, but are piecewise constant;*
- (ii) *several distant light sources can illuminate the scene;*
- (iii) *the scene can comprise several smooth surfaces;*
- (iv) *the surfaces are matte;*
- (v) *the image is formed by an orthographic projection.*

4. THE GENERAL SHAPE FROM SHADING PROBLEM

Given

- (i) *the tangent field;*
- (ii) *the shading flow field.*

Recover

- (i) *the surface shape;*
- (ii) *the illuminant direction.*

The scope of this shape from shading problem is more general than the classical shape from shading setting (see Problem 3.1). The assumptions made about the scene are less constraining. The only given variables are the geometric structures of the image — these can be extracted from the image intensities, as we now show.

CHAPTER 5

Computing the Shading Flow Field

Our shape from shading approach is based on the shading flow field. Therefore, it relies on an accurate initial estimation of these local properties of the shading. We have studied the problem of extracting this information from the image.

In this chapter, we describe how the shading flow field can be extracted from the image.

1. Introduction

Since the shading flow field is the “dual” of the normalized gradient field, we look at how one could extract the gradient of an image.

The problem of computing the gradient of an image would be straightforward if the image intensities were simply a mapping of $\mathbb{R}^2 \rightarrow \mathbb{R}$. It would suffice to compute the directional derivative in the x and y directions:

$$\nabla I(\vec{x}, y) = \left(\frac{\partial}{\partial x} I(x, y), \frac{\partial}{\partial y} I(x, y) \right) .$$

But, the image intensity is not a mapping of $\mathbb{R}^2 \rightarrow \mathbb{R}$. The image is discrete. It is a mapping of $\mathbb{Z}^2 \rightarrow \mathbb{Z}$. The situation becomes complicated by the fact that the notion of differentiability is not defined on integers. To make matters worse, the image is also intrinsically noisy.

So, it is necessary to redefine the problem. We are not computing the gradient of the image intensities directly, but seek to infer the gradient of a function that captures the shading of the scene as projected on the image plane. The image intensities provide a sampling of this function.

1.1. Fuzzy Derivatives. Traditionally, the gradient of an image $\vec{\nabla}I(x, y)$ is computed by smoothing the image intensity function $I(x, y)$ before estimating directional derivatives. The smoothing is usually done by the convolution of a Gaussian kernel over the image.

$$I_{\sigma}(x', y') \approx \sum_{(x, y)} I(x, y) G_{\sigma}(x - x') G_{\sigma}(y - y') ,$$

where

$$G_{\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} ,$$

and σ is the variance.

Using the convolution property of Gaussian kernels, these equations can be rewritten in terms of “fuzzy derivatives” [74]:

$$\frac{\partial I_{\sigma}}{\partial x} \approx I * G'_{\sigma}(x) G_{\sigma}(y) ,$$

$$\frac{\partial I_{\sigma}}{\partial y} \approx I * G_{\sigma}(x) G'_{\sigma}(y) ;$$

where $G'_{\sigma}(x)$ is the first derivative of the Gaussian. The gradient estimate follows immediately, and the isoluminance direction is simply perpendicular to the gradient.

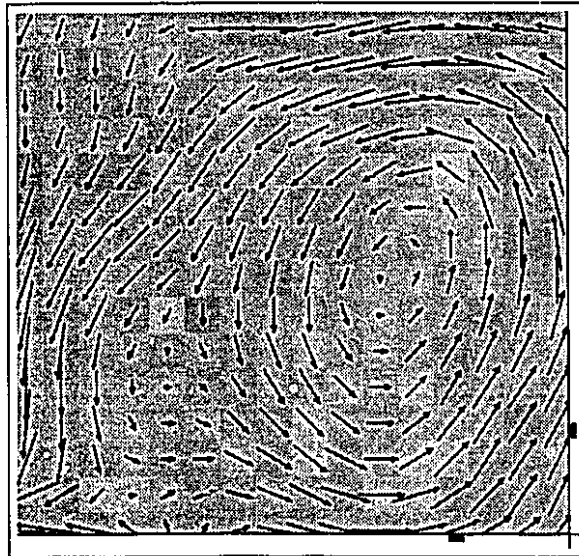


FIGURE 5.21. In this figure, we provide an example of the use of “fuzzy derivatives”. The data used is the gray-level image in background, and the resulting shading flow field, augmented with the gradient magnitude, is shown in foreground.

The one limitation of this approach is that (depending on the magnitude of σ) it always infers a smooth gradient field, even when the underlying image is non-continuous.

1.2. Limitation and Difficulties. The notion of continuity is an important underlying condition for differentiability. By smoothing the image with a Gaussian kernel, one ensures that the function is continuous and that its derivatives are continuous. But by doing so, we are also masking other important structures.

Consider the image formation process for a smooth matte surface under a constant lighting condition. It yields a smooth intensity function. Its projection on the image plane should also yield a smooth image intensity function¹. It should normally result in a smooth gradient field (See Fig. 5.21). This is fine.

However, often the imaged scene contains discontinuous features. We shall now consider two such images. They both consist of two smooth surfaces: a sphere in front of a plane.

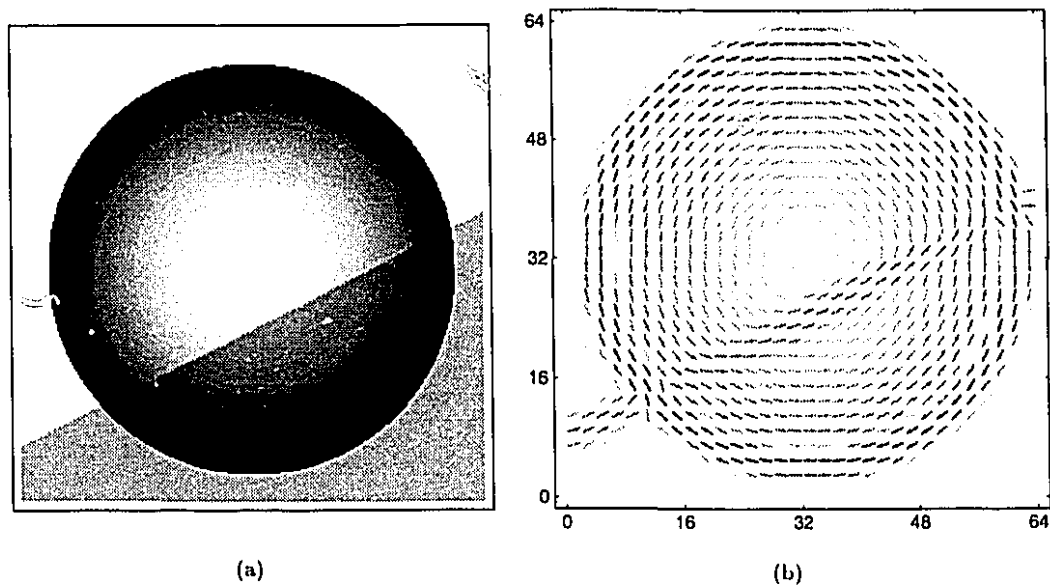


FIGURE 5.22. The technique of fuzzy derivatives produces a smooth shading flow field even though the underlying image intensity function has marked discontinuities. Here, (a) the image of a scene with intensity discontinuities yields (b) a continuous shading flow field. In this figure, the gradient magnitude is coded in gray levels — darker meaning larger gradient.

In the first image (Fig. 5.22(a)), the lighting condition remains constant, but one part of the sphere and the plane is of a different albedo than the other. Both surfaces generate a smooth intensity function. Each surface should independently exhibit a certain shading flow pattern. But since the sphere partially occludes the plane, the projection onto the image plane would typically yield a discontinuous intensity function and a discontinuous gradient

¹Here we assume that the surface is not self-occluding

field. The discontinuity should lie on a curve. On both sides of that curve, the intensity function and the gradient field should be smooth. Figure 5.22(b) shows the shading flow field obtained with the fuzzy derivative approach. It is continuous everywhere. Near an intensity discontinuity (it is particularly striking around the albedo change), the flow field's distortion reflects the size of the convolution kernel. This undesirable result occurs because the technique of fuzzy derivatives involves smoothing the intensity function everywhere without consideration for the discontinuities that occur at various scales. Therefore, it masks an important image feature and consequently, it introduces errors in the estimation of the directional derivatives.

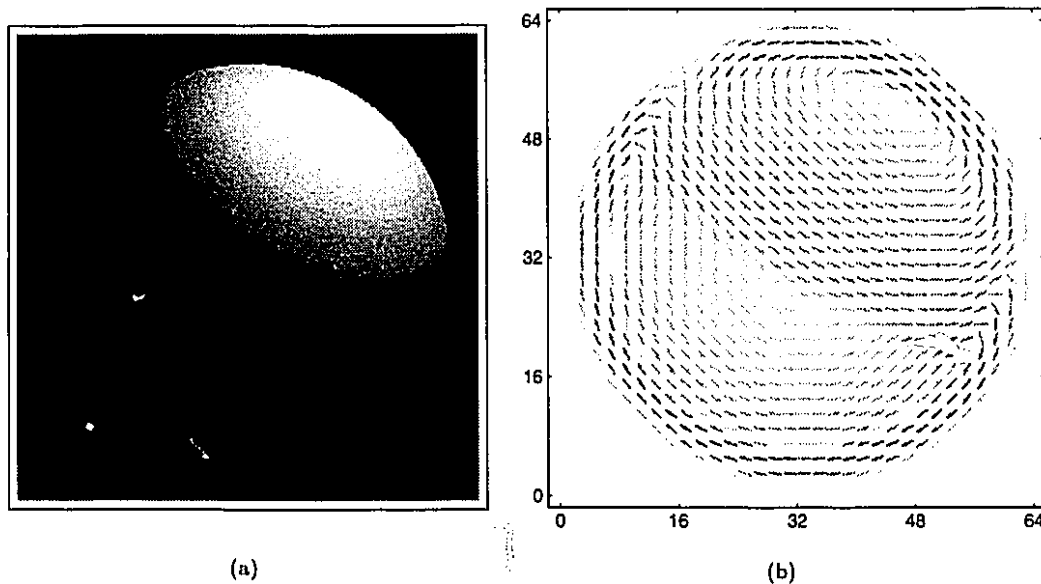


FIGURE 5.23. The technique of fuzzy derivatives produces a smooth shading flow field even though the underlying image intensity function should lead to a shading flow field with marked discontinuities. Here, (a) the image of a scene with a discontinuity that should be reflected in the shading flow field yields (b) a continuous shading flow field. In this figure, the gradient magnitude is coded as term of gray levels — darker meaning larger gradient.

In the second image (Fig. 5.23(a)), the albedo remains constant, but the lighting conditions change since one of the light sources present is not visible from everywhere on the sphere. Where both light sources shine, one should observe a different shading flow pattern than where only one light shines. We should observe a discontinuity in the shading flow field. Figure 5.23(b) shows the shading flow field obtained with the fuzzy derivative approach. Again, it is continuous everywhere. This undesirable result occurs because the technique of fuzzy derivatives involves smoothing the intensity function everywhere without

consideration for potential underlying shading flow field discontinuities. It therefore masks another important image feature and consequently, it introduces errors in the estimation of the directional derivatives.

OBSERVATION 5.1. *Smoothing over discontinuities distorts the gradient field and the shading flow field.*

The fuzzy derivative approach is thus inappropriate in the presence of discontinuities.

2. Piecewise Continuous Flow Field

To extract the gradient from an image, we consider a different method. For each image point, the intensity function is locally modeled by a smooth function ($\mathbb{R}^2 \rightarrow \mathbb{R}$). Since only the first derivatives are needed, we use a linear function:

$$I(x, y) \approx a x + b y + c .$$

A least squares fit provides the parameters of the modelling function from which the gradient is computed. Since this is a local model, we assume that the confidence in the data decreases with the distance to the pixel (p_0) over which the gradient is estimated; $\Delta(x, y) = G_\sigma(r)$ with $\sigma = 2$. For computational simplicity, we ignore data points further than 4.5 pixels, as their contributions are not noticeable — the integral over the remaining pixels' weight represents less than 0.4% of the total integral. The pixels used for locally fitting a continuous intensity function are shown in Fig. 5.24(a).

This “fitting” approach also infers a smooth gradient field. As with the “fuzzy derivatives”, if there is a discontinuity in the neighbourhood of a pixel, the gradient at this point will be distorted. Up to now, this approach is a first order approximation to the intensity function $I(x, y)$ around p_0 . It assumes that the function and its first derivatives are continuous over the data points. This is not always the case. Next, we review how discontinuities are uncovered, and after, we show how we adapt the “fitting” approach to take advantage of this newly acquired information.

2.1. Computing Intensity Discontinuities. The discontinuities in intensity form image curves. These image curves are characterized both by their tangential properties (orientation $\theta^{(t)}$, curvature $\kappa^{(t)}$) and their cross-sectional properties (edge, positive and negative contrast lines).

The problem of edge detection in images appeared in the first computer vision system [115]. Edges are said to be, by definition, transitions between two markedly dissimilar

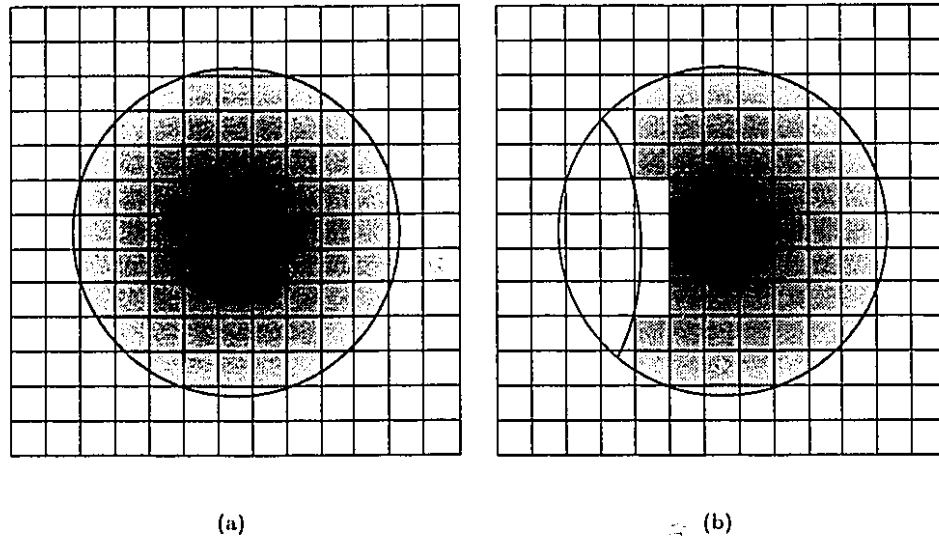


FIGURE 5.24. For a given pixel, the data points used for fitting a local smooth function, lies in its neighbourhood. (a) The confidence given to a data point depends on its distance from where the model is fitted as illustrated the radial shading. For computational simplicity, we use only the pixels whose centre are within the circle. (b) The pixels, on or beyond an edge or a line passing through the neighbourhood, are not used to compute the gradient.

intensities. In terms of the image intensity function, the edge is a region in the $x - y$ plane where $I(x, y)$ has a gradient of large magnitude. The Roberts cross operator was designed as a discrete approximation to the gradient magnitude. It is based on a 2×2 window and thus very sensitive to noise.

Attempts to find a less noise sensitive operator have been numerous. Sobel suggested an operator [28] using a larger window (3×3) than the Roberts cross operator in an attempt to smooth out the noise. Modelling edges in images as an ideal step edge and additive noise gave rise to an entire class of solutions [19, 25, 47, 93, 94]. The various designs proposed can be decomposed in two operations: one to smooth the noise, an other to locate the edge. These methods rely on thresholds to determine what is considered a gradient of large magnitude. This is not desirable since a single threshold value for the entire image is rarely appropriate. Depending on an arbitrary choice of the global threshold, the operator may signal the presence of edges where there are none and may signal the absence of an edge where there are some.

In his thesis, Lee Iverson [66] addresses all these problems while differentiating between the different types of image curves. Furthermore, his model takes into account the possible coexistence of multiple edges at a given image location which allows for an accurate location

and description of features such as corners and junctions. These are fundamental to the understanding of line drawings [7, 48, 126].

Edge detection is still an active research topic. Even though we haven't considered using them, it is worth mentioning that other techniques have recently been proposed such as "steerable filters" [38, 110, 111] and "edge-energy" methods [96, 111].

2.2. Defining the Shading Flow Field Operator's Domain. For the "fitting" approach to be meaningful, corrective steps are needed for both intensity discontinuities and shading flow field discontinuities.

At a discontinuity, the gradient is not defined. However, it is defined near a discontinuity — it is in fact defined on both sides of a discontinuity. Both the magnitude and the orientation can differ. Hence, we choose to acknowledge this reality by allowing multiple representations of the gradient at any given pixel location. We define a set of possible gradient orientations, and design operators which provide a confidence measure for each one.

We have investigated methods of obtaining stable *discontinuous* shading flow fields using logical/linear operators [65] to find edges and lines. By first identifying the discontinuity curves in the image, the use of data points lying on² or beyond any discontinuity can be avoided. This is precisely what we shall do as shown in Fig. 5.24(b). This corrective step prevents smoothing across an identifiable intensity discontinuity.

For shading flow field discontinuities, the difficulty is different since no *a priori* knowledge about their presence is available. To evaluate the confidence of a possible gradient orientation in the neighbourhood of a given pixel, we consider sixteen different fits. Each fit uses a subset of pixels in the neighbourhood as shown in Fig. 5.25. Each subset corresponds to a different assumption about the presence of a shading flow field discontinuity and rejects pixels beyond the hypothetical discontinuity curve. The number of fits ($N_d = 16$) reflects two constraints:

- (i) different shading flow field discontinuities should yield different domains for the fit;
- (ii) the domains for the fit should be large enough to smooth the noise.

For each orientation of a possible discontinuity, the gradient is computed.

The initial estimate of the shading flow field reflects these values. We assign a confidence, $P(\theta_i)$, to the hypothesis that the orientation of the shading flow field is a value θ_i ,

²A pixel on an image intensity discontinuity usually reflects an interpolation of the intensity on both sides of the discontinuity. It can thus induce an error in the gradient estimation.

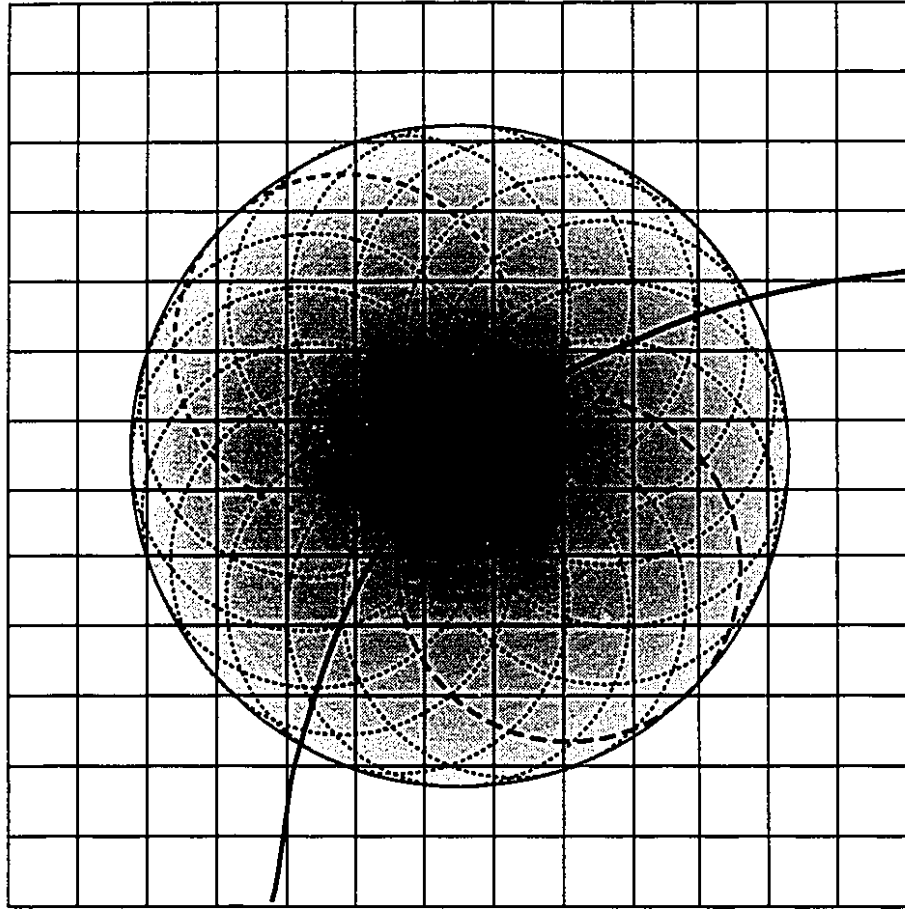


FIGURE 5.25. For a given pixel, the data points used for fitting a local smooth function, lie in its neighbourhood. The confidence given a data point depends on its distance from where the model is fitted as illustrated by the radial shading. To deal with the possible presence of shading flow field discontinuities, we subdivide the neighbourhood. A set of 16 neighbourhoods used to provide reliable estimates of the gradient are shown here as dotted circles. This partitioning of the neighbourhood insures that at least one fit will not be corrupted even if there is a shading flow field discontinuity. For example, if such a discontinuity (shown here as a solid line) passes through the pixel, the two dashed circles would provide support to the appropriate flow hypotheses as both encircle a region adjacent to the discontinuity. Two distinct shading flow hypotheses could be retained. For another example, if such a discontinuity (not shown in this figure) passes elsewhere in the neighbourhood of the pixel, there would be more dotted circles to provide support for the one appropriate flow hypothesis. As in Fig. 5.24, if the tangent field indicates the presence of an edge or a line, pixels on or beyond are not used to compute the fit.

as follows:

$$P(\theta_i) = \frac{1}{N_d} \sum_{j=1}^{N_d} e^{-\left(\frac{16}{\pi} (\theta_i - \theta_j)\right)^2}$$

where θ_j is the orientation of the gradient computed assuming a shading flow field discontinuity of orientation $\frac{i\pi}{8}$ rd.

2.2.1. Discrete Shading Flow Field Representation. The space of shading flow orientation is divided into a discrete set of regular intervals. Each interval defines an equivalence class which we represent with the central orientation of the interval: θ_i . Each one is viewed as a hypothesis that the directional derivative of the image intensity function is null in an orientation within the interval.

The shading flow field's orientation is a local measure. Hence, the need to use a neighbourhood as small as possible. The discretization of image position, (x, y) , simply reflects the tessellation in pixels of the image. The limited information available induces a limited precision in the measure. Our equivalence class simply reflects this fact. We observed that sixteen was about the number of different orientations, θ , that we could distinguish considering a circular neighbourhood with a radius of four pixels.

We represent the shading flow field augmented with the magnitude of the image intensity gradient, $|\nabla I|$. This scalar quantity is discretized in five intervals. The coarseness of this quantization is due to the uncertainty induced by the orientation quantization.

At this point, we can restate the problem as one of finding a consistent flow field out of these initial estimates. The shading flow field is no longer uniquely valued. We shall now consider the shading flow field as a union of direction fields. Such a relaxed definition allows a better representation of the shading flow field at discontinuities — both sides of a discontinuity can be represented.

3. Relation to Relaxation Labelling

Relaxation labelling is a computational method to find consistent structures in a network of nodes. Hummel and Zucker [58] have laid down the foundation of relaxation processes, and Parent and Zucker [103, 104] provided insights regarding their implementation.

To relate the problem of computing a shading flow field to the relaxation labelling paradigm, we consider each orientation as a node of the relaxation network. Since the orientation hypothesis can be either true or false, we associate two labels to each node, TRUE and FALSE. We distribute a measure p_i over each i representing confirmation of the hypothesis. In this section, we provide a brief review of the relaxation labelling paradigm.

3.1. The General Relaxation Labelling Paradigm. Let a set of nodes I be given. To each node $i \in I$, we associate a set of labels Λ_i . Each label $\lambda \in \Lambda_i$ is interpreted as a possible value to be assigned to a node i .

DEFINITION 5.1. *The confidence is a scalar value $p_i(\lambda)$ associated with each label $\lambda \in \Lambda_i$ at node $i \in I$. This value is restricted to be within zero and one, inclusive:*

$$0 \leq p_i(\lambda) \leq 1$$

When the confidence $p_i(\lambda)$ is unity, it confirms the assignment of label λ at the node i . When the confidence $p_i(\lambda)$ is zero, it disconfirms the assignment of label λ at the node i .

DEFINITION 5.2. *The labelling assignment is the triple of $(I, \Lambda, \mathbf{p})$ where $\mathbf{p}$ is an instance of a confidence measure being assigned to every label of every node and for which the following restriction holds:*

$$\sum_{\lambda \in \Lambda_i} p_i(\lambda) = 1 .$$

We define K as the space of such labelling assignments.

This restriction expresses the exclusive nature of the assignment problem. A simple interpretation of such an assignment is that each $p_i(\lambda)$ is the *confidence* that the label λ should be assigned to node i .

The labelling assignment is said to be *unambiguous* when

$$p_i(\lambda) \in \{0, 1\} \quad \forall i \in I, \lambda \in \Lambda_i .$$

Such a labelling defines a mapping

$$i \rightarrow \lambda \text{ if and only if } p_i(\lambda) = 1 .$$

We say then that λ is assigned to the node i .

Since the goal of the relaxation labelling method is to solve an assignment problem, a criterion needs to be defined. We seek a labelling which maximizes a measure of consistency.

DEFINITION 5.3. *The compatibility matrix R captures the pairwise compatibility between labels on different nodes. Its elements $r_{ij}(\lambda, \lambda')$ provide a measure of the compatibility between label λ at node i and label λ' at node j .*

A positive compatibility, $r_{ij}(\lambda, \lambda')$, between labels (λ, λ') means that the labels are consistent with each other, a negative compatibility means that the labels are inconsistent with each other, a null compatibility means that the labels are unrelated.

DEFINITION 5.4. *The support for a label λ at node i is defined as the weighted sum of the compatibilities between this label and every other possible label at every node.*

$$s_i(\lambda; \mathbf{p}) \equiv \sum_{j \in I} \sum_{\lambda' \in \Lambda_j} r_{ij}(\lambda, \lambda') p_j(\lambda') .$$

The support indicates the consistency of a label λ at a node i with labels at other nodes given a labelling assignment $\mathbf{p}$.

DEFINITION 5.5. *A consistent labelling is a labelling assignment which fulfills the condition that*

$$\forall i \in I, \forall \mathbf{v} \in K : \sum_{\lambda} p_i(\lambda) s_i(\lambda; \mathbf{p}) \geq \sum_{\lambda} v_i(\lambda) s_i(\lambda; \mathbf{p}) .$$

Formally, relaxation labelling solves the problem of finding a consistent labelling given an initial description $(I, \Lambda, \mathbf{p})$ and the compatibility matrix r_{ij} .

DEFINITION 5.6. *The average local consistency is defined as the weighted sum of the support for every possible label at every node.*

$$A(\mathbf{p}) = \sum_{i \in I} \sum_{\lambda \in \Lambda_i} p_i(\lambda) s_i(\lambda; \mathbf{p}) ,$$

or to make explicit the quadratic form of the function

$$A(\mathbf{p}) = \sum_{i \in I} \sum_{\lambda \in \Lambda_i} \sum_{j \in I} \sum_{\lambda' \in \Lambda_j} p_i(\lambda) r_{ij}(\lambda, \lambda') p_j(\lambda') ,$$

Hummel and Zucker [58] have shown that, for symmetric compatibilities, the following algorithm constitutes a gradient ascent on average local consistency which terminates at a consistent labelling.

ALGORITHM 5.1 (RELAXATION LABELLING).

- (i) *Compute an initial estimate of $\mathbf{p} = \{p_i(\lambda)\}$ which constitutes a labelling assignment. Call this $\mathbf{p}^0$.*
- (ii) *Repeat starting with $n = 0$ until $\mathbf{p}^n$ is consistent:*
 - (a) *Repeat for all $i \in I$:*
 - (i) *Compute $p_i^* = p_i^n + \delta s_i$.*
 - (ii) *Project p_i^* onto a valid labelling assignment. This new assignment is p_i^{n+1} .*
 - (b) *Set $n = n + 1$.*
- (iii) *Generate the mapping $i \rightarrow \lambda$.*

In this algorithm, δ is the quantity that controls the stepsize and its predetermined small value is chosen to facilitate convergence.

Such “computational energy” forms have become common in neural networks, Hopfield networks [49] are a special case, as are polymatrix games, under certain conditions [95].

3.2. The Two-Labels Relaxation Labelling Paradigm. Iverson and Zucker [64, 66] have considered the special case known as *two-labels relaxation labelling*. It is also the one that we consider. Here, the set of nodes I is referred to as the network of hypotheses. The set of two labels is $\Lambda_i = \{\text{TRUE}, \text{FALSE}\}$. Either a hypothesis is true or it is false.

The representation of the two-labels’ confidence can be simplified because of the complementarity of the two labels. It is only necessary to explicitly represent one label.

$$p_i(\text{FALSE}) = 1 - p_i(\text{TRUE})$$

Similarly the representation of the two-labels’ support can be simplified if we impose the design condition that evidence for a hypothesis is evidence opposing the converse.

$$s_i(\text{FALSE}) = -s_i(\text{TRUE})$$

This design condition translates into the following condition on the structure of the compatibilities $r_{ij}(\lambda, \lambda')$:

$$r_{ij}(\text{TRUE}, \text{TRUE}) = -r_{ij}(\text{TRUE}, \text{FALSE}) = -r_{ij}(\text{FALSE}, \text{TRUE}) = r_{ij}(\text{FALSE}, \text{FALSE})$$

We take advantage of these simplifications to make our notation lighter. We choose to represent only the TRUE label and thus using

$$p_i \text{ to refer to } p_i(\text{TRUE})$$

$$s_i \text{ to refer to } s_i(\text{TRUE})$$

$$r_{ij} \text{ to refer to } r_{ij}(\text{TRUE}, \text{TRUE})$$

We can now rewrite the support as

$$s_i = \sum_{j \in I} r_{ij} p_j ,$$

and the update rule as

$$p_i^{n+1} = [p_i^n + \delta s_i]_0^1 ,$$

where the function $[x]_0^1$ is x truncated to the interval $[0, 1]$.

3.3. Implementation Remarks. The relaxation labelling method uses a gradient ascent procedure which inevitably terminates on a local maximum. We have to ensure that this local maximum is meaningful. We designed the calculation of the initial estimates and the derivation of the compatibilities with this in mind. For that purpose, the support s_i should only be positive when the features that are considered essential for a valid solution can be verified locally.

4. Coherent Shading Flow Field

We use relaxation labelling processes to refine the initial estimate. The key task of this idea is to define what is a consistent labelling for a shading flow field. To reach this objective, we use the local differential properties of the flow field as a constraint.

4.1. Using Curvatures as Constraints. Consider that the local description of the shading flow field comprises the orientation of the flow field and two curvatures: the curvature of the shading flow field and the curvature of its dual, the gradient field. These curvatures capture the local variation the shading flow orientation, thus they can be used as constraints.

Iverson [66] used these quantities to constrain texture flow fields and showed that the relaxation will interpolate a dense field even from sparse inputs without arbitrarily smoothing over discontinuities. Since this is precisely where we are aiming, we first consider these quantities to design the consistency relationship for the shading flow field. Iverson considered that, for texture, flow patterns can be viewed as resulting from a combination of translations or rotations. As a consequence, the texture flow field is modelled with both a divergence component and a curl component. This is not appropriate for shading flow field as the shading flow field reflects a differentiation with respect to a scalar field (the image intensities), and the curl component of the gradient is always null.

We can derive a different constraint based solely on the definition of these curvatures. Then, the variation of the orientation of the shading flow is given by

$$\frac{\partial \theta}{\partial r} = (\kappa_s \sin \vartheta + \kappa_g \cos \vartheta) ,$$

where the variables r and ϑ are simply the local frame coordinates expressed as polar coordinates.

The principal interest of this representation is that it provides a local description to the image intensity $I(x, y)$ which remains the same³ for the image irradiance $E(x, y)$.

³Here, we assume only that the relation between $I(x, y)$ and $E(x, y)$ is independent of the image position (x, y) and that this relation preserves the sign of directional derivative.

4.2. Using Differential Invariants as Constraints. The shading flow field's representation is richer than the texture flow field's. There is an extra dimension that is available to constrain the shading flow field. It is the magnitude of the gradient, $|\vec{\nabla}I|$.

PROPOSITION 5.1. *The shading flow field, augmented with the magnitude of the intensity gradient, can be locally described by $\Upsilon^s = \{x, y, \vec{S}, \mathbf{D} \mid (x, y) \in \mathbb{R}^2, \vec{S} \in \mathbb{R}^2, \mathbf{D} \in \mathbb{R}^3\}$.*

Proof: Consider a smooth shading flow field $\vec{S}(\vec{r})$. We develop a Taylor expansion of $\vec{S}(\vec{r})$ to obtain a local description and we keep only the first terms:

$$\vec{S}(\vec{r} + d\vec{r}) = \vec{S}(\vec{r}) + \frac{\partial \vec{S}(\vec{r})}{\partial \vec{r}} d\vec{r} + \dots$$

where

$$\frac{\partial \vec{S}(\vec{r})}{\partial \vec{r}} = \frac{1}{2} \text{curl } \vec{S} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \frac{1}{2} \text{div } \vec{S} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \text{def } \vec{S} R_\varphi^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} R_\varphi$$

Since the shading flow field is related to image intensity gradient by:

$$\vec{S}(x, y) = \begin{pmatrix} -I_y(x, y) \\ I_x(x, y) \end{pmatrix} = \begin{pmatrix} |\vec{\nabla}I(x, y)| \cos \theta(x, y) \\ |\vec{\nabla}I(x, y)| \sin \theta(x, y) \end{pmatrix},$$

it follows that

$$\text{div } \vec{S} = \text{curl } \vec{\nabla}I \equiv 0,$$

the trace of the variation of the flow field is null. Thus, three quantities suffice to describe it locally:

$$(5.1) \quad \vec{S}(x + \Delta x, y + \Delta y) \approx \vec{S}(x, y) + \begin{pmatrix} -I_{xy} & -I_{yy} \\ I_{xx} & I_{xy} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}.$$

The coefficients of this matrix relate to the three components of a DISTORTION vector as follows:

$$\mathbf{D} \equiv \begin{pmatrix} -\text{def } \vec{S} \sin(2\varphi) \\ \text{def } \vec{S} \cos(2\varphi) \\ \text{curl } \vec{S} \end{pmatrix} = \begin{pmatrix} I_{xx} - I_{yy} \\ -2I_{xy} \\ I_{xx} + I_{yy} \end{pmatrix}.$$

Figure 5.26 provides a geometric picture of the relationship between the shading flow orientation and the nature of the distortion. Combined with the shading flow vector, it completes the local shading flow description space, $\Upsilon^s = \{x, y, \theta, |\vec{\nabla}I|, \mathbf{D}\}$. $\square$

If we express the variation in terms of the coordinates defined by the tangent and the normal to the shading flow at $\vec{r}$, we can relate curvatures, κ_g and κ_s , and the differential invariants:

$$\frac{\partial \vec{S}(\vec{r})}{\partial \vec{r}} d\vec{r} = \begin{pmatrix} -\kappa_g & \kappa_c \\ -\kappa_s & \kappa_g \end{pmatrix} \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix},$$

where $\Delta u = \Delta x \cos \theta + \Delta y \sin \theta$, $\Delta v = -\Delta x \sin \theta + \Delta y \cos \theta$, and $\kappa_c = -2I_{xy}I_xI_y - I_{xx}I_x^2 - I_{yy}I_y^2$. This third term can not be accounted for with the model without magnitude.

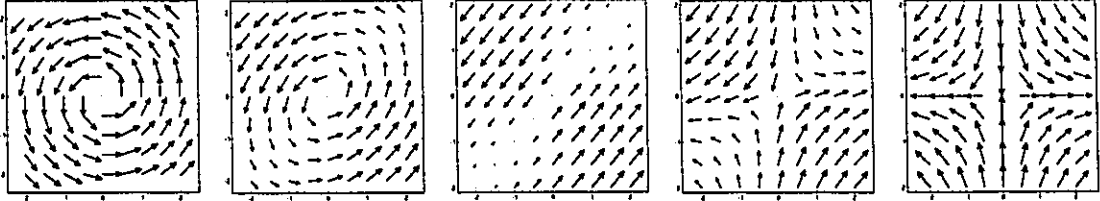


FIGURE 5.26. This sequence shows flow fields for which the distortion vector differs. The left most depicts a pure curl, $D_3 = 1$, the right most depicts a pure deformation $D_3 = 0$, and in between, the two components blend with different proportions. Whereas the curl is symmetric with respect to a point, the deformation is symmetric with respect to an axis. Note how the nature of the singularities differs as the distortion changes.

The shading flow field, augmented with the magnitude of the intensity gradient, is fully constrained. To appreciate this result, we can look at texture flow fields — they are naturally constrained with only two quantities (such as the curvatures κ_g , κ_s). Different models have been proposed to relate them with differential invariants, models that are exclusive of one another (e.g. pure rotation and expansion [66] or pure deformation [40]).

4.3. Compatibilities between Local Shading Flow Descriptions. So given two local shading flow descriptions, $\Upsilon_i^s = \{x_i, y_i, \theta_i, |\vec{\nabla}I|_i, \mathbf{D}_i\}$ and $\Upsilon_j^s = \{x_j, y_j, \theta_j, |\vec{\nabla}I|_j, \mathbf{D}_j\}$, we shall define a compatibility measure.

The transported shading flow description $\tilde{S}_i^s$ is derived from Eq. 5.1 with $\Delta \mathbf{x}$ equal to $\mathbf{x}_j - \mathbf{x}_i$ and the matrix coefficients derived directly from $\mathbf{D}_i$. The transported shading flow description $\tilde{S}_j^s$ is obtained in a similar way; $\Delta \mathbf{x}$ is equal to $\mathbf{x}_i - \mathbf{x}_j$ and the matrix coefficients are derived directly from $\mathbf{D}_j$.

The shading flow descriptors $\Upsilon_i^s = \{x_i, y_i, \theta_i, |\vec{\nabla}I|_i, \mathbf{D}_i\}$ constrain the shading flow descriptors $\Upsilon_j^s = \{x_j, y_j, \theta_j, |\vec{\nabla}I|_j, \mathbf{D}_j\}$, and reciprocally, the shading flow descriptors $\Upsilon_j^s = \{x_j, y_j, \theta_j, |\vec{\nabla}I|_j, \mathbf{D}_j\}$ constrain the shading flow descriptors $\Upsilon_i^s = \{x_i, y_i, \theta_i, |\vec{\nabla}I|_i, \mathbf{D}_i\}$. We require the compatibility to be positive only when both these constraints are satisfied. To achieve this, we use the logical/linear combinators introduced by Iverson and Zucker [65].

DEFINITION 5.7. The logical/linear combinator $\hat{\Delta}$ is given by:

$$x \hat{\Delta} y = \begin{cases} x + y, & \text{if } x > 0 \wedge y > 0 ; \\ y, & \text{if } x > 0 \wedge y \leq 0 ; \\ x, & \text{if } x \leq 0 \wedge y > 0 ; \\ x + y, & \text{if } x \leq 0 \wedge y \leq 0 . \end{cases}$$

The compatibility is then given by:

$$r_{ij} = r_{ij}^* \hat{\Delta} r_{ji}^*$$

where

$$\begin{aligned} r_{ij}^* &= G''(\theta_j - \theta_i^*) G \left(\left| \vec{\nabla} I_j \right| - \left| \vec{\nabla} I_i^* \right| \right) G (|\vec{x}_j - \vec{x}_i|) , \\ r_{ji}^* &= G''(\theta_i - \theta_j^*) G \left(\left| \vec{\nabla} I_i \right| - \left| \vec{\nabla} I_j^* \right| \right) G (|\vec{x}_i - \vec{x}_j|) . \end{aligned}$$

The orientation of the flow field is *strongly* constrained; i.e. the shading flow descriptions are said to be compatible only if the transported orientation of i falls within the Voronoi cell of the orientation j and vice-versa, the shading flow descriptions are said to be incompatible otherwise. On the other hand, the magnitude of the gradient is only *weakly* constrained; i.e. this component only modulates the strength of the compatibility, it does not affect the sign.

This compatibility function generalizes the co-circularity and the concentricity models [64]. Both these models project a center of curvature (which is a singularity for the flow field) in a direction perpendicular to the flow vector at a distance determined by the flow field's curvature. Similarly, the distortion vector projects a singularity, but the relation between the flow vector and the position of the singularity can be more complex because of the deformation component.

5. Numerical Results

In this section, we present a couple of examples for which this approach has been used to extract the shading flow field. These examples were chosen to illustrate that a piecewise smooth shading flow field can be inferred from images where such discontinuity arises.

5.1. Abrupt Albedo Change. The first example that we ran to test our algorithm was an image of a scene in which the albedo changed abruptly (Fig. 5.27(a)). It simply consists of a sphere in front of a plane.

We extracted the edge map using the algorithm of Iverson [66] and the shading flow field using the algorithm outlined in this chapter. These are shown in Fig. 5.27(c). If we compare

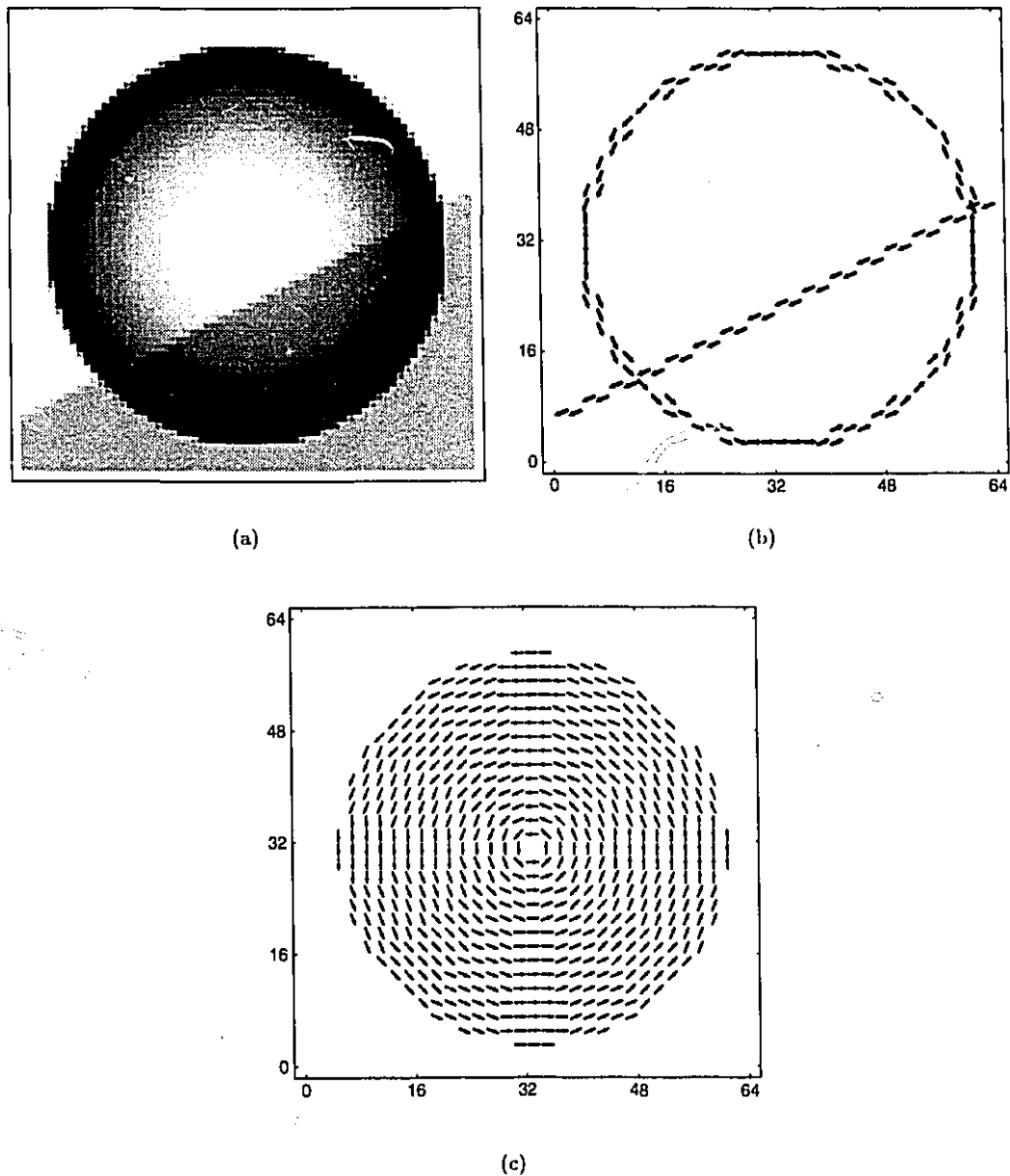


FIGURE 5.27. This test image provides an example of a shading flow field recovery. It was obtained after 5 iterations of the relaxation labelling algorithm proposed in this chapter. An ideal intensity image of a simple scene illuminated by a single distant point light source but with an abrupt change in albedo (a), the corresponding tangent field (b) and shading flow field (c) are shown in this figure. This shading flow field presents a point singularity of index one in the centre and a undefined region on the image's outskirts. For display purposes only, thresholds are used to avoid cluttering the figures with useless information and the resolution of the edge map and shading flow field are decreased by two.

this result with the shading flow field obtained with "fuzzy derivatives" (see Fig. 5.22), we note differences where the discontinuities occur. The relaxation labelling approach allowed for discontinuity in the shading flow field whereas the "fuzzy derivatives" approach does not.

5.2. Multiple Light Sources. The second example that we ran to test our algorithm was an image of a scene which is illuminated by two distant light sources (Fig. 5.28(a)). Again, it simply consists of a sphere in front of a plane.

Such as for the previous example, we extracted the edge map using the algorithm of Iverson [66] and the shading flow field using the algorithm outlined in this chapter. These are shown in Fig. 5.28(c). The only edges in this example are the curved edges that form a circle. There is an inversion of the edge's direction where the polarity of the image contrast changes. The most interesting feature in this example is the curve running across the sphere along which the shading flow field is also discontinuous. The shading flow field discontinuity is extracted even though there is no intensity discontinuity. The comparison of this result with the shading flow field obtained with "fuzzy derivatives" (see Fig. 5.23) shows again a difference where the discontinuities occur.

5.3. The Car's Fender. The third example that we ran to test our algorithm was a real image of a scene which is illuminated by two distant light sources (Fig. 5.29(a)). This time, the scene consists of a car. For the example, we use a close-up of the fender on which runs a shadow (Fig. 5.29(b)).

We show in Fig. 5.29(c), the edge map obtained using the algorithm of Iverson [66] and the shading flow field obtained using the algorithm described in this section. Even though the data is noisy, the shading flow field is usually smooth. The points where the shading flow field is not smooth correspond to curves marking discontinuities.

These results show that the shading flow field can be extracted from the image without necessarily destroying the information of discontinuities. The approach presented in this chapter produces piecewise smooth shading flow fields that preserve singularities and discontinuities.

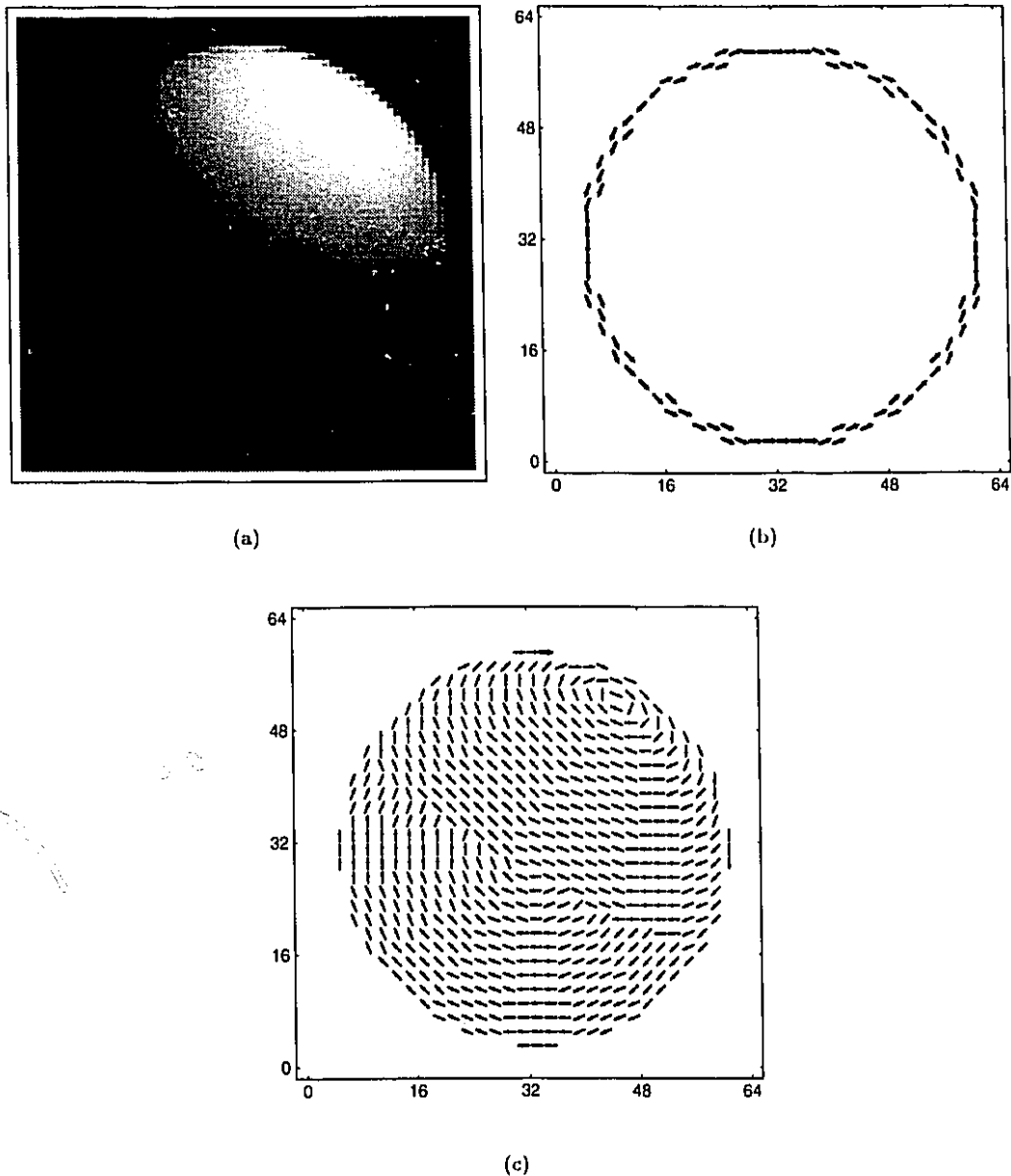


FIGURE 5.28. This test image provides an example of a shading flow field recovery. It was obtained after 5 iterations of the relaxation labelling algorithm proposed in this chapter. An ideal intensity image of a simple scene illuminated by two distant point light source (a), the corresponding tangent field (b) and shading flow field (c) are shown in this figure. Note the shading flow field presents again a point singularity of index one (near the centre of the upper right quadrant) and an undefined region. For display purposes only, thresholds are used to avoid cluttering the figures with useless information and the resolution of the edge map and shading flow field is decreased by two.

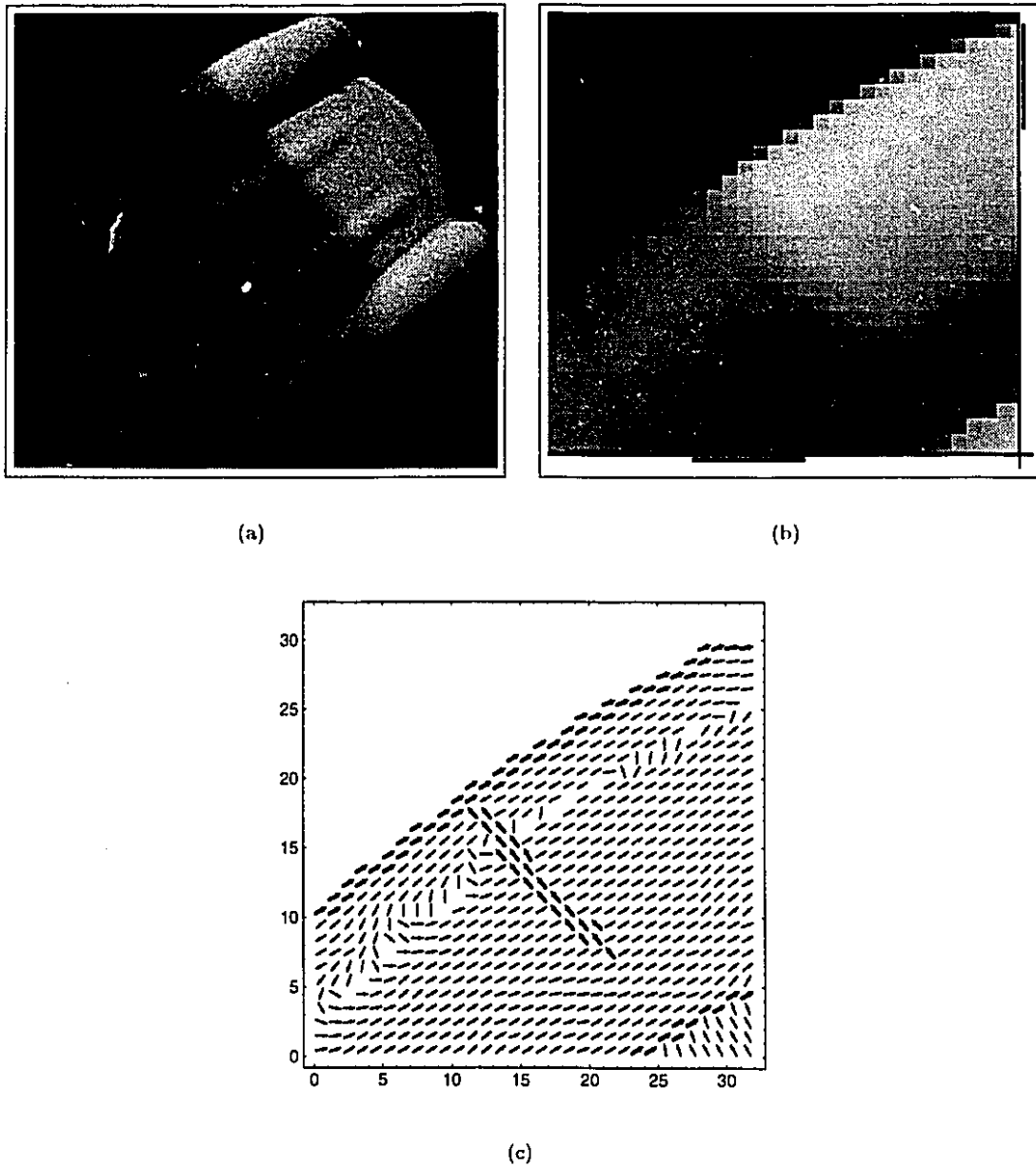


FIGURE 5.29. This test image provide an example of shape recovery for a real image. The image used (b) is a close-up of the fender of a car (a). The corresponding edge map and shading flow field (c) are used as input.

CHAPTER 6

Interpretations of Discontinuities

The general shape from shading problem introduces discontinuities. They occur in both the given initial data and the scene model. The relationships between these are at the root of the image interpretation.

Considering that any of the scene parameters (lighting condition, albedo, surface orientation) can change abruptly, can we determine which one did given the geometric properties of the image?

1. Various Cases of Scene Discontinuity

In this section, as a first step toward answering this question, we establish the signature of each scene discontinuity for generic surfaces with a generic illuminant direction and from a generic viewpoint¹

Before looking at the various cases of scene discontinuity, we begin by considering the situation when there is none. The image irradiance depends on the albedo, the lighting condition, and the surface orientation:

$$E = \rho \lambda \mathbf{L} \cdot \mathbf{N} ;$$

and the orientation of the shading flow field depends on the lighting condition, and the surface orientation:

$$\tan \theta = \frac{\mathbf{L} \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L} \cdot \frac{\partial}{\partial x} \mathbf{N}} .$$

We can derive directly an important first result.

OBSERVATION 6.1. *A smooth matte surface of constant albedo projects, under constant illumination, a smooth image irradiance and almost everywhere, a smooth shading flow field.*

¹The assumptions of "generic view" and "generic illuminant direction" states that the observer and the light sources are not in a special position relative to the scene. A small change in their position should not affect the qualitative aspect of the image [9,39].

Discontinuity of the shading flow field can occur at singularities of the direction field, that is, where the image irradiance gradient is null. Although both point and line singularities are of measure zero with respect to the domain of the shading flow field, our analysis requires that we take them into consideration.

OBSERVATION 6.2. *Now, let's consider the shading flow field of a smooth matte surface of constant albedo under constant illumination, but augmented with the gradient magnitude. There exists a representation (the Cartesian coordinate) in which the resulting vector field is continuous and differentiable everywhere.*

1.1. Shadows. Shadows occur when some light source does not shine on part of the scene. Such an event can be explained as an occlusion of the light source by an opaque object. The boundaries of shadow correspond to where the illumination abruptly changes.

Let $\lambda_{(i)} \mathbf{L}_{(i)}$ describe the illumination and the direction of light sources indexed by i . Consider a shadow caused by the occlusion of a point light source $\lambda_{(0)} \mathbf{L}_{(0)}$.

The linearity of Lambert's model allows the definition of an equivalent light source (see the illumination model in Chapt. 7) as the sum of the light source visible from the surface. Hence, the lighting condition on each side of the shadow boundary can be described by $\lambda \mathbf{L}$ and $\lambda_s \mathbf{L}_s$, the latter corresponding to the shadow side:

$$\begin{aligned}\lambda \mathbf{L} &= \sum_i \lambda_{(i)} \mathbf{L}_{(i)} , \\ \lambda_s \mathbf{L}_s &= \sum_{i, i \neq 0} \lambda_{(i)} \mathbf{L}_{(i)} ,\end{aligned}$$

hence this simple relationship:

$$\lambda_s \mathbf{L}_s = \lambda \mathbf{L} - \lambda_{(0)} \mathbf{L}_{(0)} .$$

At a point on the boundary, the onset of the occlusion can be caused either by a possibly distant opaque object or by the surface itself slanting away from the light source. We call these two cases respectively, cast shadow boundary and attached shadow boundary. These account for different image structures.

1.1.1. The Cast Shadow Boundary. We examine the behaviour of image properties where cast shadow boundaries occur. The first quantity that we consider is the image irradiance. Is it continuous? The difference in image irradiance as we enter in the shadow is given by:

$$E - E_s = \rho \lambda_{(0)} \mathbf{L}_{(0)} \cdot \mathbf{N} .$$

For a cast shadow, $\mathbf{L}_{(0)} \cdot \mathbf{N} > 0$, hence $E_s < E$. The result corresponds to our intuition.

OBSERVATION 6.3. *A cast shadow boundary is marked by a discontinuity in the image irradiance².*

The second quantity that we consider is the shading flow orientation. Again, the focus is on continuity. The shading flow orientations, on the respective sides of a cast shadow boundary, are given by:

$$\tan \theta = \frac{\mathbf{L} \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L} \cdot \frac{\partial}{\partial x} \mathbf{N}},$$

$$\tan \theta_s = \frac{\mathbf{L}_s \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L}_s \cdot \frac{\partial}{\partial x} \mathbf{N}}.$$

In general, the shading flow orientation differs from each side of the shadow boundary.

OBSERVATION 6.4. *A cast shadow boundary is generically marked by a discontinuity in the shading flow field.*

One should note that in an image of a scene, coincidental alignments are likely to occur at some points along the boundaries.

1.1.2. *The Attached Shadow Boundary.* Now, we examine the behaviour of image properties where attached shadow boundaries occur. We first determine whether or not, the image irradiance is continuous. The difference in image irradiance as we enter in the shadow is given by:

$$E - E_s = \rho \lambda_{(0)} \mathbf{L}_{(0)} \cdot \mathbf{N}.$$

Even though there is a marked decrease in surface illumination, the image irradiance is continuous since $\mathbf{L}_{(0)} \cdot \mathbf{N} = 0$; the incoming rays from the occluded source only graze the surface, hence $E_s = E$.

OBSERVATION 6.5. *When an attached shadow boundary occurs on a smooth surface, it is marked by continuous image irradiance.*

Again, the second quantity that we consider is the shading flow orientation, and focus is still on continuity. The shading flow orientations, on the respective sides of an attached

²Strictly speaking, a discontinuity occurs only if the hidden light is a point source and if there is no atmospheric dispersion (see App. D). The size of the image curve detectors defines a notion of scale. This scale relaxes the previously stated condition to allow light sources which are somewhat spatially extended and some limited amount of blurring due to atmospheric dispersion.

shadow boundary, are given by:

$$\tan \theta = \frac{\mathbf{L} \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L} \cdot \frac{\partial}{\partial x} \mathbf{N}},$$

$$\tan \theta_s = \frac{\mathbf{L}_s \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L}_s \cdot \frac{\partial}{\partial x} \mathbf{N}}.$$

Just as for the cast shadow boundary case, in general the shading flow orientation differs on each side of the shadow boundary. Again, one should note that in an image of a scene, coincidental alignments can occur at some isolated points along the boundaries.

OBSERVATION 6.6. *An attached shadow boundary is generically marked by a discontinuity in the shading flow field.*

Thus, the signature of the cast and the attached shadow boundaries differ by their image irradiance discontinuity.

1.2. Change in Reflectivity. Changes in reflectivity occur when the material properties change. If we consider a surface for which the reflectivity changes abruptly, the change will occur along a curve. Let ρ_l and ρ_d be respectively the albedo on the lighter and darker side of that curve.

Now, we examine the behaviour of the image irradiance on both sides of the curve where the reflectivity changes. If the image irradiance on the lighter side is denoted as E_l , then the image irradiance on the darker side is given by:

$$E_d = \frac{\rho_d}{\rho_l} E_l.$$

If the surface is illuminated, then $E_d < E_l$ since $\rho_d < \rho_l$. Hence, another confirmation of an intuitive result.

OBSERVATION 6.7. *A change in reflectivity is marked by a discontinuity in the image irradiance.*

The orientation of the image irradiance gradient has a different behaviour. Since it is independent of the albedo,

$$\tan \theta_d = \tan \theta_l = \frac{\mathbf{L} \cdot \frac{\partial}{\partial y} \mathbf{N}}{\mathbf{L} \cdot \frac{\partial}{\partial x} \mathbf{N}},$$

the orientation of the image irradiance gradient is continuous where only the reflectivity changes.

OBSERVATION 6.8. *A change in reflectivity is marked by a continuous shading flow field.*

1.3. Occlusions and Other Surface Discontinuities. The difference in the image irradiance at surface discontinuities is given by the following:

$$E - E' = \rho \lambda \mathbf{L} \cdot (\mathbf{N} - \mathbf{N}') .$$

The surface orientations on each side of a surface discontinuity ($\mathbf{N}$, $\mathbf{N}'$) are not related, hence $\mathbf{N} - \mathbf{N}'$ is usually not equal to zero. Since the illumination direction is not related to $\mathbf{N} - \mathbf{N}'$, the dot product $\mathbf{L} \cdot (\mathbf{N} - \mathbf{N}')$ is also usually not equal to zero.

OBSERVATION 6.9. Surface discontinuities are generically marked by discontinuities in the image irradiance.

Note that the occurrence of a point, somewhere along an occlusion, for which $\mathbf{L} \cdot (\mathbf{N} - \mathbf{N}') = 0$ is expected. For a generic situation, this point marks an inversion of contrast; Fig. 5.28 provides two examples of such points.

For the shading flow field, the consideration that the surface orientations (and its derivatives) on each side of a surface discontinuity are not related, yields the following observation.

OBSERVATION 6.10. Surface discontinuities are usually marked by discontinuities in the shading flow field.

Surface discontinuities occur when one surface partially occludes another. In such a case, when the albedos of the two surfaces are different, a change in albedo occurs at the same image location as the occlusion. The same example can be reworked for lighting conditions; in this case, the front surface could occlude some light sources from the back surface. Again, a change in lighting condition occurs at the same image location as the occlusion.

OBSERVATION 6.11. At surface discontinuities, changes in albedo and lighting conditions often occur. These are not due to a coincidental alignment, these are rather due to a common cause.

2. Summary of Results

The results obtained for generic surfaces are summarized in Table 6.3 (printed in bold-face). By bringing the previous observations together, we get the beginning of an answer to our question. At least in some circumstances, we can determine which scene parameters change abruptly from the geometric properties of the image.

disc. int.	disc. $\vec{S}$	probable interpretation			characteristic
		surface	albedo	lighting	
no	no	smooth	constant	constant	
		smooth parabolic	constant	shadow bound- ary (attached)	parallel flow field
no	yes	smooth	constant	shadow bound- ary (attached)	
yes	no	smooth	abrupt change	constant	
		smooth parabolic	constant	shadow bound- ary (cast)	parallel flow field
		aligned parabolic surfaces junction	—	—	parallel flow field
yes	yes	smooth	constant	shadow bound- ary (cast)	
		discontinuous	—	—	

TABLE 6.3. This table shows the relations between image discontinuities and scene discontinuities. The shading flow field $\vec{S}(x, y)$ is perpendicular to the normalized gradient intensity $\frac{\vec{\nabla}I}{|\vec{\nabla}I|}$.

OBSERVATION 6.12. *The different discontinuity signatures allow us to distinguish occluding boundaries from abrupt albedo changes or from attached shadow boundaries.*

For generic surfaces, there is only one ambiguity that cannot be resolved prior to shape estimation. Discontinuous surfaces and cast shadow boundaries on a smooth surface cannot be distinguished. The interpretation is nevertheless trivial once the surface shape is known (which happens in stage 2 of our process, see Chapt. 9).

Non-generic surfaces such as parabolic surfaces³ or planar surfaces introduce more ambiguities. They have however characteristic shading flow fields that allows us to sort them out.

Parabolic surfaces produce parallel shading flow fields and, furthermore, the orientation of the field is only dependent on the geometric aspect of the surface: this leads to the following consequences.

- Although the attached shadow boundary no longer exhibits its shading flow field discontinuity, the magnitude of the gradient remains discontinuous along a line parallel to the flow. It is thus possible to distinguish it from a smooth surface of uniform albedo under constant illumination condition.
- The cast shadow boundary and discontinuous aligned parabolic surfaces no longer exhibit a discontinuity in the shading flow field. But again the magnitude of the gradient remains discontinuous along a line, this time parallel to intensity discontinuity which for the cast shadow boundary's case is not necessarily parallel to the shading flow field.

These results are also summarized in Table 6.3.

Planar surfaces always produce null shading flow fields, hence they also generate ambiguities. It is not possible to distinguish between an abrupt change in albedo, a shadow boundary⁴, and a surface discontinuity where only the surface orientation changes.

The ambiguities introduced by parabolic surfaces do not completely prevent the interpretation of the geometric structures of the image. There is always some characteristic shading flow field where these ambiguities occur. Our analysis can be carried everywhere else.

³A surface is said to be *parabolic* if one of its principal curvatures is null; a surface is said to be *planar* if both principal curvatures are null.

⁴Only the cast shadow boundary can occur on a flat surface, an attached shadow boundary involves a surface curving away from a light source.

disc. $ \vec{\nabla}I $	disc. int.	disc. $\vec{S}$	probable interpretation			characteristic
			surface	albedo	lighting	
no	no	no	smooth	constant	constant	
yes	no	no	smooth parabolic	constant	shadow bound- ary (attached)	parallel flow field

TABLE 6.4. This table shows other relations between image discontinuities and scene discontinuities.

At least one such ambiguity can be resolved. An attached shadow boundary on a smooth parabolic surface of constant albedo does not demarcate itself from a smooth surface of constant albedo under constant lighting condition, given the image geometrical structures identified thus far. We examine the behaviour of the image irradiance gradient magnitude near an attached shadow boundary. It is given by:

$$\begin{aligned} |\vec{\nabla}E|_s^2 = |\vec{\nabla}E|^2 - \rho^2 & \left(\left((2\lambda\mathbf{L} - \lambda_{(0)}\mathbf{L}_{(0)}) \cdot \frac{\partial}{\partial x}\mathbf{N} \right) \left(\lambda_{(0)}\mathbf{L}_{(0)} \cdot \frac{\partial}{\partial x}\mathbf{N} \right) \right. \\ & \left. + \left((2\lambda\mathbf{L} - \lambda_{(0)}\mathbf{L}_{(0)}) \cdot \frac{\partial}{\partial y}\mathbf{N} \right) \left(\lambda_{(0)}\mathbf{L}_{(0)} \cdot \frac{\partial}{\partial y}\mathbf{N} \right) \right). \end{aligned}$$

Since $\lambda\mathbf{L} \geq \lambda_{(0)}\mathbf{L}_{(0)}$ and $\lambda_{(0)}\mathbf{L}_{(0)} > 0$, the magnitude of the image irradiance differs from each side of the shadow boundary.

OBSERVATION 6.13. *An attached shadow boundary is marked by a discontinuity in the image irradiance gradient magnitude (even for a parabolic surface).*

The discontinuity of the gradient magnitude lifts the ambiguity as shown in Table 6.4. Image curves capturing this type of discontinuity can also be extracted from the image using logical/linear operators [66]).

We conclude this chapter by stressing the usefulness of the geometrical structures of the image — the shading flow field and the tangent field used together, not only on their own — for the understanding of the scene. A lot of interpretation can be carried out prior to any shape from shading computation.

Part 3

THE SCENEL BUNDLE APPROACH

The Scenel Bundle Approach

Having developed the notion of a shading flow field and discussed several of its properties, we now present an approach to inferring shape from it. The approach is motivated by modern notions of fibre bundles in differential geometry [120]. It provides a solution to the generalized shape from shading problem that is posed as a coupled collection of “local” problems, the solution to each of which is that local scene element (or scenel¹) that captures the local image properties, and which are then coupled together to form global piecewise smooth solutions.

In this chapter, we present an outline of our generalized shape from shading approach and then we briefly introduce fibre bundles. We then establish the relationship between fibre bundles and the geometric construction on which our approach is based. The structure of fibre bundles is used to provide a clear picture of the scene’s model space and its characteristics.

1. The Outline of the Scenel Bundle Approach

We are essentially considering the generalized shape from shading problem as a coupled family of local problems. Given the formal specification of a local scene element, our approach has two requirements:

- (i) a mechanism for inferring the scene element from the image (or more precisely, from the image geometric structures);
- (ii) a mechanism for the local to global transition, i.e. for sewing the local scene patches together in a consistent fashion.

We formulate a geometric construction to structure these mechanisms and we illustrate it in the Figs. 7.30, 7.31, 7.32, and 7.33.

The local scene model comprises a local surface model and a local lighting model. These models interact to produce the geometric structures of the image (see Fig. 7.30). This relationship provides the basis for the local shape from shading inference.

¹*cf.* Pixel, voxel, ... scenel.

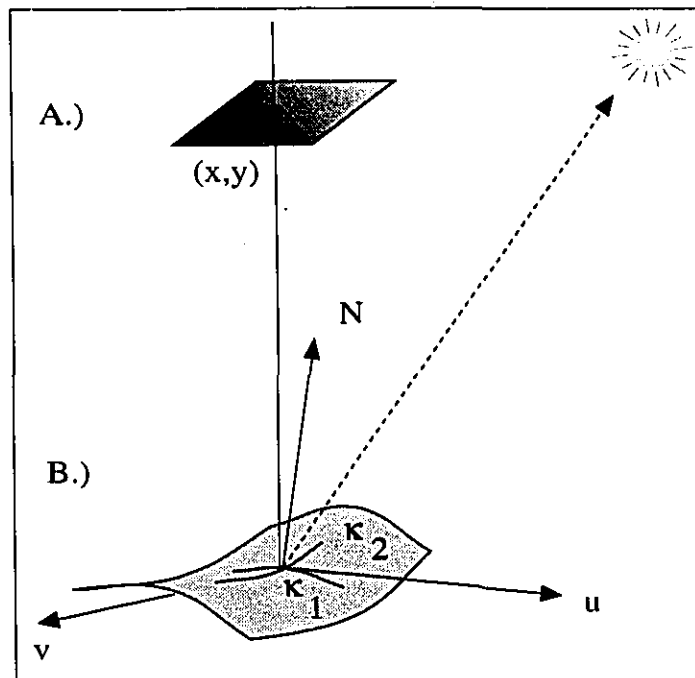


FIGURE 7.30. Depiction of an abstract scene element, or scenel, corresponding to an image patch (A). The scenel (B) consists of a surface patch, described by its image coordinates, surface normal, and curvature. Its material properties (albedo) are also represented. Finally, a virtual light source completes the photometry.

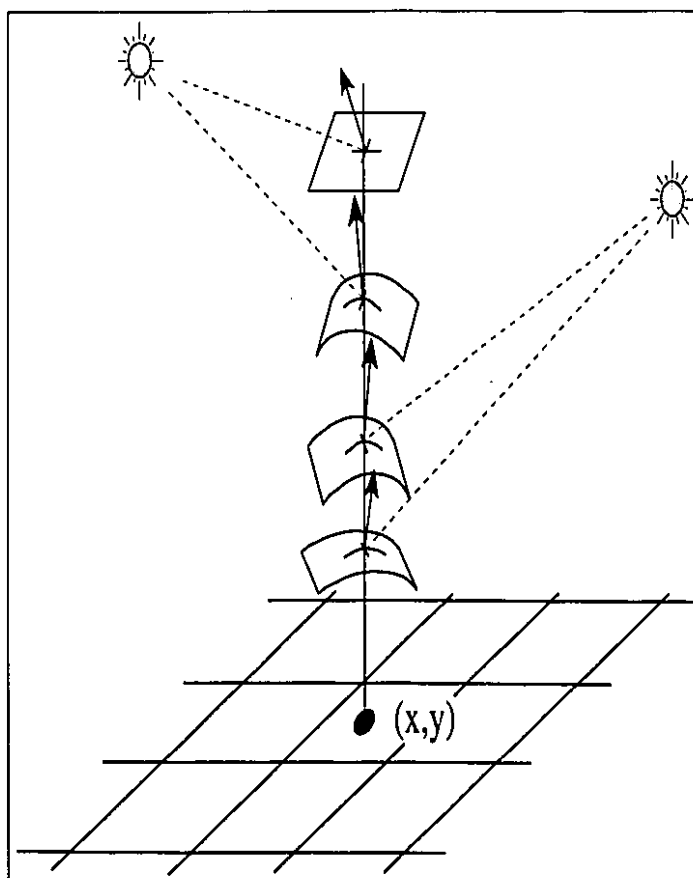


FIGURE 7.31. Depiction of a Scenel Fibre over a pixel of the image. At each point in the image there are many possible scene elements, or scenels. Each of these scenels is depicted along a fibre, or vertical space above each image coordinate. We tag a confidence measure to every scenel. This confidence measure indicates how well the scenel matches the local geometric properties of the image.

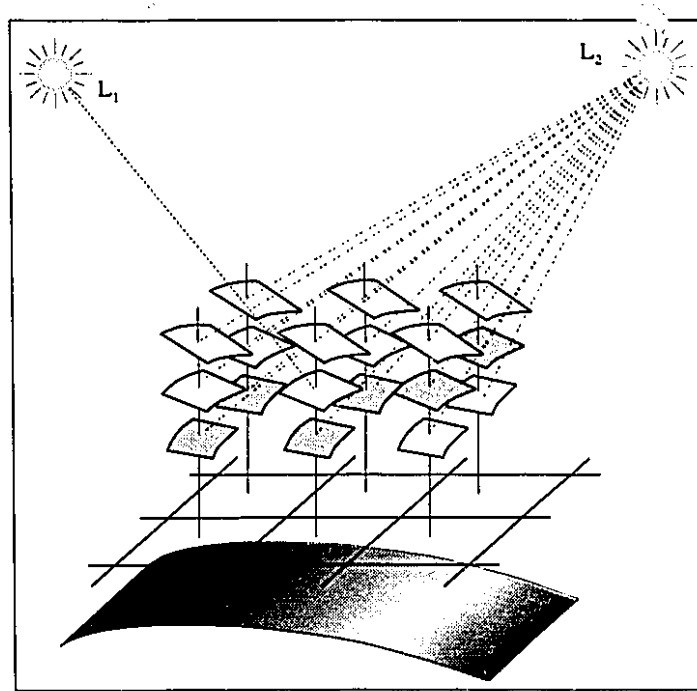


FIGURE 7.32. Depiction of a Scenel Bundle over an image. The union of scenel fibres over the entire image is called a scenel bundle. The shape from shading problem is formulated as determining sections through the scenel bundle. Such a section is depicted by the shaded scenels, and represents a "horizontal" cross-section across the bundle. Scenel participation in a horizontal section is governed by surface smoothness and material and light source constancy constraints.

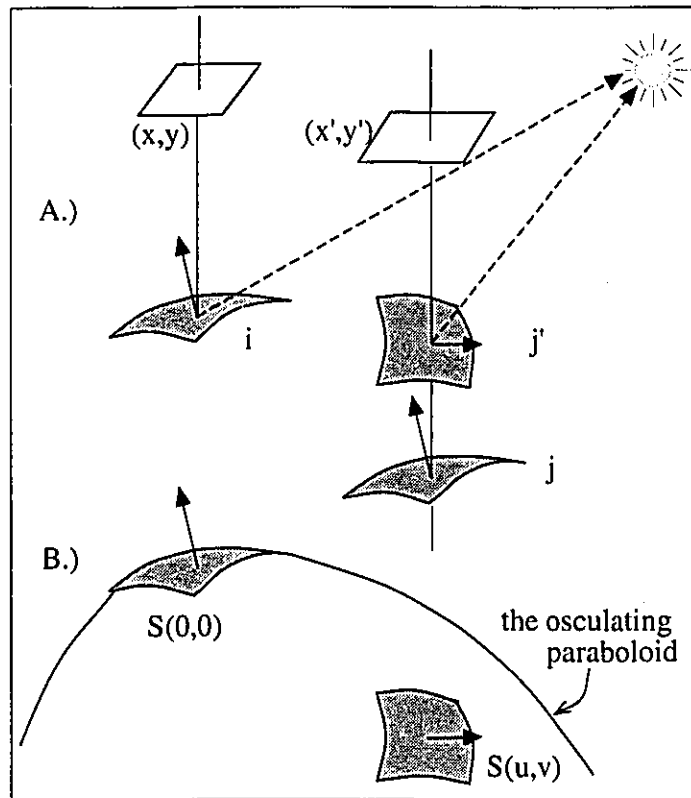


FIGURE 7.33. Illustration of the compatibility relationship for scenel consistency. Two scenels are shown on the fibre at image location (x', y') , and are evaluated against the scenel (i) at (x, y) . The surface represented in $\text{scenel}_{x,y}$ is modeled by the osculating paraboloid, and extended to (x', y') . It is now clear that one scenel (j') at (x', y') is consistent, because its surface patch lies on this paraboloid and light source and albedo agree. The other scenel (j) is inconsistent, because its surface does not match the extended paraboloid. Such osculating paraboloids are used to simulate the parallel transport of $\text{scenel}_{x',y'}$ onto $\text{scenel}_{x,y}$.

Typically, the local geometric structures from the image are consistent with an equivalence class of local scene models. This class of possible scene models becomes the building block of a geometric construction (see Fig. 7.31).

The next step is to consider the interaction between the scene models at different locations. For each pixel location, a class of local scene models is inferred from the geometric structures of the image (see Fig. 7.32). The interactions are the local constraints that we have described in Chapt. 4.

The coupling between the local scene models dictates a consistency relationship over them (see Fig. 7.33). This consistency relationship derives from two principle considerations:

- (i) A SURFACE SMOOTHNESS CONSTRAINT, which states that the surface normal and curvatures must vary according to a Lipschitz condition between pairs of scenels which project to neighbouring points in the image domain. This notion is subtle to implement, because it involves comparison of normal vectors following parallel transport to the proper position (see [117]).
- (ii) A LIGHT SOURCE CONSTRAINT, which states that the virtual light source is constant almost everywhere for pairs of scenels which project to neighbouring points in the image domain.

OBSERVATION 7.1. *The nature of the consistency relationships is purely geometric.*

The consistency relationships do not depend on the image, they depend only on the pairs of scenels.

1.1. Informal Description of Scenel Bundles. While the fibre bundle construction is quite abstract, we use it in the manner shown in Figs. 7.30, 7.31 and 7.32. This formalism offers a natural association between cross-sections of a bundle and the problem of shape inference. It provides a powerful descriptive language to deal with the relationships between the base and fibre spaces.

We take the image manifold as the base space and consider the photometry for each point on it. In a small neighbourhood around the point (x, y) , the shading information can be described by a combination of illumination model and surface model. Each of these defines a scenel (Fig. 7.30) and the space of all possible scenels defines the fibre over that point (Fig. 7.31). Together, the collection of scenel fibres defines the scenel bundle (Fig. 7.32).

The reason we introduce the bundle structure on scenels is to develop its topological, as well as its geometrical structures.

OBSERVATION 7.2. *The scenel fibre must present a topology such that continuous surface changes can be expressed by a continuous path in the fibre space.*

We seek a solution of the shape from shading problem as connected sets of scenels in which neighbours are consistent; such a solution is called a CROSS-SECTION through the scenel bundle. Formally, a cross-section assigns a member of each fibre to each position in the manifold.

OBSERVATION 7.3. *Since the surface of a solid is oriented, a consistent cross-section of the scenel bundle must also be oriented.*

A sub-bundle of the scenel bundle is the TANGENT BUNDLE, in which the fibres consist of the tangent spaces at each point and the sections correspond to vector fields; Sander and Zucker [117] previously used this bundle in their study of inferring principal direction fields on surfaces.

2. The Local Representation of the Scene

We want the *scene element* or *scenel* to be a local representation of the scene which comprises an illumination model and a surface model. In this section, we describe the information captured by the illumination model and the surface model.

2.1. The Illumination Model. A model of the illumination could provide, at every point, the brightness of every incoming ray. The spectral composition of each ray could also be described. There are several other properties of light that could also be included, for example polarity, coherence... Such a detailed model is not always necessary.

Restricting our shading analysis to the matte component of the reflection allows us to use a simple local illumination model.

ASSUMPTION 7.1. *Lambert's reflectance law holds for each scenel patch.*

We can limit our consideration to only two attributes:

- (i) **ILLUMINATION:** Denoted by λ , the illumination is the scalar quantity that indicates the amount of incoming light.
- (ii) **ILLUMINANT DIRECTION:** Denoted by $\mathbf{L}$, the illuminant direction is the unit vector that indicates from which direction the light is coming.

These two attributes are quite simple to visualize when a single distant point light source illuminates the surface patch. The model's attributes, the illumination and the illuminant

direction, correspond respectively to the illumination and the direction of a point light source. But what if several distant point sources illuminate a single surface patch?

DEFINITION 7.1. *The virtual illuminant is an imaginary distant point light source that would provide the same surface patch irradiance as the actual light sources. It comprises two attributes, λ and $\mathbf{L}$, that can be derived considering the linear property of Lambert's reflectance function.*

Consider a set of M distant point sources that illuminate the surface patch. Let this set be described by $\{\lambda_{(i)} \mathbf{L}_{(i)} : 1 \leq i \leq M\}$ where $\lambda_{(i)}$ and $\mathbf{L}_{(i)}$ are respectively the intensity and the direction of the individual point light source. The attributes of the virtual illuminant are given by:

$$\lambda \mathbf{L} \equiv \sum_i \lambda_{(i)} \mathbf{L}_{(i)} ,$$

Now consider extended light sources illuminating the surface patch. Let $\Lambda(\mathbf{L})$ denote the brightness of the light ray that is incident from direction $\mathbf{L}$. Using the linear property of Lambert's reflectance function to integrate the light rays coming from all directions in the visible hemisphere $\mathcal{H}$, we obtain:

$$\lambda \mathbf{L} \equiv \frac{1}{\pi} \int_{\mathcal{H}} \Lambda(\mathbf{L}) \mathbf{L}(\Omega) d\Omega .$$

The virtual illuminant thus provides a very convenient representation of the illumination as it allows complex illumination to be described simply by a scalar and a unit vector. It is a much more general model than that of a simple infinitely distant point source.

2.2. The Surface Model. The surface model provides a local description of what reflects the light. We consider two distinct attributes of the surface model:

- (i) **MATERIAL PROPERTIES:** The nature of the reflecting surface is what is meant to be described here. We have chosen to only consider the matte component of the reflection, thus the only quantity that we need is a scalar: the albedo ρ . Note that we have also chosen to ignore the spectral composition of the light since our input is a gray-level image. We would otherwise consider one scalar per colour channel of our sensor.
- (ii) **SURFACE SHAPE DESCRIPTORS:** We use the first and second fundamental forms' coefficients to describe surface patches. The two principal curvatures (κ_1, κ_2) describe the shape up to rotation. Two angles, slant σ and tilt τ , are needed to describe the surface tangent plane orientation with respect to the viewer's coordinate frame. An additional angle ϕ is needed to describe the principal direction of the Darboux

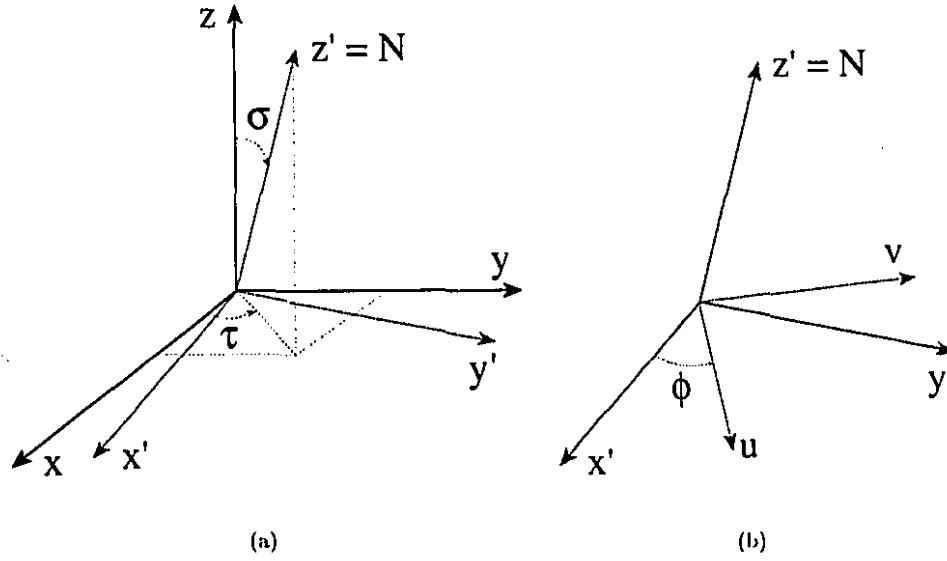


FIGURE 7.34. The world coordinates define the axes (x, y, z) where x - y is the image plane. We express the surface normal as a function of its tilt τ and slant σ : $N = (\sin \tau \sin \sigma, \cos \tau \sin \sigma, \cos \sigma)$. (a) The rotation of the frame from the z axis to the normal N defines a local coordinate system (x', y', z') where x' - y' is the surface tangent plane. (b) The principal frame defines a different coordinate system in the surface tangent plane: u - v . The rotation between these two coordinate systems defines the angle ϕ .

frame in the surface tangent plane (See Fig. 7.34). Other terms could be included to enriched this local description. In App. B, we examine such a possibility by considering the local spatial variation of the principal curvatures and the principal directions.

We seek a surface representation with the following features:

- any continuous change of the surface can be mapped by a continuous change of the surface model parameters,
- any continuous change of the surface model parameters can be mapped by a continuous change of the surface.

The representation of the surface orientation does not pose any challenge. The two angles needed to orient the surface tangent plane in space describe the surface normal

$$\mathbf{N} = (N_x, N_y, N_z) = (\cos \tau \sin \sigma, \sin \tau \sin \sigma, \cos \sigma) ,$$

where σ is the slant and τ is the tilt (Fig. 7.35). Continuous changes of the surface orientation correspond to continuous changes of the surface normal $\mathbf{N}$ and vice-versa.

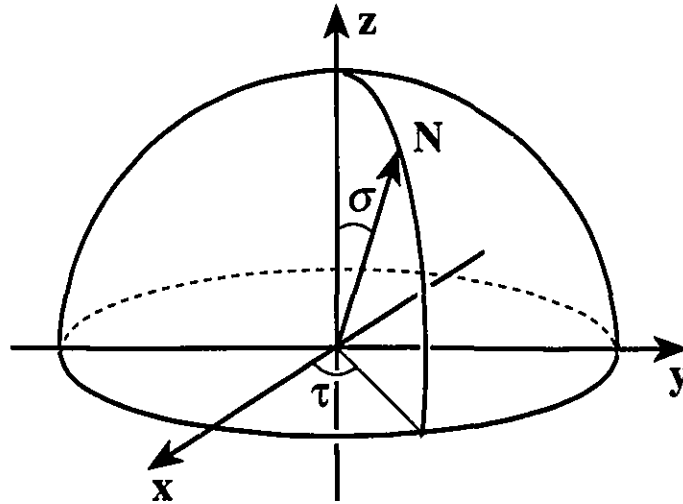


FIGURE 7.35. The normal vector $\mathbf{N}$ is defined in terms of two angles: the slant σ and the tilt τ . Since only visible surface patches are represented, the normal lies on the unit hemisphere facing the viewer.

The representation of the surface shape is not as straightforward. The geometry suggests considering the Darboux frame representation, but this has difficulties with singularities. Moreover, for umbilic surfaces², the principal directions are undefined. Consider the following example:

EXAMPLE 7.1. *Consider the surface of a paraboloid described by*

$$z = -\frac{1}{2} (\kappa_1 x^2 + \kappa_2 y^2)$$

where $\kappa_1 > \kappa_2$. The parameter characterizing the principal direction, the angle ϕ , takes the value 0.

Moreover, if one smoothly changes the surface such that only principal curvature κ_2 varies, say κ_2 increases, there is an abrupt change in the value of ϕ as κ_2 becomes greater than κ_1 . The angle ϕ suddenly takes the value $\frac{\pi}{2}$.

In light of these undesirable characteristics, we reject the Darboux frame representation in its classical form.

We have designed a new shape representation that has the desired features. We show here how it relates to the Darboux frame parameters κ_1, κ_2, ϕ . The two principal curvatures (κ_1, κ_2) are mapped into a curvedness measure

$$c = \max(|\kappa_1|, |\kappa_2|)$$

and a shape index measure

$$s = \cos^{-1} \left(\frac{\kappa_1 + \kappa_2}{2c} \right) .$$

These are analogous to Koenderink's curvedness and shape index [70], with the choice of norm for the curvedness and the spreading function for the shape index modified slightly.

OBSERVATION 7.4. *The angles 2ϕ and s are, respectively, the longitude and latitude of a spherical coordinate system covering shape variation in the tangent plane. As we stated previously, the angle ϕ represents the principal direction, while s values 0 and π represent umbilic surfaces where principal directions are not defined.*

DEFINITION 7.2. *The shape vector is a vector designed as follows:*

$$\mathbf{K} = (K_1, K_2, K_3) = (c \cos(2\phi) \sin s, c \sin(2\phi) \sin s, c \cos s) ,$$

where s is the shape index and ϕ is the angle giving the orientation of the principal directions.

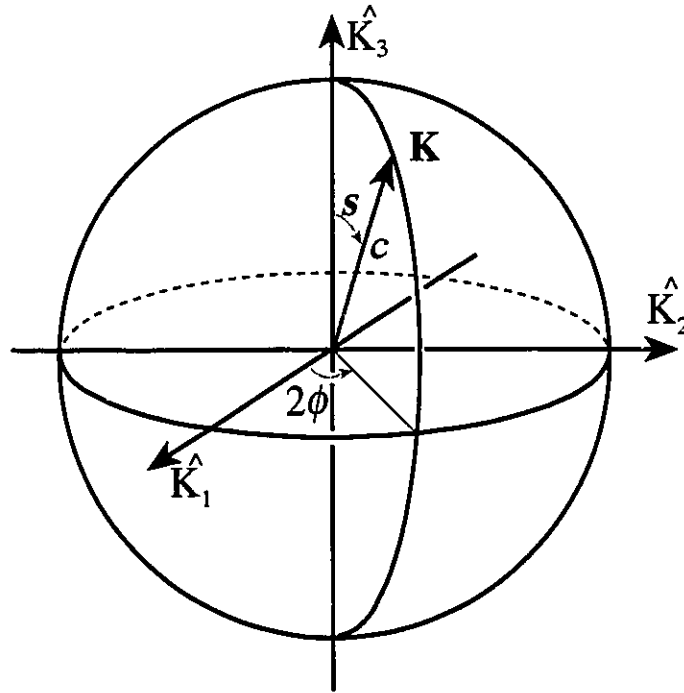


FIGURE 7.36. The shape of a surface patch is represented by $\mathbf{K}$, the “shape” vector. This vector lies in a three-dimensional space, $\mathbb{R}^3$. Its direction can be defined in terms of two angles: ϕ giving the orientation of the principal directions in the tangent plane and $s = \cos^{-1}(\frac{\kappa_1 + \kappa_2}{2c})$ giving a shape index. The length of the vector can be defined as the curvedness $c = \max(|\kappa_1|, |\kappa_2|)$. The continuity of this representation is its main interest — it is continuous even at umbilical points where the principal directions are undefined. A convex umbilical point will be characterized by $\mathbf{K} = (0, 0, c)$; a concave umbilical point, by $\mathbf{K} = (0, 0, -c)$; and a planar point, by $\mathbf{K} = (0, 0, 0)$.

This representation is interesting since nearby positions in the $\mathbf{K}$ space indicate smooth changes in either the shape or orientation of a scenel and conversely, smooth changes in either the shape or orientation of a scenel correspond to nearby positions in the $\mathbf{K}$ space. Even though the principal directions are undefined for umbilic surfaces (thus, so is ϕ), these surfaces do not cause any singularity in $\mathbf{K}$ space as they are simply represented by the shape vector $\mathbf{K} = (0, 0, \pm c)$.

Another interesting characteristic of the $\mathbf{N}$ - $\mathbf{K}$ representation is that complementary objects, whose surfaces match the solid as the mold of an object, would be oriented in the opposite direction. Such an object would also have opposite curvatures. The surface representation would thus be $(\mathbf{N}', \mathbf{K}') = -(\mathbf{N}, \mathbf{K})$.

EXAMPLE 7.2. Consider a sphere. It presents an umbilic convex surface patch facing the viewer that can be described by $\mathbf{N} = (0, 0, 1)$, $\mathbf{K} = (0, 0, 1)$. The mold of this surface would be a spherical cavity. So the corresponding complementary patch would be an umbilic concave patch pointing away from the viewer. This patch could be described by $\mathbf{N} = (0, 0, -1)$, $\mathbf{K} = (0, 0, -1)$.

EXAMPLE 7.3. Now consider a flat surface with an arbitrary orientation. Any patch on such a surface can be described by $\mathbf{N} = (N_x, N_y, N_z)$, $\mathbf{K} = (0, 0, 0)$. The mold of this surface is also a flat surface but oriented in the opposite direction. Any patch on this complementary surface can thus be described by $\mathbf{N} = (-N_x, -N_y, -N_z)$, $\mathbf{K} = (0, 0, 0)$.

EXAMPLE 7.4. Finally consider a minimal surface³ (other than the plane) with an arbitrary orientation. Such a saddle surface patch can be described by $\mathbf{N} = (N_x, N_y, N_z)$, $\mathbf{K} = (c \cos(2\phi), c \sin(2\phi), 0)$. The mold of this surface is also a minimal surface with the opposite orientation and the same shape but inverted. The complementary saddle surface patch can thus be described by $\mathbf{N} = (-N_x, -N_y, -N_z)$, $\mathbf{K} = (-c \cos(2\phi), -c \sin(2\phi), 0)$.

2.3. Fibre Bundles. Since we use fibre bundles as the framework for our approach, we introduce the terminology in this section. The notion of fibre bundles is fundamental to modern differential geometry [59].

DEFINITION 7.3. A bundle is a triple (E, π, B) , where $\pi : E \rightarrow B$ is a map. The space B is called the base space, the space E is called the total space, and the map π is called the

²An umbilic surface is characterized by a normal curvature which is independent from the direction; e.g. every point on a sphere is umbilic, $\kappa_1 = \kappa_2 = \frac{1}{\text{radius}}$.

³A minimal surface is a surface for which the mean curvature is null; i.e. $\kappa_1 = -\kappa_2$.

projection of the bundle. For each $b \in B$, the space $\pi^{-1}(b)$ is called the fibre of the bundle over $b \in B$ [59].

DEFINITION 7.4. A bundle (E', π', B') is a sub-bundle of $\pi : E \rightarrow B$ provided E' is a subspace of E , B' is a subspace of B , and $\pi' = \pi|_{E'} : E' \rightarrow B'$ [59].

DEFINITION 7.5. A cross-section of a bundle (E, π, B) is a map $s : B \rightarrow E$ such that $\pi s = 1_B$. In other words, a cross-section is a map $s : B \rightarrow E$ such that $s(b) \in \pi^{-1}(b)$, the fibre over b , for each $b \in B$ [59].

Denoting the fibre F , an example is the product bundle $\epsilon = (B \times F, \pi, B)$, which illustrates how the total space can be viewed as the base manifold crossed with the fibre space. The bundle can be thought of as a union of fibres $F(b) = \pi^{-1}(b)$ for $b \in B$ parametrized by B and “glued together” by the topology of the space E .

2.4. The Scenel Bundle. The N-K space describes the surface model and provides a representation with the desired characteristics. The virtual illuminant direction $\mathbf{L}$ and the image position (x, y) complete the local scene description.

DEFINITION 7.6. The SCENEL is a nine-dimensional object, $\mathcal{I}_i = \{x_i, y_i, \mathbf{L}_i, \mathbf{N}_i, \mathbf{K}_i\}$, which provides a hypothesis that the scene can be locally described by an image position (x_i, y_i) , an illuminant direction $\mathbf{L}_i$, a surface orientation $\mathbf{N}_i$, and a surface shape $\mathbf{K}_i$.

DEFINITION 7.7. The SCENEL BUNDLE is the triple (E, π, B) , where the total space is the set of all scenels $E = \{\mathcal{I}_i, \mathcal{I}_i \in \mathcal{I}\}$, the base space is the image plane $B = \{(x_i, y_i), (x_i, y_i) \in \mathbb{R}^2\}$, and $\pi : E \rightarrow B \mid \mathcal{I}_i \rightarrow (x_i, y_i)$

DEFINITION 7.8. For each $(x_i, y_i) \in B$, the space $\pi^{-1}(x_i, y_i) = \{x_i, y_i, \mathbf{L}_i, \mathbf{N}_i, \mathbf{K}_i \mid \mathbf{L}_i \in \mathbf{L}, \mathbf{N}_i \in \mathbf{N}, \mathbf{K}_i \in \mathbf{K}\}$ is called the FIBRE of the scenel bundle over $(x_i, y_i) \in B$.

OBSERVATION 7.5. The scenel bundle is a trivial bundle, the total space is the product of the base space and the fibre space.

OBSERVATION 7.6. In contrast, the Darboux frame field is not a trivial bundle for which visible smooth surfaces are cross-sections as the image plane does not define a global chart for the Darboux frame.

The scenel bundle structure allows us:

- (i) to relate the scene properties with the image properties in a purely local analysis;
- (ii) to connect local scene descriptions together in a way consistent with their geometry.

In the next subsection, we examine the scenel bundle from both a topological perspective and a geometric one.

2.5. Topology. We state the embedding space of the fibre: $\mathbf{L} = \{L_i, L_i \in \mathbb{S}^2\} \Rightarrow \mathbf{L} \subseteq \mathbb{S}^2$, $\mathbf{N} = \{N_i, N_i \in \mathbb{S}^2\} \Rightarrow \mathbf{N} \subseteq \mathbb{S}^2$, and $\mathbf{K} = \{K_i, K_i \in \mathbb{R}^3\} \Rightarrow \mathbf{K} \subseteq \mathbb{R}^3$. Since $\mathbb{S}^2$ is orientable, both $\mathbf{L}$ and $\mathbf{N}$ are orientable. And since $\mathbb{R}^3$ is orientable, $\mathbf{K}$ is orientable. The product of orientable spaces is also an orientable space, hence the following observation.

OBSERVATION 7.7. *The scenel total space, $\mathcal{I} \subseteq \mathbb{R}^2 \times \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{R}^3$, is orientable.*

We now turn to orientation of sections. Solids have surfaces that are oriented manifolds. We follow the convention that the normal points away from the solid. The projection of the virtual illuminant direction on the normal has to be positive for the light to shine on the surface.

The scenel bundle $\mathcal{I} = \{x, y, \mathbf{L}, \mathbf{N}, \mathbf{K}\}$ is a trivial bundle over the base space, any section s is a globally defined map $\pi^{-1} : \mathbb{R}^2 \rightarrow \mathcal{I}$ and thus constitutes a global chart for visible scene $\pi^{-1}(\mathbb{R}^2)$. 2-Manifolds with global charts are necessarily orientable [21].

OBSERVATION 7.8. *Any continuous cross-section of the scenel space constitutes an oriented manifold.*

The scene description that constitutes a solution to the shape from shading problem is a viewer centered description. Only one side of a surface is visible. Non-orientable surfaces such as the Möbius band, do not pose any special problems. The problem of global non-orientability does not arise since such surfaces are never entirely visible.

The scenel bundle topology thus provides a weak constraint on the compatibility between scenels. The orientation of a scenel bundle's cross-section must be maintained. The Jacobian determinant of the coordinate transformation between compatible scenels must always be positive. If we consider two visible surface patches, compatibility requires that:

- for both surface patches, the inner product of the normal and the viewer's direction be of the same sign (positive);
- for both surface patches, the inner product of the normal and the virtual illuminant's direction be of the same sign (positive).

OBSERVATION 7.9. *The scenel bundle, $\mathcal{I} = \{x, y, \mathbf{L}, \mathbf{N}, \mathbf{K}\}$, possesses a topology, $(x, y) \subset \mathbb{R}^2$, $\mathbf{L} \subseteq \mathbb{S}^2$, $\mathbf{N} \subseteq \mathbb{S}^2$, $\mathbf{K} \subseteq \mathbb{R}^3$, consistent with the topology of the space of surfaces.*

This observation is important as it validates the relation between equivalence class and the Voronoi cell. The neighbourhood in the fibre space carries a meaning of “closeness” in the scene domain.

2.6. Geometry. By looking at the geometry of the scenel bundle, we establish an even more restrictive condition for the scenel compatibility.

Let $I = \{Q_I : \mathbb{R}^2 \rightarrow \mathbb{R} \mid Q_I(x, x) \geq 0\}$ denote the space spanned by the first fundamental forms, and $II = \{Q_{II} : \mathbb{R}^2 \rightarrow \mathbb{R} \mid Q_{II}(x, x) \geq 0\}$ denote the space spanned by the second fundamental forms, where Q denotes quadratic forms,

$$Q(x, x) := x^t \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} x .$$

Let $N^+ = \{(N_x, N_y, N_z) \in \mathbf{N} \mid N_z > 0\}$ denote the space of visible surface orientations.

PROPOSITION 7.1. *Given a scale, there exists a continuous mapping from the visible N^+ - $\mathbf{K}$ space into the product space $I \times II$ of first and second fundamental forms.*

Proof: From the normal vector, $N_i \in N^+$, we can define two vectors in the tangent plane, $\mathbf{X}' = (1, 0, -\frac{N_x}{N_z})$ and $\mathbf{Y}' = (0, 1, -\frac{N_y}{N_z})$. These tangent vectors are defined for all visible surface, i.e. $N_z > 0$. The inner products $\langle \mathbf{X}', \mathbf{X}' \rangle$, $\langle \mathbf{X}', \mathbf{Y}' \rangle$, and $\langle \mathbf{Y}', \mathbf{Y}' \rangle$ define the coefficients of a first fundamental form,

$$Q_I(x, x) = x^t \begin{pmatrix} \langle \mathbf{X}', \mathbf{X}' \rangle & \langle \mathbf{X}', \mathbf{Y}' \rangle \\ \langle \mathbf{X}', \mathbf{Y}' \rangle & \langle \mathbf{Y}', \mathbf{Y}' \rangle \end{pmatrix} x .$$

Thus, we can define a mapping $\xi : (N^+, \mathbf{K}) \rightarrow I$.

In the previous section, we have expressed a continuous mapping from κ_1, κ_2, ϕ to $\mathbf{K}$. Now, we establish the existence of an inverse mapping, from $\mathbf{K}$ to κ_1, κ_2, ϕ . The principal curvatures κ_1 and κ_2 can be recovered from a shape vector $\mathbf{K}$ as follows:

$$\begin{aligned} \kappa_1 &= \begin{cases} c & \text{if } K_3 \geq 0 \\ c + 2K_3 & \text{otherwise} \end{cases} \\ \kappa_2 &= \begin{cases} -c + 2K_3 & \text{if } K_3 \geq 0 \\ -c & \text{otherwise} \end{cases} \end{aligned}$$

where $c = \sqrt{K_1^2 + K_2^2 + K_3^2}$. These mappings are continuous even when $K_3 = 0$, for $K_3 \rightarrow 0^+$ and $K_3 \rightarrow 0^-$, both yield $\kappa_1 = -\kappa_2 = c$.

The principal frame is defined with the normal and the principal directions. The latter, when defined, are represented by the angle ϕ which can also be recovered as follows:

$$\phi = \frac{1}{2} \tan^{-1} (K_1, K_2) .$$

To get the second fundamental form coefficients, $\{q_{11}^{\text{II}}, q_{12}^{\text{II}}, q_{22}^{\text{II}}\}$ in terms of the firsts, $\{q_{11}^{\text{I}}, q_{12}^{\text{I}}, q_{22}^{\text{I}}\}$, and $\mathbf{K}$, we use the following system of equations:

$$\begin{aligned} \kappa_1 \kappa_2 &= \frac{q_{11}^{\text{II}} q_{22}^{\text{II}} - q_{12}^{\text{II}^2}}{q_{11}^{\text{I}} q_{22}^{\text{I}} - q_{12}^{\text{I}^2}} \\ \kappa_1 + \kappa_2 &= \frac{q_{22}^{\text{I}} q_{11}^{\text{II}} - 2q_{12}^{\text{I}} q_{12}^{\text{II}} + q_{11}^{\text{I}} q_{22}^{\text{II}}}{q_{11}^{\text{I}} q_{22}^{\text{I}} - q_{12}^{\text{I}^2}} \end{aligned}$$

$0 = (q_{12}^{\text{I}} q_{22}^{\text{II}} - q_{22}^{\text{I}} q_{12}^{\text{II}}) \sin^2 \phi + (q_{11}^{\text{I}} q_{22}^{\text{II}} - q_{22}^{\text{I}} q_{11}^{\text{II}}) \sin \phi \cos \phi + (q_{22}^{\text{I}} q_{12}^{\text{II}} - q_{12}^{\text{I}} q_{22}^{\text{II}}) \cos^2 \phi$
 where $\kappa_1 + \kappa_2 = 2K_3$ and $\kappa_1 \kappa_2 = c(2|K_3| - c)$. When the angle ϕ is undefined, $K_1 = K_2 = 0$, the third equation reduces to:

$$\frac{q_{11}^{\text{I}}}{q_{11}^{\text{II}}} = \frac{q_{12}^{\text{I}}}{q_{12}^{\text{II}}} = \frac{q_{22}^{\text{I}}}{q_{22}^{\text{II}}}$$

Thus, for visible surfaces, there exists a continuous mapping $(\mathbf{N}^+, \mathbf{K}) \rightarrow (\text{I}, \text{II})$. $\square$

OBSERVATION 7.10. *From a cross-section of the sub-bundle $\{x, y, \mathbf{N}^+, \mathbf{K}\} \subset \mathcal{I}$, the surface can be recovered up to position if Cartan's equations of compatibility are satisfied.*

This observation follows directly from Prop. 7.1 and the *Fundamental theorem of surfaces* [27, 70]. It suggests the basis for the scenels' compatibility relationships that we develop in Chapt. 8.

We have only a weak topological constraint on the virtual illuminant. The stronger constraint that we will impose derives from the underlying assumption of the shape from shading problem.

CHAPTER 8

The Scene Element

The key idea in the step of inferring shape from a shading flow field is to consider it as a coupled family of local problems. Each of these problems is a “micro”-version of the shape from shading problem in which a lighting and a surface model interact to produce the locally observed shading structure. We call each of these different models a scene element, or *scenel*. Whereas each scenel deals with the local consistency with the data, the coupling between scenels on different fibres provides the means to assure global consistency.

1. The Discrete Scenel Space

On a digital computer, both the data and the result of the computation are represented by a finite number of bits. Our data, the intensity image, is represented by a function for which both the domain and the range are discrete values. This affects any attempts to recover shape from shading with such a device. A consequence is that quantities such as the local orientation of a contour can only be derived with finite precision [66], thus limiting the precision of the shading flow field as well. We acknowledge this and exploit it to our advantage.

An analogy can be drawn with significant digits in experimental data — we rarely write more than one non-significant digit when an error is associated with a measure (e.g. gravitational acceleration is equal to $9.8 \pm 0.1 \text{ m/sec}^2$). It is in this sense that we are using the term “precision”.

OBSERVATION 8.1. We can model the scene with only a finite number of different scenels without affecting the precision of the result.

1.1. The Quantization of the Scenel Attributes. To model the scene by a finite number of different scenels, each scenel attribute is quantized:

- (i) **IMAGE POSITION:** The image pixels are given as a set of discrete values of x and y position.
- (ii) **ILLUMINATION DESCRIPTORS:** Our illumination model comprises two components describing a virtual light source:

- (a) The possible virtual light source directions map onto a unit sphere. We sample this sphere as uniformly as possible to get a discrete set of virtual illuminant directions; see Fig. 8.37. In viewer-centered coordinates, these unit vectors are given as

$$\mathbf{L} = (L_x, L_y, L_z) .$$

- (b) The illumination from the virtual light source λ is not represented in the scene description as it is not used in our shape from shading analysis.

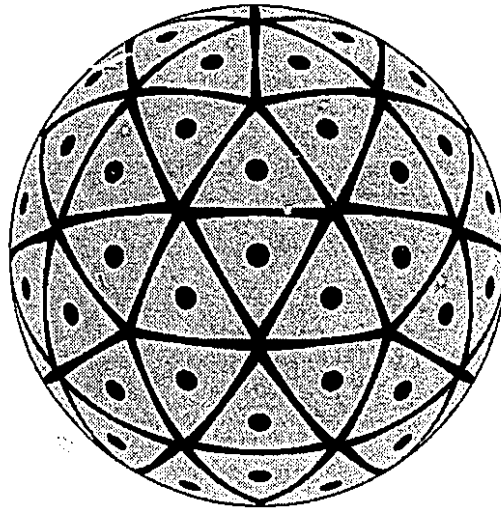


FIGURE 8.37. The discrete representation of a unit vector such as the illumination direction starts with a sampling of the unit sphere. The sampled points on the sphere are chosen to be regularly spaced if possible or something close to it otherwise. These points form a set that defines a Voronoi tessellation. The Voronoi cell consists of every point on the sphere that is closer to a given sampled point than any other. In our representation, the vector corresponding to a sample point represents the class of vectors corresponding to the Voronoi cell.

- (iii) **SURFACE DESCRIPTORS:** In the previous section, the surface was said to be represented by the albedo ρ , the normal $\mathbf{N}$, and the shape vector $\mathbf{K}$.

- (a) The set of all normals forms a unit sphere. The surface of this sphere is sampled as uniformly as possible to derive a discrete set of normals. In viewer-centered

coordinates, these unit vectors are given as

$$\mathbf{N} = (N_x, N_y, N_z) .$$

Of these, only the ones in the hemisphere facing the viewer are used; the others are not visible.

- (b) Although the shape vector lies in $\mathbb{R}^3$, we prefer to quantize it as if it was lying in $\mathbb{S}^2 \times \mathbb{R}^+$. The shape vector can be decomposed in a magnitude (the curvedness c) and a direction (a unit shape vector $\hat{\mathbf{K}}$). The set of all unit shape vectors also forms a unit sphere. Again, by sampling the surface of the sphere uniformly we derive a discrete set of parameters which cover all variations of smooth, oriented shape in the tangent plane. Augmenting this with the curvedness index provides a complete, discretely sampled shape descriptor. In the $\hat{\mathbf{K}}$ space, these unit vectors are given as

$$\hat{\mathbf{K}} = (\hat{K}_1, \hat{K}_2, \hat{K}_3) .$$

- (c) The albedo of the surface ρ is not represented in the scenel description. As with illumination, the albedo is not needed for our shape from shading analysis.

1.2. The Set of Scenels. Given the discrete sampling of the scene attributes as defined above, we derive a set of scenel labels

$$\mathcal{I} = \{x, y, \mathbf{L}, \mathbf{N}, \mathbf{K} \mid (x, y) \in \text{image}, \mathbf{L} \in \mathbb{S}^2, \mathbf{N} \in \mathbb{S}^2, \mathbf{K} \in \mathbb{S}^2 \times \mathbb{Z}\}$$

which represent all potential assignments of these scene attributes. Thus each i represents the hypothesis that the scene can be locally described by the scenel $(x_i, y_i, \mathbf{L}_i, \mathbf{N}_i, \mathbf{K}_i)$.

DEFINITION 8.1 (ALTERNATE). *Each scenel is viewed as the hypothesis that, at the given image position, the scene can be locally described by a surface of a given shape, oriented in a given direction, and illuminated from a given direction.*

Recall that the total space of the scenel bundle is a product bundle composed of the image plane (a finite set of pixels) as the base space and the scenel fibre. The scenel space is thus quantized. A scenel i is meant to model the equivalence class of local scene whose descriptions lie “closer to” it than to any other scenel j . The collection of all such local scene descriptions forms the Voronoi cell of scenel i . The scenel space is thus depicted as a Voronoi tessellation corresponding to our chosen set of scenels.

The Voronoi tessellation implies that the scenel space is a *metric space*. Consider the mapping $\varrho : \mathcal{I} \times \mathcal{I} \rightarrow \mathbb{R}$ such that:

$$\varrho(\mathcal{I}_i, \mathcal{I}_j) = \left((x_j - x_i)^2 + (y_j - y_i)^2 + \left(\frac{\cos^{-1}(\mathbf{N}_j \cdot \mathbf{N}_i)}{\Delta\gamma^{\mathbf{N}}} \right)^2 + \left(\frac{\cos^{-1}(\hat{\mathbf{K}}_j \cdot \hat{\mathbf{K}}_i)}{\Delta\gamma^{\hat{\mathbf{K}}}} \right)^2 + \left(\frac{c_j - c_i}{\Delta c} \right)^2 \right)^{\frac{1}{2}}$$

where $\Delta\gamma^{\mathbf{N}}, \Delta\gamma^{\hat{\mathbf{K}}}, \Delta c$ are the average distance between neighbouring scenels for the given attribute. Since

$$\begin{aligned} \varrho(\mathcal{I}_i, \mathcal{I}_j) &\geq 0, \\ \varrho(\mathcal{I}_i, \mathcal{I}_j) &= 0 \quad \text{iff } \mathcal{I}_i = \mathcal{I}_j, \\ \varrho(\mathcal{I}_i, \mathcal{I}_j) &= \varrho(\mathcal{I}_j, \mathcal{I}_i), \\ \varrho(\mathcal{I}_i, \mathcal{I}_k) &\leq \varrho(\mathcal{I}_i, \mathcal{I}_j) + \varrho(\mathcal{I}_j, \mathcal{I}_k), \end{aligned}$$

the mapping $\varrho(\mathcal{I}_i, \mathcal{I}_j)$ defines the distance. Each local descriptor is treated as an independent dimension and is normalized by the descriptor's quantization. The image coordinates, x and y , are already given in unit pixels.

To relate the scenel bundle approach to the relaxation labelling paradigm, we consider each scenel as a node of the relaxation network. Since the scenel hypothesis can be either true or false, we associate to each node, two labels: TRUE and FALSE. We distribute a measure $p_i(\lambda)$ over these labels $\lambda \in \Lambda$ for each scenel i . Since there are only two labels, it suffices to explicitly represent one label as $p_i(\text{TRUE}) = p_i(\text{FALSE})$. We choose to represent only the TRUE label. This allows us to simplify the notation: p_i shall be used from now on to represent confirmation of the scenel hypothesis.

2. The Inference of the Scenel Fibre

Our approach is data-driven. We have already observed that the scenels cannot be directly derived from the image irradiance if we allow the surface albedo and illumination to vary. We have to consider the local geometric structures of the image that do not depend on the surface albedo and illumination: the tangent field and the shading flow field.

Thus, from the initial data, a static intensity image, we extract relevant image geometric structures. These, in turn, become our data. For each image position, we infer the set of possible scenels that can account for these geometric structures. Hence, at each image position, the local geometric structures yield a sampling of the fibre of possible scenels.

When we examine the relation between the local image structures and the scenel, there are two important considerations to keep in mind.

- (i) Since the scenel represents a range of scene descriptors, a scenel can account for several different image geometrical structures.
- (ii) Furthermore, several different scenels can account for the same image geometrical structures.

In this section, we show how to assign the confidence measure to a scenel given the available data.

DEFINITION 8.2. *The confidence measure associated to each hypothesis reflects the compatibility of the corresponding scenel and the observed structures in the image.*

At the onset of the shape from shading problem, the confidence measure's assignment is pixel-wise. The discrete nature of both the scenel representation and the image local geometric properties, allows us to build a look-up table to capture the link between the image local geometric properties, $\Upsilon = \Upsilon^s \cup \Upsilon^t$, and the confidence measures for the scenels.

It takes the form of a look-up table in which, to build this look-up table,

- (i) we need to establish first the relationship between the scene descriptors and the geometrical structures of the image (see below);
- (ii) we over-sample the scenel space as uniformly as possible and tag each sample by marking the Voronoi cell to which it belongs;
- (iii) then for each sample, we compute the expected geometrical image structures and add the sample to a *data* bin corresponding to the appropriate image geometrical structure Voronoi cell — this takes the form of a histogram whose dimensionality reflects the dimensionality of the image geometrical structures $\Upsilon = \Upsilon^s \cup \Upsilon^t$;
- (iv) the weight p_i associated with scenel i will reflect the ratio of the samples tagged with i in the appropriate data bin over all samples tagged with i .

For every given image geometrical structure, Υ_l , a list of possible scenels (each with an associated weight) is provided.

2.1. The Local Image Structures generated by the Scenel. To generate geometrical image structures from the scenel, we consider the lighting condition and the surface locally described by the scenel i . We choose to locally approximate the surface model by a paraboloid because it allows us to fully exploit the information provided by the scenel

$(x_i, y_i, \mathbf{L}_i, \mathbf{N}_i, \mathbf{K}_i)$. The paraboloid is a curved surface that obeys the following local parametric form:

$$\left(u, v, -\frac{1}{2} \kappa_1 u^2 - \frac{1}{2} \kappa_2 v^2 \right),$$

where the u and v are the parameters corresponding respectively to the two principal directions, and κ_1 and κ_2 correspond to the two principal curvatures at the origin.

Note that this approximation can find its ancestry in Pentland's local shading analysis [31,106]. In his algorithm, Pentland used the sphere (an umbilic surface) to locally approximate the surface shape. The paraboloid is a generalization over the sphere that enjoys important qualitative differences. In particular, it can model both elliptic and hyperbolic surfaces as κ_1 and κ_2 can take positive and negative values.

There are a few different coordinate systems that we consider when establishing the relation between scenels and the local image structures.

- The viewer's coordinates $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ are defined such that the vector $(0, 0, 1)$ points toward the viewer. Note that, for image formation, we will assume orthographic projection, hence, the image lies in the x - y plane.
- The surface patch local coordinates $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ are defined such that, at a given point, the vector $(0, 0, 1)$ indicates the direction of the surface normal, and the vectors $(1, 0, 0)$ and $(0, 1, 0)$ correspond to the surface principal directions.
- The principal frame field $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ is defined such that, at every point, the vector $(0, 0, 1)$ indicates the direction of the surface normal, and the vectors $(1, 0, 0)$ and $(0, 1, 0)$ correspond to the surface principal directions.

The viewer's coordinates and the surface patch local coordinates are related by the linear transformation matrix:

$$\mathbf{M} = \begin{pmatrix} e_1^{(x)} & e_2^{(x)} & e_3^{(x)} \\ e_1^{(y)} & e_2^{(y)} & e_3^{(y)} \\ e_1^{(z)} & e_2^{(z)} & e_3^{(z)} \end{pmatrix}$$

where $e_i^{(x)}, e_i^{(y)}, e_i^{(z)}$ are respectively the x, y, z components of the i^{th} vector of the surface patch local coordinate system. This transformation is only a rotation matrix. Hence the inverse of the matrix $\mathbf{M}$ is simply its transpose.

The matrix $\mathbf{M}$ can be recovered from the scenel i descriptors, more precisely, from the normal $\mathbf{N}_i$ and the shape vector $\mathbf{K}_i$ as follows:

$$\mathbf{M} = \mathbf{M}_n \mathbf{M}_c$$

where M_n describes the rotation depicted in Fig. 7.34(a) and M_c describes the rotation depicted in Fig. 7.34(b). These matrices are related to the scenel descriptors as follows:

$$M_n = \begin{pmatrix} X'_x & Y'_x & N_x \\ X'_y & Y'_y & N_y \\ X'_z & Y'_z & N_z \end{pmatrix} = \begin{pmatrix} \frac{N_x^2 N_z + N_y^2}{N_x^2 + N_y^2} & \frac{N_x N_y N_z - N_x N_y}{N_x^2 + N_y^2} & N_x \\ \frac{N_x N_y N_z - N_x N_y}{N_x^2 + N_y^2} & \frac{N_y^2 N_z + N_x^2}{N_x^2 + N_y^2} & N_y \\ -N_x & -N_y & N_z \end{pmatrix}$$

and

$$M_c = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

where $\phi = \frac{1}{2} \tan^{-1} \left(\frac{K_2}{K_1} \right)$.

Now that we have shown the relation between the paraboloid model and the scenel, we use this model to derive the expected shading flow field and the expected edge map.

2.1.1. The Expected Shading Flow Field. Given a sample of the scenel space, we want to compute the shading flow field. The relationship between the scene descriptors and the shading flow field is based on the image irradiance equation:

$$I = \rho \lambda \mathbf{L} \cdot \mathbf{N},$$

where ρ , λ , $\mathbf{L}$ are locally constant functions of (x, y) , and $\mathbf{N}$ is a locally smooth function of (x, y) .

The orientation of the shading flow field θ and the shading flow field's normalized distortion vector $\tilde{\mathbf{D}}$ can be derived from the first and second spatial partial derivatives:

$$\tan \theta = \frac{I_y}{I_x};$$

$$\tilde{\mathbf{D}} = \frac{1}{\sqrt{I_x^2 + I_y^2}} \begin{pmatrix} I_{xx} - I_{yy} \\ -2I_{xy} \\ I_{xx} + I_{yy} \end{pmatrix}.$$

Using the paraboloid as the local surface model, we evaluate the first spatial partial derivatives¹:

$$I_x = \rho \lambda \mathbf{L} \cdot (e_1^x (-k_1 \mathbf{f}_1) + e_2^x (-k_2 \mathbf{f}_2))$$

$$I_y = \rho \lambda \mathbf{L} \cdot (e_1^y (-k_1 \mathbf{f}_1) + e_2^y (-k_2 \mathbf{f}_2))$$

¹Since the point considered is the origin of the surface patch local coordinates, the components of principal frame field $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ are equal to $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

and the second spatial partial derivatives:

$$\begin{aligned} I_{xx} &= -\rho\lambda\mathbf{L} \cdot (e_1^x e_1^x \kappa_1^2 + e_2^x e_2^x \kappa_2^2) \mathbf{f}_3 \\ I_{xy} &= -\rho\lambda\mathbf{L} \cdot (e_1^x e_1^y \kappa_1^2 + e_2^x e_2^y \kappa_2^2) \mathbf{f}_3 \\ I_{yy} &= -\rho\lambda\mathbf{L} \cdot (e_1^y e_1^y \kappa_1^2 + e_2^y e_2^y \kappa_2^2) \mathbf{f}_3 \end{aligned}$$

As mentioned in Chapt. 4, the useful properties of the shading flow field are only dependent on the scenel descriptors $\mathbf{N}$, $\mathbf{K}$, $\mathbf{L}$ — the product $\rho\lambda$ always appears in both the numerator and the denominator and thus always cancels itself.

Hence, we have a relationship between the scenel description $\mathcal{I} = \{x, y, \mathbf{L}, \mathbf{N}, \mathbf{K}\}$ and the shading flow field: $\Upsilon^s = \{x, y, \theta, \tilde{\mathbf{D}}\}$.

2.1.2. The Expected Edge Map. Among the image curves that we extract from the data, there are edges which corresponds to occluding contours. These occluding contours are curves that lie in the scene domain. We exploit three of their features:

- (i) Occluding contours lie on surfaces.
- (ii) At the occluding contour, the surface is perpendicular to the line of sight.
- (iii) Occluding contours projected on the image plane form edges.

The occluding contour is a three-dimensional curve that can be represented in the viewer's coordinate system $(x(t), y(t), z(t))$ or the local surface patch coordinate system $(u(t), v(t), w(t))$ with the two being related as follows:

$$(x(t), y(t), z(t)) = \mathbf{M} (u(t), v(t), w(t)) .$$

If we consider again the paraboloid as our local surface model, the surface normal vector is given in terms of a local surface patch coordinate system:

$$\mathbf{N} = \left(\frac{\kappa_1 u}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{\kappa_2 v}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{1}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}} \right)$$

Since at the occluding boundaries, the surface becomes perpendicular to the line of sight, we express the normal vector in image coordinates.

$$\mathbf{N}^{(xyz)} = \mathbf{M} \mathbf{N}^{(uvw)}$$

And thus occluding boundaries are characterized by the z component of the normal vector being null.

$$\mathbf{N}^{(z)} = \frac{e_1^z \kappa_1 u + e_2^z \kappa_2 v + e_3^z}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}} = 0$$

which provide a first constraint to define the curve describing the occluding contour.

$$(8.1) \quad e_1^z \kappa_1 u(t) + e_2^z \kappa_2 v(t) + e_3^z = 0$$

Another constraint on the curve describing the occluding contour is that it lies on the paraboloid, thus

$$w(t) = -\frac{1}{2} \kappa_1 u^2(t) - \frac{1}{2} \kappa_2 v^2(t) .$$

We can project this curve on the image plane and obtain a planar curve $\mathbf{x} = (x(t), y(t))$ that describes the corresponding edge:

$$\begin{aligned} x(t) &= e_1^x u(t) + e_2^x v(t) + e_3^x w(t) , \\ y(t) &= e_1^y u(t) + e_2^y v(t) + e_3^y w(t) . \end{aligned}$$

Successive differentiations by the arclength yield the tangent vector and the curvature vector. The tangent vector of this curve is given by:

$$\mathbf{t} = (\sin(\theta(t)), \cos(\theta(t))) = \frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{dt} \frac{dt}{ds} ,$$

where s is the arclength parametrization. And the curvature vector of this curve is given by:

$$\mathbf{k} = \kappa \mathbf{n} = \frac{d\mathbf{t}}{ds} .$$

For every sample, we consider whether or not an occluding contour could be visible at the same pixel location. An occluding contour is said to be visible at the same pixel location if the corresponding image curve can pass through the pixel in which the origin of the paraboloid model lies.

Hence, we have a relationship between the scenel description $\mathcal{I} = \{x, y, \mathbf{L}, \mathbf{N}, \mathbf{K}\}$ and the tangent field corresponding to occluding contours: $\Upsilon^t = \{x, y, \theta(t), \kappa(t)\}$.

2.2. The Labelling. The discretized scenel space used for our implementation is detailed in App. E. The question of how coarse or how fine to sample each of the dimensions of the scenel space is subtle, as there are opposing considerations. For computational considerations, the fewest scenels possible is preferable, whereas for representational considerations, the more scenels the better. However, because of the limited precision of the data, there is a limit on the precision of the scenel representation.

OBSERVATION 8.2. *This number of scenels will be dictated by the discriminability of the scenels given the shading flow field.*

In Fig. 8.38, we illustrate the variation of the expected shading flow that occurs for scene descriptors within an equivalence class. The discrimination between these is impractical because of the limited precision of the shading flow field representation.

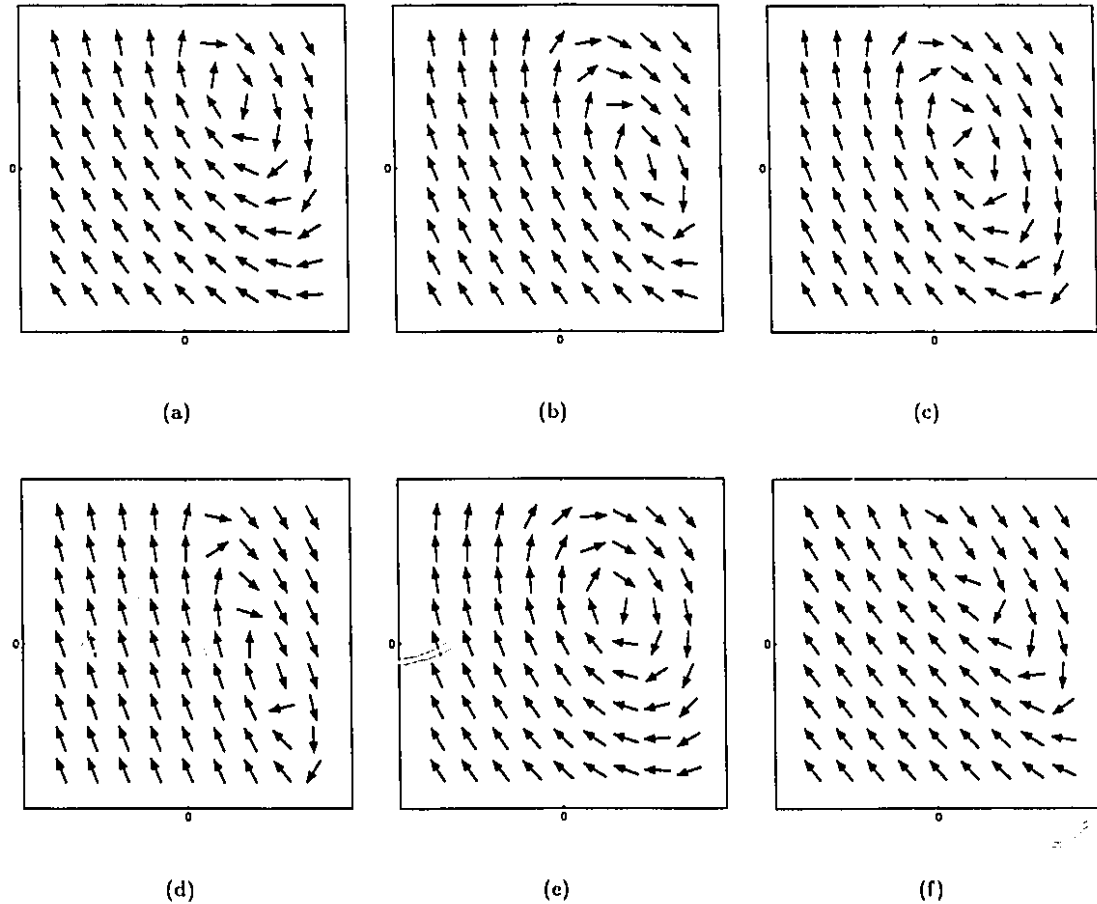


FIGURE 8.38. This figure illustrates the variation of the expected shading flow field within an equivalence class of scenes. In (a), (b), (c), the normal changes. There is an angle difference of approximately 10° between any pair of normals shown. In (d), (e), (f), it is the shape vector which changes. There is an angle difference of approximately 20° between any pair of shape vectors shown. The normal and shape vectors used are depicted in Fig. 8.40. The curvedness is 0.15 and the light source is (0.3035, 0.5257, 0.7946).

In Fig. 8.39, we illustrate the variation of the expected shading flow that occurs for scene descriptors belonging to neighbouring equivalence classes. The discrimination between these is possible even given the limited precision of the shading flow field representation.

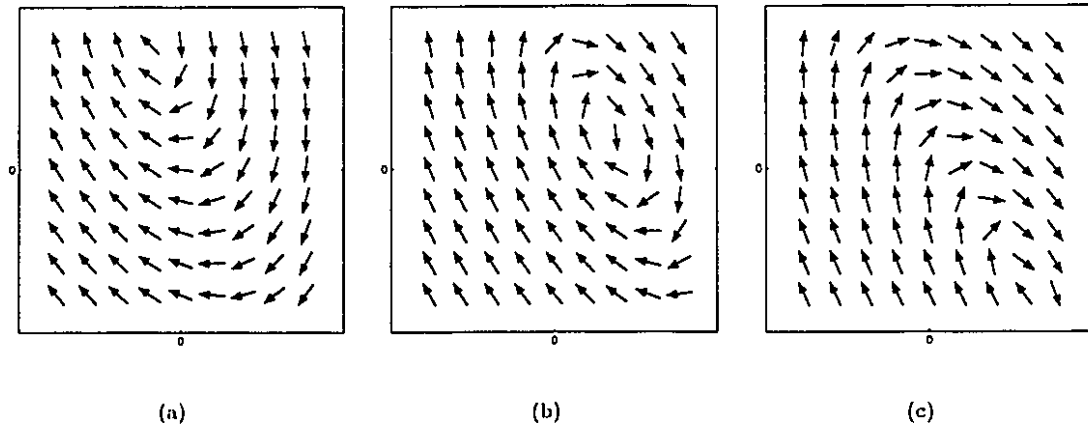


FIGURE 8.39. This figure illustrates the variation of the expected shading flow field for neighbouring scenels. The normal changes and all the other scenel descriptors remain constant. The normal and the shape vectors used are shown in Fig. 8.40; the curvedness is 0.15 and the light source is (0.3035, 0.5257, 0.7946). There is an angle difference of approximately 30° between the normals of (a) and (c), and an angle difference of approximately 20° between the normals of (b) and either of (a) or (c).

The expected shading flow field computed from a scenel's descriptors is projected on the discretized shading flow field space. The discretized shading flow field must yield back the scenel.

3. The Consistent Scenel Labelling

Up to now, we have dealt with the scenels' compatibility with the local geometric structures inferred from the image. In this section, we consider the interaction of scenels on neighbouring fibres, thereby making the transition from local to global.

3.1. The Solution: a Cross-Section of the Scenel Bundle. Viewed globally, the solution we seek consists of sections in which a single (equivalent) light source illuminates a collection of surface patches with constant material properties but whose shape properties vary smoothly. The above constraints are embedded into a functional, and consistent sections through the scenel bundle are stationary points of this functional. More specifically, the constraints are expressed as compatibility relationships between pairs of neighbouring estimates within a relaxation labelling process.

$$S_i = \sum_{j \in N(i)} r_{ij} p_j .$$

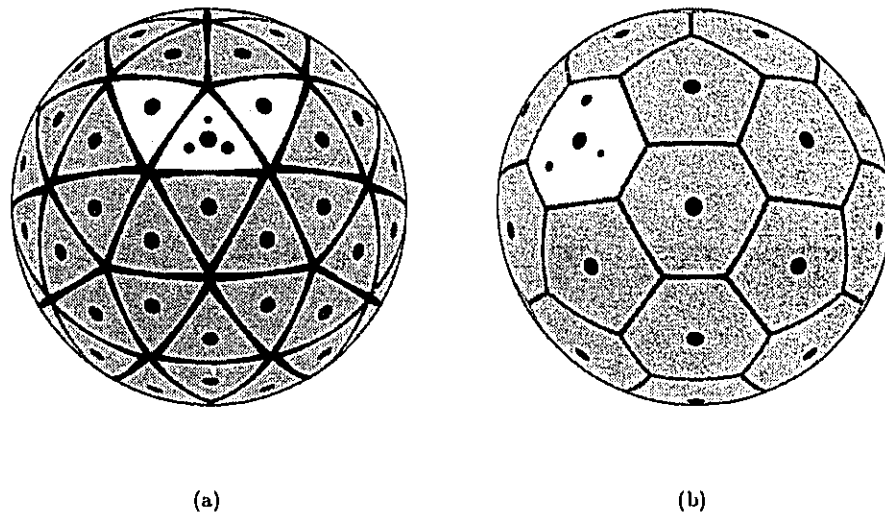


FIGURE 8.40. These figures show the tessellations of (a) the normal space and of (b) the shape vector space. Represented as black dots of various sizes are the normals and the shape vectors used to illustrate the variation of the shading flow within an equivalence class and between equivalence classes. We show as white Voronoi cells (in the normal and the shape vector spaces), the equivalence class from which examples are taken for Fig. 8.38. In that figure's first row, (a)–(c), the normal varies taking the values indicated respectively by, the smallest, the second smallest, and the third smallest dot depicted in the normal's white Voronoi cell. The shape vector remains constant, it keeps the value indicated by the largest black dot in the shape vector's white Voronoi cell. In the second row, (d)–(f), the shape vector varies taking the values indicated respectively by, the third smallest, the second smallest, and the smallest dot depicted in the discretized shape vector's white Voronoi cell. The normal remains constant, it keeps the value indicated by the largest dot in the normal's white Voronoi cell. In Fig. 8.39, we show examples of shading flows from neighbouring equivalence classes. We use two immediate neighbours along the normal dimension (the Voronoi cells shown in light gray). The normal varies, it corresponds to the largest dot in the white and light gray Voronoi cells. The shape vector remains constant, it keeps the value indicated by the largest black dot in the shape vector's white Voronoi cell.

The final labelling is selected such that it locally maximizes the average local support

$$A(p) = \sum_{i \in I} p_i S_i = \sum_{i \in I} \sum_{j \in N(i)} p_i r_{ij} p_j .$$

Such a labelling is said to be consistent [58].

3.2. Compatibility between Neighbouring Scene Elements. As we mentioned in Chapt. 7, the coupling between the local scenel problems dictates a consistency relationship over them. This consistency relationship derives from two principle considerations:

- (i) A SURFACE SMOOTHNESS CONSTRAINT, which states that the surface normal and curvatures must satisfy Cartan's equation for pairs of scenels which project to neighbouring points in the image domain.
- (ii) A LIGHT SOURCE CONSTRAINT, which states that the virtual light source is constant for pairs of scenels which project to neighbouring points in the image domain.

For the relaxation labelling process, these translate into the following:

- A scenel j is *compatible* with the scenel i if they have the same virtual illuminance direction $\mathbf{L}$, and if scenel j 's surface descriptors fall on scenel i 's extrapolated surface at the corresponding relative position.
- A scenel j is *incompatible* with the scenel i if they have the same virtual illuminance direction, and if there exists another scenel j' , neighbouring scenel j along the fibre, that better fits the extrapolated surface from scenel i than scenel j . Observe that this incompatibility serves to localize information along each fibre.
- otherwise a scenel j is *unrelated* to the scenel i .

Using these guiding principles, we assign a value to the compatibility r_{ij} between two scenels i and j . This compatibility will be positive for compatible hypotheses, negative for incompatible hypotheses, and zero otherwise. In general, variation in r_{ij} is assumed to be smooth between nearby points in the parameter space $\mathcal{I}$. The process is illustrated in Fig. 8.41.

We now derive the exact form of the scenel compatibility function from these principles. Consider the paraboloid $S_i(u, v)$ such that the neighbourhood of $S_i(0, 0)$ is described by the surface parameters of scenel i .

OBSERVATION 8.3. *The paraboloid $S_i(u, v)$ is uniquely defined by the scenel i .*

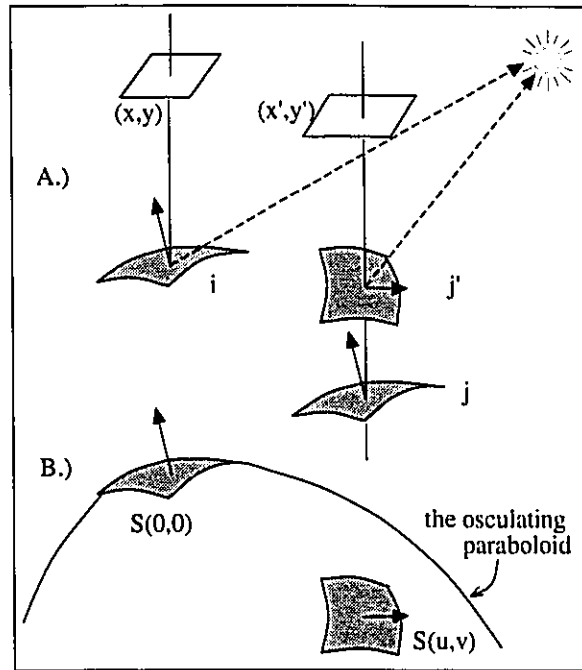


FIGURE 8.41. Illustration of the compatibility relationship for scenel consistency. Two scenels are shown on the fibre at image location (x',y') , and are evaluated against the scenel (i) at (x,y) . The surface represented in scenel $_{x,y}$ is modeled by the osculating paraboloid, and extended to (x',y') . It is now clear that one scenel (j') at (x',y') is consistent, because its surface patch lies on this paraboloid and light source and albedo agree. The other scenel (j) is inconsistent, because its surface does not match the extended paraboloid. Such osculating paraboloids are used to simulate the parallel transport of scenel $_{x',y'}$ onto scenel $_{x,y}$.

The paraboloid provides a model for defining the parallel transport of the scenel on the surface. The surface orientation in the scenel's local coordinates is given by:

$$\mathbf{N}(u, v) = \left(\frac{\kappa_1 u}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{\kappa_2 v}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{1}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}} \right)$$

The principal curvatures and directions are obtained as follows:

$$\kappa_1(u, v) = \frac{\kappa_1 (1 + \kappa_2^2 v^2)}{(1 + \kappa_1^2 u^2 + \kappa_2^2 v^2)} ;$$

$$\kappa_2(u, v) = \frac{\kappa_2 (1 + \kappa_1^2 u^2)}{(1 + \kappa_1^2 u^2 + \kappa_2^2 v^2)} ;$$

$$\frac{dv}{du}(u, v) = \frac{\kappa_1 \kappa_2 uv}{(1 + \kappa_2^2 v^2)} .$$

We find the point (u^*, v^*) for which the local descriptors $(x_i^*, y_i^*, \mathbf{N}_i^*, \hat{\mathbf{K}}_i^*, c_i^*)$ of the paraboloid $S_i(u, v)$ are "closest to" scenel j according to the distance measure, ρ .

As we have said before, each scenel j defines a Voronoi cell. Each point in this cell is closest to the scenel j than to any other scenel j' . So if the descriptors $(x_i^*, y_i^*, \mathbf{N}_i^*, \hat{\mathbf{K}}_i^*, c_i^*)$ fall into the Voronoi cell of scenel j , then r_{ij} is positive. If the descriptors fall into a neighbouring Voronoi cell, then r_{ij} is negative. If the descriptors fall further, then r_{ij} is null. Figure 8.42 illustrates the desired behaviour for the compatibilities along a dimension, such as the normal, which maps onto a unit sphere.

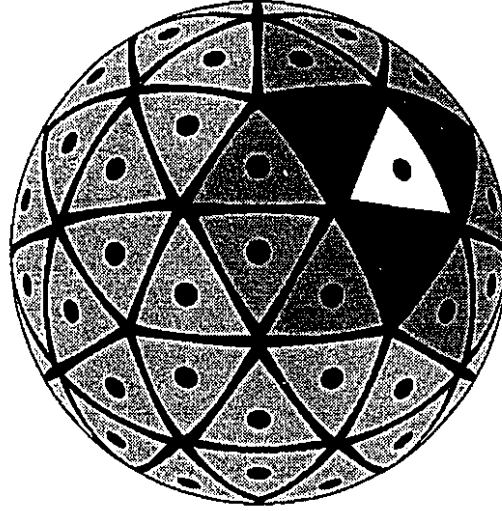


FIGURE 8.42. If the closest representation is the Voronoi cell in white, this scenel would have a strong positive compatibility (white), whereas its immediate neighbours would have a strong negative compatibility (dark gray) and the compatibility of the furthest neighbours would be null (light gray).

Building on the observation that the different dimensions of the scene are independent, we consider these separately to localize the compatible scene in the fibre. Note that we realize that it is not possible to localize the scene along the curvedness dimension because of the coarseness of its representation — there are only three curvedness for our implementation. The normal vector space and the shape vector space are the two remaining dimensions. Both are defined on unit spheres, $\mathbb{S}^2$.

The quantization of the unit sphere leads usually to an irregular and anisotropic tessellation. To avoid the related difficulties, we approximate the Voronoi cell by a disk whose radius, ζ , corresponds to the furthest point in the cell for each of these dimensions. The Gaussian, G_ζ , provides a measure over the equivalence class of objects. We use its second derivative, G''_ζ , to localize compatible measures in a given scene dimension. This function, with a sign reversal, mimics the desired behaviour for the compatibility; it is positive as long as the point lies within the approximated Voronoi cell, it becomes negative if the point lies outside of the approximated Voronoi cell, the inhibitory lobe being stronger for the immediate neighbours. Hence, we shall build our localization functions as follows:

$$Q_{ij}^N = \delta_{L_i, L_j} \cdot G(\gamma_{ij}^x) \cdot -G''_{\zeta_j^N}(\gamma_{ij}^N) \cdot G_{\zeta_j^K}(\gamma_{ij}^K) \cdot G_{\zeta_j^c}(\gamma_{ij}^c) ,$$

$$Q_{ij}^K = \delta_{L_i, L_j} \cdot G(\gamma_{ij}^x) \cdot G_{\zeta_j^N}(\gamma_{ij}^N) \cdot -G''_{\zeta_j^K}(\gamma_{ij}^K) \cdot G_{\zeta_j^c}(\gamma_{ij}^c) .$$

where

$$\begin{aligned} \gamma_{ij}^x &= \sqrt{(x_j - x_i^*)^2 + (y_j - y_i^*)^2} \\ \gamma_{ij}^N &= \cos^{-1}(\mathbf{N}_j - \mathbf{N}_i^*) \\ \gamma_{ij}^K &= \cos^{-1}(\hat{\mathbf{K}}_j - \hat{\mathbf{K}}_i^*) \\ \gamma_{ij}^c &= c_j - c_i^* \end{aligned}$$

are the different components of the distance between the transported scene and scene j , and the variances, $\zeta_j^N, \zeta_j^K, \zeta_j^c$, correspond to the Voronoi cell's radius along the $\mathbf{N}$, $\hat{\mathbf{K}}$, and c dimensions respectively.

The function Q_{ij}^N returns a positive value if and only if the extrapolated normal from scene i falls within the approximated Voronoi cell. This function will tend to take a small value if in any dimension, there is a poor match between the scene j and the descriptors derived from the paraboloid $S_i(u, v)$ at the point (u^*, v^*) . The function Q_{ij}^K behaves similarly but consider the shape vector instead of the normal.

For scenes i and j to be compatible, we require that both the paraboloid surface's orientation and curvatures need to be within the equivalence class defined by the scene j .

The scenels i and j are not considered compatible if any of the normal or the shape vector does not match — e.g. if the normal does not match, the scenels shall not be considered compatible regardless of how well the shape vector matches. To achieve this goal, we use the logical/linear combinators introduced by Iverson and Zucker [65].

DEFINITION 8.3. *The logical/linear combinator $\mathbb{A}$ is given by:*

$$x \mathbb{A} y = \begin{cases} x + y, & \text{if } x > 0 \wedge y > 0 ; \\ y, & \text{if } x > 0 \wedge y \leq 0 ; \\ x, & \text{if } x \leq 0 \wedge y > 0 ; \\ x + y, & \text{if } x \leq 0 \wedge y \leq 0 . \end{cases}$$

The compatibility between scenel i and scenel j can be expressed as a logical/linear combination of the normal's localization function $Q_{ij}^{\mathbf{N}}$, and of the shape vector's localization function, $Q_{ij}^{\mathbf{K}}$:

$$r_{ij} = Q_{ij}^{\mathbf{N}} \mathbb{A} Q_{ij}^{\mathbf{K}}$$

where r_{ij} is positive if and only if both $Q_{ij}^{\mathbf{N}}$ and $Q_{ij}^{\mathbf{K}}$ are positive. Hence, the scenel is localized in both $\mathbf{N}$ and $\mathbf{K}$ dimensions.

Notice that these values depend only on the relationship between scenel i and scenel j , which are fixed and constant throughout the computation. Therefore, these compatibilities can be calculated once and then stored in either a lookup table or as the weights in some sort of network.

Implementation

The general shape from shading problem as we have defined it is definitely not an easy problem to solve. Inevitably our solution to it involves intense computational burden. Our approach has the merit of being fully parallel and it is in principle well suited to implementation on a massively parallel machine. This is precisely what we have done. The machine we have is a MasPar-1, it is a SIMD (single instruction, multiple data) type machine with 2048 processors forming an toroidal array of 32×64 with 16 K-Bytes of memory per processor.

We have found that the most severe constraint on the machine was its limited memory capacity. Even with small images (64×64), when computing the local support, the number of possible scenels on each fibre is such that the number of pairwise combinations exceeds the machine's storage capacity. The relevant compatibility matrix elements cannot be all stored on the machine's RAM. Hence the machine swaps its memory with a disk which tremendously slows down the actual computation. This has severely limited our experimentation.

1. Numerical Results

We have chosen only a few examples but we chose these to illustrate that our algorithm can resolve the limitations and difficulties that were shown to be typical for the classical shape from shading algorithms.

1.1. Abrupt Albedo Change. The first example that we ran to test our algorithm was an image of a scene in which the albedo changed abruptly (Fig. 9.43(a)). It simply consists of a sphere in front of a plane. This same example was used in Chapt. 3 when we were exposing the limitations and difficulties related to classical shape from shading algorithms (See Fig. 3.16).

We extracted the edge map using the algorithm of Iverson [66] and the shading flow field using the algorithm outlined in Chapt. 5. These are shown in Fig. 9.43(b). The straight edges are interpreted correctly according to Table 6.3 as an abrupt change in albedo, since

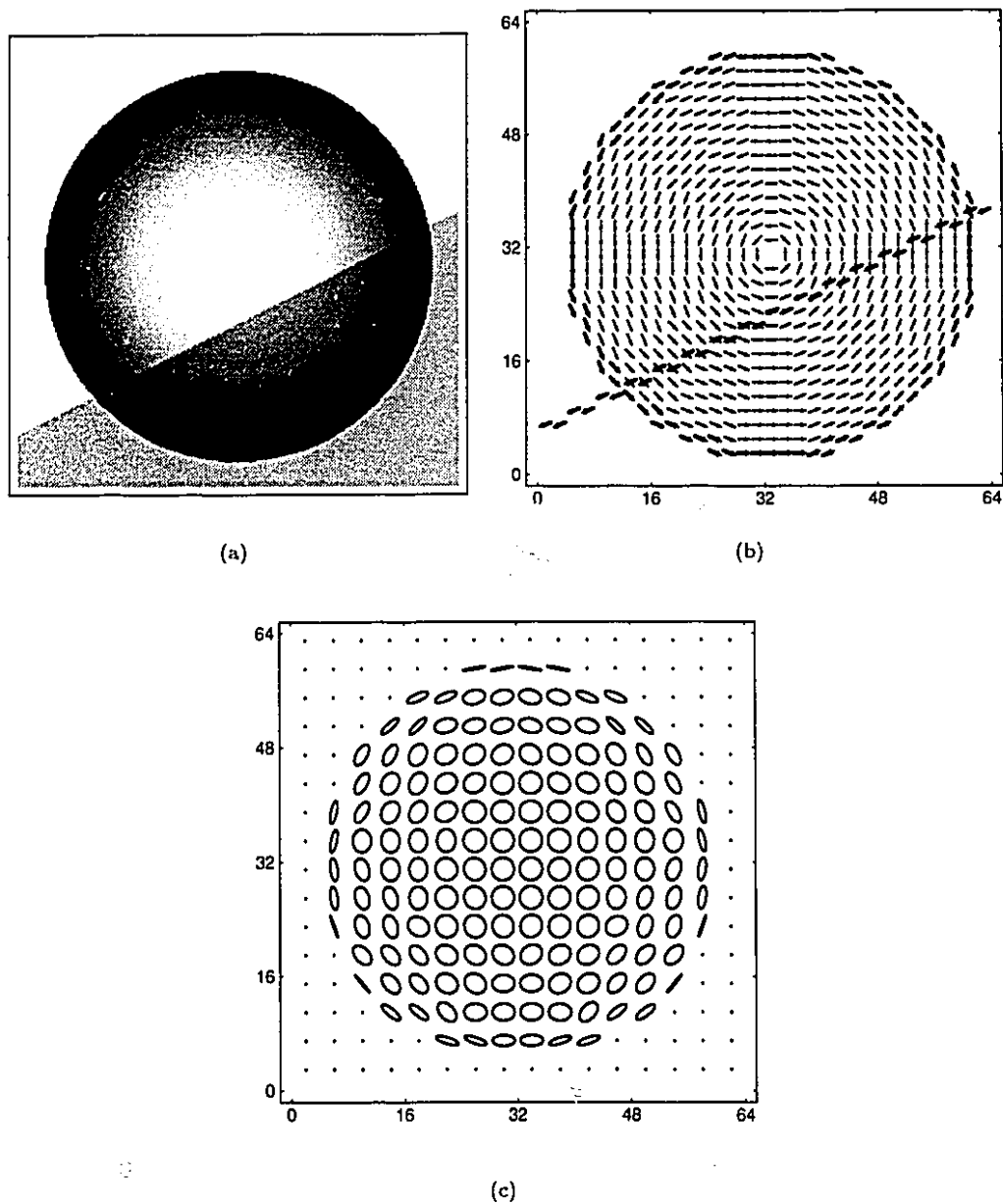


FIGURE 9.43. This test image provides an example of a shape recovery. We implemented an algorithm using the generalized shape from shading assumptions. An ideal intensity image of a simple scene illuminated by a single distant point light source but with an abrupt change in albedo (a) and the corresponding shading flow field (b) are shown in this figure. The shape recovery is illustrated by the surface normals (c). For display purposes only, thresholds are used to avoid cluttering the figures with useless information and the resolution of the edge map and shading flow field is decreased by two and the resolution of the surface normal is decreased by four.

a discontinuity in intensity is present while the shading flow field is continuous. The curved edges that form a circle carry two interpretations since discontinuities in both the intensity and the shading flow field are present:

- (i) either they mark a cast shadow boundary;
- (ii) or they mark a surface discontinuity.

While the artificial scene is not meant to be characterized by a change in lighting condition, such interpretation is not incompatible with the image. The second interpretation is the one which we intend to exploit as these edges mark in fact an occluding contour.

Our shape from shading algorithm yields the result shown in Fig. 9.43(c) after five iterations. The visible surface of the sphere is accurately recovered as well as the illumination direction of the light shining on it. The description of the back plane remains ambiguous as multiple cross-sections of the scene bundle are equally supported. Each cross-section indicates the presence of a planar surface which is accurate. The ambiguity (which cannot be resolved) lies only in the planar surface's orientation.

1.2. Multiple Light Sources. The second example that we ran to test our algorithm was an image of a scene which is illuminated by two distant light sources (Fig. 9.44(a)). Again, it simply consists of a sphere in front of a plane.

As with the previous example, we extracted the edge map using the algorithm of Iverson [66] and the shading flow field using the algorithm outlined in Chapt. 5. These are shown in Fig. 9.44(b). The only edges in this example are the curved edges that form a circle. Again, as in the previous example, they can support two interpretations since discontinuities in both the intensity and the shading flow field are present:

- (i) either they mark a cast shadow boundary;
- (ii) or they mark a surface discontinuity.

Again, while the artificial scene is not meant to be characterized by a change in lighting condition, such an interpretation is not incompatible with the image. The second interpretation is the one which we intend to exploit as these edges mark in fact an occluding contour.

There are discontinuities in the shading flow field along a curved line. Since these are not accompanied by intensity discontinuities, the curved line is correctly interpreted, according to Table 6.3 interpreted as an attached shadow boundary — the surface is thus inferred to be continuous.

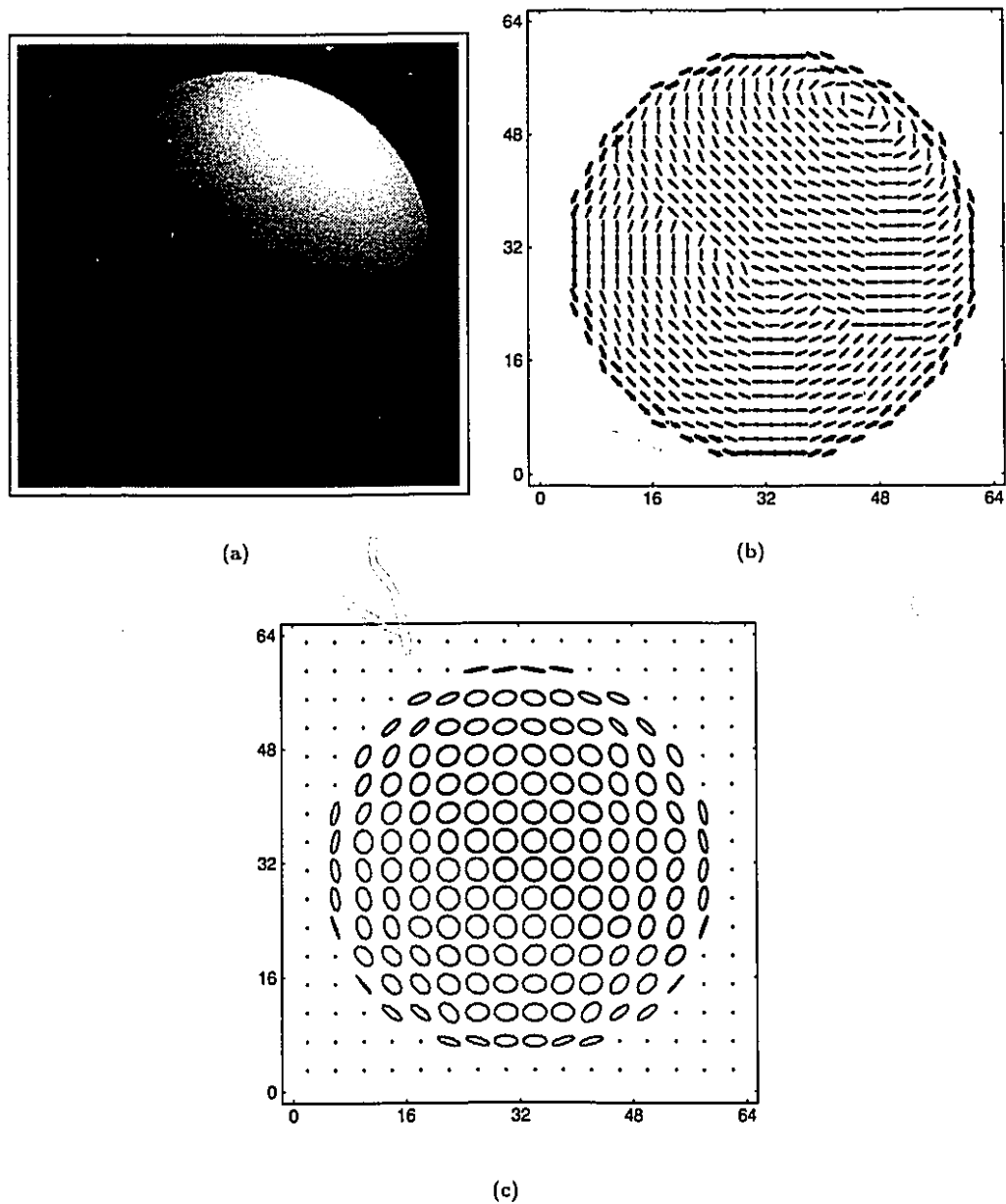


FIGURE 9.44. This test image provides an example of a shape recovery. We implemented an algorithm using the generalized shape from shading assumptions. An ideal intensity image of a simple scene illuminated by two distant point light sources (a) and the corresponding shading flow field (b) are shown in this figure. The shape recovery is illustrated by the surface normals (c). The gray levels code different illuminant directions.

Our shape from shading algorithm yields the result shown in Fig. 9.44(c) after five iterations. The visible surface of sphere is again accurately recovered as well as the illumination direction of the light shining on it. The difference between this result and the previous example is not immediately visible on the figure since the surface orientations are exactly the same — it is only the light source orientation which differs.

As previously, the description of the back plane remains ambiguous as multiple cross-sections of the scene bundle are equally supported. Each cross-section indicates the presence of a planar surface which is accurate. The ambiguity (which cannot be resolved) lies only in the planar surface's orientation.

1.3. The Car's Fender. The third example is a real image of a scene (Fig. 9.45(a)). Consequently, there is sensor noise on top of the omnipresent quantization noise. This time, the scene consists of a car which is illuminated by two distant light sources. We selected a sub-image because of practical constraints, and focused on a close-up of the fender (Fig. 9.45(b)) because of the wealth of features in this small region: there is an occlusion; the fender is partially hiding another surface; there is a cast shadow that runs across the fender; and there is an attached shadow that runs along the surface.

We show in Fig. 9.45(c), the edge map obtained using the algorithm of Iverson [66] and the shading flow field obtained using the algorithm outlined in Chapt. 5. After five iterations, our shape from shading algorithm yields the result shown in Fig. 9.45(d). Unlike the two previous examples, we do not know what the ground truth is. This inconvenience highlights the necessity of being able to generate realistic looking images of artificial scenes. We show in App. D our implementation of a new parallel algorithm that attempts to palliate this inconvenience.

Meanwhile, note that the scene of this example has several qualitative features that allow us to evaluate our results:

- There are edges on both sides of the fender where there are also discontinuities in the shading flow field. As stated for the two previous examples, such instances can support two interpretations:
 - (i) either they mark a cast shadow boundary;
 - (ii) or they mark a surface discontinuity.

The correct one includes a surface discontinuity. Our result also suggests that the normal slants away from the viewer as it approaches the edge from the fender's side. On the other side of the edge, our result gives multiple consistent cross-sections.

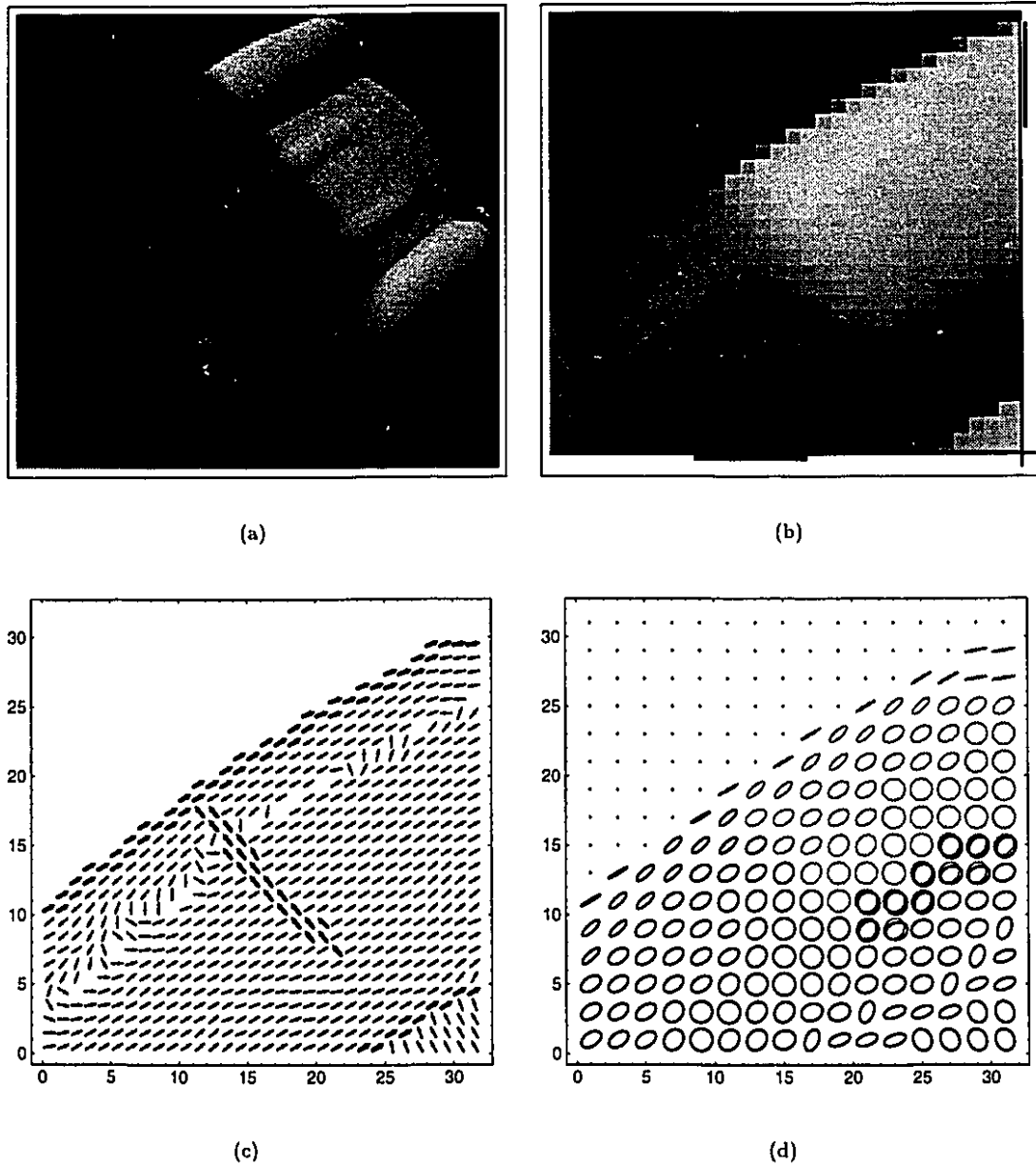


FIGURE 9.45. (a) A real image is used for this last example. The shape recovery is done with (b) a close-up of the fender of a car. In this image, there are several noteworthy features, such as occlusion, cast and attached shadow boundaries. (c) The corresponding edge map and shading flow field are used as input. The edges are depicted as bold arrows, and the shading flows, as thin arrows. (d) The shape recovery is illustrated by the surface normals (c). The gray levels code different illuminant directions.

All these cross-sections correspond to planar surfaces of various orientations and lighting conditions. But planar surfaces are not compatible with the curving fender, hence the surface discontinuity.

Anywhere a surface discontinuity occurs, both the albedo and lighting condition can independently change. The back plane being black, this issue cannot be resolved. Our result is no different in this regard, none of these possibilities are rejected.

- The line singularity running along the fender reveals the parabolic nature of the car's fender. A visual assessment of the surface shape yields a result consistent with this prediction — that is, the fender is roughly cylindrical.
- There are also edges running across the fender. Along most of it there is no discontinuity in the shading flow field but unlike the first example, the shading flows are parallel. This allows an additional possible interpretation according to Table 6.3, which is correct for this example:

(i) The edges mark a cast shadow boundary on a smooth parabolic surface.

The shading flow field is discontinuous only where the edges running across the fender meet the singularity lines. These image geometric structures are consistent with

- (i) either a cast shadow boundary;
- (ii) or a surface discontinuity.

The cast shadow boundary is the correct interpretation.

Our result is consistent with the correct interpretation. The surface is continuous all along this edge, but the illuminant direction differs.

- The attached shadow presents an interesting special case for this example. There is no edge and no shading flow field discontinuity. The parallel shading flow is consistent with either:

- (i) a smooth surface with constant albedo and lighting conditions;
- (ii) or an attached shadow boundary on a parabolic surface.

Our result shows that the surface is continuous where we perceive an attached shadow boundary. It also shows a partial overlap between cross-sections with different illuminant directions. The attached shadow is thus localized in the region defined by the overlap. The localization of the attached shadow on parabolic surface patches could be improved by localizing discontinuities in the magnitude of the intensity gradient. This geometric structure is a type I structure and could, in principle, be extracted from the image [66].

CHAPTER 10

Discussion

By relaxing the constraints of the classical shape from shading problem, we have given shape from shading an entirely new structure.

The data that serve as input to classical shape from shading algorithms have been shown to be inappropriate when the classical constraints are relaxed. The albedo and the illumination are not known since they are functions of (x, y) , not global quantities. Consequently, the photometric values cannot be used because of their dependence on the albedo and the illumination.

We have shown that in order to address the new generalized shape from shading problem, we can make use of other properties of the image — in particular, we exploited the geometrical structures of the image. Both the image curves and the shading flow field can be extracted reliably from the image. The integration of these two cues is shown to be necessary.

We have developed a new computational framework, designed to extract the shape of the surfaces, which integrates different cues: specifically, it integrates the shading flow field and the tangent field. This framework exploits our new local representation of the surface orientation and shape for which continuous transformations of the surface are mapped to continuous paths in a five-dimensional space. It also uses the relationships that we have established between the local descriptors of the scene and the local geometric properties of the images, and between the different neighbouring local scene models.

Our three examples provide an indication that our framework is sound and that the resulting algorithm works. These examples also show the importance of the interpretation of discontinuities as it allows us to relate the geometrical structures of the image to scene properties.

Generalized shape from shading can deal with many situations for which algorithms based on the classical constraints would inevitably fail. There is obviously a much more severe computational burden associated with the generalized problem, and we have tackled this issue by designing a massively parallel algorithm.

1. Future Directions

The work presented in this thesis has provided a new direction in shape from shading. As such, it has answered some questions, but it has also raised new ones. In this section, we describe the future directions that we think might be insightful to explore further.

1.1. Implementation and Experimentation. The images used in our implementation (Chapt. 9) have been chosen to provide a proof of concept. The images depict scenes that contain discontinuities in surface, lighting, and albedo.

However, a more extensive implementation is desirable. Experimentation on more realistic images is needed to characterize the algorithm's behaviour. For this, a different machine is needed, one with enough memory to store the relevant compatibilities for each fibre, and preferably, massively parallel.

1.2. Consequences of the Shading Flow Field. The shading flow field is proposed as an intermediate structure for the generalized shape from shading problem.

- Mathematical properties — If we consider the shading flow field to be a vector field on a manifold, what are its mathematical properties? How do they relate with the texture flow field's properties? What is the relation between the index of shading flow singularities and scene properties such as highlights?
- Psychophysical properties — Does the shading flow field have any psychophysical reality?
 - (i) Geometric style computation. In App. A, we have begun to explore the idea that the primate brain uses a geometrical style of computation. Further ongoing psychophysical experiments focus more precisely on the shading flow field.
 - (ii) Sensitivity issues. How accurate are we in perceiving the various parameters of the shading flow? Can we perceive and accurately localize shading flow discontinuities?

1.3. Integration of other Cues. Experiments supported the sufficiency of the computational framework for integrating the information provided by the tangent field and the shading flow field. What about the other cues available in a single static image?

1.3.1. Visual Texture. Shading and visual texture can be found together [109]. The intensity gradient can be computed within the texture elements in regions where the scale of the visual texture is large. If, in a region, the scale of the visual texture is small, then this part of the image would yield no shading flow field. Consequently, we know that in this part of the image, our shape from shading yields no result either. This is interesting, as it opens the door for integrating another complementary cue to infer shape [43]: visual texture.

1.3.2. Highlights and Shadows. Evidence indicated that the singularities of the shading flow field could provide cues to unveil highlights and shadows. These features also provide other evidence about surface shape as the extrema of the surface luminance have been shown to cling to parabolic lines [72] for matte surfaces.

Although we are mainly concerned with matte surfaces in this thesis, we need to be aware that other types of surface exist. An important class is glossy surfaces. In computer graphics, glossy surfaces are often modelled as the superposition of a matte reflection and a specular reflection. In Fig. 10.46, we show the geometry of the specular reflection. The unit vectors $\mathbf{R}$ and $\mathbf{V}$ show respectively the direction of the pure mirror-like reflection and a viewpoint. For a given surface patch and a given viewpoint, the Phong model [17] describes the intensity of the specular component due to a point light source. When $\mathbf{R} \cdot \mathbf{V} > 0$, the intensity is given by:

$$I = k_d \rho \lambda \mathbf{N} \cdot \mathbf{L} + k_s \lambda (\mathbf{R} \cdot \mathbf{V})^n .$$

The coefficients k_d and k_s give respectively the fractions of the incoming light involved in matte reflection and in specular reflection. The specular-reflection exponent, n , depends on the surface material being simulated. Values of n typically vary from one to several hundred [35]. This equation contains an extra term which can also give rise to singularities in the shading flow field. This would occur when $\mathbf{R} \cdot \mathbf{V} \rightarrow 1$, i.e. the light source, the surface, and the viewer aligned such that there is a mirror-like reflection, $\mathbf{N} \cdot \mathbf{L} = \mathbf{N} \cdot \mathbf{V} = \sqrt{\frac{1}{2}(1 + \mathbf{L} \cdot \mathbf{V})} > 0$.

Torrance and Sparrow [124] developed a geometrical optic model for specular reflection by rough surfaces. Beckmann and Spizzichino [4] on the other hand developed physical optics model for the reflection of plane waves on smooth and rough surfaces. Nayar et al. [98]

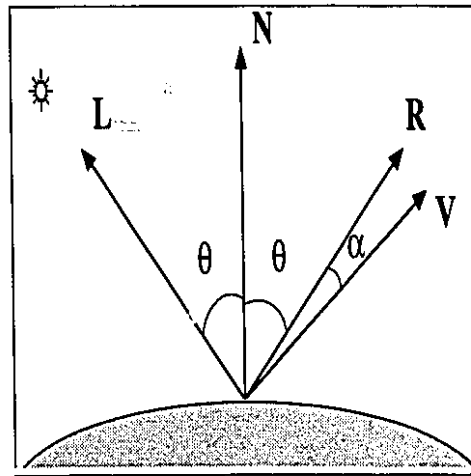


FIGURE 10.46. The vector $\mathbf{R}$ is the direction of the mirror-like reflection of an incoming ray from the direction $\mathbf{L}$ for a surface whose orientation is provided by the normal vector $\mathbf{N}$.

studied the behaviour of both these models and proposed a physical model comprising three reflection components: a diffuse lobe, a specular lobe and a specular spike. The diffuse lobe correspond to Lambert's reflection model of a matte surface.

Since the specular reflection has different spectral and geometrical properties than the matte reflection, it is sometime possible to identify specularities [89]. The shading flow field's singularities may provide a new way to localize and analyze them.

1.4. Improving the Scenel Model. Algorithms fail when underlying assumptions are not all met. To improve upon a model, it is interesting to know when the assumptions break. It is particularly insightful when the algorithm itself can recognize such a situation as it limits the image interpretation to a level where results are reliable.

1.4.1. Nearby and Extended Light Sources. We used the virtual illuminant to model light sources. For lighting conditions such as a point light source, a collection of point light sources, an extended light source that is entirely visible from the surface, or a collection of these, the virtual illuminant is a piecewise constant function. This is consistent with our underlying assumption.

However, nearby and extended light sources may, in some circumstances, generate shading patterns that are inconsistent with our main assumption. For extended light sources, this would occur in the penumbra (see Fig. 10.47). For a scene in which the light source's extent is infinite, Langer and Zucker [83,84] reported that the solid angle of visible light source (or the aperture) then becomes a dominating quantity. The use of our illumination

model uncovers a key characteristic of all these situations: the virtual illuminant direction changes wildly.

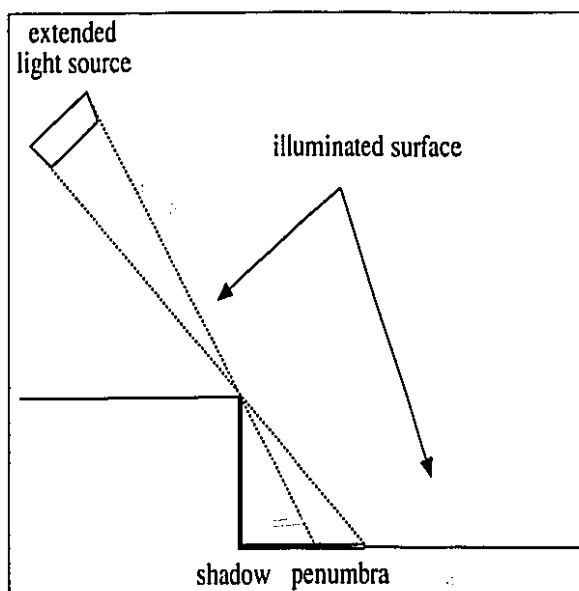


FIGURE 10.47. An extended light source can cause a smooth variation in intensity even if the surface orientation remains constant. This occurs in the penumbra as the amount of the extended light source that is visible from the surface is gradually changing.

The phenomenon of mutual illumination between surface patches (see Fig. 10.48) can become significant if the albedo of the surface approaches unity [36]. When considering inter-reflections, other surface patches are treated as light sources — local extended light sources. The variation of virtual illuminant direction may then provide an indication of the significance of the inter-reflections' contribution.

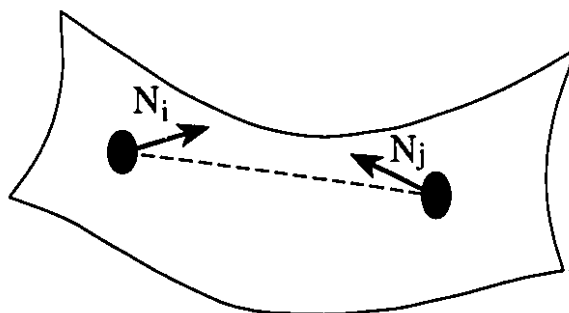


FIGURE 10.48. When a surface patch i and a surface patch j can be joined by a straight line in ambient space (that is not through a solid shape), then the surface patch are visible one from the other. Therefore, if some light shines on the surface patch i , a fraction can be reflected in the direction of patch j and vice-versa. This phenomena is called mutual illumination.

Since the cross-section corresponding to a scene with a varying virtual illuminant direction is said to be inconsistent, we conjecture that for such an image our algorithm would yield a null result.

- (i) A null result is interesting because it constitutes a statement to the effect that our assumption does not apply for this scene, hence a distributed penumbra or some other effect is detected.
- (ii) Observe that in such a situation, classical approaches to shape from shading simply give wrong answers.

There is a caveat to the previous conjecture when a scene that breaks our assumption mimics a scene that doesn't. Such situations could occur naturally (e.g. camouflage of animals) or artificially (e.g. Trompe-l'oeil paintings on buildings). They are considered *illusions* and can potentially confuse both human and machine vision systems.

Consider the canonical example: a photograph (where the albedo varies continuously), a projected slide (where the illumination on the screen varies continuously) and the scene itself. Any shape from shading process should reconstruct shape information identically. In absence of other cues, this ambiguity cannot be resolved and thus, we have intentionally chosen to concentrate our efforts on complex surfaces but with simple albedo and illumination variations.

1.4.2. *Light Interaction with the Ambient Medium.* Our setting of shape from shading assumes that light interacts with surfaces and is captured to form the image. But light can also interact with the medium in which it is travelling. Diffusion and absorption of light are two common phenomena that could occur when light interacts with fog, dust or smoke. Here again, these phenomena can cause a smooth variation of intensity.

In order to pursue our investigation on these limiting cases, we need images for which we have complete knowledge of scene; i.e. at every point where light interact with matter, the shape and reflectance properties of the surface element or the absorption and diffusion characteristic of the medium in which light travel, the lighting condition. We have begun work for the synthesis of such image on a massively parallel computer. Our current implementation is detailed in App. D.

Appendices

APPENDIX A

Flows and Scalar Field: Evidence for Multiple Pathways

In this thesis, we have considered separately the photometric and the geometric properties of the image. To emphasize this distinction here, we use a parallel with texture analysis. Density is to texture what intensity is to shading. Whereas we have shading flows for shading analysis, we have texture flows for texture analysis.

This appendix describes psychophysics experiments designed to examine an hypothesized representation of scalar quantities and vector properties in the visual cortex.

1. Introduction

The notion of parallel processing pathways has been integral to psychophysical modeling for some time, and has led to various “independence” hypotheses in early vision. Early physiological evidence was provided by the separation of ON and OFF channels in the retina, but more recently the emphasis has been on functional specialization [134]. Perhaps the most explicit such proposal has been made by Cavanagh [20], who posited separate channels for luminance, motion, binocular disparity, colour, and texture following the striate cortex. These channels are based in part on the belief that V4 is an area specialized for colour [132], and have been strengthened by the discovery of cytochrome oxidase blobs in visual area V1, coupled with the anatomy of connections through V2 to V4 [133]. We shall focus on this functional specialization; namely that the blobs support a color system, while the “interblob” pathways support a luminance system, because it has become rather prominent in the literature [90, 131]. However, we argue that, while the color/luminance distinction is attractive pragmatically, the evidence is that the blob/interblob distinction is much more subtle functionally. Allman and Zucker [1] summarized it as follows: (i) cells within blobs are sensitive to contrast as well as color; (ii) some animals with color vision appear to be lacking blobs; and (iii) some animals without color vision appear to have them.

Thus, while the colorization process may be pleasing for viewing classical motion pictures, its evolutionary value is not at all clear.

New proposals regarding blob/interblob function are required, and we review one in the next section. To stress the functional specialization beyond the domain of contrast and color, we enlarge the argument to show how it suggests a novel view of the representation of texture in visual cortex. A series of psychophysical experiments are then reported in support of the view. The conclusion is that a more appropriate model of blob/interblob functional specialization transcends the above pragmatically-defined categories to others that are more abstract. The different classes of feature are categorized mathematically into those that are scalar as opposed to those that are (differentially, or locally) geometric.

1.1. The Allman/Zucker Hypothesis. Allman and Zucker [1] have advanced the hypothesis that blob cells are not only coding different information from interblob cells, but they are encoding it differently. They suggest in particular that blob cells are selective for scalar information and encode it by frequency of firing, while interblob cells are selective for geometric variables and encode a strength of match. Representational differences emerge as well, as follows.

A classical view of the visual cortex is as columnar organizations of cells responsive to specific stimulus properties. Key among these is orientation, one of the most prominent *geometric* features of our environment. Orientation hypercolumns consist of cells tuned to different orientations, and can be modeled as representing (a sampling of) all orientations at each retinotopic position. The firing rate of such cells can be interpreted in proportion to how well the stimulus orientation matches the cell's preferred orientation; i.e., as a strength of match. However, since strength of match is confounded with stimulus contrast, such cells typically saturate within about one order of magnitude of log contrast [92]. This facilitates finding the border around and within natural objects, given the variations imposed by lighting in natural scenes. Their firing rate rarely exceeds about 100 spikes/second.

Cells within the cytochrome oxidase blobs, however, respond to other stimulus properties, such as contrast and color. Observe that these are *scalar* properties, in that they can be represented by a single number. (Geometric properties, such as orientation, require a vector.) Now, what defines the blobs is an energy measure: cytochrome oxidase is an enzyme that indicates energetic capacity. How is it, then, that cells within the blobs stain for more cytochrome oxidase activity than those between blobs (in the orientation hypercolumns)?

The Allman/Zucker proposal specifically claims that, if scalar variables such as contrast were encoded by firing rate within blob cells, then such cells would require the energetic

capacity to sustain a broad dynamic range. This will lead to a high concentration of cytochrome oxidase. While firing rates for cells within blobs are just now being studied, the indirect evidence is that cells that are likely candidates for being located within blobs respond to over 200 spikes/sec and exhibit contrast saturation after 3 log orders of magnitude. In addition, contrast is available at every location within the image, so contrast-signaling cells will always be active to some extent. Orientation, on the other hand, will be single valued at most image points, so the majority of cells within an orientation hypercolumn will be quiet. Hence the difference in energetic capacity.

1.2. Two Classes of Textures. With this background, we can now state the formal observation motivating our psychophysical experiments: the scalar vs. geometric variable difference arises within textures, as well as between boundaries and contrasts. In particular, (i) there are those texture patterns, such as stars on a clear night, points of light shining through foliage, or grains of sand on a beach, whose arrangements are characterized primarily by variables such as density; and (ii) there are those texture patterns, such as fur, hair, grass, or wheat, that are characterized by their orientation structure, or flow. We shall refer to these classes of patterns as (i) texture point fields and (ii) texture flows, respectively, to stress the differences. Our current interest is to determine whether texture point fields and texture flows exhibit different psychophysical characteristics, as a prolegomena to determine whether one is carried by the blob system and the other by the interblob system.

We shall use different types of dot patterns to represent these two classes of textures, with random dot patterns obviously representing texture point fields. Texture flows will be represented with random dot Moiré patterns (RDMP), or Glass patterns [44], composed as follows: begin with an original random dot pattern, and make a copy of it. Now, shift each dot in the copy by a given transformation, and superimpose the shifted copy onto the original. If the transformation were a lateral displacement, for example, then the composite would depict a linear flow; locally the flow is carried by the pairs of transformed dots. Other transformations might include rotations, expansions, etc. Finally, a second copy can be made, transformed from the first copy, and then superimposed onto the original plus first copy; the resultant even richer pattern will be composed of triples of points arranged according to the given transformation. We refer to the number of dot patterns comprising the RDMP as the *path length*; note that longer path lengths correspond to more spatial structure in the flow pattern. A path length of 1 is, of course, a texture point field.

While the psychophysical category of texture includes both texture point fields and texture flows, we shall focus on the differences. Texture point fields are dominated by their

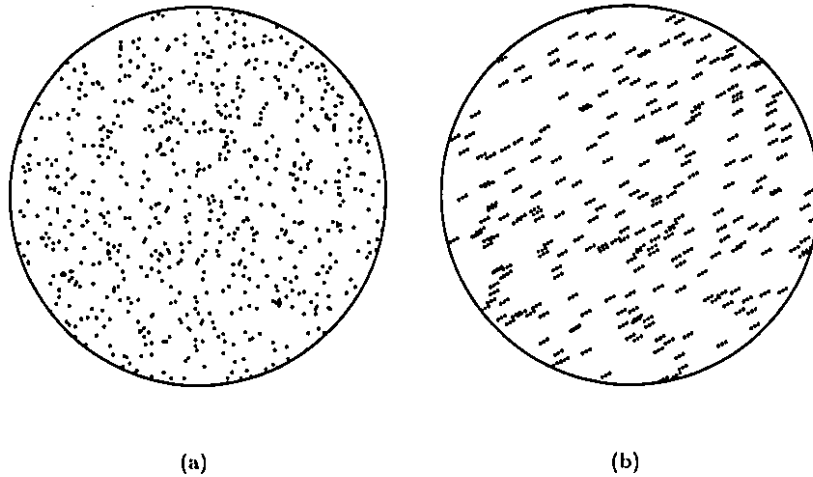


FIGURE A.49. This figure provides an illustration of the two types of texture patterns used in the experiments. (a) A random dot texture point field, characterized by its density. (b) A random dot Moiré (or Glass) pattern [44], constructed as follows: begin with an original random dot pattern, and make a copy of it. Now, shift each dot in the copy by a given transformation, and superimpose the shifted copy onto the original. The transformation shown is a linear displacement, and the composite is a linear flow; other transformations are also possible, e.g. rotations, expansions, etc. Finally, a second copy is made, transformed from the first copy, and then superimposed onto the original plus first copy; the resultant pattern is now composed of triples of points arranged according to the given transformation. We refer to the number of dot patterns comprising the random dot Moiré pattern as the *path length*; longer path lengths correspond to more spatial structure in the flow pattern. A path length of 1 is, of course, a texture point field. The advantage of using such dot patterns is that density can be held constant while they are arranged to carry increasingly more geometric structure.

first-order statistics, or density, while an important second-order term is introduced for RDMPs with path length = 2. Barlow [3] showed that, for texture point fields, density was the key informational variable, and that the information in specific pattern structures, or arrangements of dots, was not utilized. Julesz [67], in his texton theory, made them distinct, and included them among other texture primitives.

2. Preliminary Observations

There is a classical dictum attributed to the early Disney artists that, to make convincing cartoons, effort should be concentrated on borders rather than color overlays; gaps and smudges will be perceptually filled in by the interior color. This informal observation was confirmed by Kolers and von Grünau [77] (discussed below), and provides the basis for our first informal examination of texture filling-in phenomena: since textures, like colours, depict surface coverings, one would expect them both to fill-in moving contour boundaries. Our first display confirmed this expectation, but also provided a hint that texture point fields and texture flows are not treated identically.

Displays were created on a Silicon Graphics IRIS Personal Workstation and they were shown on a Silicon Graphics Color Monitor model #CM2086A3SG at approximately 80 cm from the subject. The refresh rate is 60 hertz, the horizontal and vertical resolutions of the monitor are respectively 3765 and 3793 pixels/meter, and each RGB channel has 8 bits.

The figure consisted of a circular border, approximately 2 degrees to 4 degrees in diameter, and which was filled with a jittering RDMP. The filled, circular figure underwent motion, as follows. The entire figure followed a path around another circle, approximately 6 degrees in diameter. The center of the border circle followed the path exactly, but the interior RDMP jittered around this path, so that at times there were gaps between the interior texture and the border, and at times the interior texture went outside of rotating border. The subject was instructed to fix his view at the center of the 6 degree path, and to report when either a gap appeared within the border, or an "arm" of texture broke through it. The path length of the interior RDMP was varied from 1 – 5.

OBSERVATION A.1. From the response of our subjects, it is immediately apparent that gaps and arms are more visible for the longer path lengths than for the shorter ones.

Whereas texture point fields seem to "fill in" like colours in Disney's observation, texture flows do not. The Glass pattern appeared more detachable than a random dot pattern

with the same average density. Because the question is confounded by prominent outliers, this experiment is difficult to quantify. So we move to a different paradigm.

Unlike colour and contrast, texture filling-in has never been systematically studied. There is indirect evidence that intensity and texture density are confounded variables [97, 127] and that texture fields can be segregated from texture flows [46]. The implications are larger than variable confounding, however, because of the possibility of parallel processing streams would arise if the representation of texture fields were decoupled from texture flows.

3. Experimental Paradigm

We now explore the differences between texture fields and texture flows psychophysically, using a paradigm developed by Kolers and von Grünau [77]. Building on an earlier observation that the boundary of an object deforms smoothly during apparent motion [76], they explored how the color interior to a shape changed during such apparent motion. The experiment placed two shapes in apparent motion, with color₁ in shape₁, and color₂ in shape₂. If shape₁ were, say, a square, and shape₂ a triangle, then the boundary percept would be of a square deforming smoothly into a triangle. Their real concern was with the interior color, and two observations are relevant:

- (i) the color appeared to stay within the deforming contour, completely filling it but not extending outside of it;
- (ii) when color₁ differed from color₂, the change was abrupt from color₁ to color₂.

Thus, unlike the boundary, color did not deform smoothly. Similar results hold for different contrasts as well.

This difference between abrupt color changes and smooth boundary changes suggests a psychophysical measure that discriminates between those properties carried by the blob system and those carried by the interblob (hypercolumn) system: when the information from these systems is integrated, changes in blob variables (color or contrast) appear abrupt, while changes in interblob variables (boundary) appear smooth. With this operational assumption in place, we can now formulate our specific hypothesis: (scalar) information about texture fields (density) is carried by the blob system, while (geometric) information about texture flows is carried by the interblob system. We thus predict that changes in texture density will be abrupt during apparent motion, while changes in texture flow orientation will be smooth.

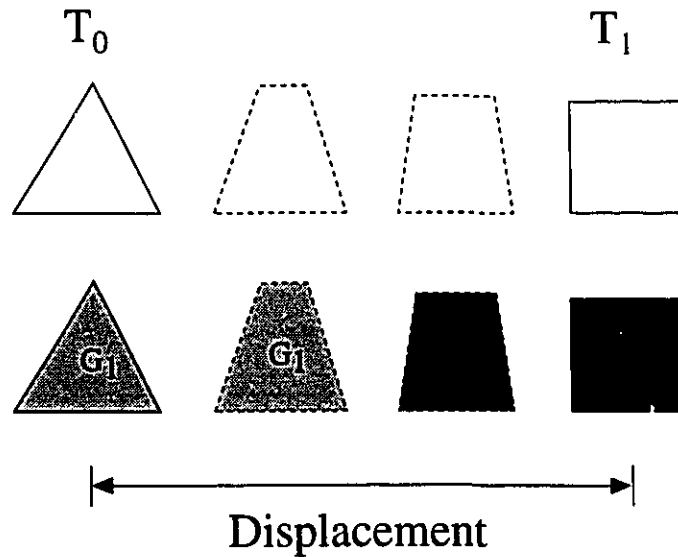


FIGURE A.50. This figure illustrates how boundaries change during apparent motion, and how the color interior to a shape fills in the changing contour. (*top*) Depiction of an apparent motion sequence in which a triangle deforms smoothly into a square while moving from left to right. The percept is that of a triangle deforming smoothly into a square while moving from left to right. The figures that are actually displayed are shown in solid lines, while a rendition of the apparent shape at intermediate positions is shown dotted. (*bottom*) When the initial and final figures are filled with either a single color or contrast, the same color or contrast appears to fill the intermediate apparent contours completely. However, when the initial and final colors or contrasts differed, the appearance was of an abrupt change in the color within the moving, apparent contour. This is depicted in the bottom sequence in which gray level abruptly switches from G_1 to G_2 at about the midpoint.

4. Experimental Procedure

Our subjects were either graduate students in our research centre or visitors. None of them had prior knowledge about the experience or the hypothesis being tested. The experiments took place in our laboratory in normal use but with dim light condition. The stimuli were created and displayed in the same set-up as the previous paradigm.

The subject was asked to fix a point at the centre of the screen and then a pair of stimuli were presented. The first stimulus appears at the left of the fixation point for only 150 msec; then the screen remains blank for 50 msec; then finally the second stimulus appears at the right of the fixation point for again only 150 msec. The apparent displacement is approximately 3° of visual angle.

The subject was initially shown identical stimuli to get familiar with apparent motion. Once this stage was completed, the subject was shown pair of stimuli that often differed. The subject was asked whether the filling change abruptly or not.

The first types of filling tested were intensity and colour to reproduce and complement Kolers and von Grünau results.

(i) We considered two series of stimuli with intensity:

- (a) On a black background, we use a white stimulus as the reference and the other stimulus varying from white to dark gray.
- (b) On a white background, we use a black stimulus as the reference and the other stimulus varying from black to light gray.

(ii) We considered three series of stimuli with colour:

- (a) On a gray background, we use a red stimulus as the reference and the other stimulus varying from red to green (via yellow).
- (b) On a gray background, we use a blue stimulus as the reference and the other stimulus varying from blue to red (via magenta).
- (c) On a gray background, we use a blue stimulus as the reference and the other stimulus varying from blue to yellow (via cyan and green).

For the colour stimuli, we determined a set of isoluminant colours for each individual subject just prior to the experiment. This was done by rapidly alternating two colours; the first colour is the reference and the second is adjusted by the observer until the scintillation ceases or is minimal. This was done in order to avoid the confusion between hue and intensity.

We designed the second types of filling to test the prediction of the Allman/Zucker hypothesis with respect to texture point field and texture flow.

(i) We considered two series of stimuli with texture point field:

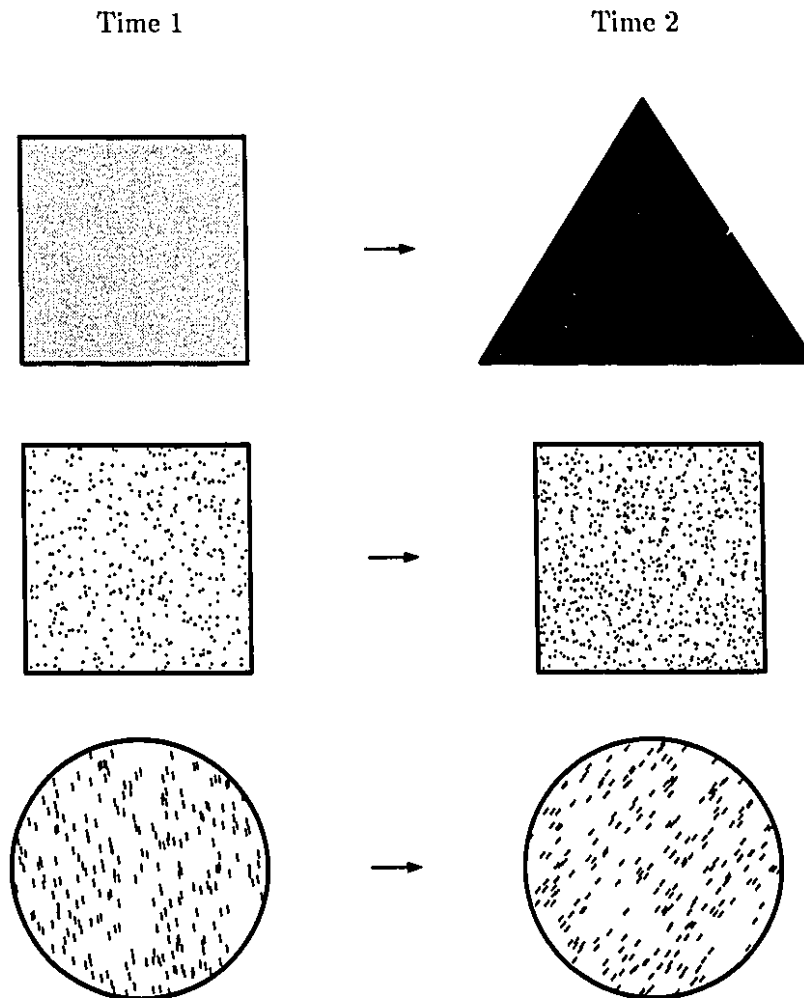


FIGURE A.51. The paradigm is identical to Fig. A.49, in which (*top*) we confirmed the Kolers results. We then extended the interior patterns to textures, using both (*middle*) texture fields and (*bottom*) texture flows. The (*middle*) pattern illustrates a change in texture density. Two sets of trials: for the first, one stimuli had a density of 480 dots in 32,400 pixels (the figure was in a 180 x 180 window), and the other was more dense; for the second, one stimuli had 2560 dots and the other was less dense. The (*bottom*) pattern illustrates a change in texture flow orientation. The initial flow orientation was taken as 0°, 10°, 20°, 30°, and 45°, and the final orientation was taken as 45°, 30°, 20°, 10°, 0°, -10°, -20°, -30°, -45°. Trials with Glass patterns of path lengths equal to 3 (shown) and 5 (not shown) were carried out. For our experiments we used objects sub-tending 3 degrees in visual angle, displayed on a Silicon Graphics Personal Iris (color display model no. CM2086A35G) in a dimly illuminated room. Subjects had no direct knowledge of the experimental questions, and all had normal or corrected vision.

- (a) On a black background, we use as the reference a random dot pattern of a fixed dot density and for the other stimulus any of a range of other random dot patterns with a varying dot density — in this experiment we keep the fixed dot luminance constant.
- (b) On a black background, we use as the reference a random dot pattern of a fixed dot density and for the other stimulus any of a range of other random dot patterns with a varying dot density — in this experiment we keep the fixed average luminance constant.
- (ii) We considered one series of stimuli with texture flow:
 - (a) On a black background, we use as the reference a random dot Moiré pattern of with a given orientation and for the other stimulus any of a range of other random dot Moiré patterns with varying orientation (up to 90°).

A difficulty arises because the average intensity and average contrast of random dot patterns are correlated with dot density. In an attempt to assess if the potential bias is significant, we look at both extreme instances: constant average intensity and constant average contrast.

5. Results

All graphs plot the fraction of displays in which subjects reported an abrupt change in the interior region (ordinate) vs. the stimulus dimension of interest (abscissa). In each experiment the initial display was filled with a pattern given by the leftmost dot along the abscissa. For no change in the pattern, all subjects reported no abrupt changes (first data point). For significant differences between the initial and final displays, all experiments show a psychophysical curve with a steep slope, indicating the rapid onset of apparent abrupt changes with interior differences.

In Fig. A.52, we present the result of four subjects for both experiments in which only the luminance differs in the two stimuli presented. For both negative and positive contrast, an abrupt change is usually perceived when the stimuli are quite different.

In Fig. A.53, we present the result of respectively five, five and three subjects for the experiments in which only the hue differs (red $\rightarrow$ red-green, blue $\rightarrow$ blue-red, blue $\rightarrow$ blue-yellow) in the two stimuli presented. Again, for all combinations, an abrupt change is usually perceived when the stimuli are quite different.

We note that our results for the luminance and colour experiments are in agreement with Kolers and von Grünau's result.

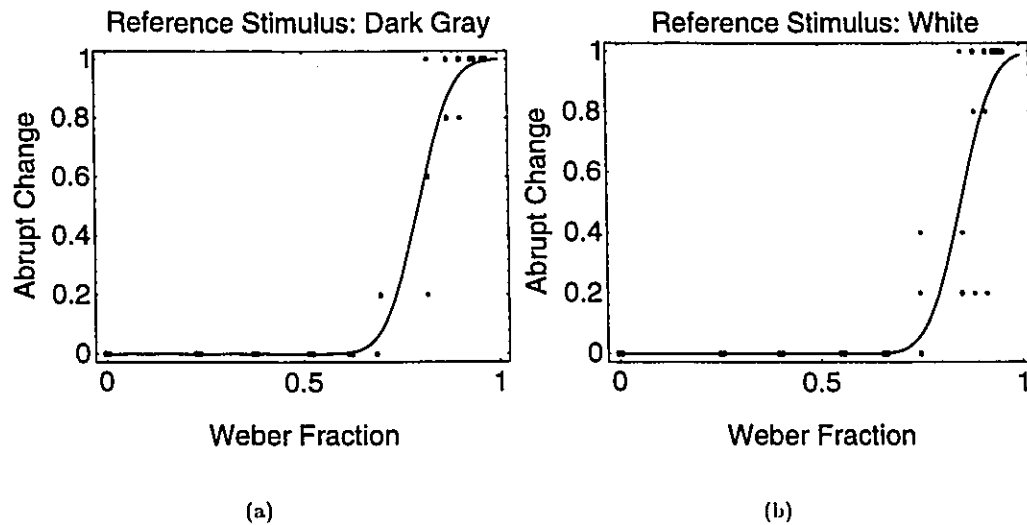


FIGURE A.52. Results of four experiments of apparent motion. Two stimuli are successively shown. In (a), the colour of the first was red and the colour of the other varied in hue but was of the same perceived luminance. In (b), the first was dark and the relative luminance of the other varied.

In Fig. A.54, we present the result of eight subjects for both experiments in which the dot density differs in the two stimuli presented. Whether the luminance is constant or the contrast is constant, an abrupt change is usually perceived when the stimuli are quite different. Our perception of texture density is thus similar to our perception of colour and luminance.

Finally, no data are graphed for the texture flow experiment, as no subject reported an “abrupt” response under any condition examined. We note here that the perception of texture flows is more like the shape boundary — a smooth deformation is perceived. This result indicates that the perception of texture flows is quite different than the perception of texture fields.

6. Discussion

Our psychophysical results provide further evidences to support the claim that the cortex treats differently the scalar properties and geometrical properties. This claim suggests that for computer vision, both the scalar and geometrical images properties be considered, and possibly on their own.

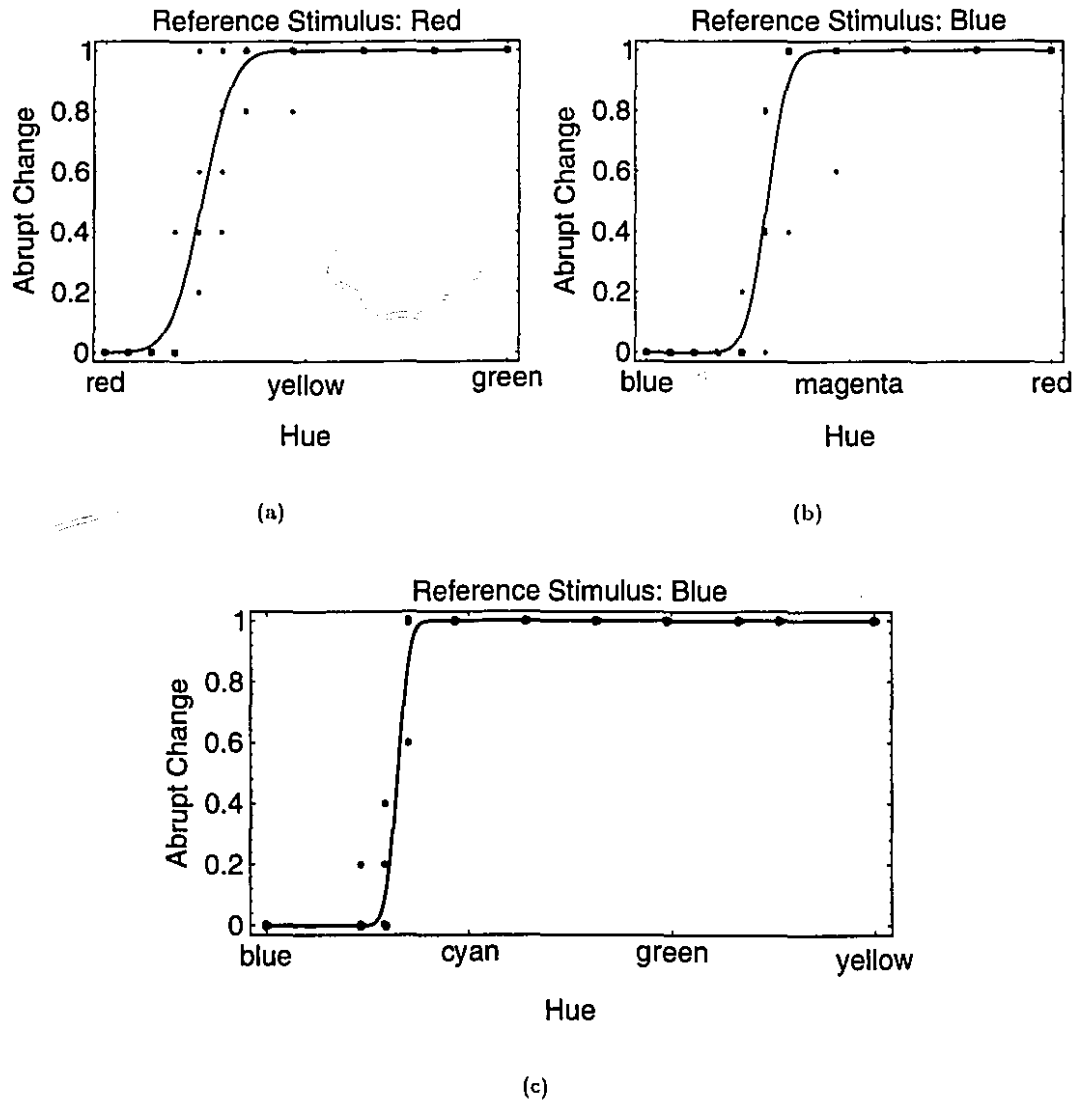


FIGURE A.53. Results of series experiments of apparent motion. Two stimuli are successively shown. In (a), the colour of the first was red and the colour of the other varied in hue but was of the same perceived luminance. In (b), the first was dark and the relative luminance of the other varied. In both (c).

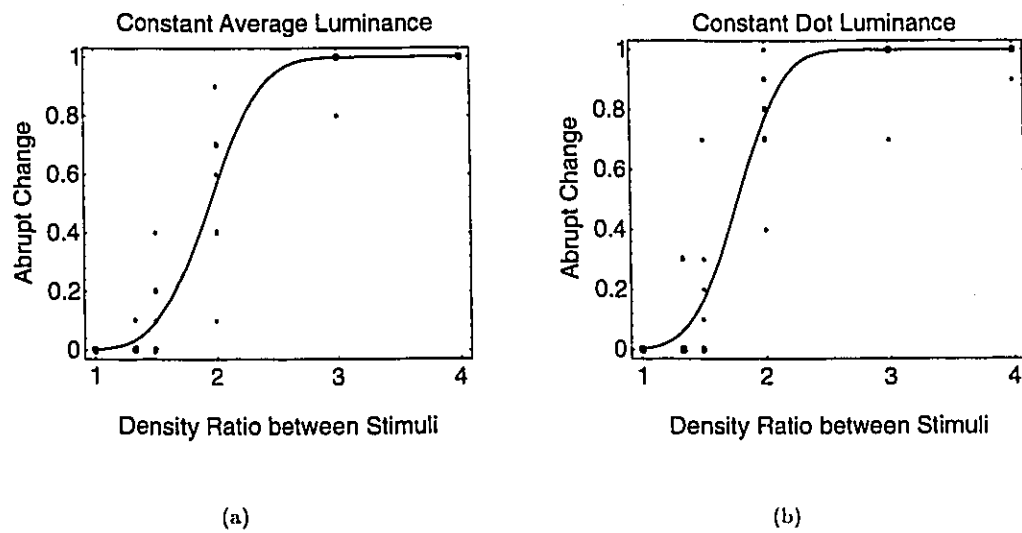


FIGURE A.54. Results of two experiments of apparent motion. Two stimuli are successively shown. In both (a) and (b), it is the dot density that varies. In (a), the average luminance remains constant whereas in (b), the dot luminance remains constant.

APPENDIX B

Higher Order Constraints

The local representation of the surface could be enriched to include the variation of the principal curvatures κ_1, κ_2 in the two principal directions.

The spin forms and the variation of the principal curvatures could be used to further constrain the surface. The spin forms provide a constraint on both the principal directions and the principal curvatures.

These descriptors are necessary to predict the curvature of the gradient and the curvature of the shading flow field from the scenes.

To implement a system making use of these quantities, more memory would be needed. The implementation presented in the body of the thesis already suffers from the current computer hardware limitations. It is nevertheless interesting to look at the relation between intensity image properties and other scene descriptors.

1. Definitions of the Coordinate Systems

In the following sections, we will consider a few different coordinate systems.

- The viewer's coordinates (x, y, z) are defined such that the vector $(0, 0, 1)$ points toward the viewer. Note that, for the image formation, we will assume orthographic projection, hence, the image lies in the x - y plane.
- The surface patch local coordinates (e_1, e_2, e_3) are defined such that, at a given point, the vector $(0, 0, 1)$ indicates the direction of the surface normal, and that the vectors $(1, 0, 0)$ and $(0, 1, 0)$ correspond to the surface principal directions.
- The principal frame field (f_1, f_2, f_3) is defined such that, at every point, the vector $(0, 0, 1)$ indicates the direction of the surface normal, and that the vectors $(1, 0, 0)$ and $(0, 1, 0)$ correspond to the surface principal directions.

2. Shading and Spin forms

Our basic hypothesis for shape from shading is that all smooth variations of the image intensity are due to smooth variations of the surface orientation. First we describe smooth variations of the surface orientation and then, we relate the smooth variations of the surface orientation and the image intensity.

2.1. The Partial Derivative of a Surface Patch Normal. On a smooth surface, the normal is a smooth vector field so it is differentiable. For any point (x_i, y_i, z_i) on such a surface, we can define the surface patch local coordinates $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. The normal of the surface can then be expressed as a function of these coordinates $\mathbf{N}(\mathbf{e}_1, \mathbf{e}_2)$ such that $\mathbf{N}(0,0) = \mathbf{e}_3$. In this section, we look at both the first and second order partial derivatives of the normal with respect to the surface patch local coordinates. We start here with the first:

$$\frac{\partial}{\partial \mathbf{e}_i} \mathbf{N} = \frac{\partial \mathbf{f}_1}{\partial \mathbf{e}_i} \frac{\partial}{\partial \mathbf{f}_1} \mathbf{N} + \frac{\partial \mathbf{f}_2}{\partial \mathbf{e}_i} \frac{\partial}{\partial \mathbf{f}_2} \mathbf{N} ,$$

In differential geometry, the study of how the principal frame of a surface changes is classical. In the nineteenth century, several mathematicians contributed to this field — among them Cartan, Codazzi, Gauss, Mainardi, Rodriguez, Weingarten obtained results very relevant for this section. Their results are nicely explained in some differential geometry textbooks; see [34, 70, 102, 120].

To evaluate the partial derivatives of the principal frame (of which the normal is the third component $\mathbf{f}_3$), we use the Cartan matrix. More precisely, we use the independent components $\tilde{\omega}^{ij}$ which are called the connection forms of the frame field.

DEFINITION B.1. *The connection form $\tilde{\omega}^{ij}(\mathbf{f}_k)$ is the rate of turn of the frame vector $\mathbf{f}_i$ toward the frame vector $\mathbf{f}_j$ when the point of application moves in the direction of the frame vector $\mathbf{f}_k$*

Considering that

$$\frac{\partial}{\partial \mathbf{f}_i} \mathbf{f}_j = \sum_k \tilde{\omega}^{jk}(\mathbf{f}_i) \mathbf{f}_k ,$$

in conjunction with Rodriguez formulas,

$$\begin{aligned} \tilde{\omega}^{13}(\mathbf{f}_1) &= -\tilde{\omega}^{31}(\mathbf{f}_1) = k_1 , \\ \tilde{\omega}^{23}(\mathbf{f}_2) &= -\tilde{\omega}^{32}(\mathbf{f}_2) = k_2 , \end{aligned}$$

give

$$\begin{aligned}\frac{\partial}{\partial \mathbf{f}_1} \mathbf{f}_3 &= \tilde{\omega}^{31} \langle \mathbf{f}_1 \rangle \mathbf{f}_1 + \tilde{\omega}^{32} \langle \mathbf{f}_1 \rangle \mathbf{f}_3 = -k_1 \mathbf{f}_1 , \\ \frac{\partial}{\partial \mathbf{f}_2} \mathbf{f}_3 &= \tilde{\omega}^{31} \langle \mathbf{f}_2 \rangle \mathbf{f}_1 + \tilde{\omega}^{32} \langle \mathbf{f}_2 \rangle \mathbf{f}_3 = -k_2 \mathbf{f}_3 .\end{aligned}$$

The relation between the local coordinates and the frame field is trivial for the first order differentiation.

$$\frac{\partial \mathbf{e}_i}{\partial \mathbf{f}_j} = \delta_{ij}$$

Thus the first order partial derivatives are

$$\begin{aligned}\frac{\partial}{\partial \mathbf{e}_1} \mathbf{N} &= -k_1 \mathbf{f}_1 , \\ \frac{\partial}{\partial \mathbf{e}_2} \mathbf{N} &= -k_2 \mathbf{f}_2 .\end{aligned}$$

there are two parameters needed to describe them and these are the two principal curvatures k_1, k_2 .

The second order partial derivatives of the normal is slightly more complicated to obtain.

$$\begin{aligned}\frac{\partial^2}{\partial \mathbf{e}_j \partial \mathbf{e}_i} \mathbf{N} &= \frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{e}_j \partial \mathbf{e}_i} \frac{\partial}{\partial \mathbf{f}_1} \mathbf{N} + \frac{\partial \mathbf{f}_1}{\partial \mathbf{e}_i} \left[\frac{\partial \mathbf{f}_1}{\partial \mathbf{e}_j} \frac{\partial^2}{\partial \mathbf{f}_1^2} \mathbf{N} + \frac{\partial \mathbf{f}_2}{\partial \mathbf{e}_j} \frac{\partial^2}{\partial \mathbf{f}_2 \partial \mathbf{f}_1} \mathbf{N} \right] \\ &+ \frac{\partial^2 \mathbf{f}_2}{\partial \mathbf{e}_j \partial \mathbf{e}_i} \frac{\partial}{\partial \mathbf{f}_2} \mathbf{N} + \frac{\partial \mathbf{f}_2}{\partial \mathbf{e}_i} \left[\frac{\partial \mathbf{f}_1}{\partial \mathbf{e}_j} \frac{\partial^2}{\partial \mathbf{f}_1 \partial \mathbf{f}_2} \mathbf{N} + \frac{\partial \mathbf{f}_2}{\partial \mathbf{e}_j} \frac{\partial^2}{\partial \mathbf{f}_2^2} \mathbf{N} \right]\end{aligned}$$

We use the independent components $\tilde{\omega}^{ij}$ of the Cartan matrix — these are called the connection forms of the frame field.

This time, we need all the connection equations. These equations allow us to describe the smooth variation of the principal frame field in terms of four parameters: the principal curvatures k_1, k_2 , and the spins s_1, s_2 .

$$\begin{aligned}
\tilde{\omega}^{12} \langle \mathbf{f}_1 \rangle &= -\tilde{\omega}^{21} \langle \mathbf{f}_1 \rangle = s_1 \\
\tilde{\omega}^{13} \langle \mathbf{f}_1 \rangle &= -\tilde{\omega}^{31} \langle \mathbf{f}_1 \rangle = k_1 \\
\tilde{\omega}^{23} \langle \mathbf{f}_1 \rangle &= -\tilde{\omega}^{32} \langle \mathbf{f}_1 \rangle = 0 \\
\tilde{\omega}^{12} \langle \mathbf{f}_2 \rangle &= -\tilde{\omega}^{21} \langle \mathbf{f}_2 \rangle = -s_2 \\
\tilde{\omega}^{13} \langle \mathbf{f}_2 \rangle &= -\tilde{\omega}^{31} \langle \mathbf{f}_2 \rangle = 0 \\
\tilde{\omega}^{23} \langle \mathbf{f}_2 \rangle &= -\tilde{\omega}^{32} \langle \mathbf{f}_2 \rangle = k_2
\end{aligned}$$

$$\frac{\partial}{\partial \mathbf{f}_i} \mathbf{f}_j = \sum_k \tilde{\omega}^{jk} \langle \mathbf{f}_i \rangle \mathbf{f}_k$$

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{f}_1} \mathbf{f}_1 &= \tilde{\omega}^{12} \langle \mathbf{f}_1 \rangle \mathbf{f}_2 + \tilde{\omega}^{13} \langle \mathbf{f}_1 \rangle \mathbf{f}_3 = s_1 \mathbf{f}_2 + k_1 \mathbf{f}_3 \\
\frac{\partial}{\partial \mathbf{f}_1} \mathbf{f}_2 &= \tilde{\omega}^{21} \langle \mathbf{f}_1 \rangle \mathbf{f}_1 + \tilde{\omega}^{23} \langle \mathbf{f}_1 \rangle \mathbf{f}_3 = -s_1 \mathbf{f}_1 \\
\frac{\partial}{\partial \mathbf{f}_1} \mathbf{f}_3 &= \tilde{\omega}^{31} \langle \mathbf{f}_1 \rangle \mathbf{f}_1 + \tilde{\omega}^{32} \langle \mathbf{f}_1 \rangle \mathbf{f}_2 = -k_1 \mathbf{f}_1 \\
\frac{\partial}{\partial \mathbf{f}_2} \mathbf{f}_1 &= \tilde{\omega}^{12} \langle \mathbf{f}_2 \rangle \mathbf{f}_2 + \tilde{\omega}^{13} \langle \mathbf{f}_2 \rangle \mathbf{f}_3 = -s_2 \mathbf{f}_2 \\
\frac{\partial}{\partial \mathbf{f}_2} \mathbf{f}_2 &= \tilde{\omega}^{21} \langle \mathbf{f}_2 \rangle \mathbf{f}_1 + \tilde{\omega}^{23} \langle \mathbf{f}_2 \rangle \mathbf{f}_3 = s_2 \mathbf{f}_1 + k_2 \mathbf{f}_3 \\
\frac{\partial}{\partial \mathbf{f}_2} \mathbf{f}_3 &= \tilde{\omega}^{31} \langle \mathbf{f}_2 \rangle \mathbf{f}_1 + \tilde{\omega}^{32} \langle \mathbf{f}_2 \rangle \mathbf{f}_2 = -k_2 \mathbf{f}_2
\end{aligned}$$

The first partial derivations introduce the two principal curvatures. Therefore, on the second partial derivations, we have to take into account the variation of the principal curvatures.

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{f}_1} k_2 &= (k_1 - k_2) \tilde{\omega}^{12} \langle \mathbf{f}_1 \rangle = (k_2 - k_1) s_2 \\
\frac{\partial}{\partial \mathbf{f}_2} k_1 &= (k_1 - k_2) \tilde{\omega}^{12} \langle \mathbf{f}_2 \rangle = (k_1 - k_2) s_1
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{f}_1} k_1 &= t_1 \\
\frac{\partial}{\partial \mathbf{f}_2} k_2 &= t_2
\end{aligned}$$

Two other parameters t_1, t_2 are needed to describe the variation of the principal curvatures. The relation between the local coordinates and the frame field:

$$\begin{aligned}
\frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{e}_1^2} &= 0 & \frac{\partial^2 \mathbf{f}_2}{\partial \mathbf{e}_1^2} &= -s_1 \\
\frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{e}_1 \partial \mathbf{e}_2} &= s_1 & \frac{\partial^2 \mathbf{f}_2}{\partial \mathbf{e}_1 \partial \mathbf{e}_2} &= 0 \\
\frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{e}_2 \partial \mathbf{e}_1} &= 0 & \frac{\partial^2 \mathbf{f}_2}{\partial \mathbf{e}_2 \partial \mathbf{e}_1} &= s_2 \\
\frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{e}_2^2} &= -s_2 & \frac{\partial^2 \mathbf{f}_2}{\partial \mathbf{e}_2^2} &= 0
\end{aligned}$$

And thus the second order partial derivatives are

$$\begin{aligned}
\frac{\partial^2}{\partial \mathbf{e}_1^2} \mathbf{N} &= -t_1 \mathbf{f}_1 - s_1 (k_1 - k_2) \mathbf{f}_2 - k_1^2 \mathbf{f}_3, \\
\frac{\partial^2}{\partial \mathbf{e}_1 \partial \mathbf{e}_2} \mathbf{N} &= (k_1 - k_2) (-s_1 \mathbf{f}_1 + s_2 \mathbf{f}_2), \\
\frac{\partial^2}{\partial \mathbf{e}_2 \partial \mathbf{e}_1} \mathbf{N} &= (k_1 - k_2) (-s_1 \mathbf{f}_1 + s_2 \mathbf{f}_2), \\
\frac{\partial^2}{\partial \mathbf{e}_2^2} \mathbf{N} &= -t_2 \mathbf{f}_2 - s_2 (k_2 - k_1) \mathbf{f}_1 - k_2^2 \mathbf{f}_3.
\end{aligned}$$

2.2. The Partial Derivatives of the Image Irradiance. We assume that the reflectance can be modeled by Lambert's reflectance function, and that furthermore the image is obtained by an orthographic projection. First we relate the image coordinate system and the local surface patch coordinate system. Then we evaluate the partial derivatives of the image irradiance.

2.2.1. Viewer's Coordinates and Surface Patch Local Coordinates. The viewer's coordinates and the surface patch local coordinates are related by the linear transformation matrix:

$$\mathbf{M} = \begin{pmatrix} e_1^{(x)} & e_2^{(x)} & e_3^{(x)} \\ e_1^{(y)} & e_2^{(y)} & e_3^{(y)} \\ e_1^{(z)} & e_2^{(z)} & e_3^{(z)} \end{pmatrix}$$

where $e_i^{(x)}, e_i^{(y)}, e_i^{(z)}$ are respectively the x, y, z components of the i^{th} vector of the surface patch local coordinate system.

This transformation is only a rotation matrix. Hence the inverse of the matrix $\mathbf{M}$ is simply its transpose. We can relate the representation of a vector — e.g. the illuminant direction — in the viewer's coordinate frame with the representation of that vector in the surface patch local coordinate frame as follows:

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \mathbf{M} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} \text{ or inversely } \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \mathbf{M}^t \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}.$$

$$\begin{aligned} L_x &= e_1^{(x)} L_1 + e_2^{(x)} L_2 + e_3^{(x)} L_3 \\ L_y &= e_1^{(y)} L_1 + e_2^{(y)} L_2 + e_3^{(y)} L_3 \\ L_z &= e_1^{(z)} L_1 + e_2^{(z)} L_2 + e_3^{(z)} L_3 \end{aligned}$$

Two rotations relate the viewer's coordinates to the surface patch local coordinates. The first rotation takes care of the surface orientation and the second, of the principal directions. In matrix form, it can be written as

$$\mathbf{M} = \mathbf{M}_n \mathbf{M}_c$$

where

$$\mathbf{M}_n = \begin{pmatrix} \cos^2 \tau \cos \sigma + \sin^2 \tau & \sin \tau \cos \tau \cos \sigma - \sin \tau \cos \tau & \cos \tau \sin \sigma \\ \sin \tau \cos \tau \cos \sigma - \sin \tau \cos \tau & \sin^2 \tau \cos \sigma + \cos^2 \tau & \sin \tau \sin \sigma \\ -\cos \tau \sin \sigma & -\sin \tau \sin \sigma & \cos \sigma \end{pmatrix}$$

and

$$\mathbf{M}_c = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The axis of the first rotation is the axis perpendicular to the plane defined by the $\mathbf{z}$ axis and the surface normal $\mathbf{N}$. This rotation matrix can be rewritten in terms of the surface normal components to provide a better intuition.

$$\mathbf{M}_n = \begin{pmatrix} \frac{N_x^2 N_z + N_y^2}{N_x^2 + N_y^2} & \frac{N_x N_y N_z - N_x N_y}{N_x^2 + N_y^2} & N_x \\ \frac{N_x N_y N_z - N_x N_y}{N_x^2 + N_y^2} & \frac{N_y^2 N_x + N_x^2}{N_x^2 + N_y^2} & N_y \\ -N_x & -N_y & N_z \end{pmatrix}$$

The axis of the second rotation corresponds to the surface normal $\mathbf{N}$.

2.2.2. *The Image Irradiance.* Since we assume that the reflectance properties of a surface can be modeled by Lambertian law,

$$I = \rho\lambda\mathbf{L} \cdot \mathbf{N} ,$$

the first order directional derivatives of the intensity function can be computed as follows.

$$I_x = \rho\lambda\mathbf{L} \cdot \frac{\partial}{\partial \mathbf{x}} \mathbf{N}$$

$$I_y = \rho\lambda\mathbf{L} \cdot \frac{\partial}{\partial \mathbf{y}} \mathbf{N}$$

If we consider the surface normal as a function of the local surface patch coordinates and apply the chain rule, we obtain

$$I_x = \rho\lambda\mathbf{L} \cdot \left(\frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} \frac{\partial}{\partial \mathbf{e}_1} \mathbf{N} + \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} \frac{\partial}{\partial \mathbf{e}_2} \mathbf{N} \right)$$

$$I_y = \rho\lambda\mathbf{L} \cdot \left(\frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} \frac{\partial}{\partial \mathbf{e}_1} \mathbf{N} + \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} \frac{\partial}{\partial \mathbf{e}_2} \mathbf{N} \right)$$

And since

$$\frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} = e_1^x \quad \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} = e_2^x$$

$$\frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} = e_1^y \quad \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} = e_2^y$$

thus obtain

$$(B.1) \quad I_x = \rho\lambda\mathbf{L} \cdot (e_1^x (-k_1 \mathbf{f}_1) + e_2^x (-k_2 \mathbf{f}_2))$$

$$(B.2) \quad I_y = \rho\lambda\mathbf{L} \cdot (e_1^y (-k_1 \mathbf{f}_1) + e_2^y (-k_2 \mathbf{f}_2))$$

The second order directional derivatives of the intensity function can be computed as follows.

$$I_{xx} = \rho\lambda\mathbf{L} \cdot \frac{\partial^2}{\partial \mathbf{x}^2} \mathbf{N}$$

$$I_{xy} = \rho\lambda\mathbf{L} \cdot \frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{y}} \mathbf{N}$$

$$I_{yy} = \rho\lambda\mathbf{L} \cdot \frac{\partial^2}{\partial \mathbf{y}^2} \mathbf{N}$$

Just as the first partial derivatives, we consider the surface normal as a function of the local surface patch coordinates and apply the chain rule, and we obtain

$$\begin{aligned}
I_{xx} &= \rho\lambda L \cdot \left(\frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} \frac{\partial^2}{\partial \mathbf{e}_1^2} N + 2 \frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} \frac{\partial^2}{\partial \mathbf{e}_1 \partial \mathbf{e}_2} N + \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} \frac{\partial^2}{\partial \mathbf{e}_2^2} N \right) \\
I_{xy} &= \rho\lambda L \cdot \left(\frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} \frac{\partial^2}{\partial \mathbf{e}_1^2} N + \left(\frac{\partial \mathbf{e}_1}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} + \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} \right) \frac{\partial^2}{\partial \mathbf{e}_1 \partial \mathbf{e}_2} N + \frac{\partial \mathbf{e}_2}{\partial \mathbf{x}} \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} \frac{\partial^2}{\partial \mathbf{e}_2^2} N \right) \\
I_{yy} &= \rho\lambda L \cdot \left(\frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} \frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} \frac{\partial^2}{\partial \mathbf{e}_1^2} N + 2 \frac{\partial \mathbf{e}_1}{\partial \mathbf{y}} \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} \frac{\partial^2}{\partial \mathbf{e}_1 \partial \mathbf{e}_2} N + \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} \frac{\partial \mathbf{e}_2}{\partial \mathbf{y}} \frac{\partial^2}{\partial \mathbf{e}_2^2} N \right)
\end{aligned}$$

These equations combined with equations B.1 and B.2 provide the link between local surface patch properties and the variation of the image irradiance.

2.2.3. The Shading Flow Field. For recovering shape from shading, we propose to use some characteristics of the shading flow field. We consider the most important to be the orientation of the flow:

$$\tan \theta = \frac{I_y}{I_x}.$$

which is independent from both the albedo and the illumination:

$$\tan \theta = \frac{\mathbf{L} \cdot (e_1^y (-k_1 \mathbf{f}_1) + e_2^y (-k_2 \mathbf{f}_2))}{\mathbf{L} \cdot (e_1^x (-k_1 \mathbf{f}_1) + e_2^x (-k_2 \mathbf{f}_2))}.$$

We note that it is not the case for magnitude of image irradiance gradient:

$$|\nabla I| = \sqrt{I_x^2 + I_y^2},$$

or explicitly

$$|\nabla I| = \rho\lambda \sqrt{\left(\mathbf{L} \cdot (e_1^x (k_1 \mathbf{f}_1) + e_2^x (k_2 \mathbf{f}_2)) \right)^2 + \left(\mathbf{L} \cdot (e_1^y (k_1 \mathbf{f}_1) + e_2^y (k_2 \mathbf{f}_2)) \right)^2}.$$

Curvatures are other useful characteristics of the shading flow field. Just like the orientation, they are independent from both the albedo and the illumination. The curvature of the isoluminance line is given by:

$$\kappa_i = \frac{2I_x I_y I_{xy} - I_x^2 I_{yy} - I_y^2 I_{xx}}{(I_x^2 + I_y^2)^{\frac{3}{2}}}.$$

And the curvature of the gradient line is given by:

$$\kappa_g = \frac{I_y^2 I_{xx} - I_x^2 I_{yy}}{(I_x^2 + I_y^2)^{\frac{3}{2}}}.$$

APPENDIX C

Surface Shape and Occluding Contours

In this appendix, we derive the relationship between the differential properties of the occluding contours and the surface shape.

1. Differential properties of a planar curve

Let's consider a planar curve with an arbitrary parametrization:

$$\mathbf{x} = (x(t), y(t))$$

The tangent vector of this curve is given by

$$\mathbf{t} = \frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{dt} \frac{dt}{ds}$$

where s is the arclength parametrization, thus

$$\frac{ds}{dt} = \left| \frac{d\mathbf{x}}{dt} \right| = \left(\left(\frac{dx(t)}{dt} \right)^2 + \left(\frac{dy(t)}{dt} \right)^2 \right)^{\frac{1}{2}}$$

or reciprocally

$$\frac{dt}{ds} = \left| \frac{d\mathbf{x}}{dt} \right|^{-1} = \left(\left(\frac{dx(t)}{dt} \right)^2 + \left(\frac{dy(t)}{dt} \right)^2 \right)^{-\frac{1}{2}}$$

And the curvature vector of this curve is:

$$\mathbf{k} = \kappa \mathbf{n} = \frac{d\mathbf{t}}{ds}$$

with

$$\frac{d\mathbf{t}}{ds} = \frac{d\mathbf{t}}{dt} \frac{dt}{ds} = \frac{d^2\mathbf{x}}{dt^2} \left(\frac{dt}{ds} \right)^2 + \frac{d\mathbf{x}}{dt} \frac{dt}{ds} \frac{d}{dt} \left(\frac{dt}{ds} \right) = \frac{d^2\mathbf{x}}{dt^2} \left(\frac{dt}{ds} \right)^2 + \mathbf{t} \frac{d}{dt} \left(\frac{dt}{ds} \right)$$

and

$$\frac{d}{dt} \left(\frac{ds}{dt} \right) = - \left(\frac{dx(t)}{dt} \frac{d^2x(t)}{dt^2} + \frac{dy(t)}{dt} \frac{d^2y(t)}{dt^2} \right) \left(\left(\frac{dx(t)}{dt} \right)^2 + \left(\frac{dy(t)}{dt} \right)^2 \right)^{-\frac{3}{2}}$$

2. Self Occluding Edges

For our surface model, the paraboloid, the surface normal vector is given in term of a local surface patch coordinate system:

$$\mathbf{N} = \left(\frac{\kappa_1 u}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{\kappa_2 v}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{1}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}} \right)$$

Since at the occluding boundaries, the surface becomes perpendicular to the line of sight, we express the normal vector in image coordinates.

$$\mathbf{N}^{(xyz)} = \mathbf{M} \mathbf{N}^{(uvw)}$$

And thus occluding boundaries are characterized by the z component of the normal vector being null.

$$\mathbf{N}^{(z)} = \frac{e_1^z \kappa_1 u + e_2^z \kappa_2 v + e_3^z}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}} = 0$$

which provide a first constraint to define the curve describing the occluding contour.

$$e_1^z \kappa_1 u + e_2^z \kappa_2 v + e_3^z = 0$$

This leads to the following arbitrary parametrizations: if either $\kappa_1 = 0$ or $e_1^z = 0$ and either $\kappa_2 = 0$ or $e_2^z = 0$ then no occlusion is visible; else if $\kappa_1 = 0$ or $e_1^z = 0$, then

$$\begin{aligned} u(t) &= t \\ v(t) &= -\frac{e_3^z}{e_2^z \kappa_2} \end{aligned}$$

or if $\kappa_2 = 0$ or $e_2^z = 0$, then

$$\begin{aligned} u(t) &= -\frac{e_3^z}{e_1^z \kappa_1} \\ v(t) &= t \end{aligned}$$

otherwise

$$\begin{aligned} u(t) &= \frac{1}{\kappa_1} \left(-e_2^z t - \frac{e_3^z}{2 e_2^z} \right) \\ v(t) &= \frac{1}{\kappa_2} \left(e_1^z t - \frac{e_3^z}{2 e_1^z} \right) \end{aligned}$$

Another constraint on the curve describing the occluding contour is that it lies on the paraboloid, thus

$$w(t) = -\frac{1}{2} \kappa_1 u^2(t) - \frac{1}{2} \kappa_2 v^2(t)$$

We can project this curve on the image plane and obtain the following planar curve.

$$x(t) = e_1^x u(t) + e_2^x v(t) + e_3^x w(t)$$

$$y(t) = e_1^y u(t) + e_2^y v(t) + e_3^y w(t)$$

$$\frac{dx(t)}{dt} = e_1^x \frac{du(t)}{dt} + e_2^x \frac{dv(t)}{dt} + e_3^x \left(-\kappa_1 u(t) \frac{du(t)}{dt} + -\kappa_2 v(t) \frac{dv(t)}{dt} \right)$$

$$\frac{dy(t)}{dt} = e_1^y \frac{du(t)}{dt} + e_2^y \frac{dv(t)}{dt} + e_3^y \left(-\kappa_1 u(t) \frac{du(t)}{dt} + -\kappa_2 v(t) \frac{dv(t)}{dt} \right)$$

$$\frac{du(t)}{dt} = \frac{-e_2^z}{\kappa_1}$$

$$\frac{dv(t)}{dt} = \frac{e_1^z}{\kappa_2}$$

$$\frac{dx(t)}{dt} = -\frac{e_1^x e_2^z}{\kappa_1} + \frac{e_2^x e_1^z}{\kappa_2} - e_3^x \left(\left(e_2^z t + \frac{e_3^z}{2 e_2^z} \right) \frac{e_2^z}{\kappa_1} + \left(e_1^z t - \frac{e_3^z}{2 e_1^z} \right) \frac{e_1^z}{\kappa_2} \right)$$

$$\frac{dy(t)}{dt} = -\frac{e_1^y e_2^z}{\kappa_1} + \frac{e_2^y e_1^z}{\kappa_2} - e_3^y \left(\left(e_2^z t + \frac{e_3^z}{2 e_2^z} \right) \frac{e_2^z}{\kappa_1} + \left(e_1^z t - \frac{e_3^z}{2 e_1^z} \right) \frac{e_1^z}{\kappa_2} \right)$$

$$\frac{dx(t)}{dt} = -\frac{e_1^x e_2^z}{\kappa_1} - \frac{e_3^x e_3^z}{2\kappa_1} + \frac{e_2^x e_1^z}{\kappa_2} + \frac{e_3^x e_3^z}{2\kappa_2} - \left(\frac{e_2^z e_2^z}{\kappa_1} + \frac{e_1^z e_1^z}{\kappa_2} \right) e_3^x t$$

$$\frac{dy(t)}{dt} = -\frac{e_1^y e_2^z}{\kappa_1} - \frac{e_3^y e_3^z}{2\kappa_1} + \frac{e_2^y e_1^z}{\kappa_2} + \frac{e_3^y e_3^z}{2\kappa_2} - \left(\frac{e_2^z e_2^z}{\kappa_1} + \frac{e_1^z e_1^z}{\kappa_2} \right) e_3^y t$$

$$\frac{d^2 x(t)}{dt^2} = - \left(\frac{e_2^z e_2^z}{\kappa_1} + \frac{e_1^z e_1^z}{\kappa_2} \right) e_3^x$$

$$\frac{d^2 y(t)}{dt^2} = - \left(\frac{e_2^z e_2^z}{\kappa_1} + \frac{e_1^z e_1^z}{\kappa_2} \right) e_3^y$$

The orientation of the occluding edge is thus:

$$\tan(\Theta) = \frac{\frac{dy(t)}{dt}}{\frac{dx(t)}{dt}}$$

$$\tan(\Theta) = \frac{-e_1^y e_2^z \kappa_2 - \frac{1}{2} e_3^y e_3^z \kappa_2 + e_2^y e_1^z \kappa_1 + \frac{1}{2} e_3^y e_3^z \kappa_1 - (e_2^z e_2^z \kappa_2 + e_1^z e_1^z \kappa_1) e_3^y t}{-e_1^x e_2^z \kappa_2 - \frac{1}{2} e_3^x e_3^z \kappa_2 + e_2^x e_1^z \kappa_1 + \frac{1}{2} e_3^x e_3^z \kappa_1 - (e_2^z e_2^z \kappa_2 + e_1^z e_1^z \kappa_1) e_3^x t}$$

APPENDIX D

Shading from Shape

In order to test our shape from shading algorithm, it is necessary to have prior knowledge of the illumination and shape of every surface patch viewed in the image. For this purpose, we have implemented an algorithm that solves the “shading from shape” problem. This problem is often referred to as the forward problem whereas the shape from shading problem is referred to as the inverse problem. Depicting a realistic image of a scene is a computationally intensive task. Massively parallel computers have entered computer graphics research, raising the issue of how to embed the radiosity problem [45, 119] into massively parallel SIMD computing architectures. We have shown that the Langer’s parallel radiosity algorithm [81] could be implemented efficiently on a SIMD machine with a two-dimensional array of processors and first-neighbour connectivity. We extended the algorithm to deal with scenes in which a participating media such as fog or smoke is present. This generality is a strong improvement over previous radiosity algorithms where each additional scene parameter (eg. specularities [63], fog [116]) required a significant amount of extra programming effort.

1. Previous Work

Surface inter-reflections have been traditionally modelled by expressing the radiance of each surface point as a weighted sum of the radiances of all the other surface points. Specifically, the surfaces in a given scene are represented as a set of n inter-reflecting planar facets, and the radiosity equation is approximated as a set n linear equations. The equations have n^2 coefficients, or *form factors* [119] which must be computed.

The main bottleneck in computing the form factors is to determine which surface facets are visible from which. This visibility problem is solved by computing a perspective view of the scene sequentially for each surface facet. Once a visibility function $V(\mathbf{x}, \mathbf{x}^*)$ is computed, a linear system of equations that relate the surface radiosities is constructed. For

Lambertian surfaces and sources, the continuous radiosity equation is

$$R(\mathbf{x}) = R_{emit}(\mathbf{x}) + \frac{\rho(\mathbf{x})}{\pi} \int R(\mathbf{x}^*) \frac{\cos \theta_1 \cos \theta_2}{|\mathbf{x} - \mathbf{x}^*|^2} V(\mathbf{x}, \mathbf{x}^*) d\mathbf{x}^*,$$

where $R(\mathbf{x})$ is the radiance of $\mathbf{x}$, and θ_1 and θ_2 are the angles between the line segment $\overline{\mathbf{x}\mathbf{x}^*}$ and the unit surface normals $\mathbf{N}(\mathbf{x})$ and $\mathbf{N}(\mathbf{x}^*)$.

When either directional reflectance or participating media are present, the above equations require significant revision. Directional reflectance may be included by partitioning the hemisphere of directions above each surface patch into a finite set of solid angles[63], effectively multiplying the number of form factors by m^2 (where m is the size of the partition). Participating media may be included by explicitly representing the surface-volume and volume-volume exchange of light. This requires an enormous increase in the number of form factors[116].

Langer [81] reformulated the radiosity equation in terms of the light rays in a scene, rather than in terms of pairs of surface facets. In an important sense, the set of light rays in a scene is smaller than the set of pairs of surface facets: while each light ray corresponds to a pair of surface facets (namely the points of origin and termination), many pairs of surface facets may not be joined by a straight line through free space. Indeed, the visibility function may be thought of as a labelling of a line segment as "a light ray" or "not a light ray".

Let $R(\mathbf{x}, \mathbf{L})$ denote the *radiance of a light ray* arriving at $\mathbf{x}$ in direction $\mathbf{L}$. Radiance has units Watts per square meter per steradian, and is defined as follows. Consider viewing a scene through a narrow straight tube (see Fig. D.55). Suppose that the near end of the tube is positioned at a point $\mathbf{x}$ in free space, and that the central ray through the tube has direction $\mathbf{L}$. Let d^2E denote the light energy passing through the tube and let $d\mathbf{L}$ denote the solid angle subtended by the far end of the tube. Let $d\mathbf{a}$ denote the cross sectional area of the tube. Then, the radiance of the light ray passing through $\mathbf{x}$ in direction $\mathbf{L}$ is

$$R(\mathbf{x}, \mathbf{L}) \equiv \frac{d^2E}{d\mathbf{a} d\mathbf{L}}.$$

Let $\mathcal{H}(\mathbf{x})$ denote the hemisphere of directions pointing out of the surface at $\mathbf{x}$. For Lambertian surfaces and sources, the radiance of ray leaving a surface does not depend on direction, and thus may be written $R(\mathbf{x})$. The radiosity equation may now be expressed in terms of the radiance of light rays as

$$R(\mathbf{x}) = R_{emit}(\mathbf{x}) + \frac{\rho(\mathbf{x})}{\pi} \int_{\mathcal{H}(\mathbf{x})} R(\mathbf{x}, -\mathbf{L}) \mathbf{N}(\mathbf{x}) \cdot \mathbf{L} d\mathbf{L}.$$

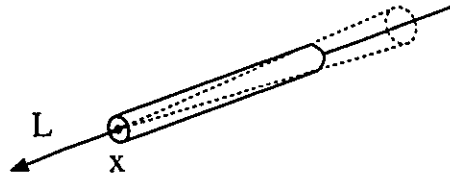


FIGURE D.55. A light ray passes through the point $\mathbf{x}$ in direction $\mathbf{L}$. The radiance of the ray is defined as the energy that passes through the tube per unit area cross section per steradian subtended by the far end of the tube.

This system of equations may be solved numerically by using a Jacobi iteration,

$$(D.1) \quad R^{k+1}(\mathbf{x}) := R_{emit}(\mathbf{x}) + \frac{\rho(\mathbf{x})}{\pi} \sum_{\mathbf{L} \in \mathcal{H}(\mathbf{x})} R^k(\mathbf{x}, -\mathbf{L}) \mathbf{N}(\mathbf{x}) \cdot \mathbf{L} \Delta \mathbf{L} .$$

Surface radiance at $\mathbf{x}$ is just the radiance of a light ray whose point of *origin* is $\mathbf{x}$. To apply (D.1), however, requires that light rays are indexed by their points of *termination*. Langer [81] designed an algorithm for reparameterizing the set of light rays in a scene, from their points of origin to their points of termination. One of the key advantages of this algorithm is that it may be embedded into a massively parallel SIMD architecture, thus implicitly solving the visibility problem but in a fashion that is naturally parallel.

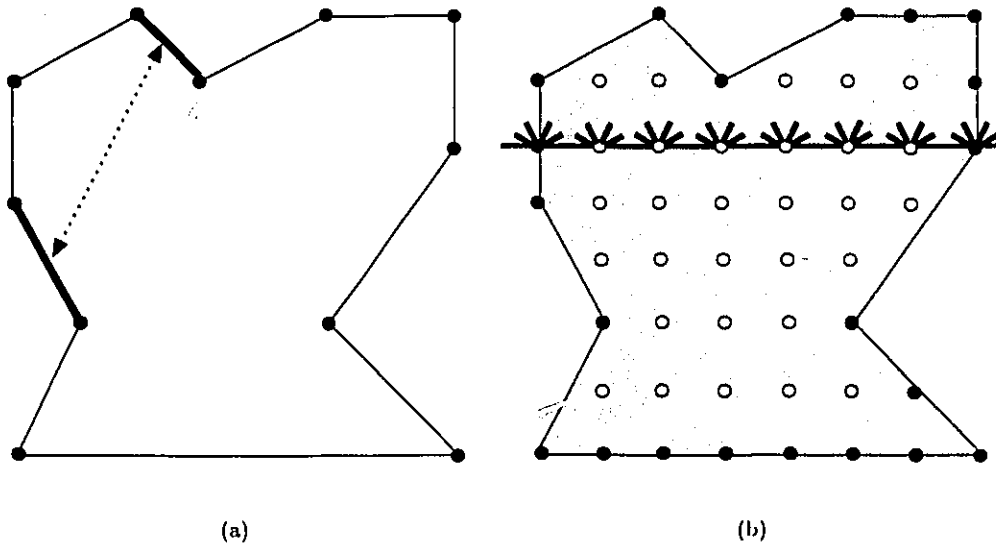


FIGURE D.56. A comparison of the traditional and new formulations of the radiosity equation. (a) In the traditional formulation, scene geometry is represented by a set of surface facets and by a visibility function on pairs of facets. (b) In the new formulation, scene geometry is represented by a cubic lattice of nodes, and by a local coordinate system on the light ray manifold.

2. Light Ray Manifold

In the absence of scattering by fog or dust, light travels along straight lines in free space. These oriented lines, or light rays, may be either finite, semi-infinite, or infinite in both directions. Let $\mathcal{M}$ denote the set of light rays in a given scene. Langer [81] showed that this set is a four dimensional manifold, which he called the *light ray manifold* (see [27] for the general definition of a manifold.). He notes that an important property of the light ray manifold is that *in the absence of a participating medium, radiance is constant along a ray* [42]. *That is, radiance is a positive real valued function on the light ray manifold,*

$$R : \mathcal{M} \rightarrow \mathbb{R}^+.$$

This property is independent of the surface reflectance. Moreover, it allows us to identify the radiance at the point of origin of a ray with the radiance at the point of termination.

2.1. Local Coordinates: Continuous Case. Local coordinate systems on a light ray manifold $\mathcal{M}$ may be defined in a variety of ways. For example, consider a local patch $\mathbf{x}(u, v)$ of a surface in the scene. For each point $\mathbf{x}$ on the patch, the set of light rays that originate from $\mathbf{x}$ may be parameterized by a hemisphere of unit vectors $\mathcal{H}(\mathbf{x})$. These vectors specify the directions of the outgoing rays. Since both a hemisphere and a surface patch are two dimensional sets, it follows that the set of light rays that originate from the surface patch $\mathbf{x}(u, v)$ is four dimensional. Notice that a similar local coordinate system is defined by parameterizing the light rays that terminate at the surface patch. These local coordinate systems of $\mathcal{M}$ are widely used in traditional radiosity algorithms. Surfaces in a scene are defined by a set of planar facets, and for each facet, a hemicube of incident [23] or reflected [22] rays is defined.

Langer [81] introduced an alternative local coordinate system for $\mathcal{M}$ is introduced. (See Fig. D.57.) Consider a two-dimensional plane passing through the scene, for example, the plane $z \equiv z_0$. A given point on this plane is either in free space, inside an object, or on the surface of an object. Each light ray that intersects this plane is specified by four coordinates: two determine where the ray intersects the plane, and two determine the direction of the ray. For example, the light ray that passes through a point (x_0, y_0, z_0) in direction (ϕ, θ) may be parameterized by (x_0, y_0, ϕ, θ) .

Observe that, when the plane is swept through space, a coordinate evolution on $\mathcal{M}$ is obtained. Suppose that a light ray passes through both planes $z \equiv z_0$ and $z \equiv z_1$. Let (x_0, y_0, ϕ, θ) and (x_1, y_1, ϕ, θ) be the two resulting parameterizations of this ray, and further observe that $(x_1, y_1) \rightarrow (x_0, y_0)$ as $z_1 \rightarrow z_0$. Thus, *the coordinate evolution is continuous.*

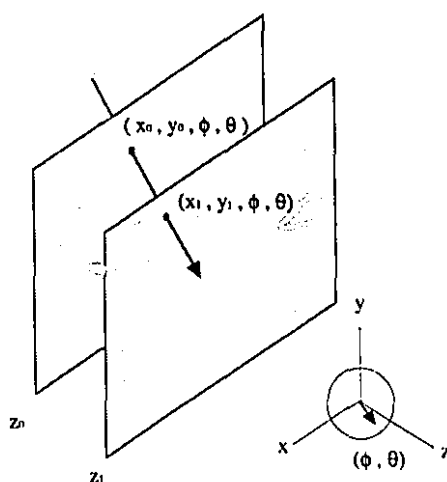


FIGURE D.57. The parameterization of a light ray depends on the coordinate plane. A given ray is parameterized by (x_0, y_0, ϕ, θ) in the plane $z = z_0$. In the plane $z = z_1$, the same ray is parametrized by (x_1, y_1, ϕ, θ) . Observe that $(x_1, y_1, \phi, \theta) \rightarrow (x_0, y_0, \phi, \theta)$ as $z_1 \rightarrow z_0$.

This continuity property is crucial because it allows us to compute a coordinate evolution on $\mathcal{M}$ using a massively parallel computing architecture (in particular, using a two dimensional array of processing elements that have only local connectivity).

2.2. Local Coordinates: Discrete Case. As in discrete ordinate methods [32, 33, 57, 85], Langer [81] considered that the nodes in space and the light rays that pass through space are discretized as follows. Space is represented by a $N \times N \times N$ cubic lattice. Nodes in this lattice are of four types: FREE nodes, SURFACE nodes, SOURCE nodes, and SOLID nodes. Light rays travel through FREE nodes. Light is absorbed and reflected at SURFACE nodes. Light is emitted at SOURCE nodes. Light does not reach SOLID nodes.

Light rays are restricted to travel in a finite set of directions through the space lattice. For each FREE, SURFACE, or SOURCE node $\mathbf{x}$, Langer defined the finite set of light rays that may pass through $\mathbf{x}$ by the nodes on a small cube that is centered at $\mathbf{x}$. These directions are defined by the line segments joining $\mathbf{x}$ to points on the six faces of the cube. (See Figure D.58.) The cube is analogous to the hemicube of Cohen and Greenberg[23]; however, there are two important differences. First, the half width of our cube is much smaller ($M = 5$ vs. $M = 50$) than that of Cohen and Greenberg. Second, Langer defined a light ray cube at each FREE node, as well as at each SURFACE and SOURCE node, whereas Cohen and Greenberg only defined it at each SURFACE node. It is by making these additional hemicubes explicit, that the visibility problem becomes implicit and the algorithm becomes naturally parallel.

Local coordinates on the sampled light ray manifold may be represented as follows. (See Fig. D.59.) For a given face F of the light ray cube (i.e. there are 6 faces), consider the i th plane in the space lattice that is parallel to F . The local coordinate system $\mathcal{C}_i^F \subset \mathcal{M}$ is the set of light rays that are specified by F and that pass through plane i ,

$$\mathcal{C}_i^F \equiv \{ (x, y, \mathbf{L}) : (x, y, i) \text{ is a FREE or SURFACE node, and } \mathbf{L} \in F \}.$$

Neighboring local coordinate systems may overlap. For example, consider a FREE node $\mathbf{x} = (x, y, i)$, and a ray passing through this node in direction $\mathbf{L} = (L_x, L_y, M)$. This ray could be parameterized in at least three different ways:

$$(x, y, \mathbf{L}) \in \mathcal{C}_i^F, (x + L_x, y + L_y, \mathbf{L}) \in \mathcal{C}_{i+M}^F, \text{ or } (x - L_x, y - L_y, \mathbf{L}) \in \mathcal{C}_{i-M}^F.$$

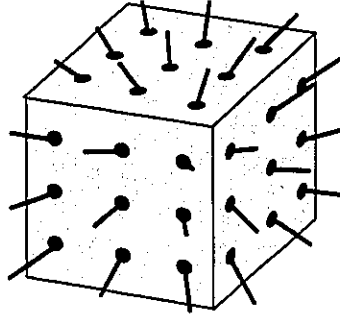


FIGURE D.58. Light rays are restricted to travel in a finite set of directions through the space lattice. These directions are defined by the nodes on a small cube, specifically, the directions of the line segments joining the center of the cube to points on the six faces of the cube.

3. Coordinate Evolution

To solve Equation (D.1) requires that rays be parameterized by their point of termination. For that purpose, we use Langer's algorithm. This algorithm consists of two nested procedures. The first is a local coordinate transformation (depicted in Fig. D.60) from one coordinate system $\mathcal{C}_i^F$ to its neighbor $\mathcal{C}_{i+M}^F$. The second is a sequence of local transformations, or coordinate evolution, that reparameterizes all rays in the light ray manifold.

Consider the case of two neighboring coordinate systems $\mathcal{C}_i^F$ and $\mathcal{C}_{i+M}^F$. Let $R_0(x, y, \mathbf{L})$ and $R_1(x, y, \mathbf{L})$ be radiance functions that are defined on $\mathcal{C}_i^F$ and $\mathcal{C}_{i+M}^F$, respectively. Given an estimate $R^k(\mathbf{x})$ of the surface radiance and given $R_0(x, y, \mathbf{L})$, the radiance $R_1(x, y, \mathbf{L})$ is

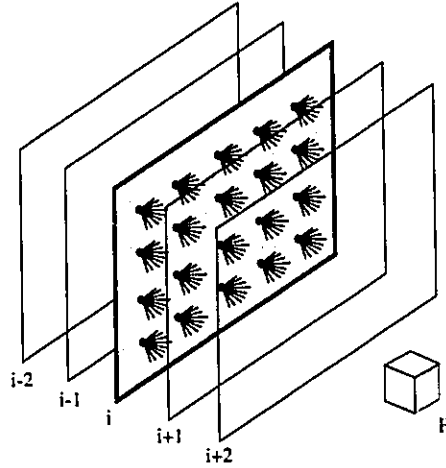


FIGURE D.59. A local coordinate system on the light ray manifold is defined follows. For a given cube face F , consider the i th plane in the space lattice that is parallel to face F . The local coordinate system C_i^F is the set of light rays that are in the directions specified by F and that pass through plane i .

computed by transforming from C_i^F to C_{i+M}^F . In Section 5, this algorithm will be modified slightly to include the presence of a participating medium.

```

Local Transformation(  $F, i, R_0, R_1, R^k$  ) {
  for all  $(x, y)$  in parallel
     $\mathbf{x} := (x, y, i)$ ;
    for all  $\mathbf{L} := (L_x, L_y, M)$ 
      case  $\{\mathbf{x} - \mathbf{L}\}$ 
      SOLID :  $R_1(x, y, \mathbf{L}) := 0$ ;
      FREE :  $R_1(x, y, \mathbf{L}) := R_0(x - L_x, y - L_y, \mathbf{L})$ ;
      SURFACE : if  $(\mathbf{L} \cdot \mathbf{N}(\mathbf{x} - \mathbf{L}) > 0)$ 
         $R_1(x, y, \mathbf{L}) := R(\mathbf{x} - \mathbf{L}, \mathbf{L})$ ;
        else  $R_1(x, y, \mathbf{L}) := 0$ ;
      SOURCE :  $R_1(x, y, \mathbf{L}) := R_{emit}(x - L_x, y - L_y, \mathbf{L})$ ;
}

```

A coordinate evolution is defined by a sequence of local coordinate transformations along the three orthogonal axes of the cubic lattice, first in the positive direction and then in the negative direction of each axis. In each of these six “sweeps”, the rays defined by a particular face of the cube of light ray directions are considered.

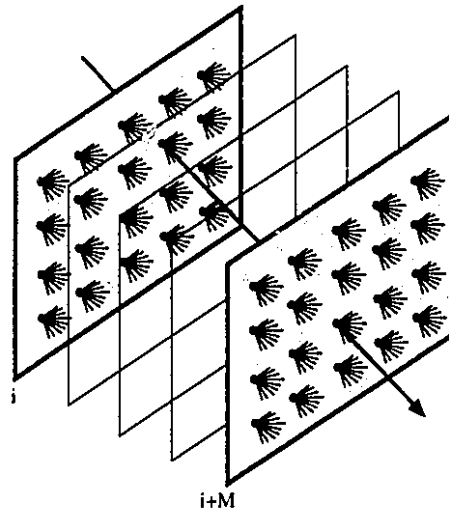


FIGURE D.60. In the procedure **Local Transformation**, the parameterization of the light rays is transformed from one local coordinate system to its neighbor a distance M away. Because of the continuity property discussed in Fig. D.57, a light ray passes through nearby nodes in two neighboring coordinate planes.

```

Coordinate Evolution( $R^k$ ) {
  for all cube faces  $F$ ,
    initialize  $R_0$  on  $C_j^F$ ;
    for ( $i = 0$ ;  $i < N$ ;  $i := i + 1$ ) {
      Local Transformation( $F, i, R_0, R_1, R^k$ );
       $R_0 := R_1$ ;
    }
  }

```

Once the rays have been reparameterized in terms of their point of termination, the surface radiance is updated using (D.1). In our implementation, the surface radiances are updated within the procedure **Local Transformation** whenever a ray terminates at a SURFACE node.

Two observations should be made. First, the coordinate evolution may interpreted as the propagation of radiance along rays, so that the algorithm is just a simulation of the physics of radiosity. The number of iterations of the algorithm corresponds to the number of surface inter-reflections that are considered. Second, the algorithm does not depend on the reflectance properties of the surface. If the reflectance was non-Lambertian, then the surface radiance function would be non-isotropic. This case will be discussed in detail in a future paper.

4. Parallel Implementation

A coordinate evolution is a sequence of local coordinate transformations on the light ray manifold. This evolution is implemented on a massively parallel computer, a MasPar-1, which is a SIMD machine having 2048 4-bit processing elements. The PEs form a two dimensional array of size $N_x \times N_y = 64 \times 32$. Each PE is directly connected to its eight neighbors, and the boundary is connected in a toroidal topology.

The width of a scene to be rendered is typically larger than the width of the PE array. To accommodate this difficulty, a coordinate plane is mapped to the PE array by wrapping the plane around the array as many times as needed. The result is that, at each PE, many space columns are represented.

For example, the images shown in Fig. D.61 correspond to a scene of width $N = 128$. This scene is wrapped onto the PE array, 2 times for the x direction and 4 times for y direction. Each PE thus represents 8 space columns, each of which is 128 nodes deep. Hence, 1028 space nodes are represented at each PE.

4.1. Memory Costs. For each space node, a variety of state variables are represented. For SURFACE nodes, these state variables include the space type, the albedo, surface radiance, surface normal and emittance. For FREE nodes, the state variables include the scattering and absorption coefficients, and the luminescence (see Sec. 5).

If $N_x = N_y = N$, then N space nodes are represented at each PE. Within the procedure **Local Transformation**, $O(M^2)$ rays are required at each PE to represent the radiance of the rays. Additional temporary memory of size $O(N)$ is also needed to rotate the (x, y, z) axes within the procedure **Coordinate Evolution**. Thus, the local memory cost to each PE is $O(N + M^2)$.

When many space columns are represented at each PE, the memory costs are multiplied by the number of columns at each PE, which is $N^2/(N_x N_y)$. Each PE on the MASPAR-1 has a local memory of 16 Kbytes, which imposes a maximum scene width of $N = 128$. More recent models of the MasPar machine have over 16 K processors (forming an array of 128×128) and 64 Kbytes of memory per processor. On such a machine, it would be possible to generate 512×512 images. To include directional effects such as specularities, refraction, and non-isotropic scattering, the memory costs are multiplied by a factor $O(M^2)$.

4.2. Time Costs. The basic time cost of the algorithm per iteration is $O(NM^2)$. **Local Transformation** is $O(M^2)$, and **Coordinate Evolution** multiplies this cost by $O(N)$. If the scene is wider than the PE array, then the time cost of **Local Transformation**

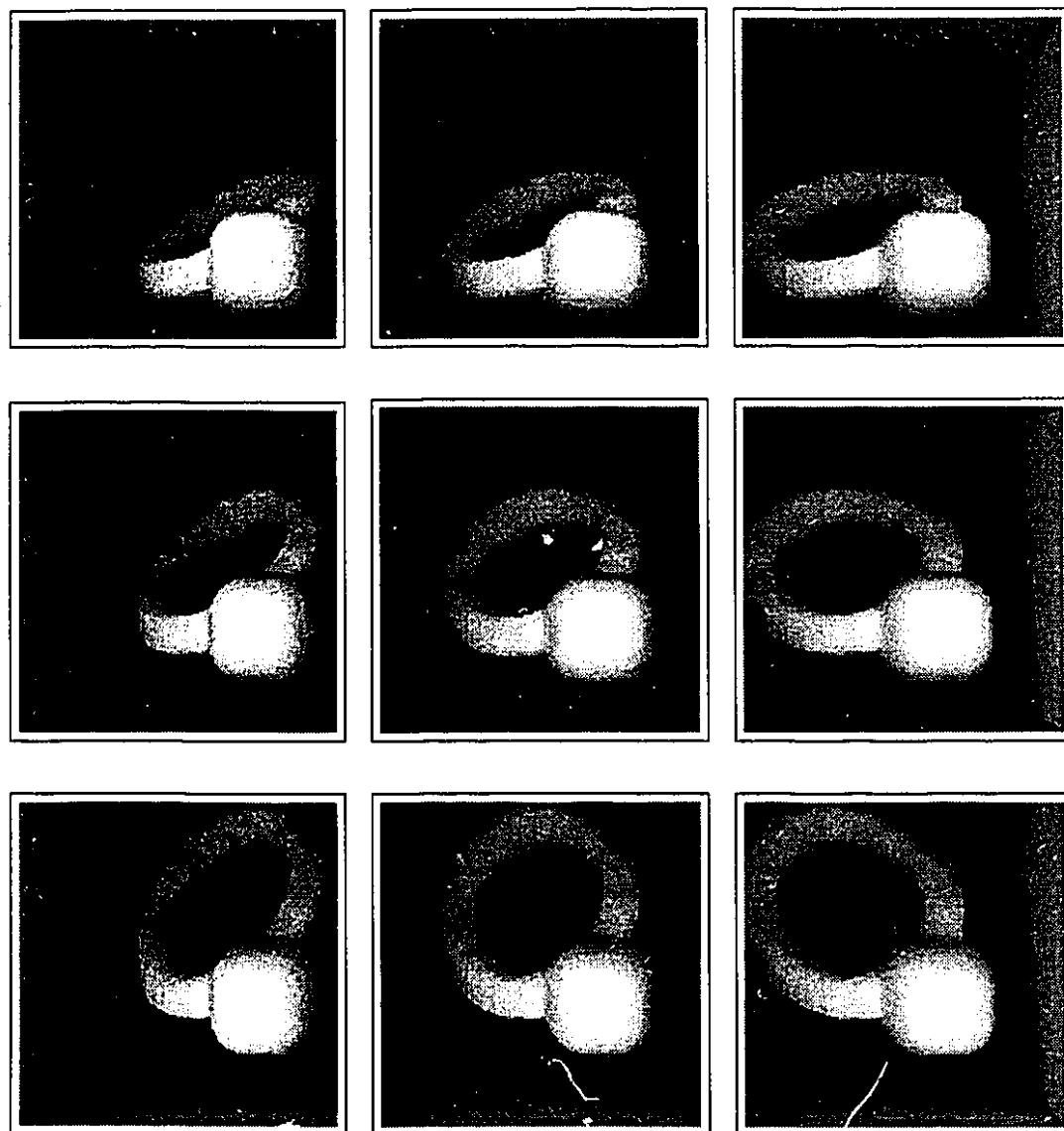


FIGURE D.61. A ring floating in a box and viewed through the top. No participating media is present. The albedo of all surfaces is 0.9. The light source has two components, ambient and directed, both of which enter the box through the open roof. The ambient component is a uniform hemispheric sky. The directed component is oblique, from the left. Twelve iterations are used. The scene is viewed from nine different directions, centered around the vertical.

is multiplied by $N^2/(N_x N_y)$. For our machine, the procedure **Local Transformation** requires roughly 0.6 seconds when $N = 128$, resulting in a total computation time of roughly 8 minutes per iteration of the algorithm.

We also note that time is required to rotate the (x, y, z) axes within the procedure **Coordinate Evolution**, between "sweeps". On a serial machine, the time cost of this rotation is $O(N^3)$. We have developed an efficient parallel algorithm to solve the problem. The algorithm has a time complexity of $O(NN_x)$ when $N = N_x$, and $O(N^3/N_x)$ when $N > N_x$. The operations required for the axis rotation are quite simple, and in practice the time cost is insignificant. In particular, when $N = 128$, each rotation takes 0.3 seconds.

5. Scattering and Absorption in Free Space

Until now, only the case of a vacuum has been discussed. In many interesting situations, however, a participating medium such as smoke or fog will be present [11], and light will be scattered and absorbed. In this section, we show how to generalize algorithm of Sec. 3 to account for isotropic scattering and absorption effects. An important result is that the space and time costs of the algorithm are unaffected by this generalization. An example is shown in Figure D.62.

Isotropic scattering and absorption may be modelled by assigning a scattering coefficient $\zeta(\mathbf{x})$ and an absorption coefficient $\gamma(\mathbf{x})$ to each point in free space. Consider a ray passing through $\mathbf{x}$ in direction $\mathbf{L}$. The radiance of rays passing near $\mathbf{x}$ satisfy

$$\frac{dR(\mathbf{x} + \epsilon\mathbf{L}, \mathbf{L})}{d\epsilon} = (-\zeta(\mathbf{x}) - \gamma(\mathbf{x})) R(\mathbf{x}, \mathbf{L}) + \frac{\zeta(\mathbf{x})}{4\pi} \int_{S^2} R(\mathbf{x}, \mathbf{L}^*) d\mathbf{L}^* .$$

In particular, radiance is conserved along a ray over any region of free space in which $\zeta(\mathbf{x})$ and $\gamma(\mathbf{x})$ vanish.

When a participating medium is present, radiance is not a unique function defined on the manifold $\mathcal{M}$. Rather, different radiance functions are defined on different local coordinate systems of $\mathcal{M}$. The above equation relates the radiance functions of neighboring local coordinate systems.

Scattering is incorporated into our algorithm by modelling it as isotropic luminescence. Specifically, energy is temporarily accumulated in a state variable $\Lambda^k(\mathbf{x})$ whenever a light ray passes through $\mathbf{x}$. The energy that is accumulated during iteration k is scattered isotropically in iteration $k + 1$. The **Local Transformation** algorithm must be modified slightly. A box is placed around the modified lines of pseudo-code.

Observe that the time and space costs of the algorithm are unchanged from the case of the vacuum. This is in sharp contrast to the method proposed in [116], where the computational costs increase dramatically in the presence of a participating medium.

```

Local Transformation(  $F, i, R_0, R_1, R^k$  ) {
  for all  $(x, y)$  in parallel {
     $\mathbf{x} := (x, y, i)$ ;
    for all  $\mathbf{L} := (L_x, L_y, M)$ ,
      case  $\{\mathbf{x} - \mathbf{L}\}$ 
        SOLID :  $R_1(x, y, \mathbf{L}) := 0$ ;
        FREE :  $R_1(x, y, \mathbf{L}) := (1 - \gamma(\mathbf{x}) - \zeta(\mathbf{x})) R_0(x - L_x, y - L_y, \mathbf{L}) + \frac{\zeta(\mathbf{x})}{4\pi} L^k(\mathbf{x}) ;$ 
            $\Lambda^{k+1}(\mathbf{x}) += \gamma(\mathbf{x}) R_0(x - L_x, y - L_y, \mathbf{L}) \Delta \mathbf{L} ;$ 
        SURFACE : if  $(\mathbf{L} \cdot \mathbf{N}(\mathbf{x} - \mathbf{L}) > 0)$ 
            $R_1(x, y, \mathbf{L}) := R(\mathbf{x} - \mathbf{L}, \mathbf{L})$ ;
           else  $R_1(x, y, \mathbf{L}) := 0$ ;
        SOURCE :  $R_1(x, y, \mathbf{L}) := R_{emit}(x - L_x, y - L_y, \mathbf{L})$ ;
  }
}

```

6. Summary

The radiosity equation has been reformulated in terms of the radiances of the set of light rays in a scene, rather than in terms of the the radiosity exchanged by each pair of surface facets. Form factors and visibility functions are not computed in the new formulation. The traditional computational bottleneck of computing the visibility between surface facets is replaced by the problem of parameterizing the set of light rays by their points of termination, rather than by their points of origin. This parameterization is solved by a coordinate evolution algorithm.

This new formulation has two main advantages over the traditional one. The first is its generality. The presence of a participating medium requires only a slight modification of the algorithm and no change to either the space or time complexity. (The memory requirements for the case of non-isotropic surface reflectance and scattering will be presented in a future paper.)

The second advantage is that the algorithm is massively parallel; it may be implemented on a SIMD architecture having local connections between nodes. Machines that are of appropriate size are now entering the marketplace. The existence of rendering algorithms

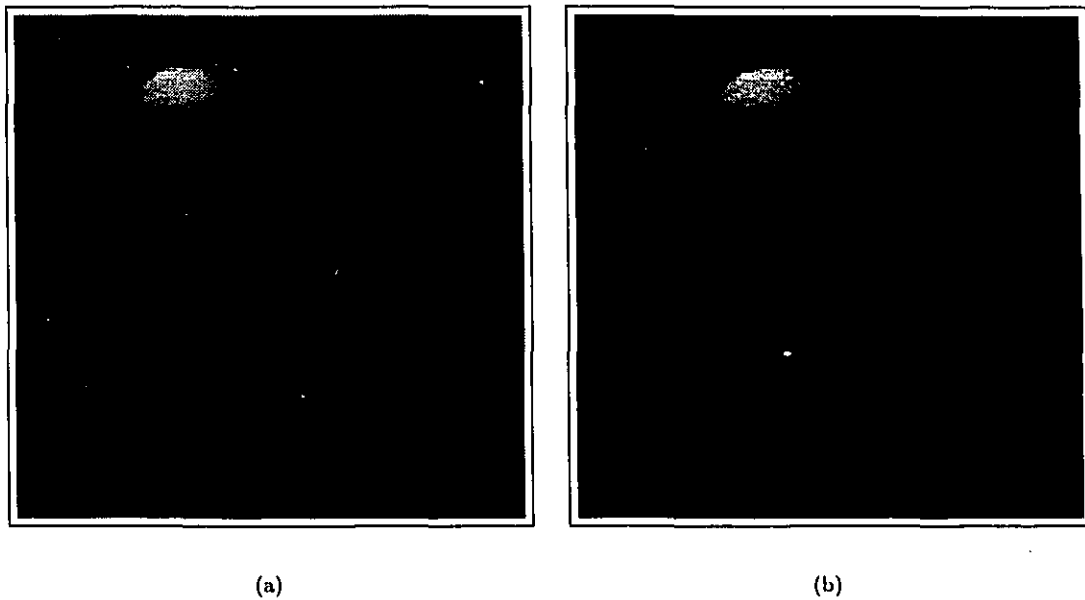


FIGURE D.62. Three balls floating in a box that is filled with fog. The scene is illuminated by a directed light source that is from above but slightly oblique. The shadows cast by the balls are visible both in space (as a tube) and on the surface. Observe that the deepest ball is the darkest since light is absorbed as it passes through the fog.

that make use of these machines is a further technological incentive for the development of massively parallel SIMD graphics hardware.

APPENDIX E

Labelling

The surface orientation is represented by unit normal facing the viewer. We attempt to sample this space as uniformly as possible to get forty cells. The cells are listed in Table E.5.

(0.338, 0.939, 0.059)	(0.644, 0.762, 0.059)	(-0.982, 0.176, 0.059)
(-0.982, -0.176, 0.059)	(0.644, -0.762, 0.059)	(0.338, -0.939, 0.059)
(-0.212, 0.939, 0.270)	(-0.707, 0.653, 0.270)	(0.919, 0.285, 0.270)
(0.919, -0.285, 0.270)	(-0.707, -0.653, 0.270)	(-0.212, -0.939, 0.270)
(0.127, 0.934, 0.333)	(0.745, 0.577, 0.333)	(-0.872, 0.356, 0.333)
(-0.872, -0.356, 0.333)	(0.745, -0.577, 0.333)	(0.127, -0.934, 0.333)
(-0.440, 0.762, 0.473)	(0.880, 0.000, 0.473)	(-0.440, -0.762, 0.473)
(0.233, 0.756, 0.610)	(0.538, 0.580, 0.610)	(-0.772, 0.176, 0.610)
(-0.772, -0.176, 0.610)	(0.538, -0.580, 0.610)	(0.233, -0.756, 0.610)
(-0.333, 0.577, 0.745)	(0.666, 0.000, 0.745)	(-0.333, -0.577, 0.745)
(0.005, 0.580, 0.814)	(0.500, 0.294, 0.814)	(-0.505, 0.285, 0.814)
(-0.505, -0.285, 0.814)	(0.500, -0.294, 0.814)	(0.005, -0.580, 0.814)
(0.170, 0.294, 0.940)	(-0.340, 0.000, 0.940)	(0.170, -0.294, 0.940)
(0.000, 0.000, 1.000)		

TABLE E.5. This table gives the sampling of the surface normal that we use in our implementation.

The shape of the surface is locally represented by the shape vector and the curviness. The curviness is very coarsely sampled as we show in Table E.6. We attempt to sample the shape vector space as uniformly as possible to get thirty-two cells. The cells are listed in Table E.7.

The direction of the illumination is represented by a unit vector. In our implementation, we use the same thirty-two cells defined by the sampling of the shape vector space.

0.2 0.08 0.05

TABLE E.6. This table gives the sampling of the surface curveness that we use in our implementation.

(0.000, 0.000, -1.000)	(0.607, 0.000, -0.794)	(-0.303, 0.525, -0.794)
(-0.303, -0.525, -0.794)	(0.333, 0.577, -0.745)	(-0.666, 0.000, -0.745)
(0.333, -0.577, -0.745)	(-0.127, 0.934, -0.333)	(0.872, 0.356, -0.333)
(-0.745, 0.577, -0.333)	(-0.745, -0.577, -0.333)	(0.872, -0.356, -0.333)
(-0.127, -0.934, -0.333)	(0.491, 0.850, -0.187)	(-0.982, 0.000, -0.187)
(0.491, -0.850, -0.187)	(-0.491, 0.850, 0.187)	(0.982, 0.000, 0.187)
(-0.491, -0.850, 0.187)	(0.127, 0.934, 0.333)	(0.745, 0.577, 0.333)
(-0.872, 0.356, 0.333)	(-0.872, -0.356, 0.333)	(0.745, -0.577, 0.333)
(0.127, -0.934, 0.333)	(-0.333, -0.577, 0.745)	(-0.333, 0.577, 0.745)
(0.666, 0.000, 0.745)	(-0.607, 0.000, 0.794)	(0.303, 0.525, 0.794)
(0.303, -0.525, 0.794)	(0.000, 0.000, 1.000)	

TABLE E.7. This table gives the sampling of the surface shape vector that we use in our implementation.

REFERENCES

- [1] J. Allman and S. Zucker. Cytochrome oxidase and functional coding in primate striate cortex: A hypothesis. *Cold Spring Harbor Symposia on Quantitative Biology*, 55:979-982, 1990.
- [2] Y. Aloimonos. Combining sources of information in vision: I. computing shape from shading and motion. *Perception*, pages 738-741, 1987.
- [3] H. Barlow. The efficiency of detecting changes of density in random dot patterns. *Vision Research*, 18:637, 1979.
- [4] P. Beckmann and A. Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Pergamon Press, 1963.
- [5] J. M. Beusmans, D. D. Hoffmann, and B. M. Bennett. Description of solid shape and its inference from occluding contours. *Journal of the Optical Society of America*, 4(7):1155-1167, July 1987.
- [6] M. Bichsel and A. P. Pentland. A simple algorithm for shape from shading. In *Proceedings, Conference on Computer Vision and Pattern Recognition* [60], pages 459-465.
- [7] I. Biederman. Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing*, 32(1):29-73, October 1985.
- [8] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94:115-147, 1987.
- [9] A. Blake. Improbable views. *Nature*, 368:498-499, April 1994.
- [10] A. Blake, A. Zisserman, and G. Knowles. Surface descriptions from stereo and shading. *Image and Vision Computing*, 3(4):183-191, 1985.
- [11] J. F. Blinn. Light reflection functions for simulation of clouds and dusty surfaces. *Computer Graphics*, 16:21-29, 1983.
- [12] P. Breton, L. A. Iverson, M. S. Langer, and S. W. Zucker. Shading flows and scene bundles: A new approach to shape from shading. In *ECCV92* [29], pages 135-150.
- [13] M. J. Brooks. Surface-normals from closed path. In *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, volume 1, pages 98-101, Tokyo, Japan, August 20-23 1979. International Joint Conference on Artificial Intelligence.
- [14] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pages 932-936, Los Angeles, CA, August 1985.
- [15] A. R. Bruss. *The image irradiance equation: its solution and application*. PhD thesis, Massachusetts Institute of Technology, June 1981.
- [16] A. R. Bruss. The eikonal equation: Some results applicable to computer vision. *Journal of Mathematical Physics*, 23(5):890-896, 1982.

- [17] P. Bui-Thong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311-317, June 1975.
- [18] H. H. Bülthoff and H. A. Mallot. Interaction of depth modules: Stereo and shading. *Journal of the Optical Society of America*, 5:1749-1758, 1988.
- [19] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679-698, 1986.
- [20] P. Cavanagh. Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity and shape. *Computer Vision, Graphics, and Image Processing*, 37:171-195, 1987.
- [21] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick. *Analysis, Manifold and Physics*. North-Holland, Amsterdam, revised edition, 1982.
- [22] M. F. Cohen, S. E. Chen, J. R. Wallace, and D. P. Greenberg. A progressive refinement approach to fast radiosity image generation. *Computer Graphics*, 22(4):75-84, August 1988.
- [23] M. F. Cohen and D. P. Greenberg. The hemicube: A radiosity solution for complex environments. *Computer Graphics*, 19(3):31-40, July 1985.
- [24] Computer Society of the IEEE. *Proceedings, 3RD International Conference on Computer Vision*, Osaka, Japan, Dec. 1990. IEEE Computer Society Press.
- [25] R. Deriche. Using canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1(2):167-187, 1987.
- [26] J. V. Digellen. A photometric investigation of the slopes and heights of the ranges of hills in the maria of the moon. *Bulletin of the Astronomical Institute of the Netherlands*, 11(423), July 1951.
- [27] M. P. DoCarmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.
- [28] R. O. Duda and P. E. Hart. *Pattern classification and scene analysis*. Wiley-Interscience, New York, 1973. p. 271.
- [29] *Computer Vision - ECCV 92*, Santa Margherita Ligure, Italy, 19-22 May 1992. Springer-Verlag.
- [30] F. P. Ferrie and J. Lagarde. Curvature consistency improves local shading analysis. *Computer Vision, Graphics, and Image Processing*, 55(1):95-105, January 1992.
- [31] F. P. Ferrie and M. D. Levine. Where and why local shading analysis works. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(2):198-206, 1989.
- [32] W. Fiveland. Discrete ordinates solutions of the radiative transport equation for rectangular enclosures. *Journal of Heat Transfer*, 106:699-706, November 1984.
- [33] W. Fiveland. Discrete ordinate methods for radiative heat transfer in isotropically and anisotropically scattering media. *Journal of Heat Transfer*, 109:809-812, August 1987.
- [34] H. Flanders. *Differential Forms*. Academic Press, 1963.
- [35] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley, Reading, Mass., second edition, 1990.
- [36] D. Forsyth and A. Zisserman. Reflections on shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:671-679, July 1991.
- [37] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):439-451, 1988.
- [38] W. Freeman and E. Adelson. The design and use of steerable filters for image analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991.

- [39] W. T. Freeman. The generic viewpoint assumption in a framework for visual perception. *Nature*, 368:542-545, April 1994.
- [40] J. Gårding. *Shape from Surface Markings*. PhD thesis, Royal Institute of Technology, University of Stockholm, Stockholm, Sweden, May 1991.
- [41] P. R. Garabedia. *Partial Differential Equations*. Chelsea, New York, second edition, 1986.
- [42] A. Gershun. The light field. *J. Math Phys.*, 18:51-151, 1939.
- [43] J. Gibson. *The Perception of the Visual World*. Houghton Mifflin, Boston, 1950.
- [44] L. Glass. Moire effect from random dots. *Nature*, 243:578-580, 1969.
- [45] C. M. Goral, K. E. Torrence, D. P. Greenberg, and B. Battaile. Modelling the interaction of light between diffuse surfaces. *Computer Graphics*, 18(3):213-222, July 1984.
- [46] Y. Hel Or and S. W. Zucker. Texture fields and texture flows: Sensitivity to differences. *Spatial Vision*, 4(2/3):131-139, 1989.
- [47] E. C. Hildreth. *Implementation of a theory of edge detection*. PhD thesis, Massachusetts Institute of Technology, April 1980.
- [48] D. Hoffman and W. A. Richards. Parts of recognition. *Cognition*, 18:65-96, 1985.
- [49] J. J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Journal of the Proceedings of the National Academy Science*, 81:3088-3092, 1984.
- [50] B. K. Horn and M. Brooks. *Shape from Shading*. MIT Press, Cambridge, Massachusetts, 1989.
- [51] B. K. P. Horn. *Shape from Shading: a method for obtaining the shape of a smooth opaque object from one view*. PhD thesis, Massachusetts Institute of Technology, June 1970.
- [52] B. K. P. Horn. *The Psychology of Computer Vision*, chapter Obtaining Shape from Shading Information. McGraw-Hill, New York, 1975.
- [53] B. K. P. Horn. Understanding image intensities. *Artificial Intelligence*, 8:201-231, 1977.
- [54] B. K. P. Horn. *Robot Vision*. MIT Press, Cambridge, Mass., 1986.
- [55] B. K. P. Horn. Height and gradient from shading. *International Journal of Computer Vision*, 5(1):37-75, 1990.
- [56] B. K. P. Horn and M. J. Brooks. The variational approach to shape from shading. *Computer Vision, Graphics, and Image Processing*, 33:174-208, 1986.
- [57] J. Howell. Thermal radiation in participating media: The past, the present, and some possible futures. *Journal of Heat Transfer*, 110:1220-1229, November 1988.
- [58] R. A. Hummel and S. W. Zucker. On the foundation of relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5:267-287, 1983.
- [59] D. Husemoller. *Fibre Bundles*. Springer, New York, 1966.
- [60] IEEE-CS. *Proceedings, Conference on Computer Vision and Pattern Recognition*, Champaign, Illinois, June 15-18 1992. IEEE Computer Society Press.
- [61] K. Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(6):661-669, November 1981.
- [62] K. Ikeuchi and B. K. P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17:141-184, 1981.
- [63] D. S. Immel, M. F. Cohen, and D. P. Greenberg. A radiosity method for non-diffuse environments. *Computer Graphics*, 20(4):133-142, August 1986.

- [64] L. Iverson and S. W. Zucker. Orientation selection to optical flow: A computational perspective. In *Proceedings of the IEEE Workshop on Computer Vision*, pages 184–189, Miami, FL, 1987. IEEE.
- [65] L. Iverson and S. W. Zucker. Logical/linear operators for measuring orientation and curvature. Technical Report TR-CIM-90-06, McGill University, Montréal, 1990.
- [66] L. A. Iverson. *Toward discrete geometric models for early vision*. PhD thesis, McGill University, 1993.
- [67] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290:91–97, 1981.
- [68] D. C. Knill and D. Kersten. Learning a near-optimal estimator for surface shape from shading. *Computer Vision, Graphics, and Image Processing*, 50:75–100, 1990.
- [69] J. J. Koenderink. What does the occluding contour tell us about solid shape? *Perception*, 13:321–330, 1984.
- [70] J. J. Koenderink. *Solid Shape*. MIT Press, 1990.
- [71] J. J. Koenderink and A. J. van Doorn. The singularities of the visual mapping. *Biological Cybernetics*, 24:51–59, 1976.
- [72] J. J. Koenderink and A. J. van Doorn. Photometric invariants related to solid shape. *Optica acta*, 27(7):981–996, 1980.
- [73] J. J. Koenderink and A. J. van Doorn. The shape of smooth objects and the way contours end. *Perception*, 11:129–137, 1982.
- [74] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 63:291–297, 1987.
- [75] J. J. Koenderink, A. J. van Doorn, and A. M. Kappers. Surface perception in pictures. *Journal of Perception and Psychophysics*, 52(5):487–496, 1992.
- [76] P. A. Kolers and J. R. Pomerantz. Figural change in apparent motion. *Journal of Experimental Psychology*, 87:99–108, 1971.
- [77] P. A. Kolers and M. von Grünau. Shape and colour in apparent motion. *Vision Research*, 1:329–335, 1976.
- [78] J. Lambert. *Photometria sive de mensura de gratibus luminis, colorum et umbrae*. Eberhard Klett, Augsburg, 1760.
- [79] E. H. Land. Recent advances in retinex theory. *Vision Research*, 26(1):7–21, 1986.
- [80] E. H. Land and J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61:1–11, 1971.
- [81] M. S. Langer. *Shading Computations on the Radiation Manifold*. PhD thesis, McGill University, Montréal, June 1994.
- [82] M. S. Langer and S. W. Zucker. Qualitative shape from active shading. In *Proceedings, Conference on Computer Vision and Pattern Recognition* [60], pages 713–715.
- [83] M. S. Langer and S. W. Zucker. Diffuse shading, visibility fields, and the geometry of ambient light. In *Proceedings of the 4th International Conference on Computer Vision*, pages 138–147, Berlin, Germany, 1993.
- [84] M. S. Langer and S. W. Zucker. Shape from shading on a cloudy day. *Journal of the Optical Society of America*, February 1994.
- [85] K. Lathrop. Ray effects in discrete ordinates equations. *Nuclear Science and Engineering*, 32:357–369, 1968.
- [86] Y. G. Leclerc and A. F. Bobick. The direct computation of height from shading. In *Proceedings, Conference on Computer Vision and Pattern Recognition*, pages 552–558, Lahaina, Maui, Hawai, 1991. Computer Society of the IEEE, IEEE Computer Society Press.

- [87] C.-H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, 26(2):125-143, 1985.
- [88] K. M. Lee and C.-C. J. Kuo. Shape from shading with a linear triangular element surface model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):815-822, August 1993.
- [89] S. W. Lee and R. Bajcsy. Detection of specularities using color and multiple views. In *ECCV92* [29], pages 99-114.
- [90] M. Livingstone and D. Hubel. Anatomy and physiology of a color system in primate visual cortex. *Journal of Neuroscience*, 4:309-356, 1984.
- [91] E. Mach. Über die physiologische Wirkung räumlich vertheilter Lichtreize. *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, 54(2):393-408, 1866.
- [92] W. Maguire and J. Baizer. Luminance coding of briefly presented stimuli in area 17 of the rhesus monkey. *Journal of Neurophysiology*, 47:128, 1982.
- [93] D. Marr. *Vision*. Freeman, San Francisco, 1982.
- [94] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society*, 207:187-217, 1980.
- [95] D. A. Miller and S. W. Zucker. Efficient simplex-like methods for equilibria of non-symmetric analog networks. *Neural Computation*, 4:167-190, 1992.
- [96] M. Morrone and D. Burr. Feature detection in human vision: a phase dependent energy model. *Proceedings of the Royal Society*, 235:221-245, 1988.
- [97] J. Mulligan and D. MacLeod. Reciprocity between luminance and dot density in perception of brightness. *Vision Research*, 28:503, 1988.
- [98] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: physical and geometrical perspectives. *Patt. Recogn. Machine Intell.*, 13(5):185-212, 1991.
- [99] J. Oliensis. Shape from shading as a partially well-constrained problem. *Computer Vision, Graphics, and Image Processing*, 54:163-183, 1991.
- [100] J. Oliensis. Uniqueness in shape from shading. *IJCV*, 1991.
- [101] J. Oliensis and P. Dupuis. A global algorithm for shape from shading. In *Proceedings of the 4th International Conference on Computer Vision*, pages 692-701, Berlin, Germany, 1993.
- [102] B. O'Neill. *Elementary differential geometry*. Academic Press, 1966.
- [103] P. Parent and S. W. Zucker. Radial projection: An efficient update rule for relaxation labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):886-889, August 1989.
- [104] P. Parent and S. W. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):823-839, August 1989.
- [105] A. Pentland. Finding the illuminant direction. *Journal of the Optical Society of America*, 72:448-455, 1982.
- [106] A. Pentland. Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):170-187, 1984.
- [107] A. Pentland. Linear shape from shading. *International Journal of Computer Vision*, 4:153-162, 1990.
- [108] A. Pentland. Photometric motion. In *ICCV90* [24].
- [109] A. P. Pentland. Shading into texture. *Artificial Intelligence Journal*, 29:147-170, June 1986.

- [110] P. Perona. Steerable-scalable kernels for edge detection and junction analysis. In ECCV92 [29], pages 3-18.
- [111] P. Perona and J. Malik. Detecting and localizing edges composed of steps, peaks and roofs. In ICCV90 [24], pages 52-57.
- [112] F. Ratliff. *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. Holden-Day, San Francisco, CA, 1965. p. 285.
- [113] W. Richards, J. J. Koenderink, and D. Hoffman. Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America*, 4(7):1168-1175, July 1987.
- [114] T. Rindfleisch. Photometric method for lunar topography. *Photogrammetric Engineering*, 32(2):262, March 1966.
- [115] L. G. Roberts. *Optical and Electro-Optical Information Processing*, chapter Machine perception of three-dimensional solids, pages 159-197. MIT Press, Cambridge, Mass., 1965.
- [116] H. E. Rushmere and K. E. Torrance. The zonal method for calculating light intensities in presence of a participating medium. *Computer Graphics*, 21(4):293-302, July 1987.
- [117] P. Sander and S. W. Zucker. Inferring surface trace and differential structure from 3-d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:833-854, 1990.
- [118] B. V. H. Saxberg. *A modern differential geometric approach to shape from shading*. PhD thesis, Massachusetts Institute of Technology, 1989.
- [119] E. M. Sparrow and R. D. Cess. *Radiation Heat Transfer*. McGraw-Hill, New-York, 1978.
- [120] M. Spivak. *A Comprehensive Introduction to Differential Geometry*. Publish or Perish, Berkeley, 1979.
- [121] T. M. Strat. A numerical method for shape from shading from a single image. Master's thesis, Massachusetts Institute of Technology, 1979.
- [122] D. Terzopoulos. Multilevel computational processes for visual surface reconstruction. *Computer Vision, Graphics, and Image Processing*, 24:52-96, 1983.
- [123] D. Terzopoulos. *Image Understanding*, chapter Multigrid relaxation methods and the analysis of lightness, shading, and flow, pages 225-262. Ablex, Norwood, NJ, 1984.
- [124] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *JOSA*, 57(9), 1967.
- [125] H. von Helmholtz. *Treatise of Physiological Optics*. J.P.C. Southall, Dover, 1962. reprint.
- [126] D. Waltz. *The Psychology of Computer Vision*, chapter Understanding line drawings of scenes with shadows, pages 19-91. McGraw-Hill, New York, 1975.
- [127] T. Watanabe and P. Cavanagh. Texture and motion spreading, the aperture problem, and transparency. *Perception and Psychophysics*, 50(5):459-464, 1991.
- [128] L. B. Wolff. Shape understanding from lambertian photometric flow fields. In *Proceedings, Conference on Computer Vision and Pattern Recognition*, pages 46-52, Washington, June 1989. Computer Society of the IEEE, IEEE Computer Society Press.
- [129] R. J. Woodham. Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. *Proc. SPIE*, 155:136-143, 1978.
- [130] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139-144, 1980.
- [131] S. Zeki. Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *Journal of Physiology (London)*, 277:273-290, 1978.
- [132] S. Zeki. The representation of colours in the cerebral cortex. *Nature*, 284:412-418, 1980.

REFERENCES

- [133] S. Zeki. Segregation of pathways leading from area v2 to areas v4 and v5 of macaque monkey cortex. *Nature*, 315:322-324, 1985.
- [134] S. Zeki. *Colour vision and functional specialisation in the visual cortex*, volume VI of *Discussion in Neuroscience*. Elsevier, March 1990.
- [135] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Analysis of shape from shading techniques. In *Proceedings, Conference on Computer Vision and Pattern Recognition*, pages 377-384, Seattle, Washington, June 21-23 1994. IEEE-CS, IEEE Computer Society Press.
- [136] S. W. Zucker. Early orientation selection: tangent fields and the dimensionality of their support. *Computer Vision, Graphics, and Image Processing*, 32:74-103, 1985.
- [137] S. W. Zucker. *Encyclopedia of Artificial Intelligence*, volume 1, chapter Early Vision, pages 394-420. Wiley Interscience, New York, second edition, 1992.