Technical Report 1112

On Interpreting Stereo Disparity

Richard Patrick Wildes

MIT Artificial Intelligence Laboratory

This blank page was inserted to preserve pagination.

On interpreting stereo disparity

by
Richard Patrick Wildes

Submitted to the Department of Brain and Cognitive Sciences
on February 27, 1989 in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

Abstract

The problems under consideration center around the interpretation of binocular stereo disparity. In particular, the goal is to establish a set of mappings from stereo disparity to corresponding three-dimensional scene geometry. Stereo disparity is represented as a vector field derived from differential projection of a three-dimensional scene onto a pair of two-dimensional imaging surfaces. The resulting disparity field is analysed with the aid of mathematical tools from classical field theory. This analysis shows how disparity information can be interpreted in terms of three-dimensional scene properties, such as surface depth, discontinuities and orientation. These theoretical developments have been embodied in a set of computer algorithms for the recovery of scene geometry from input stereo disparity. The results of applying these algorithms to several disparity maps are presented. Finally, comparisons are made to the interpretation of stereo disparity by biological systems.

Thesis Supervisors:

Dr. W. Eric L. Grimson
Title: Associate Professor of Electrical Engineering and Computer Science

Dr. Whitman A. Richards
Title: Professor of Brain and Cognitive Sciences

Acknowledgements

I thank Eric Grimson and Whitman Richards for being my co-advisors during the course of my thesis research. I also extend thanks to William Thompson for serving as my outside examiner. During my stay at M.I.T. I have had a number of interesting professional interactions; in addition to my interactions with my advisors I note those that I have enjoyed with Heinrich Bülthoff and Ellen Hildreth. Fred Attneave, Jacob Beck and Kent Stevens deserve mention for sparking my interest in computational vision while I was an undergraduate at the University of Oregon. Jan Ellertsen gets thanks for watching over M.I.T. course IX graduate students. Emilio Bizzi gets thanks for providing me with financial coverage during the last months of this research. During the majority of my graduate study I have been funded by an NSF graduate fellowship. The first year of my study was funded by an NIH training grant.

I thank Bonnie for being my friend throughout.

Support for research at the M.I.T. Artificial Intelligence Laboratory is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124.

Contents

1 Introduction 7
- 1.1 Motivation . . . 7
- 1.2 Related work . . . 10
  - 1.2.1 Surface fitting . . . 10
  - 1.2.2 Differential imaging . . . 13
  - 1.2.3 Less related approaches to recovering discontinuities from disparity . . . 14
  - 1.2.4 Distinguishing features of the research presented in this thesis 17
- 1.3 Outline of chapters . . . 18

2 Planar surfaces 20
- 2.1 Analysis of disparity . . . 21
  - 2.1.1 Basic differential projection . . . 22
  - 2.1.2 Recovering view . . . 32
  - 2.1.3 Recovering geometric surface parameters . . . 35
  - 2.1.4 Recovering surface discontinuities . . . 36
  - 2.1.5 Recapitulation . . . 37
- 2.2 Stability analysis . . . 38
  - 2.2.1 Degeneracies . . . 38
  - 2.2.2 Error analysis . . . 41
  - 2.2.3 Operating in the face of perturbed data . . . 50
  - 2.2.4 Recapitulation . . . 56
- 2.3 Computer implementation . . . 57
  - 2.3.1 Description of algorithm and implementation . . . 57
  - 2.3.2 Experiments . . . 61
  - 2.3.3 Recapitulation . . . 66

3 Curved surfaces 74
- 3.1 Analysis of disparity . . . 74
  - 3.1.1 Recovering surface discontinuities . . . 75
- 3.2 Stability analysis . . . 80
  - 3.2.1 Degeneracies . . . 80
  - 3.2.2 Error analysis . . . 82
  - 3.2.3 Operating in the face of perturbed data . . . 84
  - 3.2.4 Recapitulation . . . 85

3.3 Computer implementation 85
3.3.1 Description of algorithm and implementation 86
3.3.2 Experiments 89
3.3.3 Recapitulation 90

4 Biological considerations **96**
4.1 Literature 96
4.2 Experiment 105
4.3 Recapitulation 112

5 Conclusions and suggestions for further research **121**
5.1 Summary and conclusions 121
5.2 Suggestions for further research 124

A Recovering view **128**
A.1 Full perspective method 128
A.2 Orthographic approximation 129
A.3 Recovering view with absolute scale 130
A.4 Considerations of stability 133

B The decomposition of discontinuous disparity fields **138**

C Surface curvature from disparity **142**

D Extension to motion based disparity **148**

Bibliography **151**

STOP! DON'T SWEAT IT. SIMPLY MOVE A FEW INCHES LEFT OR RIGHT TO GET A NEW VIEW POINT. Look...

Sometimes a Great Notion, Ken Kesey

This empty page was substituted for a blank page in the original document.

Chapter 1

Introduction

1.1 Motivation

Humans are quite adept at using visual information to infer the three-dimensionality of their surrounding world. Interestingly, this inference takes place in face of the fact that the inputs to the visual system (the retinal projections) are inherently two-dimensional. In order to understand this mapping from the two-dimensional retinal projections to inferences about a three-dimensional world most researchers have broken the task into a set of functional modules. For example, one finds studies of visual motion, binocular stereopsis and the various shape-from-x paradigms (e.g., shape from shading, texture, etc.). Following this model for vision research, this thesis shall be concerned with certain aspects of binocular stereopsis. In particular, this research is concerned with interpreting the disparity information that results from the correspondence of two retinal images.

Consider the paradigm within which stereopsis is currently studied: The basic

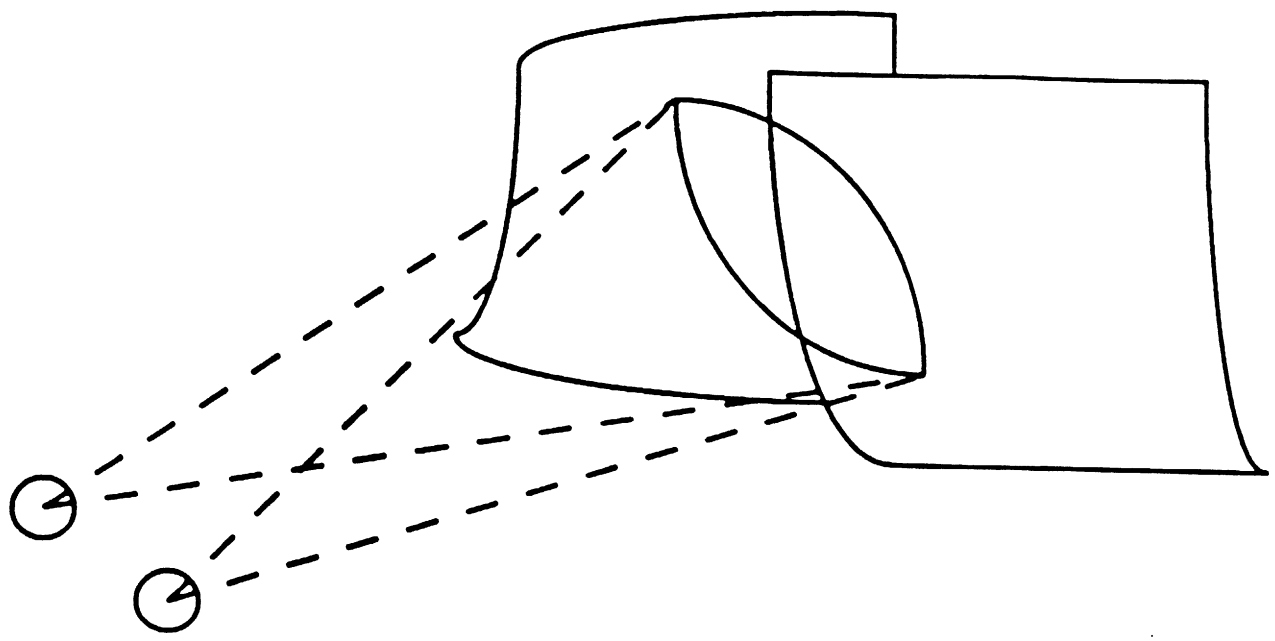


Figure 1.1: The basic situation for binocular stereopsis.

situation leading to stereopsis is illustrated in Figure 1.1. Here, an arrangement of surfaces in the three-dimensional world project differentially onto a pair of two-dimensional retinae. To understand stereopsis would be to understand how the corresponding inverse mapping can take place. That is, given a pair of two-dimensional projections of a three-dimensional world, how is it possible to exploit the geometry of the situation to recover useful properties of the geometry of that world. In our current state of understanding of stereopsis, it is convenient to break the problem into two relatively independent parts: (1) the correspondence problem and (2) the disparity interpretation problem. The correspondence problem consists of matching those elements in the two views that are projections of the same element in the three-dimensional world. Defining disparity as the difference in projective coordinates of

matched elements, it is seen that the output of the correspondence process can be considered a disparity map.[1] The disparity interpretation problem is to infer from the disparity map the three-dimensional properties of the viewed scene.

Typically, it has been thought that the difficult part of stereopsis was the solution of the correspondence problem. With the disparity map recovered, it has been assumed that (with knowledge of the relative orientation of the two views) the interpretation was a simple matter of triangulation. If the stereo data points are sparse the triangulated distance values can be interpolated. Such an approach is adequate (in theory) to specify the distance from the viewer to every point of the visible surfaces of the viewed scene. (See e.g., Barnard & Fischler [8] for a review of computational stereo vision studies within this paradigm.)

Now, consider the following questions: Is the distance to the visible surfaces in a scene the only (or even the most) desirable output of stereopsis? In particular, can the interpretation of the disparity map yield more sophisticated information than point by point distance? As alternatives, consider the possibility of directly interpreting stereo disparity in terms of surface orientations and surface discontinuities as well as distance. Intuition suggests that information concerning these latter properties would be more useful to subsequent visual processes (e.g., object recognition and passive navigation) than would simple point by point distance from the viewer. With these possibilities in mind, the goal of this thesis is to take a deeper look at the disparity interpretation problem.

[1]The relation between this definition of disparity and the classical angular disparity will be clarified in Section 2.1 of this thesis.

The particular approach taken shall be the computational approach (Marr [72]). Here, one initially attacks a problem as an abstract information processing problem. This abstraction allows one initially to focus attention on the formal nature of the problem under consideration and on constraints over its solution space. In the case of understanding stereo disparity this approach leads to considering the basic mathematical structure of the disparity map. From this study constraints shall be derived that allow one to make relatively sophisticated inferences about three-dimensional scene geometry from a corresponding input disparity map.

1.2 Related work

This section provides an overview of computational vision studies related to interpreting stereo disparity. When useful, this survey will also mention studies in interpreting motion based disparity. Much of this literature can be usefully broken into two categories: (i) surface fitting and (ii) studies of differential imaging. Also considered will be several miscellaneous studies related to the specific problem of recovering surface discontinuities from disparity. The section closes with a discussion that serves to distinguish the research presented in this thesis from other work in disparity interpretation.

1.2.1 Surface fitting

In its simplest form the idea behind the surface fitting approach is to interpolate (or approximate) the (possibly sparse) disparity values resulting from the correspondence

process with a smooth surface. Technically the disparity values should be first converted to depth values; in practice the disparity values are often employed directly. The result of such a surface fit is either a point by point depth map or the parameters of an algebraic surface patch. Such a representation does not necessarily make surface orientation explicit. Also, unless precautions are taken, the approach will allow surface discontinuities to be smoothed over during the interpolation process.

The surface fitting idea has been instantiated in at least two forms: (i) minimization of spline functionals and (ii) directly fitting polynomial based surface patches. The intuitive idea behind minimizing spline functionals is simple enough: Fit an elastic plate or membrane to the given data points and allow it to achieve equilibrium. The resulting representation is of point by point depth. The nontrivial technical details of applying this approach to disparity information has been the focus of much research (Blake [11], Boult [16], Grimson [37] and Terzopoulos [121]). The polynomial based approaches proceed by directly fitting a polynomial to the available depth data. For example, Eastman & Waxman [25] and Hoff & Ahuja [49] have used least squares to fit low-order (up through quadratic terms) Taylor series to depth data. Other polynomial bases could be used for this purpose; apparently this has not been investigated. However, in the area of interpolating and approximating grey-tone image intensity, Haralick's "Facet Model" has fostered much experimentation with fitting various polynomial forms to intensity values (Shapiro, et al. [112]). It is likely that some of the methods developed within the "Facet Model" could be carried over to distance interpolation.

Various attempts have been made to extend the surface fitting approach to deal

with such properties as surface discontinuity, orientation and curvature. Consider first, studies toward making surface orientation and curvature explicit. Within the spline based methods two paths have been followed. The first is to operate on the point by point distance representation and compute orientation and curvature through numerical differentiation (Brady et al. [17], Medioni & Nevatia [82]). The second path has been to couple the recovery of orientation and curvature to depth recovery via a cascade of differencing operations that are in effect during the spline minimization (Harris [47] and Terzopoulos [123]). Recovery of surface orientation and curvature from the polynomial based methods can be accomplished in some cases. For example, if a Taylor series is used the coefficients may have natural interpretations as surface gradients and curvatures (Eastman & Waxman [25], Hoff & Ahuja [49]).

Attention has also been given to allowing for discontinuous surfaces. These extensions can be grouped into two classes. The first class seeks to first interplolate and then look for likely areas where a discontinuity has been smoothed over. The second class attempts to recover a piecewise smooth surface while simultaneously allowing for discontinuity formation. Within the "interplolate and look" class several approaches have appeared: Grimson [38] proposed applying an edge detector (e.g., the Canny edge detector [18] or the Marr-Hildreth edge detector [74]) to the interpolated surface to discover discontinuities. This attack met with little empirical success (Grimson [41]). Terzopoulos [122] proposed the heuristic that points of high tension in the interpolated surface (marked by inflection points and-or steep gradient) should be considered for discontinuities. Finally, approaches have also been founded on the idea that loci of high residual in an approximating surface may indicate an underlying

discontinuity (Eastman & Waxman [25], Hoff & Ahuja [49], Grimson & Pavlidis [43] and Lee & Pavlidis [66]).

The joint recovery of surface and discontinuities has also received much attention. The idea is to allow discontinuities to form in a piecewise smooth surface at a penalty to a global energy functional. The resulting functional to be minimized is nonconvex. Several approaches to solving this problem have been proposed, both deterministic (Blake & Zisserman [12]) and probabalistic (Koch et al. [57] and Marroquin [75]) in nature. However, these methods are not guaranteed to find a global minimum (if one even exists).

1.2.2 Differential imaging

Studies in differential imaging seek to understand the relation between scene geometry and an infinitesimal change of viewpoint. Analysis proceeds by first specifying a locally analytic form for a surface and then developing the difference equation for the surface's projection onto image planes related via an infinitesimal change of coordinates. The study of the resulting vector field can explicitly relate surface geometry (e.g., distance, orientation and curvature with respect to the viewer) to the structure of projected disparity.

Differential imaging has been studied with reference to optical flow (e.g., Kanatani [54], Koenderink & van Doorn [58, 60], Longuet-Higgins & Pradzny [70], Pradzny [102], Subbaro [120], Waxman & Ullman [131] and Waxman & Wohn [132]) as well as stereo vision (e.g., Eastman & Waxman [25], Longuet-Higgins [68], Mayhew [78], Mayhew & Longuet-Higgins [80], Rogers [109], Stevens & Brookes [118] and Wein-

shall [133]). Most often, this work has limited consideration to recovering surface geometry only through first order. However, some consideration of surface curvature has occurred: Waxman and his associates ([131, 132, 25]) have developed algebraic relations between disparity and curvature. Also, Rogers [109] and Stevens & Brookes [118] have independently noted that second order differences of stereo disparity yield a surface curvature measure that is (supposedly) independent of distance. The question of surface discontinuity has received little attention in the differential imaging paradigms. An exception to this comment is Eastman & Waxman [25] where high residuals in the fit of difference equations to available data are taken as indication of surface discontinuity. Unfortunately, the difference equations relating surface geometry to disparity are highly nonlinear and the stability of their solution may be suspect (Barron et al. [9], Koenderink & van Doorn [61] and Wohn & Wu [135]).

Recently, it has been pointed out that similar work has been carried out for some time in the field of photogrammetry (Horn [50, 51] and *Manual of Photogrammetry* [71]). It is worth noting that the common thread to these analyses is that they are based in the application of tensor analysis to the classical field theory of mathematical physics (see Truesdell & Toupin [126]).

1.2.3 Less related approaches to recovering discontinuities from disparity

While the material in this thesis is not closely related to any of the approaches described below, it is nonetheless useful to provide an overview of alternative approaches

to the particular subproblem of recovering surface discontinuites. Four different types of studies are presented: (i) edge detection, (ii) correlational, (iii) general statistical and (iv) analysis of occlusions.

There have been some attempts to apply edge detection to disparity fields. Clocksin [20] showed the relation between surface discontinuities and discontinuities in a disparity field for the case of a purely translational differential view. This result was generalized to arbitrary infinitesimal differential view in Thompson et al [125]. In order to implement these ideas Thompson et al. [124] broke the disparity field into x and y scalar fields and convolved each component separately with a Laplacian operator. Discontinuities were found by combining the component wise Laplacians into a vector field and searching for the vector analog of a zero-crossing. Schunk [111] discusses interlacing an edge detection procedure with an iterative disparity field recovery algorithm. These techniques met with success in the analysis of optic flow. Stevens [118] suggests using a finite difference type mechanism to find discontinuities is a stereo disparity map. However, it appears that as yet there has been little attempt to study the feasibility of this idea either through a stability analysis or actual implementation.

One approach to establishing correspondence is to correlate regions (or features) between two images (e.g., Barnard & Fischler [8] and Moravec [89]). When such a process attempts to correlate across a discontinuity it is quite likely that the correlation will break down. This idea has been exploited to make surface discontinuites explicit in the analysis of both stereo (Smitly & Bajcsy [114] and optic flow (Anandan [2]). Such an approach is capable of making discontinuities explicit very early dur-

ing the stages of processing. Interestingly, Marr & Poggio [74] discuss how matching statistics should proceed if correspondence is being established properly by their algorithm; however, there apparently has been no attempt to turn their analysis around to recover likely regions of discontinuity. Along these lines, Nishihara [95] has provided an error analysis of a stereo matcher (related to the Marr-Poggio algorithm) that could likewise be used for discontinuity detection.

The idea that disparity field statistics should differ across a region corresponding to a surface discontinuity has been pursued by Spoerri & Ullman [116]. In this case the statistics of adjacent regions are compared after the correspondence has been established. These researchers report some success in applying these ideas to both stereo and optic flow based disparity maps.

Finally, consider the following notion: when viewing a discontinuous surface one eye is likely to see some surface detail that is not visible to the other eye. That is, due to the geometry of the situation one eye's view is occluded with respect to the other. This situation has been analysed for optic flow by Mutch & Thompson [90]. Resulting algorithms have been applied to motion sequence images. The application to stereo disparity is clear and is likely to yield a powerful approach. As yet this extension has not taken place.

1.2.4 Distinguishing features of the research presented in this thesis

The research that is presented in the body of this thesis bears some resemblance to several of the studies that have just been reviewed. Most of the analytic developments presented in this thesis are based in differential imaging. Therefore, the closest relatives to the presented work are naturally found in earlier studies of differential imaging. However, the current work makes a number of novel contributions to the disparity interpretation problem. The most significant points of distinction are:

- This thesis emphasizes the recovery of surface geometry (i.e., orientation, curvature, discontinuities, in addition to relative distance) directly from stereo disparity, as opposed to the surface fitting approaches where higher order surface geometry typically is derived only indirectly from distance information.

- Novel relations between the differentially projected orientation of surface detail (e.g., texture) and underlying three-dimensional surface geometry are presented. These relations are used to motivate new and numerically stable methods for recovering three-dimensional surface orientation, distance and stereoscopic viewing parameters from binocular stereo disparity.

- The analysis of stereo disparity that is developed in this thesis also lends insight into the recovery of the discontinuties in distance to three-dimensional surfaces in a viewed scene. In particular, a method for recovering surface discontinuities founded on local disparity based measurements is proposed, implemented and

tested on natural and synthetic stereo data.

- An extensive stability analysis is presented for each of the proposed methods for recovering surface geometry from stereo disparity. This type of detailed analytic stability analysis is uncommon in the computational vision literature.

- The results of the stability analysis indicate not only the requirements for the accurate recovery of surface geometry, but also how disparity interpretation algorithms can monitor the reliability of their own output.

- An empirical psychophysical study is presented that is motivated directly on the analysis of stereo disparity developed in this thesis.

1.3 Outline of chapters

Chapter 1 has served to motivate the problem of understanding stereo disparity as well as provide an overview of related work from the computational vision literature. Chapter 2 unfolds in three sections: The first section (2.1) presents an analysis of stereo disparity resulting from the differential projection of planar surfaces into a pair of images. Section 2.2 studies the stability of this analysis. In section 2.3 a computer program that is based on these analyses is described. The program recovers three-dimensional surface discontinuities from input disparity maps. Chapter 3 extends the analyses and results of chapter 2 to curved surfaces; its three sections parallel those of chapter 2. In chapter 4 some relevant aspects of biological visual systems are presented and discussed. Chapter 5 provides conclusions. Finally, a series of

appendices offer some extensions to the proposed theory.

Chapter 2

Planar surfaces

This chapter is concerned with the analysis of stereo disparity due to the differential projection of planar surfaces onto a pair of two-dimensional imaging surfaces. The goal of this analysis is to develop an understanding of the relations between the geometric structure of a stereo disparity map and the corresponding geometry of a stereoscopically viewed scene. Ultimately it will be shown how it is possible to interpret stereo disparity information in terms of three-dimensional scene geometry. In particular, the stereo information will be used to recover measures of relative distance, surface orientation and surface discontinuities. The developments unfold in three main sections: The first section develops a formal understanding of the disparity field. The second section studies the numerical stability of the relations defined in Section 1. Section 3 describes a set of computer algorithms based on these analyses. The algorithms recover surface discontinuities from stereo disparity. The results of applying these algorithms to several disparity maps are presented.

2.1 Analysis of disparity

In this section a formal analysis of stereo disparity will be presented. The first part of the analysis is concerned with understanding the forward process of differentially projecting a three-dimensional world onto a pair of two-dimensional retinae. This shall lead to defining in turn the stereo disparity field as well as the stereo disparity gradient tensor. The stereo disparity field is a two-dimensional vector field. Horizontal disparity serves to define one component of this field, while vertical disparity serves to define the second component. Horizontal and vertical disparity will be defined in terms of the differential horizontal and vertical position of corresponding elements in the two projected views. The gradient tensor of disparity is a representation of the rate of spatial change in a disparity field. This tensor will lead to the definition of a third type of disparity, orientational disparity. Orientational disparity is the differential orientation of linear elements as imaged in the stereoscopic views.

The latter parts of this section are concerned with the inverse process of recovering three-dimensional scene geometry given a corresponding disparity field. Methods for recovering differential viewing parameters, surface depth, orientation and discontinuities will be developed. The recovery methods will employ only measures of horizontal and orientational disparity. Vertical disparity is not employed due to the fact that its relatively small magnitude leads to numerical instability (see Appendix A). However, it is necessary to introduce vertical disparity in the developments as it serves in the definition of the disparity gradient tensor. Following these formal developments the section closes with a recapitulation of its main results.

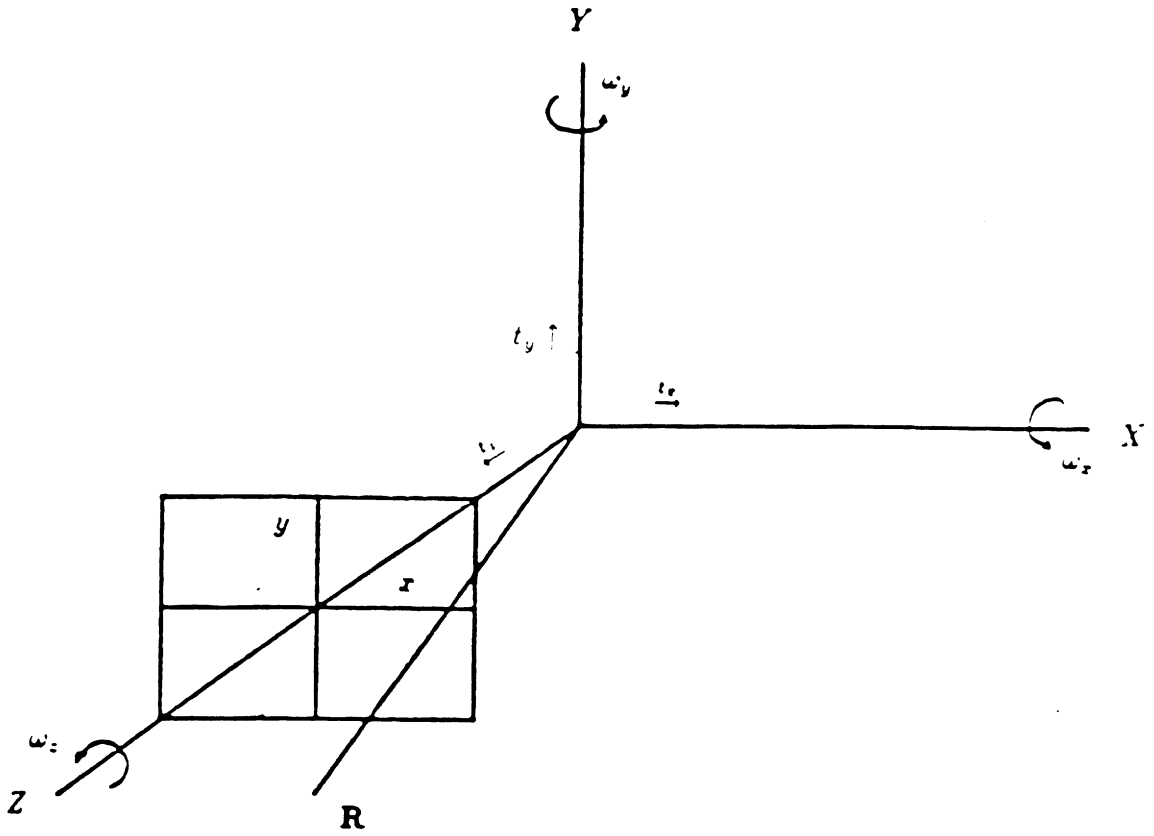


Figure 2.1: A general infinitesimal change of coordinates is composed of a translation $\mathbf{T} = (t_x, t_y, t_z)$ and a rotation $\Omega = (\omega_x, \omega_y, \omega_z)$. A point $\mathbf{R} = (X, Y, Z)$ undergoes perspective projection onto a plane located at $Z = 1$.

2.1.1 Basic differential projection

Given a general change in coordinate systems the corresponding change to a point $\mathbf{R}$ can be described as

$$\delta\mathbf{R} = -\mathbf{T} - (\Omega \times \mathbf{R}) \tag{2.1}$$

where the symbols are described with reference to Figure 2.1.[1] Now, for the case of stereo vision it is not necessary to deal with the most general change of coordinates

[1] A few comments on notation: Throughout this presentation bold-font shall be used for vectors. Upper case letters X, Y and Z will denote world coordinates; while lower case x and y will denote image coordinates. Subscripts will be used for vector components, not to denote differentiation.

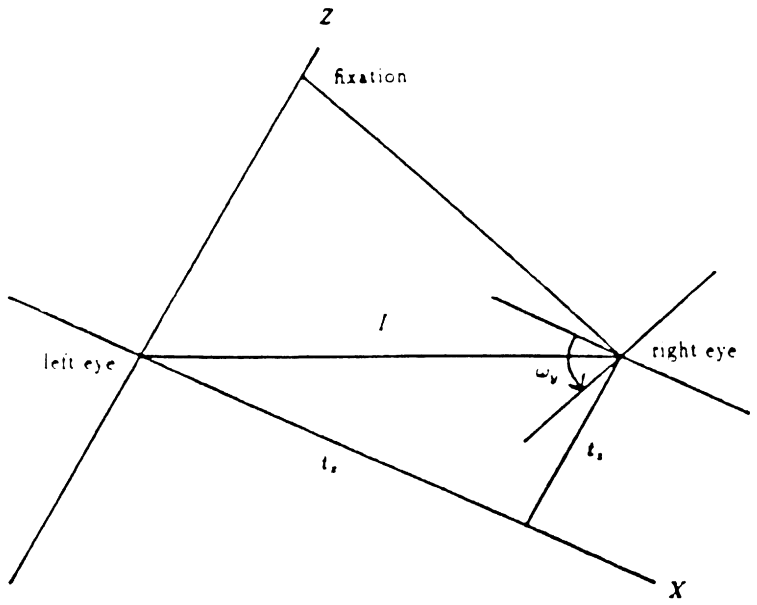


Figure 2.2: A model of stereo viewing geometry.

as described by (2.1). Instead, consideration can be restricted to the model of stereo geometry as given in figure Figure 2.2. This system is related to a coordinate system defined at the optical center of the left eye. The translation components are confined to the plane defined by the view direction and the axis connecting the two eyes; thus, $t_y = 0$. The rotation is confined to rotation about the y-axis; thus, $\omega_x = \omega_z = 0$. This is not to say that elevation of the eyes is not permitted. Rather, the coordinate system is simply always moved with the elevation.[2] For this situation substitution into (2.1) yields

$$\delta\mathbf{R} = -(t_x + \omega_y Z, 0, t_z - \omega_y X). \tag{2.2}$$

Perspective projection serves as the model of how the world projects into an image

[2]From a biological point of view, this model may be considered inadequate as it ignores torsional movements of the eyes about their optical axes. However, if it is desired to include them they would be uniquely defined by the other viewing parameters via Donder's and Listing's laws; see Helmholtz [48].

plane. The laws of perspective give (with appropriate units)

$$x = \frac{X}{Z} \quad y = \frac{Y}{Z}. \tag{2.3}$$

To understand how a point in space changes in its projective coordinates from one view to another let

$$\chi = (\chi_x, \chi_y) = (\delta x, \delta y). \tag{2.4}$$

Considering (2.3) it is found that

$$\chi = \left(\frac{\delta X}{Z} - X\frac{\delta Z}{Z^2}, \frac{\delta Y}{Z} - Y\frac{\delta Z}{Z^2} \right). \tag{2.5}$$

Then upon substituting (2.2) into (2.5)

$$\chi = \left(\frac{1}{Z}(xt_z - t_x) - (x^2 + 1)\omega_y, \frac{1}{Z}(yt_z) - xy\omega_y \right). \tag{2.6}$$

Equation (2.6) is then the basic first-order relation for horizontal and vertical stereo disparity.

Notice that the definition of disparity embodied in (2.6) is somewhat different from the "classical" definition of stereo disparity as presented in, e.g., Ogle [97]. The geometric relation between these two definitions can be clarified with reference to Figure 2.3. This figure depicts a stereoscopic observer fixating a point P_1. The point is projected onto the left and right imaging surfaces via the optical nodes O_l and O_r. The optical nodes in Figure 2.3 correspond to the points labeled "left eye" and "right eye" in Figure 2.2; the stereo baseline I is also the same in both figures. Now, consider the point labeled P_2 in Figure 2.3. The classical definition of disparity for the point P_2 with reference to P_1 would be the difference in the angles δ_2 and δ_1. In contrast, the

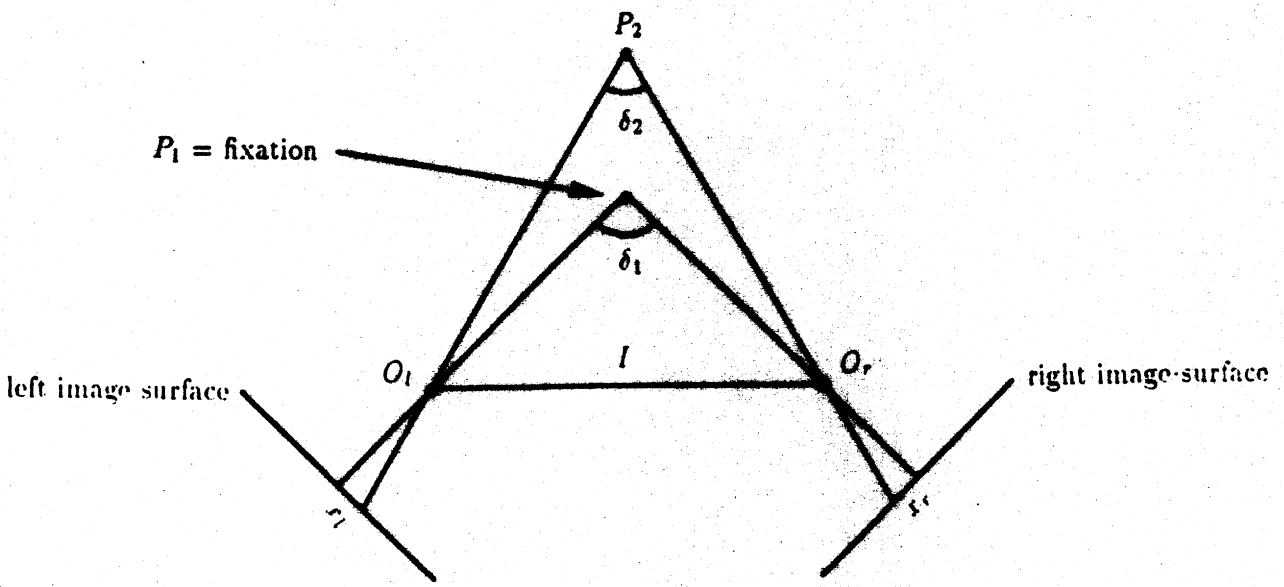


Figure 2.3: The geometric relation between the classical definition of stereo disparity and the definition used in this thesis.

definition of disparity employed in this thesis would assign the difference in projected coordinates x_l and x_r as the horizontal disparity associated with the point P_2.

With the definition of stereo disparity in hand, attention is now directed to the gradient of disparity. This study will lead to further relations between the variables of interest. In particular, from an understanding of the disparity gradient it will be possible to derive relations that concern the gradient of distance (that will later allow the recovery of surface orientation). This gradient is a first-order tensor of the following form

$$\chi' = \begin{pmatrix} \frac{\partial \chi_x}{\partial x} & \frac{\partial \chi_x}{\partial y} \\ \frac{\partial \chi_y}{\partial x} & \frac{\partial \chi_y}{\partial y} \end{pmatrix} \tag{2.7}$$

where

$$\begin{aligned} \frac{\partial \chi_x}{\partial x} &= \frac{t_z}{Z} + \left(\frac{\partial}{\partial x}\frac{1}{Z}\right)(xt_z - t_x) - 2\omega_y x \\ \frac{\partial \chi_x}{\partial y} &= \left(\frac{\partial}{\partial y}\frac{1}{Z}\right)(xt_z - t_x) \\ \frac{\partial \chi_y}{\partial x} &= \left(\frac{\partial}{\partial x}\frac{1}{Z}\right)(yt_z) - \omega_y y \\ \frac{\partial \chi_y}{\partial y} &= \frac{t_z}{Z} + \left(\frac{\partial}{\partial y}\frac{1}{Z}\right)(yt_z) - \omega_y x. \end{aligned} \tag{2.8}$$

To further interpret the relations (2.8) it is necessary to decide upon a representation for the depth parameter, Z. Recalling that the current developments are restricting attention to planar surfaces, consider the standard first-order representation

$$Z = pX + qY + r \tag{2.9}$$

where $(p, q) = \nabla Z$ is the surface gradient and r is the distance along the Z-axis. In terms of image coordinates, (x, y), equation (2.9) becomes

$$\frac{1}{Z} = \frac{1 - px - qy}{r} \tag{2.10}$$

and therefore

$$\frac{\partial}{\partial x}\left(\frac{1}{Z}\right) = -\frac{p}{r} \quad , \quad \frac{\partial}{\partial y}\left(\frac{1}{Z}\right) = -\frac{q}{r} \ . \tag{2.11}$$

Upon substituting (2.11) into relations (2.8) and retaining only first-order terms it is now found that

$$\frac{\partial \chi_x}{\partial x} = \frac{t_z}{Z} - \frac{p}{r}(xt_z - t_x) - 2\omega_y x$$

$$\frac{\partial \chi_x}{\partial y} = -\frac{q}{r}(xt_z - t_x)$$

$$\frac{\partial \chi_y}{\partial x} = -\frac{p}{r}(yt_z) - \omega_y y$$

$$\frac{\partial \chi_y}{\partial y} = \frac{t_z}{Z} - \frac{q}{r}(yt_z) - \omega_y x.$$

As a final step in simplifying the representation of χ' it is useful to choose a coordinate system that is oriented such that the Z-axis is oriented along the line of regard. In this system $x = y = 0$ while $Z = r$.[3] Therefore, the disparity gradient tensor can be written as

$$\chi' = \begin{pmatrix} \frac{1}{r}(pt_x + t_z) & \frac{q}{r}t_x \\ 0 & \frac{t_z}{r} \end{pmatrix}. \tag{2.12}$$

Recalling that an eventual goal is the recovery of useful relations involving planar geometric surface parameters p, q and r, it is pleasing to see these terms appear in the final form of χ' given in (2.12).

[3]Notice that the appropriate change of coordinates is given in terms of Euler angles by transforming the original system according to

$$\begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix}$$

where θ and ϕ are the spherical polar coordinates of the point of regard; see, Goldstein [36] or Korn & Korn [64].

For purposes of analysis it is convenient to split χ' into its symmetric, χ'_+, and antisymmetric, χ'_-, parts. This gives

$$\chi' = \chi'_+ + \chi'_- = \frac{1}{2}\begin{pmatrix} \frac{2}{r}(pt_x + t_z) & \frac{q}{r}t_x \\ \frac{q}{r}t_x & 2\frac{t_z}{r} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & \frac{q}{r}t_x \\ -\frac{q}{r}t_x & 0 \end{pmatrix}.$$

Physically, χ'_+ describes the nonrigid change in shape as an object is differentially projected; while χ'_- describes how an object is rigidly rotated through differential imaging. This interpretation follows directly from the Cauchy-Stokes decomposition theorem of tensor analysis (Aris [6]). For most of the rest of this paper, attention will be restricted to the properties of χ'_+ as it has proven to give the most insight into interpreting the disparity field.

In order to understand the nature of χ'_+ it is useful to perform an eigen-decomposition. (Intuitively speaking, this analysis will yield information about the direction and magnitude of the nonrigid transformation embodied in χ'_+.) The characteristic equation, $det(\chi'_+ - \lambda\mathbf{I}) = 0$ (where $\mathbf{I}$ is the identity matrix), corresponding to χ'_+ is

$$\lambda^2 - \frac{1}{r}(pt_x + 2t_z)\lambda + \frac{1}{r^2}\left(pt_xt_z + t_z^2 - \frac{(qt_x)^2}{4}\right) = 0$$

the roots of which, and hence the eigenvalues, are

$$\lambda = \frac{1}{2r}[pt_x + 2t_z \pm (p^2 + q^2)^{\frac{1}{2}}]. \tag{2.13}$$

For each eigenvalue, λ_i, the equation $(\chi'_+ - \lambda_i\mathbf{I})\xi_i = 0$ yields the corresponding eigenvector ξ_i. This yields

$$\xi = \frac{[p + (p^2 + q^2)^{\frac{1}{2}}, q]}{(p^2 + q^2)^{\frac{1}{2}}} \tag{2.14}$$

as the eigenvector corresponding to the positive root of (2.13). The eigenvector corresponding to the negative root is found to be orthogonal to (2.14). This completes

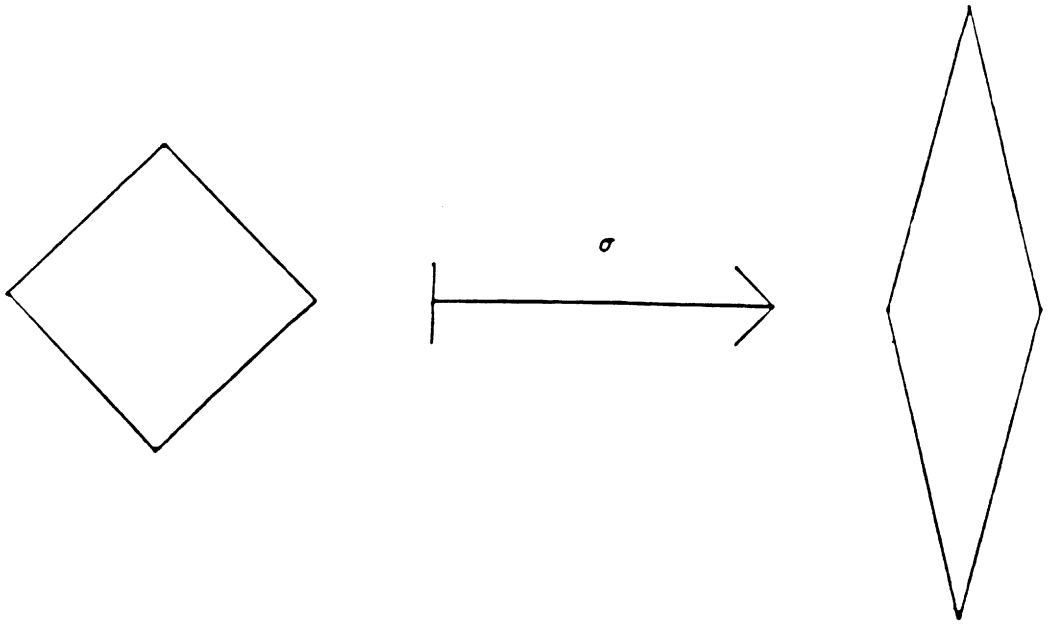


Figure 2.4: The difference of the eigenvalues, σ, of the symmetric part of the disparity gradient tensor, χ'_+, corresponds to a nonconformal but area preserving transformation.

the algebra of the eigen-decomposition. The standard interpretation of such results says that χ'_+ operates on an object by stretching it an amount λ_i along the direction specified by ξ_i.

Should the two values assigned to λ by (2.13) be unequal the deformation embodied by χ'_+ is nonisotropic. To make this notion precise define

$$\sigma = \lambda_{max} - \lambda_{min} = \frac{t_x}{r}(p^2 + q^2)^{\frac{1}{2}}. \tag{2.15}$$

Physically, σ accounts for an area-preserving, but nonconformal transformation between differentially projected images. It may be interpreted as a contraction along the direction of one of the eigenvectors, (2.14), with a corresponding expansion along the other eigenvector, see figure 2.4. Most interestingly for present concerns is that (2.15) is the product of the magnitude of the surface gradient, $(p^2 + q^2)^{\frac{1}{2}}$, and the

depth scaled view translation along the X-axis, $\frac{t_x}{r}$. (Similar results are reported in Koenderink & van Doorn [60].)

For the final series of developments in this section, consider the following intuition: In so far as σ is a nonconformal transformation, it should be possible to relate it explicitly to a change in how angles appear in the differential projections. This would clearly be a desirable result as a change in angles should be directly measurable from a pair of images. Now, the change in orientation of a linear segment is due to the operation of χ'. To understand this it is helpful to express χ' as

$$\chi' = \frac{1}{2}\begin{pmatrix} 0 & \frac{q}{r}t_x \\ -\frac{q}{r}t_x & 0 \end{pmatrix} + \frac{1}{2}(\hat{\xi}_1, \hat{\xi}_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}(\hat{\xi}_1, \hat{\xi}_2)^{-1}$$

where $\hat{\xi}_i$ are the normalized eigenvectors. Next, suppose that the normalized eigenvectors are represented in terms of an angle θ that represents their orientation with reference to the image coordinates. Then $2\chi'$ can be written as

$$\begin{pmatrix} 0 & \frac{q}{r}t_x \\ -\frac{q}{r}t_x & 0 \end{pmatrix} + \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \qquad (2.16)$$

Now, to obtain a relation concerning how a linear segment changes orientation between views: First, apply (2.16) to an oriented segment $(\cos\psi, \sin\psi)$. Second, take the cross-product of the result with the same oriented segment. After some amount of algebraic manipulation it is found that to first order the sine of the angle between the initial segment and the transformed segment is

$$\frac{1}{2}[\sigma \sin 2(\psi - \theta) - qt_x]. \qquad (2.17)$$

Relation 2.17 serves as the definition of orientational disparity, the difference in the orientation of a linear element between two projected views. By taking the difference

of two such measurements, that is a difference in projected angles, the effects of the rigid rotation, qt_x, are discounted. Thus, the suspicion that a change in angles mediated by differential projection should directly reflect the effects of σ is confirmed.

Finally, consider the relation of the vector quantities ξ_i to disparity based measurements. Following through on the difference of two orientational disparities as defined in (2.17) yields

$$\frac{\sigma}{2}[\cos\theta(\sin\psi_1 - \sin\psi_2) + \sin\theta(\cos\psi_2 - \cos\psi_1)] \tag{2.18}$$

where as before $(\cos\theta, \sin\theta)$ specifies the direction of the axes ξ_i and $(\cos\psi_j, \sin\psi_j)$, $j = 1, 2$, specify the directions of the two differentially projected oriented segments. Notice that only the directions of the ξ_i are important. Therefore, an additional pair of orientational disparities allows the unique determination of the eigenvectors ξ_i.

This section has outlined several derivations involving stereo disparity and its gradient. Before proceeding it is useful to pause and emphasize several points:

- Three different types of disparity have been defined: horizontal disparity (2.6), vertical disparity (2.6) and orientational disparity (2.17).

- These disparity mesures along with equations (2.10), (2.14), (2.15) and (2.18) provide relations between stereo disparity, stereo viewing parameters and geometric surface parameters p, q and r.

- In following developments, these key results will lead to relatively straightforward methods for recovering three-dimensional scene geometry given stereo disparity information.

The presentation now turns to deriving these recovery methods.

2.1.2 Recovering view

This subsection presents an approach to recovering the differential viewing parameters t_x, t_z and ω_y which relate a pair of stereo views. The formulation shall exploit horizontal and orientational but not vertical disparity. The restriction from using vertical disparity is motivated by the suspicion that it will not be possible to accurately recover their extremely small values in a real world imaging situation (see, Appendix A). The presented method works with the assumption that the magnitude of the interocular separation is a known value, say I.[4] In the end, the method recovers the viewing parameters only up to an arbitrary scaling factor. This is due to the fact that the distance to some point in the world is assigned an arbitrary value in the course of the solution.

To begin these developments, substitute (2.10) into the horizontal disparity relation from (2.6) to obtain

$$\chi_x = \left(\frac{1 - px - qy}{r}\right)(xt_z - t_x) - (x^2 + 1)\omega_y. \tag{2.19}$$

Now, notice that at the fixation point $(x, y) = (0, 0)$ equation (2.19) reduces to

$$0 = \frac{1}{r}(-t_x) - \omega_y$$

[4]The assumption that this quantity is known a priori is not justified for the general "second view problem". However, for a machine stereo system it is a one time measurable and thus seems reasonable to assume it to be a known quantity.

or

$$\omega_y = \frac{-t_x}{r}. \tag{2.20}$$

Substitution of (2.20) into (2.19) allows for the elimination of one of the view parameters, ω_y.

The next step is to use orientational disparity to eliminate the surface orientation parameters p and q from (2.19). To accomplish this goal notice that (2.14) implies that the ray defined by ξ_1 is half way between the rays defined by (p,q) and the X-axis. (Speaking more generally, ξ_1 is half way between (p,q) and the angular part of $\mathbf{T}$, $\mathbf{T}_{ang}$.) This observation leads one to note that

$$(1,0) \cdot \frac{(p,q)}{(p^2+q^2)^{\frac{1}{2}}} = \hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2$$

and

$$(1,0) \times \frac{(p,q)}{(p^2+q^2)^{\frac{1}{2}}} = 2\hat{\xi}_{1x}\hat{\xi}_{1y}$$

or

$$\frac{(p,q)}{(p^2+q^2)^{\frac{1}{2}}} = (\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2, 2\hat{\xi}_{1x}\hat{\xi}_{1y}) \tag{2.21}$$

where $\hat{\xi}_1 = (\hat{\xi}_{1x}, \hat{\xi}_{1y})$ is ξ_1 normalized.[5] Now rewriting (2.15) as

$$(p^2+q^2)^{\frac{1}{2}} = \frac{r\sigma}{t_x}$$

allows substitution into (2.21) for the term $(p^2+q^2)^{\frac{1}{2}}$ with the result that the surface orientation parameters, p and q, can be substituted for in (2.19) as

$$\begin{aligned} p &= (\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2)\frac{r\sigma}{t_x} \\ q &= (2\hat{\xi}_{1x}\hat{\xi}_{1y})\frac{r\sigma}{t_x} \end{aligned}. \tag{2.22}$$

[5]Notice that in three-dimensions the operation $\times$, the cross product of two vectors, yields a vector quantity. However, here in the two-dimensional case $\times$ is the rotational; it yields a scalar quantity.

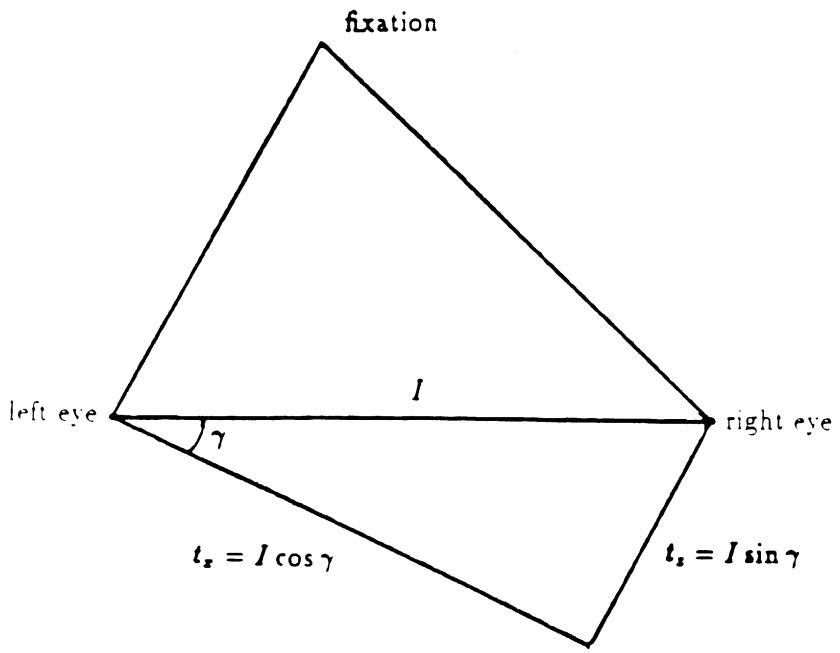


Figure 2.5: The definition of γ

Now, the viewing parameters t_x and t_z can be related in a relatively simple equation with one further manipulation: allow for an arbitrary depth scale and set the remaining surface parameter, r, to an arbitrary value of unity. This yields a relation of the form

$$a_0 = a_1 t_x + a_2 t_z + a_3 \frac{t_x}{t_z} \tag{2.23}$$

where the a_i consist entirely of known (or measurable) values. Explicitly,

$$a_0 = \chi_x - \sigma[x(\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2) + 2y\hat{\xi}_{1x}\hat{\xi}_{1y}]$$

$$a_1 = x^2$$

$$a_2 = x$$

$$a_3 = -\sigma[x^2(\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2) + 2xy\hat{\xi}_{1x}\hat{\xi}_{1y}].$$

Relation (2.23) could be used to solve for the desired parameters t_x and t_z in a number of ways. Here, the system is solved by making use of several substitutions and a small angle approximation. Let γ be defined as shown in Figure 2.5. From

figure 2.4 it is clear that

$$\begin{aligned} t_x &= I\cos\gamma \\ t_z &= I\sin\gamma \\ \tfrac{t_z}{t_x} &= \tan\gamma \end{aligned} \qquad (2.24)$$

Substituting (2.24) into (2.23) results in

$$a_0 = a_1 I\cos\gamma + a_2 I\sin\gamma + a_3\tan\gamma. \qquad (2.25)$$

The next step is to take standard first-order trigonometric substitutions so that γ may be solved for as

$$\gamma = \frac{a_0 - a_1 I}{a_3 + a_2 I} \qquad (2.26)$$

with I known.[6] With γ recovered the original view parameters t_x, t_z and ω_y are easily obtained with reference to (2.24) and (2.20).

Reviewing these results, it is found that the view parameters have been recovered using only a single horizontal disparity and a pair of orientational disparities.

2.1.3 Recovering geometric surface parameters

With the viewing parameters recovered consideration can be turned to recovering the geometric surface parameters p, q and r. Notice with the viewing parameters recovered the Z value of any point can be easily recovered by consideration of either of the relations from equation (2.6). Now adopt a change of coordinates such that the new Z-axis points toward the point of consideration (as discussed in section 2.1.1).

[6]Notice that under the vast majority of real world viewing conditions (e.g., observer fixating not too eccentrically and at a moderate distance), γ will in fact be small.

In this new coordinate system the recovered Z value can be interpreted as r. It now remains to recover only p and q. But of course, the necessary relations are already in hand. Equations (2.22) derived to eliminate p and q for the recovery of viewing parameters can (now that the view has been recovered) be used to recover these same values. Thus, minimal requirements for the recovery of the surface parameters p, q and r are the observation of a single horizontal disparity as well as three orientational disparities which all derive from the same surface patch.

2.1.4 Recovering surface discontinuities

Suppose that the geometric surface parameters corresponding to two adjacent surface patches have been recovered as (p_1, q_1, r_1) and (p_2, q_2, r_2), respectively. Then there is a trivial test for surface discontinuities for the case of planar surfaces: Specifically, require that

$$(p_1, q_1, r_1) = (p_2, q_2, r_2). \tag{2.27}$$

If the test (2.27) fails a surface discontinuity is necessarily present. Notice that, strictly speaking, any triad of surface parameters computed by the methods proposed earlier in this section are actually defined only in a local coordinate system. This system was taken with the Z-axis pointing along the corresponding line of regard. In order to actually make sensible computations involving parameters derived for separate systems the triads must be appropriately rotated into a common coordinate system. Use of the inverse of the matrix presented in footnote 3 is appropriate.

2.1.5 Recapitulation

This section has presented an analysis of stereo disparity due to the differential projection of a three-dimensional scene onto a pair of two-dimensional imaging surfaces. The developments have been restricted to the projection of planar surfaces arranged in three-space.

The section began by developing the basic relations for differential projection. In this light, results relating horizontal and vertical disparity to stereo viewing parameters and three-dimensional depth were derived, equation (2.6). The second major development was to derive relations involving surface gradient, the gradient of disparity and orientational disparity. These relations are embodied in equations (2.14), (2.15) and (2.17). With these basic results in hand it was possible to turn attention to inverting the disparity information to recover properties of the differentially projected world. Specifically, relations were derived for recovering the differential viewing parameters, surface depth, gradient and discontinuity. Equation (2.26) was derived for the recovery of surface viewing parameters from horizontal and orientational disparities. Equations (2.22) were derived for recovering surface gradient. Finally, relations defined with reference to (2.27) gave a method for recovering surface discontinuities from disparity. The key to the strong results obtained for recovering three-dimensional properties from disparity lay in first developing an understanding of the disparity field itself.

2.2 Stability analysis

At this point it is useful to analyze the numerical properties of the proposed recovery methods. In turn, this section shall consider degenerate sets of measurements, sensitivity to measurement errors and an approach to operating in the face of noise corrupted data. Finally, the section closes with a recapitulation of the main results.

2.2.1 Degeneracies

It is possible that certain combinations of measured disparities and image coordinates will lead to situations where the proposed recovery methods will be undefined. Such situations shall be referred to as degenerate. In this subsection these situations will be analyzed. Of particular interest shall be those data combinations which lead to a ratio becoming undefined as its denominator tends to zero.

Consider first the key relation for defining the viewing parameters, (2.23). Relation (2.23) will become undefined as its denominator approaches zero. Thus, it is necessary to consider the condition

$$0 = a_3 + a_2 I$$

or, upon appropriate substitution

$$0 = xI - \sigma[x^2(\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2) + 2xy\hat{\xi}_{1x}\hat{\xi}_{iy}].$$

Examination of this quantity indicates that the image line $x = 0$ is degenerate. Continuing by making the substitutions implied by (2.15) and (2.19) and cancelling

appropriately yields

$$0 = I - \frac{t_x}{r}(px + qy)$$

or

$$1 = \frac{t_x(px + qy)}{Ir} \tag{2.28}$$

as a degenerate condition. In words: The numerator of (2.28) is the product of two factors: The first factor, t_x, is the projection of the stereo baseline on the X-axis. The second factor, $(px + qy)$, is the radial distance from the point of regard to the Z-intercept of the corresponding plane. The denominator of (2.28) is the product of the stereo baseline and the Z-intercept of the surface of regard. These two quantities must be equal for the viewing solution (2.23) to be undefined. It is quite unlikely that such a configuration should occur generically. For intuition, notice that in typical viewing conditions $\|t_x\| \approx \|I\|$. Therefore, the degenerate condition demands that the surface of regard is viewed at a point where it is approximately the same radial distance from its Z-intercept as the Z-intercept is from the viewer. See figure 2.6.

Now, turn attention to degeneracies related to the recovery of the components of the surface gradient $\nabla Z = (p, q)$ defined in (2.19). Two conditions present themselves. First, should the plane of consideration pass throught the origin (i.e., the optical center of the left eye) the solution will not apply. In this situation the plane appears as a line to the left eye. Second, should $t_x = 0$ then (2.19) is undefined. For stereo vision this a mechanical impossibility as it requires one eye to be directly behind (and hence see through) the other eye. Recalling that the method for recovering surface discontinuities (2.27) is directly related to (2.19), leads to the conclusion that it shares

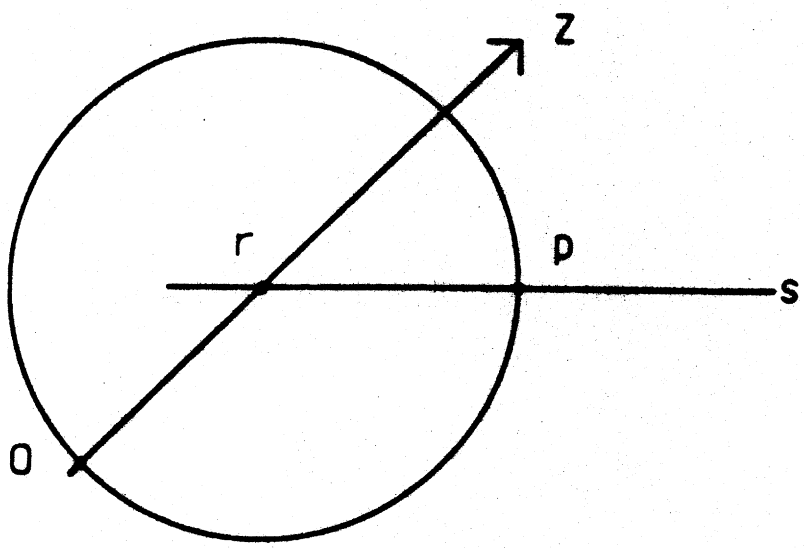


Figure 2.6: A geometric configuration of surface and viewer leading to a degeneracy of the proposed method for recovering view and surface geometry. An observer, o, views a point, p, on a planar surface, S. Suppose that S intercepts the Z-axis at r. The degenerate condition is $\|or\| = \|pr\|$. The points o and p must lie on a circle centered at r.

the same degeneracies.

Happily, it is possible to draw positive conclusions concerning the degeneracies of the proposed recovery methods. Specifically:

- The degenerate conditions of the solution methods embodied in (2.19) and (2.27) are either unlikely to occur or impossible for a binocular stereo system.
- The degeneracy of practical importance for (2.23) is the image line $x = 0$. This condition can be easily diagnosed during the recovery process.

2.2.2 Error analysis

Now consider the effects of applying perturbations to the data upon which the three-dimensional recovery methods operate. The general approach shall be to outline those conditions and choices of image measurements which lead to especially stable or unstable solutions. As a measure of stability the "generalized error-propagation formula" (Dahlquist & Björk [21]) will be used. This measure, which tells the local rate of change of a solution with respect to perturbed data, can be written

$$\Delta y \approx \sum_{i=1}^{n} \frac{\partial y}{\partial x_i}(\tilde{\mathbf{x}}) \cdot \Delta x_i$$

and hence,

$$\|\Delta y\| \lesssim \sum_{i=1}^{n} \|\frac{\partial y}{\partial x_i}(\tilde{\mathbf{x}})\| \cdot \|\Delta x_i\| \tag{2.29}$$

where

$$y = y(\mathbf{x})$$

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

and the perturbation to x_i is Δx_i resulting in

$$\tilde{x_i} = x_i + \Delta x_i$$

$$\tilde{\mathbf{x}} = (\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n})$$

$$\Delta y = y(\tilde{\mathbf{x}}) - y(\mathbf{x}).$$

Clearly, small values for $\|\Delta y\|$ correspond to stable solutions.

Imaging perturbations

Consider the effects of error perturbations applied in the disparity map to the measures χ_x, σ and ξ.[7] The general approach shall be to derive the form of the general error propagation formula (2.29) for each of the recovery methods. With the derivation in hand, each term in the summation can be examined separately for stable conditions (which correspond to small magnitude). The intersection of these conditions for all the terms will be stable for the combined form.

First, turn attention to the stability of the view recovery method (2.26). Then the parameters of (2.29) become $\mathbf{x} = (\chi_x, \sigma, \theta)$ and $\tilde{\mathbf{x}} = \mathbf{x} + (\Delta\chi_x, \Delta\sigma, \Delta\theta)$, with $(\cos\theta, \sin\theta)$ specifying the direction of ξ. Considering (2.29), the goal is to understand the conditions where

$$\|\frac{\partial\gamma}{\partial\chi_x}(\tilde{\mathbf{x}})\| \cdot \|\Delta\chi_x\| + \|\frac{\partial\gamma}{\partial\sigma}(\tilde{\mathbf{x}})\| \cdot \|\Delta\sigma\| + \|\frac{\partial\gamma}{\partial\theta}(\tilde{\mathbf{x}})\| \cdot \|\Delta\theta\| \tag{2.30}$$

is small. To begin, notice that trivially (2.30) can have arbitrarily small magnitude as $(\Delta\chi_x, \Delta\sigma, \Delta\theta) \rightarrow (0,0,0)$. More realistically, the perturbations, $(\Delta\chi_x, \Delta\sigma, \Delta\theta)$, need

[7] To make the error analysis managable it shall be assumed that errors in the assessment of the visual directions (i.e., (x,y)) are negligible as compared to those in the disparity measurements. Hence, the following developments will only consider perturbations to the disparity measures.

to be small compared to the denominators of the corresponding partial derivatives. These denominators shall now be examined in some detail. The term $\frac{\partial \gamma}{\partial \chi_x}(\tilde{\mathbf{x}})$ can be expanded (with the aid of double angle formulas) as

$$-x[\tilde{\sigma}(x,y)\cdot(\cos 2\theta, \sin 2\theta) + I]^{-1}$$

or

$$-x[\tilde{\sigma}\|(x,y)\|\cos\psi + I]^{-1} \tag{2.31}$$

where ψ is the angle between (x,y) and $(\cos 2\tilde{\theta}, \sin 2\tilde{\theta})$. From (2.31) it can be concluded that the error due to the first term of (2.30) can be made small given three conditions: (i) σ, the magnitude of stereoscopic shear, is large, (ii) I, the magnitude of the stereo baseline, is large, (iii) (x,y) is chosen in the direction of twice θ (i.e., twice the directions of the σ-axis, ξ). Using similar procedures and nomenclature, the second term of (2.30) can be written

$$\frac{\|(x,y)\|\cos\psi[\tilde{\chi}_x - I(1+x^2) - 2\tilde{\sigma}\|(x,y)\|\cos\psi]}{x(\tilde{\sigma}\|(x,y)\|\cos\psi + I)^2}. \tag{2.32}$$

By inspection it is possible to conclude that (2.32) has small magnitude when σ and I are relatively large. In the limit, as $\|(x,y)\| \to \infty$ l'Hopital's rule (Korn & Korn [64]) suggests that taking (x,y) in the direction 2θ is useful provided that σ is relatively large. For more moderate $\|(x,y)\|$ case analyses still suggest this is the appropriate direction for (x,y), particularly for $\tilde{\chi}_x \approx -[\tilde{\sigma}\|(x,y)\| + I(1+x^2)]$. The last term of (2.30) can be written

$$\frac{2\tilde{\sigma}\|(x,y)\|\sin\psi[\tilde{\chi}_x - I(1+x^2)] - 4\tilde{\sigma}\|(x,y)\|^2\sin 2\psi}{x[\tilde{\sigma}\|(x,y)\|\cos\psi + I]^2}. \tag{2.33}$$

Consideration of (2.33) reveals that σ and I large with (x, y) chosen in the direction 2θ leads to its small magnitude. It is also useful for $\chi_x \approx -I(1 + x^2)$. Notice in particular that the numerator of (2.33) contains terms of $\sin\psi$. This heightens the importance of choosing (x, y) in the direction 2θ.

Thus, it is possible to conclude that three conditions are particularly important for (2.30) to have stable solutions:

- The magnitude of stereoscopic shear, σ, should be relatively large. In terms of world conditions this condition (i), implies that the magnitude of the surface gradient is large while the viewing distance is not too great. This is just another way of saying that the differential perspective information must be salient.

- The image coordinates, (x,y), should be chosen in the direction of twice θ (i.e., twice the directions of the σ-axis, ξ). In the world this means that the image coordinates should be chosen in the direction of ∇Z, see (2.21). Intuitively, the data points are best when chosen in the direction of the projection of the surface gradient.

- The magnitude of the stereo baseline, I, should be relatively large. Again, this condition is related to making the disparity information as salient as possible.

Notice that the final condition can be satisfied as a one time design constraint on a stereo system. Similarly notice that the first two conditions can be monitored by an intelligent disparity processing algorithm. This last observation deserves emphasis:

- This analysis indicates that the view recovery method can guide its own behavior in order to recover a stable solution given input disparity information.

Now, consider how errors in the measurements affect the recovery of the local depth parameter, r. By an appropriate local transformation (e.g., the matrix of footnote 4) r is recovered in terms of Z. Therefore, the appropriate relation is (2.6). For the following analysis the potential sources of error will be in the horizontal disparity, χ_x, as well as the previously recovered view parameters t_x, t_z and ω_y. Thus the parameters of (2.29) become $\mathbf{x} = (\chi_x, t_x, t_z, \omega_y)$ and $\tilde{\mathbf{x}} = \mathbf{x} + (\Delta\chi_x, \Delta t_x, \Delta t_z, \Delta\omega_y)$. Specializing (2.29) with respect to (2.6) leads to

$$\|\frac{\partial Z}{\partial \chi_x}(\tilde{\mathbf{x}})\| \cdot \|\Delta\chi_x\| + \|\frac{\partial Z}{\partial t_x}(\tilde{\mathbf{x}})\| \cdot \|\Delta t_x\| + \|\frac{\partial Z}{\partial t_z}(\tilde{\mathbf{x}})\| \cdot \|\Delta t_z\| + \|\frac{\partial Z}{\partial \omega_y}(\tilde{\mathbf{x}})\| \cdot \|\Delta\omega_y\|. \quad (2.34)$$

The term $\frac{\partial Z}{\partial \chi_x}(\tilde{\mathbf{x}})$ evaluates to

$$\frac{t_x - x t_z}{[\chi_x + (x^2+1)\omega_y]^2}. \quad (2.35)$$

Inspection of (2.35) shows that its contribution to (2.34) will be small if three conditions are met: (i) the horizontal disparity, χ_x, is relatively large, (ii) the rotation about the Y-axis, ω_y is relatively large and (iii) the difference of the two relative view translations, t_x and t_z, is relatively small. Intuitively, these observations suggest that stable situations result from those viewing conditions that tend to maximize the difference in the two stereo views. Similarly, the $\frac{\partial Z}{\partial \omega_y}(\tilde{\mathbf{x}})$ term of (2.34) evaluates to

$$\frac{t_x(x^2+1) - t_z(x^3+x)}{[\chi_x + (x^2+1)\omega_y]^2}$$

which can be seen to have the same conditions for small magnitude as does (2.35). Finally, $\frac{\partial Z}{\partial t_x}(\tilde{\mathbf{x}})$ and $\frac{\partial Z}{\partial t_z}(\tilde{\mathbf{x}})$ yield

$$-[\chi_x + (x^2+1)\omega_y]^{-1}$$

and

$$x[\chi_x + (x^2 + 1)\omega_y]^{-1},$$

respectively. For these last two cases only the conditions that both χ_x and ω_y are relatively large are necessary for stability. On the whole it can be concluded that the magnitude (2.34) can be kept small (and hence the recovery of r stable) provided that viewing conditions are chosen to make the difference in the two stereo views salient.

The final developments of this section consider the stability of the surface orientation measures as embodied in (2.22). For these cases the parameters of (2.29) become $\mathbf{x} = (\sigma, \theta, r, t_x)$ and $\tilde{\mathbf{x}} = \mathbf{x} + (\Delta\sigma, \Delta\theta, \Delta r, \Delta t_x)$. (As earlier $(\cos\theta, \sin\theta)$ specifies the direction of the eigenvector ξ.) The measure of stability for p can be written

$$\|\frac{\partial p}{\partial \sigma}(\tilde{\mathbf{x}})\| \cdot \|\Delta\sigma\| + \|\frac{\partial p}{\partial \theta}(\tilde{\mathbf{x}})\| \cdot \|\Delta\theta\| + \|\frac{\partial p}{\partial r}(\tilde{\mathbf{x}})\| \cdot \|\Delta r\| + \|\frac{\partial p}{\partial t_x}(\tilde{\mathbf{x}})\| \cdot \|\Delta t_x\|. \quad (2.36)$$

The terms of interest in (2.36) evaluate to

$$\frac{\partial p}{\partial \sigma}(\tilde{\mathbf{x}}) = \cos 2\tilde{\theta}\frac{\tilde{r}}{\tilde{t}_x}$$

$$\frac{\partial p}{\partial \theta}(\tilde{\mathbf{x}}) = -2\sin 2\tilde{\theta}\frac{\tilde{r}\tilde{\sigma}}{\tilde{t}_x}$$

$$\frac{\partial p}{\partial r}(\tilde{\mathbf{x}}) = \cos 2\tilde{\theta}\frac{\tilde{\sigma}}{\tilde{t}_x}$$

$$\frac{\partial p}{\partial t_x}(\tilde{\mathbf{x}}) = -\cos 2\tilde{\theta}\frac{\tilde{r}\tilde{\sigma}}{\tilde{t}_x^2}$$

Similarly, the relation for q substituted into (2.29) leads to consideration of

$$\|\frac{\partial q}{\partial \sigma}(\tilde{\mathbf{x}})\| \cdot \|\Delta\sigma\| + \|\frac{\partial q}{\partial \theta}(\tilde{\mathbf{x}})\| \cdot \|\Delta\theta\| + \|\frac{\partial q}{\partial r}(\tilde{\mathbf{x}})\| \cdot \|\Delta r\| + \|\frac{\partial q}{\partial t_x}(\tilde{\mathbf{x}})\| \cdot \|\Delta t_x\|. \quad (2.37)$$

For (2.37) it is found that

$$\frac{\partial q}{\partial \sigma}(\tilde{\mathbf{x}}) = \sin 2\tilde{\theta}\frac{\tilde{r}}{\tilde{t}_x}$$

$$\frac{\partial q}{\partial \theta}(\tilde{\mathbf{x}}) = 2\cos 2\tilde{\theta}\frac{\tilde{r}\tilde{\sigma}}{\tilde{t}_x}$$

$$\frac{\partial q}{\partial r}(\tilde{\mathbf{x}}) = \sin 2\tilde{\theta}\frac{\tilde{\sigma}}{\tilde{t}_x}$$

$$\frac{\partial q}{\partial t_x}(\tilde{\mathbf{x}}) = -\sin 2\tilde{\theta}\frac{\tilde{r}\tilde{\sigma}}{\tilde{t}_x^2}$$

Examining the expansions of the terms in (2.36) and (2.37) leads to the conclusion that stable solutions are reached when r is relatively small while t_x is relatively large. The requirement that t_x have relatively large magnitude is consistent with earlier results on the importance of keeping the stereoscopic differences salient. The importance of r not being too large reflects the fact that differences in relative surface orientation become less salient as viewing distance increases.

The observations made on the stable recovery of the geometric surface parameters are in accord with the earlier stability results. The general conclusion is worth highlighting:

- The key conditions leading to stable recovery of both view and surface geometry center around making the differential viewing information salient; good stereo viewing conditions lead to good solutions.

Three-dimensional perturbations

Implicit in the developments up to this point has been the assumption that the disparity measurements (horizontal and orientational) derive from the differential projection of surface detail that lies upon the surface of concern (e.g., the stripes of an animal lie upon the animal's surface). In many real world situations surface

texture may extend away from the underlying surface (e.g., the quills of a porcupine extend away from the animal's surface). It is reasonable to ask if this sort of three-dimensional perturbation can be considered in the light of image perturbations. The answer hinges upon how adequate it is to express the effects of the three-dimensional perturbations as simple additive error components to the disparity measurements. Consider in turn the effects of such perturbations on horizontal and orientational disparity.

The effects of adding a three-dimensional perturbation to horizontal disparity, χ_x, can be conceptualized as follows: Suppose that a point along a texture element projects into the images to give rise to a horizontal disparity. Let ΔZ be the amount that the point of consideration extends away from its surface, Z. Recalling (2.6) the perturbed horizontal disparity can be written

$$\tilde{\chi_x} = \frac{1}{Z + \Delta Z}(xt_z - t_x) - (1 + x^2)\omega_y.$$

Expressing this as a sum of the unperturbed disparity and a component due to the three-dimensional perturbation, ΔZ, yields

$$\chi_x + \Delta\chi_x = \frac{1}{Z}(xt_z - t_x) - (1 + x^2)\omega_y + \frac{\Delta Z(t_x - xt_z)}{Z(Z + \Delta Z)}. \tag{2.38}$$

Equation (2.38) shows that the three-dimensional perturbation interacts in a nonlinear fashion with parameters which are to be recovered, t_x and Z.

Now consider the effects of three-dimensional perturbations on orientational disparity. Suppose that a texture element extends from a plane $\mathbf{P} = (p, q, r)$. This element can be considered as embedded in a plane related to $\mathbf{P}$ as $\tilde{\mathbf{P}} = \mathbf{P} + (\Delta p, \Delta q, \Delta r)$. Next, recall that the differentially projected orientation of a linear element $(\cos\psi, \sin\psi)$

is due to the operation of the disparity gradient tensor χ'. The effects of the three-dimensional perturbations on orientational disparity can be formalized by using the parameters of $\check{\mathbf{P}}$ to form $\check{\chi}'$ and applying the result to $(\cos\psi, \sin\psi)$. With the aid of (2.12) this results in an orientation of

$$\frac{1}{r+\Delta r}\left([(p+\Delta p)t_x + t_z]\cos\psi + (q+\Delta q)t_x \sin\psi, t_z \sin\psi\right).$$

This can be usefully rewritten as a sum of two terms. The first term is entirely due to the effects of an element lying on the plane $\mathbf{P}$. The second term is due to the perturbation $\Delta\mathbf{P}$. Specifically, the first term is

$$\frac{1}{r}\left[(pt_x + t_z)\cos\psi + qt_x \sin\psi, t_z \sin\psi\right]$$

while the second (error) term is

$$\frac{1}{r(r+\Delta r)}\left((\cos\psi, \sin\psi)\cdot[t_x(r\Delta p - p\Delta r) + t_z(r-\Delta r), t_x(r\Delta q - q\Delta r)], t_z\Delta r \sin\psi\right). \tag{2.39}$$

Again, the results of writing the effects of three-dimensional perturbations as sum of unperturbed and perturbed terms results in a nonlinear error term. Significantly, the nonlinearities of the error involve not only the perturbations, but also the variables to be recovered.

Thus, even in the raw disparities, (2.38) and (2.39), three-dimensional perturbations combine in a nonlinear fashion with the parameters to be recovered. From this observation two conclusions can be drawn:

- Three-dimensional perturbations can not be adequately characterized as analogous to additive image perturbations.

- Three-dimensional perturbations can well lead to unstable performance in algorithms that operate on horizontal and orientational disparities. The nonlinear nature of perturbations in the disparity data will only be compounded as these values are operated upon.

An algorithmic approach to dealing with three-dimensional texture could attempt to infer an underlying surface in a piecemeal fashion by locally fitting surface patches (with, e.g., least-squares) to the endpoints of the texture elements. However, such an approach goes against the main philosophy of the present approach: the recovery of higher-order surface geometry directly from stereo disparity. A deeper approach would be founded in a theory of three-dimensional texture. The development of such a theory of three-dimensional texture is beyond the scope of this thesis. However, if an algorithm based on the theory presented in this thesis is to exhibit robust behavior in the real world it must take account of such perturbations in some fashion. These problems are considered in the following subsection.

2.2.3 Operating in the face of perturbed data

Having developed a feel for the behavior of the recovery methods in the face of perturbed data, it is useful to develop strategies for the recovery of the desired parameters in spite of such perturbations. Given that the implementation reported in section 3 concentrates on recovering view and surface discontinuities, this subsection shall be limited to considering those cases. Specifically, approaches to combining redundant data shall be considered as well as the selection of thresholds in the detection of

surface discontinuities.

Recovering view parameters

The recovery of viewing parameters via (2.23) does not require the selection of any thresholds. The issue is how to combine redundant sensing measures. That is, suppose it is possible to acquire multiple measures of horizontal disparity, χ_x, stereoscopic shear, σ, as well as the σ-axis, ξ. While only one set of measures is needed to define a solution for γ, it is desirable to combine multiple measures for the sake of robustness. Several paths are possible.

Perhaps the most well founded approach would be to derive optimal filters to minimize the noise of the disparity data values. Two facts make this an impractical goal: First, the nonlinear fashion in which the data measures interact make much of standard estimation theory nonapplicable (but see, Oppenheim et al. [99]). Second, due to the complex effects of three-dimensional texture based noise it is not possible to invoke the typical assumptions of estimation theory (e.g., stationary noise, etc).

At the other extreme of sophistication might be to apply a pseudo-inverse based solution (Albert [1]). (For the following discussion define $\alpha_i = a_{3i} + a_{2i}I$ and $\beta_i = a_{0i} - a_{1i}I$ with the a's defined as in (2.20). Also, let the subscript i representing the ith measurement set of the disparities.) Briefly, the idea is to minimize the sum of the squared errors

$$\epsilon_i = \gamma\alpha_i - \beta_i$$

which can be written in matrix form as

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} (\gamma) = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

and more concisely as $\mathbf{A}\Gamma = \mathbf{B} + \epsilon$. The solution, in terms of the pseudo-inverse $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is

$$\Gamma = \mathbf{A}^\dagger(\mathbf{B} + \epsilon).$$

The problem with this solution is that it implicitly assumes that there is only error in the terms β_i. For the present problem it is at least as likely that the terms α_i are noise corrupted. This assumption seems too naive for present concerns. The approach also makes rather simple assumptions about the distribution of the error (zero-mean, Gaussian, random noise) that are probably not appropriate.

The approach adopted here is to use an eigenvector-value based approach to combining multiple data sets. This approach as implemented makes the same (naive) assumptions with regard to the distribution of errors as does the pseudo-inverse method. However, this method is more appropriate in that it recognizes the possibility of error in both the α_i and the β_i terms. In this case, the squared error is minimized along the direction Γ^T, with Γ being $(\gamma, -1)$ normalized. Correspondingly, the error becomes

$$\epsilon_i = \frac{\gamma\alpha_i - \beta_i}{[(\gamma, -1)(\gamma, -1)^T]^{\frac{1}{2}}}.$$

Minimizing the sum of such squared errors leads to consideration of the matrix equa-

tion

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \vdots & \vdots \\ \alpha_n & \beta_n \end{pmatrix} \Gamma^T = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

which can be rewritten as

$$\mathbf{D}\Gamma^T = \epsilon.$$

It can be shown (e.g., Koopmans [63]) that the summed squared error is minimized by selecting Γ as the eigenvector corresponding to the smallest eigenvalue of $\mathbf{D}^T\mathbf{D}$. This is the method of estimating γ that shall be used. It is seen as a compromise between the prohibitive cost of the nonlinear estimation problem and the simple pseudo-inverse method. Intuitively, the chosen measure is minimizing the perpendicular distance of a line plotted through the values (α_i, β_i). It is worth noting that the nonlinear forms of the data may not be too nonlinear given that the values for σ and ξ should vary minimally for a local neighborhood from a planar surface.

Recovering surface geometry

The issues surrounding the recovery of surface discontinuities in the face of noise perturbed data are more difficult than the recovery of view. This is due to the fact that not only is it necessary to combine redundant data, but also thresholds must be set in the comparisons of adjacent regions of the disparity map. It would be possible to apply the eigen approach of the previous subsection to combine data. However, due to the nature of the error term its interpretation in terms of a threshold would be du-

bious. On the other hand, a nonlinear optimal estimation approach is prohibitively complex. In the face of unknown three-dimensional texture perturbations, it may not be possible at all unless ad hoc assumptions are made about the corresponding noise distributions. A principled approach to combining redundant data and establishing thresholds while making minimal assumptions about the form of the error and data can be found in nonparametric statistical methods. In this subsection an approach is developed based on histogramming local measurements and applying the nonparametric method of Kolmogorev-Smirnov (Siegel [113]).

The Kolmogorev-Smirnov method is a two-sample test of whether two samples have been drawn from the same source. A large deviation between two sample cumulative distributions is taken as evidence that the samples were drawn from distinct sources. More precisely: Let $x_1 < x_2 < \ldots < x_n$ and $y_1 < y_2 < \ldots < y_m$ be the ordered samples from two sources that have continuous cumulative distribution functions $F(z)$ and $G(z)$. Also, let $S_n^x(z) = \frac{k}{n}$ with k the number of samples less than or equal to z for the set x_i. Similarly, let $S_m^y(z) = \frac{l}{m}$ for the set y_i. Then the measure

$$D = \max_z \|S_n^x(z) - S_m^y(z)\| \tag{2.40}$$

can be used to decide if $F(z) \equiv G(z)$. For small sample size and $n = m$ the probability that $D \leq \frac{h}{n}$, $h = \max \|k - l\|$, can be derived via a set of recurrence relations. The results of this computation are commonly available in sources such as Siegel [113] or Korn & Korn [64]. The Kolmogorev-Smirnov test is a particularly good tool for dealing with small sample sizes. It can be shown that compared to the t-test it has a power-efficiency of 96 per cent for small values of n (Dixon [23]). Power in the

face of small sample size is important in the current application to recovering surface discontinuities. If measures that have been derived locally can be compared with confidence, then the ability to accurately localize differences will be improved.

Thus, the proposed method for recovering surface discontinuities from inaccurate data is to use the Kolmogorev-Smirnov method to compare locally histogrammed values of the surface parameters p, q and r. Interestingly, if the same confidence level is used for comparing all three histogram pairs the local differences in distance, r, will be judged as more important than local differences in surface gradient, p and q. This result seems to be in accord with intuition on the relative importance of these parameters in recovering surface (distance) discontinuities. The differential weighting falls out of the fact that distance can be recovered via disparate points , while orientation is recovered via disparate linear segments. Thus, the observation of one segment would always allow for the observation of multiple points. Therefore, for a given area of the disparity map the count in the distance histogram would be greater than that in either component of the gradient.

As the final consideration in this section, turn attention to the confidence measure D. This measure serves to set the threshold for discontinuity detection. In this thesis the value of D will remain as a free parameter in the recovery of surface discontinuities. However, it is worth noting three possible approaches to setting the value of D in a more well founded fashion. First, suppose the error in the disparity measures was assumed to have a known and simple distribution (e.g., a simple exponential distribution). In this case it might be possible to derive the value of D from residual measures of disparity. Unfortunately, from a theoretical stand it is not at all clear what the

form of disparity errors should be. Further, it appears unlikely that any reasonably simple distribution will adequately capture stereo disparity errors. Second, it might be possible to derive values for D by bounding the likely disparity errors. However, this type of approach often leads to a result where even tight and conservative bounds on the input yield weak bounds in the computed error (e.g., Grimson [42]). Finally, the possibility exists that appropriate values for D could be derived by empirical study of the discontinuity recovery method in the face of natural stereo data. While this last approach does not yield a first principles solution, it might yield the most practical results.

2.2.4 Recapitulation

This section has focused on developing an understanding of how the proposed approach to the disparity interpretation problem can be expected to behave in the face of imperfect data. The discussion began by considering the possibility of degenerate sets of image measurements that would not allow the computations to be defined. It was concluded that for general stereo viewing, degenerate conditions are quite limited and easily diagnosed or avoided. The second set of developments considered the numerical stability of the proposed computations. For this purpose the "Generalized Error Propagation Formula" was used to understand the extent that error in the input disparity measurements will result in error in the recovered parameters of view and surface geometry. This analysis resulted in the intuitively pleasing result that stereo viewing conditions that make disparity salient will lead to good stability in the recovery methods. Significantly, these results indicate that an algorithm

for recovering scene geometry from stereo disparity can practically monitor its own performance. Finally, approaches to combining redundant data and thresholding for discontinuity recovery were presented. Redundant data is exploited via eigenvector analysis for viewing parameters and via histogramming for discontinuities. Thresholding for discontinuity detection is based on the nonparametric Kolmogorev-Smirnov method.

2.3 Computer implementation

This section describes the computer implementation and testing of an aspect of the proposed theory for disparity interpretation. In particular, the proposed method for recovering the discontinuities of planar surfaces has been embedded in a set of computer algorithms that have been applied to stereo disparity data. The discussion unfolds in three parts: First, the algorithms are described. Second, the results of applying the algorithms to synthetic stereo data are presented. The final section provides a brief recapitulation.

2.3.1 Description of algorithm and implementation

The proposed method for recovering the discontinuities of planar surfaces has been instantiated in a set of algorithms. In turn, these algorithms have been the subject of corresponding software implementations. The remainder of this subsection provides an overview of these developments. It will be seen that the algorithms and implementations proceed from the earlier developments of this chapter in a most

straightforward fashion.

At a coarse grain of analysis the algorithm for recovering surface discontinuities has the following three steps:

Algorithm for recovering surface discontinuities

1. Recover the local values of the geometric surface parameters p, q and r from an input disparity map.

2. Combine the recovered values of p, q and r into separate local histograms and smooth.

3. Compare adjacent histograms for each surface parameter with the Kolmogorev-Smirnov test. If the value of the test exceeds a specified value then assert a discontinuity in the region between the histograms.

This description of the algorithm clearly leaves quite a bit of detail to the imagination. In the ensuing discussion each step will be specified at a finer grain of analysis.

To begin, consider Step 1 of the proposed algorithm. Step 1 itself decomposes into three subparts: First, locally select three pairs of horizontal and orientational disparity measures to serve as input to the p, q and r recovery. A simple approach based on eight-connectivity serves to define these local groupings of disparity measures: Scan by rasters until a line segment is located in the left image; define the line's position by its top left pixel. Then search the eight-connectivity neighborhood pixels for another line segment. If no more segments are found scan the eight-connected neighbors of the last scanned set, and so on until the desired number of inputs, 3, are acquired.

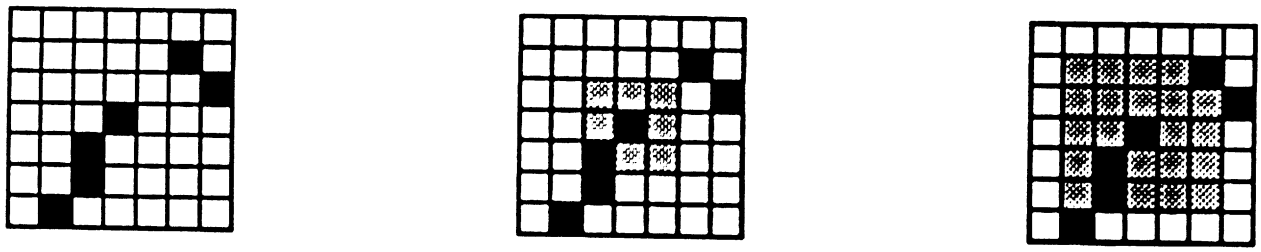


Figure 2.7: A simple algorithm for selecting local regions for input to the surface parameter and histogram computations: Expand eight-connectivity regions about a central pixel until the desired number of inputs have been scanned. In this example, it takes two iterations of the algorithm to locate the second line segment. For this figure, line segments are depicted with black; expanding serach regions are depicted with grey.

See Figure 2.7. Now, consider the second part of Step 1. In this part the three local orientational disparity measures are used as input to relation (2.18). This allows for the local recovery of σ and its axis of contraction ξ_1. Finally, part three of Step 1 is to use the recovered values of σ and ξ_1, along with local measures of horizontal disparity, χ_x, as input to relations (2.6) and (2.22). This specifies the local values of p, q and r.

Now turn attention to Step 2 of the algorithm for recovering surface discontinuities. The formation of the histograms is a relatively straightforward task. The selection of the measures serving as input to a local triad of histograms (one each for p, q and r) employs the same algorithm used to define local inputs to the surface parameter computation. Specifically, scan for the first p, q and r (located at the position of the first line element). Then, iterate the eight-connectivity scheme until n such points are located, where n is the histogram size. In forming subsequent histograms, points that have already been scanned are excluded. When all the local histograms have been formed, the disparity map will be divided into $\frac{m}{3n}$ regions, where m is the number of line elements in the disparity map. The values in the histogram are then smoothed by forming the serial products of the histogram buckets with the one-dimensional mask [0.5,1.0,0.5]. This smoothing operation helps to reduce discretization errors that are due to the noncontinuous nature of the histogram buckets.

Finally, consider Step 3 of the algorithm: Test for the significance of differing p, q and r between adjacent histogrammed regions. This step consists of applying the Kolmogorev-Snirnov test to the neighboring histograms. For each histogram: First, form a cumulative distribution. Second, compute the maximum difference, D, be-

tween neighboring cumulative distributions corresponding to each surface parameter. Third, if any value of D exceeds the specified level of significance then assert a discontinuity. For present purposes a discontinuity is asserted to lie in the region between the support areas of the neighboring histograms. Neighboring regions are again selected using the eight-connectivity approach. The specification of significance level and other parameters, being empirical matters, will be delayed until the next portion of this report: experiments with the algorithm.

This completes the description of the algorithm for recovering surface discontinuities from horizontal and orientational stereo disparity. The entire algorithm has been the subject of a software implementation in Zetalisp running on a Symbolics Lisp Machine. The result of applying this implementation to several disparity maps is documented in the next subsection of this report.

2.3.2 Experiments

The described algorithm and implementation have been applied to both synthetic and natural image disparity maps. The results of these experiments are the subject of this subsection. The first part of the discussion focuses on the results for the synthetic test cases; the second part presents the results for applying the method to a natural image stereo disparity map.

Experiments with synthetic stereo data

To begin this discussion, the particular parameters employed for the Kolmogorev-Smirnov test as applied to the synthetic images are now delineated: The histograms

n	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$
1	1.0000					
2	0.6667	1.0000				
3	0.4000	0.9000	1.0000			
4	0.2286	0.7714	0.9714	1.0000		
5	0.1270	0.6428	0.9206	0.9921	1.0000	
6	0.0693	0.5260	0.8571	0.9740	0.9978	1.0000

Table 2.1: Probability that $D \leq \frac{h}{n}$ for the Kolmogorev-Smirnov statistic.

are formed to be of size $n = 5$. The histogram bucket width is ± 1 pixel about the fixation for r and recovered values of 0.3 for p and q. For intuition, notice e.g., that $p = 0$ is a fronto-parallel plane while $p = 1$ is a plane rotated 45° about the vertical. The performance of the algorithm is demonstrated on all the displays with a significance level of $D = .6428$. To examine the effect of increasing the significance level, one display is also tested with $D = .9206$. For reference, the relevant portion of a table giving values of D for small n is given in Table 2.1.

The implementation with these parameter settings has been tested on five synthetic stereo disparity maps. Each disparity map corresponds to a random line stereogram created by randomly tossing 700 lines of dimension 20×1 pixels on one of a pair of 512×512 pixel arrays. The second array is generated from the first by differentially projecting the line segments to correspond to a simple three-dimensional scene. See, e.g., the top half of Figure 2.7. The horizontal and orientational disparity

corresponding to each line is made available to the algorithm. The position of the line segment is indexed on the basis of its upper left hand corner.

The particular "scenes" that served as input to the algorithm are depicted via random line stereograms in the top halves of Figures 2.8-2.12. A description of each scene is provided in the caption to the figure. The scenes were selected to illustrate the algorithms performance in a number of interestingly different situations. In particular, different magnitudes and direction of surface gradient and distance are illustrated in each case. Also, the performance of the algorithm when applied to the same data but with differing values of D is demonstrated in Figure 2.10. The bottom half of each figure shows the regions of discontinuity that were recovered by the algorithm in each case. The relative distance of the forward most surface along the discontinuity is coded in terms of grey level. The coding has areas of increasing intensity as closer to the viewer.

Several observations can be made about the performance of the algorithm:

- In general, it is clear that the algorithm performs well in all of the tested cases.
- The algorithm is not at loss when operating in the face of only p or only q differences, Figures 2.8 and 2.9.
- The algorithm is capable of recovering surface discontinuities in the face of both large and small surface gradients. Figures 2.10 and 2.11.
- The performance of the algorithm on these noise free examples is not dependent on the value of the significance level D, Figure 2.10.

- When faced with a slightly more complex scene, the algorithm still performs adequately, Figure 2.12.

Experiment with natural stereo data

The algorithm and implementation for recovering the discontinuities of planar surfaces has also been applied to a natural image disparity map: The top half of Figure 2.13 shows a pair of stereoscopic aerial photos of the University of British Columbia. A disparity map corresponding to these photographs was provided by Eric Grimson of the MIT AI lab. The bottom left panel of Figure 2.13 shows the linear segments that were used to derive the horizontal and orientational disparity for input to the surface discontinuity recovery process. This particular disparity map is a difficult case for the discontinuity algorithm for several reasons: First, the density of texture in the data is low. This challenges the algorithm as it requires several linear segments for each local area where it performs the discontinuity computation. Second, the overall disparity range present in the disparity map is rather compressed. This naturally leads to a poor signal to noise ratio in the data that is used to drive the algorithm. Finally, it is important to note that a simple threshold based on absolute disparity would fail to find many of the interesting discontinuities for this test case; the data does not simply correspond to a set of fronto-parallel planes arranged at various distances from the viewer. Examination of the disparity data shows that there is strong gradient of disparity across the map (roughly from the lower left corner to the upper right corner). (This gradient corresponds to the fact that the buildings in this stereo pair are actually situated on a hill.) Thus, any attempts to set an overall

disparity threshold would lead to either a high miss rate in regions where the average disparity has small magnitude or a high false alarm rate in regions where the average disparity has a large magnitude.

The parameters used in applying the discontinuity algorithm to the natural image test case were identical to those used for the synthetic image test case, with two exceptions: In order to cope with the low texture density and the poor signal to noise ration, the number of local regions used to form the histograms was reduced from 5 to 3 and the significance level was set at $D = .90$. The result of applying the algorithm to the test data is displayed in the lower right panel of Figure 2.13. As earlier, the recovered depth along the discontinuities is rendered in terms of grey levels with lighter regions corresponding to points that are nearer the observer. Inspection of these results allows for several observations:

- The important discontinuities in distance that are present in the disparity data are recoverd by the algorithm. Further, there are few false alarms signaled by the algorithm.

- Certain regions of discontinuity that e.g., a human might infer from the stereo pair are not recovered by the algorithm. However, the data used to drive the algorithm does not derive directly from the stereo pair, but rather from the disparity associated with the edge image shown in the lower left panel. In this regard, it is seen that "missing" regions of discontinuity (e.g., parts of the outline of the central building) correspond to regions of the disparity map that provide little or no information input to the algorithm. Thus, the limitation does not

lie with the algorithm; rather, the limitation is in the data made available to the algorithm.

- Certain of the recovered discontinuities do not neatly outline a contour of discontinuity. For example, the tops of some of the buildings are not outlined, but rather are signaled by a single patch of discontinuity. This result can be accounted for by the lack in the texture density at the corresponding regions of the disparity map. There are not enough disparity measures available to drive the several separate patch comparisions that would be needed to outline the entire contour. Again, the limitation is not in the algorithm per se, but rather in the low density of information that has been provided to the algorithm.

- The recovery of relative depth along the discontinuities (as coded by grey level) is in good accord with the corresponding percept resulting from stereoscopic viewing of the stereo pair.

2.3.3 Recapitulation

This section has documented the results of computer based experiments with an aspect of the proposed theory for disparity interpretation. The particular aspect of the theory that has been studied is the method for recovering the discontinuities of planar surfaces. The first part of the section described the details of instantiating the method in an algorithm and corresponding implementation. The latter part of the section described a series of experiments with the algorithm and its implementation. The experiments centered around the performance of the algorithm in the face of

synthetic disparity maps. The results of the experiments are positive; the algorithm performs well in recovering the surface discontinuities for all the presented cases.

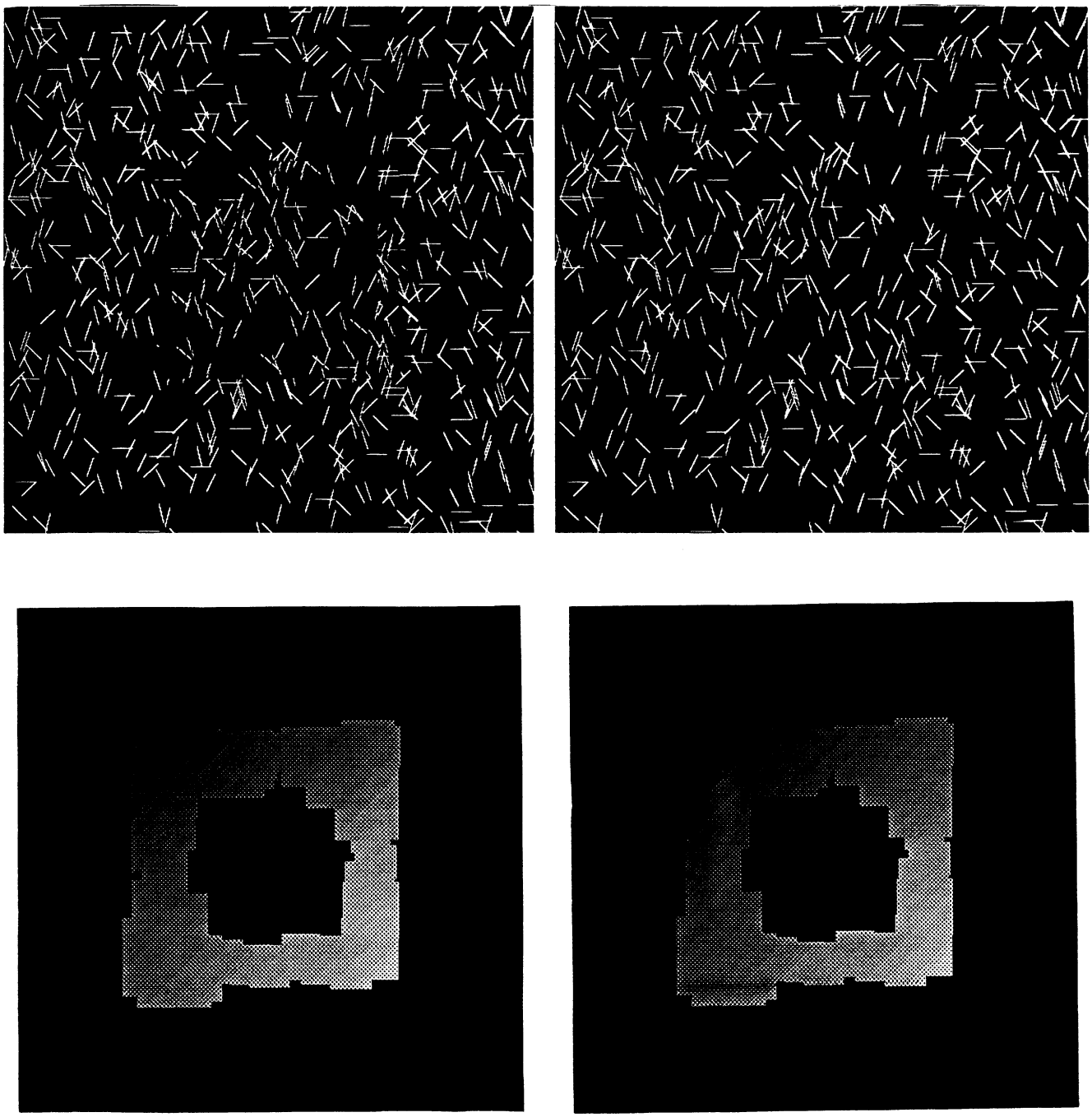


Figure 2.10: Results of a computer experiment with the proposed method for recovering the discontinuities of planar surfaces. The top half of the figure displays a random line stereogram corresponding to the input disparity information. The disparity corresponds to a central planar region that has been rotated about an oblique axis by $> 45°$. The bottom half of the figure shows the recovered regions of discontinuity. The bottom left and right panels show results for $D = .6428$ and .9206, respectively. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

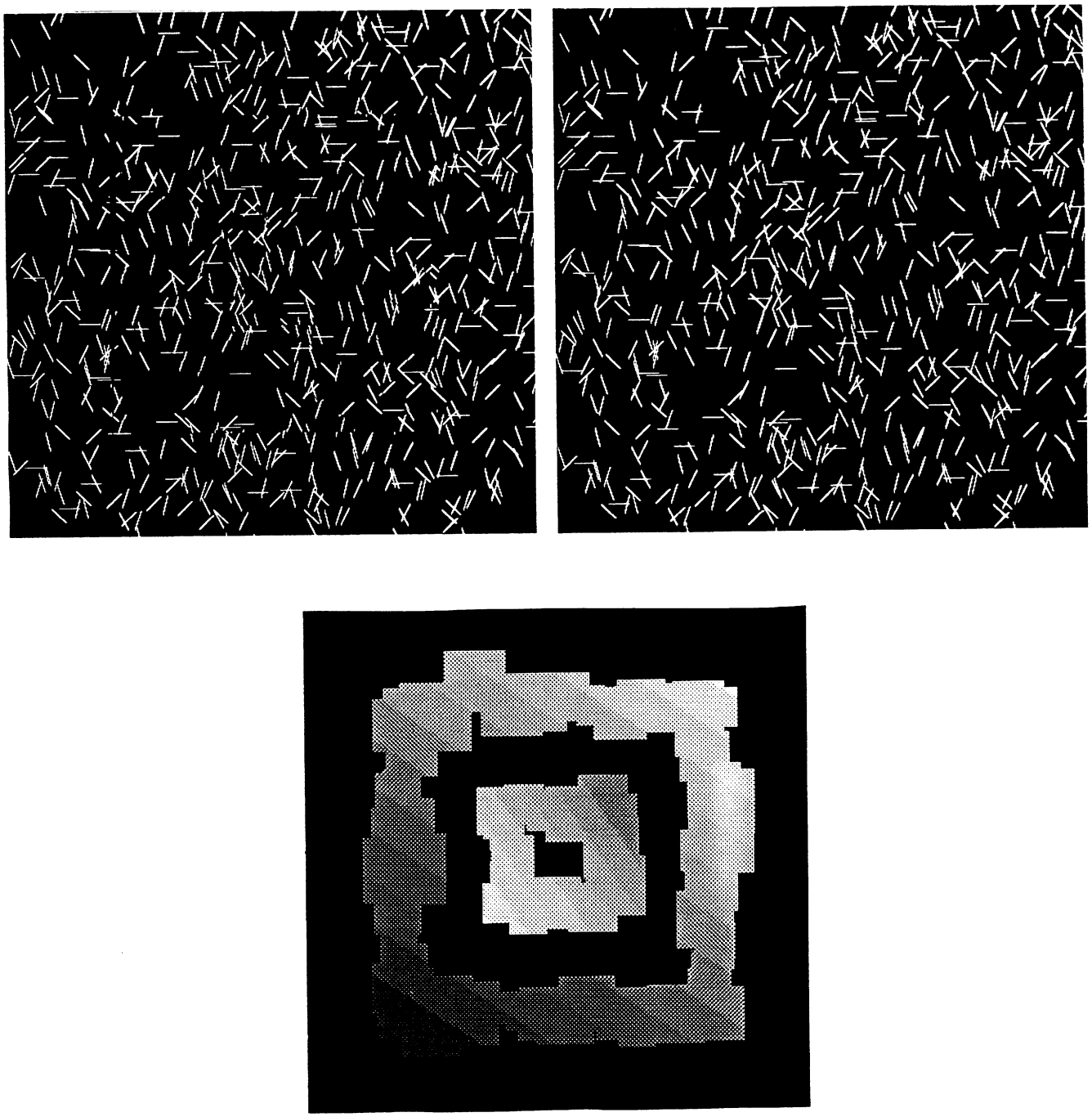


Figure 2.12: Results of a computer experiment with the proposed method for recovering the discontinuities of planar surfaces. The top half of the figure displays a random line stereogram corresponding to the input disparity information. The disparity corresponds to a pair of concentric planar regions that have been differentially rotated about oblique axes. The bottom half of the figure shows the recovered regions of discontinuity. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

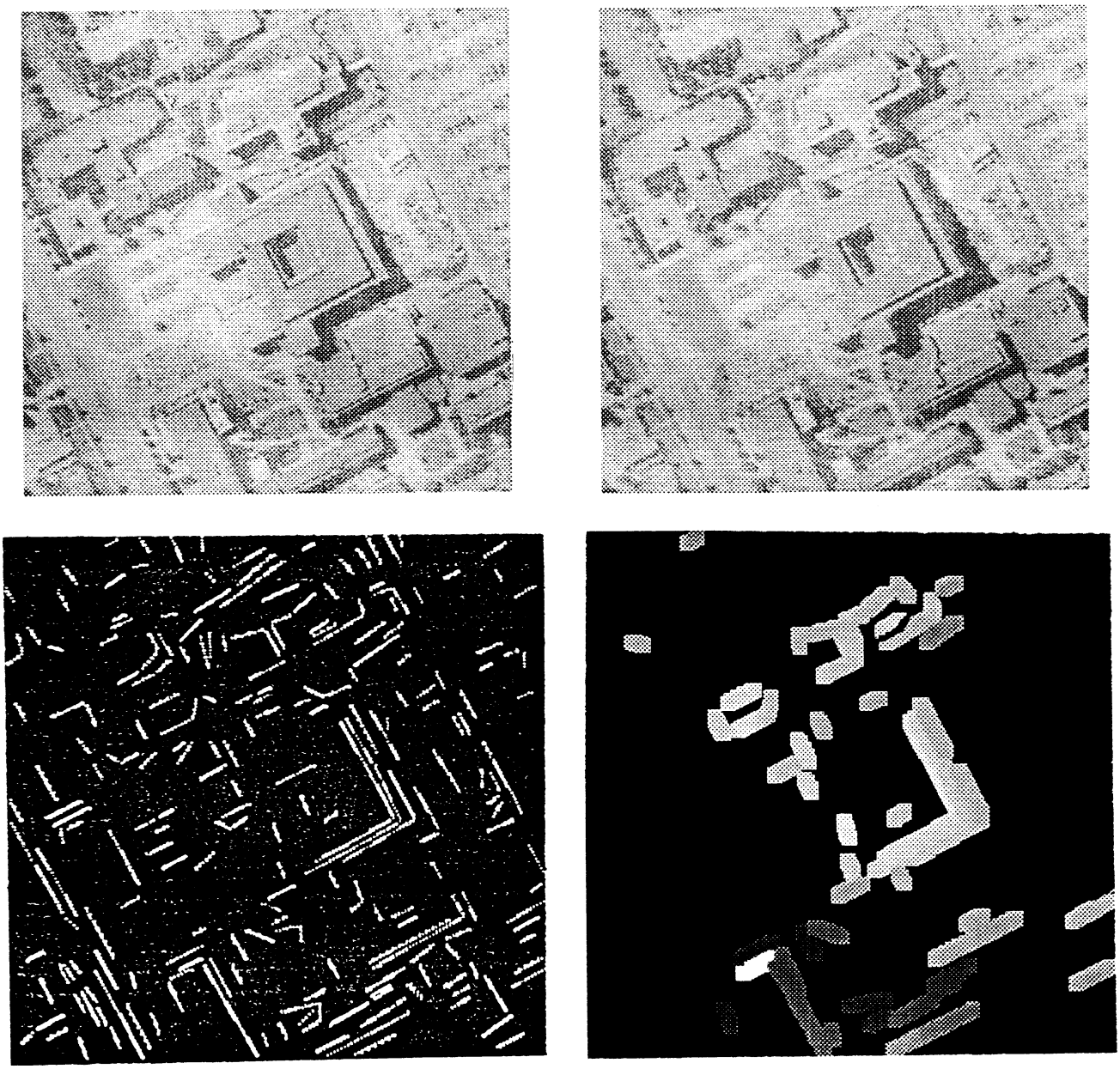


Figure 2.13: Results of a computer experiment with the proposed method for recovering the discontinuities of planar surfaces. The top half of the figure displays a natural image stereo pair. The bottom left panel shows the corresponding linear segments from which the horizontal and orientational disparity information was derived. The bottom right panel shows the recovered regions of discontinuity. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

Chapter 3

Curved surfaces

This chapter presents an extension of the results presented in Chapter 2. Specifically, the results for recovering surface discontinuities from stereo disparity are extended toward dealing with discontinuous curved surfaces. The organization is as follows: The first section presents the basic theoretical extensions. The second section studies the numerical stability of the relations defined in Section 1. Section 3 describes a set of computer algorithms for recovering the discontinuities of curved surfaces described by disparity. These algorithms are then applied to several disparity maps.

3.1 Analysis of disparity

In this section an approach is developed for recovering the discontinuities of curved surfaces from stereo disparity. The approach builds in a straightforward fashion on the results obtained in the previous chapter for planar surfaces. The proposal is based on approximating a curved surface by a collection of planar patches. In this

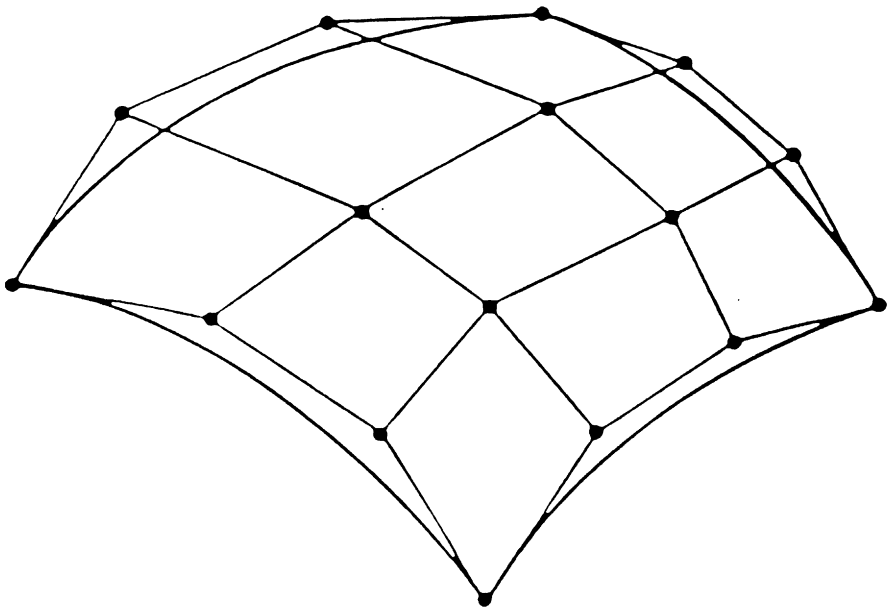


Figure 3.1: A curved surface can be approximated by a collection of planar patches. Neighboring patches intersect along dihedral edges.

representation, neighboring planar patches will intersect along an edge in space. More specifically, the neighboring patches will come together along a dihedral edge. See Figure 3.1. The remainder of this section discusses and expands upon these notions. The developments unfold in two parts: The first part defines the dihedral edge where adjacent surface patches meet. The definition is given in terms of the geometric parameters used to define the surface patches. Following this definition, an analysis of how the dihedral edge projects into an image will be presented. The second part of this section turns this analysis around to show how conditions on surface discontinuities can be recast as conditions on dihedral edges and their differential projections.

3.1.1 Recovering surface discontinuities

Suppose that a curved surface is approximated by a collection of planar patches as illustrated in Figure 3.1. Let the distance (e.g., with respect to an observer) along

two neighboring patches be denoted as Z_1 and Z_2. Using the standard representation of a planar surface patch, these two distance values can be expanded as

$$Z_1 = p_1 X + q_1 Y + r_1$$

and

$$Z_2 = p_2 X + q_2 Y + r_2,$$

where r_i is the radial distance to patch i while (p_i, q_i) are the corresponding surface gradient terms. This representation of a planar surface patch is the same as that presented in Chapter 2, equation (2.9). Now, for a continuous surface these adjacent patches meet along a dihedral edge. Along this edge the distance values of the two neighboring patches are clearly the same. That is to say

$$Z_1 = Z_2.$$

This relation can be expanded in terms of the proposed representation for planar patches to yield

$$p_1 X + q_1 Y + r_1 = p_2 X + q_2 Y + r_2. \tag{3.1}$$

Equation (3.1) defines a line in three-space where the planes embedding the patches meet. The definition of the line has followed from the structure of the chosen representation for curved surfaces.

The ultimate goal of the current developments is to effect the recovery of three-dimensional surface information from stereoscopically imaged information. Therefore, attention is now turned to the imaging properties of approximating planar patches as well as their dihedral edges of intersection defined in (3.1). Recall that image

coordinates are related to world coordinates by $(x, y) = (\frac{X}{Z}, \frac{Y}{Z})$ and appropriate units. Then (3.1) can be rewritten in image coordinates as

$$\frac{1 - p_1 x - q_1 y}{r_1} = \frac{1 - p_2 x - q_2 y}{r_2}. \tag{3.2}$$

While (3.2) defines a line in the image plane, it is somewhat difficult to interpret due to its nonstandard form. This line can be recast into a more useful form by simple algebraic manipulation. By setting $y = 0$ it is seen that the x-intercept of the line is

$$\frac{r_2 - r_1}{r_2 p_1 - r_1 p_2}.$$

Similarly, setting $x = 0$ shows that the y-intercept of the line can be written as

$$\frac{r_2 - r_1}{r_2 q_1 - r_1 q_2}.$$

Further algebraic manipulation shows that the equation of the line can be written in the standard form

$$ax + by + c = 0 \tag{3.3}$$

where $a = r_2 p_1 - r_1 p_2$, $b = r_1 q_2 - r_2 q_1$ and $c = r_1 - r_2$. This parameterization is related to the familiar normal form of a line, $x \cos\theta + y \sin\theta - \rho = 0$, where ρ is the length of the directed perpendicular from the origin to the line and θ is the counterclockwise angle between this perpendicular and the positve x-axis, see Figure 3.2. In terms of (3.3), $\rho = \frac{\|c\|}{(a^2+b^2)^{\frac{1}{2}}}$ while $(\cos\theta, \sin\theta) = \frac{(a,b)}{(a^2+b^2)^{\frac{1}{2}}}$.

It is worth emphasizing the interpretation of equation (3.3): Suppose a continuous curved surface is approximated by a collection of planar patches. Then (3.3) describes the image projection of the dihedral edge where a pair of adjacent approximating patches meet.

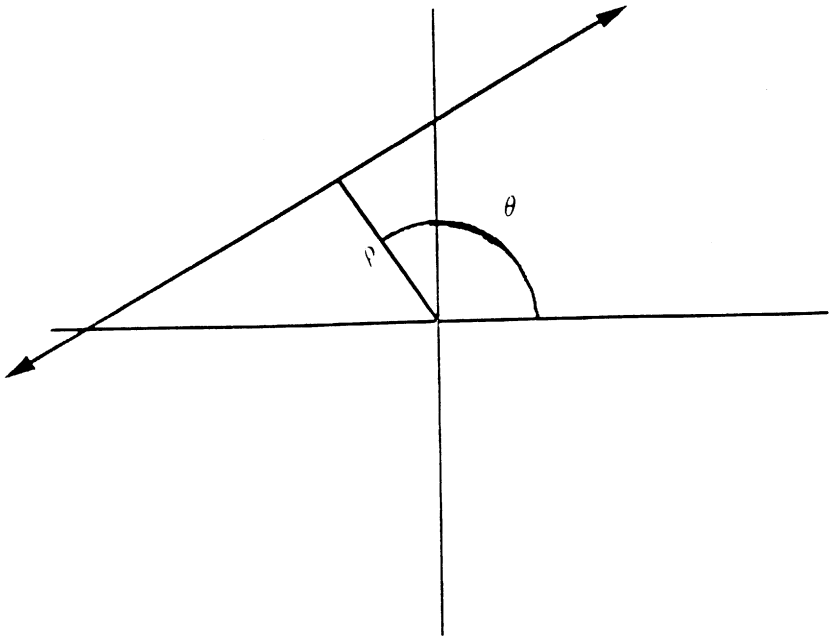


Figure 3.2: A straight line can be described by the equation $x\cos\theta + y\sin\theta - \rho = 0$. This representation avoids the singularities of the more common slope-intercept forms for a line.

The foregoing analysis has been concerned with the piecewise planar approximation to continuous curved surfaces. However, the results of this analysis can be exploited in developing an approach to recovering the discontinuities of curved surfaces from stereo disparity. Just as a continuous surface will have its approximating planar patches meet in a dihedral edge, a discontinuous surface can be taken to be implicated by adjacent patches that fail to meet along such an edge. Notice however, that the planes embedding the patches will intersect *somewhere*. Therefore, two patches failing to satisfy the image (3.3) cannot be the entire constraint on recovering discontinuities. The added observation that does allow for the recovery of discontinuities is simple enough: The projected line of intersection (3.3) must project between the projections of the two patches that defined it. If the projection is not between the patches, the surface is discontinuous.

If a projected line of intersection, (3.3), fails to fall between the two patches that defined it then the corresponding surface must be discontinuous. However, is it true that if the line projects between the patches then the surface must be continuous? Unfortunately, the answer to this question is no; there are possible "miss" situations. One potential difficulty arises if the local patches do not extend fully into the region between the patches. This situation is not too serious because without evidence to the contrary (e.g., a luminance edge) it is reasonable to enforce an extrapolation into the region between the patches (c.f., Grimson [40]). It is also possible that certain configurations of surface patches and viewer can define a line that would project between the surface patches, even in the discontinuous case (more study of this situation is needed). The possibilty of these "miss" situations is an acknowledged limitation of the proposed approach to recovering surface discontinuities.

The proposed method for recovering the discontinuities of curved surfaces can now be stated as follows: First, recover the approximating planar surface parameters (p, q, r) of adjacent patches of the disparity map via the methods of chapter 2. Second, combine the surface parameters into an equation of the form (3.3). Third, check to see if the line that has been defined passes between the adjacent regions of the disparity map. If not, assert a discontinuity as lying in that region. It is clear that this approach builds directly on the method described in Chapter 2 for planar surfaces.

3.2 Stability analysis

This section focuses attention on the numerical properties of the proposed scheme for recovering the discontinuities of curved surfaces. The developments parallel those in the corresponding section of chapter 2. Specifically, attention shall be given to degenerate sets of measurements, sensitivity to noise corrupted data and a method for operating in the face of noisy data.

3.2.1 Degeneracies

Consider those situations where the proposed method for recovering surface discontinuites is undefined. To begin, recall that the method for curved surfaces begins by recovering the local planar parameters of local depth, r, and surface gradient, (p, q), via the method of Chapter 2. Therefore, the method for curved surfaces inherits the degeneracies of the method for planar surfaces. Specifically, (as noted in Section 2.2.1) there are two conditions of practical concern: (i) Disparities that are measured along the left image line $x = 0$ will lead to an undefined solution to the recovery of the differential viewing parameters. (ii) If the corresponding planar approximation to the surface passes through the optical node of the left view (the origin of the cyclopean coordinate system) then it is not possible to recover the surface gradient. Notice that in this condition the plane under consideration degenerates to a line as seen from the left vantage point.

Next, consider if there are any new degenerate conditions introduced by the extension to curved surfaces. This matter can be considered by turning attention to the

constraint equation for curved surfaces (3.3). Three conditions of interest arise: (i) If $r_1 = r_2 = 0$ then the left side will be identically zero irrespective of the surface gradients. However, this degeneracy does not add anything new because the constraint that solutions require $r_i \neq 0$ was also present for planar surfaces. (ii) If the magnitude of both surface gradients is zero, (3.3) does not apply. This degeneracy corresponds to the fact that the potential line of intersection for the two such planar patches will not project into the image plane (i.e., the line is at projective infinity). Practically, this condition can be handled by not checking equation (3.3) if the recovered surface gradients are all zero. Instead, a simple comparison of the values of r will serve to diagnose discontinuities. (iii) If $(p_1, q_1, r_1) = (p_2, q_2, r_2)$ then the left hand side of (3.3) is again identically zero. This makes intuitive sense as all lines in the plane will indeed be lines of intersection of a single plane. Should this condition arise, recourse to the method for planar surfaces is appropriate.

In summary, consideration of degenerate conditions in recovering the discontinuites of curved surfaces has led to positive conclusions:

- No degeneracies of practical concern have been discovered that were not already known for the planar solution.

- While certain new degenerate conditions were noted, they are not of practical importance as simple diagnostics and solutions for these situations have been presented.

3.2.2 Error analysis

Attention is now focused on studying the stability of the discontinuity recovery method in the face of noise perturbed data. Following the approach to studying numerical stability that was employed in Chapter 2, the "generalized error-propagation formula" (2.29) is again used as a tool in understanding how perturbations in the input to the discontinuity recovery method effect the corresponding output. For current concerns: The inputs are two sets of planar surface parameters (p_1, q_1, r_1) and (p_2, q_2, r_2); the output is the image projection of the dihedral edge where the local planar approximations to the surface meet. Insofar as this method relies on parameters (p_i, q_i, r_i) that are derived from disparity measures, its stability will in turn rest on the stable recovery of these parameters from disparity. The stable recovery of planar surface parameters has been addressed elsewhere in this thesis and will not be reconsidered here. Rather, the present investigation will concentrate on how errors in the recovered parameters (p_i, q_i, r_i) lead to errors in the discontinuity constraint equation (3.3).

The discontinuity constraint equation (3.3) is defined by the three parameters a, b and c. Therefore, in order to understand how errors in the derived parameters p, q and r effect the accuracy of equation (3.3) the generalized error-propagation formula (2.29) is now employed in the following fashion: Let the errors in p_i, q_i and r_i be Δp_i, Δq_i and Δr_i, respectively. Then, with regard to (2.29), let $\mathbf{x} = (p_1, q_1, r_1, p_2, q_2, r_2)$. Applying the generalized error-propagation formula to a, b and c then yields

$$\|\frac{\partial a}{\partial p_1}\| \cdot \|\Delta p_1\| + \|\frac{\partial a}{\partial q_1}\| \cdot \|\Delta q_1\| + \|\frac{\partial a}{\partial r_1}\| \cdot \|\Delta r_1\| + \|\frac{\partial a}{\partial p_2}\| \cdot \|\Delta p_2\| + \|\frac{\partial a}{\partial q_2}\| \cdot \|\Delta q_2\|$$

$$+\|\frac{\partial a}{\partial r_2}\| \cdot \|\Delta r_2\|,$$

$$\|\frac{\partial b}{\partial p_1}\| \cdot \|\Delta p_1\| + \|\frac{\partial b}{\partial q_1}\| \cdot \|\Delta q_1\| + \|\frac{\partial b}{\partial r_1}\| \cdot \|\Delta r_1\| + \|\frac{\partial b}{\partial p_2}\| \cdot \|\Delta p_2\| + \|\frac{\partial b}{\partial q_2}\| \cdot \|\Delta q_2\|$$

$$+\|\frac{\partial b}{\partial r_2}\| \cdot \|\Delta r_2\|$$

and

$$\|\frac{\partial c}{\partial p_1}\| \cdot \|\Delta p_1\| + \|\frac{\partial c}{\partial q_1}\| \cdot \|\Delta q_1\| + \|\frac{\partial c}{\partial r_1}\| \cdot \|\Delta r_1\| + \|\frac{\partial c}{\partial p_2}\| \cdot \|\Delta p_2\| + \|\frac{\partial c}{\partial q_2}\| \cdot \|\Delta q_2\|$$

$$+\|\frac{\partial c}{\partial r_2}\| \cdot \|\Delta r_2\|,$$

respectively. The correpsonding evaluation of these forms leads to

$$\|p_2\| \cdot \|\Delta r_1\| + \|p_1\| \cdot \|\Delta r_2\| + \|r_2\| \cdot \|\Delta p_1\| + \|r_1\| \cdot \|\Delta p_2\|, \tag{3.4}$$

$$\|q_2\| \cdot \|\Delta r_1\| + \|q_1\| \cdot \|\Delta r_2\| + \|r_2\| \cdot \|\Delta q_1\| + \|r_1\| \cdot \|\Delta q_2\| \tag{3.5}$$

and

$$\|\Delta r_1\| + \|\Delta r_2\|. \tag{3.6}$$

as the results for the parameters a, b and c, respectively. Now, recall that small magnitudes in the generalized error-propagation indicate stable solutions. For present concerns it is important to understand how to attain small magnitudes in (3.4), (3.5) and (3.6). Clearly, these equations have small magnitude when the magnitude of the errors Δp_i Δq_i and Δr_i are small. Further, the magnitudes of (3.4) and (3.5) are dependent on the uncorrupted planar parameters, p_i, q_i and r_i. In particular, relatively small absolute values of the planar parameters lead to small magnitudes in (3.4) and (3.5). In other words, recovery of the parameters a, b and c is most stable when the viewed surface has surface gradient of small magnitude and is not

too distant from the viewer. Correspondingly, these are the conditions that allow for the robust recovery of the discontinuity constraint equation (3.3).

This investigation of the numerical stability of the discontinuity constraint equation suggest that the proposed method for recovering discontinuities can be of practical use. However, in evaluating the claims for stability it is important to keep in mind that the robust recovery of the planar surface parameters from disparity measures must be assured first.

3.2.3 Operating in the face of perturbed data

In practice there will always be some amount of error in the recovered values of the surface parameters that are used to define equation (3.3). Therefore, it is important to propose a method for combining local measures in a fashion that will allow the effects of these errors to be minimized. The most popular approach to this type of problem is to use a least-squares based approach. However, this approach is not overly appropriate for the matter at hand. As pointed out in Chapter 2, the manner that errors in the input disparity interact in the recovered surface parameters makes application of the least-squares approach rather dubious. As an alternative a simple histogramming approach is used. Specifically, local histograms are separately compiled for values of p, q and r as recovered via the methods of Chapter 2. The peaks of the smoothed histograms are then used for subsequent computation. For the matter at hand, peak values of the surface parameters from adjacent regions of the disparity map are used to specify equation (3.3). As previously discussed, if the line so defined passes between the corresponding regions, the surface is taken as locally continuous.

If the line passes outside this region, the surface is taken as locally discontinuous. This constitutes the proposed method for combining local measures in order to offset the effects of inaccuracies in the recovered surface parameters.[1]

3.2.4 Recapitulation

This section has considered the stability of the proposed method for recovering the discontinuities of curved surfaces. With regard to degeneracies of the solution method, it was concluded that little is different from the case of purely planar surfaces. Consideration of robustness in the face of noise perturbed data led to positive findings provided two types of conditions are met: First, all the considerations previously outlined for planar surfaces must be met. Second, it is best if the magnitude of the locally planar surface parameters are relatively small. Finally, a simple histogramming method has been outlined for combating noise corrupted data.

3.3 Computer implementation

This section describes the results of embedding the proposed method for recovering the discontinuities of curved surfaces in computer algorithms and implementations. The discussion unfolds in three parts: First, the algorithms are described. Second, the results of applying the algorithms to both synthetic and natural image stereo data are presented. The final section provides a brief recapitulation.

[1] In practice it may be neccesary to allow surfaces to be regarded as continuous if the line defined by (3.3) simply passes "near-by" the region under consideration. Currently, a good analysis of what constitutes near-by is lacking in the proposed method.

3.3.1 Description of algorithm and implementation

The proposed method for recovering the discontinuities of curved surfaces has been instantiated as a set of algorithms. In turn, these algorithms have been the subject of corresponding software implementations. The remainder of this subsection provides an overview of these developments. It will be seen that the algorithms and implementations follow from the earlier developments of this chapter in a fairly straightforward fashion.

The algorithm for recovering the discontinuties of curved surfaces can be seen to have the following four basic steps:

Algorithm for recovering surface discontinuities

1. Recover local values of the first-order surface parameters p, q and r from an input disparity map.

2. Combine the recovered values of p, q and r into separate local histograms and smooth.

3. Use the peaks of adjacent histograms to define the line of intersection of the corresponding first-order surface fits.

4. Check to see if the line of intersection projects into the region between the two local histograms. If not, assert a discontinuity in the region between the histograms.

To be of use this algorithm must be specified more precisely. The following discussion presents the required details.

To begin, notice that Steps 1 and 2 of this algorithm correspond exactly to Steps 1 and 2 of the algorithm presented in Section 2.3. The details are also the same and will not be repeated here. Now, consider Step 3 of the current algorithm. With the values of the surface parameters in hand this is simply a matter of plugging in the values to (3.3). The ensuing computation yields the parameters describing the desired line of intersection. Finally, consider Step 4 of the algorithm. The crucial point of this step is to have a method for determining if a line passes between two regions in a plane. In solving this problem it is useful to exploit the geometric structure of the histogram support regions. Recall that due to the algorithm used to search for members of a histogram, the support region will be rectangular (square in fact). Interestingly, a line passing between two rectangular regions must intersect both segments joining nearest corners between the regions. To allow for slack in the exact location of the intervening line, the endpoints of the segments can move away from the nearest corners along the edges of the regions. See Figure 3.3. The algorithm can be instantiated numerically as the solution to two sets of constrained linear systems. Each system solves for the potential intersection of the line defined by equation (3.3) and one of the segments e or f shown in Figure 3.3. If there is no solution then a discontinuity is asserted in the region of the disparity map between the support of the two histograms.

This completes the description of the algorithm for recovering surface discontinuities from stereo disparity. The complete algorithm has been the subject of a software implementation in Zetalisp running on a Symbolics Lisp Machine. The result of applying this implementation to several disparity maps is outlined in the following subsection.

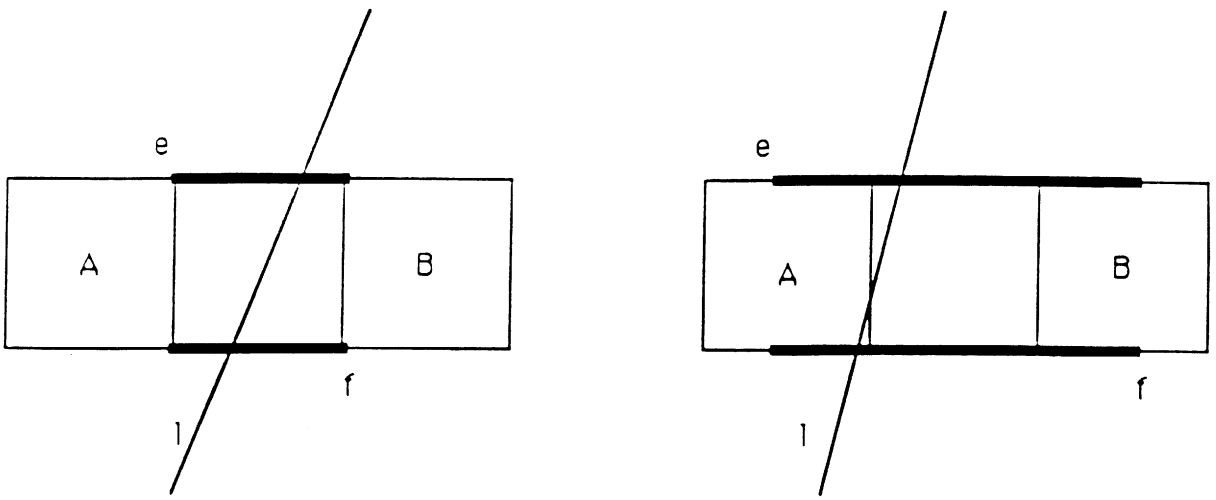


Figure 3.3: An approach to determining if a line passes between two regions. (a) A line passing between the gap that separates region A and region B must intersect both line segments e and f. (b) To allow for inaccuracies the endpoints of the segments can be allowed to extend along the edges of the regions.

3.3.2 Experiments

The described algorithm and implementation have been applied to several synthetic disparity maps.[2] The results of these experiments provide the focus for this subsection. To begin this discussion, the parameter values of the implementation are now noted. For the synthetic disparity data the histograms are formed with the same parameters used in Chapter 2. Specifically, the histogram size is $n = 5$ and the bucket widths are 0.3 for p and q and ± 1 pixel about the fixation for r. A liberal threshold is adopted for defining the region where a line of intersection can fall and still be indicative of a continuous surface. Specifically, the line can pass anywhere between or within the support regions of the histograms.

The implementation with these parameter settings has been applied to a set of four synthetic disparity maps. As in experiments described in Chapter 2, the disparity maps correspond to random line stereograms. The stereograms are defined over 512×512 pixel arrays. The disparate elements are 700 randomly distributed lines with dimensions 20×1 pixels. Both the horizontal and orientational disparity from each line serves as input to the algorithm.

[2]Preliminary attempts to test the algorithm and implementation in the face of natural image stereo data have met with limited success. Unfortunately, these studies have been hampered by the limited availability of test data that are appropriate for the curved surface algorithm. In particular, only the data for the aerial view of the UBC campus (shown in Figure 2.13) has been available for testing. This is a poor test case for three reasons: First, the surfaces in the image are largely planar, while the algorithm is based on curved surfaces. Second, the density of the texture data needed to drive the algorithm is quite low. Third, the disparity range of the UBC disparity map is rather compressed. This leads to difficulties in the signal to noise ratio.

The particular synthetic scenes used to test the algorithm correspond to four basic types of curved surface patches: planar, cylindrical, elliptic and hyperbolic. These surfaces are rendered as random line stereograms in the top halves of Figures 3.4-3.7, respectively. The bottom half of each figure shows the regions of discontinuity that were recovered by the algorithm. The relative depth of the forward most part of the surface along the discontinuity is coded in terms of grey level intensity. Areas of higher intensity correspond to regions that are closer to the viewer.

Several observations can be made concerning the results of these experiments with synthetic stereo data.

- In general, the algorithm performs well on all the synthetic examples.
- The differences in the type of curved surface (e.g., hyperbolic vs. elliptic) make little difference in the performance of the algorithm.
- Even in the face of a surface with zero curvature, the algorithm recovers the regions of discontinuity, Figure 3.4.

3.3.3 Recapitulation

This section has documented the results of computer based experiments with the proposed approach to recovering the discontinuites of curved surfaces. The first part of the section specified the details of instantiating the method in an algorithm and a corresponding implementation. The latter part of the section described a series of experiments with the algorithm and its implementation. The experiments tested the algorithm's performance in the face of synthetic stereo data. The results of the

experiments are positive; the algorithm performs adequately in recovering the surface discontinuities for all the presented cases.

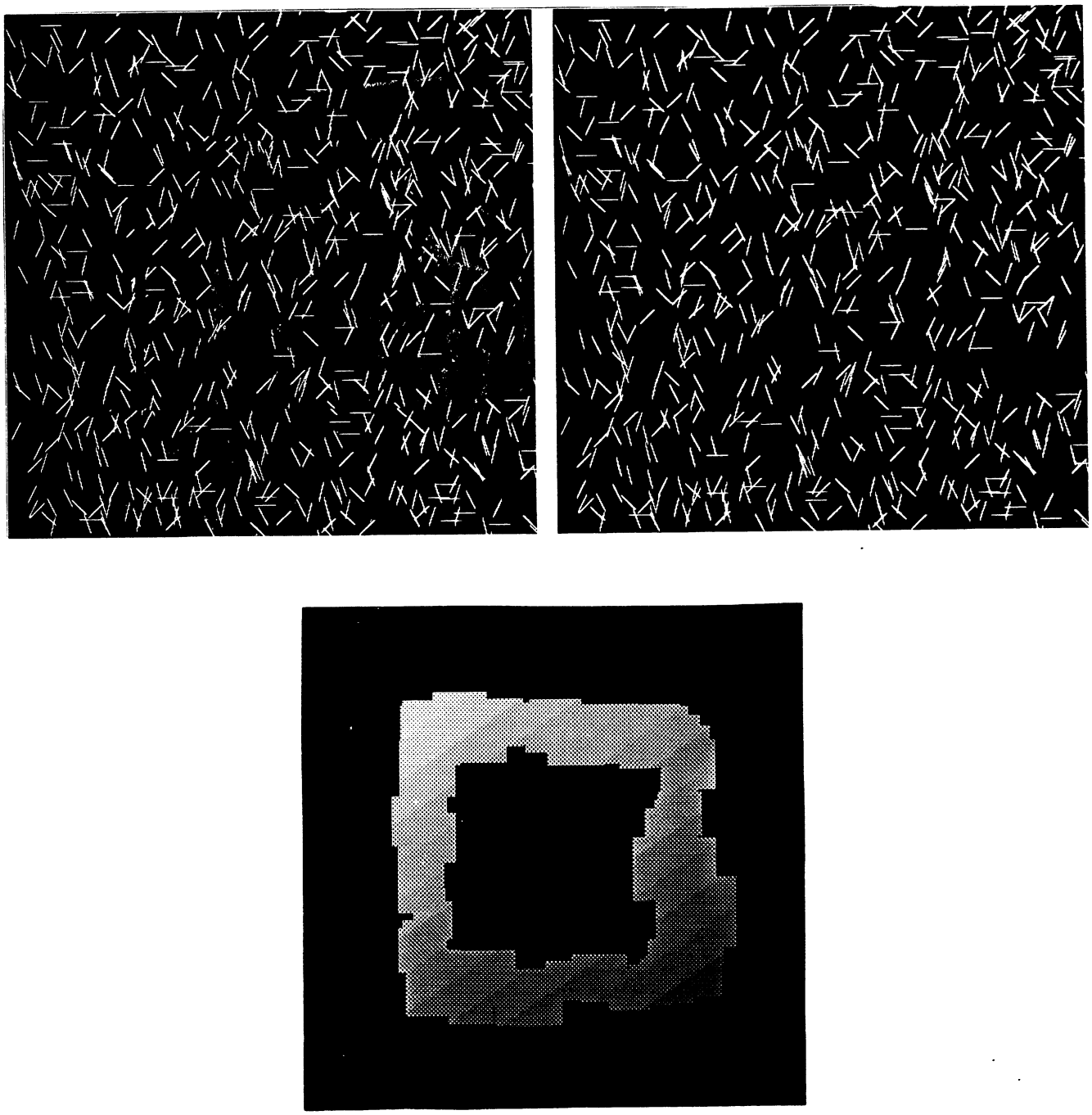


Figure 3.4: Results of a computer experiment with the proposed method for recovering the discontinuities of curved surfaces. The top half of the figure displays a random line stereogram corresponding to the input disparity information. The disparity corresponds to a central planar region that has been rotated about an oblique axis by $< 45°$. The bottom half of the figure shows the recovered regions of discontinuity. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

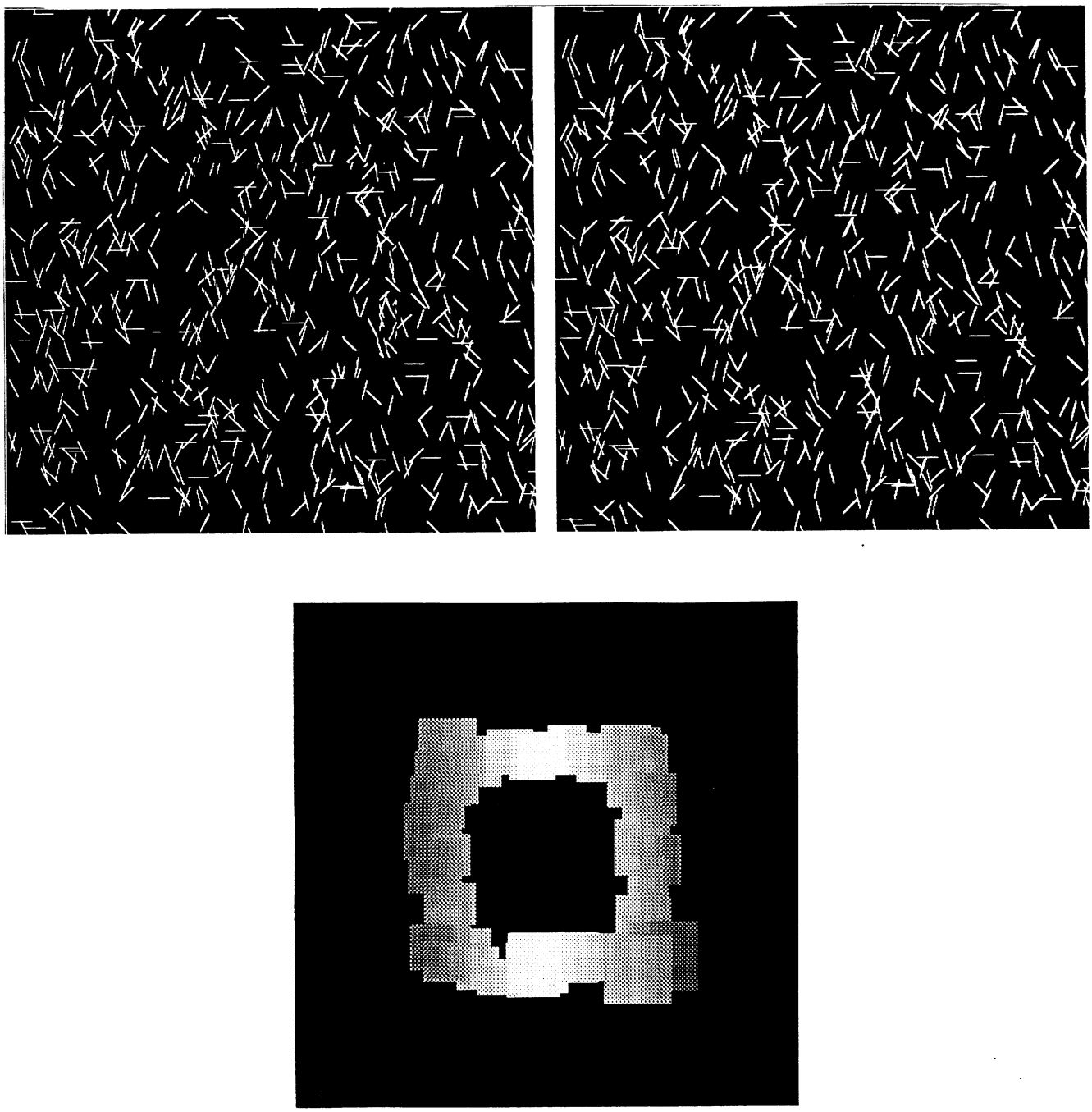


Figure 3.5: Results of a computer experiment with the proposed method for recovering the discontinuities of curved surfaces. The top half of the figure displays a random line stereogram corresponding to the input disparity information. The disparity corresponds to a central cylindrical region. The bottom half of the figure shows the recovered regions of discontinuity. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

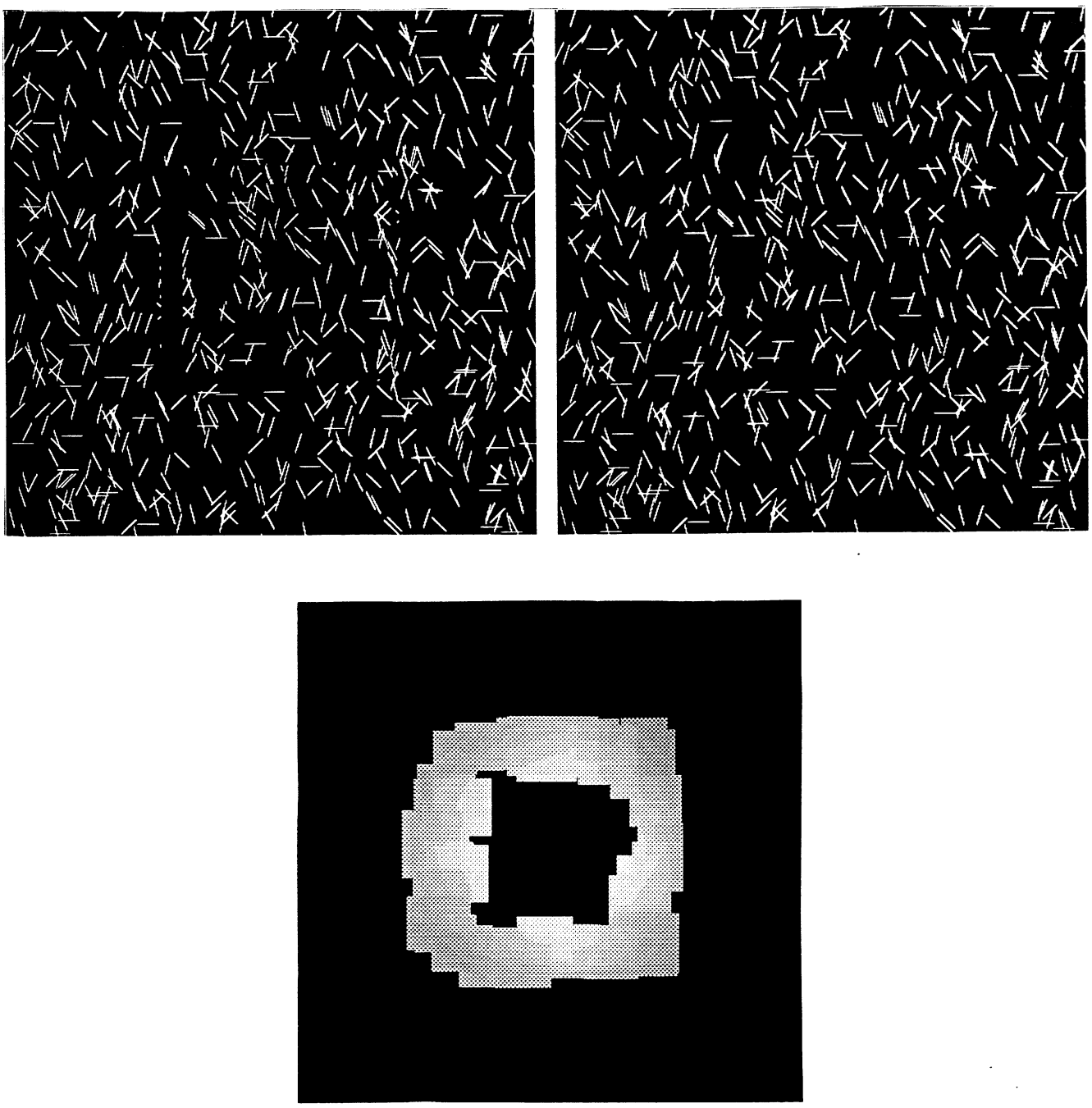


Figure 3.6: Results of a computer experiment with the proposed method for recovering the discontinuities of curved surfaces. The top half of the figure displays a random line stereogram corresponding to the input disparity information. The disparity corresponds to a central spherical region. The bottom half of the figure shows the recovered regions of discontinuity. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

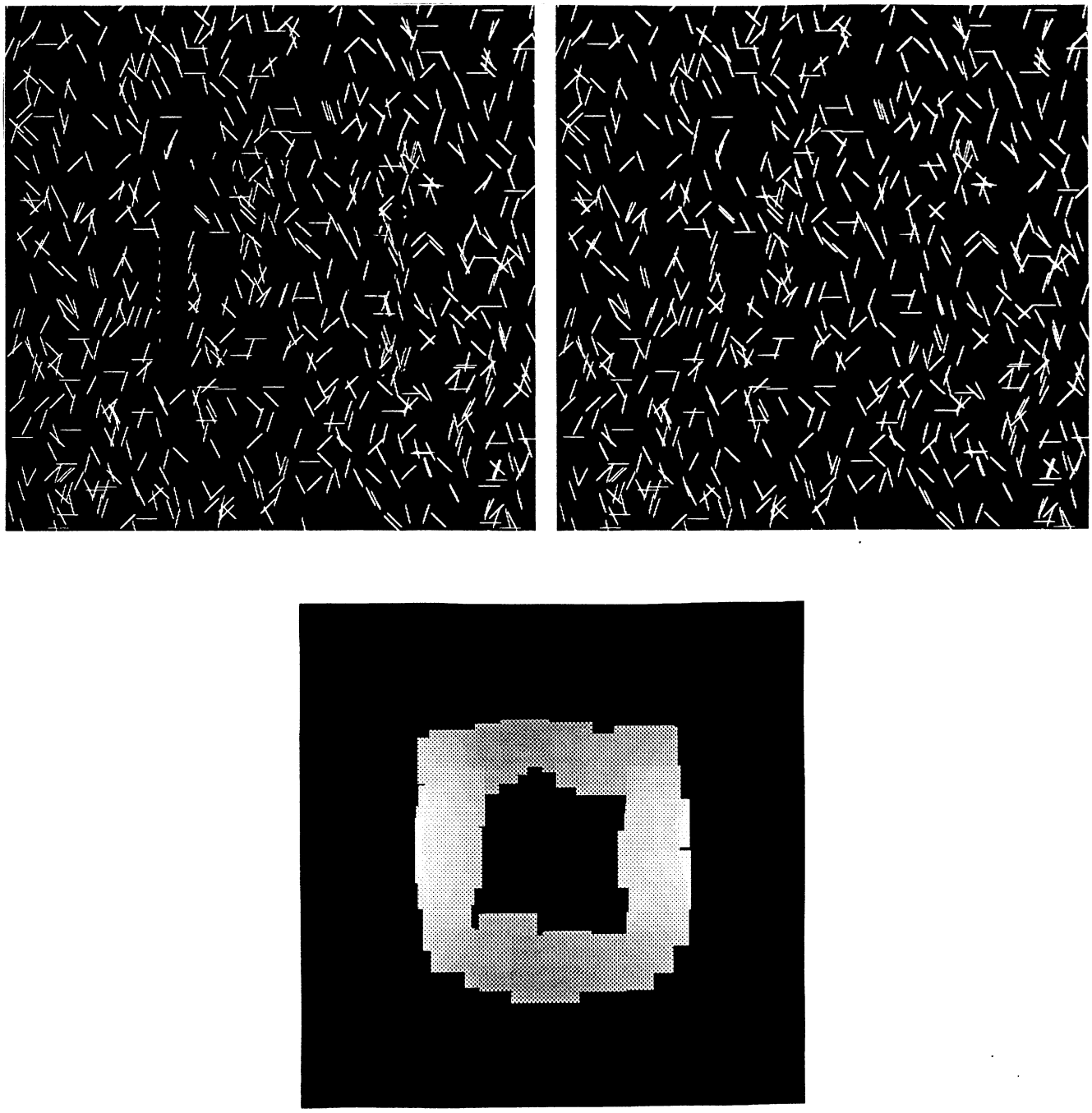


Figure 3.7: Results of a computer experiment with the proposed method for recovering the discontinuities of curved surfaces. The top half of the figure displays a random line stereogram corresponding to the input disparity information. The disparity corresponds to a central hyperbolic region. The bottom half of the figure shows the recovered regions of discontinuity. The recovered depth from the viewer is displayed in terms of grey levels with black the furthest and white the closest.

Chapter 4

Biological considerations

This chapter is concerned with issues surrounding the interpretation of stereo disparity by biological systems. Attention will be given to both psychophysical and neurophysiological studies. The first section of the chapter presents an overview of the relevant literature. The second section presents a new psychophysical study that has been motivated by the theory presented in this thesis.

4.1 Literature

This section provides a brief review of psychophysical and neurophysiological studies relevant to the interpretation of stereo disparity. For the most part, consideration will be limited to biological data that can be brought to bear fairly directly on the theory presented in this thesis. Two general issues will serve to focus the discussion: (i) To what extent is there evidence that biological stereo vision systems make use of horizontal, vertical and orientational disparity? (ii) What can be said about how

biological systems make use of the disparity that they do measure to recover three-dimensional scene geometry?

There have been many psychophysical studies that focused on horizontal disparity as an input to the human stereo vision system. Currently, there is little debate over the fact that horizontal disparity is a strong stimulus for stereopsis. Sources such as Ogle [97], Kaufman [56], Julesz [53] and Gulick & Lawson [44] (among others) copiously document the fact that differential projection along the horizontal dimension can lead to an impression of depth in a binocular viewer. There are also a number of studies that document neural sensitivity to horizontal disparity. For example: Barlow, Blakemore & Pettigrew [7], Ferster [26] and Ohzawa & Freeman [98] all report recording from cells in cat visual cortex that are sensitive to differential horizontal projection to the two eyes. In monkey, Hubel & Wiesel [52], Poggio & Fischler [101], Fischler & Poggio [27] and Poggio [100] report on cells in both striate and prestriate cortex that are sensitive to binocular horizontal disparity. Some of the evidence from monkey recordings also indicates that disparity is neurally coded into three coarse categories: cells selective for the locus of zero disparity (the horopter), cells selective for disparity that would arise from an object beyond the fixation (far cells) and cells selective for disparity that would arise from an object in front of the fixation (near cells). Interestingly, Richards [105] has reported evidence that human observers can be selectively "stereo blind" to zero, far or near disparity, thus supporting this type of coding scheme in humans.

While it is fairly clear that horizontal disparity provides some type of input to biological stereopsis, exactly how that information is used to yield a sense of three-

dimensionality is far less clear. With regard to the theory presented in this thesis, the following question is of particular interest: What is the relationship between horizontal disparity and local estimates of depth? Psychophysical studies suggest that while there is a relation between horizontal disparity and perceived depth, the relation is not overly straightforward. In particular, many studies show that when observers are presented with binocular horizontal disparity the resulting perception is not what would be predicted if the stereo system was performing a simple triangulation with all the viewing parameters known in advance; a constant binocular disparity corresponds neither to a constant perceived depth nor to a constant perceived distance ratio (see, e.g., Leibovic, Balslev & Mathieson [67] and Foley [29]). Much of this data can be accounted for by hypothesizing a scaling factor in the disparity to depth computation. There is some evidence that the setting of the scaling factor is related to extra retinal eye vergence information (Foley [29] and Foley & Richards [30]). However, at least two pieces of evidence suggest that vergence is not the only factor influencing depth scale: First, monocular cues can effect the apparent depth scale (Richards [106]). Second, the relative configuration of disparate elements in a given stereo display can effect the depth scale (Mitchison & Westheimer [84]). While there is no direct neurophysiological data available on depth scaling, it has been suggested that the necessary computation could be at least partially carried out by the lateral geniculate of the thalamus (Richards [104]).

Overall, there is currently not enough data available to specify the exact relation between horizontal disparity and perceived depth for biological systems. (Perhaps the best psychophysical review to date is provided in Foley [29].) However, the available

data is consistent with two notions put forth in the theoretical analysis of this thesis: First, local estimates of depth are related (in some fashion) to horizontal disparity. Second, the disparity to depth computation is not absolute, but involves a scaling factor. The suggestion in Chapter 2 of this thesis that the scale can be set arbitrarily is probably at odds with the available psychophysical data. Notice, however, that if an estimate of scale were availabe from some other source (e.g., vergence) it could be incorporated and used to set the scale in a nonarbitrary fashion.

At this point, attention is directed to issues surrounding the use of vertical binocular disparity in biological stereo vision. The theory proposed in this thesis does not exploit vertical disparity due to the suspicion that its relatively small magnitudes can not be accurately measured (see Appendix A). Nevertheless, there has been considerable controversy surrounding its use in the psychophysical literature. Therefore, it is appropriate to consider vertical disparity in this review. Most of the evidence in support of a role for vertical disparity comes from the so called "induced effect". Originally reported in Ogle [96], this phenomenon refers to the apparent slant of a frontal plane surface about a vertical axis resulting from the vertical magnification of the image to one eye. This general result has been replicated many times (see, e.g., Rigaudiere [108], Stenton, Mayhew & Frisby [117] and Gillam, Chambers & Lawergren [35]). Certain researchers have argued that vertical disparity can lead to the induced effect by establishing an inappropriate depth scale (see, e.g., Longuet-Higgins [69], Mayew & Longuet-Higgins [80] and Gillam & Lawergren [32]). However, a recent study that directly addressed the effects of vertical disparity on the scaling of horizontal disparity found no measurable effect (Fox, Cormack & Norman [28]). One

alternative explanation of the induced effect has been based in the observation that the differential vertical magnification of oblique lines will lead to horizontal disparities that can in turn be exploited. (See, Arditi, Kaufman & Movshon [5] and Arditi [4] for a discussion of this analysis as well as Mayhew & Frisby [78] for a rebuttal.) A similar argument can be made with the aid of the theory presented in this thesis: It appears that all of the experimental preparations used for the induced effect lead to not only vertical, but also orientational disparities. Thus, it is possible that the induced effect may have at its root the use of orientational disparity. The crucial experiment to date was reported in Westheimer [134]. In this study it was shown that vertical disparity in the absence of orientational disparity does not produce the induced effect, and in fact has no measurable effect on the available horizontal disparity.

It thus appears that the burden of proof for the use of vertical disparity in biological stereo vision is on those who would advocate its use. Before closing the discussion of vertical disparity two further points are worth noting: There is some evidence that biological systems are capable of at least detecting vertical disparity (Duwaer & van der Brink [24]). Finally, there is evidence that accurate stereopsis is actually difficult to obtain in the face of vertical disparity (Nielsen & Poggio [92]).

The final type of disparity measure considered in this thesis is orientational disparity. A number of psychophysical studies can be cited that implicate the processing of orientational disparities in humans. First, studies have demonstrated that it is possible to induce a tilt aftereffect in depth by adaptation to lines that appear differentially oriented to the two eyes (DeValois, von der Heydt, Adorjani & DeValois [22], von der Heydt [128]). While it is tempting to try and account for this result

using pointwise horizontal disparities, one observation makes this difficult: The aftereffect generalizes well to tilted patterns at other distances relative to the fixation. In another pertinent study (von der Heydt & Dursteller [130]), a binocular pattern of dynamic tilted random lines was presented to observers. Since the lines in the two eyes were randomly related to each other, the positional disparities were random in direction and amount. However, the lines were of different orientations in the two eyes and thus had a consistent orientation disparity. The pattern was perceived as tilted. Finally, studies that separately manipulate the orientational disparity and horizontal disparity at the end points of line segments find that orientational rather than horizontal disparity is more effective in conveying depth slant information (Ninio [93, 94]). Turning to the neurophysiological literature leads to a less clear situation. Blakemore, Fiorentini & Maffei [13] reported on cells in cat visual cortex that were selective for binocular orientational disparity. Due to the high variability of response in these cells, the result has been criticized as simply due to random errors of measurement (Nelson, Kato & Bishop [91] and Bishop [10]). However, a recent report on differential orientation tuning in monkey binocular cortical cells appears to be on firmer ground (Hänny, von der Heydt & Poggio [46], see also von der Heydt, Adorjani & Hänny [129]).

The theoretical analysis of orientational disparity presented in this thesis has emphasized its potential usefulness in recovering surface orientation. The psychophysical studies that have been noted in support of the general notion of orientational disparity also clearly support its role in computing surface orientation. However, other reports question the strength of binocular cues that correspond soley to a surface slanted in

depth. For example, there is psychophysical evidence that this type of stimulation can lead to percepts that build up very slowly over time (Gillam, Chambers & Russo [34]) and are easily overriden by competing cues, such as perspective (Gillam [31] and Stevens [118, 117]). Another type of result is worth mentioning with regard to orientational disparity: In two sources (Ninio [93] and Richards [107]) it is shown that differentially projected linear elements to two eyes' views can lead to a strong percept of three-dimensional quill-like textures. Interestingly, it is difficult for humans to recover the underlying surface geometry in such displays. There is also evidence for individual differences in the abilities of observers when faced with this type of display. Similarly, recall that the method proposed in this thesis for exploiting orientational disparity in recovering surface geometry is predicted to face difficulties in the face of three-dimensional texture.

In summary of this discussion of orientational disparity: It appears safe to conclude that orientational disparity plays some role in biological perception. It also appears that the theory presented in this thesis is in general accord with known data from biology.

The final set of results reviewed in this section center around the recovery of three-dimensional surface discontinuities from stereo disparity. In general, there is evidence that stereoscopic stimuli corresponding to discontinuities in depth are powerful stimuli for biological systems (Gillam et al. [33], Gillam, Chambers & Russo [34], Stevens [118, 117]). Unfortunately, the details concerning how this facility is accomplished are much less clear. The above cited authors argue in favor of something akin to discontinuity recovery based on finding discontinuities directly in the disparity infor-

mation. However, their evidence does not directly implicate this type of strategy. Rather, it is simply supportive of the general notion that the stereoscopic projection of three-dimensional discontinuities is important.

Now, recall that the theory presented in this thesis is based on the local recovery of planar surface patches from disparity. There is some evidence in support of this idea: Mitchison [83] has studied how ambiguous stereoscopic depth segmentation is accomplished using simple repetitively patterned dot stereograms. He concludes that the segmentation is based on fitting locally planar surfaces to the endpoint disparities. Further evidence along these lines has also been reported (Mitchison & McKee [85, 86, 87]). Another interesting psychophysical study in the recovery of discontinuities from disparity can be found in Anstis, Howard & Rogers [3]. Here observers were shown two flat vertical textured surfaces in the frontoparallel plane that met at a vertical boundary. At the boundary one surface curved slightly forward, while the other slightly backward. Even though the flat portions were equidistant from the observer, the entire curved forward side of the display appeared to be closer. In fact, the entire forward surface tends to cling to the front most edge. (In analogy with a similar "illusion" in luminance the authors refer to this as a Craik-O'Brien-Cornsweet illusion for depth.) This percept is interestingly similar to the way the discontinuity recovery algorithm returns a locally planar patch that clings to the locus of the discontinuity in three-space. Additional reports that the human depth illusion is anisotropic (Rogers & Graham [110]) are not explained by the model proposed in this thesis. In conclusion, the method for recovering depth discontinuities that has been proposed in this thesis is generally consistent with the known psychophysical

data.[1]

In summary, several points should be stressed:

- The bulk of the evidence concerning vertical disparity rules against its use by biological systems.
- There is both psychophysical and physiological evidence in favor of biological systems exploiting orientational disparity.
- There is psychophysical evidence for individual differences between observers and for anomolous stereo observers.
- The theory that has been proposed in this thesis for the recovery of surface discontinuities is consistent with psychophysical evidence on discontinuity recovery.

Finally, there are no available experiments to directly contrast the use of curvature vs. approximating planes in the recovery of surface discontinuities. The theory proposed in thesis predicts the use of planar information. The next section of this chapter presents a psychophysical study that directly tests this prediction.

[1] Evidence of a rather different kind is also available for the human recovery of surface discontinuities from binocular displays: Lawson & Gulick [65] have developed binocular stimuli that contain the type of monocular occlusions typical of viewing in the vicinity of a depth discontinuity. These investigators demonstrate that the occlusion information by itself can lead to a vivid impression of a discontinuous planar surface in depth.

4.2 Experiment

This section presents a study that tests the psychological validity of the method proposed in this thesis for recovering the discontinuities of curved surfaces from stereo disparity. Recall that the proposed method operates in two stages: First, it recovers the three-dimensional planar approximation that corresponds to local disparity information. Second, it checks to see if neighboring planar approximations intersect in the intervening region of the disparity map. If they do not intersect in this neighborhood then the viewed surface is taken as locally discontinuous. Guided by this theory it is relatively easy to construct stereoscopic displays that should or should not be perceived as continuous. It is also easy to contrast these predictions with those of a likely rival theory: Curved surface discontinuities are recovered by comparing neighboring surface curvature measures. Specifically, then, the experiments test whether discontinuities in distance are perceived using planar or curvature information.

A set of seven stimuli have been devised to test the proposed theory and to contrast it with a potential theory based directly on surface curvature. In this study consideration has been limited to surfaces that are singly curved, e.g., cylinders. The stimuli are shown as left and right stereograms in figures 4.1-7 at the end of this chapter. Notice that each figure has three panels rather than the minimal two needed to render the stimulus. This is done so that an observer can view the stimuli using either crossed or uncrossed stereo fusing. That is, the left most panel can be viewed by the left eye with the middle panel viewed by the right eye to yield one version of the stimulus; similarly, by viewing the middle panel with the left eye and the right

most panel with the right eye all the depth relations are reversed. Each stimulus is a random line stereogram made up of 350 randomly placed and oriented lines that have been differentially projected to correspond to a singly curved surface in depth. A cross section in depth of each stimulus is shown in the column labelled "configuration" in Tables 4.1 and 4.2. The displays all have the same slant in depth corresponding to the upper left hand corner of the display slanting away from the viewer. Each display has a blanked out diagonal region. The experimental task involves determining if the depicted surface is discontinuous or smooth within this blank region. The blank region is employed to rule out the possible confounds due to abrupt changes in the disparity where two surfaces meet.

The differential surface curvature used to generate the displays is the distinguishing feature that allows for a test of the proposed theory. Consider these stimuli one-by-one, refering to the "configuration" column in Table 4.1. Stimulus 1 corresponds to a purely planar surface that is discontinuous across the gap. It is included in order to get a crude discrimination assessment and ensure that observers are capable of performing the task. Stimulus 2 corresponds to curved surfaces of the same sign that are discontinuous across the gap. A theory based on either local planar geometry or curvature would assert a discontinuity in this case. Stimulus 3 is a case where the same signed curvatures of the two surfaces will meet across the gap while the extrapolated tangent planes will not meet. Stimulus 4 is an example where the same signed curvatures will not meet across the gap while their extrapolated tangent planes will. In stimulus 5 the same signed curvatures will not meet and while the planar extrapolations will meet, they will not meet in the region of the gap. In stimulus

6 the opposite sign curvatures meet across the gap while the planar extrapolations do not. Finally, stimulus 7 is the converse of stimulus 6.

In the experiment the stimuli were presented as red-green anaglyphs displayed on a Conrac color monitor in a darkened room while the observers wore red-green colored glasses. The polarity of the disparity in the experiment was the same as viewing the figures with the left and right eyes fixating the left and middle panels, respectively. (One observer, B in the tables of results, viewed the displays in this configuration as well as reversed. His reported perceptions were the same in either case.) The stimuli were viewed at two distances in order to address the possible role of spatial integration in this task. The far and near viewing conditions had the observers seated at eye level with the monitor at distances of 6 and 2.15 meters, respectively. The upper and lower triangular regions of the displays each measured 25.4 cm. along both the horizontal and vertical dimensions on the monitor. The diagonal separation of the upper and lower regions (i.e., the gap width) was 6.35 cm. on the monitor. The line elements that were used to comprise the display measured .8mm on the screen. In terms of degrees of visual angle, the width $\times$ height of the overall display was 8.53×6.73 and 3.15×2.42 in the near and far conditions, respectively.

The stimuli were always presented in the same order 1-7. The head and eye movements of the observers were not constrained in any way. Observers were asked to make two judgements upon viewing the displays. First, they were instructed to give their first impression as to whether the viewed surface was discontinuous or smooth within the blanked region. Second, the observer was to say if they saw the surface as extending into the gap as planar or curved. Four observers participated in the

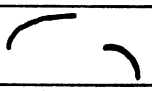
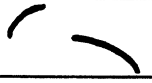
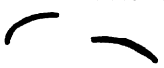
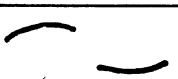
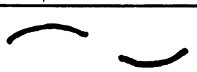
stimulus	configuration	response							
		disc.				planar			
		observer				observer			
		A	B	C	D	A	B	C	D
1		x	x	x	na	x	x	x	na
2			x	x	na	x	x	x	na
3		x	x	x	na	x	x	x	na
4					na	x	x	x	na
5		x	x	x	na	x	x	x	na
6		x	x	x	na	x	x	x	na
7					na	x	x	x	na

Table 4.1: Results for the far viewing condition.

experiment. None of the observers were completely naive as to the purpose of the experiment. Subject D was only able to take part in the near condition.

The results of the four observers are presented in Table 4.1 for the far viewing condition and Table 4.2 for the near viewing condition. To interpret the tables note that an "x" under "disc." means that the observer reported the display as discontinuous in the gapped region. An "x" under "planar" means that the observer reported that the surface under consideration extended into the gap as a plane. As an aid to understanding the results, responses that are inconsistent with the planar theory, as proposed in this thesis, have been cross-hatched in the two tables. The notation "na"

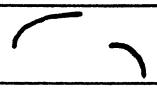
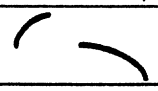
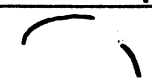
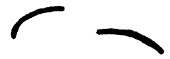
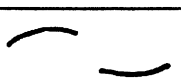
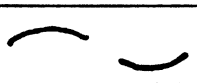
stimulus	configuration	response							
		disc.				planar			
		observer				observer			
		A	B	C	D	A	B	C	D
1		x	x	x	x	x	x	x	x
2		x	x	x	x			x	x
3				x	x			x	x
4						x	x	x	x
5		x	x	x	x	x	x	x	x
6		x	x	x	x	x	x	x	x
7		x						x	x

Table 4.2: Results for the near viewing condition.

means data not available.

Consider the results for the far viewing condition as shown in Table 4.1. The overriding result in this case is that all observers reported seeing the extension of the patches into the blanked regions as planar. All but one of the responses in the discontinuity judgement are in accord with the proposed theory. (This response is highlighted by cross-hatching in Table 4.1.) However, the response that conflicts with the planar theory also conflicts with a curvature based theory of discontinuity detection. The observer reported that in this condition the entire display appeared to flatten out. Under this perception it is not surprizing that the display was reported as continuous.

Now, consider the results for the near viewing condition as shown in Table 4.2. Notice that in this condition the stimulus is larger and a given area of the display now maps to a larger region of visual integration (this follows directly from the geometry of the situation). Therefore, it is not surprizing that there are more reports of seeing curvature in the blanked out regions of the displays. However, despite this situation observers C and D still report only planar percepts in the blank regions. Similarly, observers C and D make all their discontinuity judgements in accord with the planar theory.

The results for observers A and B are slightly more complex. To begin, notice that for stimuli 1 and 4-6 these observers also report planar percepts; their judgements of discontinuity are in exact accord with the planar theory in these cases. Stimulus 2, while reported as curved in the blanked region cannot distinguish between the planar and curvature approaches to recovering discontinuity; both methods predict the same

observed responses. The cases of most interest are the reports of A and B for stimuli 3 and 7. As shown by cross-hatching in Table 4.2, three out of four of these responses are in conflict with the planar theory. These three responses can be interpreted as evidence for a curvature based scheme for recovering discontinuities.

While three of the judgements in Table 4.2 can be taken as contrast to the theory that has been proposed in this thesis, it is not clear that the judgements are based on curvature information per se. It is also possible that the measures are based on the change of the tangent planes to the perceived surfaces. That is, the responses could be based on the change in the first order (planar) information rather than directly on second order (curvature) information. This hypothesis is strengthened by recalling that there is no evidence for curvature based judgements in the far viewing condition. In the far condition the area of visual integration for a given region of the display is smaller; thus, change of tangent plane becomes less salient. Unfortunately, the present set of displays does not adequately separate out the issues of curvature vs. change of tangent planes. The proper control is to check and see if the same absolute curvature in the far condition will interpolate the same as in the near condition.

Drawing together these results, several conclusions can be drawn:

- Strikingly, out of forty-eight potential curvature responses only three responses can be interpreted in terms of a curvature based mechanism for recovering surface discontinuities from stereo disparity. All but one of the remaining forty-five cases support the planar theory of discontinuity recovery.

- In all but one case, when the interpolated blank region is seen as planar the

discontinuity judgement is also in accord with a planar theory of discontinuity recovery.

- There are three judgements that are consistent with a curvature based mechanism. However, there is evidence that these judgements are actually based on change of tangent plane.

- Finally, it is interesting to note that the results indicate individual differences between observers. In the review section of this chapter, other evidence for individual differences in stereoscopic perception was noted (Richards [105] and Ninio [93]).

4.3 Recapitulation

This chapter was divided into two main parts. The first section provided a brief review of psychophysical and neurophysiological studies regarding the interpretation of stereo disparity. During the course of this presentation it was noted that the theory of stereo disparity interpretation that has been developed in this thesis is generally compatible with the reviewed data. The second part of the chapter presented a new psychophysical study addressing a specific aspect of the proposed theory. In particular, the study compared human performance on the recovery of the discontinuities of curved surfaces with predictions from the theory. The results of the experiment demonstrate that discontinuity measures can be based on first order surface geometry. These results leave open the question of whether curvature information is ever directly used in this task, as opposed to basing judgements on the change of first

order (i.e., planar) information.

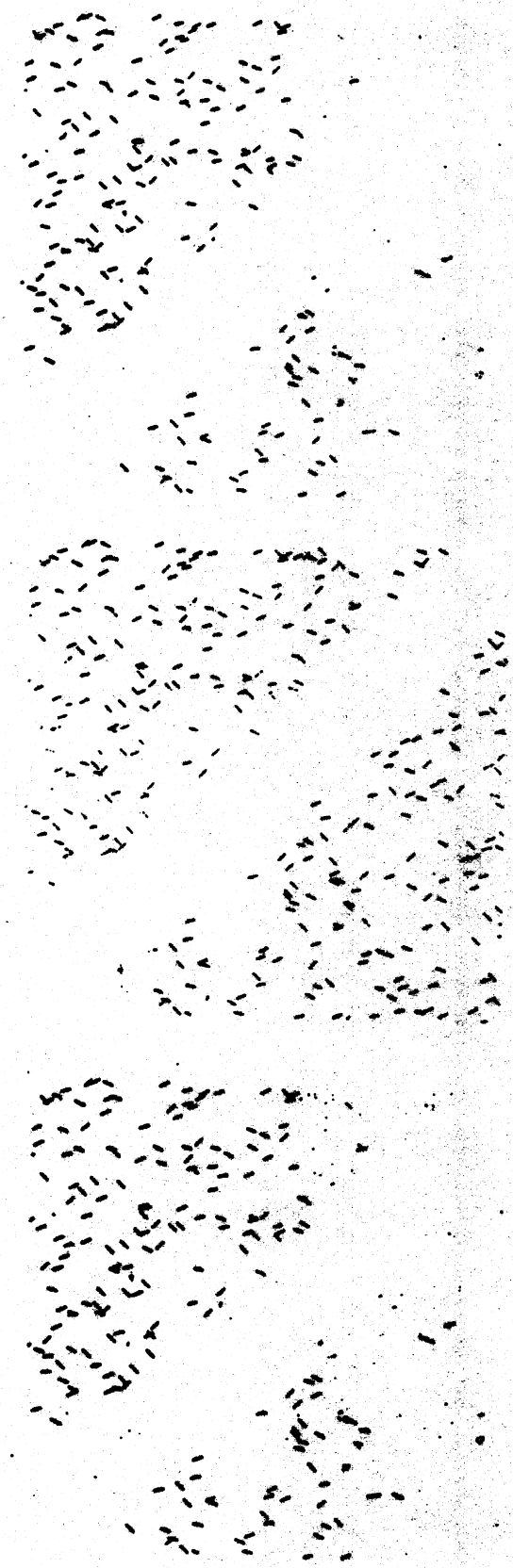


Figure 4.1: Stimulus 1.

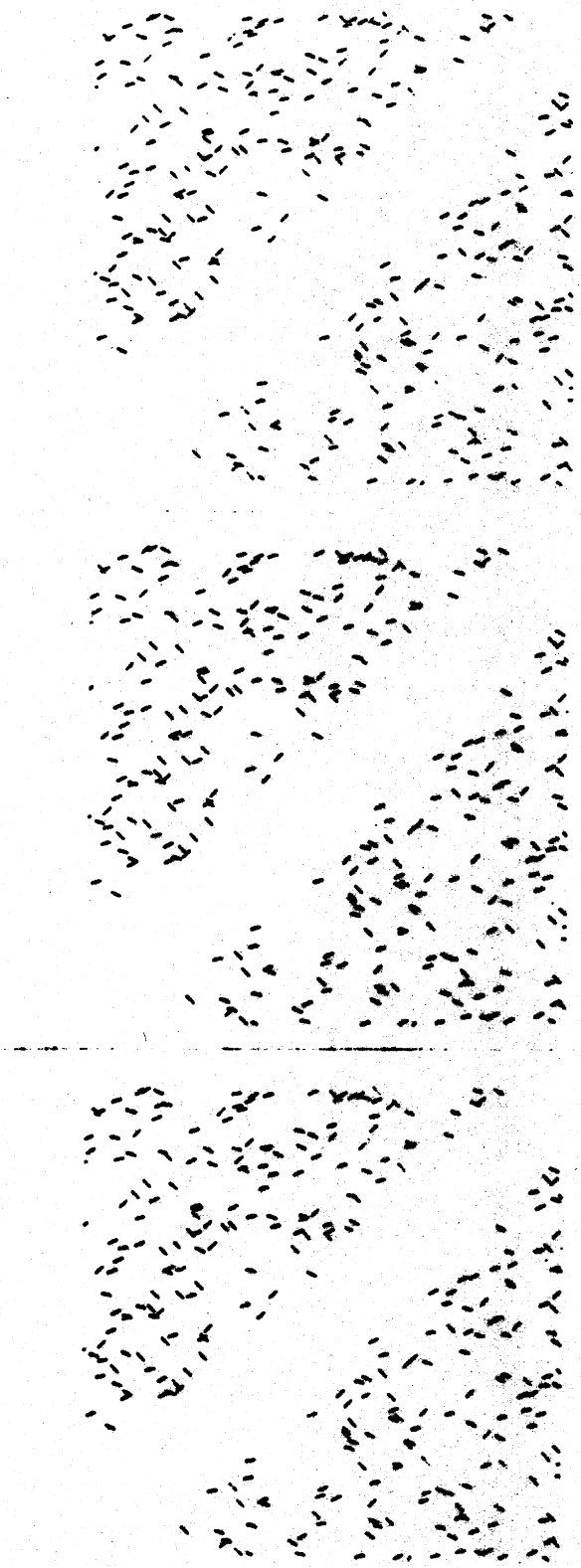


Figure 4.2: Stimulus 2.

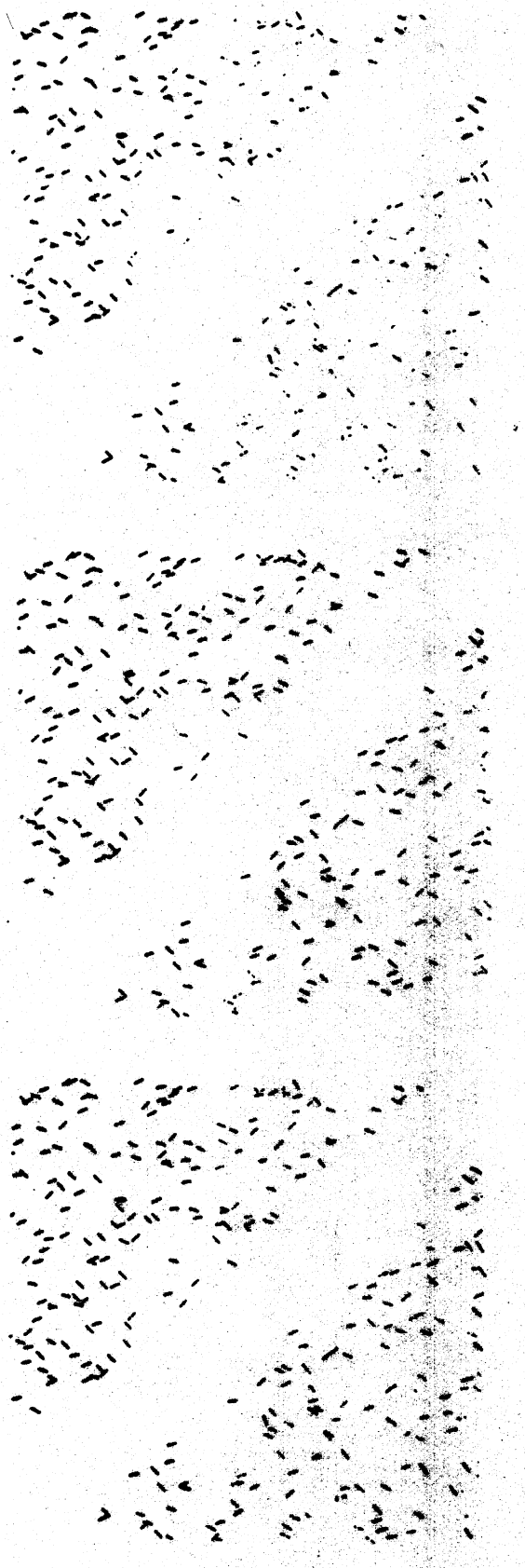


Figure 4.3: Stimulus 3.

Figure 4.4: Stimulus 4.

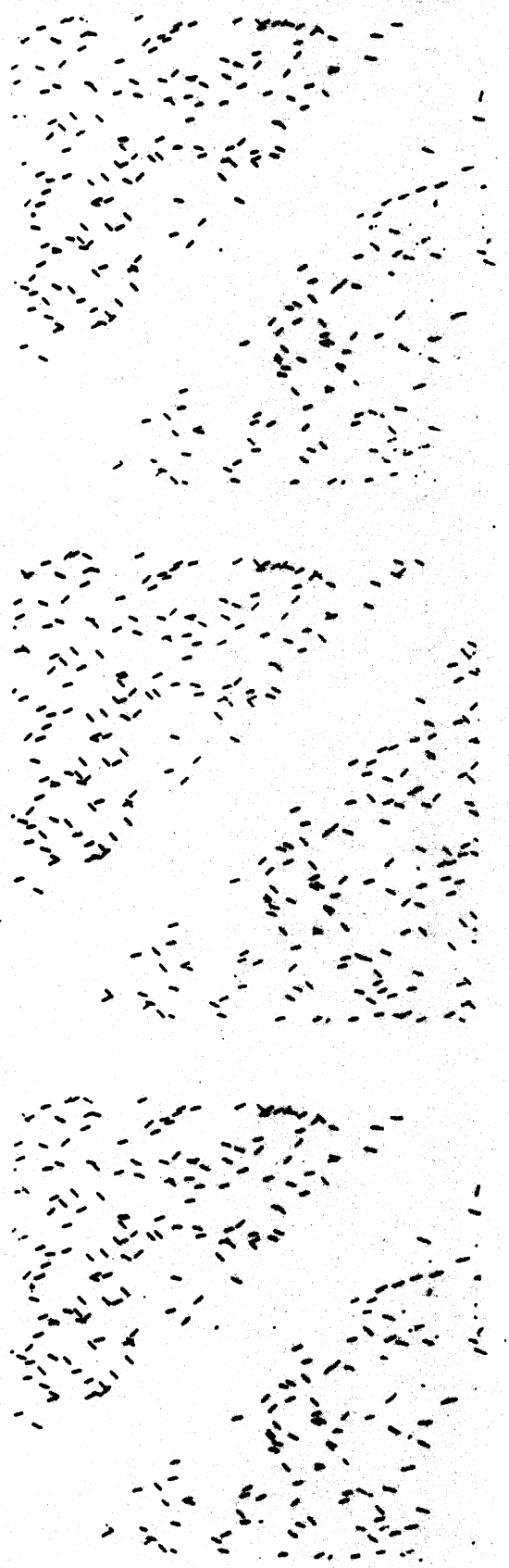


Figure 4.5: Stimulus 5.

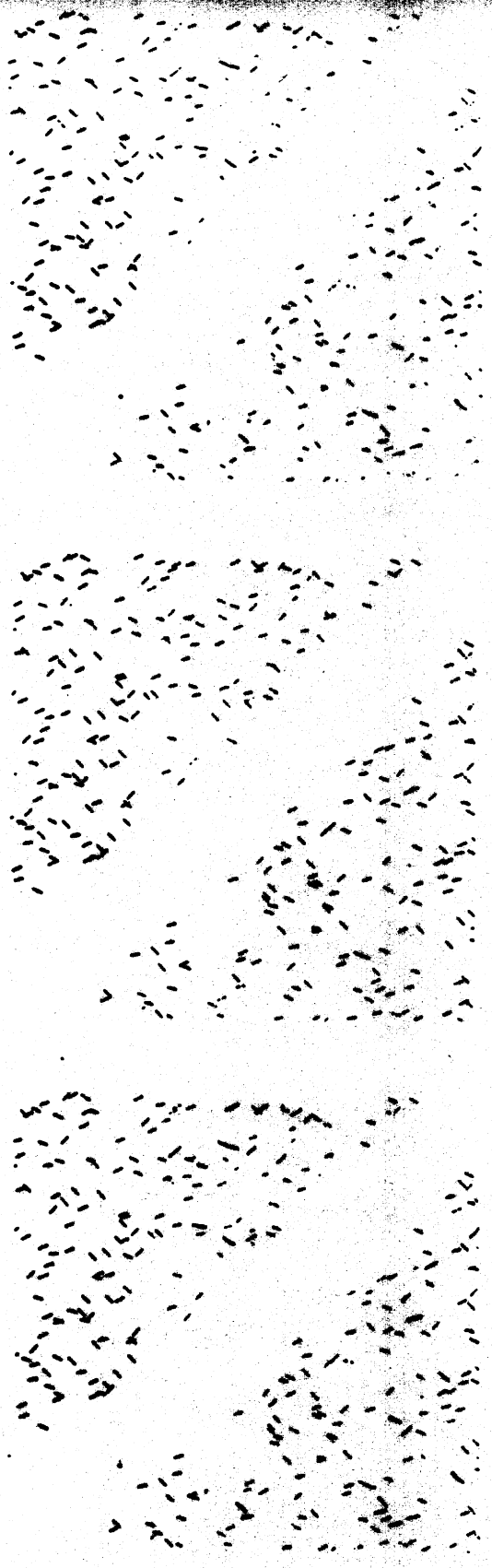


Figure 4.6: Stimulus 6.

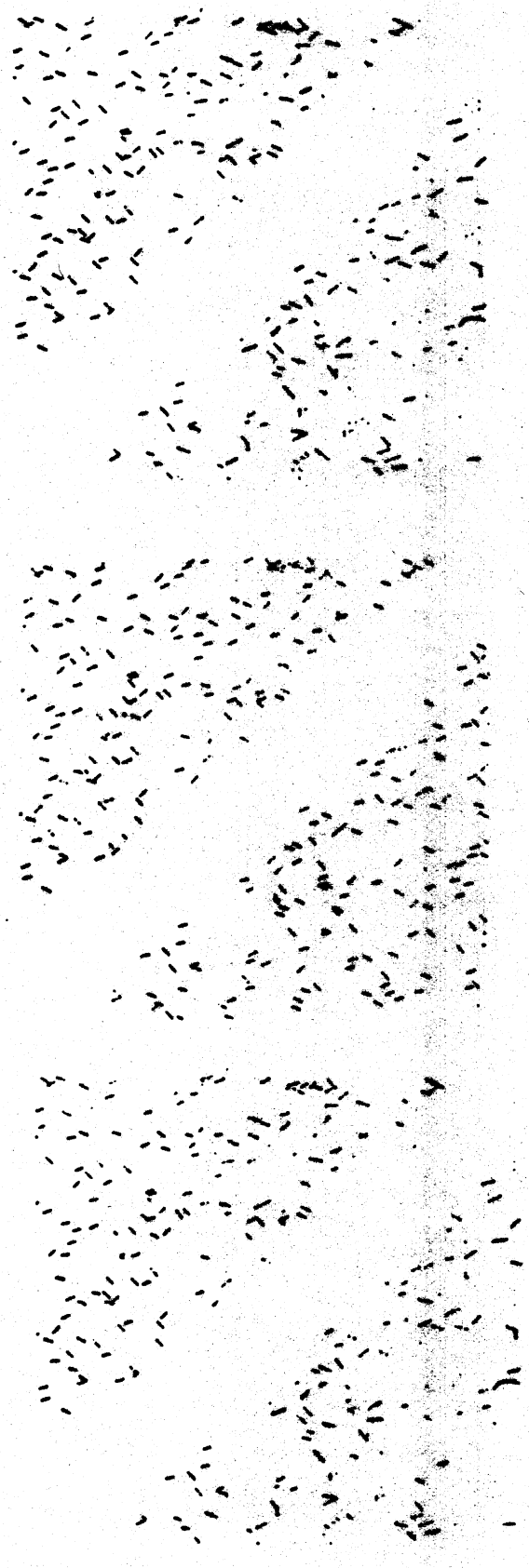


Figure 4.7: Stimulus 7.

Chapter 5

Conclusions and suggestions for further research

5.1 Summary and conclusions

This thesis has sought to develop an understanding of binocular stereo disparity. More specifically, the goal has been to develop an analysis within which one can derive a set of mappings from stereo disparity to useful descriptors of three-dimensional scene geometry. To this end, it has been shown how it is possible to recover surface depth, orientation and discontinuities directly from stereo disparity. Previous studies of the disparity interpretation problem have not explicitly related stereo disparity to surface orientation and discontinuities. A key to the success of the current study has been delaying attempts to recover three-dimensional information from disparity until a rigorous analysis of disparity itself was in hand.

Chapter 2 began the developments by considering the special case of planar sur-

faces. Stereo disparity was presented as a vector field resulting from the projection of a three-dimensional scene onto a pair of imaging surfaces related via an infinitesimal change of coordinates. With this representation in hand it was possible to make use of analytic techniques from classical field theory to effect the recovery of surface depth, orientation and discontinuities. Next, the resulting relations were studied for numerical stability. The stability analysis also provided the basis for understanding how to set thresholds for operations in the face of noise perturbed data. Chapter 2 closed by describing a set of computer algorithms for the recovery of surface discontinuities from stereo disparity. The algorithms were based on the theory proposed in this thesis. The results of applying the algorithms to several images were reported.

Chapter 3 presented an extension of the theory for planar surfaces to curved surfaces. The particular extension that is developed is the recovery of surface discontinuities for curved surfaces. The analysis was based on approximating a curved surface with locally planar patches. The key constraint on surface discontinuity recovery was derived by considering the projection of dihedral edges. Specifically: if the edge which would connect adjacent patches does not project between the patches then the surface is discontinuous. Following this analysis, the approach was studied for stability and a method for operating with inexact data was developed. Finally, a corresponding set of computer algorirthms and their results applied to disparity fields were described.

Chapter 4 presented relevant empirical data from visual psychophysics and neurophysiology. The chapter began by briefly reviewing the psychological and biological literature concerned with disparity interpretation. In general, the theory present in

this thesis is not at odds with the known literature. The second part of the chapter presented a new psychophysical study that supports the theory of surface discontinuity recovery presented in this thesis. Specifically, human observers presented with random line stereograms of cylindrical surfaces may base their judgments of surface discontinuity on depth and first-order surface geometry rather than on second-order surface geometry.

At this point it is important to step back and ask about the significance and status of the research that has been presented in this thesis. Several points should be emphasized:

- The purely theoretical sections of the thesis have presented an indepth analysis of the stereo disparity interpretation problem. The relations between stereo disparity and first order scene geometry have been precisely defined.

- Extensive numerical analysis of the disparity relations shows that they are typically quite stable. Significantly, this analysis indicates how algorithms based on these relations can monitor the stability of their own behavior.

- The understanding of stereo disparity gained from these analyses has led to a method for attacking an important problem in the processing of visual information: Recovering the discontinuities in distance to the surfaces in a viewed scene. This method has been implemented in computer algorithms and successfully tested on synthetic and natural stereo disparity data.

- The method for recovering surface discontinuities could be put to practical use as an integral part of a vision system. The early recovery of surface disconti-

nuities can serve as useful information for a number of vision tasks: (i) Surface reconstruction, where discontinuities serve to define boundary conditions; (ii) Passive naviagation, where discontinuities provide information about the configurations of obstacles in the world; (iii) Shape recognition, where discontinuities can cluster data belonging to a single object and (if precise enough) define the outline of an object.

- The understanding of binocular stereopsis that has been gained in this study can be used to make precise psychophysical predictions. This can motivate both psychological and neurophysiological investigations. An example of such a psychology experiment was presented in Chapter 4.

- The approach that has been developed to studying disparity is quite general and could be used to investigate other types of disparity information, e.g., motion.

5.2 Suggestions for further research

Several directions for further research can be discerned. Consider in turn (i) further theoretical developments and (ii) further empirical research.

One possibility is to reconsider the analysis of disparity due to the projection of planar surfaces. In particular, note that not all of the information available in the disparity gradient tensor has been exploited. In fact, only that portion due to the untraced part of the symmetric component has been used (i.e., equations (2.14) and (2.15)). It would be interesting to study the information available in the unsymmetric portion (a curl) and in the traced part of the symmetric portion (a divergence) of the

disparity gradient. Both these components have interesting interpretations in terms of the differential imaging of surface detail. The curl can be captured via the relative rotation of corresponding elements; the divergence as a relative isotropic expansion of elements.

A second set of theoretical developments can be motivated by giving further attention to the disparity due to the projection of curved surfaces. An interesting research problem would be to attempt the recovery of surface curvature from disparity. Several paths present themselves: Rodgers [109] has suggested that surface curvature can be recovered as the second differential of disparity; however, this proposal seems suspect in the light of sparse and noisy data. Keeping within the framework followed in this thesis, it may be interesting to study the relations between surface curvature and the disparity curvature tensor. Still another approach would be to attempt to recover surface curvature by extending the local geometry manifest in the disparity gradient tensor. This can be done in a well founded fashion via the connection equations of differential geometry (Prakash [103] and Spivak [115]). Of particular use for this case would be the Gauss-Weingarten equations of classical surface theory. These equations relate local first-order geometry to the coefficients of the first and second fundamental forms of a curved surface. (Koenderink & Richards [62] have used a similar approach to derive stable two-dimensional curvature operators.) It is also interesting to think in more qualitative terms for the recovery of curved surface properties. For example: Is it possible to map the disparity information directly into qualitative descriptors of surface geometry (e.g., the differential geometer's parabolic, elliptic and hyperbolic patches (Prakash [103] and Spivak [115]) or the topographist's watersheds, hills and

dales (Cayley [19] and Maxwell [77])). As an initial attack it may be possible to effect such a qualitative recovery through the study of the residuals of disparity to the planar analysis of chapter 2.

Another future theoretical development would reconsider the recovery of surface discontinuities for the case of curved surfaces. The analysis presented in this thesis is founded in approximating polyhedra and difference geometry. This analysis can be naturally extended by letting the difference equations pass to the limit and become differential equations. It would be interesting to couple the resulting equations with the Mainardi-Codazzi equations of integrability (Prakash [103] and Spivak [115]) to see what new insights could be derived. For example, violations of the Mainardi-Codazzi equations would indicate that a surface discontinuity was present.

Future research could also serve to further the stability analysis of the recovery methods. In particular, it would be useful to consider the relationship between the stability of a method and the actual expected errors that might arise in application. This leads into a need to develop an understanding of stereo matching errors. (There is some existing work on this topic, e.g., Blostein & Huang [15, 14], McVey & Lee [81], Mohan, Medioni & Nevatia [88], Nishihara [95], Verri & Torre [127].) A better understanding of these errors would also be of use in setting thresholds for the discontinuity recovery method.

Finally, much consideration should be given to further empirical testing of the current version of the theory as well as any future developments. Of particular interest is to further test the theory with corresponding computer algorithms applied to natural image data. The performance of a computational vision theory in the face of

natural imagery is in a sense the "acid test" for the discipline. Its importance should not be underestimated.

Appendix A

Recovering view

This appendix presents four additional approaches to recovering the differential viewing parameters t_x, t_z and ω_y which relate the two stereo views. All methods work with the assumption that the magnitude of the interocular separation is a known value, say I. The first two methods recover the viewing parameters only up to an arbitrary depth scaling factor. The third and fourth methods recover the viewing parameters without resorting to an arbitrary scaling factor.

A.1 Full perspective method

This method is called the "full perspective method" in that it exploits the information in the full disparity field, both horizontal and vertical disparity. Recall equation (2.6). Now, if the solution is to allow for an arbitrary depth scaling factor it is permissible to set the depth value of some point arbitrarily. For convenience, suppose that this value is set to unity. Then at some point equation (2.6) provides a set of two equations in

the three unknown viewing parameters t_x, t_z and ω_y. That is,

$$\chi_x = xt_z - t_x - (x^2 + 1)\omega_y \tag{A.1}$$

$$\chi_y = yt_z - xy\omega_y. \tag{A.2}$$

A third constraint can be derived with regard to the known magnitude of the stereo base-line, I. Specifically,

$$I^2 = t_x^2 + t_z^2. \tag{A.3}$$

Together, relations (A.1), (A.2) and (A.3) allow for the recovery of the unknown viewing parameters t_x, t_z and ω_y from the measurable horizontal and vertical disparities, χ_x and χ_y up to a sign ambiguity. The sign ambiguity can be resolved by considering another image point and the corresponding disparity values. In this case, checking to make sure that the same value of Z is derived from consideration of both the horizontal and vertical disparities allows the sign ambiguity to be checked. Thus, the viewing parameters can be recovered (up to a scale factor) by consideration of the horizontal and vertical disparities at two points. (c.f., Longuett-Higgins [68] where the simultaneous observation of seven horizontal and vertical disparities are used to recover relative viewing parameters without an arbitrary scale, if the magnitude I is known.)

A.2 Orthographic approximation

The orthographic approximation derives from a special case analysis of the restricted perspective method. Specifically, suppose it is known, or the viewer is willing to

assume that the angle γ is equal to zero. That is to say, the observer is looking straight ahead. In this case relation (2.25) is found to simplify to

$$\chi_x = \frac{I(px + qy)}{r}. \tag{A.4}$$

Again, assuming I is known and setting the depth scale arbitrarily allows for the relative surface orientation to be recovered directly (assuming that x and y are chosen to ensure linear independence for the system).

This method of recovery is not as general a model of the physical situation as are the methods of full or restricted perspective projection. For this reason it shall not receive much further attention in the main body of this thesis (although it shall be reconsidered in the appendix). However, three points are worth commenting on: First, (A.4) side steps the issue of recovering viewing parameters. Second, it is a reasonable approximation for many real world viewing conditions when an observer is looking approximately straight ahead. Third, formulation (A.4) may hold particular interest from a psychological stand point. This is due to the fact that use of equation (A.4) will lead to a systematic error in the visual periphery. Interestingly, humans are increasingly inaccurate in processing stereoscopic information in the periphery (Foley [29] and Helmholtz [48]).

A.3 Recovering view with absolute scale

Consider now the possibility of recovering view without an arbitrary scale factor, but while still restricting consideration to only horizontal and orientational disparity. In the three previous formulations for recovering view the scale was set arbitrarily by

assigning a depth value (e.g., unity) to some point. In the present analysis, the depth value will be cancelled by dividing horizontal and orientational disparities at a point. Two formulations shall be presented. One of these formulations will require that the view parameters, t_x, t_z and ω_y, are recovered in tandem with the surface orientation parameters, p and q, for a single planar patch.

Both formulations begin by noticing that substituting (2.20) into (2.19) and making the substitutions implied by (2.24) allow for horizontal disparity to be written as

$$\chi_x = \left(\frac{I\cos\gamma}{r}\right)[(1 - px - qy)(x\tan\gamma - 1) + (1 + x^2)]. \tag{A.5}$$

Next, similar geometric substitution allows (2.15) to become

$$\sigma = \frac{I\cos\gamma}{r}(p^2 + q^2)^{\frac{1}{2}}. \tag{A.6}$$

Then, dividing (A.5) by (A.6) (roughly, dividing horizontal by orientational disparity) yields

$$\frac{\chi_x}{\sigma} = (p^2 + q^2)^{\frac{-1}{2}}[(1 - px - qy)(x\tan\gamma - 1) + (1 + x^2)]. \tag{A.7}$$

This manipulation has accomplished the goal of eliminating the depth parameter r.

The first attack on solving (A.7) for view and surface orientation proceeds as follows: Relation (A.7) can be forced into a quasi-linear equation relating $\tan\gamma$, p and q with one final substitution. Specifically, using the relation between p and $(p^2 + q^2)^{\frac{1}{2}}$ implied by (2.21) allows the form

$$b = xa_1 + ya_2 + xya_3 + x^2a_4 \tag{A.8}$$

with

$$
\begin{aligned}
b &= \frac{\chi_x}{\sigma(\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2)} - x \\
a_1 &= \frac{\tan\gamma}{p} \\
a_2 &= \frac{q}{p} \\
a_3 &= -\frac{q}{p}\tan\gamma \\
a_4 &= p^{-1} - \tan\gamma.
\end{aligned}
\tag{A.9}
$$

This system can be solved by observing the horizontal and orientational disparities at four points. The original variables of interest, $\tan\gamma$, p and q, can be recovered using the relations involving a_1 - a_3 and saving a_4 as added constraint.

Such a nonlinear solution leads naturally to a question of multiple solutions. Geometric reasoning applied to the $(\tan\gamma, q, p)$-solution space is of use: Notice that the relation involving a_3 constrains the solutions to lie on a saddle-like surface in the $(\tan\gamma, q)$-plane. Also notice that the relations involving a_1 and a_2 jointly define a line in this space. Further thought shows that the line will pierce the saddle in a single point (and thus make for a unique solution) with two exceptions: (i) If both p and q vanish the line will intersect the saddle in a line, which allows for arbitrary $\tan\gamma$. (ii) If both p and $\tan\gamma$ vanish the line again intersects the saddle in a line, this time allowing for arbitrary q.

A second solution to (A.7) without arbitrary scale solves for only $\tan\gamma$. For this solution, substitute into (A.7) the values for p and q implied by (2.21). Then letting $a = [x(\hat{\xi}_{1x}^2 - \hat{\xi}_{1y}^2) + 2y\hat{\xi}_{1x}^2\hat{xi}_{1y}^2]$ and rearranging yields

$$
\frac{\chi_x}{\sigma} - a = x^2 \| \bigtriangledown Z \|^{-1} - xa\tan\gamma + x\tan\gamma \| \bigtriangledown Z \|^{-1}. \tag{A.10}
$$

With two sets of observations $\| \bigtriangledown Z \|^{-1}$ can be eliminated from (A.10). Then the value

of $\tan\gamma$ can be recovered as a solution to a quadratic equation. There is, of course, a two way ambiguity inherent in this solution. Consideration should be given to the possibility of ruling out one of the solutions on the basis of e.g., the reasonableness of the resulting viewing parameters.

A.4 Considerations of stability

The numerical stability of the full perspective method for recovering view and surface parameters has been investigated empirically. This investigation was conducted by implementing the system of equations (A.1), (A.2) and (A.3) in a simple computer program. The program operates in two stages. First, input horizontal and vertical disparity values to recover the viewing parameters. Second, the recovered viewing parameters are used to solve for the surface parameters p, q and r. To accomplish this secoond step, the usual planar relation, $Z = \frac{1-px-qy}{r}$ is used. In order to asses the effects of noise on the recovery method the input horizontal and vertical disparity measures have been systematically corrupted by error. Although it is common to conduct such studies by adding noise as some percentage of the "true" data value, this is not the tack taken here. It does not seem that noise proportional to data is a good model of how noise is likely to enter into this set of computations. Rather, it seems that noise should be of a similar magnitude as applied to the relatively large values of horizontal disparity as to the relatively large values of vertical disparity. That is to say, there is no reason to believe that a system should have better vertical than horizontal acuity. Therefore, the numerical experiments reported here apply

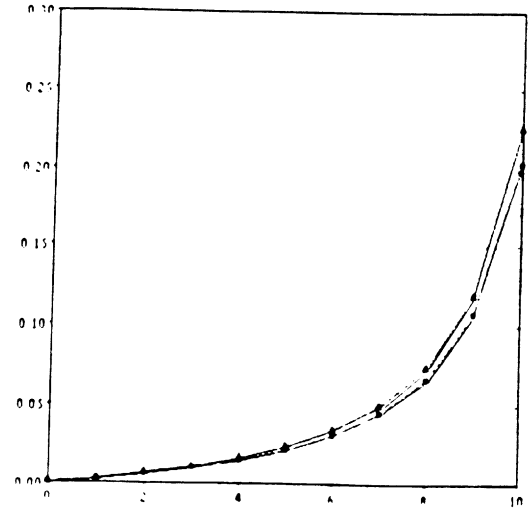
noise of equal magnitude to all disparity values.

The results of two numerical experiments are now reported. For these cases the noise to the disparity values has been incremented in a linear fashion. The error is reported as error in the recovered surface parameters p, q and r. The error is reported as a percentage of the baseline error. For the sake of comparison, the results of the same testing of the method proposed in Chapter 2 are also presented. Recall that the method presented in Chapter 2 recovers view and surface geometry while using only horizontal and orientational disparity. Since this method is restricted from using vertical disparity it will be referred to as the restricted perspective method.

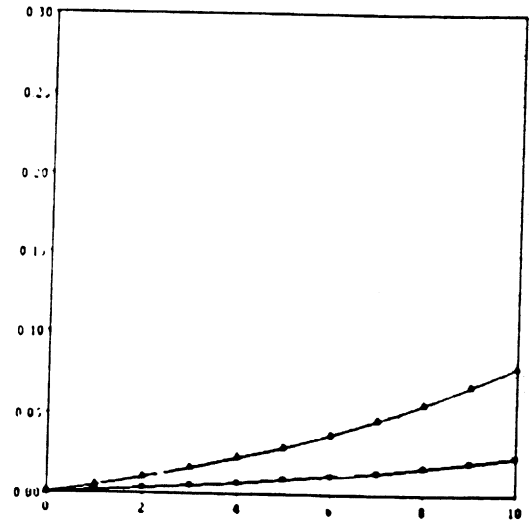
For the first experiment the simulated viewer is fixated at a point on a planar surface 50 cm. away. The view is $\frac{\pi}{8}$ radians off center and the surface makes an angle of $\frac{\pi}{4}$ radians with respect to the line of regard. The results of the full and restricted perspective methods are shown in Figures A.1.a and A.1.b, respectively. Begin by considering the results of the experiment as applied to the full perspective recovery method. Two observations can be made. First, the error trend is of higher than linear order. Second, the experiment rapidly reaches a point where the error in the computation becomes very large. In contrast, the method of restricted perspective shows error that increases approximately linearly with input noise to the data. For the second experiment the simulated viewer is fixating a point on a planar surface 200 cm. away. The view is 0 radians off center and the surface makes an angle of $\frac{15\pi}{16}$ radians with respect to the line of regard. The results for the full and restricted perspective method are shown in Figures A.2.a and A.2.b, respectively. The results are seen to be quite similar to those of experiment 1. Again, the method of full

perspective leads to rather unstable solutions; the method of restricted perspective is relatively stable.

The result that the restricted perspective method exhibits good numerical stability in these empirical tests should come as no surprize. A formal analysis was presented in Chapter 2 that indicated the stability of this system. The instability of the full perspective method can be explained as follows: It is a basic result of numerical analysis that the most stable systems of equations are those whose coefficients all have roughly the same magnitude. However, the vertical disparities used in the full perspective method are of much smaller magnitude than either the horizontal or orientational disparities, this leads to very unstable behavior. From the results of these experiments it is concluded that the method of full perspective will not lead to algorithms that are able to recover local surface geometry in a robust fashion.

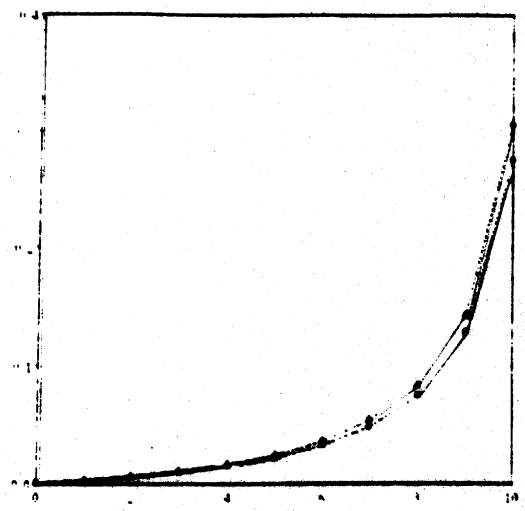


(a)

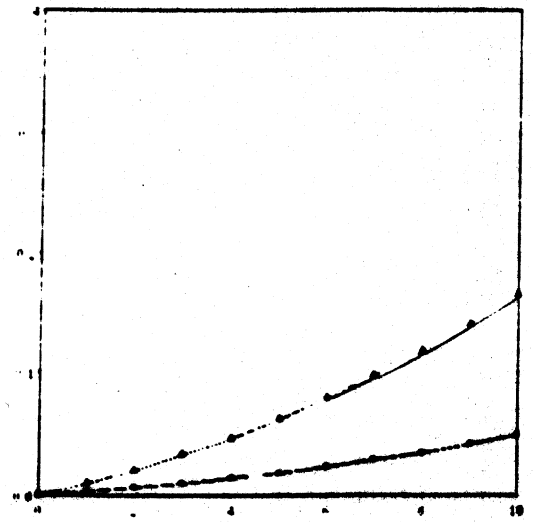


(b)

Figure A.1: Results of an empirical numerical study. (a) The full perspective method leads to error that grows rapidly as noise is added to the input data. (b) The restricted perspective method demonstrates relative stability. The horizontal axis is marked in units of noise in vertical disparity units. The vertical axis shows percent error in the surface parameters: triangles symbolize orientation while squares symbolize relative depth.



(a)



(b)

Figure A.2: Results of an empirical numerical study. (a) The full perspective method leads to error that grows rapidly as noise is added to the input data. (b) The restricted perspective method demonstrates relative stability. The horizontal axis is marked in units of noise in vertical disparity units. The vertical axis shows percent error in the surface parameters: triangles symbolize orientation while squares symbolize relative depth.

Appendix B

The decomposition of discontinuous disparity fields

In general, discontinuities of surfaces in the world will differentially project into discontinuities in a disparity field. If the disparity information is very dense, it may be possible to detect these discontinuities directly in the disparity field. Along these lines several researchers have proposed applying edge-detection techniques to both stereo (Stevens [118]) and motion (Thompson, Mutch & Berzins [124], Schunk [111]) based disparity fields. In this regard one needs to decide how to perform edge-detection in a vector field. In chapter 1, the work of Thompson, Mutch and Berzins [124] was given as an example. Recall that these researchers looked for the edges in the separate x and y scalar components of the vector field and then combined the results. In this appendix, another pair of scalar fields are noted to be useful for representing vector field discontinuities. In particular, it is shown that the divergence and rotational fields of a two-dimensional vector field capture the discontinuites of the original field. A

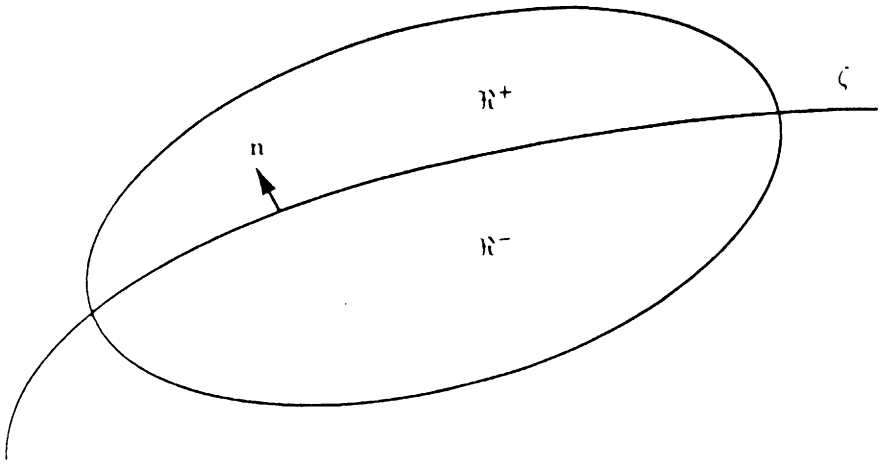


Figure B.1: The field χ in the neighborhood of a discontinuity.

representation in terms of divergence and rotational has the nice property of having coordinate system independent geometric interpretations. The divergence captures the local degree of expansion; the rotational gives a local measure of rotation.

Before presenting the analysis it is necessary to introduce some terminology as well as a classical result of Hadamard [45]. Consider the field $\chi(x, y)$ in the neighborhood of a discontinuity as depicted in Figure B.1. Assume that χ is continuous in the regions $\Re^+$ and $\Re^-$, but discontinuous on the boundary curve, ζ. Let $\mathbf{n}$ be the normal to ζ (a function of position along ζ) and take its positive sense as pointing into $\Re^+$. Further, assume that χ approaches definite limiting values as it approaches either side of ζ. Denote the limiting values of χ from $\Re^+$ and $\Re^-$ as χ^+ and χ^-, respectively. Also, denote the jump across ζ as $\chi^+ - \chi^- = [\chi]$. Finally, inorder to justify the ensuing calculations the following result is required:

Lemma B.1 (Hadamard) *Let χ be defined and continuously differentiable in the interior of a region $\Re^+$ with smooth boundary ζ, and let χ and $\partial_i\chi$ approach finite limits χ^+ and $\partial_i\chi^+$ as ζ is approached upon paths interior to $\Re^+$. Let $\mathbf{x} = \mathbf{x}(\ell)$ be a*

smooth curve upon ζ, and assume that χ^+ is differentiable on this path. Then

$$\frac{d\chi^+}{d\ell} = \partial_i \chi^+ \frac{dx^i}{d\ell}.$$

In essence, the lemma states that the theorem of the total differential (see, e.g. Korn & Korn [64]) holds as ζ is approached from one side only. The reader is referred to Hadamard [45] for a proof.

With the considerations of the previous paragraph in hand it easy to show that: (i) The normal component of $[\chi]$ is preserved in the divergence of χ, $\nabla \cdot \chi$. (ii) The transverse component of $[\chi]$ is preserved in the rotational of χ, $\nabla \times \chi$. To show that these claims are true, consider a small area, $\mathcal{A}$, of χ centered about a point along ζ. For the first assertion, recall that the two-dimensional Divergence Theorem (see, e.g., Korn & Korn [64]) states that for a vector field χ

$$\oint_{\partial \mathcal{A}} \mathbf{N} \cdot \chi = \int\int_{\mathcal{A}} \nabla \cdot \chi \tag{B.1}$$

with $\partial \mathcal{A}$ the boundary of $\mathcal{A}$ and $\mathbf{N}$ the normal along this boundary. Hadamard's Lemma allows the evaluation of the integrals in (B.1) to proceed independently in the regions $\Re^+$ and $\Re^-$. As $\mathcal{A}$ tends to an infinitesimal area it is found that

$$\nabla \cdot \chi = \mathbf{n} \cdot [\chi]. \tag{B.2}$$

Thus, the normal component of the jump $[\chi]$ is present in the divergence $\nabla \cdot \chi$. Similarly for the second assertion, recall that Green's Theorem (Korn & Korn [64]) can be stated:

$$\oint_{\partial \mathcal{A}} \chi = \int\int_{\mathcal{A}} \nabla \times \chi. \tag{B.3}$$

Again appealing to Hadamard's Lemma and allowing the region of integration to become vanishingly small shows that (B.3) evaluates to

$$\nabla \times \chi = \mathbf{n} \times [\chi]. \tag{B.4}$$

Relation (B.4) has established the second of the desired results: The transverse component of $[\chi]$ is present in the rotational $\nabla \times \chi$.

In this appendix it has been shown that the normal and tangential jumps of a discontinuous two-dimensional vector field are preserved in the divergence and rotational of the field. It is suggested that this representation may prove useful for further investigations aimed at recovering the discontinuities of visual disparity fields. The appeal of this representation is based in its coordinate system independent geometric interpretations. The ultimate usefulness of this analysis may be limited by the ability to recover the divergence and rotational components of χ for the case of visual disparity.

Appendix C

Surface curvature from disparity

This appendix provides preliminary results on the recovery of three-dimensional surface curvature from stereo disparity. In particular, it will be shown that the differentially imaged curvature of surface markings can be used to recover three-dimensional surface curvature. For the purposes of these developments several simplifying assumptions will be exploited: First, it is assumed that the optical axes are pointed straight ahead and parallel to one another. Thus, retaining the nomenclature of earlier developments, $\mathbf{\Omega} = (\omega_x, \omega_y, \omega_z) = (0,0,0)$ and $\mathbf{T} = (t_x, t_y, t_z) = (I,0,0)$, where I is the stereo baseline. Under these conditions the basic disparity relations reduce to

$$\chi = (\chi_x, \chi_y) = \left(\frac{I}{Z}, 0\right). \tag{C.1}$$

The second set of assumptions deal with the geometry of the viewed surface: The analysis focuses on the differential projection of surface detail (e.g., contours, texture or markings on the surface) in the neighborhood of the fixation point, image coordinates $(x,y) = (0,0)$. Further, it is assumed that the surface normal at this

point of regard is alligned with the optical axis. Thus, the surface gradient has zero magnitude, $\bigtriangledown Z = (0,0)$. Then, given that the surface is curved, it can be locally represented in world coordinates (X, Y, Z) in terms of a Taylor series evaluated at the origin

$$Z = r + \kappa_{xy}XY + \frac{1}{2}\kappa_{xx}X^2 + \frac{1}{2}\kappa_{yy}Y^2 \tag{C.2}$$

where r is the radial distance from the viewer while

$$\kappa_{xy} = \frac{\partial^2 Z(0,0)}{\partial X \partial Y}, \quad \kappa_{xx} = \frac{\partial^2 Z(0,0)}{\partial X^2}, \quad \kappa_{yy} = \frac{\partial^2 Z(0,0)}{\partial Y^2}$$

are surface curvature terms. In image coordinates (C.2) becomes

$$\frac{1}{Z} = \frac{1}{r} - \kappa_{xy}xy - \frac{1}{2}\kappa_{xx}x^2 - \frac{1}{2}\kappa_{yy}y^2. \tag{C.3}$$

For the remainder of this appendix consideration will be limited to those restricted viewing and surface geometrys that are embodied in (C.1) and (C.3).

Attention is now directed to an analysis of the differentially projected curvature of surface markings. Consider the case where a point along a curved surface contour is fixated. In the left image coordinate system let this contour be described as a biquadratic that passes through the origin

$$0 = y + ax + bxy + cy^2 + dx^2. \tag{C.4}$$

In order to simplify the calculation of the imaged curvature, consider a rotation of the image coordinate system that aligns the tangent at the origin with the x-axis. Let the new coordinate system (u,v) be related to the old by

$$\begin{aligned} x &= u\cos\alpha - v\sin\alpha \\ y &= u\sin\alpha + v\cos\alpha \end{aligned} \tag{C.5}$$

with α the counterclockwise angle of rotation. Substituting into (C.4) and rearranging in terms of u and v yields

$$0 = (a\cos\alpha + \sin\alpha)u + (\cos\alpha - a\sin\alpha)v + [b(\cos^2\alpha - \sin^2\alpha) + 2(c-d)\sin\alpha\cos\alpha]uv$$
$$+(b\sin\alpha\cos\alpha + d\cos^2\alpha + c\sin^2\alpha)u^2 + (-b\sin\alpha\cos\alpha + d\sin^2\alpha + c\cos^2\alpha)v^2. \tag{C.6}$$

The desired rotation requires that

$$0 = (a\cos\alpha + \sin\alpha)$$

or

$$-a = \tan\alpha. \tag{C.7}$$

Now, in this new coordinate system, the imaged curvature at the point of fixation, $(x, y) = (0, 0)$, can be conveniently calculated from (C.6) by evaluating the curvature formula

$$\kappa = \frac{\frac{d^2y}{dx^2}}{\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{\frac{3}{2}}} \tag{C.8}$$

and computing the required derivatives implicitly. Some amount of calculation shows that

$$\kappa = 2\cos\alpha(b\sin\alpha\cos\alpha + d\cos^2\alpha + c\sin^2\alpha) \tag{C.9}$$

is the resulting imaged curvature.

Now, consider how the contour (C.4) appears in the other image. In order to understand this transformation it is useful to adopt an Eulerian viewpoint (Goldstein [36]): Consider a new coordinate system (μ, ν) that is the same as the old (x, y) system except now the contour is viewed after it has been deformed by the operations

of differential projection (c.f., Waxman & Wohn [132] where the Eulerian viewpoint is exploited to analyze motion parallax).[1] The relation between points (x, y) on the original contour (C.4) and points (μ, ν) on the deformed contour is specified by

$$(\mu, \nu) = (x + \Delta_x(x, y), y + \Delta_y(x, y)) \tag{C.10}$$

where $\mathbf{\Delta} = (\Delta_x, \Delta_y)$ specifies the operation of disparate projection. For current purposes it is convenient to represent $\mathbf{\Delta}$ in terms of a Taylor series expansion of disparity. Thus,

$$\begin{aligned} \Delta_x &= \chi_x + \tfrac{\partial \chi_x}{\partial x} x + \tfrac{\partial \chi_x}{\partial y} y + \tfrac{\partial^2 \chi_x}{\partial x \partial y} xy + \tfrac{1}{2} \tfrac{\partial^2 \chi_x}{\partial x^2} x^2 + \tfrac{1}{2} \tfrac{\partial^2 \chi_x}{\partial y^2} y^2 + O^2 \\ \Delta_y &= \chi_y + \tfrac{\partial \chi_y}{\partial x} x + \tfrac{\partial \chi_y}{\partial y} y + \tfrac{\partial^2 \chi_y}{\partial x \partial y} xy + \tfrac{1}{2} \tfrac{\partial^2 \chi_y}{\partial x^2} x^2 + \tfrac{1}{2} \tfrac{\partial^2 \chi_y}{\partial y^2} y^2 + O^2 \end{aligned} \tag{C.11}$$

where O^2 represent terms that involve second and higher powers of the stereo baseline, I.

This specialization of the Eulerian viewpoint analysis can now be used to understand how the contour (C.4) deforms between the two stereo images. Applying (C.10) to (C.4) allows equating

$$\nu + a\mu + b\mu\nu + c\nu^2 + d\mu^2 \tag{C.12}$$

and

$$(y + \Delta_y) + a(x + \Delta_x) + b(xy + x\Delta_y + y\Delta_x) + c(y^2 + 2y\Delta_y) + d(x^2 + 2x\Delta_x) + O^2. \tag{C.13}$$

Also, recalling the original form of (C.4) shows that (C.13) can be reduced to

$$\Delta_y + a\Delta_x + b(x\Delta_y + y\Delta_x) + 2cy\Delta_y + 2dx\Delta_x + O^2. \tag{C.14}$$

[1] This type of analysis was originally developed as a tool for understanding fluid flow. In that case the variable separating views is time and the deforming objects are patches of flow.

Equations (C.12) and (C.14) can be combined into a more useful form by noticing that to O^2: $x\Delta_x = \mu\Delta_x$, $y\Delta_y = \nu\Delta_y$, $x\Delta_y = \mu\Delta_y$ and $y\Delta_x = \nu\Delta_x$. Making these substitutions in (C.14) and combining the results with (C.12) yields

$$0 = (\nu - \Delta_y) + a(\nu - \Delta_x) + b(\mu\nu - \mu\Delta_y - \nu\Delta_x) + c(\nu^2 - 2c\nu\Delta_y) + d(\mu^2 - 2d\mu\Delta_x). \quad \text{(C.15)}$$

The final steps in developing the differentially projected version of (C.4) are to: (i) regroup (C.15) in terms of μ and ν and (ii) evaluate the terms of (C.11) in light of the viewing and surface geometrys (C.1) and (C.3). To second-order the resulting contour is

$$0 = \nu + a\mu + (b - aI\kappa_{xy})\mu\nu + (c - aI\kappa_{yy})\nu^2 + (d - aI\kappa_{xx})\mu^2. \quad \text{(C.16)}$$

The imaged curvature of (C.16) can be evaluated at the fixation $(\mu, \nu) = (0, 0)$ in the same manner used for the original contour (C.4). For the sake of brevity the procedure is simply stated and followed by the result: First, rotate the (μ, ν) coordinate system so that the μ-axis is aligned with the tangent to (C.16) at $(0, 0)$. Second, use the curvature formula (C.8) to evaluate curavature in the rotated system. Following through on the prescribed operations shows that the new curvature measure is

$$\kappa' = 2\cos\alpha[(b - aI\kappa_{xy})\sin\alpha\cos\alpha + (d - aI\kappa_{xx})\cos^2\alpha + (c - aI\kappa_{yy})\sin^2\alpha] \quad \text{(C.17)}$$

where, as earlier, $\tan\alpha = -a$.

At this point image curvature measures have been derived for two different projections of a contour on a three-dimensionally curved surface. These measures are given by relations (C.9) and (C.17). However, recall that the original goal was to develop an

analysis that showed how to use differentially projected curvature to recover surface curvature. This goal is in hand: Subtracting (C.9) from (C.17) yields the following pleasing result

$$\kappa' - \kappa = -2I \sin\alpha(\kappa_{xy} \sin\alpha \cos\alpha - \kappa_{xx} \cos^2\alpha - \kappa_{yy} \sin^2\alpha). \qquad (C.18)$$

Under the assumptions that the stereo baseline, I, is known and the tangent to the projected contour is measurable in the image, relation (C.18) provides one equation in the three unknown surface curvatures, κ_{xy}, κ_{xx} and κ_{yy}. Provided three differentially projected surface curves can be so measured the surface curvatures can be recovered.

In summary, this appendix has provided an analysis of how it is possible to recover three-dimensional surface curvatures from two-dimensional differentially imaged curvatures. The analysis has considered only the restricted case where a stereoscopic viewer is looking straight ahead with the optical axes parallel. Further, it has been assumed that the view of the surface is along the surface normal. Under these conditions it has been shown that it is (theoretically) possible to recover surface curvature by observing three differentially projected surface contours.

Appendix D

Extension to motion based disparity

In the main body of this thesis the techniques of vector and tensor analysis were exploited to develop an understanding of binocular stereo disparity. Infact, these ideas can be extended to any area of visual information processing where the input representation can be characterized as a vector field. An obvious candidate for such an analysis is the interpretation of disparity due to motion parallax. This appendix will sketch the extension of the stereo disparity analysis to the case of motion parallax.

Recall that a general infinitesimal change in coordinate systems changes the coordinates of a point $\mathbf{R}$ by

$$\delta\mathbf{R} = -\mathbf{T} - (\boldsymbol{\Omega} \times \mathbf{R}) \tag{D.1}$$

where the symbols are defined with reference to Figure 2.1. For the case of stereo vision it was possible to equate some of the components of $\mathbf{T}$ and $\boldsymbol{\Omega}$ to zero and thus

simplify the ensuing derivations. For the case of general motion of a viewer in an otherwise stationary environment this type of simplification is not allowed. Following throught the derivations of Section 2.1, but now allowing for the full generallity of (D.1) leads to

$$\chi = (\chi_x, \chi_y) = \left(-\frac{t_x}{Z} - \omega_y + \omega_z y - x(-\frac{t_z}{Z} - \omega_x y + \omega_y x), -\frac{t_y}{Z} + \omega_x - \omega_z x - y(-\frac{t_z}{Z} - \omega_x y + \omega_y x)\right) \tag{D.2}$$

as the motion parallax disparity relations.

It is also a straight forward matter to derive the gradient tensor of disparity

$$\chi' = \begin{pmatrix} \frac{\partial \chi_x}{\partial x} & \frac{\partial \chi_x}{\partial y} \\ \frac{\partial \chi_y}{\partial x} & \frac{\partial \chi_y}{\partial y} \end{pmatrix}. \tag{D.3}$$

For the case of a planar surface patch projecting along the line of regard (D.3) evaluates to

$$\chi' = \begin{pmatrix} \frac{1}{r}(pt_x + t_z) & \frac{qt_x}{r} + \omega_z \\ \frac{1}{r}(qt_y + t_z) & \frac{pt_y}{r} - \omega_z \end{pmatrix}. \tag{D.4}$$

As earlier, in order to gain greater insight into χ' it is useful to split it into symmetric, χ'_+, and antisymmetric, χ'_-, parts. This yields

$$\chi' = \chi'_+ + \chi'_- = \frac{1}{2}\begin{pmatrix} \frac{2}{r}(pt_x + t_z) & \frac{pt_y + qt_x}{r} \\ \frac{pt_y + qt_x}{r} & 2\frac{qt_y + t_z}{r} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & \frac{qt_x - pt_y}{r} + \omega_z \\ \frac{pt_y - qt_x}{r} - 2\omega_z & 0 \end{pmatrix}.$$

(Recall that χ'_- describes the rigid rotation that an object undergoes as it is differentially projected; while χ'_+ describes a nonrigid deformation.) The axis and magnitude of the nonrigid operation of χ'_+ can be recovered via an eigen-decomposition. In this case it is found that the axis of contraction (an eigenvector) is in the direction specified

by

$$\xi = \frac{\|(p,q)\|(t_x,t_y) + \|(t_x,t_y)\|(p,q)}{\|(p,q)\|\|(t_x,t_y)\|}$$

while the magnitude of deformation (i.e., the magnitude of the difference of the eigenvalues) is

$$\sigma = \frac{\|(t_x,t_y)\|\|(p,q)\|}{r}.$$

It is worth noting that this description of the motion parallax field is similar to that first presented in Koenderink & van Doorn [58].

The relations derived in this appendix for motion parallax parallel those derived for binocular stereo derived in Chapter 2. It has again been possible to express the disparity information in terms of parameters of differential viewing, $\mathbf{T}$ and Ω, and first-order surface geometry, p, q and r. With these basic relations in hand further study will be able to indicate how to invert the disparity information for the recovery of the geometric parameters of interest.

Bibliography

[1] Albert, A. (1972) *Regression and the Moore-Penrose Pseudoinverse.* New York, New York: Academic Press.

[2] Anandan, P. (1987) A unified perspective on computational techniques for the measurement of visual motion. In *Proc. ICCV*, 219-230.

[3] Anstis, S. M., Howard, I. P. & Rogers, B. (1978) A Craik-O'Brien-Cornsweet illusion for visual depth. *Vis. Res. 18*, 213-217.

[4] Arditi, A. (1982) The dependence of the induced effect on orientation. *Vis. Res. 22*, 247-256.

[5] Arditi, A., Kaufman, L. & Movshon, J. A. (1981) A simple explanation of the induced size effect. *Vis. Res. 21*, 755-764.

[6] Aris, R, (1961) *Vectors, Tensors, and the Basic Equations of Fluid Mechanics.* Englewood Cliffs, New Jersey: Prentice Hall, Inc.

[7] Barlow, H. B., Blakemore, C. & Pettigrew, J. D. (1967) The neural mechanism of binocular depth discrimination. *J. Physiol. 193*, 327-342.

[8] Barnard, S. T. & Fischler, M. A. (1982) Computational stereo. *Computing Surveys 14 (4)*, 553-572.

[9] Barron, J. L., Jepson, A. D. & Tsotsos, J. K. (1985) The sensitivity of motion and structure computations. In *Proc IJCAI*, 700-705.

[10] Bishop, P. O. (1978) Orientation and position disparities in stereopsis. In S. J. Cool & E. L. Smith (Eds.) *Frontiers in Visual Science*, 366-350. New York, New York: Springer Verlag.

[11] Blake, A. (1984) Reconstructing a visible surface. In *Proc AAAI*, 23-26.

[12] Blake, A. & Zisserman, A. (1987) *Visual Reconstruction.* Cambridge, Massachusetts: MIT Press.

[13] Blakemore, C., Fiorentini, A. & Maffei, L. (1972) A second neural mechanism of binocular depth discrimination. *J. Physiol. 226*, 725-749.

[14] Blostein, S. D. & Huang, T. S. (1987) Error analysis in stereo determination of 3-D point positions. *IEEE-PAMI 9 (6)*, 752-765.

[15] Blostein, S. D. & Huang, T. S. (1988) Correction to "Error analysis in stereo determination of 3-D point positions. *IEEE-PAMI 10 (5)*, 765.

[16] Boult, T. E. (1987) *Reproducing Kernels for Surface Interpolation.* Columbia University AI memo, New York, New York.

[17] Brady, J. M., Ponce, J., Yuille, A. & Asada, H. (1985) *Describing Surfaces.* MIT AI Memo No. 822, Cambridge, Massachusetts.

[18] Canny, J. F. (1983) *Finding Lines and Edges in Images.* MIT AI Tech. Rep. No. 720. Cambridge, Massachusetts.

[19] Cayley, A. (1859) On contour lines and slope. *The London, Edinburgh & Dublin Philosophical Magazine & J. of Science 18 (120)*, 264-268.

[20] Clocksin, W. F. (1980) Perception of surface slant and edge labels from optical flow. *Perception 9*, 253-269.

[21] Dahlquist, G. & Björk, Å(1974) *Numerical Methods.* Englewood Cliffs, New Jersey: Prentice Hall.

[22] De Valois, K. K., von der Heydt, R., Adorjani, C. S. & De Valois, R. L. (1975) A tilt aftereffect in depth. *Assoc. Res. Vis. Ophthal. 15*, 90.

[23] Dixon, W. J. (1954) Power of several nonparametric tests. *Ann. Math. Statis., 25,* 610-614.

[24] Duwaer, A. L. & van der Brink, G. (1982) Detection of vertical disparities. *Vis. Res. 22*, 467-478.

[25] Eastman, R. D. & Waxman, A. M. (1987) Using disparity functionals for stereo correspondence and surface reconstruction. *CVGIP 39*, 73-101.

[26] Ferster, D. (1981) A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *J. of Physiol. 311*, 623-633.

[27] Fischer, B. Poggio, G. F. (1979) Disparity selectivity in cortical neurons. *Proc. Roy. Soc. Lond. B 204*, 409-419.

[28] Fox, R., Cormack, L. & Norman, F. (1987) The effect of vertical disparity on depth scaling. *Invest. Ophthal. & Vis. Sci. Supl. 28 (3)*, 293.

[29] Foley, J. M. (1980) Binocular distance perception. *Psych. Rev. 87 (5)*, 411-434.

[30] Foley, J. M. & Richards, W. A. (1972) Effects of voluntary eye movements and convergence on the binocular appreciation of depth. *P&P 11*, 423-427.

[31] Gillam, B. (1968) Perception of slant when stereopsis & perspective conflict. *J. Exp. Psych. 72*, 299-305.

[32] Gillam, B. & Lawergren, B. (1983) The induced effect, vertical disparity and stereoscopic theory. *P & P 34*, 121-130.

[33] Gillam, B., Flagg, T. & Finlay, D. (1984) Evidence for disparity change as the primary stimulus for sterepscopic processing. *P & P, 36 (6)*, 559-564.

[34] Gillam, B., Chambers, D.& Russo, T. (1988) Postfusional latency in stereoscopic perception and the primitives of stereopsis. *JEPHP&P 14 (2)*, 163-175.

[35] Gillam, B., Chambers, D. & Lawergren, B. (1988) The role of vertical disparity in the scaling of stereoscopic depth perception: An empirical study and theoretical study. *P & P 44 (5)*, 473-483.

[36] Goldstein, H. (1980) *Classical Mechanics, 2nd Edition.* Reading, Massachusetts: Addison-Wesley.

[37] Grimson, W. E. L. (1981) A computer implementation of a theory of human stereo vision. *Phil. Trans. Roy. Soc. Lond., B. 292*, 217-253.

[38] Grimson, W. E. L. (1981) *From Images to Surfaces.* Cambridge, Massachusetts: MIT Press.

[39] Grimson, W. E. L. (1982) A computational theory of visual surface interpolation. *Phil. Trans. Roy. Soc. Lond. B 298*, 395-427.

[40] Grimson, W. E. L. (1983) The implicit constraints in the primal sketch. *CVGIP 24*, 28-51.

[41] Grimson, W. E. L. (1986) Personal Communication.

[42] Grimson, W. E. L. (1987) Errors of position as a function of errors in measurement. *Robotics and Automation*

[43] Grimson, W. E. L. & Pavlidis, T. (1985) Discontinuity detection for visual surface reconstruction. *CVGIP 30*, 316-330.

[44] Gulick, W. L. & Lawson, R. B. (1976) *Human Stereopsis: A Psychophysical Analysis.* New York, New York: Oxford University Press.

[45] Hadamard, J. (1949) *Leçons sur la Propagation des Ondes et les Équations de l'Hydrodynamique.* New York, New York: Chelsea.

[46] Hänny, P., von der Heydt, R. & Poggio, G. (1980) Binocular neuron responses to tilt in depth in the monkey visual cortex: Evidence for orientation disparity processing. *Exp. Brain Res. 41*, A26.

[47] Harris, J. G. (1987) The coupled depth slope approach to surface reconstruction. In *Proc. ICCV*, 277-283.

[48] Helmholtz, H. von (1910) *Handbook of Physiological Optics.* New York, New York: Dover.

[49] Hoff, W. & Ahuja, N. (1987) Extracting surfaces from stereo images. In *Proc. ICCV*, 284-294.

[50] Horn, B. K. P. (1987a) Closed-form solution of absolute orientation using unit quaternions. *JOSA A 4(4)*, 629-642.

[51] Horn, B. K. P. (1987b) *Relative Orientation* MIT AI Memo No. 994, Cambridge, Massachusetts.

[52] Hubel, D. H. & Wiesel, T. N. (1970) Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature 225*, 41-42.

[53] Julesz, B. (1971) *Foundations of Cyclopean Perception.* Chicago, Illinois: University of Chigo Press.

[54] Kanatani, K. (1985) Structure from motion without correspondence: General principle. In *Proc. IJCAI*, 886-888.

[55] Kaplan, G. A. (1969) Kinetic disruption of optical texture: The perception of depth at an edge. *P&P 6 (4)*, 193-198.

[56] Kaufman, L. (1965) Some new stereoscopic phenomena and their implications fot theories of stereopsis. *Amer. J. Psychol. 78*, 1-20.

[57] Koch, C., Marroquin, J. & Yuille, A. (1985) *Analog "Neural" Networks in Early Vision.* MIT AI Memo No. 751, Cambridge, Massachusetts.

[58] Koenderink, J. J. & van Doorn, A. J. (1975) Invariant properties of the motion parallax field due to the mvement of rigid bodies relative to an observer. *Optica Acta 22*, 773-791.

[59] Koenderink, J. J. & van Doorn, A. J. (1976) Geometry of binocular vision. *Biol. Cyb. 21*, 29-35.

[60] Koenderink, J. J. & van Doorn, A. J. (1976) Local structure of movement parallax of the plane. *J. Opt. Soc. Am. 66 (7)*, 717-723.

[61] Koenderink, J. J. & van Doorn, A. J. (1987) Facts on optic flow. *Biological Cybernetics 56*, 247-254.

[62] Koenderink, J. J. & Richards, W. A. (1988) Two-dimensional curvature operators. *JOSA A 5 (7)*, 1136-1141.

[63] Koopmans, T. (1937) *Linear Regression Analysis of Economic Time Series.* The Netherlands: DeEverven F. Bohm.

[64] Korn, G. A. & Korn, T. M. (1961) *Mathematical Handbook for Scientists and Engineers.* New York, New York: McGraw-Hill.

[65] Lawson, R. B. & Gulick, W. L. (1967) Stereopsis and anomolous contour. *Vis. Res. 7*, 271-297.

[66] Lee, D. & Pavlidis, T. (1987) One-dimensional regularization with discontinuities. In *Proc. ICCV*, 572-577.

[67] Leibovic, K. N., Balslev, E. & Mathieson, T. A. (1971) Binocular visionn and pattern recognition. *Kybernetik 8*, 14-23.

[68] Longuet-Higgins, H. C. (1981) A computer algorithm for reconstructing a scene from two projections. *Nature 293*, 133-135.

[69] Longuet-Higgins, H. C. (1982) The role of the vertical dimension in stereoscopic vision. *Perception 11*, 377-386.

[70] Longuet-Higgins, H. C. & Pradzny, K. (1980) The interpretation of a moving retinal image. *Proc. Roy. Soc. Lond. B 208*, 385-397.

[71] *Manual of Photogrammetry* (1966) Washington, D. C.: Amer. Soc. Photogram. (no author or editor named).

[72] Marr, D. *VISION: A Computational Investigation in the Representation and Processing of Visual Information.* San Francisco, California: W. H. Freeman.

[73] Marr, D. & Hildreth, E. C. (1980) Theory of edge detection. *Proc. Roy. Soc. Lond. B 207*, 187-217.

[74] Marr, D. & Poggio, T. (1979) A theory of human stereo vision. *Proc. Roy. Soc. Lond. B 204*, 301-328.

[75] Marroquin, J. L. (1984) *Surface Reconstruction Preserving Discontinuities.* MIT AI Memo No.792, Cambridge, Massachusetts.

[76] Massey, F. J. (1951) The distribution of the maximum deviation between two sample cumulative step functions. *Ann. Math. Statist., 22,* 125-128.

[77] Maxwell, J. C. (1870) On hills and dales. *The London, Edinburgh & Dublin Philosophical Magazine & J. of Science 40 (269)*, 421-425.

[78] Mayhew, J. E. W. (1982) The interpretation of stereo disparity information. *Percpetion 11*, 387-403.

[79] Mayhew, J. E. W. & Frisby, J. P. (1982) The induced effect: Arguments against the theory of Arditi, Kaufman and Movshon. *Vis. Res. 22*, 1225-1228.

[80] Mayhew, J. E. W. & Longuet-Higgins, H. C. (1982) A computational model of binocular depth perception. *Nature 297*, 376-378.

[81] McVey, E. S. & Lee, J. W. (1982) Some accuracy and resolution aspects of computer vision distance measurements. *IEEE-PAMI 4 (6)*, 646-649.

[82] Medioni, G. & Nevatia, R. (1984) Description of three-dimensional surfaces using curvature properties. In *Proc. DARPA IU Workshop*, 291-229.

[83] Mitchison, G. (1987) *Planarity and Segmentation in Stereoscopic Matching.* MIT CBIP Memo No. 23, Cambridge, Massachusetts.

[84] Mitchison, G. J. & Westheimer, G. (1984) The perception of depth in simple figures. *Vis. Res. 24 (9)*, 1063-1073.

[85] Mitchison, G. J. & McKee, S. P. (1985) Interpolation in stereoscopic matching. *Nature 315*, 402-404.

[86] Mitchison, G. J. & McKee, S. P. (1987a) Interpolation and the detection of fine structure in stereoscopic matching. *Vis. Res. 27*, 295-302.

[87] Mitchison, G. J. & McKee, S. P. (1987b) The resolution of ambiguous stereoscopic matches by interpolation. *Vis. Res. 27*, 285-294.

[88] Mohan, R., Medioni, G. & Nevatia, R. (1989) Stereo error detection, correction and evaluation. *IEEE-PAMI 11 (2)*, 113-1210.

[89] Moravec, H. (1979) Visual mapping by a robot rover. In *Proc. IJCAI* 598-600.

[90] Mutch, K. M. & Thompson, W. B. (1985) Analysis of accretion and deletion at boundaries in dynamic scenes. *IEEE-PAMI 7 (2)*, 133-138.

[91] Nelson, J. I., Kato, H. & Bishop, P. O. (1977) Discrimination of orientation and position disparity by binocularly activated neurons in cat striate cortex. *J. Neurophsiol. 40*, 260-283.

[92] Nielsen, K. R. & Poggio, T. (1984) Vertical image registration in stereopsis. *Vis. Res. 24*, 1133-1140.

[93] Ninio, J. (1985) Orientational versus horizontal disparity in the stereoscopic appreciation of slant. *Perception 14*, 305-314.

[94] Ninio, J. (1988) Manuscript submitted to *Vis. Res.*

[95] Nishihara, H. K. (1984) Practical real-time imaging stereo matcher. *Optical Engineering 23 (5)*, 536-545.

[96] Ogle, K. N. (1938) Induced size effect I: A new phenomenon in binocular vision associated with the relative sizes of the images in the two eyes. *Arch. of Ophthal. 20*, 604.

[97] Ogle, K. N. (1950) *Researches in Binocular Vision.* Philadelphia, Pennsylvania: Saunders.

[98] Ohzawa, I. & Freeman, R. D. (1986) The binocular organization of simple cells in the cat's visual cortex. *J. Neurophys. 56 (1)*, 221-242.

[99] Oppenheim, A. V., et al. (1968) Nonlinear filtering of multiplied and convolved signals. *Proc. IEEE 56*, 1264-1291.

[100] Poggio, G. (1988) Binocular sensitivity of striate cortical cells. In *Proc. Neurosci. Soc.*, 1072.

[101] Poggio, G. & Fischer, B. (1977) Binocular interaction & depth sensitivity of striate & prestriate cortical neurons of the behaving rhesus monkey. *J. Neurophysiol. 40*, 1392-1405.

[102] Pradzny, K. (1980) Egomotion and relative depth map from optical flow. *Biol. Cyb., 36,* 87-102.

[103] Prakash, N. (1981) *Differential Geometry: An Integrated Approach.* New York, New York: McGraw-Hill.

[104] Richards, W. A. (1968) Spatial remapping in the primate visual system. *Kybernetik 4 (4)*, 146-156.

[105] Richards, W. A. (1971) Anomolous stereoscopic depth perception. *JOSA 61*, 410-414.

[106] Richards, W. A. (1977) Stereopsis with and without monocular cues. *Vis. Res. 17*, 967-969.

[107] Richards, W. A. (1988b) *Selections in Natural Computation.* Cambridge, Massachusetts: MIT Press.

[108] Rigaudiere, F. (1975) Fusion binoculaire et localisation spatiale de mires verticles et horizontales de frequences spatiales differentes. *Vis. Res. 15*, 931-938.

[109] Rogers, B. (1987) Paper presented at ICCV.

[110] Rogers, B. & Graham, M. E. (1983) Anisotropies in the perception of three-dimensional surfaces. *Science 221*, 1409-1411.

[111] Schunk, B. (1986) The motion constraint equation for optical flow. In *Proc. Int. J. Conf. Pat. Rec.*, 20-22.

[112] Shapiro, L. G., Haralick, R. M. & Phong, T. C. (1988) The use of the facet model and the topographic primal sketch in image analysis. In C. Brown (Ed.) *Advances in Computer Vision II*, 1-46. New York, New York: Academic Press.

[113] Siegel, S. (1956) *Nonparametric statistics.* New York, New York: McGraw-Hill.

[114] Smitley, L. & Bajcsy, R. (1984) Stereo processing of aeriel urban images. In *Proc. ICPR.*

[115] Spivak, M. (1975) *A Comprehensive Introduction to Differential Geometry, Volumes III & IV.* Berkley, California: Publish or Perish.

[116] Spoerri, A. & Ullman, S. (1987) Early detection of motion boundaries. In *Proc. ICCV*, 209-218.

[117] Stenton, A., Mayhew, J. E. W. & Frisby, J. P. (1984) Vertical disparity pooling and the induced effect. *Nature 309*, 622-623.

[118] Stevens, K. A. & Brookes, A. (1987) Depth reconstruction in stereopsis. In *Proc. ICCV*, 682-686.

[119] Stevens, K. A. & Brookes, A. Manuscript submitted to *Vis. Res.*.

[120] Subbarao, M. (1988) Interpretation of image flow: Rigid curved surfaces in motion. *IJCV 2 (1)*, 77-96.

[121] Terzopoulos, D. (1983) Multilevel computational processes for visual surface reconstruction. *CVGIP 24*, 52-96.

[122] Terzopoulos, D. (1986) Regularization of inverse problems involving discontinuities. *IEEE-PAMI 8 (4)*, 413-424.

[123] Terzopoulos, D. (1986) Integrating visual information from multiple sources. In S. Pentland (Ed.) *From Pixels to Predicates.* New Jersey: Ablex.

[124] Thompson, W. B., Mutch, K. M., & Berzins, V. A. (1982) Edge detection in optical flow fields. *Proc. AAAI*, 26-29.

[125] Thompson, W. B., Mutch, K. M., & Berzins, V. A. (1985) Dynamic occlusion analysis in optical flow fields. *IEEE-PAMI 7 (4)*, 374-383.

[126] Truesdell, C. & Toupin, R. A. (1960) The classical field theories. In S. Flügge (Ed) *Handbuch der Physik*, 227-793. Berlin: Springer Verlag.

[127] Verri, A. & Torre, V. (1986) Absolute depth estimate in stereopsis. *JOSA A 3 (3)*, 297-299.

[128] von der Heydt, R. (1979) *Stereoskopische Wahrnehmung der Orientierungsdisparation*, Doctoral dissertation, Eidgenossische Technische Hochschule, Zurich, Switzerland.

[129] von der Heydt, R., Adorjani, Cs. & Hänny, P. (1977) Neuronal mechanisms of stereopsis: Sensitivity to orientational disparity. *Experientia 33*, 786.

CS-TR Scanning Project
Document Control Form

Date : 8 / 3 / 95

Report # AI-TR-1112

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☒ Technical Report (TR) ☐ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 158 (166 - IMAGES)
Not to include DOD forms, printer instructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other:_____

Check each if included with document:

☒ DOD Form (2) ☐ Funding Agent Form ☒ Cover Page
☒ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages (by page number): _____

Photographs/Tonal Material (by page number): 68-73, 92-95

Other (note description/page number):

Description : Page Number:

(A) IMAGE MAP (1-158) PAGES #'ED 2-6, UN#'ED BLANK 7 -158
(159-163) SCANCONTROL, COVER, SPINE, DOD (2)
(164-166) TARGET'S (3)

(B) PAGE #3 HAS A CUT & PASTE

Scanning Agent Signoff:
Date Received: 8/3/95 Date Scanned: 8/15/95 Date Returned: 8/17/95

Scanning Agent Signature: Michael W. Cook

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AI-TR 1112	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On Interpreting Stereo Disparity		5. TYPE OF REPORT & PERIOD COVERED technical report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard Patrick Wildes		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0124
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, MA 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE April 1989 13. NUMBER OF PAGES 191
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research Information Systems Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

image understanding
stereo
surface representation
3D vision

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The problems under consideration center around the interpretation of binocular stereo disparity. In particular, the goal is to establish a set of mappings from stereo disparity to corresponding three-dimensional scene geometry. Stereo disparity is represented as a vector field derived from differential projection of a three-dimensional scene onto a pair of two-dimensional imaging surfaces. The resulting disparity field is analysed with the aid of mathematical tools from classical field theory. This analysis shows how disparity information can be interpreted in terms of three-dimensional

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 cont.

scene properties, such as surface depth, discontinuities and orientation. These theoretical developments have been embodied in a set of computer algorithms for the recovery of scene geometry from input stereo disparity. The results of applying these algorithms to several disparity maps are presented. Finally, comparisons are made to the interpretation of stereo disparity by biological systems.

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**

Scanned

Date: 8/15/1995

M.I.T. Libraries
Document Services