The Description of Image Curves: Discrete Forms of Continuity in Space

Lee A Iverson

B Eng. (Princeton University) 1984

Department of Electrical Engineering

McGill University

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Master of Engineering

December 15, 1988

Abstract

Early visual processing is seen as the search for certain kinds of abstract structure within visual stimuli, and one of these—image curves—is the focus of this work. The general context is early vision considered as a problem of inferring locally consistent vector fields from a (retinal) image. Particular reference is made to articulating the necessary local conditions for the existence of an image curve. These conditions are incorporated into the design of a local operator by formulating a non-linear decomposition of a standard linear one. The output of this operator then provides the initial estimates of confidence over a discrete and extensive description of possible image curves. A relaxation labelling system serves as an inference procedure for discovering consistent patterns in the initial estimates by exploiting the geometric structure of piecewise smooth plane curves. The calculation of the relaxation compatibilities is formulated as a closest point problem, and it is shown how the same logical conditions for the existence of an image curve can be incorporated into them.

The system developed is a highly parallel computational method. Its robustness is demonstrated in several realistic experiments and some implications for biological visual processing are explored. Finally, the ease of extending these results to other domains of early vision is indicated as a future area of research.

Résumé

La vision de bas niveau consiste à rechercher certaines structures abstraites à partir d'un stimulus visuel Ce travail s'attarde à une de celles-ci les courbes dans une image Le contexte général est celui de vision de bas niveau, laquelle étant vue comme un problème d'inférence de champs de vecteurs localement consistants obtenus à partir d'une image Un effort particulier a été fait pour essayer de formuler les conditions locales nécessaires à l'existence de telles courbes Ces conditions permettent alors de développer un opérateur local obtenu à partir d'une décomposition non-linéaire d'un opérateur linéaire standard Les résultats produits par cet opérateur procureront alors une estimation initiale de la confiance accordée à une description discrète des courbes potentielles dans l'image Un système de "relaxation d'étiquettes" sert de procédure d inférence afin de découvrir une configuration consistante à partir des estimations initiales en exploitant la structure géométrique de courbes planes lisses par morceaux Le calcul des compatibilités de relaxation est formulé comme un problème de point le plus proche et nous mentionnons comment les mêmes conditions logiques pour l'éxistence d'une courbe peuvent y être incorporées

Le système développé est hautement parallèle La robustesse du système est eprouvée à travers plusieurs expériences réalistes et quelques implications au niveau du processus visuel sont explorées Finalement, la facilité d'étendre ces résultats à d'autres domaines de vision de bas niveau est présentée comme une future voie de recherche

Acknowledgements

The first note of thanks should go to my advisor Steve Zucker who provided not only the intellectual foundation for this work but created and nurtured an exciting and inspiring intellectual environment at CVaRL and McRCIM

The combination of a strong social and intellectual atmosphere in the lab has also been essential for the success of this research and the individual contributions to this cannot go unacknowledged Most notable has been Allan Dobbins whose intellectual rigor and friendship have contributed directly to this work in many ways Benjie Kimia, Benoit Dubuc and Frank Ferrie have made similar contributions as we proceed along this path as true 'co-conspirators

In many ways, the quality of writing herein is above my usual standards, which can only be ascribed to the fact that it is not my standards it is meeting, but those of Alison Phinney She deserves a great deal of credit for managing to plow through this and bringing it up to a level where it depends less on jargon and more on simple, clear exposition. Her emotional support and the simple fact that she was the one who introduced me to this environment in the first place were both essential and much appreciated contributions

I thank both NSERC and the Department of Electrical Engineering for the financial support which has allowed me to continue this research In addition, a great debt must be paid to those who have contributed to making the Symbolics Lisp Machine the stimulating and productive creative environment that it is.

Table of Contents

Chapter 1 Introduction

Vision is properly seen as a constructive search for structure in the world from the visual stimuli which enter an eye or camera The earliest stage of such processing is the transduction of these stimuli into a set of simple measurements What has been called *early vision* is the attempt to take such crude measurements and begin to abstract from them a set of assertions about the structure of the information contained in this retinal image Beyond this, so called *later* or *high level* visual prcoesses take these descriptions and attempt to abstract from them a description of the world

My long-term research goal is to articulate a methodology for approaching problems in early vision In this thesis I focus on the inference of curves from grey-level images, one of the tasks that I believe is necessary for this larger theory My hope is that this emerging computational theory of early vision will, in addition to being a succesful engineering achievement, have relevance to visual neurophysiology and psychology

1.1 Edge/Line Detection: The Description of Image Curves

The literature on edge and line detection is incredibly varied and voluminous (for a recent review see [Rosenfeld 84]), and is based on a single guiding principle, reliably find the points in the image where an intensity edge or line exists, then describe such points in terms of planar curves Such curves are formed as the loci of points which fulfill certain logical conditions for the existence of a line or edge and are most appropriately described in terms of the local differential geometry of plane curves. *Detection* is not the sole issue, the process must eventually provide an accurate *description* of the curves discovered. Does it make sense then to decompose this process into separate stages of detection and description? I will show that this decomposition is unsuitable for describing curves in images.

Two kinds of planar curves exist in intensity images—those associated with edges and those associated with lines. In this thesis I will take *lines* to be those curves in an image which would have been drawn by a pen or pencil. *Edges* are instead the curves which separate between lighter and darker areas of the image—the perceived discontinuities in the intensity surface. *Image curves* are then either of these objects, the kinds of image events which can be described in the geometric language of planar curves.

The initial problem in arriving at a robust method for selecting and describing only edges and lines from among all possible visual events is that these concepts must be defined. I propose that lines be considered as the piecewise continuous loci of laterally extremal image intensities (the cross-section normal to the curve is locally extremal). Positive and negative contrast lines correspond to maxima and minima respectively. Similarly, edges are the loci of laterally maximal first derivatives of intensity, taken in a direction normal to the curve. I will develop these ideas first in principle then explain how they will be used in practice. Formally, for an analytic intensity surface $\{ I(x,y) : x,y \in R \}$ a smooth segment of curve **s** is defined as a differentiable mapping $\mathbf{s} : T = (t_0,t_1) \to R^2$ given by $\mathbf{s}_t = (x(t),y(t))$ with the normal vector $\mathbf{n}_t$ a unit vector in the direction $(x''(t),y''(t))$. Three kinds of image curve are then defined[†]:

$$\mathbf{s} \text{ is a Positive Contrast Line} \Longleftrightarrow \mathbf{s} \text{ smooth on } T \bigwedge$$

$$\forall(t \in T)\ \exists(\epsilon > 0)\ 0 < |\delta| < \epsilon \Rightarrow I(\mathbf{s}_t) > I(\mathbf{s}_t + \delta\mathbf{n}_t) \qquad (1.1.1)$$

[†] Note that the definition of a line is sign specific, whereas the edge is defined as a locus of local maxima *or* minima. There is no qualitative difference between edges that go from light to dark moving towards their centre of curvature and those of the opposite contrast; these events cannot be named differently.

$$\mathbf{s} \text{ is a Negative Contrast Line} \Longleftrightarrow \mathbf{s} \text{ smooth on } T \bigwedge$$

$$\forall(t \in T)\ (\exists \epsilon > 0)\ 0 < \delta < \epsilon \Rightarrow I(\mathbf{s}_t) > I(\mathbf{s}_t - \delta \mathbf{n}_t) \qquad (1.1.2)$$

$$\mathbf{s} \text{ is an Edge} \Longleftrightarrow \mathbf{s} \text{ smooth on } T \bigwedge$$

$$\forall(t \in T)\ (\exists \epsilon > 0)\ 0 < \delta < \epsilon \Rightarrow \frac{dI}{d\mathbf{n}}(\mathbf{s}_t) > \frac{dI}{d\mathbf{n}}(\mathbf{s}_t + \delta \mathbf{n}_t). \qquad (1.1.3)$$

These definitions can be taken as the necessary logical conditions for the existence of an image curve. Clear from this formulation is the chicken-and-egg nature of the problem. Since the conditions for existence of the curve depend on its differential structure (specifically the normal **n**), a candidate curve can only be tested *after* it is described. Hence separating the process into stages of detection *followed by* description is misguided.

Fortunately in practice there is a way out of this dilemma, based on the observation that all of the tests are defined locally. By truncating and quantizing the continuous formulation, the existence of a candidate curve can be tested everywhere in the image for all local and finite differential structures. This is combinatorially feasible and the techniques for accomplishing it form the body of this thesis. More specifically, my goal is to infer a quantized description of the local properties of curves—their trace, tangent, and curvature—from images. This process is called *orientation selection*, after [Zucker 85].

1.2 Problem Definition

Referring back to the curve **s**, consider the projection of this curve into a parameter space $X \times Y \times \Theta \times K$ representing local position, orientation and curvature. The problem is to represent this curve in terms of a set of discrete points distributed throughout this four-dimensional space—the hypergraph $\mathcal{I} = \{\hat{\imath} = [X_i, Y_i, \Theta_i, K_i] \mid X_i, Y_i \in R, \Theta_i \in \Theta, K_i \in \mathbf{K}\}$. This description* is obtained by selecting only that subset of discrete points $\mathcal{S} \subset \mathcal{I}$ for which some point $\mathbf{s}_t \in \mathbf{s}$ exists whose description $[x_t, y_t, \theta_t, \kappa_t]$ is closer to

* In fact this is a description of a family of curves of which **s** is guaranteed to be a member.

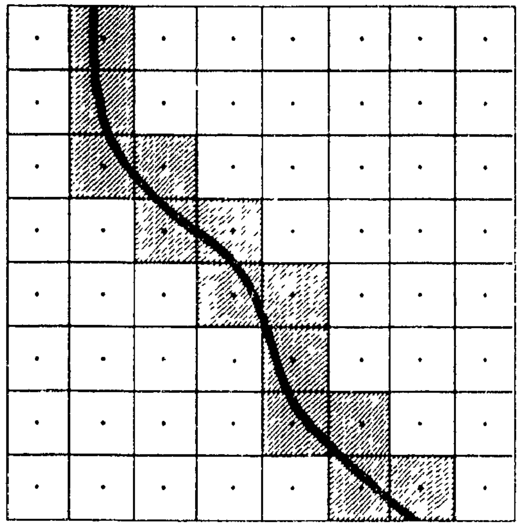


Figure 1.1 Discrete representation of continuous curve in terms of the set of pixels which the curve intersects. In this two-dimensional analogy to the real four-dimensional representation the set of grey pixels is a position-based analog of the set S of equation (1.2), each pixel being the Voronoi cell $H_{\hat{\imath}}$ around pixel center $\hat{\imath}$.

$[X_{\hat{\imath}}, Y_{\hat{\imath}}, \Theta_{\hat{\imath}}, K_{\hat{\imath}}]$ than any other discrete point, formally

$$S = \{\hat{\imath} \in I \mid \exists(t \in T) : \forall(\hat{\jmath} \in I) : \|[x_t, y_t, \theta_t, \kappa_t] - \hat{\imath}\|_2 \leq \|[x_t, y_t, \theta_t, \kappa_t] - \hat{\jmath}\|_2 \} \quad (1.2)$$

Considered as a labelling problem, the description must select the label $\lambda(\hat{\imath})$ corresponding to the discrete point $\hat{\imath} \in I$ if and only if the point $\hat{\imath}$ fulfills the condition in (1.2). If $H_{\hat{\imath}}$ is the convex Voronoi cell associated with the element $\hat{\imath}$, then the description must set the label

$$\lambda(\hat{\imath}) = \begin{cases} \text{TRUE}, & \text{if } (s \cap H_{\hat{\imath}}) \neq \emptyset, \\ \text{FALSE}, & \text{otherwise} \end{cases} \quad (1.3)$$

This representation of a curve in terms of a set of discrete units of position, orientation and curvature is related to the computer graphics representation of a plane curve as the connected set of image pixels through which it passes (see Figure 1.1). Furthermore, multiple image curves are represented as distinct connected subsets of the hypergraph so defined.

The goal of the system developed in this thesis is to infer such a description from an image. For each curve which satisfies the logical conditions of one of the definitions

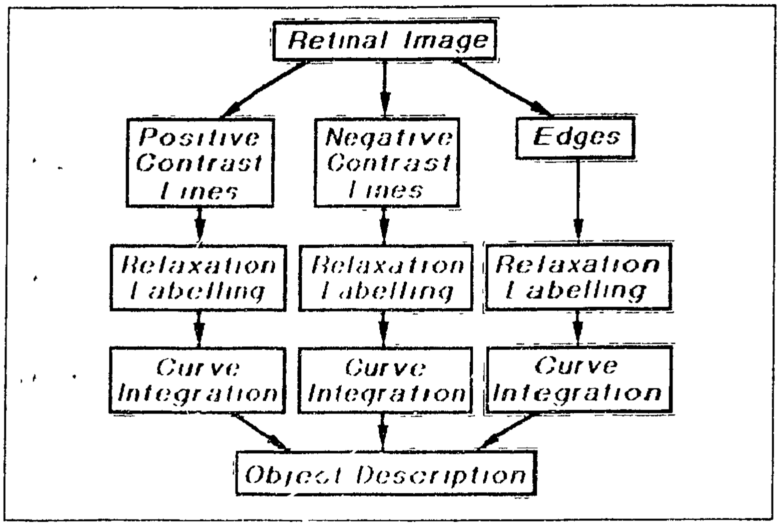


Figure 1.2 Block diagram of curve description system.

in (1 1), there must be a connected subset of the hypergraph $\mathcal{I}$ for which each label is TRUE by (1 3) Finally, no other nodes in the hypergraph should be selected

1.3 Methodology

Building on an approach previously described in [Zucker *et. al.* 77] and [Parent & Zucker 85], there are two steps involved in inferring the local differential geometry of the curves which comprise the image scene

Chapter 2 A set of simple local operators test for the existence of image curves at every point in the image This is done for a set of discrete orientations and curvatures, using a model of end-stopped and non-end-stopped simple cells in visual cortex ([Dobbins *et al* 87] and [Dobbins *et al* 88]) To improve the selectivity of these operators, I develop a non-linear decomposition of their components that tests a set of necessary conditions for the existence of a curve at the specific position, orientation and curvature

Chapter 3 A relaxation labelling system then refines these local estimates by testing for continuity and smoothness of the inferred local differential structure Each description of the local differential structure of an image curve forms a hypothesis in this network, and the compatibility between nearby hypotheses is based on an assumption of locally constant curvature A computer implementation of this system is described

Chapter 4 The results of the system are shown and described.

Chapter 5 I present a preliminary analysis of some of the extensions of this work to other domains in early vision The way in which this class of models relates to neurophysiological and psychophysical models and data is also explored.

1.4 Contributions

- A logical decomposition of a linear operator is derived which allows testing of two necessary conditions for the existence of an image line a local extremum in cross section and the continuity of the line through the operator center
- This line operator then forms the excitatory component of a Dobbins ES operator which matches local orientation and signed curvature of an image curve A method for tuning such operators is described
- The calculation of relaxation labelling compatibilities for a two-label (TRUE/FALSE) geometric hypothesis verification problem is formulated as a closest point problem in the geometric parameter space
- The consistent labelling for orientation selection is defined in terms of a connected discrete description of an image curve in a differential space
- The smoothing and localization conditions for this consistent labelling are translated into a linear/logical support network for relaxation labelling The non-linearities in this network are based on the same existence conditions as the initial operators, as well as an argument from convergence
- The general principles developed in the solution of orientation selection are outlined and extensions of these principles to texture flow and optical flow are briefly described

Chapter 2 Initial Measurements

2.1 Introduction

The most important characteristic of a measurement system is the degree to which it reflects the quantity it measures while remaining insensitive to other potentially confounding influences In images, there are many different kinds of events, some of which are irrelevant to the discovery of image curves It is the goal of this chapter to develop an operator for measuring the local orientation and curvature of image curves while remaining as insensitive to both noise and nearby curves as possible Two of the necessary logical conditions for the existence of an image curve are used to develop a decomposition of a standard linear line operator The resulting operator behaves identically to the linear one as long as all of the conditions are satisfied, however, it only responds positively when they all are Thus the analysis which forms the body of this chapter constitutes a departure from traditional operator designs—while other designs are based solely on guaranteeing responses to characteristic image structures, this operator is designed to not respond when the characteristic image structure is not present Hence it is well-behaved when faced with uncharacteristic input as well

The most damaging failure of a measurement system is to systematically respond to events or quantities unrelated to the quantity being measured, we refer to these systematic false responses as *aliasing* [⊥] In particular, such aliasing can cause catastrophic effects when fed into other systems which are unable to distinguish between aliases and 'true' responses In this chapter we will demonstrate that although standard linear operators have significant aliasing problems, it is possible to overcome them

Two different kinds of image events give rise to curves lines and edges Both lines[‡] and edges are piecewise continuous curves which can be defined in terms of the shape of their cross-sections (see equations (1 1)) The necessary logical conditions on the cross-section must be satisfied and in each case we will show that a linear operator tests them incompletely This problem will be resolved by introducing a cross sectional decomposition of the linear operator which allows testing of these missing conditions

A second necessary condition for the existence of an image curve is continuity By specifying that image curves are piecewise and not necessarily closed the existence of discontinuities and end-points on the curve is expressly considered Such discontinuities are fundamental to a description of curves since they are invariant under rotation and translation and even under certian kinds of deformation A linear operator fails to ensure that local smoothness is satisfied before giving a positive response Instead it smooths over curve discontinuities and responds positively in many places which are beyond line endings This is resolved by introducing a second decomposition of the linear operator which allows for testing of local continuity and smoothness by testing logical conditions for such continuity

A broad set of goals and constraints thus directs this analysis The overall goal is to develop an operator that maps images into hypotheses about the existence of image curves such that it

- estimates both positional location and local differential structure of the image curve,

---

⊥ Realize that while this use of the word *aliasing* is related to the use in sampling theory, it is not the same This is a broader use of the term

‡ In this context a line is not necessarily a *straight* line

- avoids smoothing over discontinuities in image curves,
- exhibits a stable, well-localized response cut-off at the end-point of an image curve;
- operates stably in the presence of multiple image curves,
- avoids characteristic responses to stimuli which do not belong to the desired class of input stimuli,
- is independent of the tiling or regularity of the detector grid,
- degrades gracefully in the presence of noise, and
- is computable in parallel

2.2 The Operator

Referring back to the definition of image curves in §1 1, there are three qualitatively different kinds positive contrast lines, negative contrast lines and edges There are thus three distinct sets of logical conditions for the existence of image curves, the curve description process should respect this distinction and prevent aliasing between them (see Figure 1 2) I ensure that the individual conditions are satisfied by testing them, measuring the existence or non-existence of an image curve of the appropriate type passing through a given image position The design will concentrate on lines and not edges (the reasons for this will be made clear in §2 3.1)

We begin by adopting a standard oriented, linear line operator similar to one arrived at in [Canny 86] Canny adopted the assumption of linearity to facilitate noise sensitivity analysis and optimization, and arrived at a line operator which is similar to standard ones with a gaussian second-derivative cross-section Neurophysiologists have adopted linear models for the responses of cells in the primary visual cortex [Hubel & Weisel 65, Orban 84], and such models seem to capture many of their functional properties These models are attractive from a strictly functional point of view because they exhibit most of the properties required of a measurement operator for image curves However, they suffer greatly from aliasing effects (shown below in Figure 2 3)

In order to limit this aliasing the assumption of linearity is dropped and the necessary local conditions for an image curve are tested explicitly The resulting operator appears to be linear to many of the standard 'probe stimuli' used by visual neurophysiologists, but it has the additional property that its response is only positive when these conditions are fulfilled. The tests are incorporated by decomposing the operator into components which represent the quantities needed for the logical tests.

2.2.1 Decomposition

The decomposition must be conducted in a principled fashion so that no information is lost in the process, and the recomposition must take advantage of the information inherent in both the original operator and in each of the components. Given the optimality of the original operator in certain circumstances, it is also important that the non-linear operator reduce to the original linear case if certain logical conditions are satisfied. Thus two properties for the decomposition are essential:

- The decomposition of the initial linear operator kernel must create a set of kernels whose sum is identical to the initial kernel.
- The combination of the component responses must exhibit both logical and linear properties, testing the necessary conditions for a local curve and producing a result which is the sum of the component responses when these tests succeed.

The first condition suggests that this decomposition resembles a *partition of unity* [Spivak 7?]. For a linear operator $f(\mathbf{x})$ and a set of component functions $g = \{g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_n(\mathbf{x})\}$ all defined over the range of values $\mathbf{x} \in R$, the set g consititutes a partition of unity if and only if

$$\forall(\mathbf{x} \in R) \sum_{1 \le i \le n} g_i(\mathbf{x}) = f(\mathbf{x}) \tag{2.1}$$

Many such partitions are clearly possible; I choose the partition such that the components reflect the underlying logical features of the linear operator.

Associating TRUE with positive values and FALSE with negative, "semi-linear" analogs for $\bigwedge$ (and) and $\bigvee$ (or) may be defined as

$$x \mathrel{\dot\wedge} y \;\triangleq\; \begin{cases} x+y, & \text{if } x > 0 \wedge y > 0, \\ y, & \text{if } x > 0 \wedge y \le 0, \\ x, & \text{if } x \le 0 \wedge y > 0, \\ x+y & \text{if } x \le 0 \wedge y \le 0. \end{cases} \tag{2.2.1}$$

$$x \mathbin{\vee\!\!\!\!+} y \quad \stackrel{\triangle}{=} \quad \begin{cases} x+y, & \text{if } x > 0 \wedge y > 0, \\ x, & \text{if } x > 0 \wedge y \leq 0, \\ y, & \text{if } x \leq 0 \wedge y > 0, \\ x+y & \text{if } x \leq 0 \wedge y \leq 0. \end{cases} \tag{2.2.2}$$

The propriety of these operators can be verified by examining their truth/addition tables

Truth/Addition Tables
for Semi-Linear Logical Operators

$\wedge\!\!\!\!+$		
	+1	1
+1	+2	1
1	1	2

$\vee\!\!\!\!+$		
	+1	−1
+1	−2	+1
−1	+1	−2

If only the sign of the result is taken as the output, then these are exactly equivalent to the logical operators (assuming positive $\Rightarrow$ TRUE and negative *implies* FALSE) These linear/logical combinators are examined more closely in Appendix A

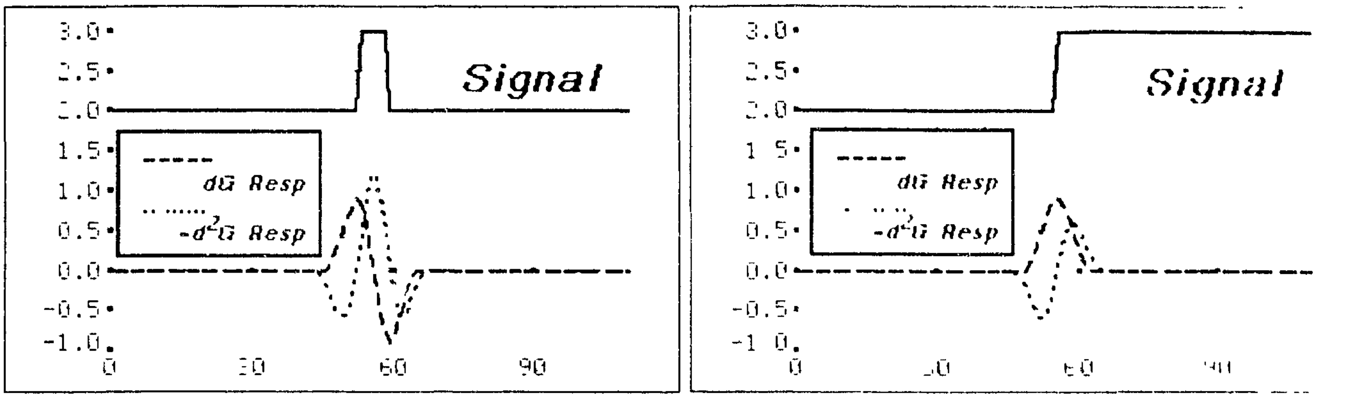
2.2.2 The Cross-Section Condition: Lateral Maxima

Examination of the cross-sections of intensity lines in images reveals that a necessary condition for the existence of such a line is a local extremum in intensity At an intensity edge, the cross-section exhibits an extremum in the directional derivative of intensity perpendicular to the orientation of the edge (refer to equations (1 1) and Figure 2 1, a display of typical 1D cross-sections of these intensity phenomena) In addition, the second (or third) directional derivative at these extremal points can be used as a measure of the significance of the line (or edge) strictly in terms of local contrast

A local extremum in a one-dimensional signal $S(x)$[*] exists only at those points p where

$$\left.\frac{dS}{dx}\right|_p = 0 \qquad \text{and} \qquad \left.\frac{d^2 S}{dx^2}\right|_p \neq 0. \tag{2.3}$$

[*] Assume $S(x)$ is sufficiently differentiable



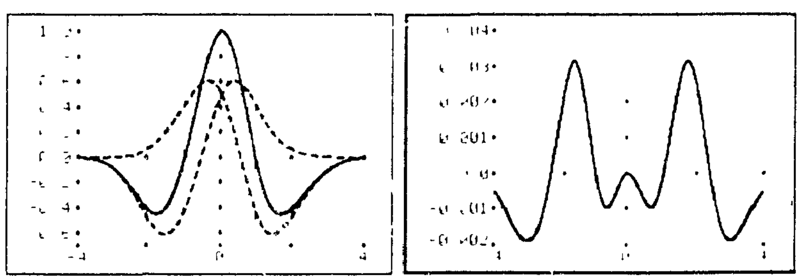
(a) (b)

Figure 2.1 Cross sections of image lines and edges. A line in an intensity image (a) is located at the peak of its cross-section. Note that this coincides with a zero in the derivative (dS) and a negative second derivative (d^2S). An intensity edge (b) occurs at peaks in the derivative (dS) of the cross-section. The derivatives shown are derived from convolution by dG and d^2G operators with $\sigma = 3$

Estimating the location of zeroes in the presence of noise is normally achieved by locating zero-crossings, or the points p where

$$\left.\frac{dS}{dx}\right|_{p-\epsilon} > 0 \quad \text{and} \quad \left.\frac{dS}{dx}\right|_{p+\epsilon} < 0 \tag{2.4}$$

(this direction of change corresponds to an extremum at which $\frac{d^2S}{dx^2} < 0$, a maximum). An operator which can reliably restrict its responses to only those occasions when these necessary conditions hold will only respond to local maxima in a one-dimensional signal.



(a) (b)

Figure 2.2 Approximation of d^2G by two dG operators. (a) Shows the linear approximation of the d^2G operator (solid line) from the sum of two dG operators (dotted lines). (b) Shows the approximation error for this sum which is always less than 3.5 parts per thousand.

A set of noise-insensitive linear derivative operators (or 'fuzzy derivatives' [Koenderink & van Doorn 86]) are the various derivatives of a Gaussian envelope

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{2.5.1}$$

$$dG_\sigma(x) = -\frac{x}{\sigma^2} G_\sigma(x) \tag{2.5.2}$$

$$d^2G_\sigma(x) = \frac{x^2-\sigma^2}{\sigma^4} G_\sigma(x) \tag{2.5.3}$$

When convolved over a one-dimensional signal these give noise-insensitive measures of the derivatives of the signal Thus two convolutions will determine $S'(x)$ and $S''(x)$ where

$$S'_\sigma(x) = dS(x) * G_\sigma(x) = S(x) * dG_\sigma(x) \tag{2.6.1}$$

$$S''_\sigma(x) = d^2S(x) * G_\sigma(x) = S(x) * d^2G_\sigma(x) \tag{2.6.2}$$

Using these measures and equations (2 3) and (2.4) extrema can be located by simply finding those points which satisfy the necessary and sufficent conditions for the existence of an extremum

$$S'_\sigma(x-\epsilon) > 0 \quad \text{and} \quad S'_\sigma(x+\epsilon) < 0 \quad \text{and} \quad S''_\sigma(x) < 0 \tag{2.7}$$

The loci of points for which this condition holds are distinct segments along x with widths $\leq 2\epsilon$ The parameter σ determines the amount of smoothing used to reduce noise-sensitivity

Comparing the operators dG and d^2G, it occurs that a linear combination of two of the dG operators might give d^2G In fact, using central limits, it is observed that.

$$\frac{df}{dx} = \lim_{\epsilon \to 0} \frac{f(x+\epsilon/2) - f(x-\epsilon/2)}{\epsilon} \tag{2.8}$$

Using the fuzzy derivative operators, one would therefore expect to find something like [†]

$$d^2G_\sigma(x) \approx \frac{1}{\sigma}\left(dG_{\sigma'}(x+\delta) - dG_{\sigma'}(x-\delta)\right) \tag{2.9}$$

[†] The $1/\sigma$ scaling factor can be derived by examining the respective d^2G and dG scaling coefficients

If this holds, then all that is necessary to evaluate the conditions in equation (2.7) is the derivative of the signal $S'_\sigma(x)$—the linear combination gives $S''_\sigma(x)$. Figure 2.2 shows that an excellent approximation is in fact possible. Forcing equality of the approximation at $x = \sigma$ and $x = 0$ gives the system of equations

$$0 = dG_{\sigma'}(\sigma + \delta) - dG_{\sigma'}(\sigma - \delta) \qquad (2.9) \text{ @ } x = \sigma \qquad (2.10.1)$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} = dG_{\sigma'}(\delta) - dG_{\sigma'}(-\delta) \qquad (2.9) \text{ @ } x = 0 \qquad (2.10.2)$$

Solving initially for σ' in (2.10.1) determines that

$$\sigma' = \sqrt{\frac{2\sigma\delta}{\log((\sigma+\delta)/(\sigma-\delta))}} \qquad (2.11)$$

and substituting into (2.10.2) gives

$$\log((\sigma+\delta)/(\sigma-\delta)) = \sqrt{\frac{2\delta}{\sigma\log((\sigma+\delta)/(\sigma-\delta))}} \exp\left(\frac{\delta\log((\sigma+\delta)/(\sigma-\delta))}{4\sigma}\right) \qquad (2.12)$$

Since every expression in which δ appears in this equation depends only on the ratio δ/σ, call this ratio b. Then

$$\log((1+b)/(1-b)) = \sqrt{\frac{2b}{\log((1+b)/(1-b))}} \exp\left(\frac{b\log((1+b)/(1-b))}{4}\right). \qquad (2.13)$$

This is independent of σ and can be solved using Newton's method. Substituting back into (2.11) gives the result

$$\begin{aligned} \sigma' &\approx 0.95443850\,\sigma = a\sigma \\ \delta &\approx 0.49816102\,\sigma = b\sigma \end{aligned} \qquad (2.14)$$

Referring back to Figure 2.2, it can be seen that with this approximation, the sum of the two dG kernels never strays from the true d^2G kernel by more than three parts in a thousand.

The use of these dG convolutions allows testing of all three conditions in (2.7) simultaneously, for the constituent convolutions provide the two offset estimates of $S'_\sigma(x)$ and their sum is the $S''_\sigma(x)$. With the linear/logical combinators of §2.2.1 we are thus able to define an operator which has a positive response only at local minima (within a range of σ of the actual minimum)

$$R_\sigma(S, x) \;=\; (S(x) * dG_{a\sigma}(x + b\sigma)) \wedge (S(x) * -dG_{a\sigma}(x - b\sigma)) \qquad (2.15)$$

where a and b are the constants from (2.14). Likewise, for local maxima, use

$$R_\sigma(S, x) \;=\; (S(x) * -dG_{a\sigma}(x + b\sigma)) \wedge (S(x) * dG_{a\sigma}(x - b\sigma)) \qquad (2.16)$$

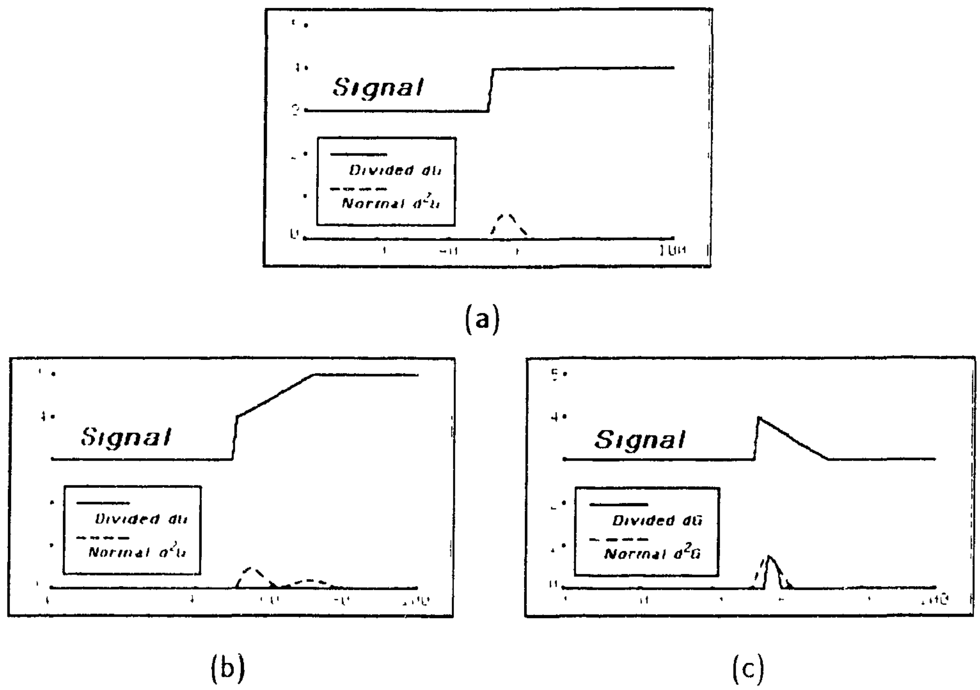


Figure 2.3 Responses of non-linear and linear operators near step. Rectified responses of the d^2G operator and divided dG operators are shown for three step edges with varying slopes of the upper region. It can be seen the the non-linear operator blocks the unwanted response near a step which is not also a local maximum (a & b) but that when it is a local maximum (c) it does respond. The d^2G operator however responds in each of these cases exhibiting consistent displacement of the peak response.

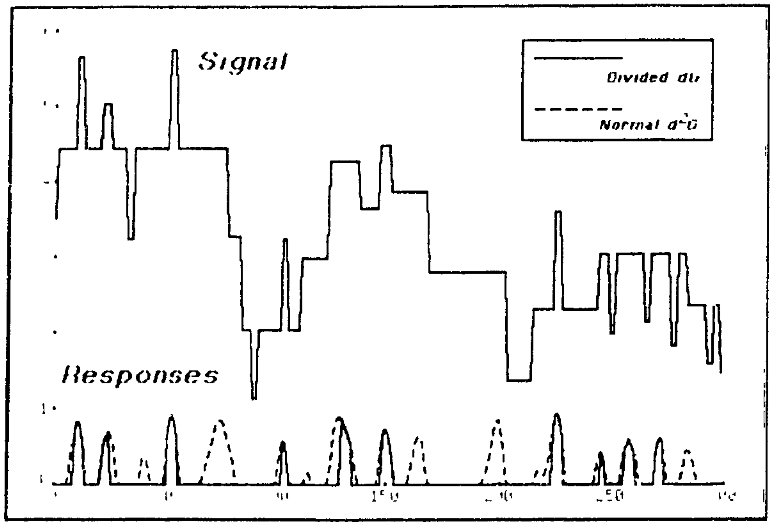
It might be asked what has been gained by introducing this level of complexity into what was a simple linear operator? The gain is considerable, and harkens back to the initial criteria for an acceptable measurement operator established in §2.1. The linear d^2G operator has the particularly undesirable characteristic of exhibiting consistent patterns of aliasing responses The simplest example of such an aliasing effect is the response near a step (see Figure 2 3) The linear operator displays a characteristic peak in response when the edge is centered over one of the zeroes in the operator profile The non-linear operation blocks this response since both dG halves of the operator register derivatives in the same direction and so do not fulfill the necessary conditions of equation (2 4) Examining the alternatives when the slope above the step is non zero it is clear that only when the slope is negative (thus making the transition point a local maximum) will the non-linear operator respond positively

In the displays of Figure 2 4 the responses from a linear d^2G operator are contrasted with the responses from its non-linear derivative Notice that with the original (noise-free) signal the new operator does a much better job of locating the peaks in the signal, it does not falsely respond near step edges (including one caused by the negative contrast line at 32) It is not until a noise level of 0 8 (S/N ratio of 1 25 1) that an appreciable number of false positives appear in the responses of the new operator An important point to note is that the positive regions in the response are all less than 4 positions wide This is a design factor, related to the size of the operator, which ensures that the peak position will always be localized to the resolution of the operator and not to the width of the signal

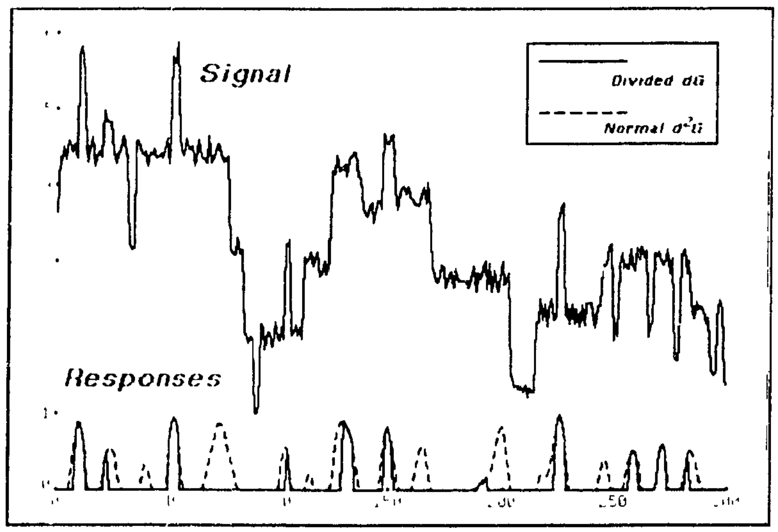
2.2.3 The Tangential Condition: Continuity

So far only the structure of an image normal to the curve has been discussed The second aspect of the definitions in (1 1) is smoothness Without this, the local orientation and thus the normal direction would not be defined An estimate of the local orientation of the curve is thus required, and we must develop an oriented operator to provide it Incorporating the preceding analysis a perpendicular cross-section as described in equation (2 16) or (2 15) will be used

The tangential structure of the operator will be developed in the same stages as the cross-sectional structure was above We begin with the assumption that the image



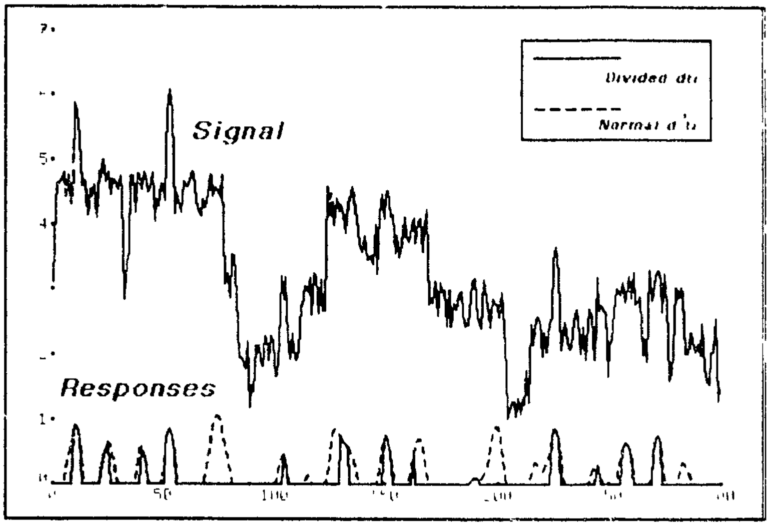
(a)



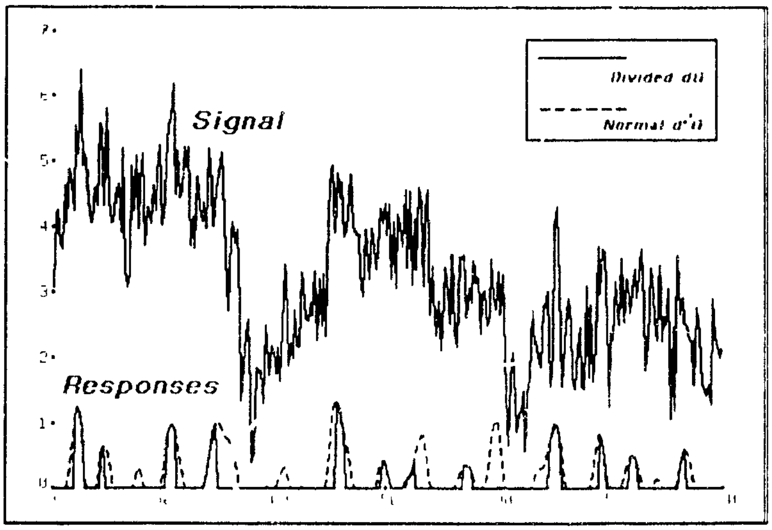
(b)

Figure 2.4 Responses of non-linear and linear operators to noisy cross-sections. Signal degraded with additive uniform noise with (a) $\sigma = 0.0$ and (b) $\sigma = 0.2$ Note that the d^2G operator exhibits many more false positives for each level of noise These can be attributed to regions of negative local curvature which do not coincide with zeroes in the first derivative

curve is everywhere smooth and corrupted by additive gaussian noise This suggests that the response should be filtered along its length, and a gaussian envelope is the natural choice for such a filter The first candidate operator is thus a pair of component operators (referred to as the *lateral components* of the operator) whose responses will be combined



(c)



(d)

Figure 2.4 Responses of non-linear and linear operators to noisy cross-sections. (Continued) Signal degraded with additive gaussian noise with (c) $\sigma = 0.4$ and (d) $\sigma = 0.8$

by the $\mathbb{A}$ operation described above.

$$W_1(x,y) = G_{\sigma_x}(x)\ dG_{a\sigma_y}(y + b\sigma_y), \qquad (2.17.1)$$

$$W_2(x,y) = -G_{\sigma_x}(x)\ dG_{a\sigma_y}(y - b\sigma_y) \qquad (2.17.2)$$

This candidate operator, which is similar to previous approaches (e.g. [Canny 86]) except for the nonlinear cross-section, fulfills all of the criteria in §2.1 but two. It

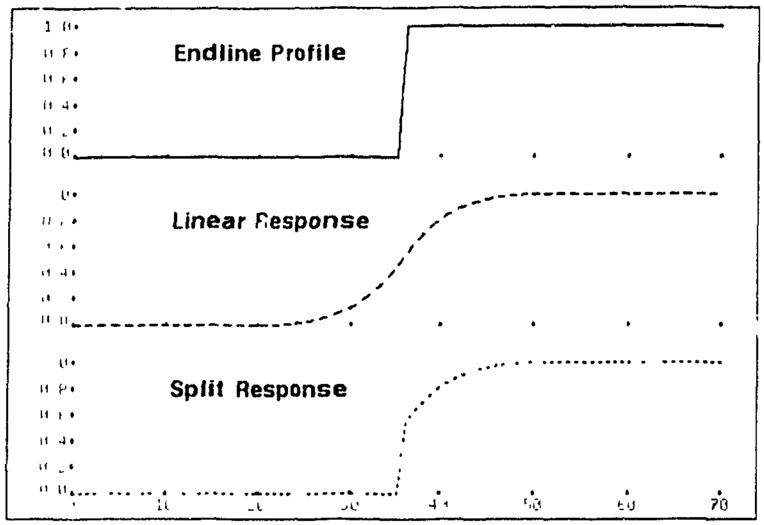
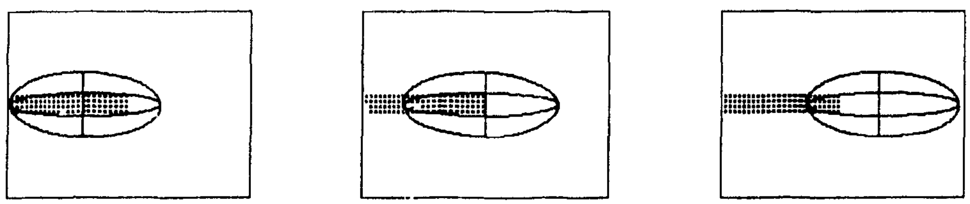


Figure 2.5 Responses of linear and split operators near end of line. The signal is a parallel cross-section of an image line Note that the semi-linear split operator drops off suddenly as its center moves off of the end of the line, while the linear operator exhibits a smooth attenuation of response



(a) (b) (c)

Figure 2.6 Schematic of the half-field decomposition and line endings. The elliptic regions in each figure represent the operator positions as the operator moves beyond the end of an image line In (a) the operator is centered on the image line and the line exists in both half-fields In (b) the operator is centered on the end-point and whereas the line only exists in one half-field, the other half-field contains the end-point In (c) the operator is centered off the line and the line only exists in one half-field

tends to smooth over discontinuities in the image curve, and has a smooth drop-off in response as it passes over an end-point in a line (see Figure 2 5) These undesirable characteristics are based on a single problem, the gaussian profile along the length of the curve treats all areas indiscriminately, smoothing over any discontinuities along the length. This smoothing should only take place when the curve is *known* to be locally continuous. A second decomposition into two tangential half-fields around the operator centre allows testing of this prior condition

Notice that if an operator is centered over a continuous section of an image curve, then the curve must be similar in both half-fields of the operator Moving beyond an end-point, the curve will exist within only one of these half-fields (see Figure 2 6) Thus, in order to verify that the operator is centered *on* a line and is not just experiencing the influence from a nearby line, both half-fields must be similarly active As long as there is some overlap between these two half-fields, there will be a partial response at a line-ending (where only one half-field is actually contributing to the response) but none once the operator is centered off of the line Near a line end point the half field that is not over the line will be much more sensitive to noise and confounding stimuli than the half field on the line (it will be matching something that does not exist where it is looking, as opposed to something which does) It may therefore be essential to divide the operator into more than just these two regions and test for *consistent* reponses from all of the regions

A mathematical response to this goal is to develop a partition of unity which divides the operator kernel into regions along its length, a partition which separates smoothly and symmetrically around equally spaced points along a line Notice first that the sigmoid

$$\sigma_\rho(x) = \frac{1}{1+e^{-\rho x}} \tag{2.18}$$

can be used to partition a function into two regions around c by the equation:

$$F^-(x) = F(x)\,\sigma_\rho(c-x) \tag{2.19.1}$$

$$F^+(x) = F(x)\,\sigma_\rho(x-c) \tag{2.19.2}$$

such that $F^-(x) + F^+(x) = F(x)$ for all x (easy to verify since $\sigma_\rho(c-x) + \sigma_\rho(x-c) = 1$) Thus F^- and F^+ form a partition of unity of F around c (see equation (2 1))

This can be generalized to $n > 2$ regions by combining these sigmoid partitions for a set of $n-1$ partition points The condition in (2 1) is easily verified for the kernels defined by (2 20) over the strictly increasing set of separation points $\{c_i \; 1 \le i < n\}$

$$W_i(x) = \left(\prod_{1\le j<i} \sigma_\rho(c_j - x)\right)\left(\prod_{i\le j<n} \sigma_\rho(x - c_j)\right). \tag{2.20}$$

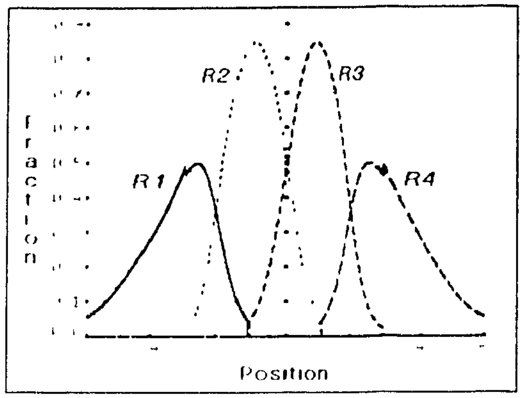


Figure 2.7 Splitting of 1D gaussian operator into four linear components using the basis functions of equation 2 20 The centers of division (c_i) are equally spaced and centered around the operator center The degree of separation ρ is 3

Each component F_i of the partition is then specified by

$$F_i(x) \;=\; W_i(x)\, F(x) \tag{2.21}$$

The separation of a gaussian envelope into four equal width regions (which we will refer to as the *length components* of the operator) by this method is shown in Figure 2 7.

The continuity condition above requires that one of the separation points be the center of the operator, and we opt for equal spacing on either side (see Figure 2 7) What remains to be resolved is the appropriate combination of the regions to ensure end line stability and a consistent line model over the length of the operator In [Davis *et al.* 73] a scheme similar to this one was proposed but the resulting operator gave a positive response only when more than half of its length components were positive I propose a combination which requires that all of either half-field to be positive and any subregion of the other half-field This replaces the threshold in the Davis operator with a more structured form expressed in terms of the linear/logical operators Labelling the length component responses as $\mathbf{L} = [r_1, \ldots, r_{n/2}]$ and $\mathbf{R} = [r_{n/2+1}, \ldots, r_n]$ (the left half-field and right half-field respectively), this condition becomes

$$W \;=\; \left(\bigwedge_{l \in \mathbf{L}} l \;\wedge\; \bigvee_{r \in \mathbf{R}} r \right) \vee \left(\bigwedge_{r \in \mathbf{R}} r \;\wedge\; \bigvee_{l \in \mathbf{L}} l \right). \tag{2.22}$$

I have thus far left open the question of how to choose the parameters of these operators. The size of the operator (determined by σ_l and σ_y), its orientation specificity (determined by σ_x σ_y), and the number of regions to divide into (the parameter n). These parameters are determined by *a priori* factors which arise from the specific structure of the measurement problem the operators are to be used for, we conduct such an analysis in §2.3 below

2.2.4 Curvature

The final stage in the development of our operator follows the work of Allan Dobbins [Dobbins *et al.* 87 and 88], which described a computational model of end stopped cells in the primary visual cortex. His key observation was that end stopping could be related to curvature, and that models of end-stopped neurons could be derived from taking differences between oriented linear operators. This model is easily extended to incorporate the stable, non-aliasing operators developed above

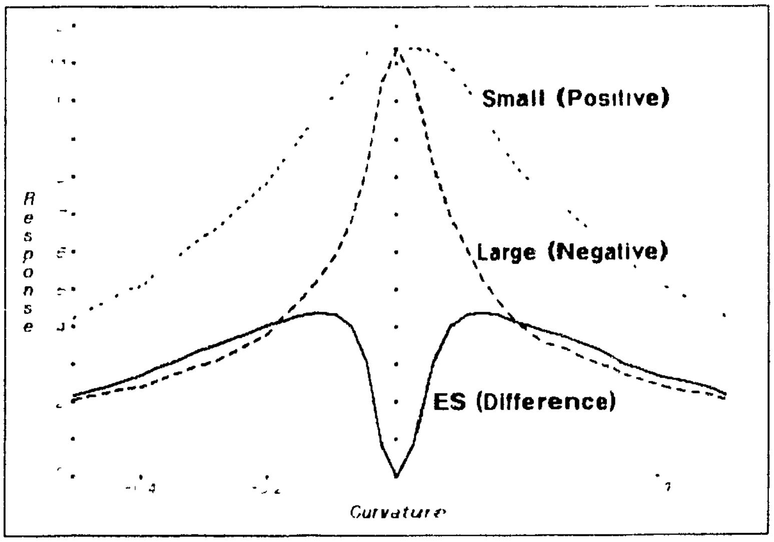


Figure 2.8 Curvature responses of simple cells of two different lengths and their difference. The response profiles for the small operator (length = 10) and the large operator (length = 20) do not coincide so when the difference of the two response curves is taken (appropriately weighted) the aggregrate response shows a peak response to a non-zero curvature

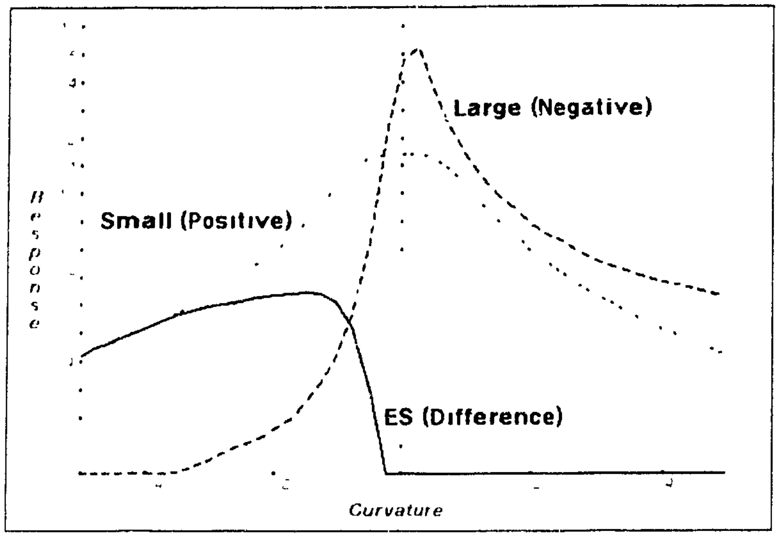


Figure 2.9 Curvature responses of even- and odd-symmetric simple cells of two different lengths the longer having odd-symmetric cross-section Because the curvature reesponse of the large operator is asymmetric its inhibitory influence on the aggregate (difference) of the two responses effectively inhibits responses for one sign of curvature and not the other The combination is thus band-pass for both sign and magnitude of curvature

Given responses from two similarly oriented line operators R_s (short) and R_l (long) and weighting factors to compensate for the differences in their relative areas(c_s and c_l), the end-stopped (ES) operator is defined to respond as

$$R_{ES} = \phi(c_s\ \phi(R_s) - c_l\ \phi(R_l)) \tag{2.23}$$

where $\phi(x)$ is a half-wave rectification If the excitatory and inhibitory components are closely matched in both spatial frequency bandwidth (in the normal direction) and orientation bandwidth then the end-stopped operator will have a characteristic maximum response to a *curved* line with some non-zero curvature Since this system depends on *signed* curvature, I have adopted a model with the linear/logical line operator described above as the excitatory component and an operator with a dG cross-section as the inhibitory component The odd-symmetric nature of such an operator ensures that it inhibits response for one sign of curvature and not the other (see Figure 2 9) This allows us to develop an end-stopped line operator which is tuned for both magnitude and sign of curvature

The analysis used to arrive at the current operator is based entirely on arguments from analytic images and convolutions An operational computational system involves discretely sampled images and discrete convolutions, so we must translate from one set of terms to the other. The use of the "fuzzy" derivatives is essential to this translation with the specifics described in Appendix B. For the purposes of simplicity, assume that the continuous convolution kernel $f(x, y)$ is approximated by the two dimensional mask $F(i, j)$ such that

$$F(i,j) \;=\; \begin{cases} f(i,j), & \text{if } f(i,j) \quad f_{\min}, \\ 0, & \text{otherwise} \end{cases} \tag{2.24}$$

for some convenient threshold $f_{\min}$, typically 1% of the maximum.

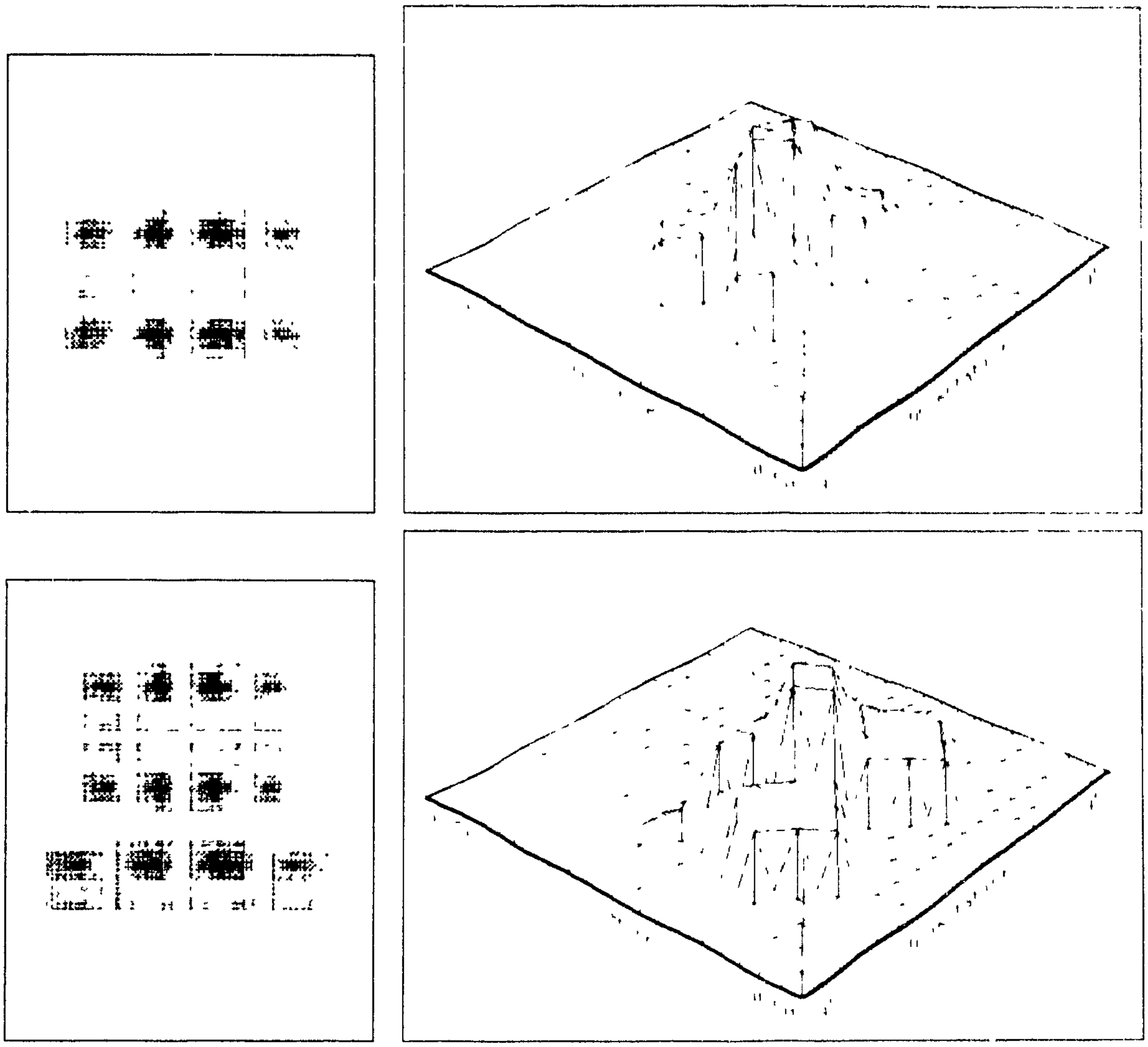
2.3 Tuning the Operator

The structure that we have established to this point defines a class of line operators which may or may not be tuned for curvature (we do, of course, have to represent straight lines). There is a small set of free parameters which must be fixed before the operator can be instantiated. A particular choice of the representational parameters described in §1.2 should completely determine the model parameters for the operator. However the model parameters allow somewhat more flexibility than these representational constraints require. In the following analysis, the desired output has an operator tuned to each of 8 orientations and 5 curvatures. The orientations are evenly spaced between 0° and 180°: $\mathbf{\Theta} = \{0, 22.5, 45, 67.5, 90, 112.5, 135, 157.5\}$ and the curvatures are distributed around zero as $\mathbf{K} = \{\ 0.2,\ 0.1,\ 0.0,\ 0.1,\ 0.2\}$ (positive curvature means curved to the right). Thus there are 40 different operators. However, since an operator at one orientation can be retuned to a different orientation by simply rotating around it center, the 40 operators will be based on only five sets of tuning parameters, one for each discrete curvature. Furthermore, there is symmetry around zero curvature so we need design only *three* tuned operators.

The first matter to resolve is what the optimal line width of the operator should be. For the combination of lateral components the optimal line width is 2σ of the associated d^2G operator. These operators can be scaled by simply varying the σ_u of the operator but there is a relationship between the range of curvatures which can be represented and the size of such an operator. The length of the operator is limited above by the radius of the maximum curvature represented, however the operator must be elongated in order to achieve any orientation specificity at all thus limiting its length below Finally, the increased noise-sensitivity of small operators must also be considered These factors force the operator design in conflicting directions Focussing our attention, for the moment, on the smallest scale of image features, I choose an operator width of 2 pixels This is as large as possible for adequately representing curves with curvature as high as 0 2 (radius of 5 pixels)

The structure of the length components of the operators must still be resolved as two parameters of this decomposition are still free, the degree of separation (ρ in equation (2 18)) and the number of components (n in equation (2 20)) It has been determined experimentally that optimal end-line stability in the presence of noise is achieved with $\rho \approx 3$ The number of length components n is at least 2 and must be even (the overlap between the two central components determines the amount of end-line stability) Furthermore, the length of these components must not be too great (greater than two or three times the width) or end-line stability becomes dependent on non-local events (e g a line that passes nearby) Assuming as much simplicity as possible in this case, I divide the cells into four regions unless this interferes with the orientation specificity of the components, in which case I use two With these criteria and a few simplifying assumptions the operators can be tuned

For simplicity assume that the inhibitory component of the ES operator is twice as long as the excitatory component (σ_x is twice the value), that both components have the same width, and that the center of the inhibitory component is laterally displaced by the same amount as the lateral shift of one of the lateral components of the excitatory operator (this is to ensure that the excitatory region of the inhibitory component is centered on the operator) Require further that the responses of each operator when presented with its design stimulus be normalized to 1 0, thus fixing the overall weight of the operator This reduces the number of parameters to tune for each cell to two σ_x (the length of the excitatory component), and the ratio of weights of the excitatory and inhibitory components

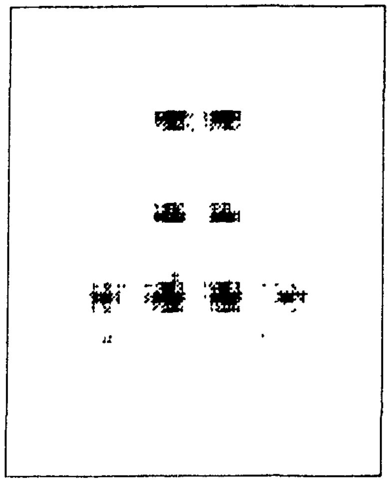


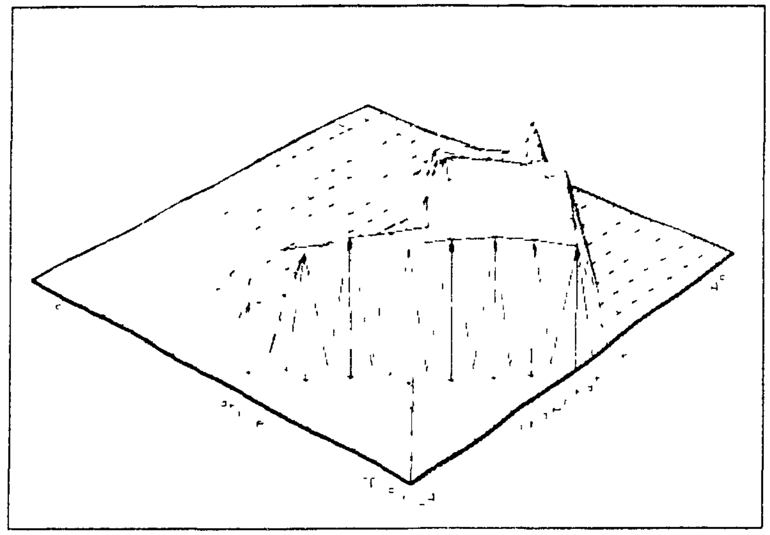
(a)

(b)

Figure 2.10 Response profiles of three end-stopped simple cells. The cells are shown exploded (components are separated) and are designed to respond optimally for curves with tangent orientation of 0° at the cell's center. The response envelopes shown are for anti-aliased high-contrast lines of width 2 pixels. Curvature and orientation are simultaneously varied over the ranges shown. The cell shown in (a) is designed for curvature 0.0 (straight lines); its response profile shows excellent locality when varying both orientation and curvature. The cell shown in (b) is designed for curvature -0.1 (arcs with radius 10 pixels below operator center). Note that for higher curvatures it becomes more difficult to obtain the same degree of locality of response.

(c_s/c_l in equation (2.23)). The tuning is done manuall using anti-aliased curves as probe stimuli. I plot tuning curves as the free parameters are varied and choose the parameters which exhibit the greatest conformity to the desired response envelope. In the future I hope to develop a set of numerical optimality criteria so that even this part of the procedure may





(c) (Operator) (Responses)

Figure 2.10 Response profiles of three end-stopped simple cells (contd.) The cell shown in (c) is designed for curvature -0 2 (arcs with radius 5 pixels below operator center) Note the decrease in locality of response associated with the higher curvature

be done automatically The cell parameters and response envelopes for the case we have considered here are shown in Figure 2 10

2.3.1 Intensity Edges

While the method developed is specific to image lines (especially the cross-section analysis), it can be easily used for image edges as well As was stated previously (§1 1) an image edge is the locus of points which are local extrema of the directional derivative of image intensity As the line operator is specific to local maxima, all that is necessary to create a similarly robust edge operator is to do a directional derivative perpendicular to the operator direction before using the line operator The line operator's responses to this new image will reflect the loci of maxima in this direction and thus the image edges By taking the derivatives in opposite directions (or equivalently by using either positive or negative contrast edge operators) the entire 360° range of orientation can be represented with exactly the same set of line operators as used for the direct discovery of lines

To ensure that the directional derivatives used are appropriate to the line operators which follow it is only necessary to restrict their size in such a way that they do not interfere with the localization behaviour of the line operator Thus, the directional derivative operator used has the same width σ_y as one of the cross sectional components of the line operator This gives a response region around a perfect step edge with the same width as the line operator The operator is twice the length of a single length component of the line operator and is decomposed into just two length components around the centre In this way, it is as large (and thus noise-insensitive) as possible without intefering with lateral localization and maintains the end-line stability of the pure line operator

2.4 Results

For the following examples the three base operators are

Table 1: Parameters for initial operators

	ES?	ES Component	σ_x	σ_y	# Regions	ρ	Weight
$\lvert\kappa\rvert = 0.0$	No		2 25	1 0	4	2.0	
$\lvert\kappa\rvert = 0.1$	Yes	Excitatory	1 5	1 0	4	2.0	1.0
		Inhibitory	3.0	1 0	4	2.0	1.33
$\lvert\kappa\rvert = 0.2$	Yes	Excitatory	1 0	1.0	4	3.0	1.0
		Inhibitory	2.0	1 0	4	2.0	2.0*

*Inhibitory displacement = 0 6

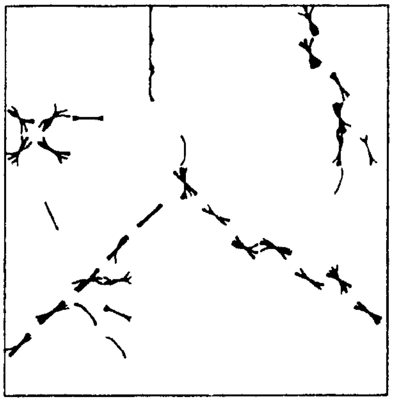
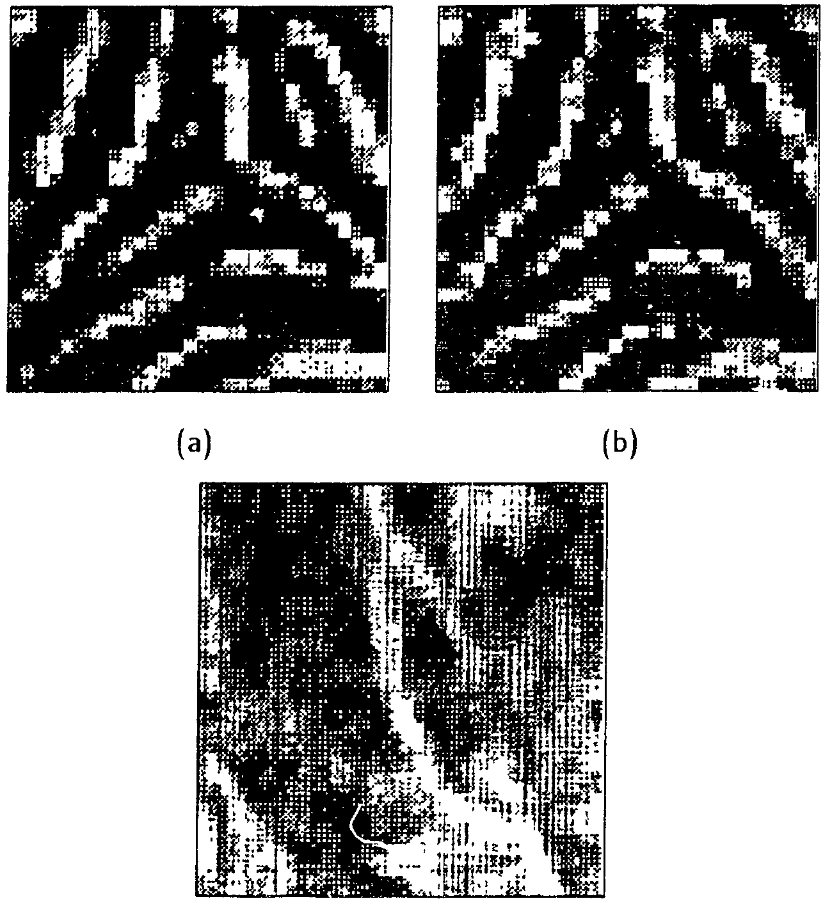


Figure 2.13 Line measurements on FPRINT (detail)



(a) (b) (c)

Figure 2.11 Test images are (a) FPRINT, a detail of a fingerprint image, (b) FPRINT degraded by additive gaussian noise and (c) ROADS a satellite image of logging roads

2.5 Discussion

The improvements in the non-linear operators over both the linear line operators and the Marr-Hildreth zero-crossings should be immediately obvious from an examination of the examples

Comparing the non-linear operator responses (Figure 2.12) with the linear version (Figure 2.14) is almost like comparing apples and oranges Even though the representation used for both outputs is identical, the non-linear operators are more specifically tuned, don't bridge the gaps between nearby lines, and don't run off the ends of the im-

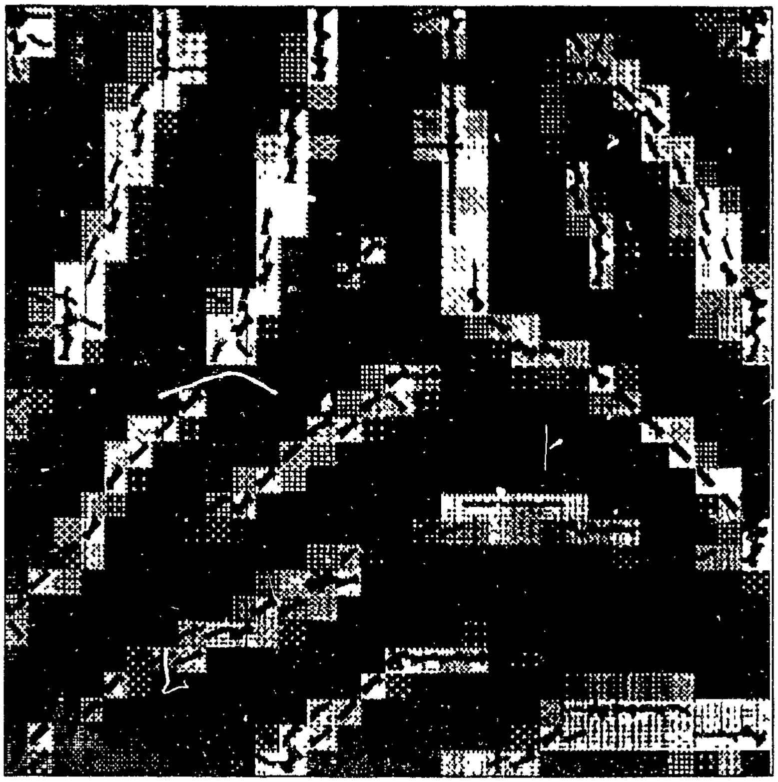


Figure 2.12 Line measurements on FPRINT a small section of a finger print image The line endings bifurcations and minimal inter curve spaces provide a serious test for the measurement operators

age lines Considering the fact that the linear operators were the starting point for the non-linear decompositions introduced in this chapter the linear/logical non linearities have clearly been effective in eliminating all but the 'best' from among the 'potential' responses The standard response to this problem (e g [Canny 86]) has been to adopt some kind of thresholding, but that leaves open the essential question how can one choose the threshold(s) in a principled manner, in advance?

The gains are similar for edges Compare the non-linear edge measurements in Figure 2 17 with the Marr-Hildreth zero-crossings of Figure 2 19 (both processes are run at exactly the same scale) Not only is the non-linear edge operator more accurate, it makes explicit a great deal of information which is simply unavailable in the zero crossing image Take connectivity, for example—how many arbitrary choices must be made in order to define a means of following the curves in the zero-crossing image? And again we are faced with the same question as with the line example discussed above—what threshold

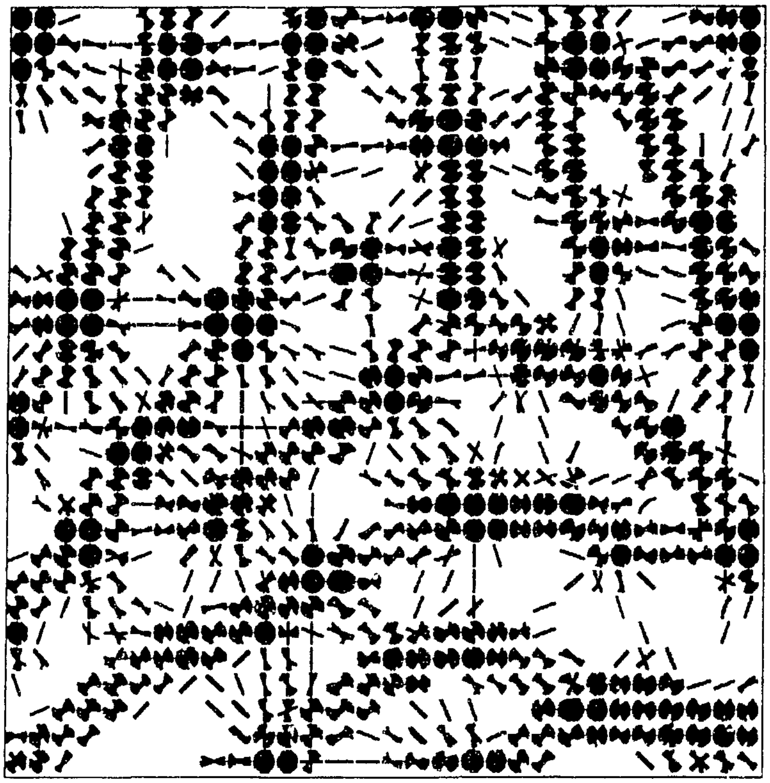


Figure 2.14 Linear line measurements on FPRINT using the linear correlates of the non-linear operators Note the excessive amount of bridging across gaps and the complete lack of specificity of response Prehaps it needs a threshold?

should one use?

While the alternatives considered are clearly inadequate, an examination of the detailed blow-ups (Figures 2 13 and 2 18) reveals both some of the strengths and inadequacies of these operators The obvious strengths (in both examples) are an accurate estimate of the positions, orientations and curvatures of the image curves Discontinuities and bifurcations in these curves are not smoothed over at all, instead they are represented as multiple tangents at the same point This is clearest in the edge image, where even inferred *edges* can stably co-exist at the same image position. However, it is clear that these measurements are still somewhat sensitive to noise and local perturbations, and do not always provide good estimates of local curvature

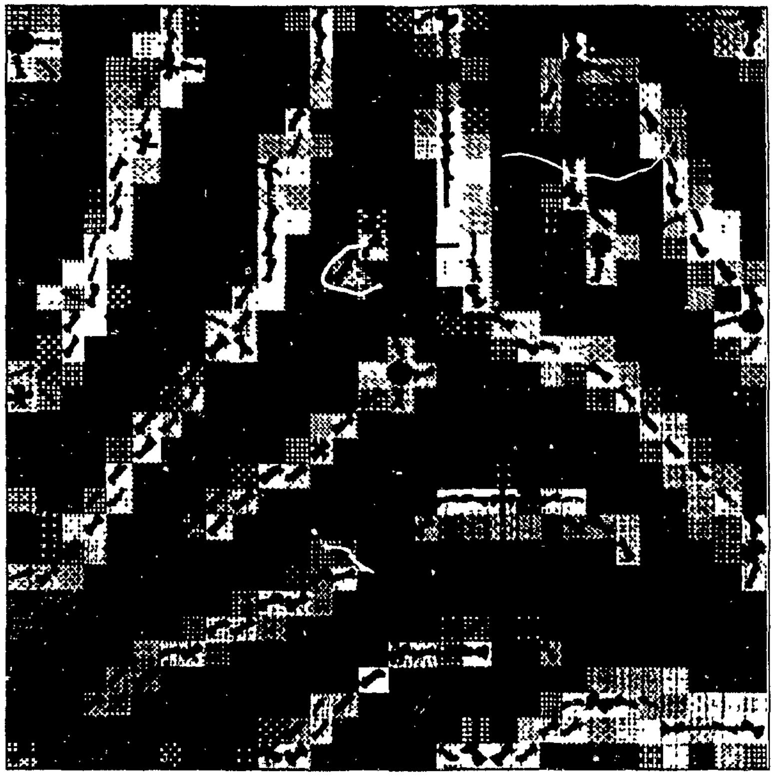


Figure 2.15 Line measurements on noisy FPRINT with additive gaussian noise ($\sigma = 5$) The image is quantized to 64 intensities

2.5.1 Biological Implications

A number of observations regarding the relationships between this model and natural visual systems are possible In general, there is nothing in the operator structure which is biologically *implausible*, in fact, some attempt was made (notably in Appendix B) to demonstrate how such a model could be instantiated in neurological terms The gross similarities between these oriented, contrast-tuned operators and cortical simple cells have already been noted The model of end-stopping which has been adopted to give these operators curvature-sensitivity was motivated largely as a biological model A number of more subtle observations can also be made

An indication that at least one of the non-linearities proposed may have a neurological correlate comes from the research of [Hammond & Mackay 83 & 85] In doing single-cell recordings using illuminated bars as stimuli, they observed that a small opposite

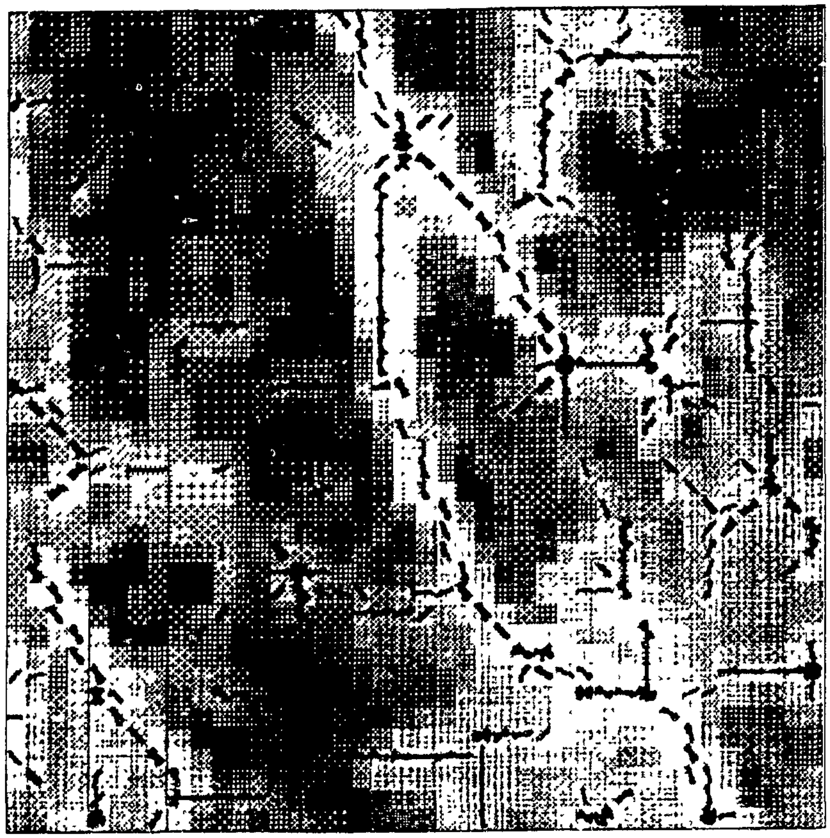


Figure 2.16 Line measurements on ROADS, a small section of a satellite image of logging roads in northern Québec More general then the fingerprint image, it clearly contains features which are not lines

contrast region in the midst of a preferred stimulus was able to inhibit the response of a simple cell much more effectively than would be predicted by a linear model The opposite contrast region was, in effect able to turn the cell off This is exactly the behaviour caused by the tangential partitioning of the operator

The reader may have noticed that in Figure 2.3(c) the response region is offset from the actual peak of the 1D signal by a small amount In light of the insistence that the operator respond only at peaks, this offset may seem problematic until the significance of an observation made in [Watt & Morgan 83] and [Whitaker & Walker 88] is recognized. It was observed that when a discrimination task involved locating the position of dot clusters or lines with non-constant intensity cross-section, the locations chosen were best described as the centroids of the intensity distributions The behaviour seen with this model is consistent with this interpretation, as the centroid of the triangular region is ≈ 56 which is

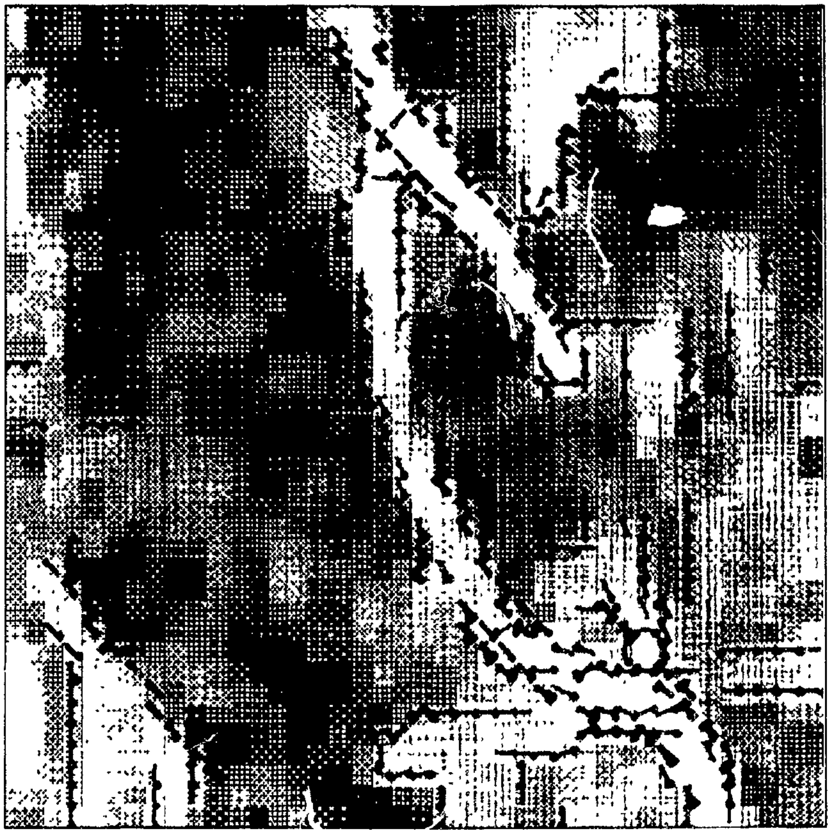


Figure 2.17 Edge measurements on ROADS using the technique described in §2 3 1 Note that the segments are now directed (if the small circle is the tail of the vector then facing along the vector the edge goes from dark to light from left to right)

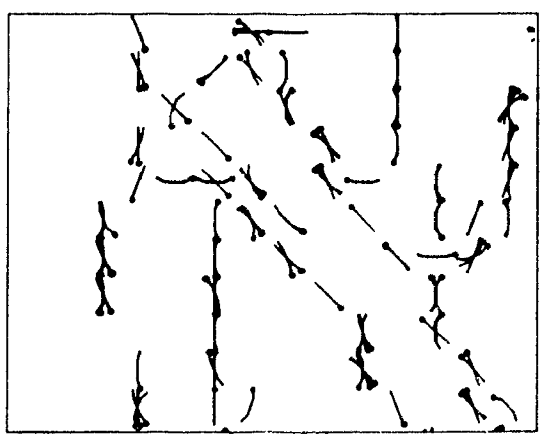


Figure 2.18 Edge measurements on ROADS (detail)

definitely inside the region of positive response. To see why the model should exhibit this behaviour, notice that the peak of a gaussian convolution with a positive signal approaches

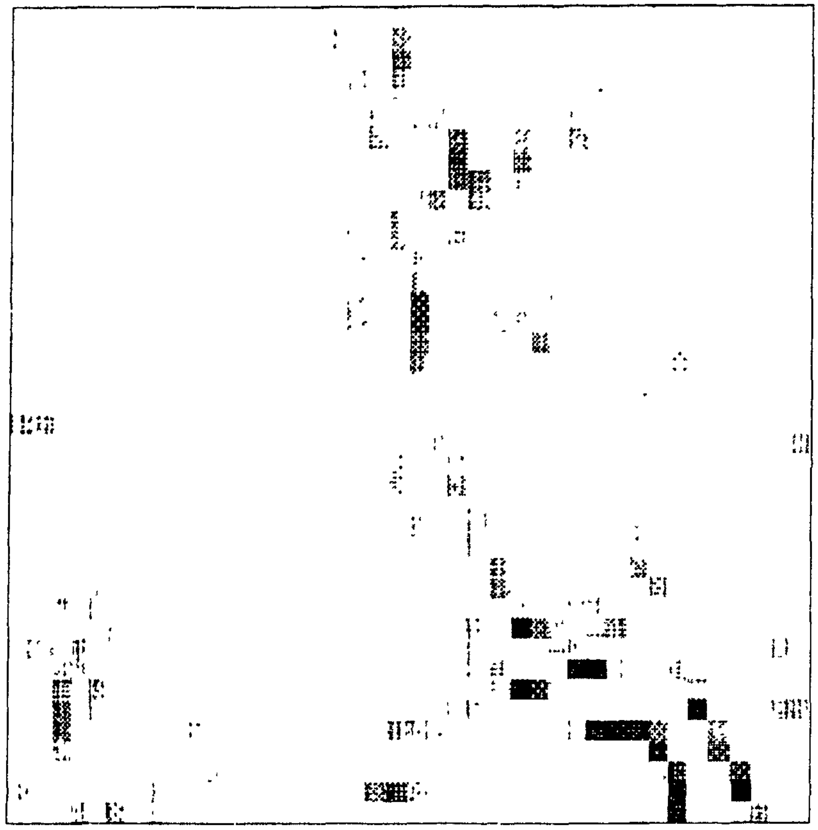


Figure 2.19 Marr-Hildreth zero-crossings on ROADS at exactly the same scale as the non-linear operators above The pixel intensities represent the significance of the zero-crossings

the centroid of the signal as the width of the gaussian increases As long as the gaussian is wide enough with respect to the signal, the peak of its response (as determined by the peak detection property of the line operator) will be close to the centroid of the intensity distribution This relationship should be investigated further

2.5.2 Conclusions

This chapter has outlined a simple, principled response to one of the major failings of linear operators in locating image curves By systematically responding to stimuli which are not within the design parameters of the operator, linear approaches cause problems for higher level processes which must distinguish between 'true' and 'false' operator responses By outlining two of the necessary conditions for the existence of an image curve, and by demonstrating an operator decomposition which allows for the efficient testing of these conditions, this outstanding issue has been neatly resolved. Furthermore, it has been demonstrated that this nonlinear operator can be used as a component in the Dobbins ES operator. The operator can thus be used as and efficient and robust mechanism for confirming the pointwise existence of image curves with an orientation and curvature as determined by the operator tuning. Care has been taken throughout to keep the formulation free of assumptions regarding the geometry or discreteness of the retinal detector grid, and a mechanism for generalizing these results to arbitrary grids was outlined.

In spite of these results, the goal of achieving the same combination of robustness and sensitivity as found in a mammalian visual system is almost certainly not obtainable with such a simple one-shot operator. These operators can however, provide the initial estimates for a set of existence hypotheses which will be examined and modified by a highly parallel relaxation network. Such a network will operate by seeking higher level consistencies based on the local differential geometry of plane curves and not simply on the patterns of light falling on a retinal array. This network is described in the proceeding chapter.

Chapter 3 Relaxation

3.1 Introduction

Any simple, direct operator for measuring local image characteristics will exhibit errors due to noise and imaging defects In a system subject to the same performance requirements as the human visual system, however, even these errors are unacceptable In Chapter 2 I developed an estimator for the local differential characteristics of image curves which limits such characteristic errors It still suffers from errors due mostly to the influence of noise Such measurement errors are therefore uncoordinated and local, and a system which refines these measurements by seeking *consistent patterns* of response should recognize and correct such errors and begin to exhibit the kind of robustness required The same structural arguments can be made with respect to any other perceptive system seeking characteristic forms and structures in its impinging stimulus In the case of image curves, the patterns represent consistent models of piecewise continuous planar curves

From a computational point of view it is important to discover a well-behaved methodology which will discover these consistent structures It is similarly important that this methodology is as parallel as possible, for speed is every bit as important as accuracy *Relaxation labelling*[1] is such a methodology A gradient ascent procedure, it is an incremental, iterative parallel update formula for constraint satisfaction It is simple

[1] Speling? Only in Canada you say? Pity

yet powerful and has been used for a variety of problems which benefit from such a highly parallel solution The only questions that remain once we adopt relaxation labelling are 'how to structure the network?', and 'how fast will it converge?' Previous approaches to designing relaxation labelling solutions to local curve description can be found in [Zucker *et al* 77] and [Parent & Zucker 85] The designs of each of these systems were specific to the curve description problem This thesis will demonstrate an approach to designing such a network that rests firmly on the geometry of this problem while still retaining enough generality to allow for its application to a range of similarly constrained problems It will also demonstrate a surprising answer to the question of convergence As was stated previously, the entire visual process ('seeing') must be completed in less than 100ms It would make little sense for a small subprocess in this pipeline to take more than a small fraction of this time, in the brain however the processing elements, the neurons, have time constants on the order of ~1ms Thus, even a very robust relaxation system which takes as few as 50 iterations to converge is completely unrealistic as a neurological model In contrast, this process converges to stable structures within as few as 3 or 4 iterations thus allowing implementation as a simple feed-forward network taking up no more than a few layers of neural circuitry and a very small fraction of the available processing time

3.2 Relaxation Labelling

Relaxation labelling is a network model which has been in use since the early 1970's Given a set of nodes I, a set of possible labels for each node Λ_i, and a measure of the compatibility $r_{ij}(\lambda, \lambda')$ between labels $\lambda \in \Lambda_i$ and $\lambda' \in \Lambda_j$ at nodes $i \in I$ and $j \in I$ respectively, relaxation labelling is a gradient ascent procedure for maximizing the consistency between all labels By defining the nodes, labels and compatibilities to reflect a given set of constraints between objects relaxation labelling will solve the constraint satisfaction problem so defined Even though the basic methodological framework has been in use since the early '70s the first formal statement of this general problem and its solution is due to [Hummel & Zucker 83]

For the *continuous* relaxation labelling processes which they investigate there is a *confidence* associated with each label, given as $p_i(\lambda)$. These confidences are restricted at each node such that

$$\forall(i \in I), (\lambda \in \Lambda_i), \; 0 \leq p_i(\lambda) \leq 1 \tag{3.1}$$

and

$$\forall(i \in I) \sum_{\lambda \in \Lambda_i} p_i(\lambda) = 1 \tag{3.2}$$

Thus each set of label confidences lies on a plane in the positive sector of an n-dimensional space, where $n = |\Lambda_i|$, the number of labels at node i.

With real valued compatibilities r_{ij} the *support* of label λ at node i is

$$s_i(\lambda) = \sum_{j \in I} \sum_{\lambda' \in \Lambda_j} r_{ij}(\lambda, \lambda')\, p_j(\lambda') \tag{3.3}$$

A *labelling assignment* is a set $\mathbf{p} = \{p_i(\lambda) \mid i \in I \wedge \lambda \in \Lambda_i\}$ such that the conditions of (3.1) and (3.2) hold. The subset of labelling assignments such that

$$p_i(\lambda) = \begin{cases} 1, & \text{if node } i \text{ maps to label } \lambda, \\ 0, & \text{if node } i \text{ does not map to label } \lambda \end{cases} \tag{3.4}$$

thus forms a convex hull of $\mathbf{p}$ which unambiguously determines a one-to-one mapping from node to label, an *unambiguous labelling assignment* $(p_i(\lambda) = 1) \Rightarrow (i \rightarrow \lambda)$. Taking an unambigous labelling assignment which maps nodes $1, \ldots, n$ to labels $\lambda_1, \ldots, \lambda_n$, a *consistent* labelling assignment is then one that fulfills the condition that

$$\forall(i \in I), (\lambda \in \Lambda_i) \cdot s_i(\lambda_i) \geq s(\lambda) \tag{3.5}$$

It is important to note that there may be many different consistent labellings for a given set of nodes, labels and compatibilities. These states correspond to local maxima in a measure known as average local consistency. Thus, the problem must be framed in such a manner that every local minimum is a point of interest with respect to the problem being solved. The system then relaxes very quickly to a local minimum near the starting point.

The entire algorithm for relaxation labelling is presented as

Algorithm 1: Relaxation Labelling

1 Compute an initial estimate of all $p_i(\lambda)$ Call this $p_i^0(\lambda)$ (N B These values must conform to the restrictions of (3 1) and (3 2))

2 Repeat starting with $n = 0$ until the labelling is unambiguous:

  1 Repeat for all $i \in I$

    1 For all $\lambda \in \Lambda_i$ compute $s_i(\lambda)$ from (3 3)

    2 Compute $\mathbf{p}_i' = \mathbf{p}_i^n + \delta \mathbf{s}_i$

    3 Project $\mathbf{p}'$ into the space of valid confidences This projection is the updated vector $\mathbf{p}_i^{n+1}$

  2 Set $n = n + 1$

3 Compute the mapping $i \to \lambda_i$ such that $\forall(\lambda \in \Lambda_i)\ p_i(\lambda_i) \geq p_i(\lambda)$

The projection in step 2 1 3 is, in its general form too complicated to describe here fully We will instead (in §3 3 1) describe the projection for the specific case of two label systems

The primary concern in the design of the operators in Chapter 2 was the prevention of aliasing It now becomes clear why such effort was expended The relaxation labelling process uses gradient ascent to discover a consistent labelling near the initial position If this initial position is derived from a set of image operators which have *consistent* patterns of aliasing responses, the relaxation process cannot recover from these initial errors It will simply incorporate these patterns into its solution If instead the operator aliasing is limited so that the only spurious responses are due to noise and small local perturbations then these events will not propagate through the relaxation but will instead be rejected as inconsistent with neighbouring hypotheses

One of the major advantages of relaxation labelling is that the update can be computed entirely in parallel. The sums of equation (3 3) are independent of order of evaluation and the supports for each of the labels could, therefore all be calculated simultaneously Assuming n nodes, m labels at each node, and a support network of k non-zero compatibilities for each label, this would allow a speedup of $n \times m \times k$ for the support calculations Because it is performed at the node level, this parallelism would reduce to n for the projection (step 2 1 3), but the lion's share of the computation is in the support calculations, which is where we have the most to gain from parallelism The entire relaxation system is thus highly amenable to parallelization or implementation in specialized hardware

3.3 Representation

In instantiating a relaxation labelling system, one is faced with the definition of factors the nodes labels and compatibilities The nodes and labels are determined by the structure of the solution being sought, while the compatibilities depend on the internal constraints of the problem In this case, the geometric structure of planar curves first determines the representation and then the compatibilities

In this case the labels in the relaxation correspond to the discretization of the differential space of the image as described in §1 2 Nodes in the relaxation thus correspond to specific positions in the sensor array, and the node positions tile the area covered by the sensors The labels at each node correspond to hypotheses regarding the existence/non existence of image curves with a set of local differential properties (orientation and curvature) at that image position

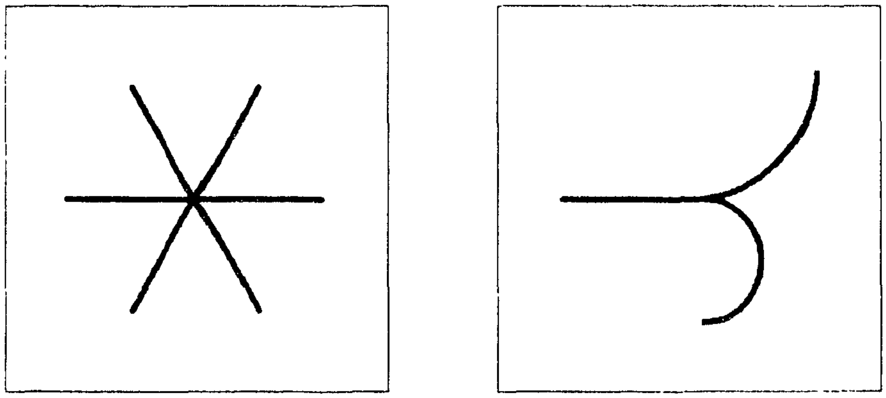
The importance of bringing the local differential geometry of plane curves to bear on this problem has already been stated—the operators of Chapter 2 reflect this conviction They provide an initial measurement of the confidence in each of the labels in the system The choice of how to organize these labels is more problematic The options are outlined as follows

1 Each image position could be a single node in the system, with the labels representing each of the different choices as to orientation (and curvature) of the curve/s passing through that image position (the option taken by [Zucker, Hummel and Rosenfeld 77])

2 Each image position could have a set of nodes corresponding to the set of allowed orientations, with the labels being associated with curvature values (or vice-versa with curvature nodes and orientation labels, though this choice seems bizarre) Something like this option was used in [Parent & Zucker 86]

3 Each image position could have a set of nodes corresponding to pairs of orientation and curvature There would be two labels at each node, which would reflect confirmation/rejection of the hypothesis that a curve passes through that position with the given pair of differential characteristics This is the option chosen

Each of these choices will adequately describe the curvature tangent field, but when con sideration is given to the limitations of each possible representation and its implications for relaxation, the reasons for the choice become clear



(a) (b)

Figure 3.1 Minimal pairs of intersecting image curves demonstrate the need to allow (a) multiple curves passing through a single point which differ only in orientation and (b) multiple curves passing through the same point which differ only in curvature

In considering the properties of planar curves we have already mentioned the importance of their local differential properties, but these are properties of curves consid ered in isolation The considerations which force a choice between our representational options arise instead from consideration of the interactions between nearby image curves Remember that in a relaxation labelling system, the final result is a one to one mapping from nodes to labels Therefore two labels for the same node cannot co exist in a consis tent labelling Consider the two configurations in Figure 3 1, each of which will arise in natural images In Figure 3.1(a) two curves cross at a point thus revealing the necessity of representing a situation in which differ minimally in orientation co-exist at the same image position This immediately rules out option 1 Figure 3 1(b) similarly illustrates the neces sity of representing a situation in which two curves that differ minimally only in curvature co-exist at the same image position This similarly obviates the use of Option 2 Neither orientation nor curvature can be considered primary, minimal pairs exist for variation in both parameters We are left only with the third option

There are a number of additional advantages to this representation which can now be elucidated. The first of these is that it removes the problem of deciding how to handle the 'no-curve' label. Relaxation labelling produces a consistent labelling as output, and in every consistent labelling exactly one of the labels at every node is selected. Since image curves do not exist at every point in an image there must be a label at each node which represents the possibility that *no curve* with the required characteristics exists in the image. It has never been clear how to incorporate such a 'no-curve' label into a label set which attempts to distinguish between another set of mutually exclusive possibilities. For example, with option 2 above, each label set would have to contain one label for each discrete curvature value *plus* a label which represents the hypothesis that no curve exists which passes through the image position with the appropriate tangent orientation. The only methods previously used rely on some sort of arbitrary thresholding (see [Zucker, Hummel & Rosenfeld 77] and [Parent & Zucker 85]). However, when the labels have TRUE/FALSE semantics, the compatibilities need only confirm or deny a *single* hypothesis associated with each node, a much simpler problem and one not prone to so many potentially arbitrary design decisions. Another substantial design advantage of a representation with only yes/no labels is that the projection operation (the least appealing aspect of the relaxation labelling framework because of its unfortunate sequentiality) can be reduced to a simple arithmetic update rule.

3.3.1 Projection with Two Labels

The projection operator upon which the theory of relaxation labelling depends is detailed in [Mohammed, Hummel & Zucker 83]. The operator determines an update vector $\mathbf{u}$ with $\|\mathbf{u}\| = 1$ which maximizes $\mathbf{s} \cdot \mathbf{u}$ and leaves $\mathbf{p}' = \mathbf{p} + \mathbf{u}$ in the space of valid $\mathbf{p}$ vectors. The direction of $\mathbf{s}$ (the gradient of $A(\mathbf{p})$) thus determines the direction of the update vector $\mathbf{u}$ as well as possible while still remaining valid. This is accomplished for general label sets by successively reducing the dimensionality of the update vector when the update extends beyond a border. However when the label vector is only two-dimensional, then the successive reductions become unnecessary, the projection can be computed in a single operation

Refer to the two labels at a node as **T**(true) and **F**(false). Suppose the support vector is $\mathbf{s}_i = (s_i(\mathbf{T}), s_i(\mathbf{F}))$ and the confidence vector is $\mathbf{p}_i = (p_i(\mathbf{T}), p_i(\mathbf{F}))$. Projecting $\mathbf{s}_i$ onto the tangent plane (r, x) gives $s_i'(\mathbf{T}) = (s_i(\mathbf{T}) - s_i(\mathbf{F}))/2$ and $s_i'(\mathbf{F}) = (s_i(\mathbf{F}) - s_i(\mathbf{T}))/2$, and the update direction depends only on the difference $s_i(\mathbf{T}) - s_i(\mathbf{F})$. This allows an enormous simplification both in storage and computation. Since $p_i(\mathbf{F}) = 1 - p_i(\mathbf{T})$ we can save space by keeping track of only $p_i(\mathbf{T})$, reducing the vector $\mathbf{p}_i$ to the scalar p_i, and computing the projected sum as $p_i' = p_i^n + \delta(s_i(\mathbf{T}) - s_i(\mathbf{F}))/2$. Ensuring that $p_i^{n+1} = \max(0, \min(1, p_i'))$ projects the result back into the valid confidence space. Assuming further that support will be positive when the label hypothesis is confirmed and negative otherwise, we can safely assume that $s_i(\mathbf{F}) = -s_i(\mathbf{T})$ (i.e. a positive confirmation implies a negative denial) and the update simplifies even further. We thus arrive at a very simple update operation which combines steps 2.1.2 and 2.1.3 of the algorithm into

$$p_i^{n+1} = \max(0, \min(1, p_i^n + \delta s_i)) \tag{3.6}$$

3.4 Compatibilities

The representation thus chosen maps retinotopic visual space onto a set of labelled nodes. Each node represents an independent hypothesis as to the existence of a curve in the retinal image which passes through the associated position with a specific orientation and curvature. Given a grid of retinal positions **R** and a set of orientations **Θ** and curvatures **K**, then each node corresponds to a four-dimensional point (X, Y, Θ, K) with $(X, Y) \in \mathbf{R}$, $\Theta \in \mathbf{\Theta}$ and $K \in \mathbf{K}$. Each node (hypothesis) i has a confidence value referred to as either p_i or $p(X_i, Y_i, \Theta_i, K_i)$. Assuming the same 8 orientations and 5 curvatures as in §2.3 the set of hypotheses which co-exist at each position in the retinal image is shown in Figure 3.2. The operators of Chapter 2 are used to provide the initial confidence estimates, each hypothesis being paired with the appropriately tuned operator. The task of the relaxation labelling is to discover certain consistent structures within the initial measurements in this four-dimensional space. Such structures are connected subsets of the discrete points which correspond to smooth segments of the image curves as described in §1.2.

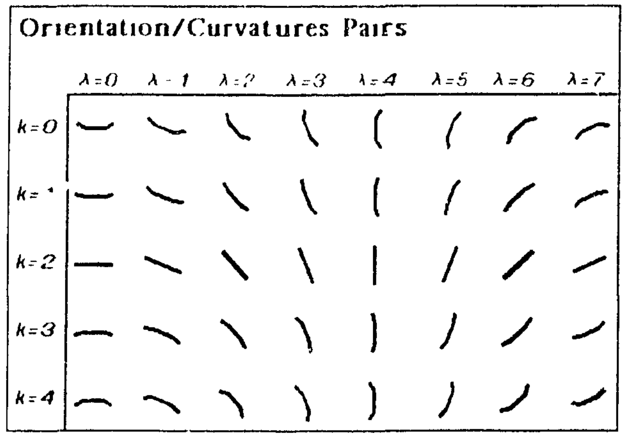


Figure 3.2 Orientation/curvature pairs for 8 discrete orientations (22.5° spacing) and 5 curvatures $\{-0.2, -0.1, 0.0, 0.1, 0.2\}$

3.4.1 The Cocircularity Constraint

Image curves as defined in §1.1 are continuous and piecewise smooth. Representing these curves in terms of a set of discrete labels, the operators of Chapter 2 measure the local structure of these curves in terms of trace (position), orientation and curvature. We have asserted however that these measurements are inadequate as they do not necessarily fulfill the strict continuity and smoothness constraints, and therefore do not necessarily represent actual image curves. These conditions are no longer phrased in terms of the intensity disributions in the image, but as constraints describing the relationships between nearby hypotheses as to the local differential properties of image curves. Relaxation labelling, as described above, is the means by which these more abstract constraints are satisfied. In order to represent the constraints in a relaxation network it is necessary to translate the natural relationships between the differential properties at different points along a planar curve into measures of the *compatibility* between neighbouring discrete labels. We begin this development by describing the geometric property of *cocircularity*, first elucidated in [Parent & Zucker 85].

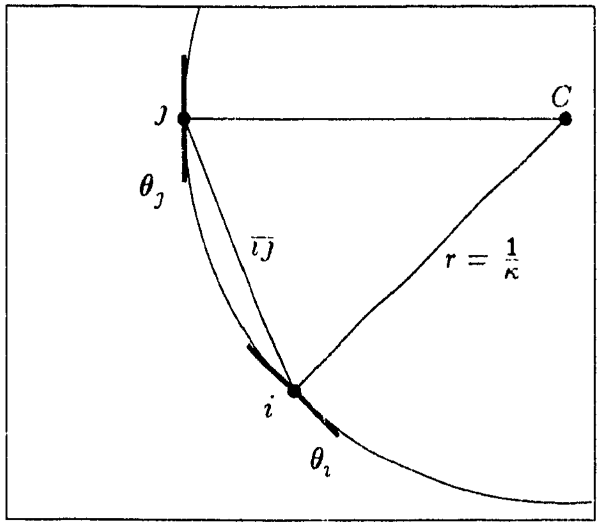


Figure 3.3 Geometric cocircularity constraint relates the orientation θ_i and curvature κ_i at point i to θ_j and κ_j, the same measures at point j Cocircularity of the two curved tangents depends on their being tangent to the same circle

Given two tangents in a tangent field at i and j, then these tangents are said to be *cocircular* if and only if they share a common osculating circle (see Figure 3 3) In a curved tangent field, the requirement is even more specific, for the curved tangents must share the same centre of curvature If the tangent vectors at i and j have orientations θ_i and θ_j and curvatures κ_i and κ_j respectively, this condition can be specified as a set of coupling equations between these values Notice that cocircularity does not depend on absolute position, but only on the relative positions of the two curved tangents $[x_j - x_i, y_j - y_i]$ (in polar coordinates this is $[d_{\overline{ij}}, \theta_{\overline{ij}}]$) Beginning with the orientation at i and the relative position (a tangent line and a point completely determine their osculating circle), one can successively restrict parameters until the relationship is fully defined This procedure leads to the coupling constraints†

$$\theta_j \equiv 2\theta_{\overline{ij}} - \theta_i \tag{3.7.1}$$

$$\kappa_i = 2\sin(\theta_{\overline{ij}} - \theta_i)/d_{\overline{ij}} \tag{3.7.2}$$

$$\kappa_j = \kappa_i \tag{3.7.3}$$

† In these coupling equations, it is essential to remember the cyclic nature of orientation angles Thus for two angles $\theta_1 \equiv \theta_2$ is equivalent to $\theta_1 \bmod 2\pi = \theta_2 \bmod 2\pi$

By making the relationship between these parameters explicit in this way one can investigate the geometry of the constraints themselves

3.4.2 From Constraints to Compatibilities

Given a set of problem-defined constraints, as defined in equations (3.7), these must be translated into relaxation labelling compatibilities. For a given pair of labels and their associated locations in the geometric space, the compatibility of one with the other is derived directly from the geometric constraints of the problem Even though the current analysis is specific to developing compatibilities for locally cocircular image curves, one of the goals of this work is to establish a methodology which is independent of the specific geometry of this case It should be clear that the formulation developed below is free of the sort of problem-specific factors that often plague these systems

In comparing two curved tangents, there are six free parameters (Θ_i, K_i, Θ_j, K_j and relative position X_j and Y_j) which together form a six-dimensional parameter space $\mathcal{V}$ Each constraining equation in (3.7) reduces the dimensionality of the space by one, so the space of parameter values which conform to the cocircularity constraints (called $\mathcal{C}$) forms a three-dimensional differentiable manifold in $\mathcal{V}$ One of the difficulties in translating these constraints into compatibilities now becomes apparent Each pair of relaxation labelling hypotheses to be considered for compatibility corresponds to a distinct point in $\mathcal{V}$ Observe, however, that since the codimension of $\mathcal{C}$ in $\mathcal{V}$ is 3 the intersection between the set of all hypothesis pairs (a set of dimension 0) and $\mathcal{C}$ has measure 0 This might not be a problem if we were able to choose the hypotheses for maximum intersection, instead we must allow for the case when the hypotheses are randomly distributed A test for *exact* cocircularity must, therefore, be overruled. We must instead derive a simple, general methodology for relating points in $\mathcal{V}$ to the manifold $\mathcal{C}$ that provides a *measure* of cocircularity

3.4.3 Perpendicular Projection

Such a method is based on *perpendicular projection*. A description of the relationship between a point and a manifold can be obtained by projecting the point onto the manifold; perpendicular projection in particular, takes a point to the *closest* point on the constraining manifold. Conceptually then the perpendicular projection represents the minimal perturbation of the point that makes it conform to the constraints embodied in the manifold. This is made clear by the following definition. The fixed point in $\mathcal{D}$ which represents the relationship between two curved tangents is $\mathbf{X} = [\Theta_i, K_i, \Theta_j, K_j, X_j, Y_j]$. Considering a point $\mathbf{X}' \in \mathcal{C}$, if $\mathbf{X}'$ minimizes the distance

$$D = \| \mathbf{X} - \mathbf{X}' \|_2 \tag{3.8}$$

then $\mathbf{X}'$ is the perpendicular projection of $\mathbf{X}$ onto $\mathcal{C}$. It should be clear that only when $\mathbf{X} \in \mathcal{C}$ do we have $d = 0$ and $\mathbf{X} = \mathbf{X}'$. Otherwise, intuition suggests that compatibility should be related to the distance D, with smaller distances giving larger compatibilities.

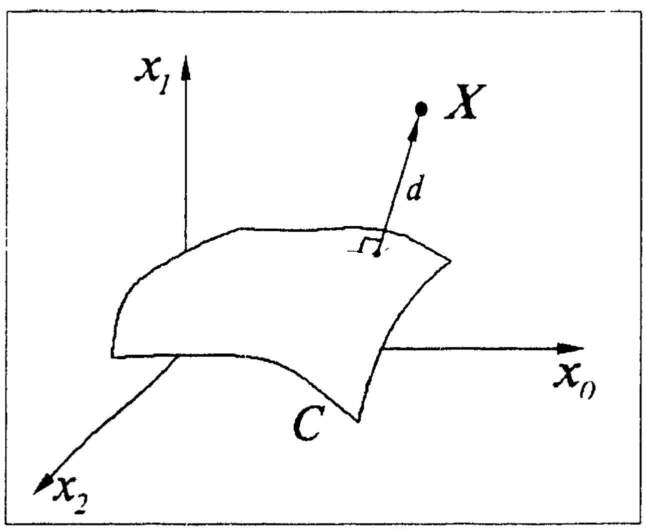


Figure 3.4 Parallel projection of point onto constraint space is shown in a three-dimensional parameter space. The point X is projected onto the manifold $\mathcal{C}$. The distance d is therefore minimal.

As a first approximation this analysis is straightforward and simple. The only difficulty arises with what should be an obvious objection, the different parameter axes can not be compared in such an offhand manner How can one compare a difference of 0.1 in curvature with a difference of 20° in orientation? It is not possible without reference to the spacing of the discrete label values in the parameter space—a perturbation of one discrete unit on each of the parameter axes is equivalent Thus if the spacing of the discrete values of the variable x can be specified with the invertible function $x(i)$, then the normalized distance between any two values x_1 and x_2 is given by

$$\Delta x(x_1, x_2) = x^{-1}(x_2) - x^{-1}(x_1) \tag{3.9}$$

For constant spacing, expressed as $x(i) = c + i\Delta x$, this simplifies to

$$\Delta x(x_1, x_2) = (x_2 - x_1)/\Delta x \tag{3.10}$$

The norm of equation (3 8) is then an L_2 norm of the difference vector made up of the set of $\Delta x(x, x')$ for each parameter x

The parallel projection above can thus be treated as the minimization of the normalized distance between the geometric values associated with the label pairs and the constrained continuous values In this case, the distance D[†] is

$$\begin{aligned} D^2 = & \left(\frac{\theta_i - (\Theta_i + 2a\pi)}{\Delta\theta_i}\right)^2 + \left(\frac{\kappa_i - K_i}{\Delta\kappa_i}\right)^2 \\ & + \left(\frac{\theta_j - (\Theta_j + 2b\pi)}{\Delta\theta_j}\right)^2 + \left(\frac{\kappa_j - K_j}{\Delta\kappa_j}\right)^2 \\ & + \left(\frac{x_j - X_j}{\Delta x_j}\right)^2 + \left(\frac{y_j - Y_j}{\Delta y_j}\right)^2 \end{aligned} \tag{3.11}$$

Details of this minimization are described in Appendix C

3.4.4 Compatibilities from Projections: Localization

Even having determined the perpendicular projection, it remains a problem to convert this description of the relationship between the pair of labels and geometric cocir-

† Because the comparison of orientation values is modulo 2π the integral factors a and b are introduced to allow differentiability They are otherwise arbitrary and thus the minimization of D^2 takes place for a and b varying over $\{-1, 0, 1\}$

cularity into a measure of compatibility. It is essential to consider the desired behaviour of the relaxation system in order to arrive at a satisfactory answer

Referring back to the representation described in §1 2 it should be quite clear that the path of an image curve through this four-dimensional space $x \times y \times \theta \times \kappa$ is a curve (and not a surface or volume), the discovery of this curve, the goal of the relaxation process, must therefore be constrained by the geometry of this curve In particular, a one dimensional object in a four-dimensional space has three dimensions in which it is localized and one in which it has non-zero extent This principle of *localization* of the inferred curve is the most important in translating the geometry into compatibilities

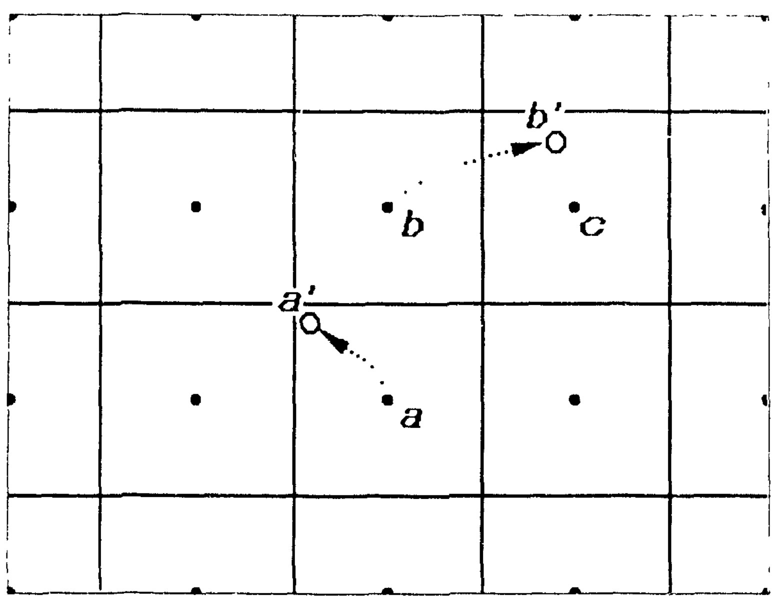


Figure 3.5 Conmpatibilities in terms of labelled points in the continuous parameter space The plane here represents the continuous space of parameter values and each point is a discrete labelled point in this space For this regular tiling the Voronoi cells of the label points are grid boxes The point a when constrained by a cocircular relationship with some other point p (not shown) projects to the point a' which is still inside the Voronoi cell defined around a The point a is thus *compatible* with the other point not shown The perturbation of b to b' however takes b out of its own Voronoi cell (b' is now closer to c) and thus b is *incompatible* with the point p

In general the compatibilities must, as shown in §3 3 1, confirm or deny the local hypotheses which they are involved in updating By considering the perpendicular projection described above in terms of local perturbations, the approach adopted below can be illustrated. Consider the description of the relationship between two labelled points in the parameter space $\mathbf{X}_{i,j} = [\Theta_i, K_i, \Theta_j, K_j, X_j, Y_j]$ The perpendicular projection of this relationship onto the constraining manifold gives the projected point $\mathbf{X}' = [\Theta'_i, K'_i, \Theta'_j, K'_j, X'_j, Y'_j]$ Translating this projection back into the image curve representation at i implies that the minimal perturbation of $\hat{\imath} = [\Theta_i, K_i, X_i, Y_i]$ that is cocircular with a similarly perturbed $\hat{\jmath}$ is $\hat{\imath}' = [\Theta'_i, K'_i, X_i + X'_j, Y_i - Y'_j]$ (see Figure 3 5 for a two-dimensional illustration of these perturbations) The label indexed by $\hat{\jmath}$ is thus compatible (has positive r_{ij}) with $\hat{\imath}$ only if the perturbed point $\hat{\imath}'$ is still in the Voronoi cell for $\hat{\imath}$ The labels are likewise incompatible (r_{ij} 0) in all other cases In practice, a threshold is chosen and only those interactions with compatibility (or incompatibility) more significant than the threshold are used in the support network Furthermore the shape of the compatibility cross-section must ensure that the relaxation converges on properly localized subsets of the labels

Image curves have a unique position and set of derivatives at every point (excepting the rare singular points)—the connected set of Voronoi cells through which the curve passes must, therefore, have a *lateral* extent of no greater than two‡ adjacent cells at any point Referring back to the trace of the positive responses of the initial estimators, no such condition is necessarily imposed Therefore, in addition to discriminating between true and spurious responses in the initial estimator outputs, the relaxation must localize the true responses on three parameter axes lateral position local orientation and local curvature

Considering each of these three localization conditions as separate one-dimensional localization problems, Appendix localize-app explores a number of possible compatibility cross-sections that could achieve the required localization through relaxation. It is demonstrated by example that the task of arriving at such localization in a two-label system is

‡ Not one but two cells This is as a result of the fact that the set S is made up of surface-connected and not corner-connected cells (analogous to the 4-connected/8-connected distinction in pixel-based images) Restricting to one-cell widths might destroy the connectedness of the representation of a continuous curve

non-trivial and depends on the addition of the same linear/logical non-linearities as were developed for the cross-sections of the operators in Chapter 2 Since such operators only respond positively within the local neighbourhood of a peak, the relaxation moves inexorably to a stable, consistent labelling centered around such peaks in the confidences --every label not in such a neioghbourhood is suppressed Since for orientation selection, the localization is required on three orthogonal axes, the compatibility structure developed below involves such a non-linear separation on three independent axes These partitions are done independently

The only difficulty with the use of such an update rule is that the linearity of the support functional which Hummel & Zucker used to prove convergence of relaxation labelling has been violated It will be necessary to follow up on this and prove that the process will still reliably converge with this non-linear support function The experiments in this thesis indicate not only convergence but exceptionally fast convergence, and the analysis above indicates why this is should be expected It remains to be proven experimentally that such a functional is *provably* convergent

3.4.5 Orientation Selection Compatibilities

The generic design which has been developed is easily applicable to the orientation selection problem Since the perpendicular projection in this domain has already been described (§3 4 3), all that remains is to describe the translation of this projection into compatibilities which will localize the confidence laterally and stably break down near the ends of curves (see §2 2 3)

The process of computing compatibilities when localization axes are considered is made simpler when such axes are aligned with parameter axes In the case of the cocircularity constraint the local orientation of the hypothesis corresponds to Θ_i and the local curvature corresponds to K_i A rotation by Θ_i around the origin further ensures that lateral position (i e perpendicular to Θ_i) aligns with the Y'_j axis Once this rotation is performed, the localization constraint can be seen as being directly reflected in the cross-section of compatibilities perpendicular to these three parameter axes.

Thus, of the six parameters which describe the relationship between two curved tangents, three (Θ, K_i, and Y_j) must be locally maximal for a label to receive positive support. The other three parameters can simply be smoothly attenuated with increasing perturbation in the projection. Given the output of the perpendicular projection $\mathbf{X}' = [\theta'_i, \kappa'_i, \theta'_j, \kappa'_j, x'_j, y'_j]$, then the displacements from exact conformity with the cocircularity constraint are described by

$$
\begin{aligned}
\Delta\theta_i &= (\theta'_i - \Theta_i)/\Delta\Theta_i \\
\Delta\kappa_i &= (\kappa'_i - K_i)/\Delta K_i \\
\Delta\theta_j &= (\theta'_j - \Theta_j)/\Delta\Theta_j \\
\Delta\kappa_j &= (\kappa'_j - K_j)/\Delta K_j \\
\Delta x_j &= (x'_j - X_j)/\Delta X_j \\
\Delta y_j &= (y'_j - Y_j)/\Delta Y_j
\end{aligned}
\tag{3.12}
$$

(N.B. The form of these equations depends on each of the discrete parameter values being equally spaced in the parameter space. Such is the case for a uniform grid and the label distributions chosen. For non-uniform grids and/or parameter spacings refer back to §3.4.3.)

The attenuation on the non-localized parameter axes is modelled as a product of perpendicular gaussians, or

$$
G^*(\mathbf{\Delta}) = \exp\left(-\frac{1}{2} \sum_{\delta \in \mathbf{\Delta}} \delta^2 \right). \tag{3.13}
$$

where $\mathbf{\Delta}$ is the set of normalized perturbations on the non-localized axes (see equations (3.12)). The localization axes, however, must be dealt with independently since the cross-section of compatibilities on each of these axes must conform to the criteria described in §3.4.4. Combining the attenuation due to variations in all of the other parameters with each

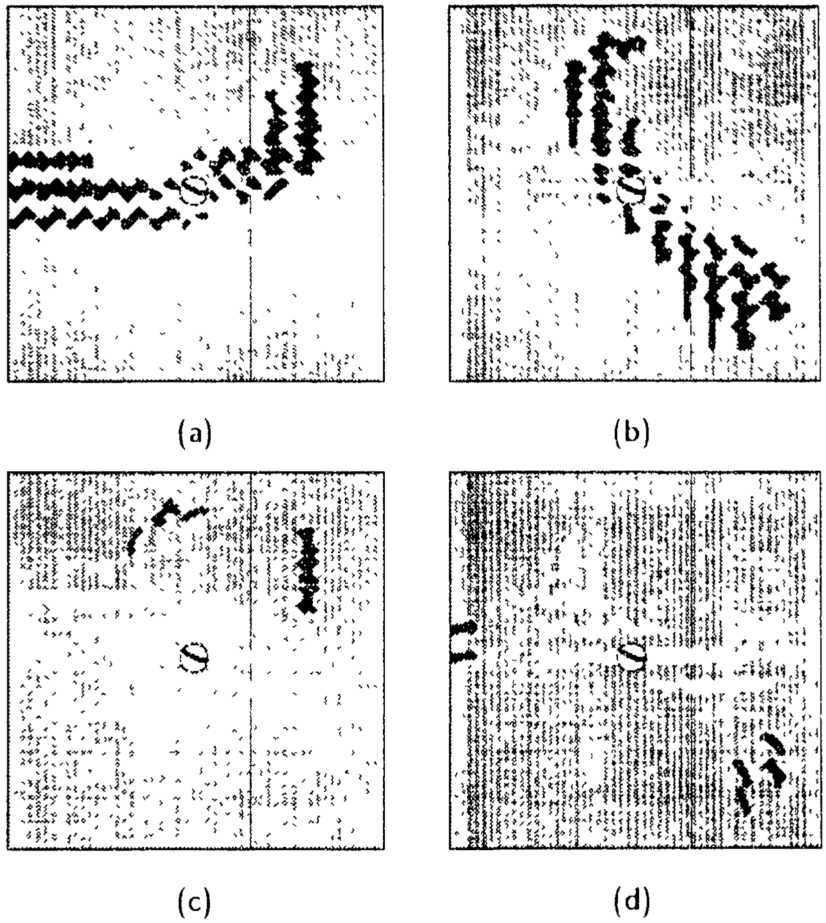
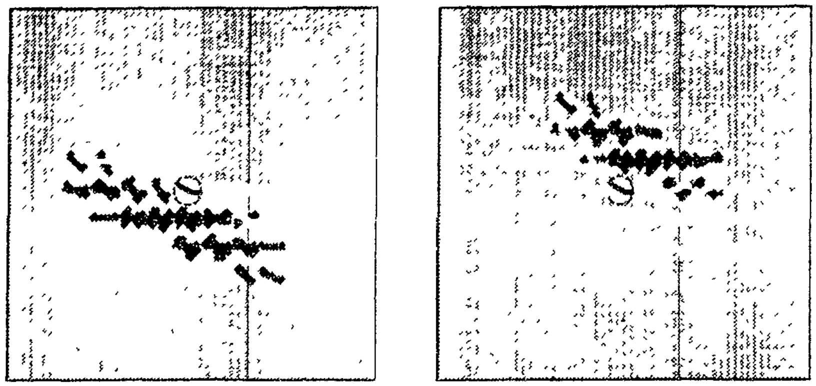


Figure 3.6 Local sets of support for a curve at 22.5° with a curvature of -0.1. The support sets shown are those which correspond to (a) $r\Theta^-$ (b) $r\Theta^+$ (c) rK^-, and (d) rK^+. Positive and negative compatibilities are white and black respectively (the degree of compatibility is not shown here)

of these localization conditions, arrives at three pairs of compatibilities, one pair associated with each localized variable

$$
\begin{aligned}
r\Theta^- \,(\Theta_i, K_i, X_j, Y_j, \Theta_j, K_j) &= dG_{\sigma'}(\Delta\theta_i - \delta)\; G^*(\Delta\kappa_i, \Delta\theta_j, \Delta\kappa_j, \Delta x_j, \Delta y_j) \\
r\Theta^+ \,(\Theta_i, K_i, X_j, Y_j, \Theta_j, K_j) &= - dG_{\sigma'}(\Delta\theta_i + \delta)\; G^*(\Delta\kappa_i, \Delta\theta_j, \Delta\kappa_j, \Delta x_j, \Delta y_j) \\
rK^- (\Theta_i, K_i, X_j, Y_j, \Theta_j, K_j) &= dG_{\sigma'}(\Delta\kappa_i - \delta)\; G^*(\Delta\theta_i, \Delta\theta_j, \Delta\kappa_j, \Delta x_j, \Delta y_j) \\
rK^+ (\Theta_i, K_i, X_j, Y_j, \Theta_j, K_j) &= - dG_{\sigma'}(\Delta\kappa_i + \delta)\; G^*(\Delta\theta_i, \Delta\theta_j, \Delta\kappa_j, \Delta x_j, \Delta y_j) \\
rY^- (\Theta_i, K_i, X_j, Y_j, \Theta_j, K_j) &= dG_{\sigma'}(\Delta y_j - \delta)\; G^*(\Delta\theta_i, \Delta\kappa_i, \Delta\theta_j, \Delta\kappa_j, \Delta x_j) \\
rY^+ (\Theta_i, K_i, X_j, Y_j, \Theta_j, K_j) &= dG_{\sigma'}(\Delta y_j + \delta)\; G^*(\Delta\theta_i, \Delta\kappa_i, \Delta\theta_j, \Delta\kappa_j, \Delta x_j)
\end{aligned}
\tag{3.14}
$$

The support associated with each of these localization sets is termed the localization



(e) (f)

Figure 3.6 Local sets of support (contd.) for a curve at 22.5° with a curvature of 0.1. Shown are (e) rY^-, and (f) rY^+

support on the X axis. To demonstrate, below we show the localization support on the Θ_i axis for the $\mathbf{i} = (X_i, Y_i, \Theta_i, K_i)$ hypothesis

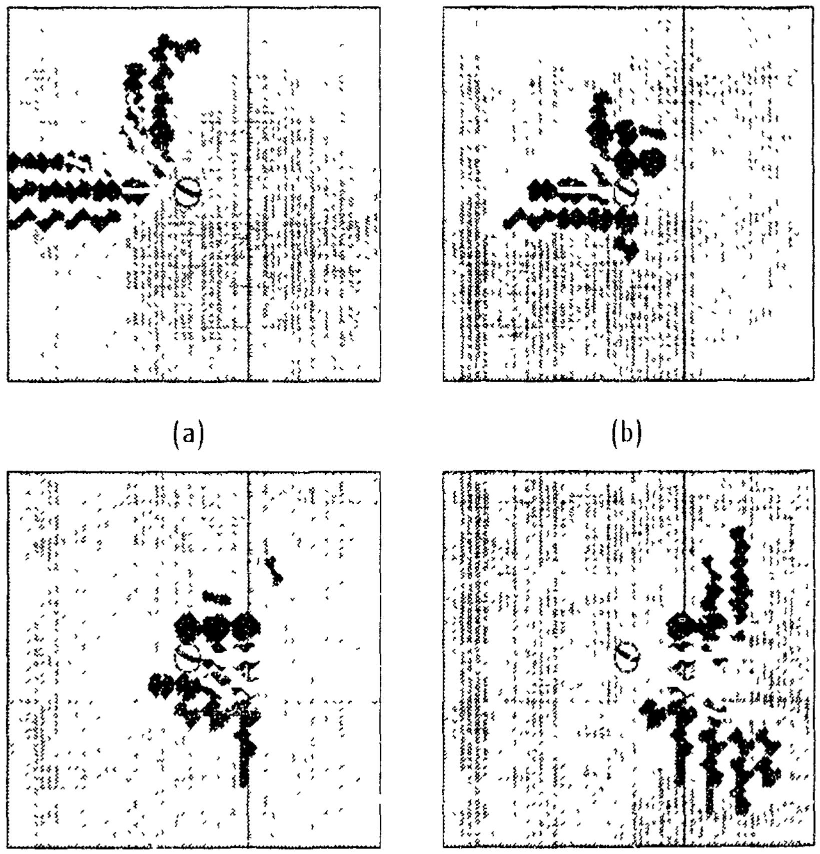
$$
\begin{aligned}
R\Theta^-(\mathbf{i}) &= \sum_{\mathbf{j}\in I} p(\mathbf{j})\, r\Theta^-(\Theta_i, K_i, X_j - X_i, Y_j - Y_i, \Theta_j, K_j) \\
R\Theta^+(\mathbf{i}) &= \sum_{\mathbf{j}\in I} p(\mathbf{j})\, r\Theta^+(\Theta_i, K_i, X_j - X_i, Y_j - Y_i, \Theta_j, K_j) \qquad (3.15) \\
R\Theta(\mathbf{i}) &= R\Theta^-(\mathbf{i}) \wedge R\Theta^+(\mathbf{i})
\end{aligned}
$$

The final support for a hypothesis pair is the $\wedge$ of the localization support for each of the three localization axes. The six support subsets for a given (Θ_i, K_i) pair are shown in Figure 3.6. Note that each support set can be seen as testing a particular condition (e.g. Figure 3.6(a) tests whether the curve is more likely to be oriented at 22.5° than at 45°.)

Note also that the negative compatibilities for the curvature options (Figures 3.6(c) and (d)) are somewhat inadequate. This is due largely to tradeoffs between the orientation and curvature spacing parameters. Because of the discretization of space, higher order derivatives are more and more coarsely measureable. For curvature, a second order measure, a discretization into five local curvature classes is near the upper limit for initial operators of this size.[*] Discretizing orientation more finely increases the number of inhibitory in-

[*] Seven curvature classes are also possible, but for a set of initial operators with optimal width of only two pixels, separating out nine curvature ranges reliably is already infeasible.

teractions in such a curvature component. It should be pointed out however, that this apparent inadequacy is only apparent—the system performs extremely stably and does not spread the curvature estimates wildly. This is due to the fact that the local curvature estimate determines the *shape* of the compatibility field even when curvature localization is not included explicitly in the decomposition.



(a) (b) (c) (d)

Figure 3.7 Compatibility subregions for a curve at 22.5° with a curvature of -0.1. Each subregion is shown as a ∠ combination of the six components of equation (3.14).

Achieving end-line stability is clearly just as important in the relaxation phase of processing as it was in the initial measurements. The relationship between the relaxation labelling support and a linear convolution again to reveals a solution. The compatibilities are decomposed into a set of length components tangential to the curve, and the support is calculated independently in each region. This is made possible by calculating the pro

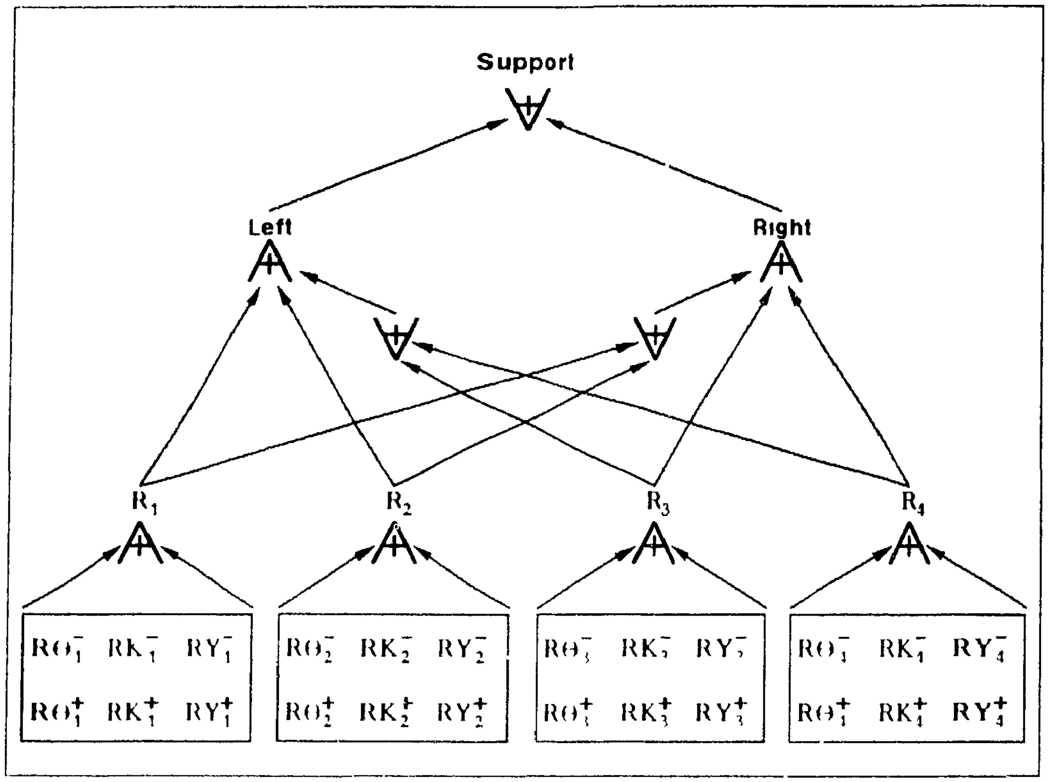


Figure 3.8 Support network used for calculating the support for a single hypothesis The lowest level consists of support for the localization subsets in each of four regions Each level above this is a semi-linear combination of the intermediate supports below it

portional contribution of each $rX^{\mp}$ to the given region (see equation (2 20)) and weighting the compatibility accordingly Thus given a tangential distance $d^{\dagger}$ between the two labels, the length components partition the compatibilities into regions of equal length Thus the contribution of a label to one of the length components of the support network is given by

$$r_i(\lambda, \lambda') = W_i(d)\ r(\lambda, \lambda') \tag{3 16}$$

As with the initial measurement operators the weights $W_i(d)$ are attenuated by a gaussian envelope The regions are then combined using the rule in equation (2.22) This decomposition is shown in Figure 3 7

† The distance d between (Θ_i, K_i) at (X_i, Y_i) and any hypothesis at (X_j, Y_j) is determined by projecting (X_j, Y_j) onto the circle with tangent (Θ_i, K_i) at (X_i, Y_i) The distance d is then the distance along the circle from (X_i, Y_i) to the projection This is distance measure is identical to the tangential distance in the initial measurement operators except that it is defined in a curved space

With the combination of the two non-linearities described above, the support calculation for an individual curved tangent hypothesis indexed as (x, y, Θ, K) can be summarized as

$$s = (R_1 \wedge R_2 \wedge (R_3 \vee R_4)) \vee (R_4 \wedge R_3 \wedge (R_2 \vee R_1)) \tag{3.17}$$

where each R_i is the aggregate response of the localization sets in region i.

$$R_i = R\Theta_i^- \wedge R\Theta_i^+ \wedge RK_i^- \wedge RK_i^+ \wedge RY_i^- \wedge RY_i^+. \tag{3.18}$$

A schematic of this support structure is shown in Figure 3.8

3.4.6 Compatibilities: Summary

We have thus developed a method for deriving the compatibilities of a relaxation labelling network which reason about the continuity and locality of curves in an image plane. The compatibilities were determined by projection of a description of the relationship between a pair of curved tangents onto a smooth manifold defined in terms of local cocircularity. The compatibilities are thus based on a measure of the size of the perturbation required to make the label pairs cocircular. In considering the requirements of the relaxation it was discovered that image curves are localized along certain of these axes of perturbation, specifically lateral position, local orientation and local curvature. A semi-linear decomposition of the compatibilities was proposed to ensure that the relaxation achieves such localization in its output. The desirability of end-line stability was reiterated and another linear/logical decomposition of the support network was proposed as a solution to this problem. The resulting support network for each hypothesis is a parallel, cascaded summation of the component supports.

In every case, the calculation of compatibilities has been based entirely on considerations of the local differential geometry of plane curves and the properties of relaxation labelling. Every decision made in the design of the support network was based on a principled analysis of alternatives. No arbitrary design factors were employed. It was necessary, however, to relax one of the conditions under which [Hummel & Zucker 83] proved convergence of relaxation labelling It has been partially demonstrated (without proof) that such convergence is retained with one of the non-linearities used It will be demonstrated experimentally that the overall convergence is stable

3.5 Conclusions

The goal of this chapter has been to develop a general methodology for refining initial measurements of local properties of image curves to reflect constraints which are assumed as to continuity and curve interactions A standard for recognition of an image curve was established which ensured that only locally twice-differentiable continuous curves would be selected as *consistent* It was further established that any reasonable algorithm for doing this must recognize and represent the probability that image curves will cross, bifurcate or otherwise interact, and a method was developed for dealing with such circumstances Finally, relaxation labelling was introduced as a massively parallel mechanism for solving the consistency constraint problem formulated

Once such constraints were in place, a general methodology was proposed for deriving relaxation labelling compatibilities from a set of algebraicly defined geometric constraints This scheme and the reasons for adopting it were illustrated using the restricted domain of local cocircularity constraints for image curves, but we assert that it has much wider application than that In fact, each of the design principles used to resolve the structure of the support network for this problem has a natural correlate in each of the other domains covered by the early visual system We will describe in some detail how such principles can be extracted from the geometry of these other problems in the section devoted to indicating future directions for research (§5 2)

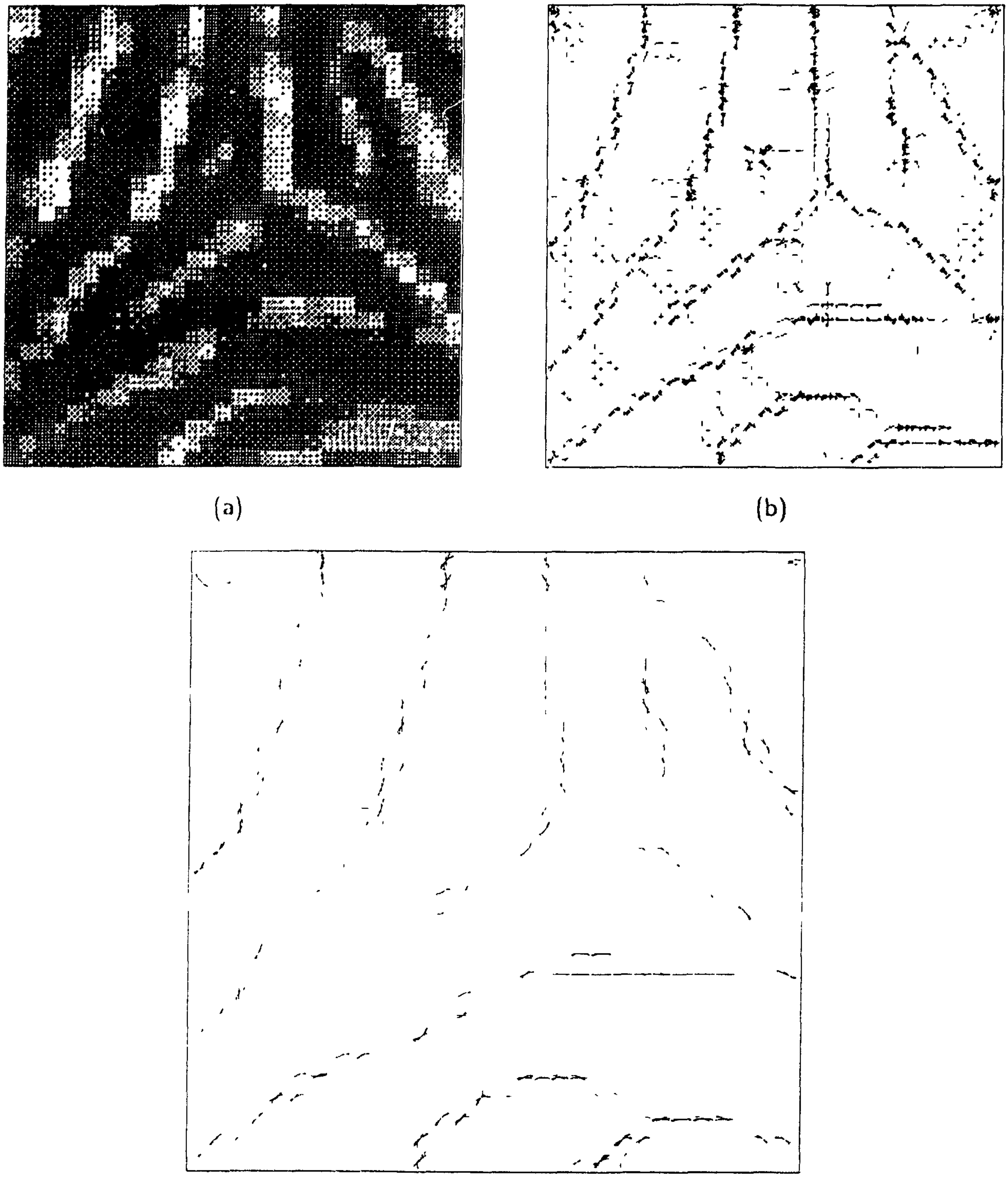
Chapter 4 Results

The results which are presented in this chapter were computed starting from the initial measurement operators described in §2.4. With these estimates of confidence in each of 40 orientation/curvature pairs at each image position, the relaxation then proceeded. The compatibilities were decomposed to provide localization for orientation, curvature, and lateral position. In addition, they were attenuated tangentially by a gaussian with $\sigma = 2$ pixels to a maximum arc distance of 5 pixels. The support network was further decomposed into 6 length components with a degree of separation $\rho = 4.0$. Only those interactions with positive compatibility 0.3 or with negative compatibility 0.07 were used.

It should be pointed out at this time that the three iterations used to demonstrate these results is not sufficient for strict convergence of the label confidences (in an unambiguous labelling all labels have a confidence of either 0 or 1). Instead, we select those labels which received positive support in the previous iteration. This is an excellent predictor of the final convergence of the system.

An important factor in evaluating these results is the evolving confidence of the labels. Unfortunately it is difficult to display the confidence, orientation and curvature of a label simultaneously in black and white. Instead, for display purposes only, the results shown below are thresholded at a confidence of 10%.

All experiments were conducted on a Symbolics 3650 Lisp Machine, running Genera 7.1 software. The system is programmed in Zetalisp, a dialect of Common Lisp.



(a) (b) (c)

Figure 4.1 FPRINT Small section of fingerprint The images are (a) input image (b) initial confidences (c) result after 3 iterations

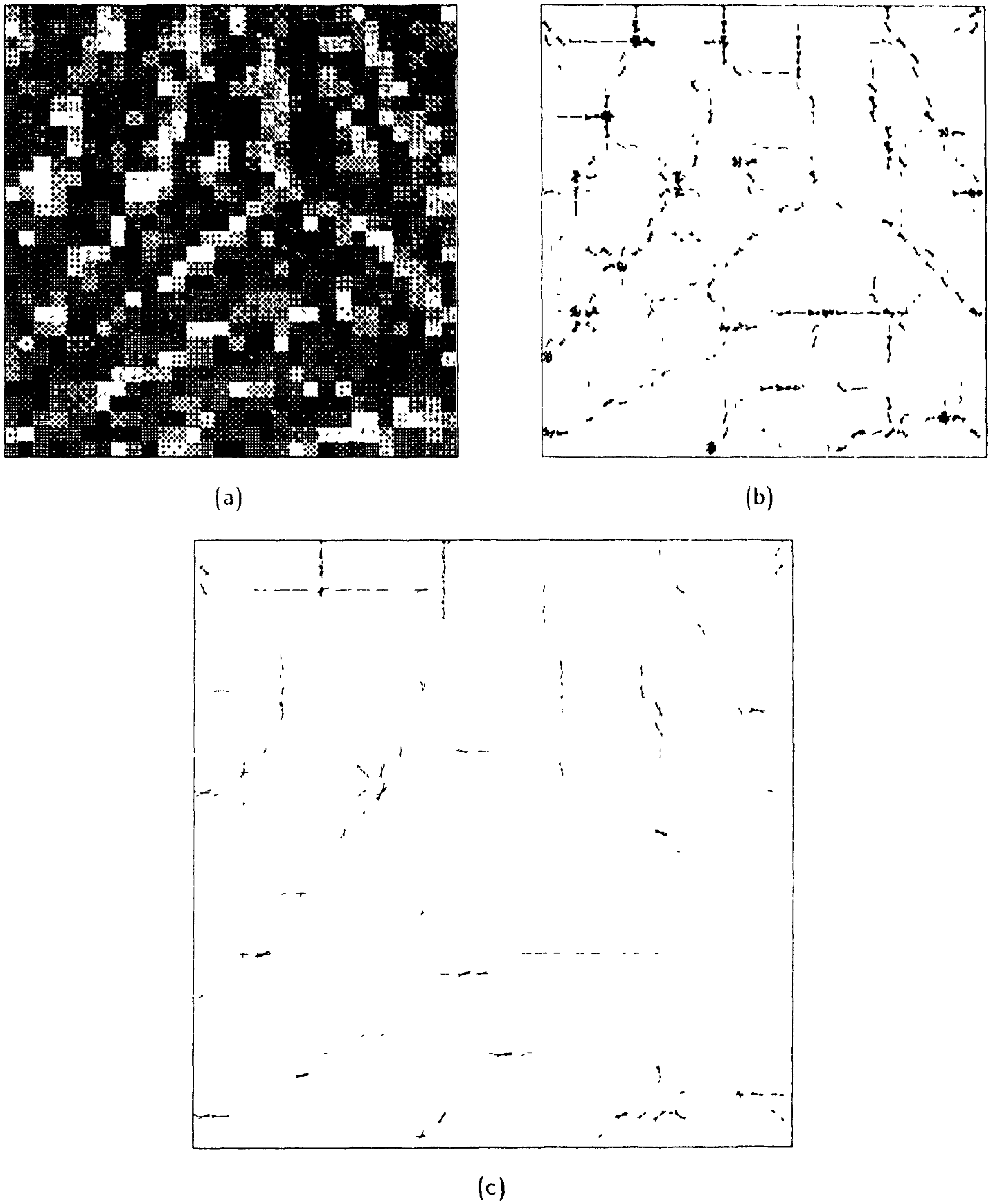


Figure 4.2 Noisy FPRINT This is just the FPRINT image degraded by additive gaussian noise with $\sigma = 10$ (on 64 grey levels) The images are (a) input image (b) initial confidences (c) result after 3 iterations

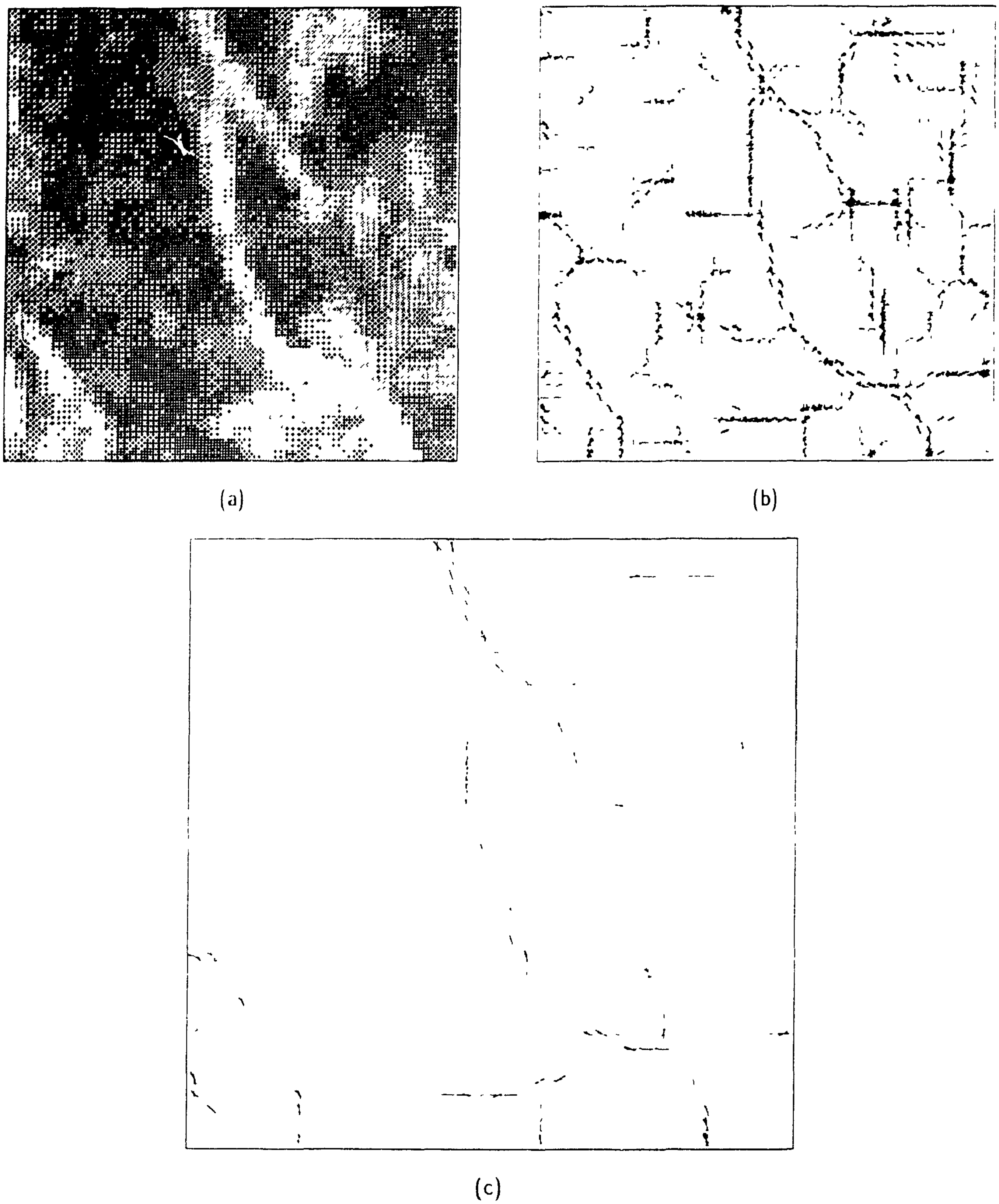


Figure 4.3 **ROADS** Satellite image of logging roads Executed with a 8 orientations and 5 curvatures the maximum radius of hypothesis interaction is 5 pixels The images are (a) input image, (b) initial confidences (c) result after 3 iterations

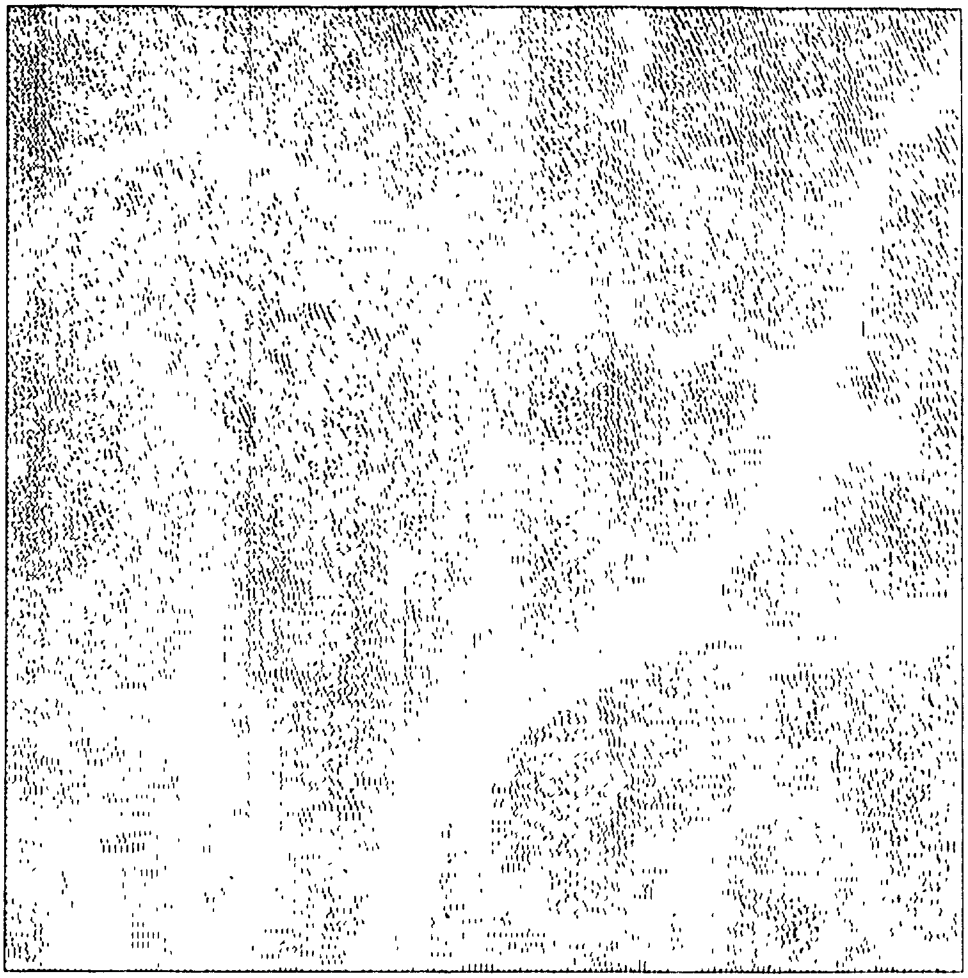


Figure 4.4 **ANGIO** Cerebral angiogram an X-ray of blood vessels in the brain

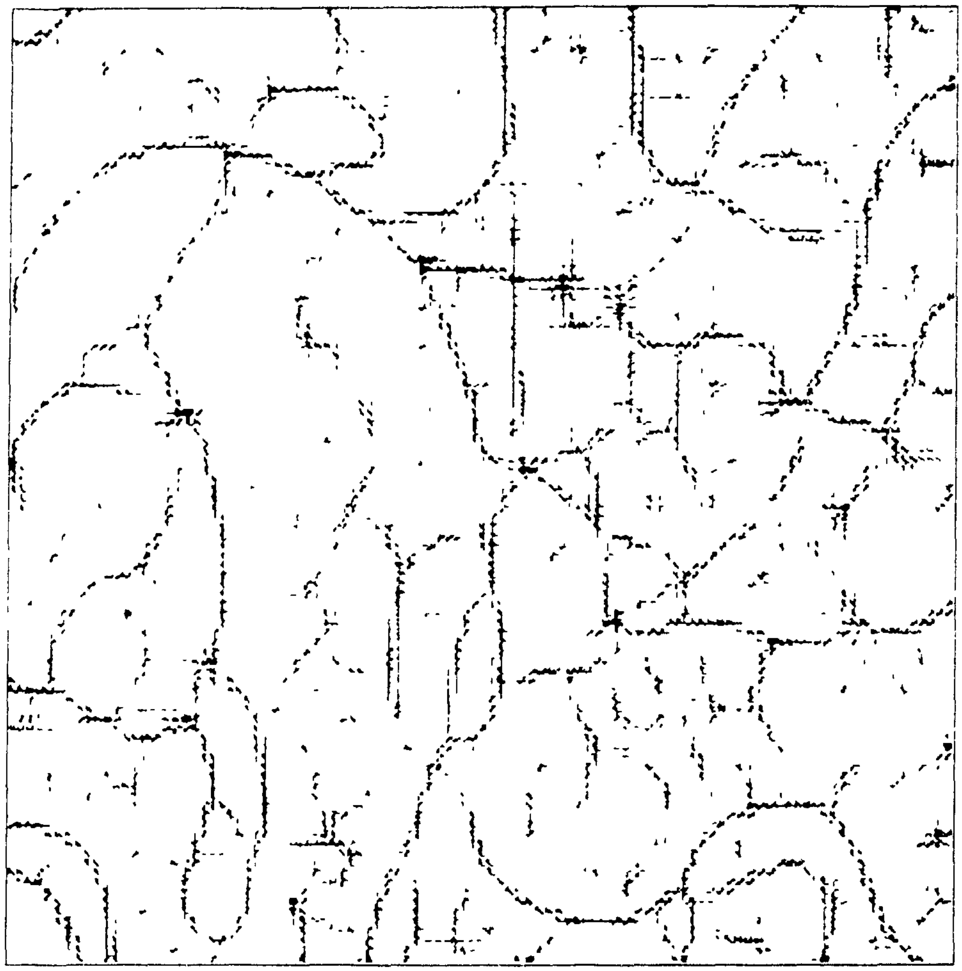


Figure 4.5 ANGIO: Initial Measurements

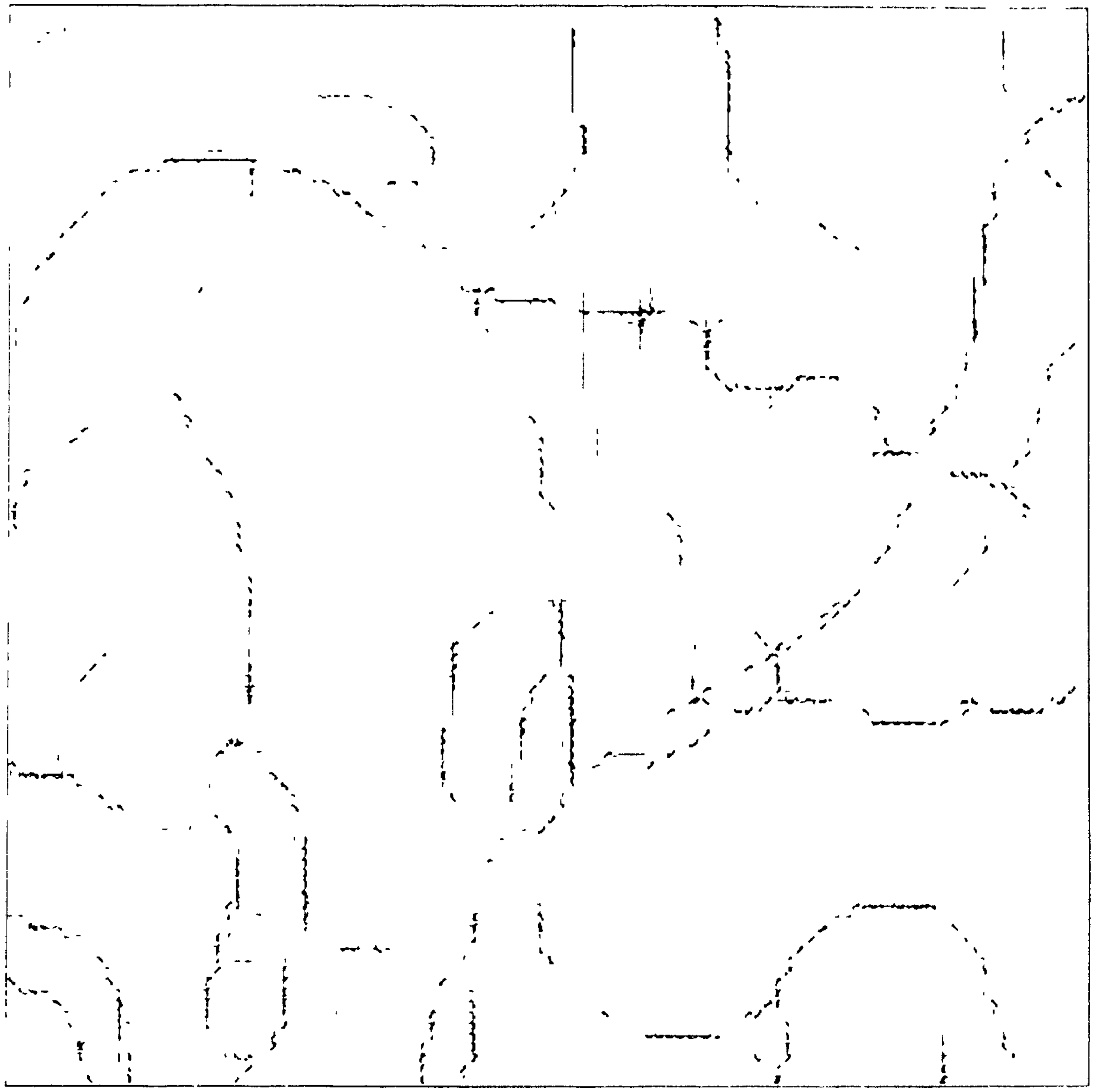


Figure 4.6 **ANGIO: Iteration 1** confidences thresholded at 10%

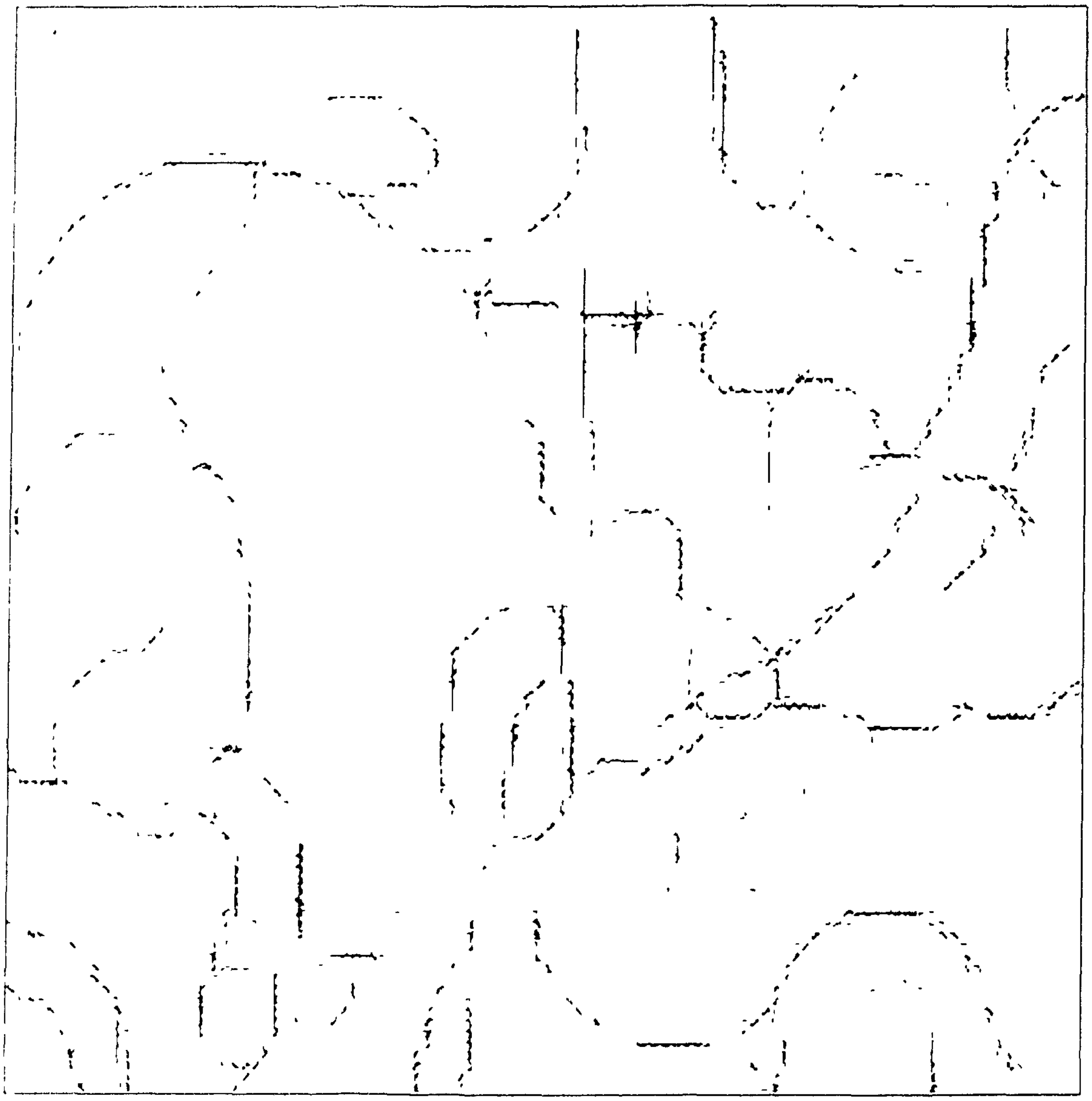


Figure 4.7 ANGIO: Iteration 2 confidences thresholded at 10%

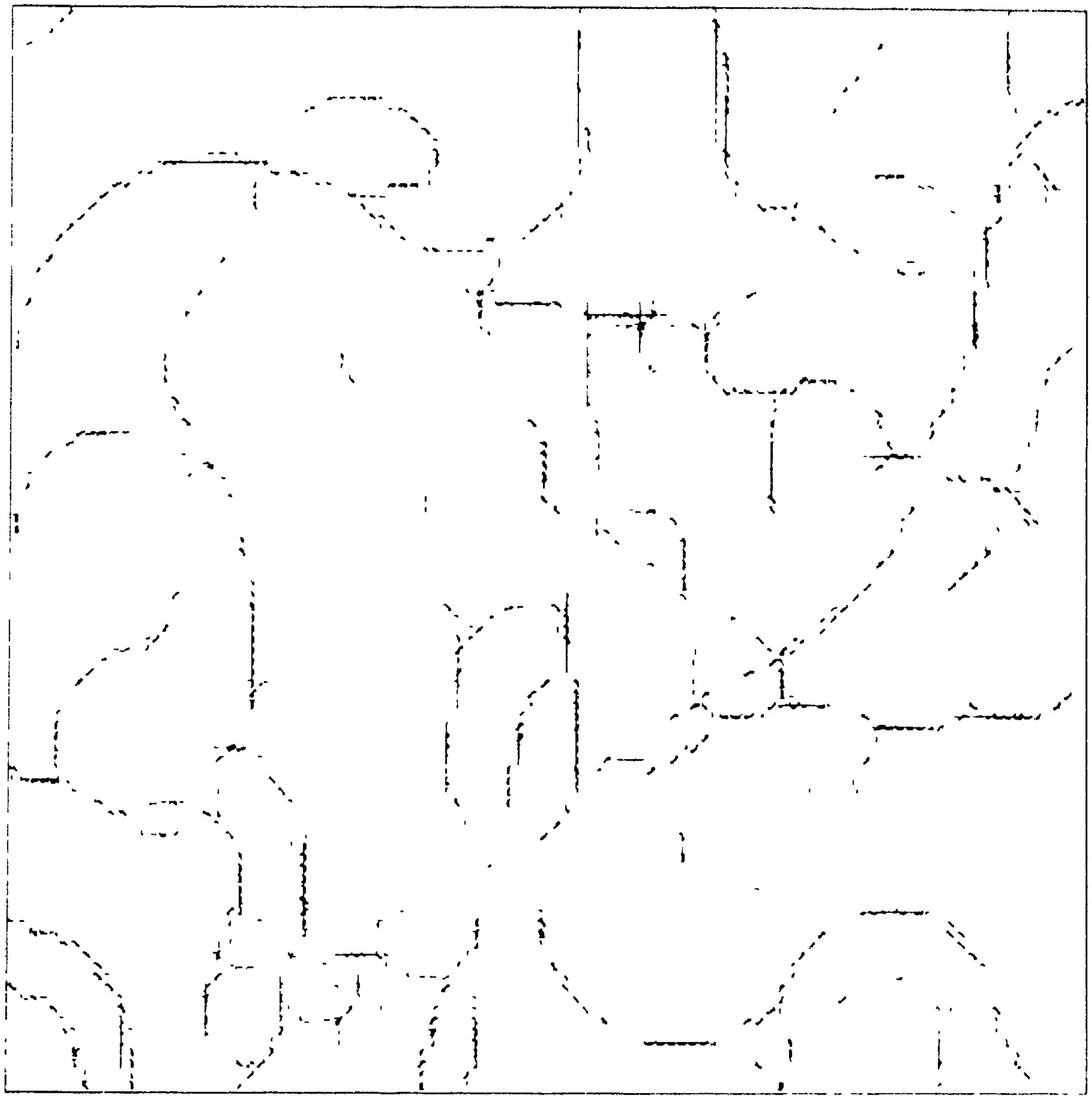


Figure 4.8 ANGIO: Final Result

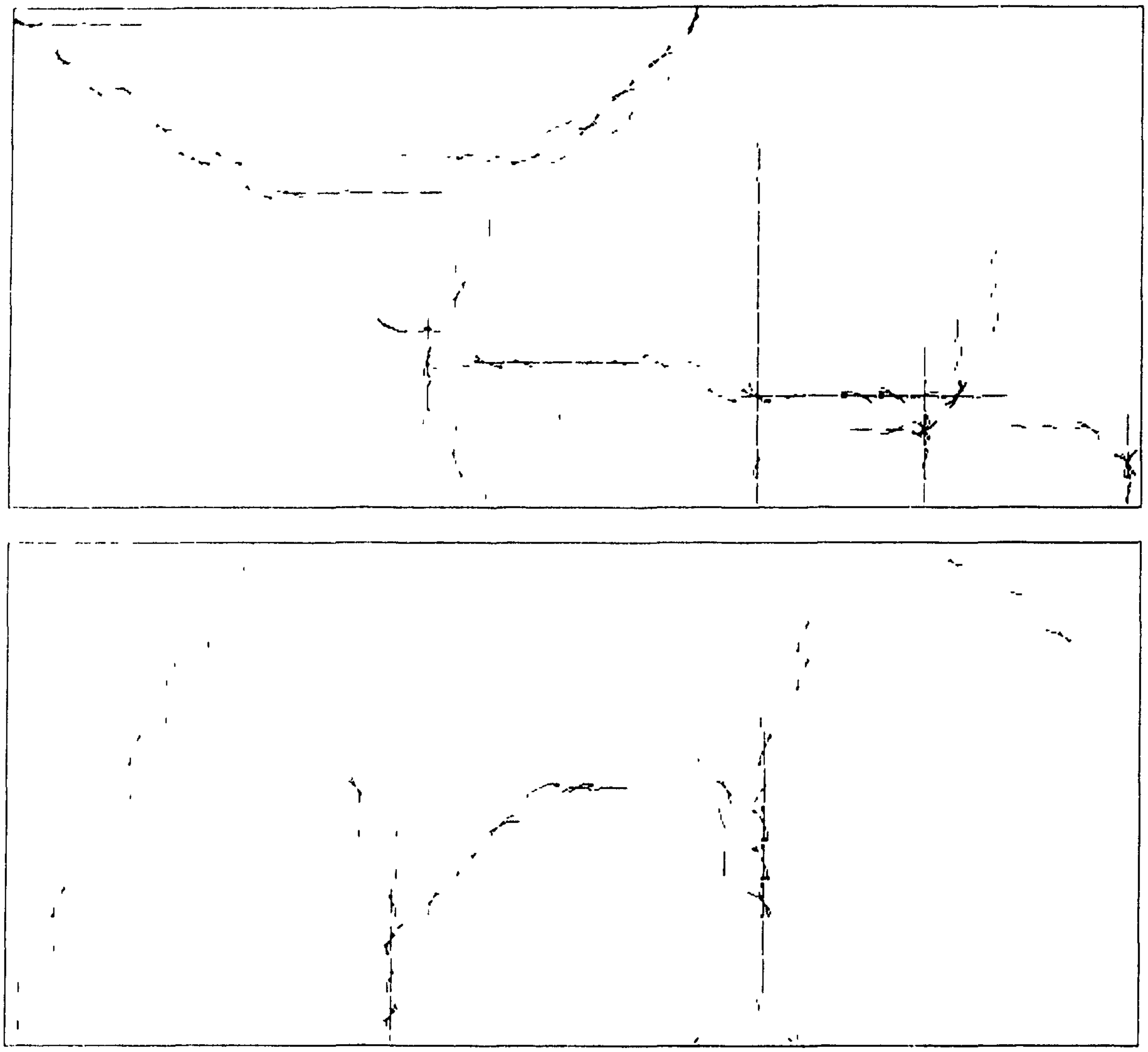


Figure 4.9 ANGIO: Final result details

Chapter 5 Conclusions

This work has established a general methodology for solving the initial measurement processes which have come to be called early vision It has generalized and formalized some of the earlier work on relaxation labelling for curve description in grey level images. and has thus exposed the possibility of treating a variety of problems in early visual processing as variations on a central theme. the discovery of consistent patterns in vector fields By keeping in mind the real time and computational constraints of a neurological system. the system developed also forms a rich and useful theory of visual processing in the visual cortex of cats. monkeys and humans

5.1 Future Directions: Early Vision

Edge and line detection. texture perception. image motion. stereo integration and colour vision are all classed as problems in early vision Each involves the discovery/inference/detection of features or structures in the retinal image which are assumed to correspond to properties of objects in the real world As a first stage in the eventual description of physical objects. early visual processing is simply a description of the structure of the retinal image

Consider then the ways in which such early processes are similar Because the coordinate systems are retinal. and not functions of real-world location there is a restricted well-defined area over which they must operate The mapping from retinal location to hardware (neurological or computational) can therefore be much simpler than when an unbounded space must be represented. Since the range of variation for local structure in images is limited – both by imaging physics and the structure of the world we live in— the kind of adaptability to wildly unusual combinations of local features which is clearly exhibited at the highest levels of visual processing are not necessary at these very early stages. The universality and perceptual constancy of many visual illusions demonstrates this dichotomy well. The early visual systems can thus specialize greatly for the kinds of universal structural quantities which are shared by all visual stimuli, no matter how semantically novel or unusual they might be.

At a more abstract level each of the standard problems in early vision can be formally described as the inference of a locally consistent vector field over the retinal image. For edge and line description, as has been shown, the vectors are the local tangents of the image curves. For stereo integration, they would be the stereo disparity vectors. For colour constancy, they would be vectors of colour description (e.g. intensity, hue, and saturation). For motion, they would be the motion vectors for optical flow. This principle of similarity between all of these early visual processes has been explored elsewhere (e.g. [Zucker 85], [Zucker & Iverson 87]), some of the computational issues involved are the ones addressed in this thesis. In each of these cases, the same issues arise. How to take crude measurements of such quantities and abstract from these a description of the vector fields so constituted? How to do this quickly enough that the computer or animal can react appropriately? How to ensure that the sorts of abstract information necessary for higher processes are made available?

It is the goal of this project to answer each of these questions in turn. A theoretical framework has been developed which exploits the similarities between these problems in such a way as to allow for efficient solutions to all of them. These solutions are not only based on an understanding of what has been discovered by visual neurologists and psychophysicists, but represent reasonable *working* models of visual processing in the brain. From a strictly engineering point-of-view, these models will work better and faster than other computational approaches

The present work demonstrates only a small facet of this larger goal In concentrating on the simplest of early vision problems, what has been called 'edge detection' this thesis explicates in a simple, straightforward manner the theoretical underpinnings of this body of research and their computer implementation The general applicability of the approach is explained at each stage in this development, with references to the other subjects of early vision Finally, this work will demonstrate how well we achieve the goal of a robust, accurate, real-time system for describing image curves, and the promise for like succes in the other domains of early visual processing

Texture flow [Zucker & Iverson 87], the problem of inferring orientation patterns in static texture images, is most closely related to curve description and here the implications of the present work are clearest. Optical flow, the problem of describing the lateral motion of patterns in the retinal image, is just a step beyond this

5.1.1 Texture Flow

The nature of texture flow is most clearly described with reference to random dot Moiré patterns—also known as Glass patterns (after [Glass 69] and [Glass 73]) (see Figure 5 1) In these patterns (and in hair patterns or any texture with a highly oriented structure) there are dense or sparse orientation cues which give an aggregate percept of an oriented, static 'flow' which exists everywhere in the image

The outline of a solution to this problem in the current framework has been previously described in [Zucker & Iverson 87] and [Iverson & Zucker 87] It is repeated here to clarify the relationship between the description of image curves and these flow patterns The most important differences between these two problems can be explained in terms of the topology of their solutions A consistent curve is a one-dimensional object while a consistent texture flow field is two-dimensional [Zucker 83] This implies two things

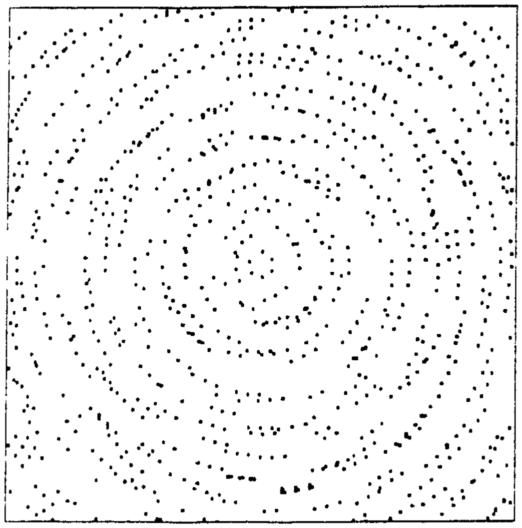


Figure 5.1 Random dot Moiré pattern demonstrates the principle of *texture flow* Even though the actual orientation cues are relatively sparse and ambiguous the overall perception is of an unambiguous circular flow field

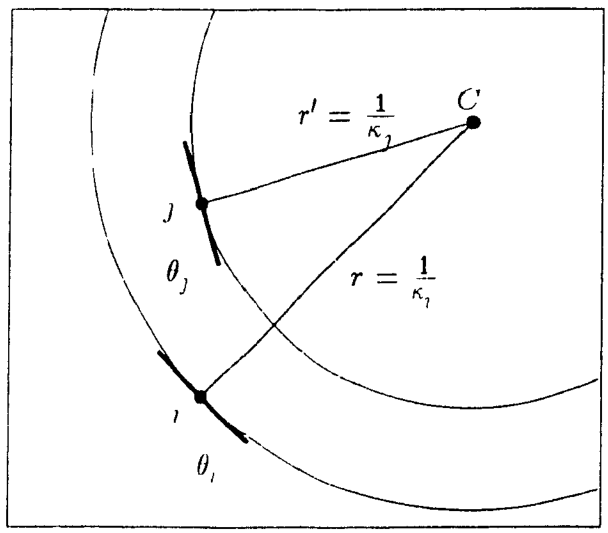


Figure 5.2 Geometric concentricity constraint relates the orientation θ_i and curvature κ_i at point i to θ_j and κ_j, the same measures at point j Cocircularity of the two curved tangents depends on their sharing the same centre of curvature C

- the manifold of geometrically consistent label pairs is four-dimensional instead of three Consistency, now defined in terms of local *concentricity* (see Figure 5 2, is thus expressed by only two coupling equations
- the inside/outside distinction so easily resolved for a one-dimensional object (see §2 2 3) is more problematic In this case, by assuming a locally straight border between distinct texture flow regions, it can be shown that rather than demonstrating consistency for the left and right half-fields, the relaxation must do so for each of four quadrants around the centre point

The coupling equations are as simple as before

$$\theta_j = \tan^{-1}\left(\frac{\sin\theta_i + \kappa_i y_i}{\cos\theta_i - \kappa_i y_j}\right)$$

$$\kappa_j = 1 \Big/ \sqrt{(\sin\theta_i - \kappa_i x_i)^2 + (\cos\theta_i - \kappa_i y_i)^2}$$

The decomposition of the support network must now proceed on the principle of localizing around only θ_i and κ_i, and of a region decomposition based on the quadrants around the central point. Each of these decompositions can proceed using the principles established in §3.4.4 and §3.4.5.

The second non-linearity in the support for the orientation selection problem was introduced for end-line stability. For texture flow, instead the problem is one of stability near discontinuities in the flow field. It might seem at first that these are truly separate problems, but in the context of relaxation labelling they are simply topological variants of the same problem—to ensure that support is only gathered from the *same* consistent region as the label being updated. For the one-dimensional orientation selection problem this was resolved by partitioning the support into left and right half fields. For two-dimensional texture flow, one must partition the support into quadrants around the centre point. The principle extends just as naturally to higher dimensions: a label's support is positive only if it is surrounded by regions which confirms its hypothesis. For n dimensional consistent structures (independent of the embedding space) this surrounding must contain at least 2^n regions (assuming that the border between consistent regions is locally planar).

5.1.2 Optical Flow

The importance of optical flow for the interpretation of visual motion has been well-established (e.g. [Koenderink & van Doorn 75], [Horn & Schunk 81]) and a great deal of effort has been expended developing methods for calculating it (see [Barron 84] for a recent review). The one overriding failure with these methods is that they do not recognize the important geometric structure of a consistent set of optical flow measurements.

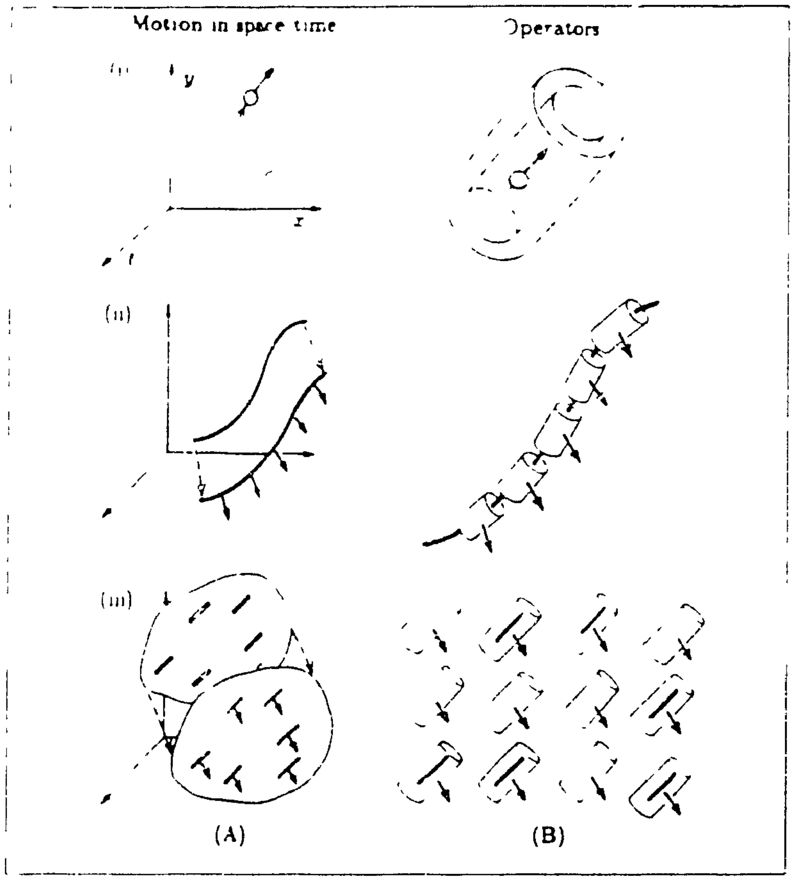


Figure 5.3 Dimensionality of optical flow is shown for (i) a moving point, (ii) a moving curve and (iii) a moving region The structures traced out by these motions in space time are the objects which a system for determining optical flow should discover

In [Zucker & Iverson 87] we described the three topologically distinct classes of optical flow Consider optical flow as defining smooth manifolds in the $X \times Y \times T$ retinotopic space-time A moving point then traces a curve in this space, a moving curve traces a surface, and a moving region traces a volume (see Figure 5 3) Assuming piecewise smooth motions, then the representations of these motions in space-time will also be piecewise smooth, so a stable relaxation labelling formulation in this space is every bit as feasible as for orientation selection and texture flow The initial measurements would be garnered from a set of directional space-time operators analogous to the ones developed in Chapter 2 Relaxation labelling would then proceed along three separate paths, each attempting to discover consistent response patterns of different dimensionality. All of the techniques developed above, which were specifically formulated to be generalizable to arbitrary dimensions are applicable to this kind of a system

5.2 Biological Observations

Throughout the development of this work care has been taken to ensure that each description, each formulation and each solution was applicable to not only the standard computational framework but to a neurophysiological substrate as well It can certainly be argued that the purely spatial processes described herein are unrealistically simplistic models of visual cortical processing, but they do represent a first (temporally invariant) candidate model for some of the processing which may be taking place

In fact, many aspects of this system resemble neurological processes more closely than traditional computational processes Appendix B demonstrates how the operators developed in Chapter 2 can be represented in a summing dendritic tree with shunting inhibition [Rall 64], thus suggesting a functional role for some of the dendro dendritic interactions observed The representation chosen is far richer in its initial stages than most computational models and a comparison with the structure of orientation hypercolumns in visual cortex reveals many similarities [Hubel & Wiesel 65]

The relaxation stage of the model has similar neurological correlates for the linear/logical summation of the support calculation is again potentially implemented in a single dendritic tree with shunting inhibition The entire three/four iteration relaxation could thus be implemented in only three layers of cortical circuitry (assuming an entirely feed-forward network) or even in one or two with some feedback A more telling point however, is combinatorial for with the current system fan ins to a single node on the order of 1000–1500 are normal This places the system much more closely to neurological parameters than those of traditional computational models

In addition to these similarities, the model is rich enough to actually generate a few simple predictions about the biology Dobbins has already done much of the work to document the curvature selectivity of cortical neurons [Dobbins *et al.* 87 88], but the other non-linearities which have been added to the operator have not been verified neurophysiologically Within the relaxation structure is embedded a description of local interactions between nearby orientation (and curvature) specific neurons Much of the debate over whether cortical simple-cell interactions are primarily excitatory or inhibitory seem superfluous now that the need for a balance of both has been revealed In fact, the cocircularity constraint and its translation to relaxation compatibilities can predict the kind and degree of interaction which should be observed between neighbouring tuned cortical simple cells

5.3 Summation

It has become clear in executing this research that a deep understanding of the constraints on the visual system and a principled application of such constraints to the solution of the geometric problems faced by this system can lead to a straightforward, robust solution It is hard to imagine how the insights gleaned from this work—which in many ways depended fundamentally on the need to develop a *work'ng* model—could have come from an approcah with less rigid modelling requirements A neurophysiologist, without the demand to produce a system which works is free tc propose any plausible solution to the actual computational problems with vision with only the requirement that it explains something of and generates useful neurological predictions This is by no means without use, but it is not often that it reveals deep truths about the inherent nature of visual processing In much the same way, a neural nets' researcher who simply blindly wires up a network and hopes that it will, when appropriately trained, reveal to him these deep inner truths may just as easily fail to understand why his model did or did not work The problem of understanding vison 's so complicated and reflects on so many different fields that only through a constructive synthesis of mathematics, computational theory, neurophysiology and psychology can one hope to provide a solution which has relevance to all fields It is hoped that this current research will make just such a constructive contribution

Chapter 6 References

[Barron 84] J Barron, "A survey of approaches for determining optic flow," Computer Science Technical Report **TR-85**, Dept of Computer Science, University of Toronto (1984)

[Blake & Zisserman 87] A Blake & A Zisserman, *Visual Reconstruction* MIT Press, Cambridge Mass (1987)

[Canny 86] J Canny, "A computational approach to edge detection," *IEEE Trans on Pattern Analysis and Machine Intelligence* **8**, pp 679–698 (1986)

[Davis, Rosenfeld & Agrawala 73] L Davis, A Rosenfeld & A K Agrawala, "On models for line detection," Computer Science Technical Report **TR-258**, Computer Science Center, University of Maryland (1973)

[Dobbins, Zucker & Cynader 87] A Dobbins, S W Zucker & M S Cynader, "Endstopping in the visual cortex as a substrate for calculating curvature," *Nature* **329**, pp 438–441 (1987)

[Dobbins Zucker & Cynader 88a] A Dobbins, S W Zucker & M S Cynader "Endstopping and curvature" McGill Research Centre for Intelligent Machines Technical Report **TR-CIM-88-3**, Dept of E E, McGill University, Montréal (1988)

[Dobbins, Zucker & Cynader 88b] A Dobbins, S W Zucker & M S Cynader "Designing endstopped neurons for curvature selectivity," McGill Research Centre for Intelligent Machines Technical Report **TR-CIM-88-4**, Dept of E E, McGill University Montréal (1988)

[Fatt & Katz 53] P Fatt and B Katz, "The effect of inhibitory nerve impulses on a crustacean muscle fibre" *J Physiology (London)* **121**, pp 374–389 (1953)

[Hammond & MacKay 83] P Hammond & D M MacKay, "Influence of luminance gradient reversal on simple cells in feline striate cortex," *J Physiology (London)* **337**, pp 69–87 (1983)

[Hammond & MacKay 85] P Hammond & D M. MacKay, "Influence of luminance gradient reversal on complex cells in feline striate cortex," *J Physiology (London)* **359**, pp 315–329 (1985)

[Haralick 82] R M Haralick, "Digital step edges from zero-crossing of second directional derivatives," *IEEE Trans on Pattern Analysis and Machine Intelligence* **6**, pp 58–68 (1982)

[Hoffman & Richards 85] D D Hoffman & W A Richards, "Parts of recognition," *Cognition* **18**, pp 65–96 (1985)

[Horn & Schunk 81] B K P Horn & B C Schunk, "Determining optical flow," *Artificial Intelligence* **17** pp 185–203 (1981)

[Glass 73] L Glass, "Perception of random dot interference patterns," *Nature* **246**, pp 360–362 (1973)

[Hubel & Wiesel 62a] D H Hubel & T N Wiesel, "Receptive fields binocular interaction, and functional architecture in the cat's visual cortex," *J Physiology (London)* **160**, pp 106–156 (1962)

[Hubel & Wiesel 62b] D H Hubel & T N Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J Physiology (London)* **195**, pp 215–243 (1962)

[Hubel & Wiesel 65] D H Hubel & T N Wiesel "Receptive fields and functional architecture in two non-striate visual areas (18 and 19) in the cat," *J Neurophysiology* **28**, pp 229–289 (1965)

[Hueckel 71] M H Heuckel, "An operator which locates edges in digital pictures," *J Association for Computing Machinery* **18**, pp 113–125 (1971)

[Hueckel 73] M H Hueckel A local visual operator which recognizes edges and lines," *J Association for Computing Machinery* **20**, pp 634–647 (1973)

[Hummel & Zucker 83] R A Hummel & S W Zucker, 'On the foundations of relaxation labeling processes " *IEEE Trans on Pattern Analysis and Machine Intelligence* **6**, pp 267–287 (1983)

[Iverson & Zucker 87] L A Iverson & S W Zucker 'Orientation selection to optical flow a computational perspective," *Proceedings IEEE Workshop on Computer Vision '87* Miami Beach, Florida, pp 184–189 (1987)

[Koenderink & van Doorn 75] J J Koenderink & A J van Doorn, "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer," *Optica Acta* **22** p 773 (1975)

[Koenderink & van Doorn 82] J J Koenderink & A J van Doorn, *Perception* **11**, pp. 129–137 (1982)

[Leclerc 84] Y Leclerc, "Capturing the local structure of image discontinuities," *Proceedings IEEE Computer Vision and Pattern Recognition '85* San Francisco, California, pp. 34–38 (1985)

[Leyton 88] M Leyton, "A process grammar for shape,' *Artificial Intelligence* **34**, pp 213–247 (1988)

[Marr 82] D Marr *Vision*, W H Freeman, San Francisco (1982)

[Marr & Hildreth 80] D Marr and E Hildreth, "Theory of edege detection," *Proc. Royal Society* **B207**, pp 187–217 (1980)

[Mohammed, Hummel & Zucker 83] J Mohammed, R A Hummel & S W Zucker, "A feasible direction operator for relaxation methods," *IEEE Trans on Pattern Analysis and Machine Intelligence* **6** pp 330-332 (1983)

[Orban 84] G Orban *Neuronal Operations in the Visual Cortex*, Springer Verlag (1984)

[Rall 64] W Rall "Theoretical significance of dendritic trees for neuronal input output relations" *Neuronal Theory and Modelling*, R Reiss ed, Stanford Univ Press Stanford pp 73-97 (1964)

[Rosenfeld 84] A Rosenfeld, "Image analysis problems, progress and prospects," *Pattern Recognition* **17** pp 3-12 (1984)

[Rumelhart & McLelland 86] D Rumelhart & McLelland *Parallel Distributed Processing*, MIT Press Cambridge Mass (1986)

[Parent & Zucker 85] P Parent & S W Zucker "Trace inference curvature consistency, and curve detection," McGill Research Centre for Intelligent Machines Technical Report **CIM-86-3**, Dept of E E McGill University Montreal (1985)

[Prewitt 70] J M S Prewitt "Object enhancement and extraction" *Picture Processing and Psychopictorics* B S Lipkin & A Rosenfeld eds, Academic Press New York, pp 75-149 (1970)

[Torre & Poggio 84] V Torre & T Poggio, "On edge detection" MIT AI Memo **AIM 768**, MIT AI Lab, Cambridge Mass (1984)

[Watt & Morgan 83] R J Watt & M J Morgan, "Mechanisms responsible for the assessment of visual location theory and evidence," *Vision Research* **23**, pp 97-109 (1983)

[Whitaker & Walker 88] D Whitaker & H Walker "Centroid evaluation in the vernier alignment of random dot clusters," *Vision Research* **28**, pp 777-784 (1988)

[Zucker 83] S W Zucker, "The fox and the forest a Type I, Type II constraint for early optical flow," CVaRL Technical Report **83-11**, Dept of E E McGill University, Montréal, presented at the *ACM Workshop on Motion Reprsentation and Control* Toronto (1983)

[Zucker 85] S W Zucker, "Early orientation selection tangent fields and the dimensionality of their support" *Computer Vision, Graphics and Image Processing* **32**, pp 74 103 (1985)

[Zucker, Dobbins & Iverson 89] S W Zucker, A Dobbins & L A Iverson, "Two stages of curve detection suggest two styles of visual computation" *Neural Computation* **1** (in press) (1989)

[Zucker, Hummel & Rosenfeld 77] S W Zucker, R A Hummel & A Rosenfeld "An application of relaxation labelling to line and curve enhancement" *IEEE Trans on Computers* **C-26**, pp 394-403 (1977)

[Zucker, Leclerc & Mohammed 81] S W Zucker Y G Leclerc & J L Mohammed "Continuous relaxation and maxima selection conditions for equivalence" *IEEE Trans on Pattern Analysis and Machine Intelligence* **3** (1981)

[Zucker & Iverson 87] S W Zucker & L A Iverson, "From orientation selection to optical flow, *Computer Vision, Graphics and Image Processing* **37** (1987)

Appendix A. Linear/Logical Operators

As was described in §2 2 1 a set of semi linear logical operators have been developed The logical structure of these operators is based on the principle that the sign of a signal (positive or negative) can be associated with a truth value (true and false, respectively) These operators then exhibit the property that if some logical condition on the inputs holds the output will appear as a linear combination of the input values

The examples already outlined are the operators analogous to logical *and* and *or* Since these are combinations of linear and logical principles a notation which reflects this has been adopted

- linear/logical *and* of x and y is $x \wedge y$.
- linear/logical *or* of x and y is $x \quad y$

The computation tables for these operations are shown below

Operation Tables for Linear/Logical Operators

$\wedge$		x 0	$x \cdot$ 0
y	0	$x+y$	x
y	0	y	$x+y$

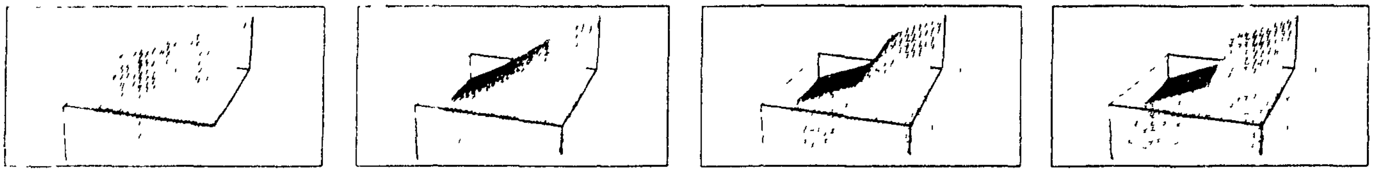
$\vee$		x 0	x 0
y	0	$x-y$	y
y	0	x	$x+y$

It is a simple matter to verify that the sign of these operations follows the associated logical rule The symmetry and associativity of these operators are also easy to demonstrate

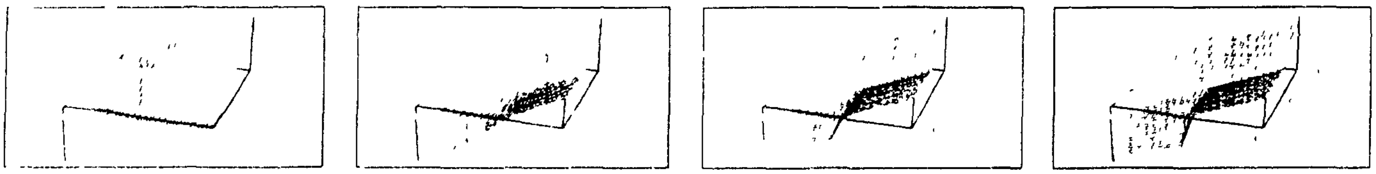
A more useful formulation of these operators is expressed in terms of the step function $\sigma(x)$

$$\sigma(x) = \begin{cases} 1, & \text{if } x \quad 0, \\ 0, & \text{otherwise} \end{cases} \tag{A 1}$$

This function is a choice operator pivoting around zero, and as such it can be used to directly define the two linear/logical operators above Consider that the operators can be



(a 1) (a 2) (a 3) (a 4)



(b 1) (b 2) (b 3) (b 4)

Figure A.1 Linear/logical functions of x and y varying p. (a 1) (a 2), (a 3) and (a 4) show $x \wedge y$ varying through 0 2 4 and ∞ (b 1) (b 2) (b 3), and (b 4) show $x \vee y$ varying through 0 2 4 and ∞

defined as.

$$x \wedge y = (x\,\textbf{unless}\,x > 0 \wedge y \le 0) + (y\,\textbf{unless}\,y > 0 \wedge x \le 0) \tag{A.2.1}$$

$$x \vee y = (x\,\textbf{unless}\,x \le 0 \wedge y > 0) + (y\,\textbf{unless}\,y \le 0 \wedge x > 0) \tag{A.2.2}$$

Expressing exactly the same logic using the identity

$$a\,\textbf{unless}\,b > 0 = a\,(1 - \sigma(b)) \tag{A.3}$$

gives

$$x \wedge y = x\,(1 - \sigma(x)\,\sigma(-y)) + y\,(1 - \sigma(y)\,\sigma(-x)) \tag{A.4 1}$$

$$x \vee y = x\,(1 - \sigma(y)\,\sigma(-x)) + y\,(1 - \sigma(x)\,\sigma(-y)) \tag{A.4 2}$$

These are derived based on the principle of selecting those cases in which either x or y contributes linearly to the output. The $1 - \sigma(a)\,\sigma(b)$ forms select all cases *except* when a and b are both positive A check back with the operation table confirms the correctness of these formulations

The key in arriving at an analytical form for these operators comes by taking successive smooth approximations to $\sigma(x)$. A simple sigmoid approximating function is the logistic function

$$\sigma_p(x) = \frac{1}{1+e^{-px}} \tag{A 5}$$

This function is infinitely differentiable and

$$\sigma(x) = \lim_{p \to \infty} \sigma_p(x). \tag{A 6}$$

Furthermore this approximation makes absolutely clear the relationship between the linear sum and these operators for $\sigma_0(x) = 1/2$ In this degenerate case, both of the operators simplify to addition

$$\begin{aligned} x \quad y &= x(1 - \sigma_0(x)\,\sigma_0(-y)) + y(1 - \sigma_0(y)\,\sigma_0(-x)) && \text{(A 7 1)} \\ &= x(1 - (1/2)\cdot(1/2)) + y(1 - (1/2)\cdot(1/2)) \\ &= 3/4\,(x+y) \\ x \,_{\cdot}\, y &= x(1 - \sigma_0(y)\,\sigma_0(-x)) + y(1 - \sigma_0(x)\,\sigma_0(-y)) && \text{(A 7 2)} \\ &= x(1 - (1/2)\cdot(1/2)) + y(1 - (1/2)\cdot(1/2)) \\ &= 3/4\,(x+y) \end{aligned}$$

Thus, the approximating functions $\cdot_p$ and $\triangledown_p$ form a continuous deformation from a linear sum to the linear/logical operation as p goes from 0 to ∞

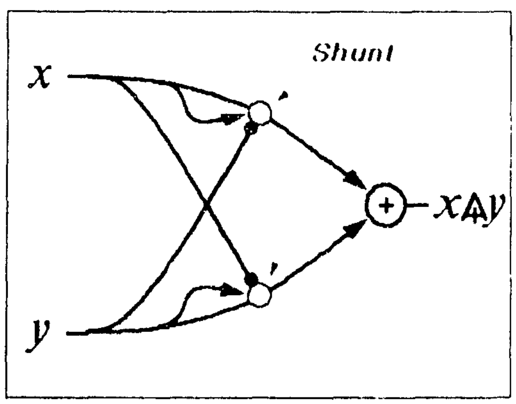


Figure A.2 Network implementation of $\mathbb{A}$ using a switching element The shunts (small circles) operate by passing their input signal unchanged unless all of the control inputs are positive in which case the output of the shunt is zero †

A local network model of a $\mathbb{A}$ circuit is shown in Figure A 2 This model is based on a simple principle, the *shunt* switch Such a switch acts as a resistive element for its input signal with the resistance controlled by the control input When the control input is negative, then the shunt is *on* and has a very low resistance, it passes its input through effectively unchanged When the control is positive the shunt is *off* and the resistance is very high, the output is effectively zero For simplicity, the shunts pictured allow multiple control inputs all of which must be positive for the shunt to turn off This mechanism is similar to the functional characteristics of a relay or transistor In neurological terms, it could be related to shunting inhibition ([Fatt & Katz 53] and [Rall 64]) in the dendritic tree of a single neuron The controls in this case would be realized as axo-dendritic or dendro dendritic synapses

Appendix B. Operators: From Continuous to Discrete

The operator designs in Chapter 2 are based entirely on analytic images and continuous convolutions In order to instantiate such a model in practice however, it is necessary to demonstrate that the model and the principles underlying the model can be translated into a framework in which the image is a set of responses from a finite number of discrete detectors This translation is straightforward since each of the continuous aspects of the model have natural discrete counterparts Continuous convolutions map to discrete convolutions and the continuous functions of the operator's components can be approximated by sampling As trivial as it may seem, specification of the parameters and assumptions of this translation has often been ignored by researchers in computer vision

Assume that the image is represented by a set of discrete detectors which have spatial transfer functions $\varphi_i \in \Phi$, a set of linear convolutions of the incident light intensity I. Consider these detector outputs as a set of basis functions from which the operator will be constructed Given a description of the linear operator as a continuous convolution with the kernel f, then we must find the vector of weights $W = \{w_1, w_2, \ldots, w_n\}$ such that $\sum_i w_i \varphi_i$ approximates f This can be accomplished by minimizing

$$d = \left\| f - \sum_{i=1}^{n} w_i \varphi_i \right\| \tag{B.1}$$

Using the inner product $(f, g) = \int fg \, d\mathbf{x}$ this minimization is accomplished by solving for W in

$$\begin{pmatrix} (\varphi_1, \varphi_1) & (\varphi_1, \varphi_2) & \cdots & (\varphi_1, \varphi_n) \\ (\varphi_2, \varphi_1) & (\varphi_2, \varphi_2) & \cdots & (\varphi_2, \varphi_n) \\ \vdots & \vdots & & \vdots \\ (\varphi_n, \varphi_1) & (\varphi_n, \varphi_2) & \cdots & (\varphi_n, \varphi_n) \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} (f, \varphi_1) \\ (f, \varphi_2) \\ \vdots \\ (f, \varphi_n) \end{pmatrix}. \tag{B.2}$$

When the basis functions φ_i are orthonormal, then the matrix collapses to the identity matrix and $w_i = (f, \varphi_i)$

The standard computational technique for converting from a continuous convolution kernel $f(x, y)$ to a discrete convolution matrix $F(i, j)$ is a specific instance of this framework For square pixels, the detector transfer function is most easily modelled as a square box Defining the square region

$$S(i,j) = \{ (x,y) : (i - 1/2 < x \leq i + 1/2) \wedge (j - 1/2 < y \leq j + 1/2) \} \tag{B.3}$$

then

$$\varphi_{i,j}(x,y) = \begin{cases} 1, & \text{if } (x,y) \in S(i,j), \\ 0, & \text{otherwise} \end{cases} \tag{B.4}$$

These functions are clearly orthonormal for integer i and j (volume is 1 and they are non-overlapping but complete) therefore the contribution from pixel (i,j) is simply the inner product $(f, \varphi_{i,j})$, or

$$F(i,j) = \iint_{S(i,j)} f(x,y)\, dx\, dy. \tag{B.5}$$

Assuming that f is planar within $S(i,j)$ allows the traditional simplification to $F(i,j) = f(i,j)$

We are not, however, limited to this instantiation This method explains how to map from continuous functions to sums of a discrete set of basis functions for any tiling of the detector plane The transfer functions φ can be round, can have square or smooth profiles and could even have the center-surround structure of a ∇^2G operator

Appendix C. Perpendicular Projection by Functional Minimization

In §3.4.3 the computation of relaxation labelling compatibilities was reduced to the perpendicular projection of label fixed points onto a manifold describing the geometric constraints of the structures to be discovered. The six-dimensional point representing the relationship between two curved tangent labels is $[\theta_i, \kappa_i, \theta_j, \kappa_j, x_j, y_j]$ and the polar description of the spatial relationship (x_j, y_j) is expressed as (d_{ij}, θ_{ij}). For the cocircularity constraint of Figure 3.3, this manifold is described in terms of the parameter coupling equations[†]

$$\theta_j = 2\theta_{ij} - \theta_i \tag{C.1.1}$$

$$\kappa_i = 2\sin(\theta_{ij} - \theta_i)/d_{ij} \tag{C.1.2}$$

$$\kappa_j = \kappa_i . \tag{C.1.3}$$

The perpendicular projection of the fixed point $[\Theta_i, K_i, \Theta_j, K_j, X_j, Y_j]$ onto the constrained manifold is then equivalent to minimizing the squared distance between the fixed point and any constrained point. Thus we minimize D^2 in[‡]

$$\begin{aligned} D^2 = & \left(\frac{\theta_i - (\Theta_i + 2a\pi)}{\Delta\theta_i}\right)^2 + \left(\frac{\kappa_i - K_i}{\Delta\kappa_i}\right)^2 \\ & + \left(\frac{\theta_j - (\Theta_j + 2b\pi)}{\Delta\theta_j}\right)^2 + \left(\frac{\kappa_j - K_j}{\Delta\kappa_j}\right)^2 \\ & + \left(\frac{x_j - X_j}{\Delta x_j}\right)^2 + \left(\frac{y_j - Y_j}{\Delta y_j}\right)^2 \end{aligned} \tag{C.2}$$

[†] Equivalent to using (C.1.2) is $x_j = (1/\kappa_i) * \left(\sqrt{1 - (\kappa_i y_j - \cos\theta_i)^2} - \sin\theta_i\right)$. This equation, however, has undefined regions and cases depending on the sign of the square root chosen, so it is more difficult to minimize. We thus opt for the other, even though having the three localization variables as the independent variables is attractive for other reasons.

[‡] The parameters a and b vary over $\{-1, 0, +1\}$ in order to deal with the problem of respecting the *congruence* of orientation instead of equality of angles.

The derivatives of this value with respect to the three independent variables θ_i, x_j and y_j are:

$$\frac{dD^2}{d\theta_i} = \frac{2\,(\theta_i - (\Theta_i + 2\pi a))}{\Delta\theta_i^2} - \frac{2\,(\theta_j - (\Theta_j + 2\pi b))}{\Delta\theta_j^2} - \frac{4\,\cos(\theta_{ij} - \theta_i)}{d_{ij}} \left(\frac{\kappa_j - K_j}{\Delta\kappa_j^2} + \frac{\kappa_i - K_i}{\Delta\kappa_i^2} \right) \tag{C.3.1}$$

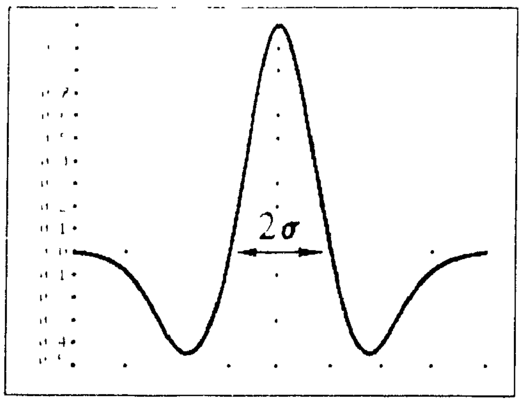
$$\frac{dD^2}{dx_j} = \frac{2\,(x_j - X_j)}{\Delta x_j^2} - \frac{4\,y_j\,(\theta_j - (\Theta_j + 2\pi b))}{\Delta\theta_j^2\, d_{ij}^2} - \frac{4\,(x_j \sin(\theta_{ij} - \theta_i) + y_j \cos(\theta_{ij} - \theta_i))}{d_{ij}^3} \left(\frac{\kappa_j - K_j}{\Delta\kappa_j^2} + \frac{\kappa_i - K_i}{\Delta\kappa_i^2} \right) \tag{C.3.2}$$

$$\frac{dD^2}{dy_j} = \frac{2\,(y_j - Y_j)}{\Delta y_j^2} + \frac{4\,x_j\,(\theta_j - (\Theta_j + 2\pi b))}{\Delta\theta_j^2\, d_{ij}^2} + \frac{4\,(x_j \cos(\theta_{ij} - \theta_i) - y_j \sin(\theta_{ij} - \theta_i))}{d_{ij}^3} \left(\frac{\kappa_j - K_j}{\Delta\kappa_j^2} + \frac{\kappa_i - K_i}{\Delta\kappa_i^2} \right) \tag{C.3.3}$$

The perpendicular projection is calculated using a standard numerical minimization technique from a starting point. For reasons of speed and accuracy (remember that this minimization must be performed for *every* pair of curved tangents in a local neighbourhood) we test a number of different staring points and use the one for which the initial D^2 is smallest. These potential staring points are chosen by projecting from the fixed point onto the constraint space along each of the three independent variable axes.

Appendix D. Relaxation Labelling Profiles for Localization

In §3.4 the importance of the localization of the selected labels for orientation selection is stressed. For general labelling problems the techniques outlined in [Zucker *et. al* 81] are appropriate, but for the two label case used throughout this work achieving such localization is more difficult. By considering a one-dimensional labelling problem, it will be demonstrated that linear compatibilities do not achieve the localization required. The localization is achieved instead by introducing a non-linearity into the support calculation. Consistent labellings are shown to obey the localization constraints as intended, and the convergence of the relaxation with such a support functional is experimentally verified.



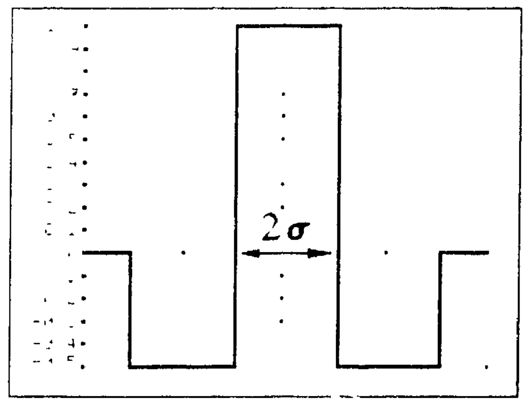


Figure D.1 Compatibility profiles for 1D relaxation labelling are (a) $\nabla^2 G_\sigma(x)$ and (b) the box operator. They both conform to the necessary constraints but differ in their convergence properties. In particular (a) does not converge while (b) converges but to a segment width much wider than itself.

In the geometric problems considered, the labellings which form meaningful solutions are discrete representations of smooth manifolds in the labelling space of the problem. The extent of such consistent labellings are therefore inherently localized on certain axes (the axes on which this is true are determined by the problem). In order to understand how such localization can be achieved by relaxation labelling in this multi-dimensional space I will show how it can be achieved for a one-dimensional labelling space and then rely on a generalization of this principle to higher dimensions.

Consider an arbitrary problem phrased in terms of locating those *isolated points* along a line which fulfill some condition. As the condition is defined locally, the line is discretely sampled and a measurement process tests each individual point along the line independently for conformance to the required conditions. The initial measurement process produces an estimate of its confidence that the conditions prescribed are satisfied at the given point. A relaxation labelling system is to be designed such that the regions over which the initial measurement produces positive results are localized to individual points. The similarities of this generic description to the orientation selection problem as defined in §1.2 should be obvious.

Notice that each node in this system is a two-label TRUE/FALSE set as described in §3.3.1 with p_i representing the confidence that position i satisfies the conditions being measured. If we adopt the notation $[v]$ to represent the scalar v clipped to the range $[0, 1]$ then the relaxation labelling update of Algorithm 1 Step 2.1 can be described in terms of the following equation

$$p_i^{n+1} = \left[p_i^n + \delta \left(\sum_{-m \le x \le m} p_{i+x}^n \, r_x \right) \right] \tag{D.1}$$

or, if $\mathbf{p} = [p_0, p_1, \ldots, p_n]$ and $\mathbf{r} = [r_{-m}, r_{-m+1}, \ldots, r_{m-1}, r_m]$

$$\mathbf{p}^{n+1} = [\mathbf{p}^n + \delta(\mathbf{p}^n \cdot \mathbf{r})] \tag{D.2}$$

where r_x is the compatibility between node i and node $i+x$ (assume that the compatibility between any two nodes depends only on their relative distance). The goal of such a relaxation labelling process is a consistent labelling with a set of isolated regions, each corresponding to a peak in the initial measurements. The problem is to find a compatibility function r_x which will achieve this goal, while minimizing the convergence time, ignoring noise, localizing the saturated regions as well as possible while allowing stable regions with as small a separation as possible. Whereas it would be extremely useful to be able to derive the compatibility structure directly from *a priori* constraints, at this point such a derivation does not seem possible. Instead, we will examine a number of possibilities and settle on the only one which appears to satisfy these goals.

A number of observations are immediately obvious The r_x's must be symmetric about r_0—looking left or right should be equivalent when locating a peak Furthermore, it is necessary to ensure that $\sum_i r_i = 0$ since the existence of a peak depends only on contrast and not absolute magnitude The stationary points of the relaxation (consistent labellings) are, as described above a set of constant width segments in which each node has unit confidence, all other nodes have zero confidence The response of the operator **r** must have a pair of zeroes at $-w/2$ and $w/2$ when convolved with a saturated segment of width w, it must also have non-negative responses between these zeroes and non-positive responses outside these bounds Furthermore the zeroes of reponse to any saturated segment wider than the desired width w must be closer to the center of the segment than the present boundaries (this is the gradient down which the relaxation will travel) The compatibility vector **r** is thus considered as a linear operator which fulfills these conditions

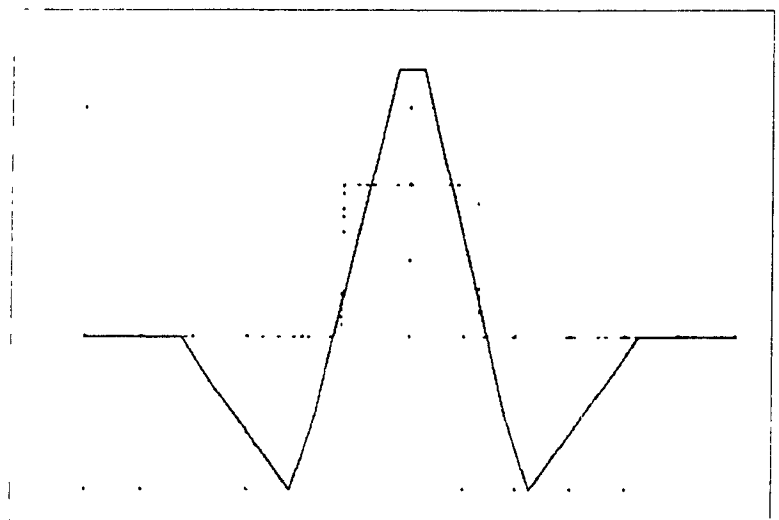


Figure D.2 Convolution of saturated segment with box operator of Figure D 1 (b) The saturated segment is defined in equation (D 3) and has width w The box operator has size described as σ This convolution shape is valid for $2\sigma < w < 3\sigma$ and the example shown is for $w = 2.5\sigma$

A close examination reveals that it may be impossible for a purely linear operator to fulfill all of these constraints simultaneously Two potential candidates for the desired operator are $d^2G_\sigma(x)$ and the box operator (see Figure D 1 (a) and (b)) For the $d^2G_\sigma(x)$ compatibility function it is straightforward to demonstrate that there is *no* stationary point for the equation (D 1) when **p** is a saturated segment Such a solitary saturated segment of width w is a stationary point iff the zeroes of $\mathbf{r} * \mathbf{p}$ are at $-w/2$ and $w/2$ However, the convolution of $d^2\mathrm{G}_\sigma(x)$ with

$$p_w(x) = \begin{cases} 0, & \text{if } x < -w/2, \\ 1, & \text{if } -w/2 \le x \le w/2, \\ 0, & \text{if } w/2 < x \end{cases} \tag{D.3}$$

is

$$d^2\mathrm{G}_\sigma(x) * p_w(x) = \frac{1}{2\sigma^2}\left((2x - w)e^{wx/\sigma^2} - 2x - w\right)e^{-(2x+w)^2/8\sigma^2} \tag{D.4}$$

which is zero when

$$\frac{2x}{w} = \frac{e^{wx/\sigma^2}+1}{e^{wx/\sigma^2}-1} \tag{D.5}$$

Setting $x = w/2$ for a stationary point gives

$$e^{w^2/2\sigma^2}+1 = e^{w^2/2\sigma^2} - 1 \tag{D.6}$$

which is never true Therefore using $d^2\mathrm{G}(x)$ as a relaxation labelling operator with two-label nodes is infeasible

The other possibility presented is the box operator of Figure D 1 (b) The convolution of this operator with (D 3) depends on the relative sizes of the segment and the operator For the range $2\sigma < w < 3\sigma$, the zero is at

$$x = \frac{3w^2 - 2\sigma w - 48\sigma^2}{2(5w - 24\sigma)} \tag{D 7}$$

Again, this is a stationary point only if $x = w/2$ and therefore when

$$w = 3\sigma \tag{D 8}$$

The difficulties arise when considering the behaviour for $w > 3\sigma$ In this case, the zero is always at $w/2$ and all segments with $w > 3\sigma$ are stationary points Because of (D 8) any segments narrower than 3σ will widen, and we have just shown that any wider than this will remain the same This is hardly the localization process described above

An approach which does work is to achieve localization by adopting exactly the same cross-sectional operator structure as was used in the initial convolutions This time we develop a partition over the cross-section of the compatibilities With the operator described in equation (2 16), the convolution (D 1) is decomposed into two convolutions which are later combined The relaxation then becomes

$$\mathbf{p}^{n+1} = [\mathbf{p}^{n} + \delta(\mathbf{p}^{n} * \mathbf{L} - \mathbf{p}^{n} * \mathbf{R})] \tag{D.9}$$

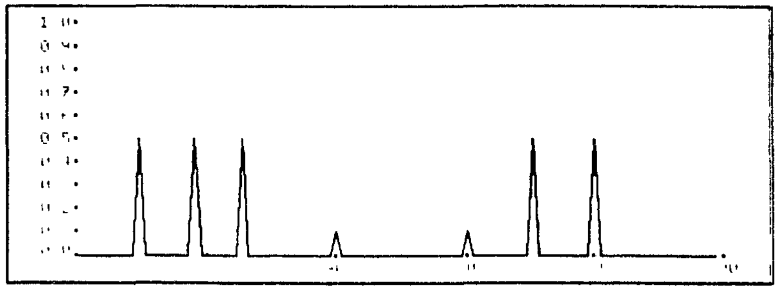
where **L** and **R** are the left- and right-hand $dG(x)$ operators of §2 2 2 Not only does this operator localize the node responses as required, but it does so extremely quickly, using only a small number of iterations As was pointed out previously, the region of positive response is confined to within σ of the actual peak In order to estimate the rate and direction of convergence it is only necessary to examine the distance between the zero of the response and the transition point at $w/2$ The convolution for one component of the operator is

$$dG_{\sigma'}(x+\delta) * p_u(x) = \sigma\left(e^{w(x+\delta)/2\sigma'^2} - 1\right)\left(e^{w(x+\delta)/2\sigma'^2} + 1\right)e^{-(2x+w+2\delta)^2/8\sigma'^2} \tag{D.10}$$

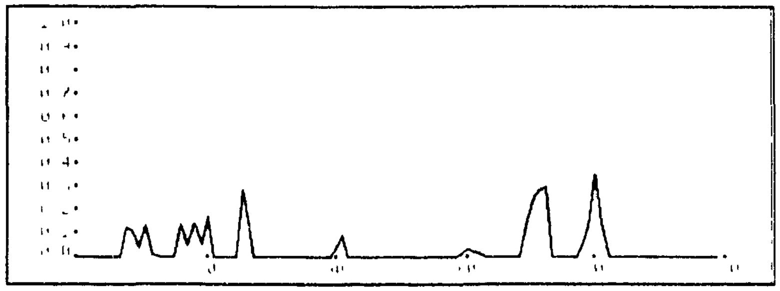
which has a single zero at $x = -\delta$[1] (By symmetry, the other component has a zero at $x = \delta$) Thus the stationary point of the relaxation is a saturated segment of width $< 2\sigma$ and the relaxation convergences extremely quickly

To demonstrate this result, we will compare the operations of these three compatibility profiles on an ideal input (Figure D 3(a)) and a degraded version of the same (Figure D 3(b))

The three diagrams in Figure D 4 show the convergence of the three candidate compatibility profiles when given the raw delta functions as initial measurements A perfect analysis should simply saturate those labels which correspond to the locations of the delta functions With $\sigma = 3$ (for noise suppression) the non-linear relaxation operator is clearly superior to the others the regions are more localized, and they are more accurately placed Not specifically the huge displacement of the peak at 60 for the 'box' operator The operation of the relaxation on the degraded input is shown in Figure D 5 Again the non-linear relaxation is clearly superior, for much the same reasons as before



(a)



(b)

Figure D.3 Input to one-dimensional relaxation whose goal is to recover the locations of the delta functions in (a) which the noisy 'measurements' in (b) were derived from

The extension of this result to higher dimensions is logical. For a smooth manifold of dimension n embedded in an m dimensional space there are locally $m - n$ orthogonal directions in which the manifold is localized. When such a space is discretized for representation in terms of discrete labels, then for each of these local axes a d^2G decomposition can be introduced into the compatibility network. The support for a given label is then the $\bigwedge$ of these component supports. An application of this approach to orientation selection is described in §3.4.5.

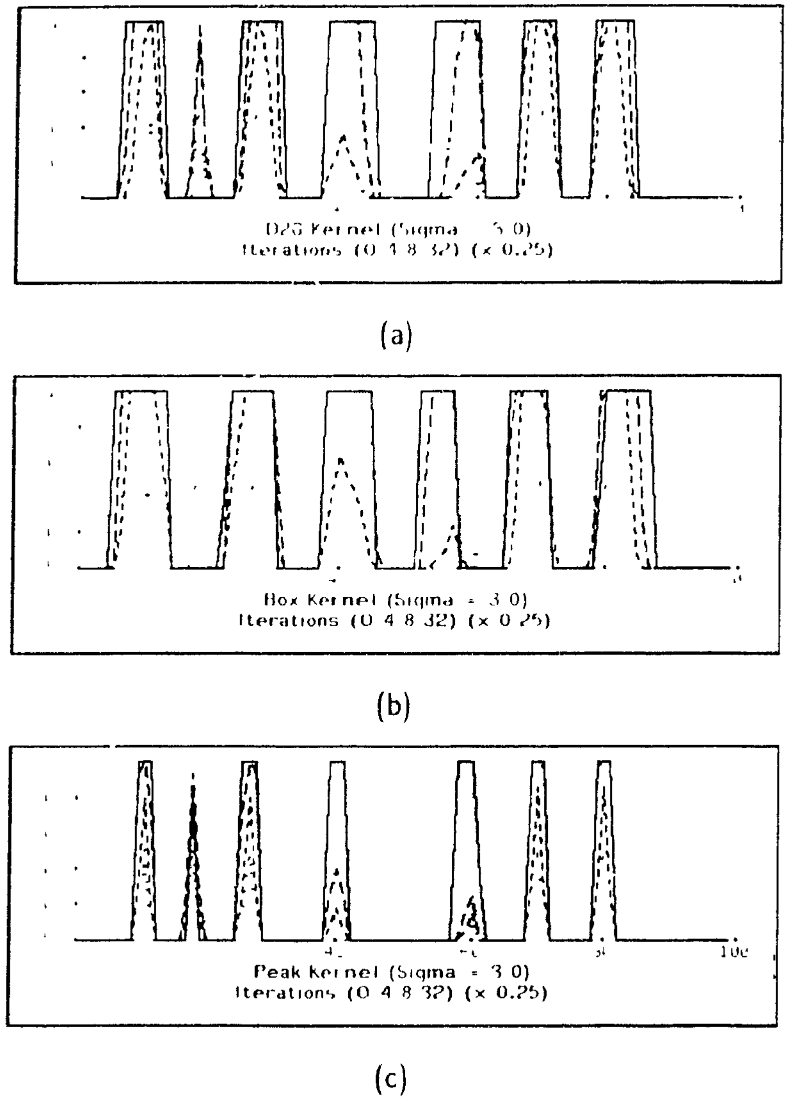
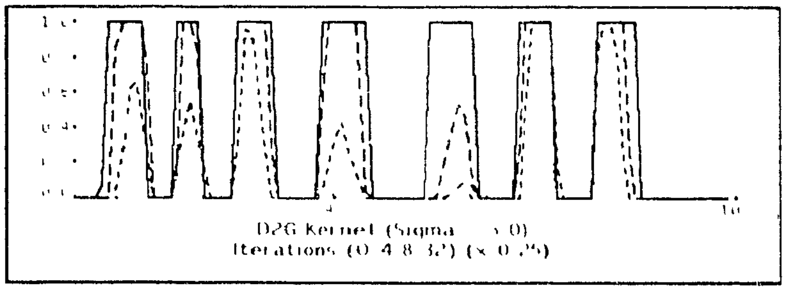
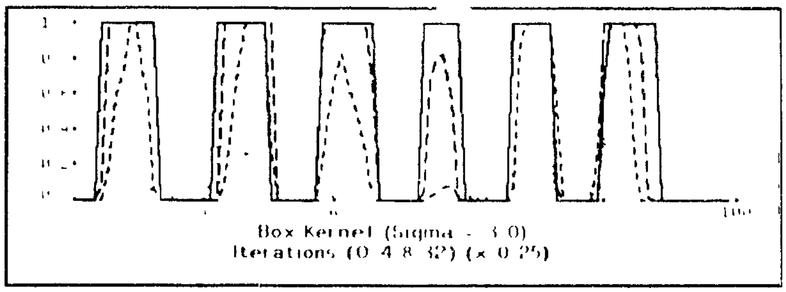


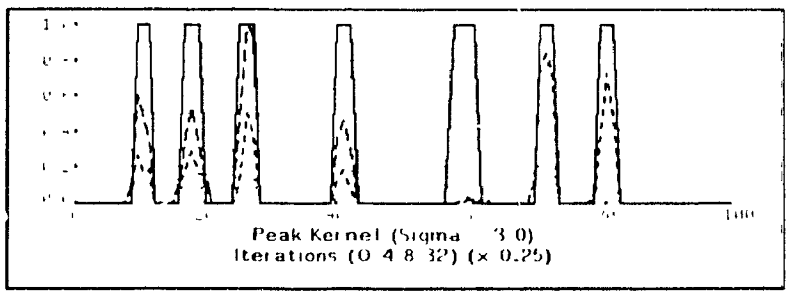
Figure D.4 Results of relaxation on the perfect input of Figure D 3(a) Three compatibility profiles are tested (a) $d^2G(x)$ (b) 'box', and (c) non-linear peak operator



(a)



(b)



(c)

Figure D.5 Results of relaxation on degraded input of Figure D 3(b) Three compatibility profiles are tested (a) $d^2G(x)$, (b) 'box and (c) non-linear peak operator