The extended Euclidean distance transform

Mark William Wright

Churchill College

Department of Engineering

June 1995

Dissertation submitted for the
Degree of Doctor of Philosophy at the
University of Cambridge

Abstract

Shape representation has always played a central role in computer vision. Skeletal shape descriptors which make symmetry explicit are an important class of shape representations. The goal of this thesis is to study the problems encountered using skeletal shape descriptors. The thesis unites three main themes of work: A filter based approach to skeletonisation, skeletonisation using parallel wave propagation and skeletonisation using an extended Euclidean distance transform.

The distance transform approach to skeletonisation computes a skeleton by identifying the so called local maxima of the distance transform. A new method has been proposed to detect these features using a filter-base approach inspired by models of processes in the human visual system. Further improvements were made by using a filter designed to detect a specific geometric feature on the distance transform which corresponded to the skeletal points. This improved the quality of skeletons obtained but could only compute restricted skeletal descriptions.

The wave propagation algorithm of Brady and Scott has been studied; they originally implemented this on a simulator of the Connection Machine. The issues of mapping the algorithm onto an array of transputers have been investigated. An efficient implementation was realised by reducing synchronisation and data transfer overheads. It was found that the algorithm could compute more general shape descriptions than the distance transform approach but the quality of skeletons produced was not as good.

Using standard techniques from singularity theory, an analysis of distance functions from object boundaries has been undertaken. This resulted in the formal definition of a new extended Euclidean distance transform. An algorithm has been devised to perform skeletonisation using the extended distance transform. This combined the advantages of the filter and wave based techniques in that it produced skeletons of a high quality which made more symmetries explicit than the standard distance transform approach. In addition, the extended distance transform provides an elegant unifying framework for work on skeletal shape descriptions.

Acknowledgements

Above all I would like to thank my supervisor Roberto Cipolla for his support, advice and enthusiasm. I am especially grateful to Maria Petrou for supplying information, data and advice on her filters. Likewise to Peter Giblin for sharing his great knowledge of singularity theory with such patience and good humour and for proof reading my work on this subject. I would like to thank my good friend Ian Green for his advice on structuring the thesis and the laborious task of proof reading the first draft. Thanks to Steve Zucker for his time looking at my work and giving advice while at the Newton Institute. Thanks to everyone in the Speech, Vision and Robotics group at the Department of Engineering, particularly Andrew Senior, Patrick Gosling, Richard Prager, Tony Robinson, Ron Daniel, K. G. Lim and all the vision group. I am indebted to the University of Edinburgh and particularly Bob Fisher for allowing me access to computer resources to complete this thesis and for his patience and understanding. Andrew Fitzgibbon, Gillian Hayes and Douglas Howie have also been very helpful. Financial support has come mainly from British Aerospace and the S.E.R.C. Other contributions from the Engineering Department and Churchill College for conferences etc, were also gratefully received. Thanks also to Patricia Gilmour and Fay Shamanski for practical help and moral support. Finally I would like to remember the late Professor Frank Fallside for his support and encouragement which was always delivered with characteristic charm and wit. My one regret at the end of my thesis is I can not thank him personally but for the record thank you Frank.

This Thesis is dedicated to my brother and sisters and to the memory of my parents.

Like all young men I set out to be a genius
but mercifully humour intervened

—Lawrence Durrell, *Clea*

Declaration

I declare that this dissertation is the result of my own original work. Where my research has drawn on the work of others, this is acknowledged at the appropriate points in the text. This dissertation has not been submitted in whole or part for a degree at any other institution.

Contents

List of Figures

Chapter 1

Introduction

The major subject area in which this work is set is that of computer vision. The representation of shape has always played a central role in this field. This thesis focuses on a particular group of shape representations called *skeletons* and the problems encountered in computing them. Informally, a skeleton is a representation which reduces a 2D or 3D shape to an axial or "stick like" form. In this chapter the motivation behind this work is put forward, then the methodology adopted is outlined and finally the contributions made by the thesis are presented.

1.1 Motivation

1.1.1 The influence of Marr

The seminal work of David Marr had a profound effect on computer vision research in the 1980s. The philosophy he advocated emphasised the rigorous treatment of vision as a computational process [62] which can be evaluated at three levels (see below). This is in contrast to previous work in vision, which was largely devoid of any general framework; work on shape description was no exception. (See Pavlidis [69] for a review of early work on shape description and Marshall [65] for a similar but more recent account.) A natural outcome of Marr's work is an emphasis on representation (Marr refers to this as a "representational" theory). The importance of skeletal representations and hence symmetry arose predominantly from this viewpoint. Marr and Nishihara [64] considered three aspects of representation design. These were the representation's coordinate system, its primitives, which are the primary units of shape information used by the representation, and finally the organisation that the representation imposes on the information. From this study they determined that *axial* or *skeletal* representations would be good candidates for shape representation as they have the following desirable traits. Firstly, they have an object-centred coordinate frame. This is more desirable than a viewpoint-centred frame as it leads to a more compact and invariant description. Secondly, they are volumetric rather than surface based. Volumetric primitives give information about the spatial distribution of a shape providing more compact and stable descriptions. Thirdly, they can be constructed so as to have a modular organisation; this allows the decoupling of stable, coarse level information from sensitive, fine level information.

1.1.2 The contemporary perspective

Over the past few years there has been a significant shift towards a more pragmatic approach to vision research. The new approach is task based in nature with an emphasis on demonstrating specific dynamic visual competences in real-time. This change in emphasis has been driven partly by theoretical problems with standard approaches and partly by technological advances making cheap real-time image acquisition and processing available [14]. In active vision [1] it is shown that many vision problems are more easily solved by an active observer than by a passive one. Animate vision [9] stresses the central role of gaze control. Symmetry is a powerful cue for directing attention: many natural and man made objects have bilateral symmetry or at least a primary axis of growth. Indeed, some of the skeletal shape descriptors that will be introduced in this thesis were inspired from analogies with biological growth. The problem of real-time hand/eye coordination is receiving increasing attention [49], symmetry and its representation using skeletal shape descriptors is again found to be of central importance in the computation of stable grasps [13].

1.2 Addressing the problems of skeletonisation

It is clear from the above that skeletal shape representations have had an important role to play in vision research over the years and continue to be of importance today. However, it is also the case that many problems exist in their use and that these need to be addressed. Skeletal representations are notoriously sensitive to noise. This can be due to quantisation and sensor noise or small variations in the shapes presented themselves. Occlusion of shapes and missing boundaries due to imperfect segmentation can produce very different skeletons. Surface markings and cluttered scenes can also disrupt descriptions. In addition to these problems, certain descriptions may require considerable computation time to acquire which is particularly important in today's real-time applications. The goal of this thesis is to address these problems in a principled and coherent manner.

1.3 Methodology

Marr made many specific contributions to vision research but perhaps one of the most enduring was to provide a framework within which vision research may be conducted in a principled way. In this thesis Marr's general framework is used as a tool to describe and provide structure to the work. This thesis makes use of two elements of the framework which are described below. Firstly, criteria for shape representation are provided; these are used in this thesis to focus on specific problems of skeletonisation algorithms. Secondly, Marr defined three conceptual levels at which one should think about issues in vision research. These levels are used in this thesis to provide structure to the solutions of the problems identified using the criteria.

1.3.1 Criteria for shape representation

One influence of Marr's work is that much work has been done in trying to deduce the nature of the internal representations used in the human visual system and to determine the criteria for useful representations in machine vision systems. Marr and Nishihara [64] presented three criteria for judging the usefulness of a representation for three-dimensional shape recognition. They are:

Accessibility: Is it possible to derive the representation from the image and can this be done within reasonable costs of computation time and memory requirements?

Scope and uniqueness: Is the representation appropriate for the class of shapes it is likely to encounter? Do these shapes have canonical descriptions in the representation and is there a single unique description for each shape encountered?

Stability and sensitivity: A good representation should be able to embody these two opposing characteristics. Stable representations stress the similarity between shapes or recognise the same shape when transformed by noise or other factors. A sensitive representation can pick out subtle differences between shapes. Marr and Nishihara suggested there should be a decoupling of stable features from fine detail.

Brady [20] made a further contribution in this area by proposing three further criteria: *Rich, local support*: “Rich” means that the representation should be information preserving. “Local support” means that it should be possible to obtain the representation through local computations.

Smooth extension and subsumption: Smooth extension is a structuring process whereby local frames should give rise to more global frames. Subsumption reduces the set of frames by rejecting locally feasible alternatives which are wholly contained within other frames.

Propagation: Frames corresponding to subparts of a shape can be propagated by inheritance or affixment whenever the corresponding frames are consistent.

Brady’s criteria identify specific key strategies that a shape descriptor should embody to be successful in real applications, in other words, these criteria are at an algorithmic level (see below). The criteria due to Marr and Nishihara are of a more generic nature suitable for the evaluation of any shape descriptor; it was therefore decided to use these rather than Brady’s to identify problems.

1.3.2 Marr’s theory: the three levels of abstraction

A crucial concept in Marr’s approach is that to understand vision we need to develop an understanding at not just one level, say at the level of neurons, but at many levels. Following this line of thinking Marr suggested that any adequate theory of vision should consist of explanations at three levels which he called computational, algorithmic and hardware.

The Computational Level

This is the highest and most abstract of the theoretical levels of understanding. Here we consider the *goal* of vision, what is its purpose. Fundamentally the goal of vision is to take an image of the world and construct a representation of the shapes and spatial arrangements of objects in the world that gave rise to that image. Marr specifies this as a mapping from the image on the retina to the internal representations of 3D objects in the scene.

But the computational level of explanation consists not just of a simple statement of the task to be achieved. It states not just what is being computed, but why, what are the constraints of the task that can be exploited in its execution. In the case of vision these constraints are borne out of the nature of the physical world and a close study of the image formation process. The absence of a distinction between "what" and "how" had previously bogged down attempts to form a rigorous theory of vision and had led to much work being lost in subjective arguments about implementation details.

The Algorithmic Level

At this level we consider by what *method* the task defined at the computational level can be achieved. In particular it is the form of the input and output representation that is of interest and the nature of the algorithm which might perform this transformation.

The Hardware Level

At this level we consider by what *means* a particular algorithm and representational scheme are physically implemented. Many hardware schemes can execute the same algorithm. Edge detection using the Marr–Hildreth operator [63], for example, could be implemented on a computer, by a dedicated electronic circuit or by the cells of an animal's retina.

1.4 Contribution

The work in this thesis unites three main themes of research which are outlined below. For each piece of work a brief description is given. The criteria addressed are stated in each case along with the conceptual level at which they are tackled. Details of any separate publications by the author are also given. Preceding the three main themes, chapter two contains a literature review which discusses skeletonisation algorithms and the problems that exist in using them. This review makes use of Marr's framework to clarify the confusion which exists in the literature between "function" and "procedure".

1.4.1 A filter based approach to skeletonisation

A standard method to perform skeletonisation is to use a distance transform. This computes a value for each pixel within the shape equal to the minimum distance of that pixel to the shape boundary. The "local maxima" of the transform correspond to the symmetric axis transform or SAT which is a skeletal representation. In chapter three a new method for detecting these "local maxima" is presented. This is done by posing the detection process as a filtering problem. Firstly, it was observed that the response of the human visual system to a grey level representation of the transform accentuates the "local maxima". Convolution of the distance transform with the Marr-Hildreth operator was proposed as this is a model of the relevant process in the human vision system. Secondly, it was suggested that a filter could be used to detect a specific geometric feature on the distance transform which corresponded to the skeletal axis. From inspection of the distance transform it was deduced that a "roof profile" would be a candidate feature. The Petrou filter was then used to detect the "local maxima" as this can be configured to detect such a feature. It was shown that this filter produced more favourable results when compared with the classic ridge finding algorithm previously used and other skeletonisation algorithms. In addition the detection filter can be tuned to different scales to give a multi-scale representation.

In terms of the criteria for shape representation given above this work tackles the issues of *stability* and *sensitivity*. The contribution of this work is to improve a subpart of an existing skeletonisation technique. This work addresses issues at an algorithmic level. It demonstrates that interesting and useful techniques can be derived from an exploration of the close links between biological and machine vision. Two papers based on this work have been published [96, 99].

1.4.2 Skeletonisation using parallel boundary propagation

The filter based approach provided some improvement in skeleton quality over the standard distance transform approach. Unfortunately it shares a limitation inherent in all distance transform based methods, namely that only the symmetric axis transform (SAT) can be computed and not more general axial representations such as smoothed local symmetries (SLS) or the symmetry set (SS). One method which overcomes this limitation is a wave based algorithm proposed by Brady and Scott which they implemented on a simulator for the Connection Machine (a powerful computer with a massively parallel architecture). In chapter four the issues involved in mapping this algorithm to an entirely different type of architecture, *i.e* a MIMD (multiple instruction multiple data) machine are investigated. The algorithm was actually implemented on an array of transputers which is an example of such a MIMD architecture. An array of transputers cannot match the huge computing power of the Connection Machine but it was found that an efficient implementation was possible if synchronisation and data transfer overheads were minimised. The key to this minimisation was to divide the wave propagation process between processors according to distance from the boundary, so called *interval decomposition*, rather than to give each processor a separate portion of boundary.

In terms of criteria for shape representation the focus of this work was *accessibility*. The contribution of this work was to show how an existing algorithm could be efficiently applied in the new domain of a different architecture and to demonstrate this on a real example of such an architecture. This chapter addresses issues at the hardware level. A paper based on this work has been published [94].

1.4.3 Extending the Euclidean distance transform

The wave propagation algorithm proposed by Brady and Scott transcends the limitations of the distance transform approach to skeletonisation. It performed well on thin elongated shapes but less well on more rounded shapes. However, in general the quality of skeletal output seemed to be poorer than that achieved using the distance transform approach. It was therefore decided that a new approach was needed which combined the strengths of the wave propagation and distance transform methods, *i.e.* which provide high quality skeletons and produce a more general axial description of shape than the SAT.

In chapter five a mathematical analysis of distance functions from smooth planar curves is presented. Using standard techniques from singularity theory, an *extended* Euclidean distance transform or *discriminant* is defined formally. The relationship between this extended distance transform, the standard distance transform, SAT and the geometry of the shape boundary is presented. In particular the relationship between the discriminant and the wave propagation approach is highlighted giving fresh insights into the results obtained using the parallel wave propagation approach. In chapter six an algorithm is presented which uses the extended Euclidean distance transform to perform skeletonisation. Using a number of examples it is shown that the algorithm produces high quality skeletons which make explicit more symmetries than the symmetric axis transform obtained using the standard distance transform.

This work again focuses on *stability* and *sensitivity* but is also concerned with issues of *accessibility* which make the algorithm more efficient particularly in terms of memory requirements. The main contribution of this work is the generalisation of an existing concept (the Euclidean distance transform) and to demonstrate the practical benefits which can stem from the application of this generalisation. A secondary contribution is that the *extended* Euclidean distance transform or *discriminant* can be used as a tool to provide an elegant unification of previously disparate elements of skeletonisation theory. The various skeletal shape descriptors, the standard distance transform, the wave propagation approach and various geometric phenomena of shape boundaries such as curvature extrema and evolutes can all be described in terms of their relationship to the *extended* distance transform. Chapter five addresses issues at the computational level whereas chapter six is concerned with algorithmic and hardware issues. A paper based on this work has been published [97] and an extended version is in press [98].

1.5 Summary

Shape representation has always played a central role in computer vision. Skeletal shape descriptors which make symmetry explicit are an important class of shape representations. The goal of this thesis is to study the problems encountered using skeletal shape descriptors. Marr's representation theory of vision has been chosen as a framework in which to work. Specifically Marr and Nishihara's criteria for shape representation have been used to identify problems. The three levels of abstraction in Marr's theory have been used to provide structure to the solution of these problems. The thesis unites three main themes: A filter based approach to skeletonisation, skeletonisation using parallel wave propagation and skeletonisation using an extended Euclidean distance transform. The significance and contribution of each piece of work has been described along with publications resulting from them.

Chapter 2

Literature survey

2.1 Introduction

In this chapter we outline the main techniques which have been used by other researchers to compute skeletal representations. Skeletonisation can be thought of as a family of shape descriptors which all have the same aim of reducing an input shape to some kind of axial representation.

A theme that we will return to later is that many algorithms, particularly thinning algorithms, are defined solely by "procedure" rather than "function". By this we mean that we are given information about the implementation of the algorithm, *e.g.* the rules for pixel removal, but we are not told what the relationship is between the initial shape boundary and the final skeleton. This makes comparison and performance measurement difficult as the exact goal of the algorithm is unclear.

We could point to two possible causes of what might seem a highly ad hoc approach to the subject. Firstly, for many of the applications of skeletonisation algorithms, such as preprocessing for optical character recognition and inspection of tracks on electronic circuit boards, the border/skeleton relationship is not so crucial. These are inputs which are highly elongated and perhaps only a few pixels wide anyway. The main requirements in these cases are that the resultant skeleton be "thin" and perhaps of most importance that it be connected. Only when we deal with more complex shapes which are not so filament like will the boundary/shape relationship be more crucial. Thinning algorithms *are* used as shape descriptors in these instances and this brings us to the second reason. It would seem that there remains a confusion between "function" and "procedure". This confusion is evident in the fact that the word "thinning" is used interchangeably to describe a function or a procedure. It is interesting to note that a similar confusion seems to have once surrounded one particular skeletal representation, *i.e.* the symmetric axis transform (SAT). In early literature the SAT was often synonymous with the "prairie fire" method proposed by Blum to compute it, so much so that it was sometimes called the "prairie fire" or "wild fire" transform. Today there is a much clearer distinction between the SAT as the skeletal descriptor or "goal" and the "wild fire" algorithm as just one "method" of computing it.

In this chapter Marr's framework is used to clarify the confusing situation outlined above. Specifically a distinction is made between the computational level description of the shape descriptor, *i.e.* its goal, and its algorithmic level discription, *i.e.* the method by which it is obtained.

2.2 Skeletonisation using bi-tangent circles

Four skeletal shape descriptors are now introduced which can be thought of as a group. These are:

1. Symmetric Axis Transform (SAT) [15, 16]
2. Smoothed Local Symmetries (SLS) [21]
3. Process-Inferring Symmetry Analysis (PISA) [59]
4. Symmetry Set (SS) [42]

This group receives particular attention because its members have a clear description at the computational level. This description is in terms of a geometric relationship between the shape boundary and skeletal points. The attribute they have in common is that they all produce their skeletal output by the use of bi-tangent or multi-tangent circles, *i.e.* circles which are tangential to the shape boundary in two or more places. The main differences between them are where the skeletal point is placed with respect to each circle and any restrictions which disqualify certain circles from contributing to the skeleton.

2.2.1 Symmetric axis transform (SAT)

The symmetric axis transform was introduced by Blum [15, 16]. The motivation behind this work was the description of biological shape and the belief that concepts of growth and symmetry were central to this goal. The SAT can be defined in a number of ways. From the computational level, each point on the skeleton is the centre of a circle which is bi-tangent to the shape boundary and does not cross the boundary. Another way to say this without reference to tangency is to say the skeleton is the locus of the centres of maximal discs, *i.e.* discs which lie within the shape but which are not contained within any other disc. The importance of the latter definition is that there is no prerequisite to explicitly compute tangents or normals as there is for other shape descriptors. Note that in the mathematical literature the SAT is often referred to as the *central set* [101].

An algorithmic level description of the SAT can be defined with reference to the analogy of a grass fire. If we imagine the shape as a grass field and a fire started all the way around its perimeter which marches uniformly to the middle, then the SAT is the locus of quench points in this process, *i.e.* where the firefronts meet.

SAT extensions

Over the years, the SAT has been modified or extended in many ways. Here we will discuss three areas of development: 3D shape description, multi-resolution shape description and greyscale image description.

3D shape description

The SAT is easily extended to 3D shapes. Bi-tangent circles or maximal discs are replaced by spheres [68, 81]. However, we should be aware that the SAT is very sensitive to missing data. Any surface information gained from a sensor such as a laser range finder on a robot arm would only give data for one side of the object. The SAT created from this incomplete "shell" of range data could be very different from the SAT generated from the complete surface of the object.

Multi-resolution shape description

Multi-resolution shape description (MRSD) is an important general technique in which shape description is built up over a range of spatial scales; this has a number of benefits over descriptions at a single scale. Firstly, MRSD can be used to tackle the problems of noise. Noise generally occurs at higher resolution leaving the coarser scale representations less affected. It also leads to a natural hierarchical description of the shape where the coarse scale features are regarded as the more perceptually relevant. Performing operations such as matching on a hierarchical description can be more computationally efficient as many candidates can be ruled out at coarser scales before final analysis at finer resolutions. There is some evidence that such a strategy could be used in the human visual system [93].

The SAT was extended to multi-resolutions by Pizer, Oliver and Bloomberg [76]. They performed 2D blurring of a binary image using Gaussians of various widths. The output of each convolution was renormalised to keep the area of the binary shape constant. The SAT is computed for each scale and then the branches of the resulting skeletons tracked (by hand, not automatically) across scales. A big issue in multi-resolution shape description is whether the smoothing is performed in 2D across the image or in 1D along the boundary. Pizer, Oliver and Bloomberg argue that region smoothing gives more intuitive results than boundary smoothing. They point out four problems with their smoothing approach. Firstly, with 2D smoothing the topology is not preserved, secondly, concavities are not made explicit, thirdly, tracing of skeletal branches is difficult and finally, hierarchical grouping of branches of similar scale is unstable as their parent/child dependencies can easily be reversed with a small change in the shape.

Grayscale image description

The idea of the SAT has been extended to the description of grey level images by J. M. Gauch using what he defines as the Intensity Axis of Symmetry or IAS [41]. The basic idea is to think of the grey level image as a three-dimensional surface and to cut this into horizontal slices at particular grey levels. These slices together constitute "level sets" which are rather like the contours marked on a geographical map. The IAS is built up by calculating the SAT for each level. This consists of a skeleton "inside" the contour, corresponding to image pixels above the given intensity level, and a skeleton "outside" the contour corresponding to pixels below the given intensity level. These individual skeletons contribute to a three-dimensional stack of skeletons which form a branching surface which is the IAS itself. Gauch presents an algorithm for efficient computation of the IAS based on an extension of the active contour model to an active surface model.

2.2.2 Smoothed local symmetries (SLS)

Smoothed local symmetries were developed by Brady and Asada to overcome some of the problems associated with the SAT. The smoothed local symmetries representation is perhaps unique as a shape descriptor as it includes both a boundary based description of significant features on the contour called the curvature primal sketch [7], and a region based skeletal description, based on the concept of local symmetry [21].

SLS boundary description

The SLS boundary representation is called the "Curvature Primal Sketch" [7]. It derives its name from the "Primal Sketch", of which it is a close analogy, advocated by Marr [62] to describe low level processes in the human visual system. The goal of the "Curvature Primal Sketch" is to compute a multi-scale representation of a shape to form a hierarchical syntactic description of the contour. This is achieved as follows. A set of basic curvature discontinuities is defined. These include two single discontinuity types called "corner" and "smooth join" and multiple types such as "end", "crank", "bump" and "dent". Parametric versions of these discontinuities are derived and analytic forms of their convolutions with the first and second derivatives of a Gaussian are calculated. These are used as a set of reference features to be matched against possible candidate curvature features. The boundary of a shape is first localised using the Canny edge detector [30]. Responses for the convolution of the contour with the first and second derivatives of a Gaussian are built up over a range of spatial scales. The local positive maxima and local negative minima are matched across scales to produce a tree-like scale space representation of the curvature discontinuities. This tree is then parsed from fine to coarse scales into instances of the curvature primitives defined above. Once these are detected there is a further localisation step to refine their estimated positions and to define knot points on the contour.

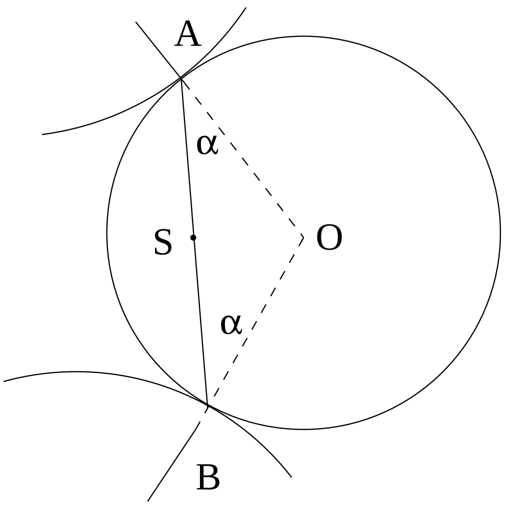
SLS region description

The definition of a point of local symmetry is shown in figure 2.1. Two points on the boundary such as A and B are deemed locally symmetric if the angles α between their inward normals and the chord joining them are equal.

The normals and chord form an isosceles triangle ABO where the two sides AO

and BO are equal. We can thus construct a circle centred at O and bi-tangent to the boundary at A and B. Although this circle is not part of the original definition of the SLS it can be constructed for any SLS skeletal point and gives us a link and means of comparison of the SLS with the SAT, PISA and symmetry set.

For example the SLS point is drawn at the midpoint of the chord AB and not the centre of the circle at O as for the SAT and symmetry set. The locus of all such points produces the SLS skeleton.



The two boundary points A and B are locally symmetric if the angles formed between their inward normals and the chord joining them are equal. The skeletal point is drawn at S, the midpoint of this chord.

Figure 2.1: Definition of a smoothed local symmetry

Rotational symmetries

Skeletal shape descriptors are poor at describing near circular shapes. Small changes in the boundary give rise to different and unstable skeletal descriptions. If we consider the SLS for a circle, all edge points are locally symmetric with all others. If all raw symmetry points were drawn in such a case (ignoring the effect of discretisation) they would fill the circle entirely. SLS recognises this weakness of skeletal shape descriptors and deals with rotational symmetry explicitly [40]. Again computation proceeds over a range of spatial scales with a search for rotational symmetries from coarse to fine scales.

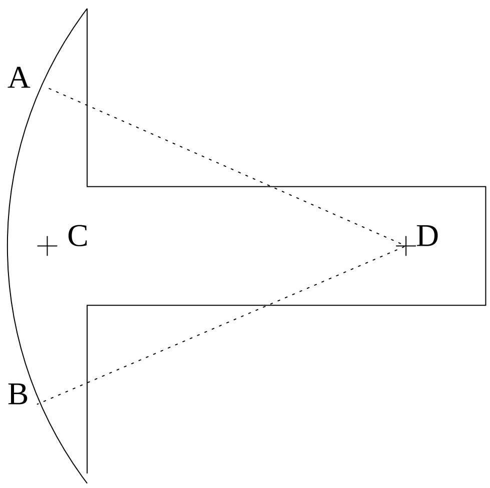
At each scale candidate points in the image are tested for their suitability as centres of rotational symmetry. A line is defined with one end fixed at the point of interest and the other somewhere on the boundary of the shape. As this line is swept around the boundary a note is kept of its length and the angle it makes with the outward normal of the boundary. For a point chosen at the centre of a circular boundary the length of the line will be constant and the angle with the normal will remain at zero. For near circular parts thresholds are applied to these functions. All boundary points within these thresholds are marked as rotationally symmetric with respect to the point of interest.

2.2.3 Comparison of the SLS with the SAT

The SAT places skeletal points at the centre of the bi-tangent circle as does the symmetry set (see below). The SLS on the other hand places the skeletal point at the midpoint of the chord between the tangent points. The midpoint of the chord was chosen so that the skeletal points would be nearer to the boundary points generating them for shapes such as that in figure 2.2. The skeletal point associated with boundary points A and B is placed at C rather than D. In fact, in the case of figure 2.2 the SAT would not allow the circle centred at point D and bi-tangent to points A and B to contribute to the skeleton as the circle is not contained entirely within the shape. However, if we consider the boundary points in figure 2.3 we see that the SAT points can vary their distance along the axis whereas the SLS points are to be found on normals to the axis, and this segments the axis is a more intuitive way.

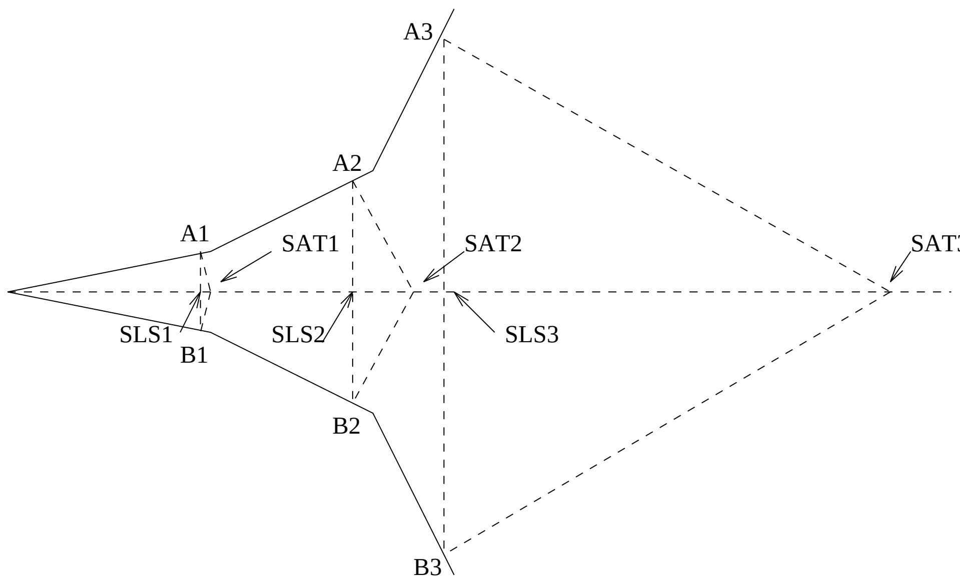
The fact the SLS points are on the chord joining boundary points ensures that the internal and external SLS skeletons are contained within the convex hull of the shape in question. Another consequence of this choice is that in general the SLS skeleton is more compact than the SAT, *i.e.* for the same length boundary the SLS will be less than or equal in length to the SAT but never longer. For a skeleton extending from a right angle corner the SLS is half the length of the corresponding SAT, and only for parallel lines are they identical. We could therefore argue that the local support for the SLS tends to be "richer" than that of the SAT. This could be a minor factor in the robustness of the SLS relative to the SAT.

The SAT requires that the bi-tangent circles contributing to the skeleton must be contained within the shape contour. In the SLS this condition is relaxed requiring



For the points A and B the SLS places the skeletal point at C rather than D.

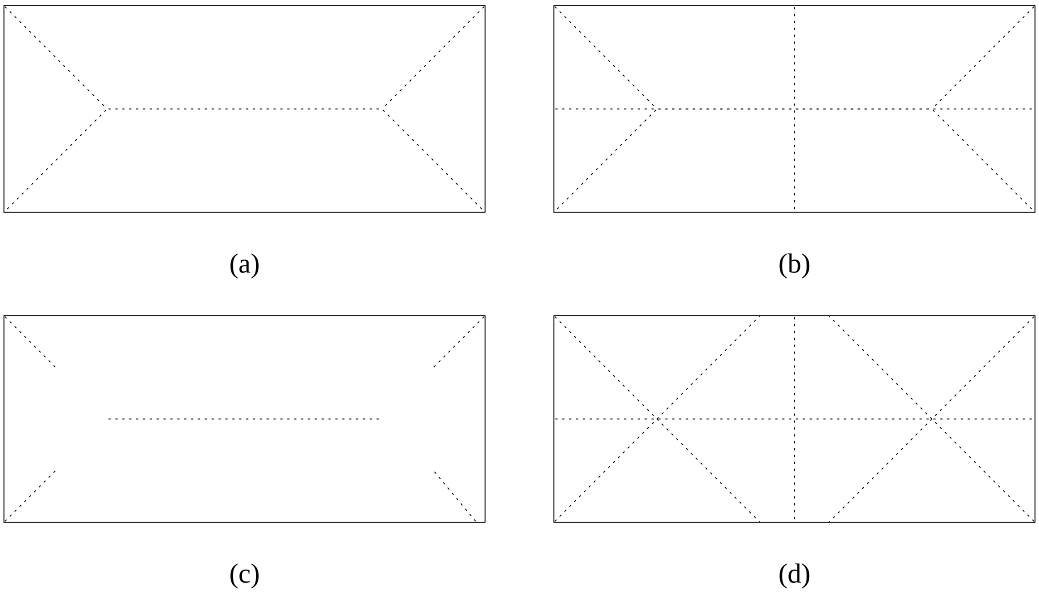
Figure 2.2: Skeletal point position of a smoothed local symmetry



The fact that the SLS places the skeletal point at the mid-point of the chord joining the boundary points means that the skeletal point stays close to the boundary points which form it. The SAT skeletal point varies its distance greatly depending on the relative angle of the boundary normals.

Figure 2.3: Comparison of skeletal point position of SLS and SAT

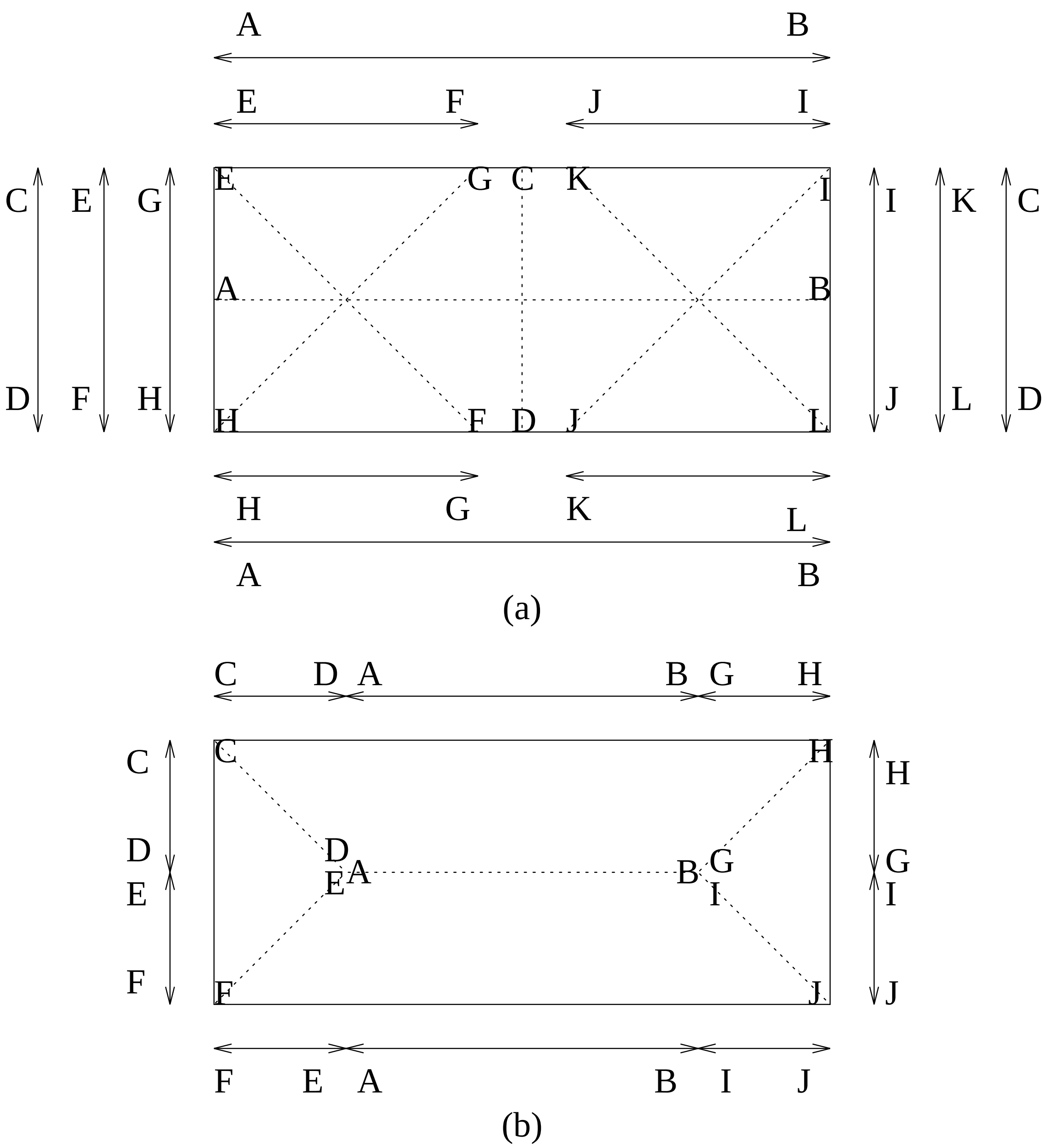
only that inward normals from the points of tangency contribute to the SLS. Of course this means that more symmetry points are generated for the SLS than for the SAT. When making comparisons between the two we must keep in mind the following point. Because the SAT marks skeletal points at the centre of the bi-tangent circle and the SLS marks them at the midpoint of the chord between the points of tangency a direct comparison of the skeletons is not possible, although of course we could convert one representation to another as in figure 2.4. In this thesis, representations are converted such that the skeletal point is marked at the centre of the circle. This is because distance transform approaches to skeletonisation (see below) discard the information regarding which boundary points contributed to the skeletal point.



The skeletons of the SAT and SLS cannot be directly compared although they can be converted into the equivalent form. Figure 2.4a shows the SAT of a rectangle and figure 2.4b shows the SLS. Figure 2.4c shows the SAT skeletal points converted to SLS form, *i.e.* moved from the centre of bi-tangent circles to the mid-point of chords joining the bi-tangent points. Figure 2.4d shows the SLS skeletal points moved to the centre of their bi-tangent circles.

Figure 2.4: Transformation between SAT and SLS representations

However, a deeper understanding is gained if we realise that, leaving aside the special cases of multi-tangent circles associated with circular arcs and skeletal branch junctions, the main job that these descriptors do is to "pair up" edge points. In figure



Skeletonisation can be seen as a process of pairing off edge points. Figure 2.5a shows the SLS of a rectangle and figure 2.5b shows the SAT of the same rectangle (dotted lines). Edge segments labeled with the same letters have been paired off, their corresponding skeletal branches are also marked. It can be seen that the SLS captures many more symmetries of the rectangle than does the SAT.

Figure 2.5: Pairing of edge points in SAT and SLS

2.5 we see what edge points are paired with each other for the case of a rectangle. In particular the SLS pairs both short and long sides along their entire length while the SAT pairs off only a section of the long sides and none of the short sides. This means the SLS captures both horizontal and vertical global symmetries, a point which Brady argues gives it more intuitive and perceptual appeal than the SAT [21]. The SLS also captures the diagonal symmetries in their entirety whereas the SAT only makes explicit half of their length.

The root cause of these deficiencies of the SAT relative to the SLS is the constraint that the bi-tangent circle must lie within the shape boundary. This means symmetry pairs are sometimes excluded from the skeleton by the proximity of other edge points in a way which can seem counter-intuitive. The SLS on the other hand effectively "decouples" the effect of other edgels from affecting the contribution of a particular symmetry pair to the skeleton. Note this also means that in the SAT an edge point can only be associated with one bi-tangent or multi-tangent circle of a single radius whereas in the SLS an edge point can, and often does, contribute to several of different radii.

This has an important consequence if the skeletons are used as shape descriptors within a complete object recognition system. The SAT will be greatly affected by occlusion, missing data or additional detail. The SLS is affected to a lesser extent, registering the relationship of new artifacts with the rest of the shape but only deleting those sections of skeleton where the supporting boundary has been erased.

Although the SLS imposes little constraint on the bi-tangent circles allowed to contribute to the skeleton the concept of "subsumption" is introduced to impose a hierarchy on the skeletal branches. Each skeletal branch is associated with the boundary points which contribute to it and the region between these points and the branch itself. These boundary points and subregion constitute the "cover" of the branch. Sometimes the "cover" of one branch is entirely contained within that of another branch and it is then said to have been "subsumed" by the branch with the larger cover. For example the diagonal branches of the SLS for a rectangle are "subsumed" by the main global axes and these are therefore accorded more importance.

2.2.4 Process inferring symmetry analysis (PISA)

Many shape descriptors such as the SAT and SLS are inspired by the idea of biological growth. In process-inferring symmetry analysis (PISA) as proposed by Leyton [59] this becomes an even more central component. PISA attempts to describe a shape in terms of processes which might have formed it. This approach is outlined for two specific situations: firstly, it is used to deduce the processes which form a single shape and secondly, given two shapes, to work out the intermediate processes which transform one into the other. These successions of shape changes are called a *process-history* and the list of possible deformations together with the rules for their combination are termed a *process-grammar*.

This inference of process-history is achieved by linking the concepts of curvature extrema to symmetry axes and, in turn, symmetry axes to the shape deformation processes themselves. The relationship between curvature extrema and symmetry axes is captured by the symmetry curvature duality theorem described in [58] and extended to the 3D case in [103]. This states that any segment of a smooth planar curve bounded by two consecutive curvature extrema of the same type, *i.e.* maxima or minima, has a unique skeletal branch which terminates at an extremum of the opposite type. The skeletal point for the bi-tangent circle is taken to be $\mathbf{Q}$, the midpoint of the arc $\mathbf{AB}$ in figure 2.6. Once the skeletal branches are defined they are linked to the deformation processes by what Leyton calls the "interaction principle". This states that the symmetry axes should be interpreted as the principal directions along which the processes act. Adding this to the curvature symmetry duality theorem means that "each curvature extremum implies a process whose trace is the unique symmetry axis associated with and terminating at that extremum" [59].

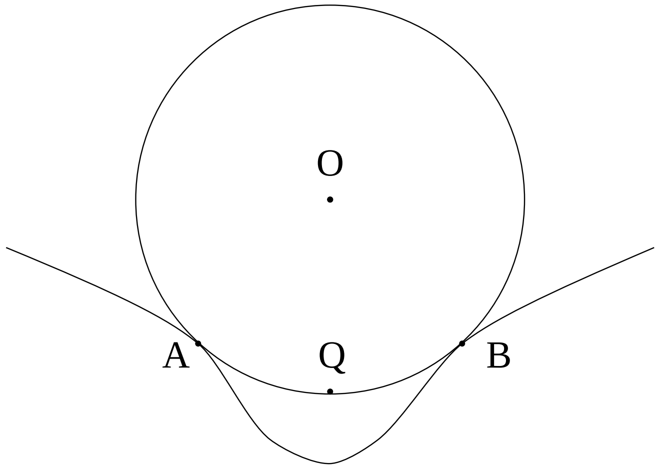
Leyton identifies four extrema which relate intuitively to specific shape deformation processes. These are given below along with their semantic meaning as a deformation process:

1. positive curvature maximum: protrusion

2. negative curvature minimum: indentation

3. positive curvature minimum: squashing

4. negative curvature maximum: internal resistance



In the PISA representation the skeletal point is drawn at the midpoint Q of the arc AB rather than at the centre O of the bi-tangent circle as in the SAT.

Figure 2.6: Definition of PISA skeletal point

The *process-grammar* is derived from the idea of the continuation and bifurcation of these semantic processes. For example the continuation of a protrusion or an indentation is trivial as these stay as they are; a squashing process on the other hand continues until it transforms into an indentation bounded by two points of zero curvature. Using this grammar a kind of directed graph can be built up where the nodes are shapes with different numbers and combinations of extrema and the branches between nodes represent the deformation processes which transform one shape into another. This graph, which Leyton calls “shape space” shows surprising structure. The task of inferring the “process history” from one shape to another is now reduced to finding a path through this graph from one corresponding node to another. As paths are often not unique, certain heuristics are required to choose between alternatives.

Leyton compares PISA to the SAT and SLS shape representations. It is important to realise that they have different goals, despite being closely related in their use of bi-tangent circles. The central goal of PISA is *process inference* whereas the SAT and SLS aim to make explicit the symmetry axes. This is why PISA defines the skeletal point as the midpoint of the circular arc as this guarantees a skeletal branch which is finite in length, unbroken and unidirectional. The choice of the arc midpoint also ensures that PISA has the remarkable property of being “extremum sensitive” in that the trace of the skeletal point approaches protrusions internally and squashed sections externally. Also, as this midpoint touches the boundary, it can be seen as the true record of the deformation process.

2.2.5 The symmetry set (SS)

A review of shape description based on bi-tangent circles would not be complete without reference to the symmetry set. The symmetry set (unlike the SAT, SLS and PISA), was not devised with a particular practical end use in mind such as shape description or process inference. The symmetry set was inspired by the SAT, of which it can be considered a superset, as one of many applications of singularity theory to differential geometry [27].

The symmetry set of a plane curve is the locus of centres c of circles bi-tangent to the curve [42]. Note that the condition for the SAT, that the circle must be internal to the curve, is relaxed so many more points are included. A similarly extended form of the SLS is introduced and called the midpoint locus [42] and in [67] Morris introduces a generalisation of the symmetry set called the rotational symmetry set (not to be confused with Fleck's rotational symmetries). In a series of interesting papers Giblin and his colleagues investigate in great detail the behaviour of the symmetry set and its relatives [42, 28, 11, 26]. Their main contribution is to outline the local structure that the symmetry set can have and in particular the various transformations which it can undergo and under what conditions these occur. Morris introduces an algorithm to compute the symmetry set based on a numerical technique of tracking around the contour while monitoring the curvature and its derivatives. He does point out that this requires the computation of up to the third derivative of curvature and so may not be robust for real data although it has been used in a robotics application [13].

The point must be made that, although the symmetry set captures more completely the local symmetries of a shape than the SAT, much of the extra detail of the symmetry set does not appear of much perceptual relevance and on the contrary appears to produce irrelevant and distracting information. (A similar state of affairs exists for the midpoint locus versus the SLS.) The main culprits for the production of these perceptually irrelevant axes are symmetries between inward and outward normals (no distinction is made between axes which are internal and external with respect to the boundary), and other symmetries which we will discuss in chapter 7.

2.3 Other skeletons with computational level descriptions

One of the few other skeletal representations with a clear computational level description, but which is not based on on the bi-tangent circle or sphere, is the generalised cones descriptor introduced by Binford [12] to model 3D objects. This consists of a curve or spine along which a 2D shape called the *cross section* is swept at a constant angle to the axis called the *eccentricity*. As this happens the cross section can be deformed according to a *sweeping rule*, the whole process sweeping out a volume which is the 3D shape of interest. It is particularly suited to the modelling of growing structures such as animals or plants and man-made objects created by extrusion or surfaces of revolution as by a lathe or potter's wheel. Its suitability as a shape descriptor is discussed at length in [20]; although desirable in many ways one of its main problems is that the decomposition of a shape into generalised cones is often not unique. Despite this it has been used as the basis for the 3D object recognition system ACRONYM [23].

Brooks introduced the variant of 2D projections of generalised cones which he called *generalised ribbons* [24]. A comparison of this representation with Blum's SAT and Brady's SLS from the point of view of generation and recovery is given by Rosenfeld [82]. Much of the analysis imposes restrictions on the ribbons, *e.g.* the cross section is normal to the spine, ends are ignored and some spines are straight. This paper confirms that the recovery of Brooks's ribbons can be difficult.

2.4 Algorithmic level classifications

One way to classify skeletonisation algorithms is at the algorithmic level, *i.e.* by the method used to compute the skeleton. There are in general four methods used for skeletonisation and these are as follows:

1. thinning
2. wave propagation
3. distance transformations
4. analytical techniques

2.4.1 Thinning

Thinning is the process of iterative removal of layers of pixels from an initial shape in the way we might peel an onion. Pixels are removed or kept according to rules dependent on the structure of pixel neighbourhoods. This process is repeated until no more removable pixels are left and only the skeleton remains. A recent survey of thinning algorithms is given in [55]. Thinning algorithms can be divided into those where the pixel removal has to be carried out sequentially and those where it can proceed in parallel, opening up the possibility of faster processing on parallel hardware. An important point is that many thinning algorithms can guarantee the connectivity of the final skeleton by recourse to the concept of the *crossing number* which is the number of times there is a transition from the foreground to the background in the traversal of the 3x3 neighbourhood of a pixel. A general problem with thinning algorithms is that they only operate on small pixel neighbourhoods (typically 3x3) and therefore cannot take account of larger scale phenomena. This means that noise spurs are easily introduced and invariance under rotation cannot be assured.

2.4.2 Diffusion

Perhaps one of the most intuitively appealing methods of skeletonisation involves the idea of wave propagation. This goes back to the early work of Blum on the symmetric axis transform [15, 16]. Blum coined the analogy of this process as a grass fire. Imagine the shape to be a dry grass field, all points on the boundary to be set on fire simultaneously and the fire to march uniformly along inward normals; at the point where one firefront meets another they are quenched and the union of all these quench points is the skeleton.

In the early sixties microcomputers were not available for Blum to implement this approach as an algorithm but he did experiment with optical and other implementations of the idea. Montanari introduced an algorithm which simulated the wavefronts analytically [66] after polygon approximation of the boundary and Xia [100] presented an algorithm to simulate propagation and extinction of fire-fronts for binary images.

Scott, Turner and Zisserman use a mixed wave/diffusion process [86]. They point out that a discrete element simulation of a simple wave process exhibits numerical instability leading to anisotropic propagation, *i.e* the wave propagates unevenly in different directions. Applying a *diffusion* process clears up the numerical instability by removing high frequency displacement. The diffusion process also has the desirable effect of broadening the wave as it travels outwards making symmetries with large radii more tolerant to scatter. Their implementation is based on convolution of the edge map with a set of "annulus" masks of different radii. Although substantial improvements to the basic wave propagation model have been made results are perhaps a little disappointing: even with perfect simulated data the axes seem to be noisy and incomplete.

Leymarie and Levine [57] simulate the grass fire wavefront using an active contour model [53]. Active contour modelling is part of a more general approach called "visual modelling" [91]. Visual modelling exploits computational physics to produce "physics-based" models which not only include geometry but also physical quantities such as force, energy and internal strain energies. These quantities can be used to initialise the models and to determine their evolution over time. Active contour models are linear, flexible entities, often given the descriptive name "snakes". These can be attached to, and track, linear features in a given data set, *e.g.* edges in an image. The essence of the technique is to drive the model by defining some form of energy function which is to be minimised.

In the context of skeletonisation Leymarie and Levine use the inverted Euclidean distance map of a given binary shape as a surface down which the contours slide until they rest in the "local minima" which correspond to the skeleton. This process is combined with boundary information by first locating significant features such as positive extrema of curvature and pinning the snakes to the boundary at these points. This ensures that the snake does not fall entirely into the well created by the inverted distance transform but leaves twin strands draping from these points thus defining the skeletal branches which touch the shape boundary. Skeletons produced by the technique are guaranteed to be connected as this is implicit in the contour model. The use of a continuous model bypasses some of the discontinuity problems of other distance transform based techniques caused by the nature of the discrete grid. This means a graph of the skeleton can easily be built up. The active contour model can produce dynamic skeletons from objects with slowly deforming boundaries such as single cells under a microscope.

In addition to the main idea of grass fire simulation by active contour, Leymarie and

Levine also address the problem of small scale perturbations of the boundary causing the formation of skeletal branches protruding to insignificant features. They introduce the concept of ridge support based on the notion of local deformation of the distance transform by small features. This can be measured using the slope amplitude of a tangent to the symmetric axis ridges of the distance transform. A minor criticism is that Leymarie and Levine fail to address the issue that in theory for a smooth curve the skeletal branch would not touch the boundary but would terminate at the cusp of the evolute. The evolute is the locus of radii of curvature (see chapter 5).

Wave propagation can be used to compute skeletal representations other than the Symmetric Axis Transform if the waves are not quenched as they meet but are allowed to continue propagation and interact with other wave fronts. Brady and Scott suggest the propagation of messages down normals [22] and devised the simulation of a massively parallel algorithm for the connection machine [47]. In this thesis this algorithm is modified for a MIMD type architecture. It should be noted that wavefronts propagated in such a fashion, as parallels to the original curve, undergo certain transformations which do not occur for the more familiar form synonymous with propagation via a membrane for example. In particular they do not obey Huygens' principle [8]. These singularities are examined briefly in [27] and more thoroughly in [25, 6].

2.4.3 Distance transformation

A distance transformation takes a shape and for each pixel within the shape computes its distance to the nearest point on the boundary. We can think of this as a three-dimensional surface where distance from the nearest point on the boundary is encoded as height above the plane. The skeleton corresponds to the creases or so called "local maxima" in this representation. Distance transformation algorithms vary as to the metric they use to calculate the distance to the boundary and whether they are sequential or parallel [19]. For the most accurate and rotationally invariant skeletons we must use the Euclidean metric [32, 78], but less accurate metrics exist which can be computed more quickly. These include, in order of increasing accuracy, city block, chess board and various so called "chamfer" metrics. See [19, 56] for comparisons and quotations of accuracy.

In his early work Blum did not use the term *distance transformation* but he does refer to a static 3D model of the wave propagation process which is a kind of distance transform [16]. The first algorithm based on this principle is due to Rosenfeld and Pfaltz [83] which used the city block metric and a simplistic ridge finder. More recently Arcelli and Sannitti de Baja [3, 5] have produced a series of algorithms using a distance transform to structure the hunt for skeletal pixels. The transform is also used to choose extra pixels for inclusion to ensure connectivity of the final skeleton. It should be stressed that these algorithms do not necessarily produce the SAT and can easily produce noise spurs.

An interesting new development is that due to Leymarie and Levine [57]. This is a hybrid between wave propagation and distance transform approaches to skeletonisation. Although very promising this algorithm shares the limitation with all approaches based on a distance transform, in that they can compute the symmetric axis transform but not the more general descriptors such as smoothed local symmetries or the symmetry set. This is because by definition they compute the distance to the *nearest* point on the boundary and disallow symmetries formed with more distant boundary points.

2.4.4 Analytic techniques

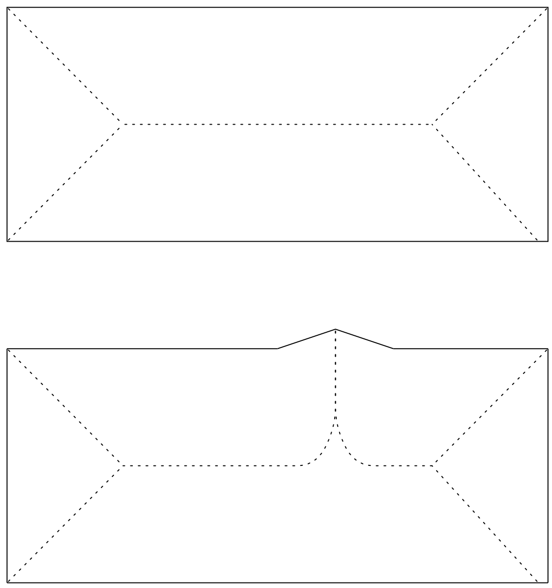
In this approach the skeleton is computed analytically after some initial model fitting stage. Brady and Asada compute smoothed local symmetries analytically after approximating the boundary with linear and circular elements [21]. More recently Saint-Marc and Medioni have used quadratic B-splines for the same purpose [85]. They fit the B-spline to the data using a least squares fitting method, they then use the fact that the spline can be represented as a series of conic sections, and simple algebra then gives another conic as the axis of symmetry between them. It is unfortunate that this is much more difficult for cubic splines with their desirable higher order continuity constraints and increased flexibility, and no analytical implementation is yet known.

2.5 Skeletonisation: problems and solutions

Having introduced the various kinds of skeletal representations and methods for their computation we now try to summarise the problems which they face and some of the general techniques which are used to tackle them.

2.5.1 Problem sources

The main problem is instability in the representation, whereby small changes to the input can produce large changes in the output which can show up in many circumstances. The classic example is shown in figure 2.7, where a small bump on the side of a rectangle has produced an extra skeletal branch.



One of the main problems of skeletonisation is instability. A small bump added to this rectangle produces a large extra branch to the skeleton.

Figure 2.7: Instability in a skeletal description

Quantisation error

The most fundamental cause of such small perturbations to a boundary is quantisation error. Most skeletonisation algorithms consider as their input a binary image. For such an image the concept of rotation is troublesome, since the pattern is only invariant for rotations of $\frac{\pi}{2}$ (for a square grid) and is undefined otherwise. For any other rotation, and subpixel translations, we need to introduce the concept of an object in continuous space which is subjected to an image formation and digitisation process. We can imagine digitising the continuous object at every conceivable location and orientation relative to the quantizing grid, which would build up a set of similar but not identical binary patterns. If the goal of skeletonisation is as a shape descriptor within a complete vision system then we require that for all of these inputs the algorithm gives us, at least qualitatively, the same output. This is something which the majority of skeletonisation algorithms ignore entirely.

Problems caused by quantisation are manifested in many other ways: algorithms based on distance transforms can give poor results if non-Euclidean metrics are used [19] and even the Euclidean algorithms are propagated from discrete boundaries. Wave based approaches are hindered by anisotropic propagation caused by the regular grid [86, 61]. This anisotropy induced by the regular grid is also present in thinning algorithms which only consider a small 3x3 neighbourhood of pixels [55]. Techniques reliant on the computation of boundary normals can fail because the normals are not dense enough and do not mesh with normals from an opposing boundary [95].

Even when all these previous problems are set aside it is not a trivial matter to apply the concepts of a skeleton in continuous space to the discrete case. See [88, 89] for a discussion of these issues.

Sensor noise

Not only is the continuous object subjected to an image formation and digitisation process but this latter process also includes sensor noise. This means that even for a single position and orientation of the object a range of different binary inputs to the skeletonisation algorithm is likely. Models for such processes exist [18, 79, 45]. Commercial CCD cameras produce quite clean signals with a typical signal to noise ratio of 40dB [48], but sensor noise can still be a problem in low contrast situations such as edges internal to the bounding contour of the object.

Preprocessing deficiencies

Many skeletonisation algorithms do not consider their role in the context of computer vision at all and those which do often make unrealistic assumptions about the quality of input they can expect from preprocessing stages. Basically they assume that the figure/ground segmentation problem has been entirely solved and the preprocessing stages deliver to them complete, uncorrupted and separate boundaries for each object. These assumptions are invalid except for synthetic images or carefully structured real environments. The segmentation process, which takes us from an input image to object outline, is still an unsolved problem [71]. Models and assumptions used by preprocessing algorithms can be quite crude, *e.g.* physical object edges do not always map to intensity step edges in images, and of course occlusion will be commonplace in any general real world application.

2.5.2 Attempted solutions

Postprocess thresholding

This approach assumes that the skeletonisation process is going to produce some irrelevant or spurious output and then attempts to clear this up after the event. The simplest criterion is to use branch length whereby all skeletal branches below a certain threshold length are pruned from the description [4]. This can mean that long branches to a small bump on the boundary are kept while short branches representing a large part of the boundary are pruned. An early attempt to remedy this by a *weighting function* is given in [17]. A limiting ratio of boundary length to axis length is defined at each axis point and this can be summed along a branch as a measure of its importance, *i.e* the relative amount of boundary it has generated. This technique comes up against the difficulty of measuring the differential quantity of the ratio in a discrete domain. Another simple measure can be generated from the idea of a symmetric axis point as the centre of a bi-tangent circle. If we take the angle between the normals to the points of tangency we see that for a straight section they are 180 degrees and for a right angled corner 90 degrees. For a small bump this angle becomes very acute and can be thresholded to filter out such insignificant features. A very similar notion is used in [56], in which the skeletal points are considered not just to have x, y position but also a z value equivalent to the radius of the bi-tangent circle at that point (an idea going back to Blum [16]). The angle at which branches of this 3D "wireframe" skeleton subtend the xy plane is bounded between zero degrees (for a boundary consisting of parallel lines) and 45 degrees (assuming a unit scaling between horizontal and vertical axes). This angle can thus be used to threshold branches.

Smoothing and multi-resolution

An alternative to post processing is to attack the problem at an earlier stage by preparing the input in some way, and one obvious candidate for this is smoothing. Smoothing amounts to low pass filtering of the image signal, and since sensor noise is mostly confined to high frequencies the filter removes this noise. (Even quantisation error can be thought of as a kind of high frequency noise.) Smoothing can be as simple as filling in or deleting certain pixel configurations [4] or be more sophisticated techniques using Gaussian kernels [21]. A particularly important idea is to compute the skeleton at multiple resolutions [76, 36] according to some variable smoothing parameter. It will be found that skeletons at coarser resolutions are more stable than their finer level counterparts. Difficulties can still be encountered in tracking branches across scales or parsing this scale-space representation [76].

Continuous approximation

In what can be seen as an attempt to escape the difficulties of dealing with noisy discrete data many researchers have attempted to use analytic descriptions to approximate the data. This generally involves fitting the data with a particular type of curve and then finding the skeleton for this continuous entity. A problem with this approach is that we can trade one kind of instability for another. There will always be a limit to how well the model can fit the data, *e.g.* polygon approximation can be unwieldy when modelling smooth objects and give different results with small changes to the data. For more sophisticated models such as splines it is not a trivial matter to find the parameters which give the best fit [70]. An idea which has had a positive impact on a wide range of problems within computer vision is the concept of active contours. These have been applied to skeletonisation by Leymarie and Levine [57] as described above. The advantage of these approaches is that once we have our continuous representation we can perform operations on them which would be difficult to perform on the discrete data directly, *e.g differentiation.* In the case of active contours the persistence of these continuous entities can increase the robustness of such processes as tracking.

2.6 Summary

At the beginning of this chapter we noted the confusion in the literature between function and procedure in the definition of skeletal shape descriptors. Marr's framework is used to resolve this confusion by labelling descriptions as either being at the computational or algorithmic level. To study a skeletal shape descriptor it is important that we understand both its computational and algorithmic level descriptions and that these should be distinct.

There are many types of skeleton and many ways that we can obtain them from images. Of particular interest are the group based on the bi-tangent circle: the symmetric axis transform, smoothed local symmetries, process inferring symmetry analysis and the symmetry set. These were chosen because they have a clear goal in terms of the geometric relationship between the shape boundary and the skeleton. Algorithmic level descriptions of skeletons were also discussed. There are four main methods: thinning, wave propagation, distance transformation and analytical techniques.

The crucial problem with skeletal shape descriptors is their sensitivity to noise and we have outlined the main sources of this problem such as quantisation, sensor noise and unrealistic assumptions about the quality of images which can be delivered by front end processing. We have then shown some of the techniques that have been tried in the face of these problems. These include post skeletonisation pruning based on some measure of the perceptual significance of skeletal branches, smoothing and multi-scale approaches and finally recent promising techniques which attempt to overcome the problems of the discrete domain by recourse to continuous approximation such as B-splines and active contours.

Chapter 3

A filter based approach to skeletonisation

3.1 Introduction

In the literature survey we saw that there are many types of skeletal shape descriptor and many ways in which they can be computed. We found that of particular note are the family of shape descriptors based on bi-tangent circles which include the Symmetric Axis Transform (SAT)[16, 15], Smoothed Local Symmetries (SLS)[21], Process Inferring Symmetry Analysis (PISA)[59] and the Symmetry Set (SS)[42]. Of the many methods which can be used to generate these and other skeletal shape descriptors we found that most can be put into one of four categories:

1. Iterative erosion of the shape boundary, *i.e.* thinning [46].

2. Wave propagation from the boundary [15].

3. Detection of "local maxima" on a distance transform [3].

4. Analytical methods, for simple shapes, following some form of function approximation, *e.g.* polygon approximation or spline fitting [21, 85].

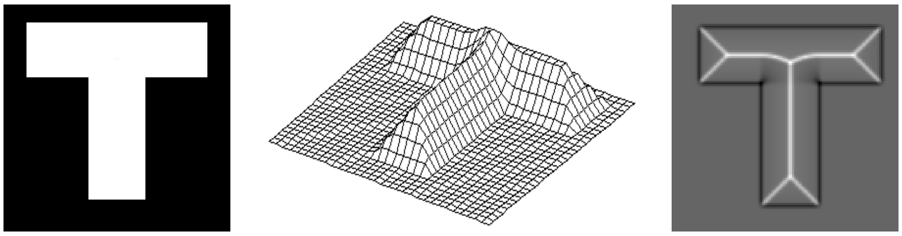
This chapter focuses on the Symmetric Axis Transform (SAT) which is a particular skeletal shape descriptor. Novel algorithms for its computation are presented based on the third approach, *i.e* the detection of "local maxima" on a distance transform. The format of the chapter is as follows; firstly, the key idea of regarding skeletonisation as a form of model-based feature extraction on the distance transform is introduced. This method has much in common with some edge detection schemes and this similarity is discussed. Two algorithms based on this approach are introduced. Firstly, an observation is made that the response of the human visual system to a grey level image of the Euclidean distance transform accentuates the "local maxima" of the transform. This leads to the extraction of skeletal features by convolution of the distance transform with the well known Marr–Hildreth operator [63] which is a model for this process. We then consider whether other operators can be envisaged to derive the skeleton by using those designed to detect instances of specific geometric models on the distance transform which correspond to the skeletal axis. This leads to the use of the Petrou filter [73] for detecting piecewise linear features in the presence of noise. This can be seen as a descendant of the filters devised by Canny for step edges and other profiles [29, 30]. Examples of these techniques are given which demonstrate that skeletons can be reliably extracted from a real image and the advantages and disadvantages of the technique are discussed.

It should be noted that other powerful methods exist for the extraction of skeletons, not least non-linear techniques such as morphological algorithms. Our approach can be seen as offering a different but complementary viewpoint. We cast skeletonisation in terms of a signal processing problem. This assumes an underlying signal (which could be continuous), which is to be extracted from noise and quantisation effects using techniques descended from Fourier analysis. This is in contrast to morphological methods which until now have started with discrete data and extracted the skeleton using methods based on set theory.

3.2 Skeletonisation as feature detection

It is well known that the "local maxima" [1] of the distance transform correspond to the skeletal points of the SAT [2]. This is demonstrated for a tee shape in figure 3.1a. The SAT of the shape shown in figure 3.1c can be seen to correspond to the "local maxima" of the distance transform shown in figure 3.1b. The main idea of this chapter is to think of the "local maxima" as features on the distance transform and to detect

[1] As shown in chapter 5 the term "local maxima" is incorrect and perhaps a better term is "singularities", however we will continue to use it here in respect of its wide acceptance in the literature.



(a) (b) (c)

Figure 3.1a shows a Tee shape. In figure 3.1b we see its distance transform, which assigns a value to each pixel within the shape equal to the minimum distance from that pixel to the shape boundary. The symmetric axis transform of a tee shape, figure 3.1c, corresponds to the "local maxima" of the distance transform. In the case of noisy images true maxima can be hard to locate and false ones can also be produced. The main idea presented in this chapter is to treat this as a signal processing problem where a filter is designed to detect instances of geometric features corresponding to the maxima which have been corrupted by noise.

Figure 3.1: The relationship of the SAT to the distance transform

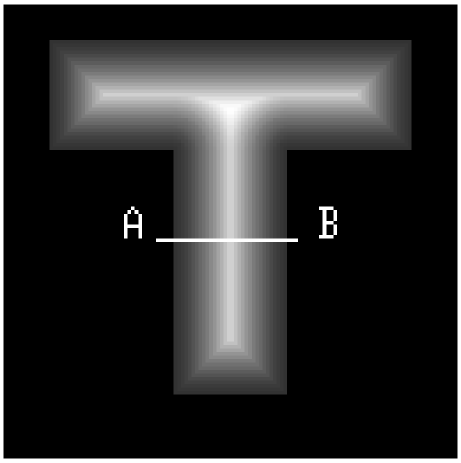
these features by convolution with a detection filter in much the same way as edges are often found in grey level images by convolution with an edge detection filter. In this work the term "model-based" is used, as in the literature on edge detection, to refer to a very simple geometric feature such as a step edge. This is in contrast to the more complex entities used to describe whole objects or classes of objects in complete "model-based" vision systems.

In edge detection image data can be thought of as a three dimensional surface where grey level intensity is encoded as height. On this surface edges or distinct boundaries between light and dark areas show up as discontinuities. The design of an edge detector often starts with a geometric model of the discontinuity to be detected, *e.g.* a step edge in Gaussian noise, and then continues with the design of an operator which when convolved with the image gives a high output when such a feature is encountered. Applying this idea to skeletonisation, we can think of the distance transform as a three dimensional surface where the height of the surface relates to the distance of the point to the nearest point on the boundary. Skeletonisation can now be thought of as a process of detecting features on this surface which correspond to the position of the skeleton. Early edge detection schemes used operators with small neighbourhoods such as the Roberts or Sobel operators [34]. These operators were poor at extracting edges from noisy data and were superseded by more sophisticated operators such as the Canny edge detector [29]. In searching for the "local maxima" on the distance map we could use a small simple operator but this might also be prone to noise in a similar way. It is therefore suggested that better results may be obtained in the presence of noise by using methods similar to those employed by the more sophisticated edge detection algorithms.

3.3 Skeletonisation by use of a receptive field model

This section presents the observation that the response of the human visual system to a grey level image of the Euclidean distance transform accentuates the "local maxima" of the transform. This leads to the extraction of skeletal features by convolution of the distance transform with the well known Marr–Hildreth operator [63] which is a model for this process. This method is illustrated by an example output from such a process.

3.3.1 Justification of receptive field model



When we look at a Euclidean distance transform encoded as a grey level image we can see the SAT directly. Consider the grey level intensity along the line AB. It appears to have a sharp central peak where in fact none is present. Some process in the visual system accentuates the grey level. If we can find a model for this process we could use it to perform skeletonisation.

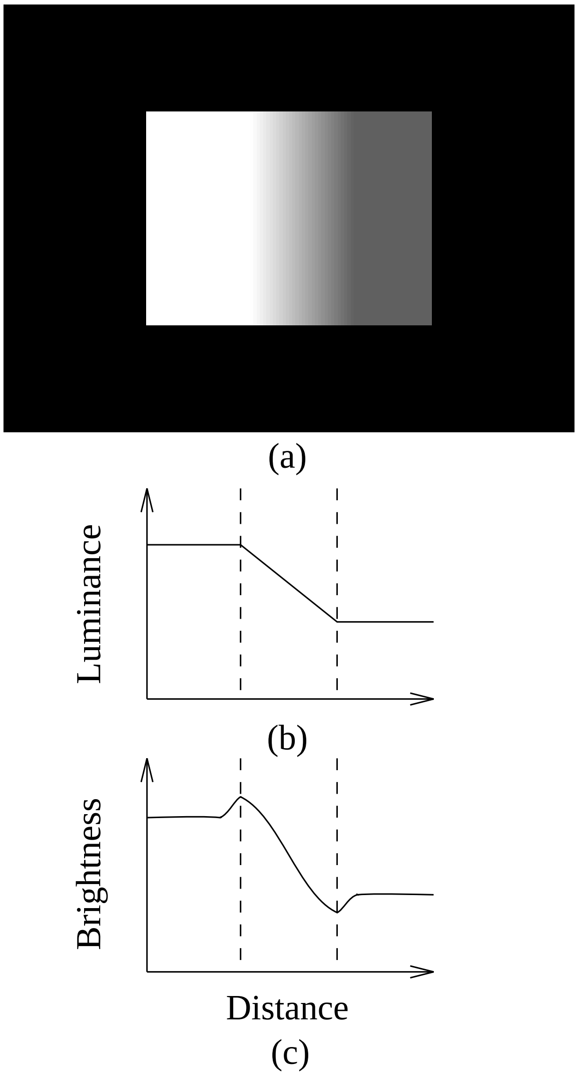
Figure 3.2: The Euclidean distance transform encoded as a grey level image

If the Euclidean distance transform is encoded as a grey level image, as in figure 3.2, an interesting effect is observed. The human visual system picks out the "local maxima" even though no large step discontinuities are present. A qualitative assessment of the image intensity, in a direction perpendicular to one of the major skeletal axes, such as along the line AB in figure 3.2, shows an apparently shallow uniform rise from the edge of the object, followed by a short intense peak in the middle, tailing off symmetrically on the other side. In actual fact there is no such central peak in the data but a uniform rise to a single maximum and then a uniform symmetrical decay. Some process in the human visual system accentuates the maximum to give rise to the subjective perception of a peak. If we can isolate this process we can use it to detect skeletal features directly.

Such a subjective overshoot or undershoot in perceived luminance is called a Mach band, and is often associated with luminance ramps of various kinds (see [38] for a review of Mach bands and other contrast related effects of the human visual system). An example of this phenomenon is given in figure 3.3. In figure 3.3a a black rectangle encloses a luminance ramp. This consists of a uniform light area on the left, another uniform dark area to the right and an intermediate area of linearly varying intensity, *i.e* a ramp. Two Mach bands appear as thin vertical bars at the borders of the ramp, a light one on the left and a dark one on the right. Figure 3.3b and figure 3.3c give a simple graphical description of the effect. Figure 3.3b shows the actual horizontal luminance profile to be two flat regions separated by a linear ramp whereas figure 3.3c shows that the perceived brightness has a distinct maximum and minimum at the ramp border. Although the commonly presented luminance profile is a simple ramp as in figure 3.3b other profiles can exhibit Mach bands, *e.g.* trapezoidal and triangular profiles [84, 39]. This is of relevance because the luminance profiles encountered on the grey level distance transform such as figure 3.2 are triangular.

3.3.2 An explanation of the Mach band effect

We now give a brief description of the physiological foundations of the Mach band effect and the mathematical models which have been put forward to account for it. Perhaps one of the most widely studied and well understood parts of the brain is the visual pathway, which extends from the retina in each eye to the primary visual cortex at the back of the head. Tracing the pathway back we first encounter the optic nerves which carry nerve fibres from the eyes. These meet at the optic chiasm where some cross



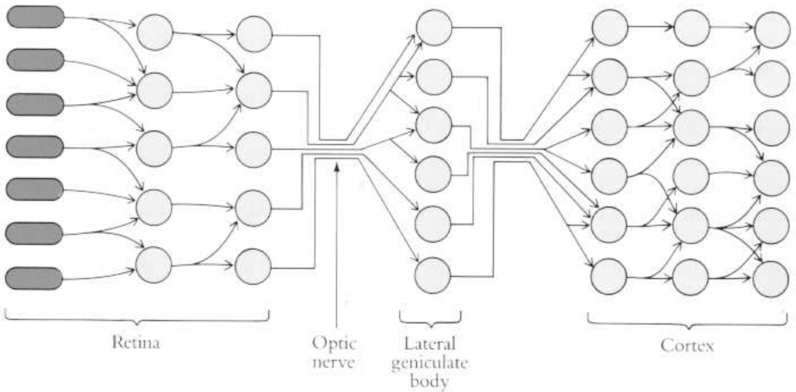
A Mach band is an overshoot or undershoot in perceived brightness with respect to the actual luminance and is associated with luminance ramps. Figure 3.3a is an image of a luminance ramp. The horizontal luminance profile of this image, figure 3.3b, consists of two flat sections connected by a linear ramp. However the perceived brightness, figure 3.3c, shows deviations at the top and bottom of the ramp. These appear as two vertical bars of light and dark in figure 3.3a.

Figure 3.3: The Mach band effect

over so fibres from the same side of the retina in each eye end up together. After the optic chiasm the pathway again divides and fibres then terminate at two small folded structures called the lateral geniculate nuclei. From here large numbers of fibres called the optic radiations project to the cortex itself.

If asked to consider the visual pathway more closely we may be forgiven for the conjecture that it would be an impossibly complex, random, "bramblescape" of interconnections. On the contrary research has found that there is rich and well defined structure throughout. The form this structure takes is that cells seem to be organised into a series of two dimensional sheets with connections between them. These sheets of cells have been described as *topographic maps* [50] because cells which are close in one layer project to cells which are close in the next layer so, although distortions may occur, the basic topology of the sheets is the same.

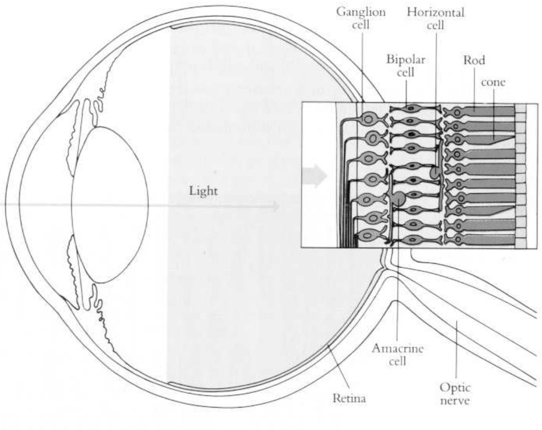
Figure 3.4 show this plate-like structure of the visual pathway in a schematic form. We can see that the first three layers are contained within the retina itself so, far from being a passive sensor such as a CCD camera which merely collects data, it actively processes the sensory information before passing it out along the optic nerve.



The initial stages of the visual system are arranged as plate-like sheets of cells, an organisation which is common to other parts of the central nervous system. The first three stages are found in the retina. (Figure courtesy of W.H.Freeman)

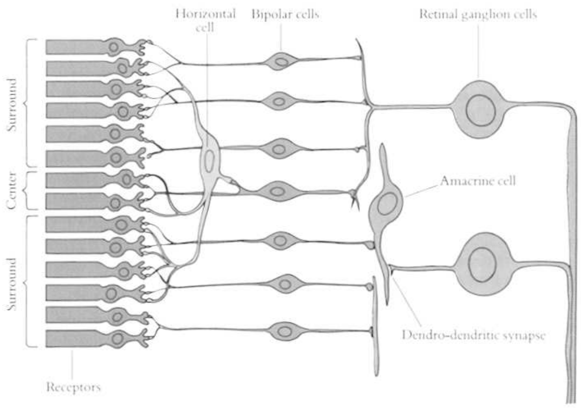
Figure 3.4: Initial stages of the mammalian visual system

Figure 3.5 shows a cross section of the eye with light entering from the lens on the



The retina lines the back of the eye ball and consists of three layers of cells, the receptors, bipolar cells and ganglion cells. There are also transversely connected cells between these layers. (Figure courtesy of W.H.Freeman)

Figure 3.5: The eye and retina



The cells of the retina are arranged into circular overlapping receptive fields. These consist of a central region, perhaps excitory, and a surrounding, perhaps inhibitory, region. This is achieved due to direct connection of the central region to the bipolar cells and indirect connection of the surround through the horizontal cells.(Figure courtesy of W.H.Freeman)

Figure 3.6: The concept of a receptive field

left and impinging on the retina which is a thin layer of cells on the back of the eye to the right. The insert of figure 3.5 shows the three layers of cells which make up the retina, the photoreceptors themselves which are called rods and cones, the bipolar cells and ganglion cells. It is interesting to note that the rods and cones face the back wall of the eye and light has to pass through the layer of other cells to reach them. Note also that there are cells between these layers connected transversely and it is these which are implicated in the Mach band effect.

Figure 3.6 shows a more detailed view of part of the retina. The receptors are arranged into circular groups called receptive fields of which one is shown here. We can identify two populations of receptors in terms of their effect on output of the ganglion cells they are connected to. Firstly there is a central region which can be excitory, *i.e.* if a spot of light shines on them they make the ganglion cell fire briskly, or inhibitory, *i.e* the light would inhibit the firing of the ganglion cell. The central region is surrounded by an annulus of cells of the opposite type, *i.e.* if the centre is excitory then the surround is inhibitory and *vice-versa*. This opposite centre surround response is achieved by virtue of direct connections of the central region to the bipolar cells whereas the surround is indirectly connected through the horizontal cells.

If we look at figure 3.7 we can see how such receptive fields could give rise to Mach bands. Figure 3.7a shows the sensitivity function of the receptive field measured along a line passing through its centre. Figure 3.7b shows a typical luminance profile of a ramp and figure 3.7c shows the channel response to the luminance profile which is obtained by convolving the sensitivity function with the luminance profile. (Compare with figure 3.3.)

Such an explanation is termed a single channel model as all the receptive fields are of the same type and size. It works well as a qualitative explanation of many contrast effects but performs less well as a quantitative predictor particularly when ramps of different scales are used. Also, even as a qualitative description it predicts that the largest Mach band effect should occur at step edges whereas in reality the Mach band effect is greatly reduced if present at all at step edges [38]. Such limitations led to the production of multi-channel models which assume that there are several sizes of receptive field; small channels would give a different response from large channels and the combination of these gives rise to the Mach band. This explains why steps (which have no implicit length scale) do not produce large Mach bands.

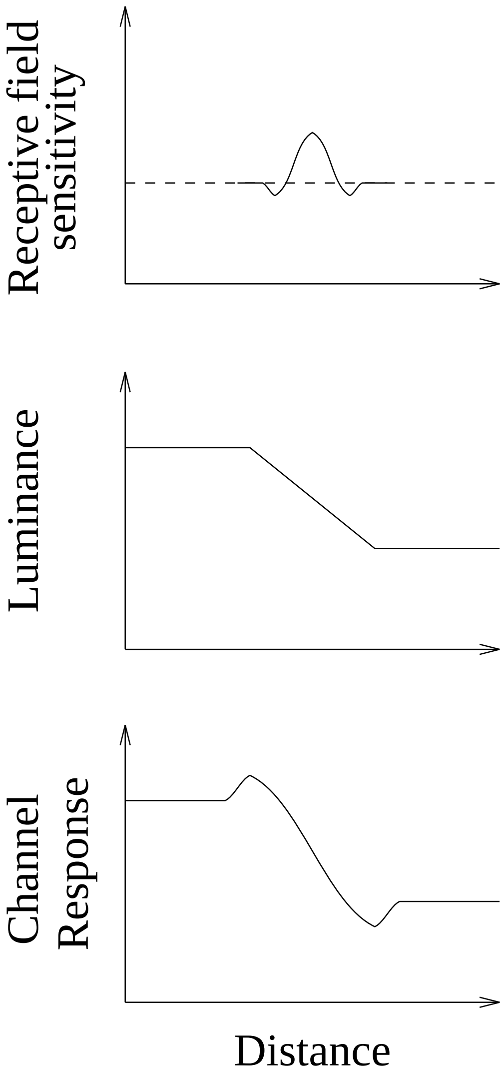


Figure 3.7a shows a cross section of the sensitivity function of a receptive field. This consists of a central excitory region surrounded by inhibitory lobes. Many similar overlapping receptive fields cover the visual field and together constitute a channel. The response of a channel to a luminance ramp such as figure 3.7b can be calculated by convolving the ramp with the sensitivity function and is shown in figure 3.7c (compare with figure 3.3c).

Figure 3.7: The role of receptive fields in the Mach band effect

Mathematical models of receptive fields

Several mathematical models have been proposed to account for the behaviour of receptive fields in the visual system. Here we describe three of the most popular; the Marr–Hildreth filter [63], the difference of Gaussians [80] and Gabor functions [33].

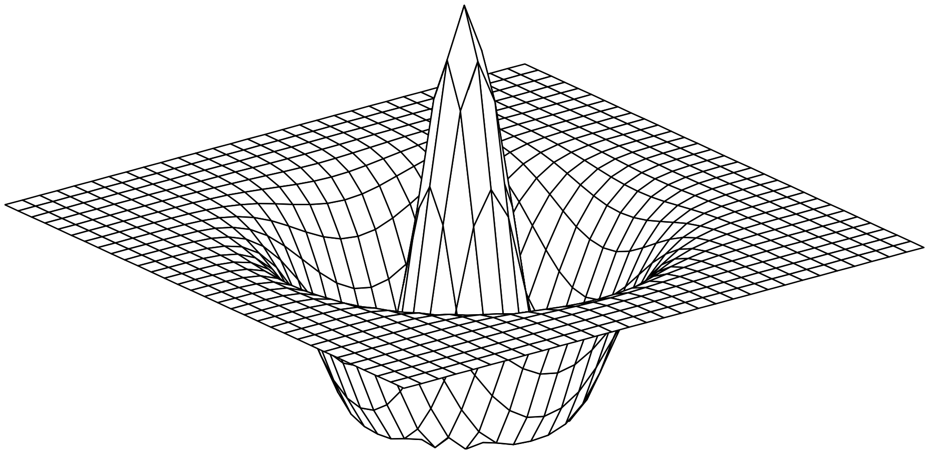


Figure 3.8: The Marr–Hildreth filter

Figure 3.8 shows the Marr–Hildreth filter [63] $\bigtriangledown^2 G$ where $\bigtriangledown^2$ is the Laplacian operator $\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ and G is a two dimensional Gaussian.

Figure 3.9 shows that the Marr–Hildreth filter (dotted line) can be closely approximated by a difference of two Gaussians, a narrow positive one and a wider negative one, shown here as a solid line. Rodieck [80] defined this filter as the function $f(x, y)$ where:

$$f(x, y) = g_1 \sigma_1^{-2} \pi^{-1}.exp(-(x^2 + y^2)/\sigma_1^{-2}) - g_2 \sigma_2^{-2} \pi^{-1}.exp(-(x^2 + y^2)/\sigma_2^{-2})$$

Finally figure 3.10 shows a Gabor function which has been used by Daugman [33] to model cortical receptive fields. A Gabor function is a sinusoid with a Gaussian envelope such as:

$$f_{cos}(x, y) = exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) cos\,(2\pi\omega x)$$

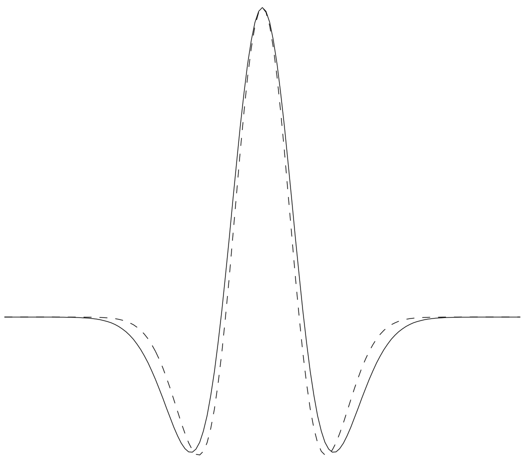


Figure 3.9: The difference of Gaussians filter

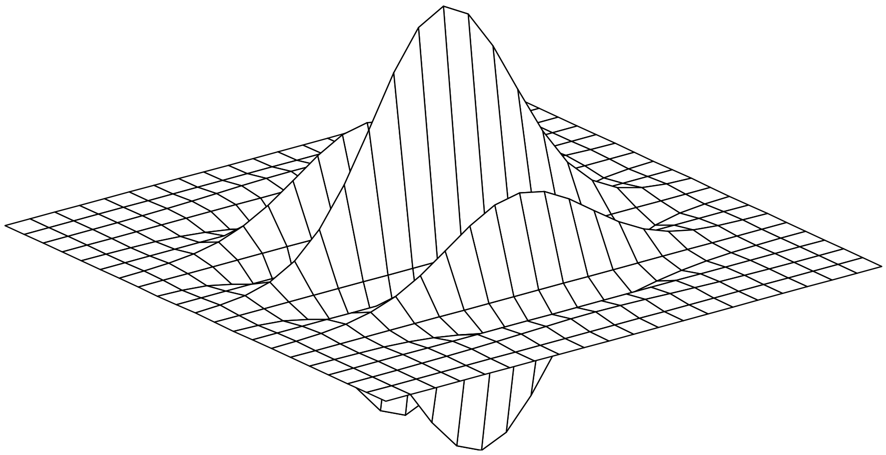


Figure 3.10: The Gabor filter

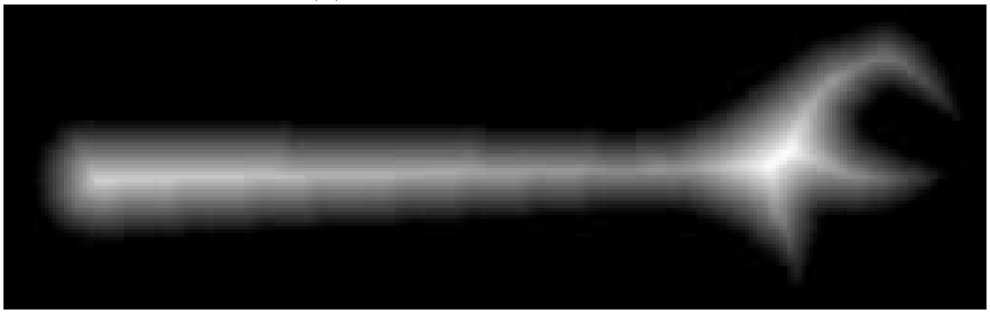
$$f_{sin}(x, y) = exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) sin\left(2\pi\omega x\right)$$

We use the Marr–Hildreth filter in this instance. If we have a Euclidean distance transform of a shape $E(x, y)$ then we convolve this with the Marr–Hildreth function thus $(\bigtriangledown^2 G) * E$.

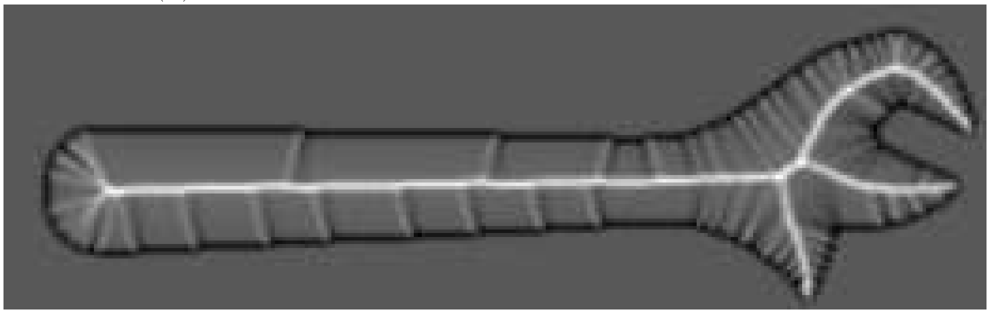
3.3.3 Output from the Marr–Hildreth operator

In figure 3.11 we can see the result of the convolution of the Euclidean distance transform with the Laplacian of a Gaussian. Figure 3.11a is the binary image of a wrench, figure 3.11b is its grey level Euclidean distance transform and figure 3.11c shows the actual output of the convolution of the distance transform with the Marr–Hildreth operator. We can see that the "local maxima" are indeed accentuated to a large degree verifying that the output from the Marr–Hildreth operator is in qualitative agreement with the subjective perception of the grey level Euclidean distance map. The major skeletal axes stand out as strong positive peaks in the output of the operator whereas discontinuities at the boundary give rise to a positive and negative peak with a central zero crossing.

(a) Binary image of a wrench



(b) Greylevel Euclidean distance transform of wrench



(c) Output from Marr–Hildreth operator

Figure 3.11a is a binary image of a wrench obtained from a back-lit greyscale image which was simply thresholded. In figure 3.11b we have applied the Euclidean distance transform due to Danielsson [32] to the binary image. When the distance transform is displayed as a greyscale image as in figure 3.11b the human visual system can pick out a central bright ridge corresponding to the Symmetric Axis Transform of the object. This illusion is due to the Mach band effect [38] which is caused by lateral inhibition in receptive fields of the visual system. One model for this process is the Marr–Hildreth operator which could be used to extract the skeleton; in figure 3.11c we see the result of the convolution of this operator with the Euclidean distance transform and can see that a skeleton is indeed picked out clearly.

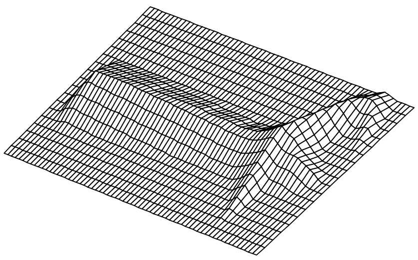
Figure 3.11: Filtering process applied to wrench image

3.4 Skeletonisation by detection of specific geometric features

Having introduced the idea of skeletonisation by convolution with a detection function we can now ask if any specific geometric feature on the distance transform exists which corresponds to the "local maxima". We could then use operators specifically designed to detect such a feature to find the skeleton. In this section we will suggest that a roof profile is a reasonable candidate for such a feature and show the result of convolving the distance transform with an operator specifically designed for roof detection.

3.4.1 Justification of roof profiles as geometric features

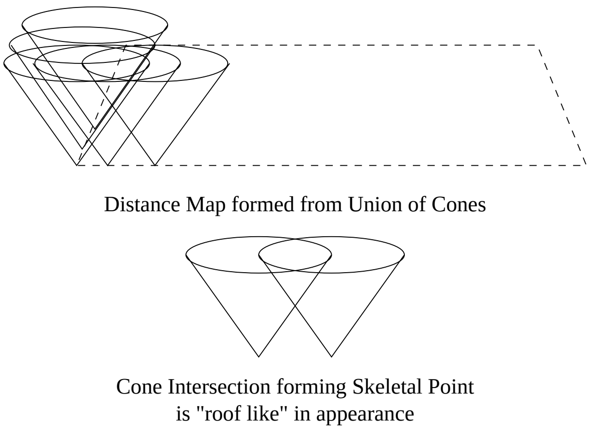
Convolution of the Euclidean distance transform with an edge detector such as a Canny [29] gives a poor response with little or no skeletal features registered in the output. One reason that any model-based detection function should give poor results is because the assumed model, in this case a step edge with Gaussian noise, is a poor approximation to the data. If we look at the distance transform as a 3D mesh as in figure 3.12 we can see that indeed no step edges are present. On further inspection we can see that the "local maxima" in this surface look like roof profiles and these would seem a good candidate for a specific geometric feature.



There are no step edges present on the Euclidean distance transform so a typical edge detector does not detect its discontinuities. A roof profile is a good candidate for a feature corresponding to the "local maxima".

Figure 3.12: Euclidean distance transform of a wrench as 3D mesh

Let us consider why roof profiles should be a common feature in the distance trans-



Blum suggested that each point on the boundary of a shape such as this rectangle can be thought of as having an inverted cone placed upon it. The union of these cones forms a surface which is in fact a distance transform. Skeletal points correspond to intersections of these cones and are thus guaranteed to be a roof profile.

Figure 3.13: Why a roof profile is a good candidate feature

form and further that they should correspond to “local maxima” and thus the SAT. In his early work Blum does not use the phrase “distance transform” but he does refer to a static 3D representation of the wave like “grass fire” process which can be used to compute the SAT. However, his 3D surface is indeed a type of distance transform where distance from the boundary is encoded as height above the image plane. Using this representation Blum tells us that in continuous space each point on the boundary can be thought of as having the apex of a right circular inverted cone placed upon it with an angle of 45 degrees between the axis and surface of the cone [15]. The distance transform can be thought of as the envelope of the surfaces of these cones as shown in figure 3.13. The contribution made by any cone to the surface of the Euclidean distance transform is that part of its surface which is not contained within any other cone. This will be a line drawn on the cone from its apex in a direction normal to the boundary tangent direction at the apex contact point to the first intersection of this line with any other cone. This intersection point will be a “local maximum” in the distance transform and a SAT point. The intersecting cone will have a similar line drawn up to the SAT point and the two lines together ensure that the surface of the distance transform near the SAT point has a qualitatively roof like profile.

3.4.2 Output of a roof detection operator

Now that we have a candidate for a geometric feature, *i.e.* a roof profile, we can use an operator specifically designed to detect such a feature. We do not have to design one ourselves as some already exist. Canny's work on edge detection also included the construction of a roof detector [29] and Petrou [73] has designed an operator to detect linear features including roof profiles.

In fact Petrou's work [74, 72, 75, 73] actually stems indirectly from that of Canny [29, 30]. Canny was one of the first to cast the design of an edge detection algorithm in terms of a set of well defined goals. These consisted firstly of a signal model of the type of feature we wish to detect and a noise model of the noise in which it is immersed. He then defined three quantitative performance criteria for the filter:

1. Good *detection*. There should be a low probability of failing to mark real edge points and a low probability of marking false edge points.

2. Good *localisation*. The edge points marked by the operator should be as close as possible to the centre of the true edge.

3. Multiple response constraint. Only one response to one edge.

Canny used the product of detection and localisation as a single criterion which was optimised subject to the multiple response constraint. The "Canny edge detector" in common use is not actually the filter designed through this process but the first derivative of a Gaussian which was suggested by Canny as a good approximation to the optimum filter and which has certain advantages in terms of implementation, *e.g. it is separable*, meaning that the filter can be split into vertical and horizontal components which can be computed simultaneously on parallel hardware. Those interested in an efficient implementation of the Canny edge detector are referred to Deriche [35] who uses a recursive implementation to perform filtering of any scale with the same low computational overhead.

Spacek [90] adopted the first two criteria of Canny's design but used a slightly different form of the third criterion which was combined with the other two to form a single performance criterion. This led to a simpler form of the differential equation from the solution of which the operator is derived. When Spacek compared his filter with Canny's he used just the product of localisation and detection as Canny had not published sufficient information for the complete performance criteria to be calculated [29], and at this time it appeared that the Spacek filter was slightly better. Ironically however, as Petrou points out, when more information was later available [30] the Spacek filter turned out to be slightly worse. One interesting point Spacek made is that a cubic spline can be used as an approximation to the optimal edge operator which is better in performance than the differentiated Gaussian suggested by Canny.

Petrou extended the work of Spacek to the case of ramp edges and other linear features which have an intrinsic length scale. Spacek held two of the six parameters of his design fixed whereas Petrou performed the optimisation correctly for all parameters.

Using a filter designed specifically to detect roof profiles picks out the skeleton very well. This also suppresses side branches to a greater extent than the Marr–Hildreth operator (compare with figure 3.11c).

Figure 3.14: Output from Petrou operator

In figure 3.14 we can see the result of convolving the Euclidean distance transform with the Petrou operator. Adopting Petrou's nomenclature [73], the value of the filter parameters used to produce this image were d = 0.5 and l = 4.0. To obtain the actual SAT we would need to perform post processing in a similar fashion to edge detectors such as non-maximal suppression and hysteresis thresholding but this is not done here. This produces a better defined skeleton than the Marr–Hildreth operator. Side branches are still present in the output but are suppressed to a much greater extent and are invisible in this figure.

3.4.3 Comparisons with other approaches

In this section we compare the output of the Petrou operator with two other standard skeletonisation algorithms over a number of images. The first algorithm is the skeletonisation algorithm used in the Khoros image processing environment [44]. The second detects maxima on the distance transform using a simple ridge finder. This ridge finding algorithm was the first algorithm to utilize the distance transform for skeletonisation [83]. In all images the parameters for the Petrou filter were $d = 1.0$ and $l = 4.0$.

The spanner of figure 3.15a, pair of pliers of figure 3.16a and hammer of figure 3.17a are all shapes which are thin and elongated. The Petrou operator gives a clean connected skeleton for each of these objects. The Khoros algorithm produces a good central skeleton with a few extraneous disconnected pixels but no spurs. The ridge finder gives a well connected central skeleton with a few gaps but is prone to short straight noise spurs. These noise spurs are particularly noticeable on the pliers in figure 3.16d and the hammer in figure 3.17d.

The file of figure 3.18a and the screwdriver of figure 3.19a have a linear axis of symmetry and some curved boundary segments. All three algorithms produce a clear central skeletal axis. The Petrou operator of figure 3.18b and figure 3.19b gives a skeleton with four pairs of symmetric branches which point to the curvature extrema on the boundary. The Khoros algorithm of figure 3.18c and figure 3.19c does very poorly at depicting the symmetric branches. The ridge finding algorithm of figure 3.18d and figure 3.19d gives a skeleton which has pairs of symmetric branches but also has noise spurs.

Figure 3.20a in an image of a spoon. The Petrou operator of figure 3.20b gives a skeleton with a single linear branch. The skeletons of the Khoros algorithm and the ridge finder are both disconnected. However, both algorithms produce forked skeletal branches at the tip of the spoon bowl which are absent from the Petrou skeleton.

A series of animal test images were used to evaluate the algorithms responses to irregular objects. These test images were a rabbit in figure 3.21a, a giraffe in figure 3.22a, a goat in figure 3.23a and an ostrich in figure 3.24a. The Petrou operator gives good skeletons for the rabbit and giraffe but displays some fragmentation for the goat and ostrich. The ridge finding algorithm gives a well connected skeleton for all animals except for a few short gaps, it does however display noise spurs which are particularly bad for the ostrich. The Khoros algorithm skeleton often consists of disconnected points, there are also large gaps present particularly for the goat and ostrich.

In general the Petrou operator produced skeletons which were connected and complete with few noise spurs. The ridge finding algorithm produced skeletons which were connected but with small gaps and many short linear noise spurs. The Khoros algorithm produced disconnected skeletons for irregular parts and did not seem to pick out fine detail as well as the other two algorithms. From the above results it is reasonable to conclude that that the Petrou operator produces better skeletons than the other two algorithms, in terms of their connectedness and completeness and the absence of extraneous skeletal points and branches.

3.5 Discussion

In this section we describe the advantages and disadvantages of using the approach of convolution with a detection function on the Euclidean distance transform as a method of skeletonisation.

3.5.1 Advantages

The advantages of this approach are as follows. Firstly major skeletal features which form the main axis of the shape give rise to a large response in the operator output, however smaller and perhaps less perceptively important skeletal branches give rise to a smaller output. This gives us the concept of ranking of skeletal elements by the magnitude of the operator output. This could be used to guide shape matching algorithms using the representation by making tentative matches using the major features before attempting further matching with the less significant features. The magnitude of the operator output could also be used to overcome the problem identified in multiresolution skeletal representations [76] that skeletal branches can disappear abruptly when the SAT is computed over a range of spatial scales. In this case the operator output would decrease markedly giving a warning of the imminent annihilation of the skeletal branch.

Secondly we could use the technique, often used in edge detection, of subpixel interpolation of edge position. If we import this idea to the domain of skeletonisation we can estimate the true skeletal position more accurately than a purely binary representation. This enables the algorithm to be less sensitive to the orientation of the object relative to the quantising grid and ensures a thin skeletal axis in all cases.

A third point, common to the literature on both edge detection and skeletonisation is the concept of a scale based representation. The smoothing parameter of the detection function can be varied and thus can be made to register the skeleton over a range of spatial scales. This helps in providing a natural hierarchical description of the shape and also helps to suppress the effects of noise. Promise of the technique in this regard is shown in figure 3.25. A test image of a Koch curve, a type of fractal with a boundary displaying regular and recursive structure over a range of spatial scales, is shown in figure 3.25a. By varying the smoothing parameter of the filter we can pick out structure at three distinct spatial scales. In this case we used the Marr–Hildreth operator with Gaussian variance 1, 30 and 100 in figures 3.25b, c and d respectively. Note that further post processing is still required to extract the final SAT associated with each scale.

A fourth point is that the convolution method is width independent and takes the same time no matter what the object of interest. This is also true for the preprocessing step of computing the Euclidean distance map which in its serial form consists of two sequential scans of the image [32].

3.5.2 Disadvantages

A number of problems still remain with the technique. Firstly although the profile of the distance transform is qualitatively a roof it is not always a pure linear roof profile in cross section as in the case of a boundary consisting of two parallel lines. The ridge of the roof can be curved and is only horizontal for the case of parallel boundaries, thus adding to complications in the operator output. The maximum response of the operator is not necessarily for roof features which are normal to the convolution direction. This is a problem in the detection of features other than step edges which have an intrinsic length scale as the feature appears broader to the operator.

The new filter based approach also shares an inherent limitation with all algorithms based on the Euclidean distance transform. Only the symmetric axis transform may be computed by this method and not more general symmetry based shape descriptors such as the smoothed local symmetries or the symmetry set.

The Matched filter

A simple *matched filter* would appear at first sight to be a good candidate for the detection of these “local maxima” in the grey level distance transform. If the noise has a constant power spectral density over the frequency range occupied by the signal then the improvement in signal to noise ratio is the best possible for any linear filter and for the case of Gaussian noise is the best possible of any filter, linear or non-linear [60]. It is interesting to ask therefore, that if this is the case, why are the “optimal” filters introduced in this chapter not of the same form as a matched filter, *i.e.* the time reverse of the signal we wish to detect? The answer to this is that the matched filter is optimal with respect to only *one* criterion, signal to noise ratio (which corresponds to the detection criterion), whereas the other filters also embody the concepts of localisation and multiple response. The matched filter also does not have to respect the boundary conditions of the optimisation process and the one it violates in particular it that there should be zero response to a constant input.

3.6 Summary

We have introduced the idea of skeletonisation by the convolution of the Euclidean distance transform [32, 78] with a detection filter, firstly by a receptive field model, the Marr–Hildreth operator [63], and then by the Petrou operator [73] which is a filter designed to detect a specific geometric feature, *i.e.* a roof profile. It is shown that improvements in noise immunity of the skeleton are possible using this technique and that it lends itself easily to a multi-resolution representation of the skeleton. Examples of the technique show that a skeleton can be reliably obtained from real data. Although a promising technique, problems still exist and in particular the fact that in general the “local maxima” of the Euclidean distance transform are not linear roof profiles with a horizontal ridge. Even if these problems can be rectified the technique still has the limitation inherent in all distance transform approaches that only the SAT can be computed and not the smooth local symmetries or the symmetry set.

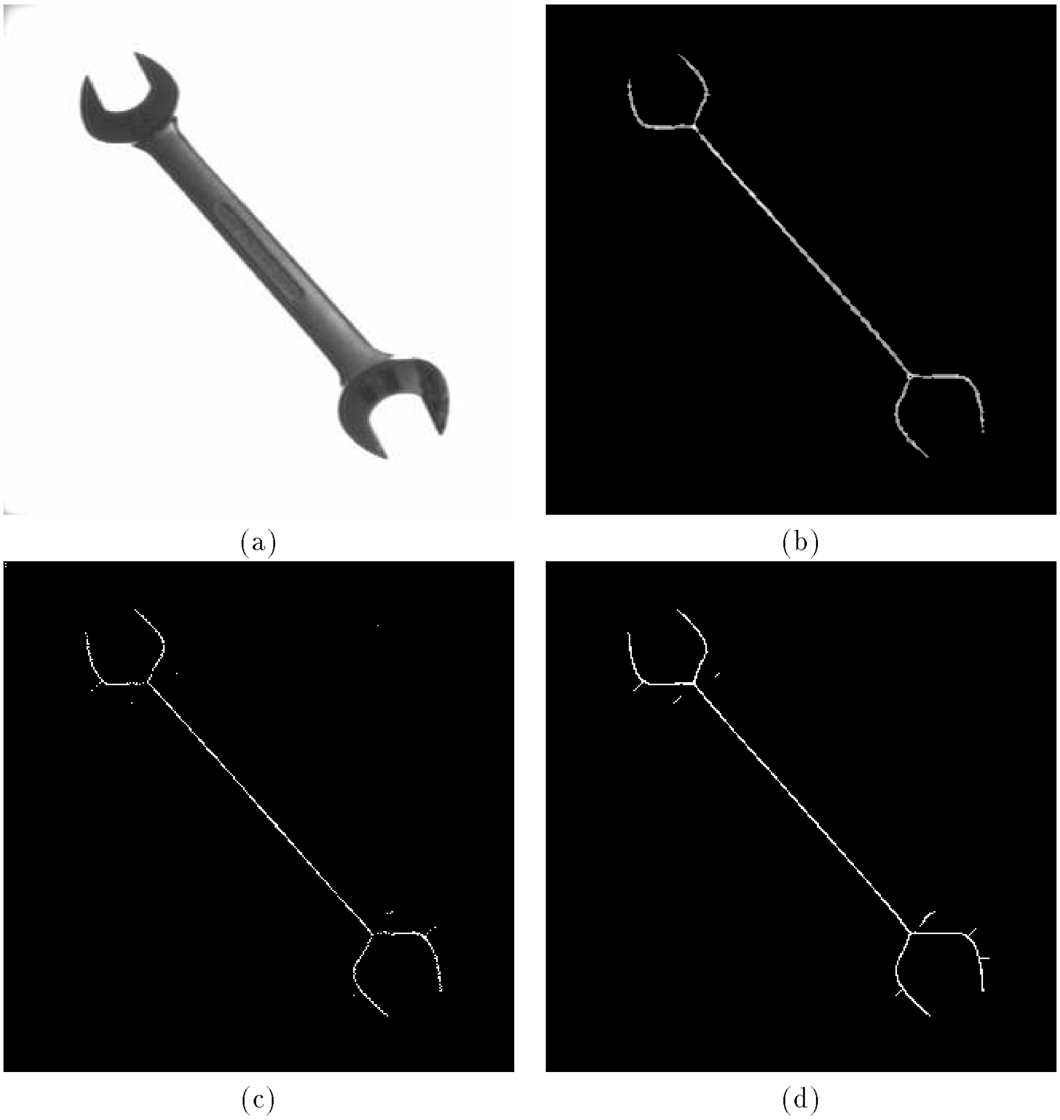
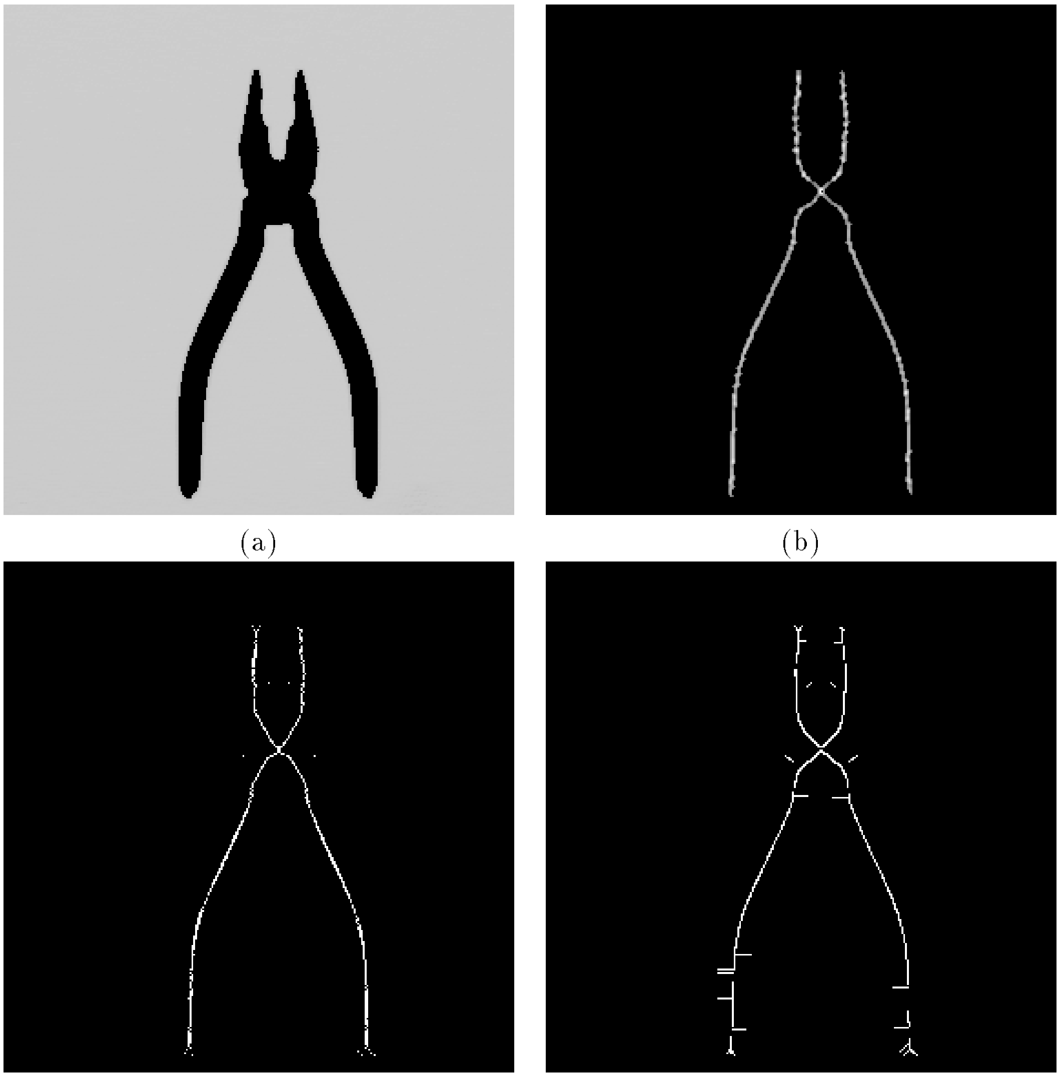


Figure 3.15a is the raw input image. Figure 3.15b is the output of the Petrou operator. Figure 3.15c is the output of the Khoros medial axis algorithm. Figure 3.15d is the output of a ridge finder on the distance transform.

Figure 3.15: Comparison of Petrou filter with other algorithms: Spanner



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

Figure 3.16a is the raw input image. Figure 3.16b is the output of the Petrou operator. Figure 3.16c is the output of the Khoros medial axis algorithm. Figure 3.16d is the output of a ridge finder on the distance transform.

Figure 3.16: Comparison of Petrou filter with other algorithms: Pliers

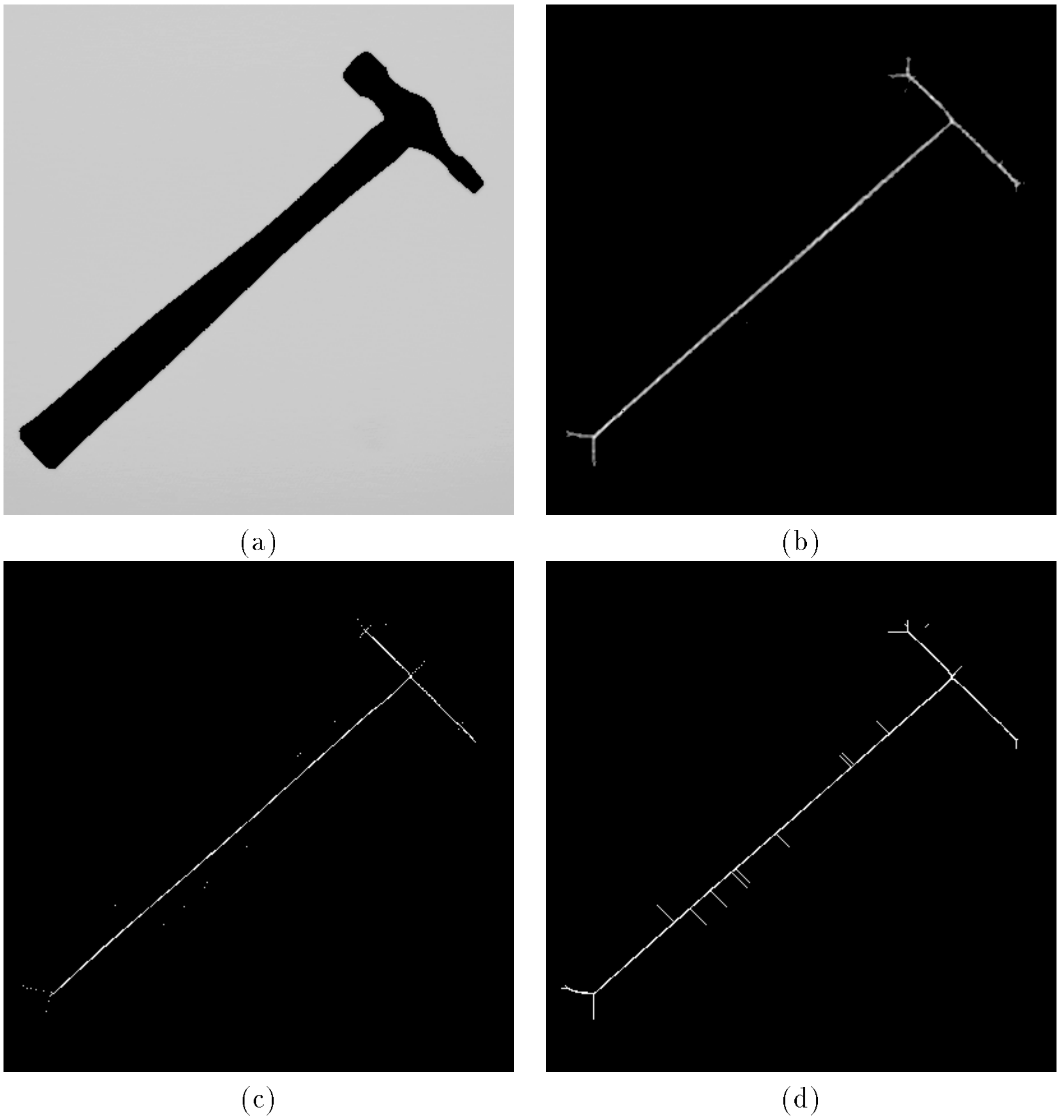
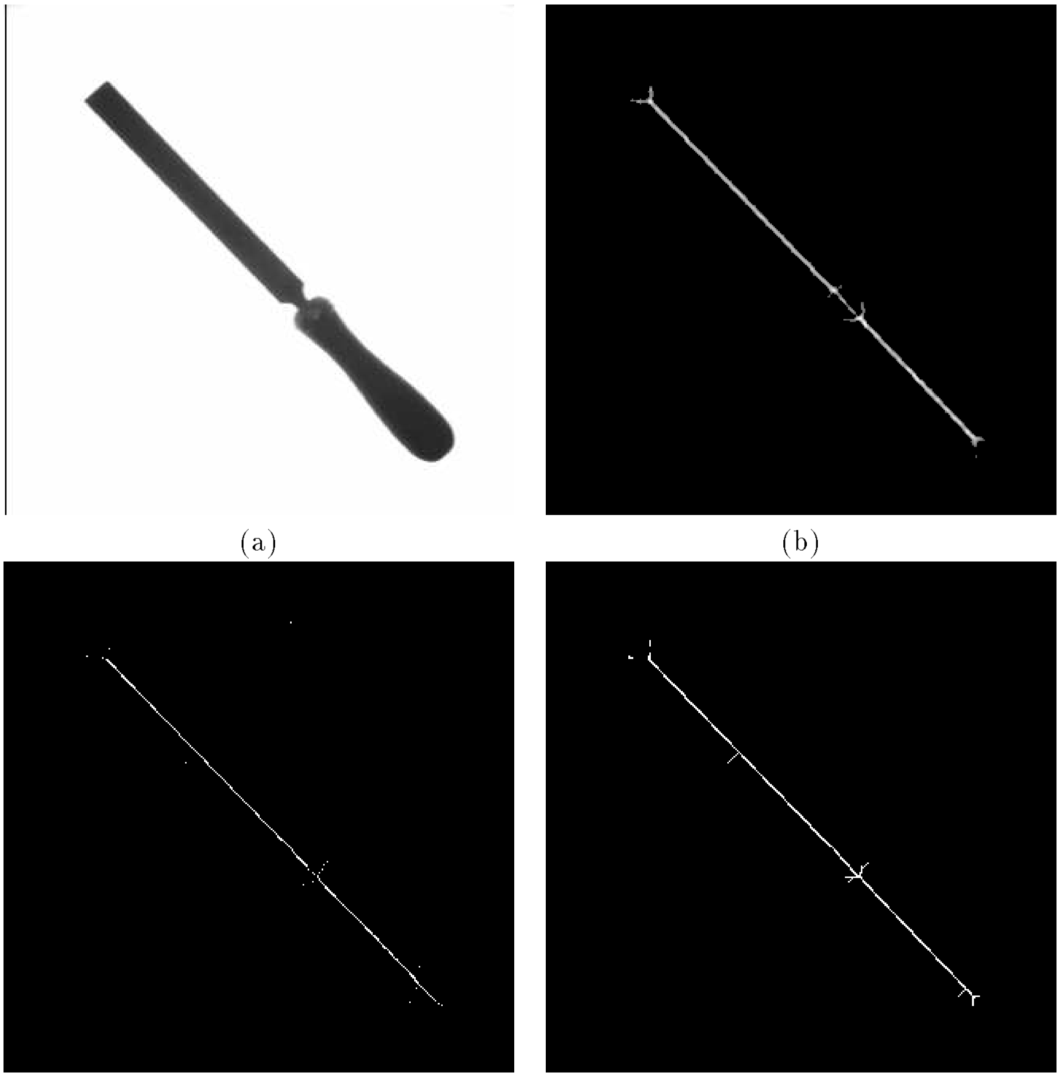


Figure 3.17a is the raw input image. Figure 3.17b is the output of the Petrou operator. Figure 3.17c is the output of the Khoros medial axis algorithm. Figure 3.17d is the output of a ridge finder on the distance transform.

Figure 3.17: Comparison of Petrou filter with other algorithms: Hammer

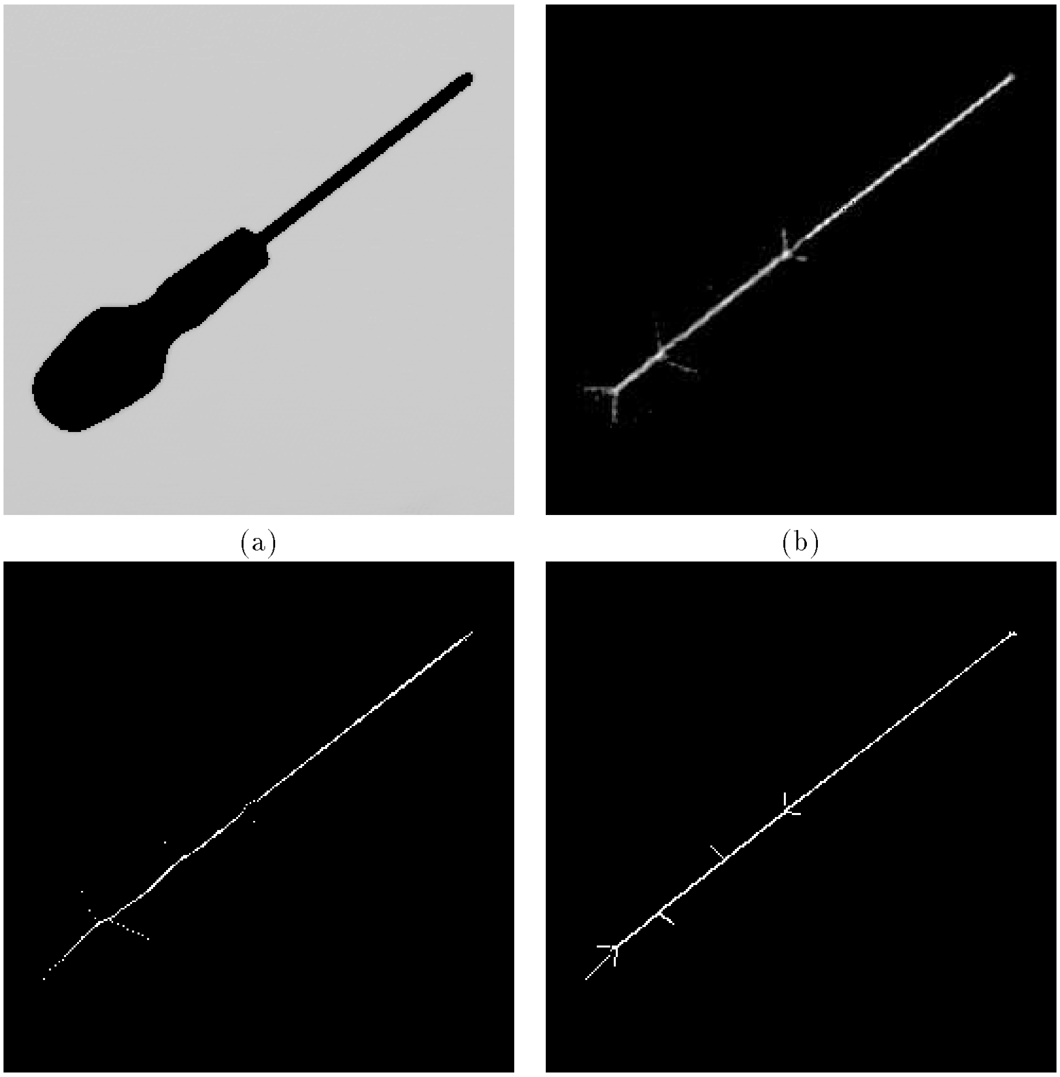


(a) (b)

(c) (d)

Figure 3.18a is the raw input image. Figure 3.18b is the output of the Petrou operator. Figure 3.18c is the output of the Khoros medial axis algorithm. Figure 3.18d is the output of a ridge finder on the distance transform.

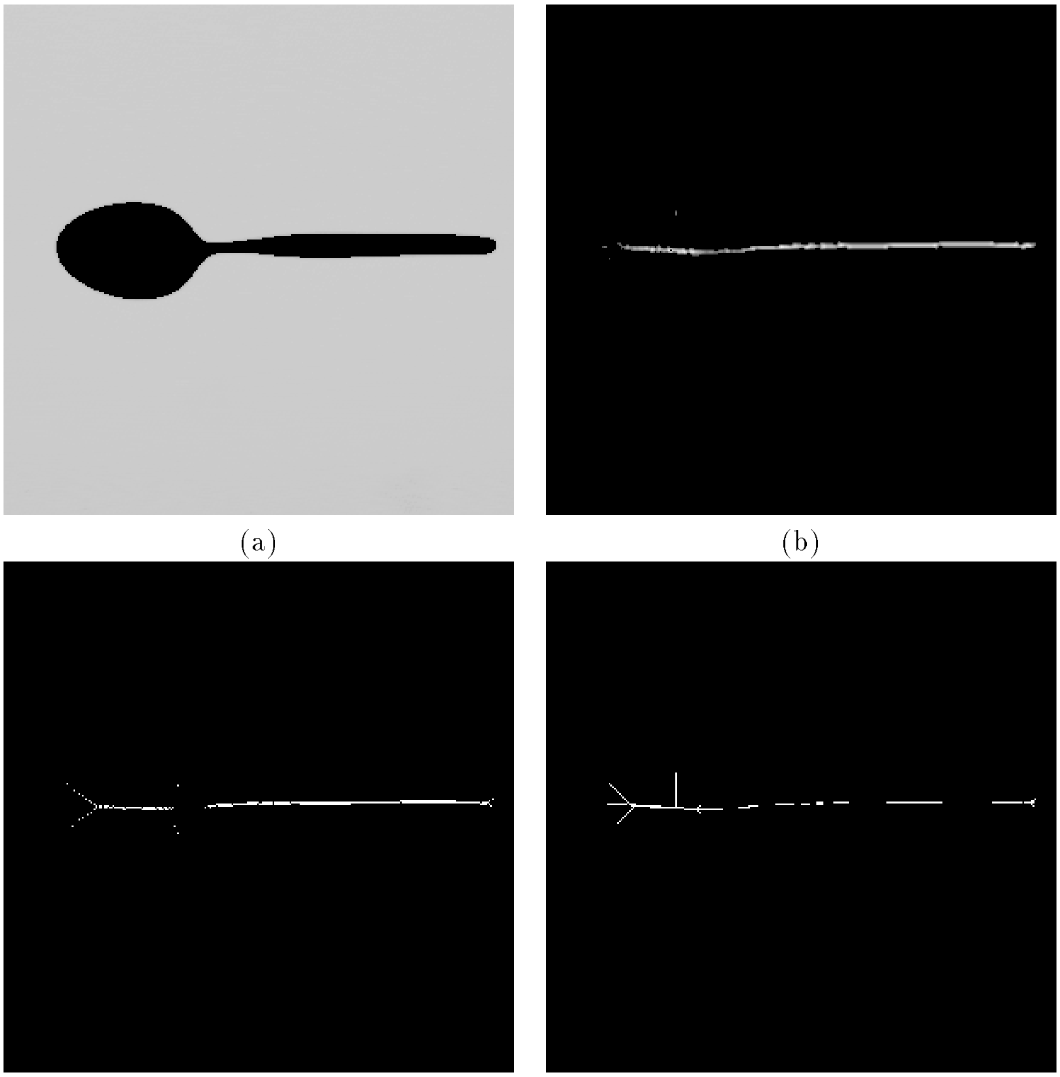
Figure 3.18: Comparison of Petrou filter with other algorithms: File



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

Figure 3.19a is the raw input image. Figure 3.19b is the output of the Petrou operator. Figure 3.19c is the output of the Khoros medial axis algorithm. Figure 3.19d is the output of a ridge finder on the distance transform.

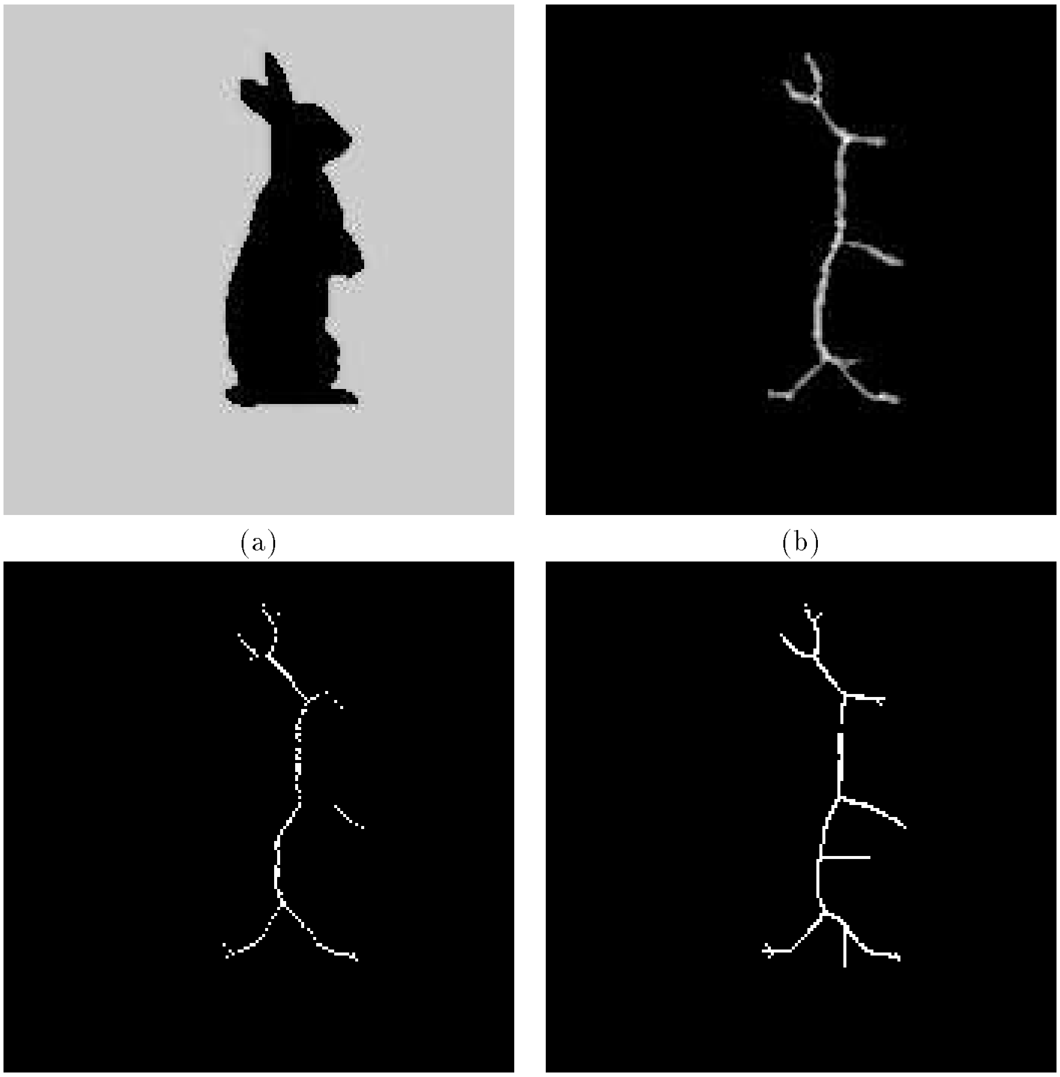
Figure 3.19: Comparison of Petrou filter with other algorithms: Screwdriver



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

Figure 3.20a is the raw input image. Figure 3.20b is the output of the Petrou operator. Figure 3.20c is the output of the Khoros medial axis algorithm. Figure 3.20d is the output of a ridge finder on the distance transform.

Figure 3.20: Comparison of Petrou filter with other algorithms: Spoon



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

Figure 3.21a is the raw input image. Figure 3.21b is the output of the Petrou operator. Figure 3.21c is the output of the Khoros medial axis algorithm. Figure 3.21d is the output of a ridge finder on the distance transform.

Figure 3.21: Comparison of Petrou filter with other algorithms: Rabbit

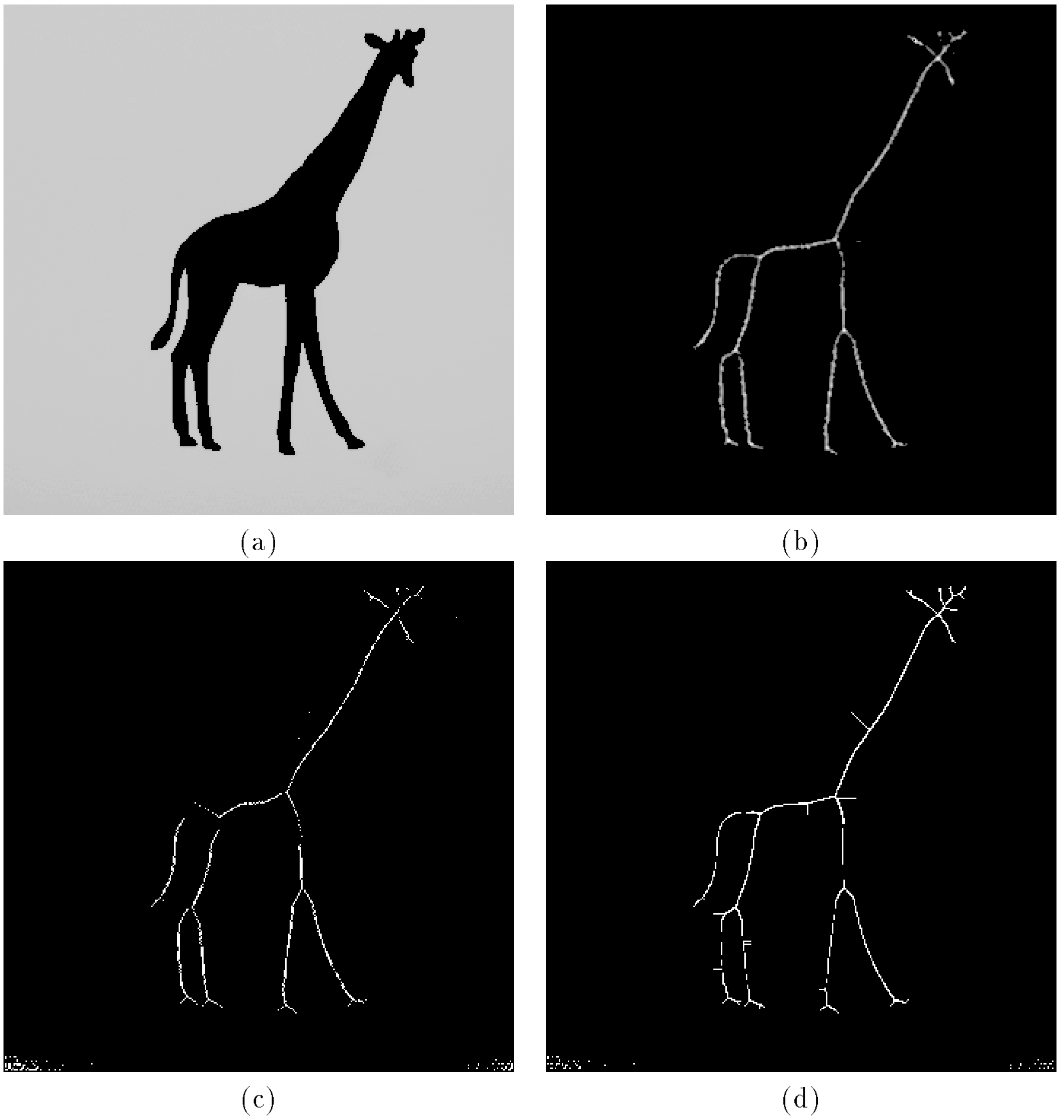
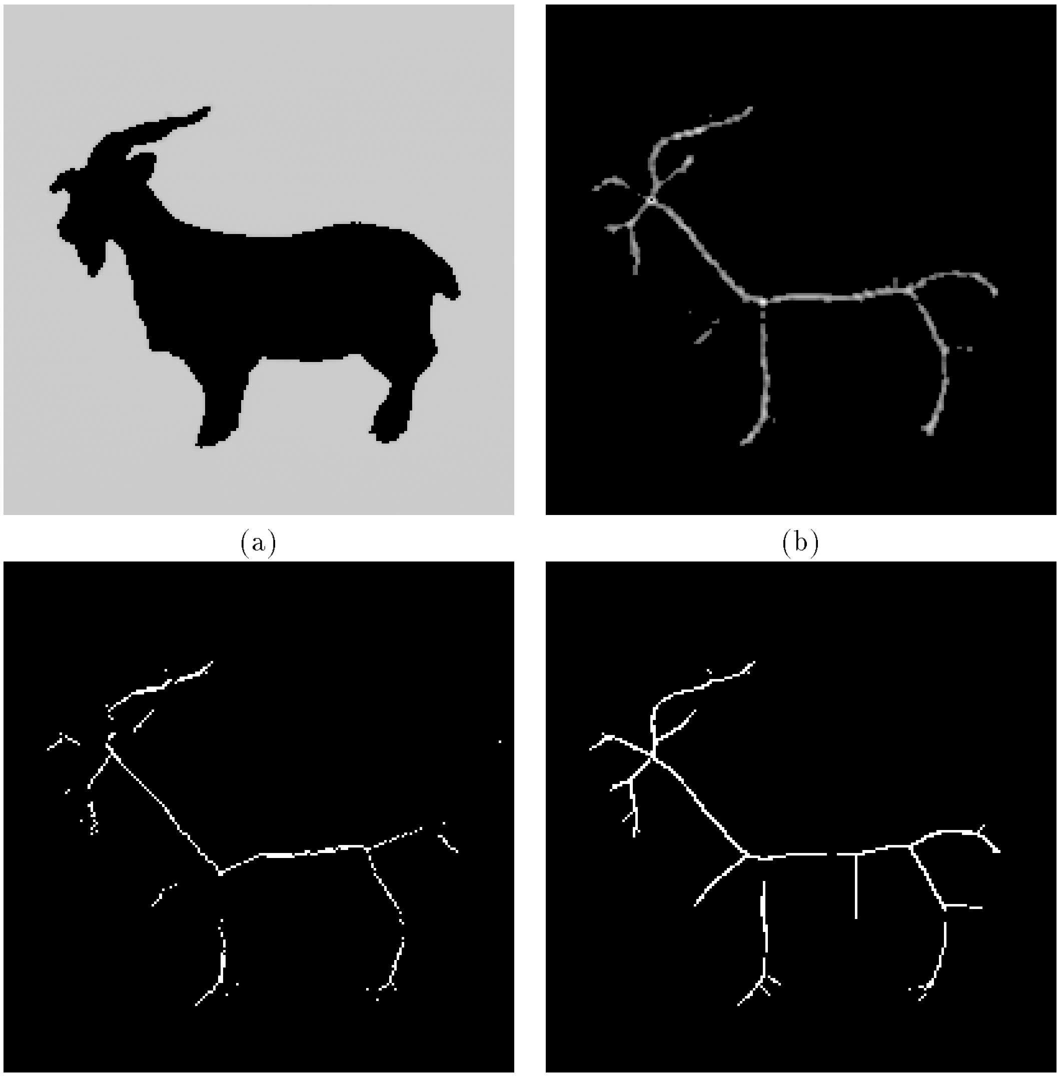


Figure 3.22a is the raw input image. Figure 3.22b is the output of the Petrou operator. Figure 3.22c is the output of the Khoros medial axis algorithm. Figure 3.22d is the output of a ridge finder on the distance transform.

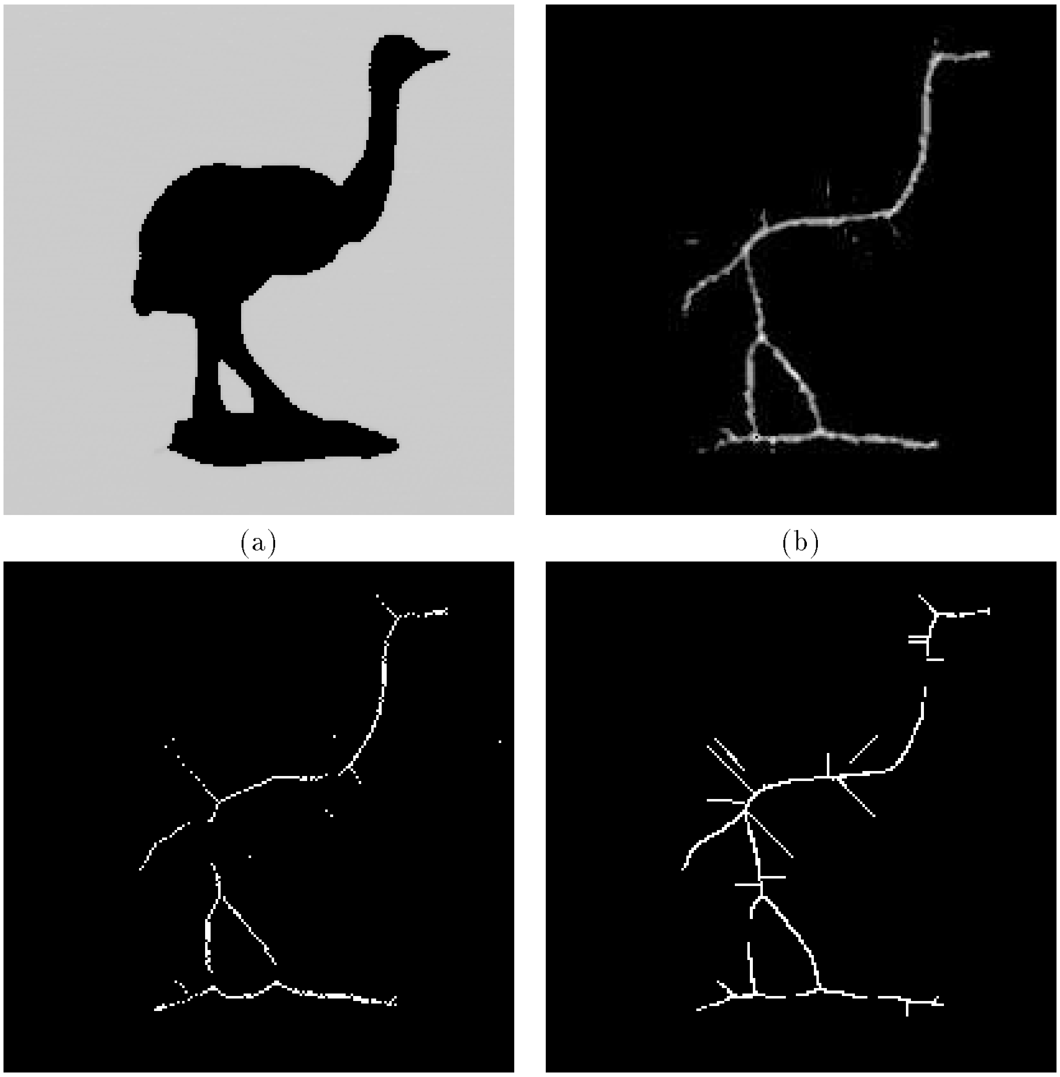
Figure 3.22: Comparison of Petrou filter with other algorithms: Giraffe



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

Figure 3.23a is the raw input image. Figure 3.23b is the output of the Petrou operator. Figure 3.23c is the output of the Khoros medial axis algorithm. Figure 3.23d is the output of a ridge finder on the distance transform.

Figure 3.23: Comparison of Petrou filter with other algorithms: Goat



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

Figure 3.24a is the raw input image. Figure 3.24b is the output of the Petrou operator. Figure 3.24c is the output of the Khoros medial axis algorithm. Figure 3.24d is the output of a ridge finder on the distance transform.

Figure 3.24: Comparison of Petrou filter with other algorithms: Ostrich

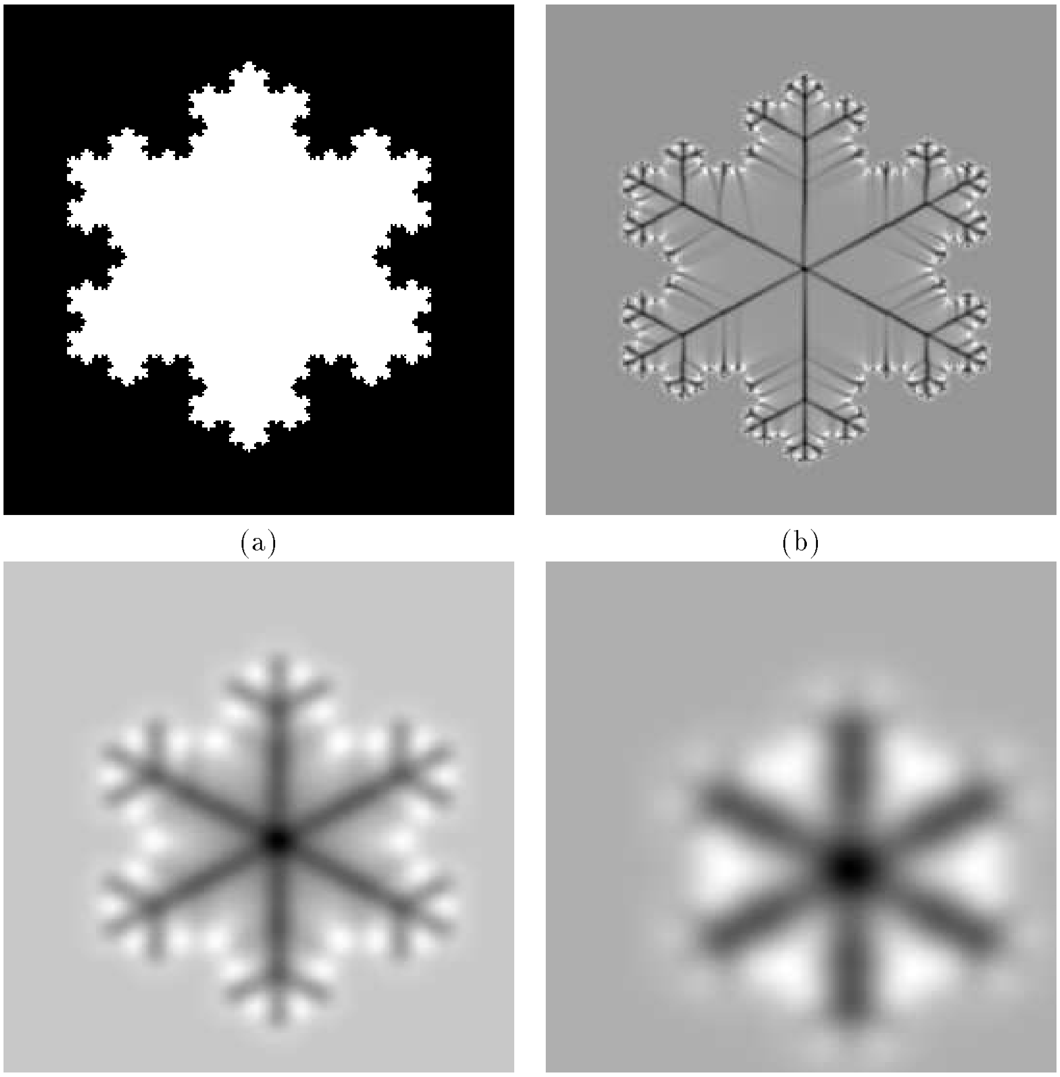


Figure 3.25a shows a Koch snowflake which is a type of fractal with regular boundary structure over spatial scale. Figures 3.25b,c and d show the effect of an increase of the smoothing which picks out detail at three distinct scales.

Figure 3.25: Multi-scale shape description: method applied to a Koch curve at a variety of spatial scales

Chapter 4

Skeletonisation using parallel boundary propagation

4.1 Introduction

An alternative to the distance transform approach to skeletonisation is wave propagation. The idea is both simple and intuitively appealing, *i.e.* a wave is propagated inward from the shape boundary and a skeletal point is marked where wavefronts collide. In this chapter we focus on a particular wave propagation algorithm due to Brady and Scott [22] for the two following reasons. Firstly, it is a parallel algorithm; this is a useful property since the generation of skeletons can be computationally intensive and therefore methods for their efficient implementation are of interest to the research community. Secondly, most wave propagation algorithms share a limitation inherent to the distance transform approach that only the symmetric axis transform can be computed and not more general descriptors such as the smoothed local symmetries or the symmetry set. This is because once the wavefronts collide with each other they do not propagate further. This is *not* the case with the Brady/Scott algorithm and so it has a major advantage over the other algorithms in this respect.

In this chapter we first introduce the Brady/Scott algorithm as it is designed for a massively parallel computer called the Connection Machine [47]. We then show how the algorithm can be adapted to an entirely different type of architecture, a MIMD (Multiple Instruction Multiple Data) machine [37] so that problems of synchronization and data transfer which arise in such a mapping are reduced. We then describe an actual implementation of the algorithm on an example of such an architecture, namely an array of transputers [52].

Results of the algorithm for a variety of object types are presented and compared with the distance transform approach of the previous chapter. The similarity of the wave based algorithm to other approaches is outlined and problems with this type of algorithm are discussed.

4.2 The Brady/Scott Connection Machine algorithm

Brady and Scott introduced a parallel algorithm [22] to compute smoothed local symmetries using a wave propagation approach. This algorithm was tested on a simulator for the Connection Machine [47] which is a computer with a massively parallel architecture with 64K processors. In their algorithm each point in the image plane is associated with a processor. Some of these processors lie on the boundary of the shape and if so store the orientation of the boundary contour at that point. The contour processors then simultaneously propagate messages along their inward normals towards the interior of the shape. Processors within the region of the shape then look for instances of messages passing "almost overhead almost instantaneously", as Brady and Scott say in their paper [22]. If such an event should occur then the two message origins on the boundary are deemed to be locally symmetric and a message is sent to the corresponding SLS point, *i.e.* the bisector of the chord between them.

Looking at figure 4.1 we can see the justification for this approach. The two points A and B are locally symmetric because the angles α between their normals and the chord which joins them are equal. This also means that A and B are vertices of an isosceles triangle together with point O, the intersection of their two normals. As the sides AO and BO of the triangle are equal, any messages sent simultaneously from A and B along their inward normals will reach O together and thus signify that A and B are locally symmetric about the point S, the bisector of the chord between them.

4.3 The adapted MIMD algorithm

The key to efficient implementation of algorithms on MIMD architectures lies in judicious problem decomposition. It will be seen in this section that the way in which

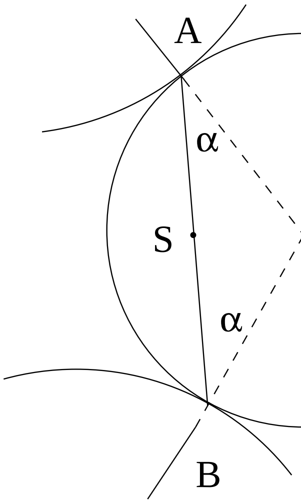


Figure 4.1: Local geometry of SLS points

the problem is divided among processors can have a great influence on the amount of computation performed and the communication and synchronization overheads that are encountered.

4.3.1 Design considerations for a MIMD array

In adapting this algorithm for execution on a MIMD architecture such as a transputer array it is important to keep in mind the nature of such an architecture, its strengths and its weaknesses. MIMD architectures are characterised by individual processors of relatively great power and versatility coupled with large amounts of local memory. They may have shared memory but often, as in this case, do not. This type of architecture performs well on problems which can be split up in a balanced way such that each processor is given its own task to do with its own slice of data and where needs for synchronization and large transfers of data are minimised [37].

Considering Brady and Scott's algorithm in light of this information it is clear that the algorithm requires adaptation for efficient execution on a MIMD architecture. Firstly, it would be absurd to implement a literal copy of the algorithm for the MIMD case. The Connection Machine has 64K processors, sufficient for a 256x256 image as in the above scheme, whereas typical transputer arrays have tens of nodes. Also, to bind each node to the processing associated with a single pixel would be an extreme underutilisation of the nodes' power and memory. Secondly, even if the task were divided more realistically by mapping the function of a few thousand Connection Machine nodes onto a single transputer, it would still involve globally synchronized propagation of messages along normals. The detection of symmetries would entail the large data transfers and synchronization overheads which we had hoped to avoid.

4.3.2 Synchronization reduction

While keeping the idea of messages propagating from contour points along inward normals, we can reduce the synchronization requirements as follows. Each pixel within the shape boundary is now considered as a bin into which data can be dropped. As a particular message passes over a particular bin it drops in two pieces of information, the distance it has travelled from its source on the boundary and the location of that source. After all the messages have propagated across the shape we can now look in the bins for distances which are approximately the same as this will signify that we have a potential local symmetry. The advantage of looking for closeness in distance rather than closeness in proximity and time is that we now do not need to propagate the messages in a globally synchronous manner. Messages on separate transputers can be propagated at any time and in any order without recourse to a global synchronization signal. The only synchronization requirement is that all the messages for a particular bin should have been delivered to it before we look through its contents for evidence of local symmetries, and even this fact does not have to result in inter-node communication if we split up the problem in a particular way as described below.

4.3.3 Data transfer reduction

Boundary decomposition

One way to split up the data is to give each node a copy of the whole shape and a unique portion of the boundary from which to propagate messages as shown in figure 4.2. The boundary portions assigned to transputers A and B are marked and we see two messages propagated from these portions and crossing at the bin marked O. This would be implemented by having local copies of bin O on both transputers. At the end of the propagation process we would need to synchronize the transputers and collect the data in both bins, and from the corresponding bins on other transputers. We would

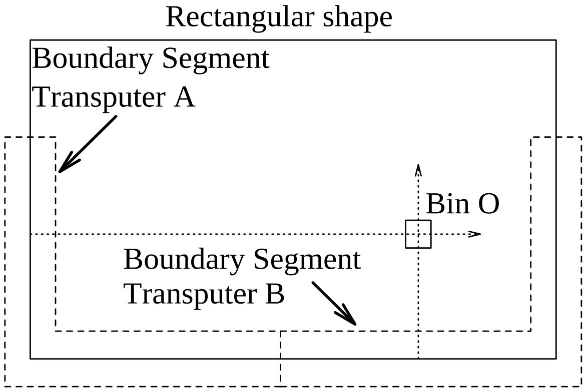


Figure 4.2: Boundary decomposition of problem

then redistribute the bin data, assigning a portion of the bins for each node to look through. This obviously requires a good deal of data transfer and synchronization. It is noted that the burden of data transfer for such a scheme would be reduced for a system with shared memory, although this would raise the problem of memory contention, *i.e.* nodes trying to write to the same memory location simultaneously, so this still leaves the need for some form of synchronization.

Interval decomposition

A better way to split up the problem is to give each node a complete copy of the boundary, but only have them propagate within small and disjoint intervals along all the normals as depicted in figure 4.3. For example, node A propagates from 0 to 5 pixels along all the normals, node B propagates from 5 to 10 pixels, *etc.* A small overlap is needed between the intervals so no potential symmetries are lost. Note that nodes do not repeat processing performed by other nodes, they jump immediately to the lower bound of their interval by use of real arithmetic. An especially interesting feature of this method is there is no need for inter-node communication as all possible local symmetry points, *i.e.* with lengths of the same order, are contained on the same node.

From a comparison of figure 4.2 and figure 4.3 we can see that this new decomposition also leads to a reduction in the total number of computations performed. The messages from transputers A and B cross at O and therefore we have evidence of a potential local symmetry. However this is not a local symmetry as the lengths the two

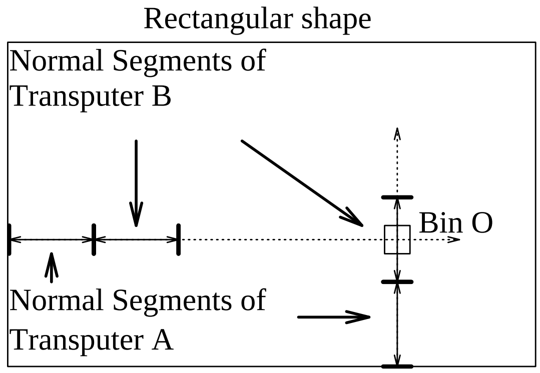


Figure 4.3: Interval decomposition of the problem

messages have travelled is greatly different. Nevertheless, in the boundary decomposition method we still take time to consider and discard them. In the new interval decomposition method we would not do this as we only consider intersecting normals of similar lengths. This effect becomes more pronounced as more processors are added to the array.

Rotational symmetries

Here we suggest that some measure of rotational symmetry can be detected at the same time as local symmetries are on a MIMD architecture. This is because points on a circular arc all have normals which cross at the centre of the circle. This was a problem on the Connection Machine because all these messages arriving simultaneously at one point swamped the finite capacity of the processors. There is no such problem on a MIMD architecture using a voting scheme as proposed above. Bins can be dynamically allocated and reach a large size without problems. A bin can easily contain all the boundary points but overall memory requirements remain low as the vast majority of bins in most shapes only contain a few votes.

Rotational and local symmetries are easily distinguished as local symmetries have only a few votes associated with them whereas rotational symmetries can have dozens or even hundreds. The criterion for rotational symmetry used here is the very simple one of the intersection of many normals at one point. Note that Brady's original formulation included the more sophisticated criteria originating from the work of Fleck [40] and can thus cope with "near circularly symmetric" shapes such as a pear or egg.

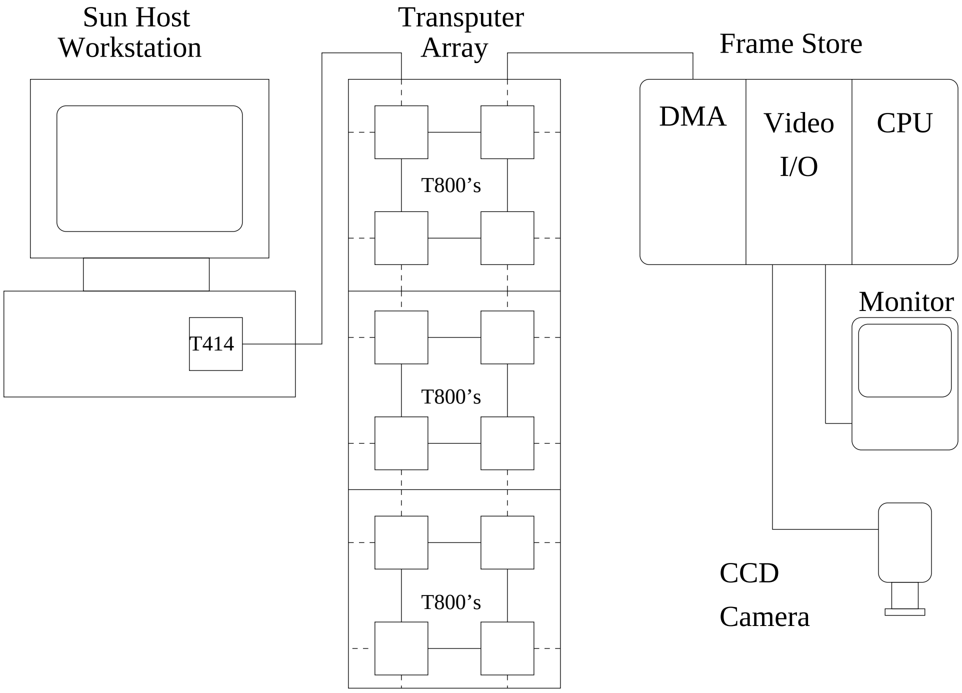


Figure 4.4: System layout

4.4 Implementation on an array of transputers

The MIMD architecture used for the implementation of the algorithm was an array of transputers [52]. The system is shown in figure 4.4. The transputer array is hosted by a SUN workstation to which is attached an interface board containing a T414 transputer, this in turn being connected to the main array of twelve T800 transputers.

Test images were downloaded from the host to the T414 which acted as the master node of the array. Here Canny edge detection was applied to give edge position and orientations. This information was then downloaded to the main array and each T800 given a copy of the shape plus an interval of normal lengths to propagate. On completion of the voting process the output was sent to the T414 where it was collated before being saved to the host for display. It was also possible to load live images either into the transputer array or to the host work station from a CCD camera and framestore if so desired.

4.5 Results

The primary motivation for using the wave propagation approach of Brady and Scott is that it makes explicit more symmetries of the shape than the distance transform approach. Figure 4.5a shows the skeleton of a synthetic tee shape produced using the distance transform and Petrou filter introduced in chapter 3. Figure 4.5b shows the skeleton produced by the wave propagation algorithm. Clearly more symmetries are made explicit using the wave based approach. Note that the skeletal point of each bi-tangent circle is drawn at the center of the circle rather than at the midpoint of the chord between the boundary points. This is to facilitate a direct comparison between the two approaches. The distance transform data could not be transformed into SLS points as the information about which boundary points contribute to the skeleton at each point is not preserved.

Figure 4.6 shows the additional symmetries made explicit by the wave propagation approach for two real objects. In figure 4.6a the reflectional symmetry of the pair of pliers is made explicit by an additional central skeletal branch which is not represented in the distance transform skeleton in figure 4.6b. In figure 4.6c additional skeletal branches for a spanner are shown. Extra branches are present in the jaws of the spanner and four branches project from the ends representing symmetries between the jaws and stem. None of these branches is present in the distance transform skeleton of figure 4.6d.

The wave propagation approach works well with elongated objects. Figure 4.7a and 4.7c show the wave based skeletons for a file and hammer. The tip of the file and end of the hammer handle show the crossing skeletal branches which would be expected for rectangular shapes. Smeared branches project from the base of the blade and handle of the file and also from the head of the hammer which are not present on the distance transform skeletons shown in figures 4.7b and 4.7d. In terms of skeleton quality the branches of the distance transform skeletons appear more connected with fewer extraneous points than the skeletons produced by the wave based approach.

Figure 4.8 shows skeletons of less elongated objects. The skeleton is clearly visible on the neck of the giraffe in figure 4.8a and that of the ostrich in figure 4.8c. The skeleton is incomplete for the bodies of both animals and there are a great many apparently extraneous skeletal points in these areas. The distance transform skeletons of figures 4.8b and 4.8d are in contrast connected and not obscured by extraneous skeletal points.

The skeletons for a goat and rabbit are shown in figure 4.9. The wave based skeletons of figure 4.9a and 4.9c contain very few connected branches. Skeletal points within the bodies of the animals are scattered. The distance transform skeletons are shown in figures 4.9b and 4.9d and are again clearly defined and for the most part a connected whole.

Some interesting output from the wave propagation algorithm is given in figure 4.10. The handle of the screwdriver and head of the spoon are both approximately rotationally symmetric. The skeletal points of these regions depicted in figures 4.10a and 4.10c are not well defined branches but they are also not randomly scattered points. The algorithm appears to be making some aspect of the objects geometry explicit but it is not obvious what this might be. Further comments on these patterns must wait until chapter 5 where a mathematical analysis of the distance transform provides the insights necessary for their explanation.

4.6 Discussion

The wave propagation technique implemented in this chapter is closely related to a number of other vision processing algorithms. Firstly as the algorithm uses "voting" or evidence accumulation it can be seen as a relative of the Hough Transform [51] and in particular the Hough Transform for circle detection [10]. If we look at figure 4.1 we can see that two locally symmetric edge points are tangents to a circle and the inward normals form radii. In propagating the normals the algorithm votes for the existence of the centre of such a circle. This will typically have two votes for a point of local symmetry but many more for a centre of rotational symmetry. Secondly the algorithm can be said to belong to a family of related techniques that could be called "propagation methods" as they all use the same basic idea of propagating wavefronts or signals down inward normals. The grass fire algorithm [15, 16] uses this idea to compute the SAT. Yuen [102] suggested that normal projection could be used to detect rotational symmetry but suggests a time/place approach as in the Connection Machine algorithm and separately from local symmetries. Ponce has suggested normal projection as a basis for computing skewed symmetries [77]. The Normal Contour Distance shape descriptor

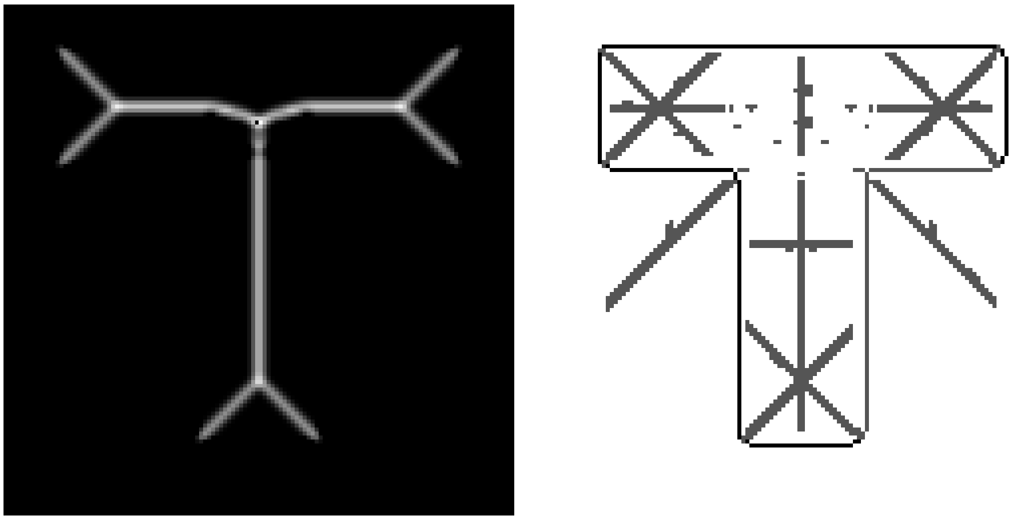
[92] creates a signature of a shape based on the distance of each point on the boundary to the opposite side of the shape along the direction of its inward normal. The normal contour distance could be computed using a wave propagation technique.

The wave propagation approach clearly has problems in computing a compact and complete skeletal description which could be caused by a number of factors. Noise and quantisation lead to positional errors in the boundary position and angular errors in the normal direction supplied by the Canny edge detector. These errors could lead to false or missed skeletal points. As propagation is on a discrete grid, normals can cross but not actually vote in the same bin thus missing each other like bishops on a chess board, although this problem is solved here by voting in not one bin but a small neighbourhood. Intersecting normals of roughly similar origin and angle can run close together for some distance either side of their true intersection point thus producing multiple skeletal points.

4.7 Summary

It has been shown that an efficient implementation of Brady and Scott's normal propagation algorithm to compute skeletal representations is possible on a MIMD architecture if the new interval decomposition is used. In addition there is an added bonus that in principle rotational symmetries can be computed simultaneously with one analysis of the shape (although this was not implemented here).

It can be seen that this method is closely related to many others and can be thought of as belonging to a generic family of shape description techniques. The algorithm is of importance because it transcends the intrinsic limitation of the distance transform approach to skeletonisation that only the SAT can be computed. However, although its performance is quite good for thin elongated objects it performs less well for wide irregular shapes. The quality of the skeletons produced by the wave based approach, in terms of connectivity or completeness, is not as good as that produced using the distance transform approach.



(a) (b)

Figure 4.5a shows the skeleton for a synthetic tee shape obtained using the distance transform and Petrou operator. Figure 4.5b shows the additional symmetries that can be obtained using the wave based approach.

Figure 4.5: Additional symmetries expressed using the wave based approach

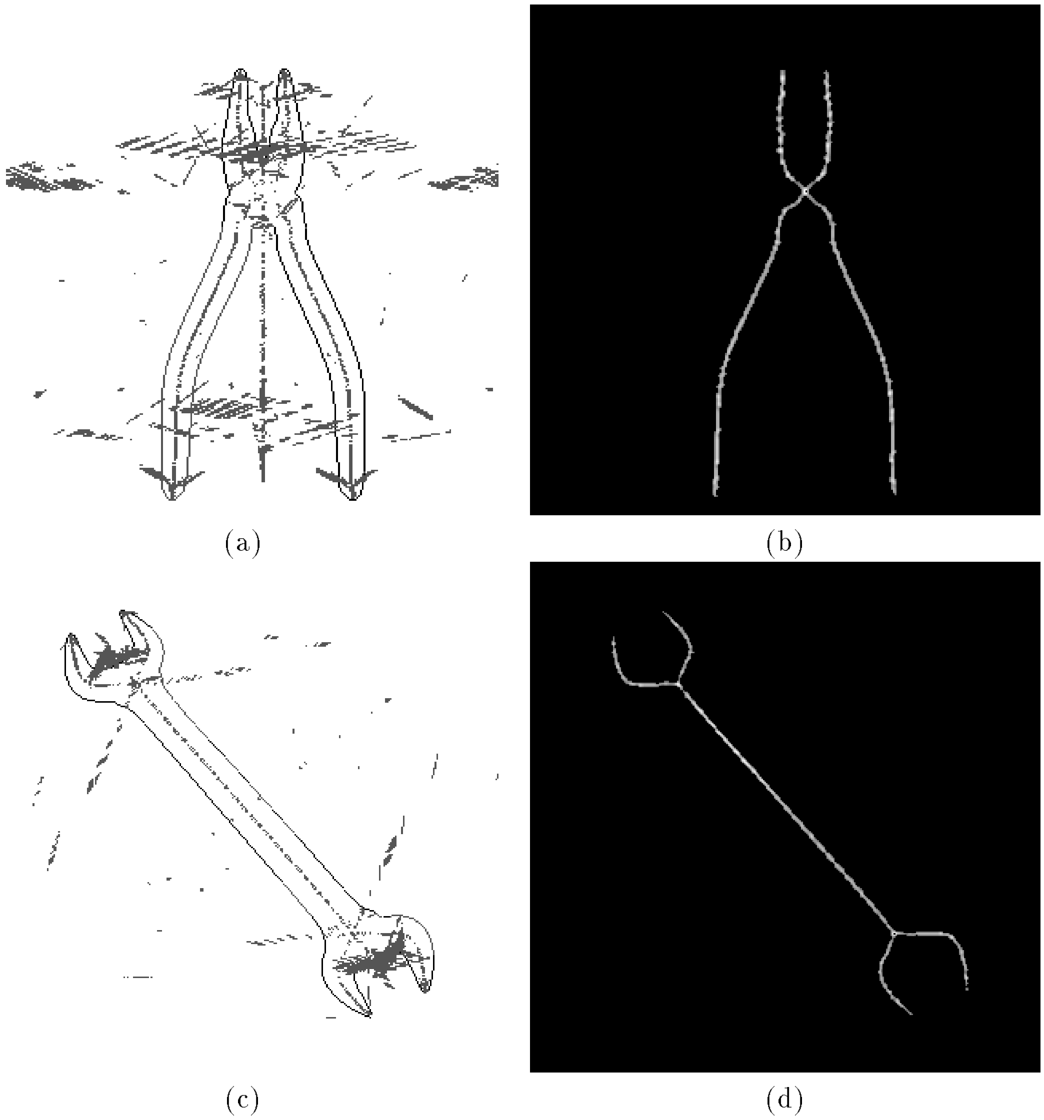
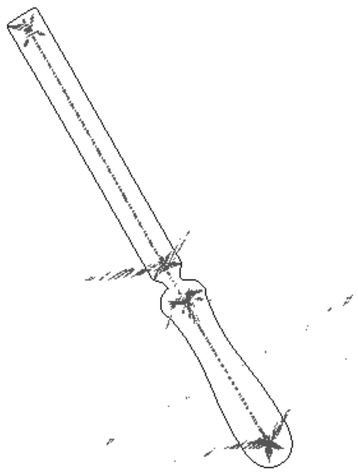
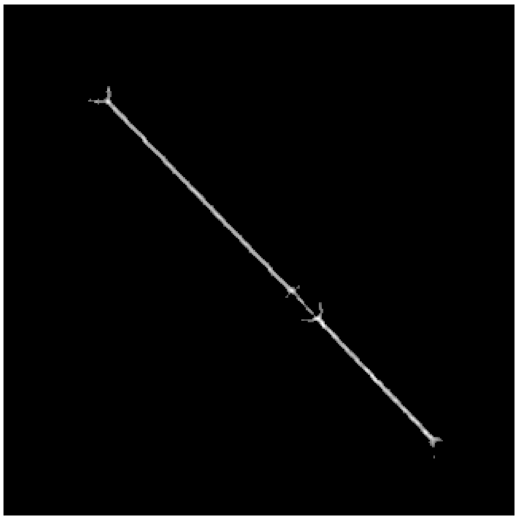


Figure 4.6a is the skeleton of a pair of pliers obtained using the wave based approach and figure 4.6b shows the distance transform skeleton. The additional symmetry between the handles is clearly visible in figure 4.6a. Figure 4.6c is the skeleton of a spanner obtained using the wave based approach and figure 4.6d shows the distance transform skeleton. In figure 4.6c additional symmetries are expressed in the jaws and two branches run outwards from each end of the spanner.

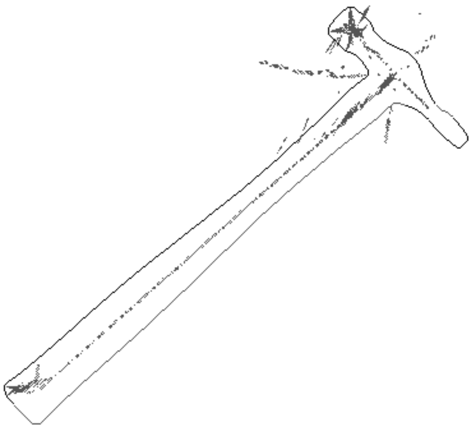
Figure 4.6: Additional symmetries expressed in real objects



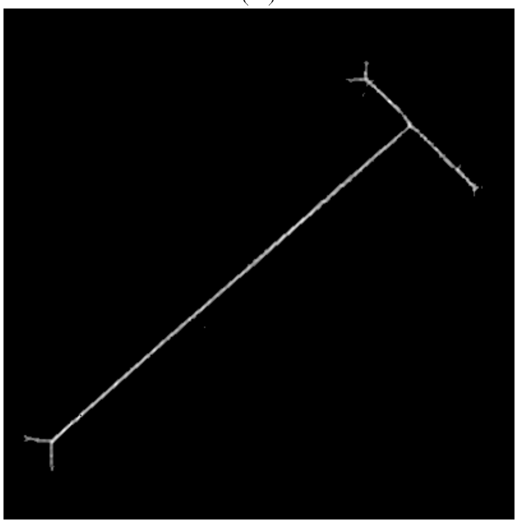
(a)



(b)



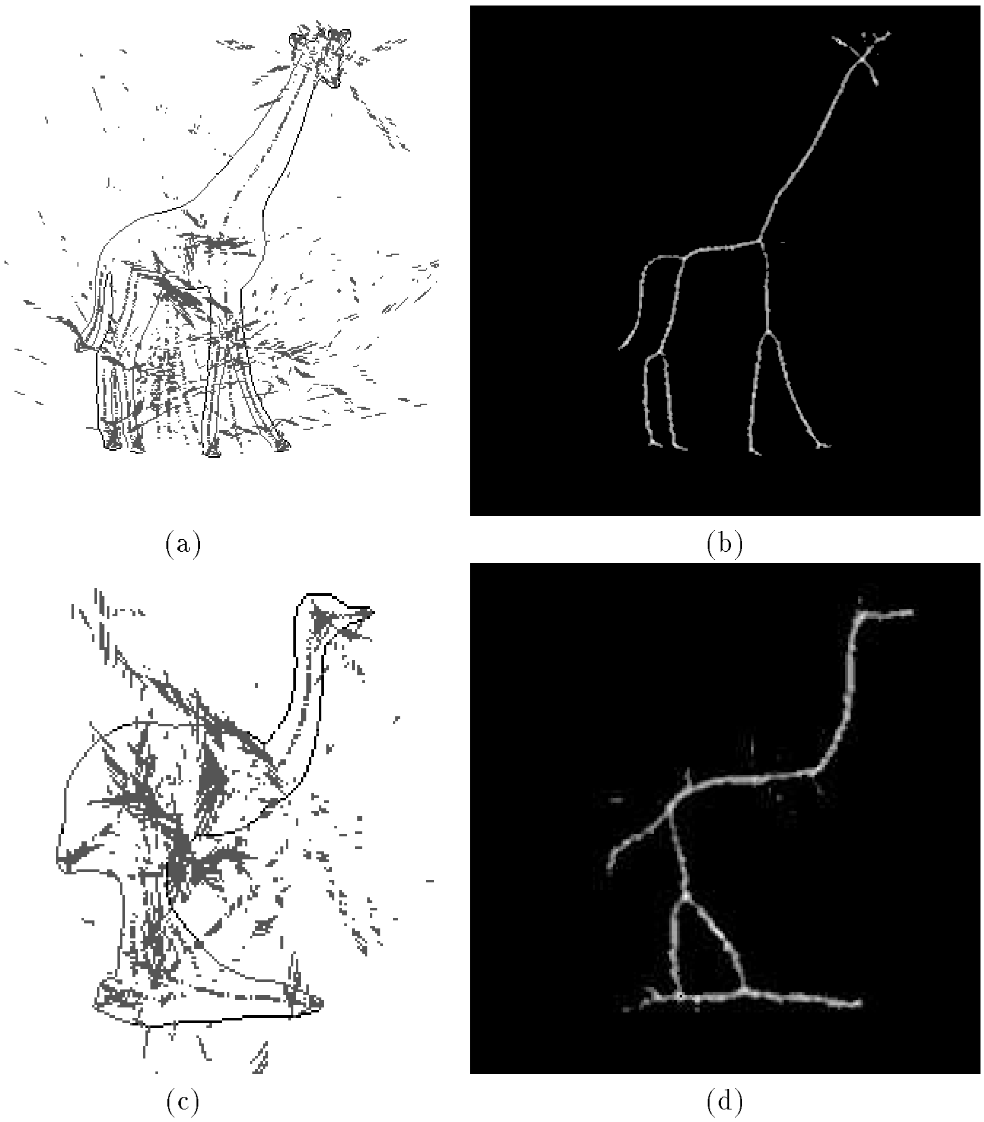
(c)



(d)

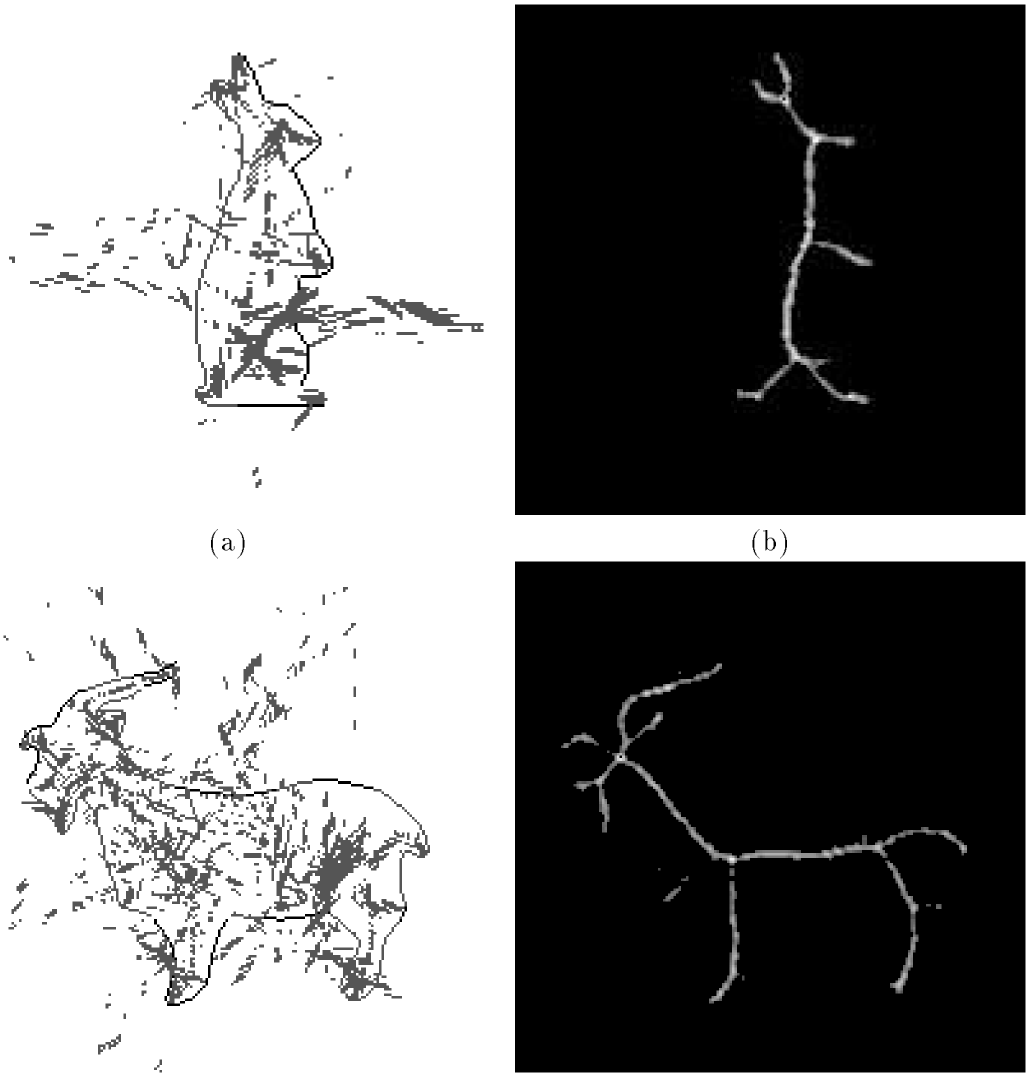
The wave propagation approach works well for elongated objects. Figure 4.7a is the skeleton of a file obtained using the wave based approach and figure 4.7b shows the distance transform skeleton. Similar results are obtained for a hammer in figure 4.7c and figure 4.7d.

Figure 4.7: Wave propagation algorithm for elongated objects



Figures 4.8a and 4.8c show skeletons produced by wave propagation for a giraffe and ostrich respectively. The skeleton is clearly visible on the neck of the animals but the bodies are obscured by a mass of skeletal points. The distance transform skeletons are shown in figures 4.8b and 4.8d for comparison.

Figure 4.8: Wave propagation algorithm: elongated irregular objects

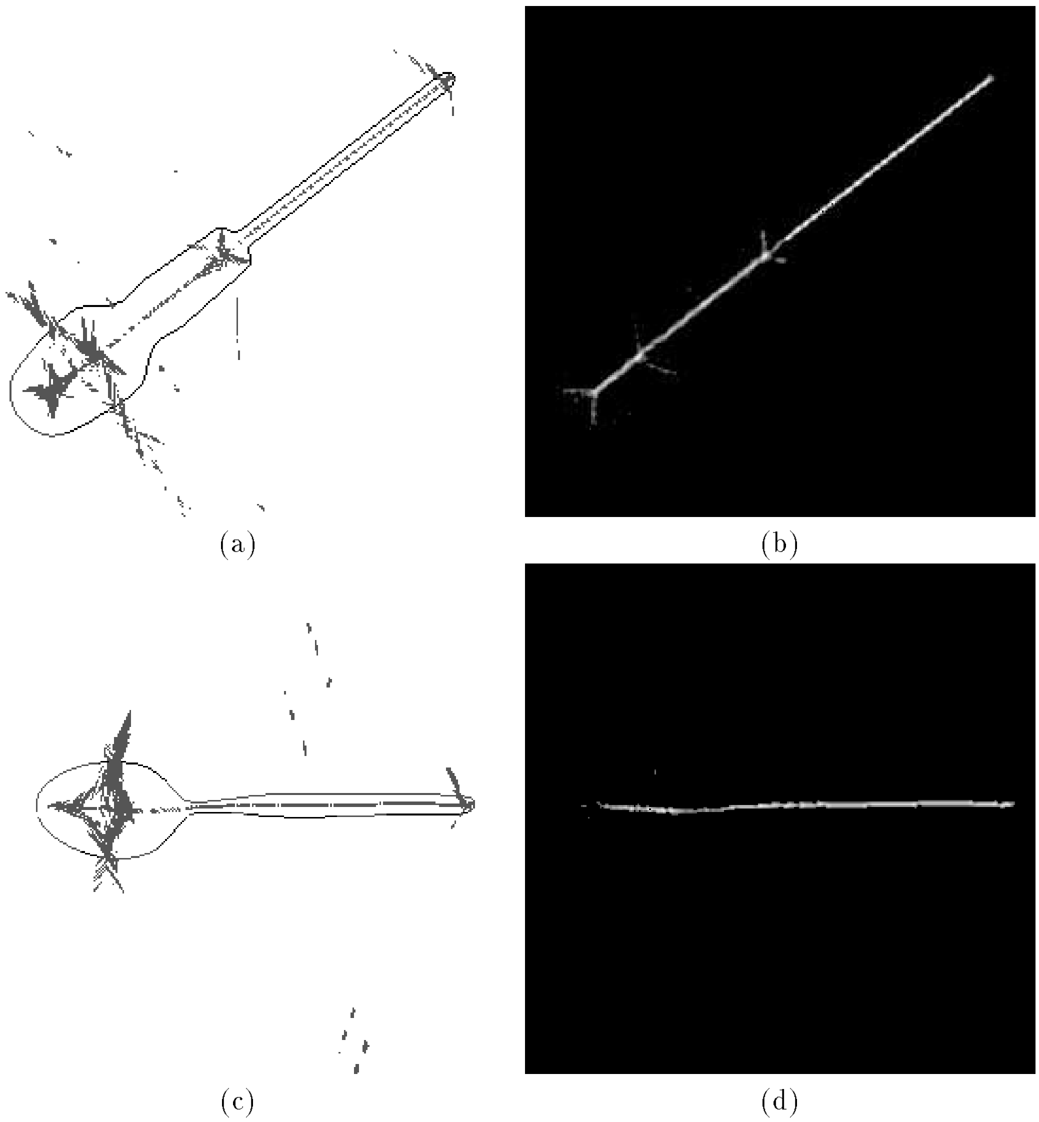


(a) (b)

(c) (d)

Figures 4.9a and 4.9c show skeletons produced by wave propagation for a rabbit and goat respectively. Very few skeletal branches are visible at all. The distance transform skeletons are shown in figures 4.9b and 4.9d for comparison.

Figure 4.9: Wave propagation algorithm: irregular objects



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

The wave propagation approach produces interesting results for objects with rotational symmetry. The handle of the screwdriver and head of the spoon in figures 4.10a and 4.10c produce smeared skeletal points but with a definite structure (see chapter 5 for an explanation). Skeletons of the objects obtained using the distance transform are shown in figures 4.10b and 4.10d.

Figure 4.10: Wave propagation algorithm: rotational/elongated objects

Chapter 5

The extended Euclidean distance transform

5.1 Introduction

Thus far we have introduced work involving two separate approaches to skeletonisation. In chapter 3 we investigated the distance transform approach to skeletonisation. The method involves first computing the Euclidean distance transform and then finding the "local maxima" in this representation. Our contribution to this approach was to cast the detection of these "local maxima" as a filtering problem and to identify a filter which could perform this task reliably. Although improvements to this technique were possible it was found that it has an intrinsic limitation in that only the symmetric axis transform could be computed and not more general descriptions such as the symmetry set. In chapter 4 we investigated an algorithm due to Brady and Scott which does not have this limitation. We first showed how this algorithm could be efficiently mapped onto a MIMD array of processors and introduced a small improvement to the skeletal point detection mechanism. The algorithm works quite well with thin, elongated shapes but in general the quality of skeletons produced is not as good as the quality of those produced by the distance transform approach.

It would be highly desirable to have a representation which exhibited the good characteristics of both of the above approaches, *i.e.* to deliver the relatively good quality skeletons of the distance transform methods but also to have the ability of algorithms like Brady and Scott's to deliver more general skeletal descriptions of shape than the symmetric axis transform. The main contribution of this chapter is to introduce the *extended* Euclidean distance transform, a representation which could form the basis of such an approach.

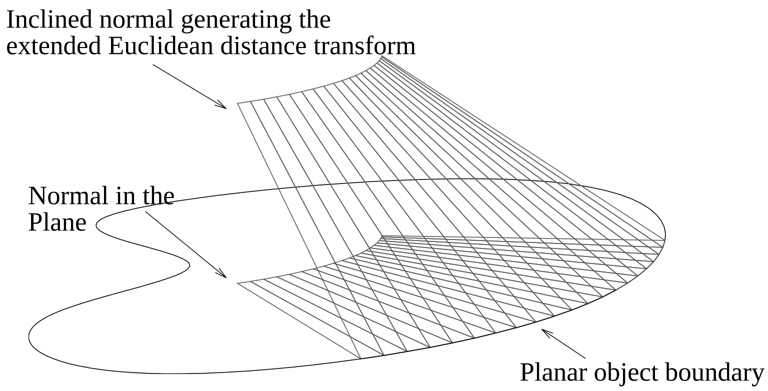
Firstly, an intuitive explanation of the extended Euclidean distance transform is given. Then a theoretical analysis of distance functions from smooth planar curves is given in which the extended Euclidean distance transform is formally defined. The singularities which occur on this surface are defined along with their relationship to the curve geometry and skeletal representations. Using this analysis interpretations of the standard distance transform and wave based methods of skeletonisation are also discussed. The analysis of this chapter is a shorter version of that to be found in appendix A; this is so the main stages of the analysis can be understood without being obscured with too much detail.

5.2 The extended distance transform

5.2.1 An intuitive explanation

A central assumption of the distance transform approach is that skeletal axes correspond to the "local maxima" of the Euclidean distance transform. A mathematical analysis of this assumption from the point of view of singularity theory [27] is given below which leads to the introduction of the extended distance transform. However, before embarking on a formal theoretical analysis it is instructive to convey an intuitive description of the extended distance transform. Consider a planar smooth curve as in figure 5.1. We can imagine tracking a point as it traverses the curve and making note of its inward normal which would sweep across the plane. (In principle this normal would project indefinitely although the normal in figure 5.1 is truncated to aid clarity.) Now consider this normal raised at an angle of 45 degrees to the plane so that its height above the plane encodes distance along the normal from the boundary. As this new inclined normal is swept along the curve it acts as a generator defining a surface in the space above the plane. It is this surface which is the *extended* distance transform or, as will be defined shortly, the *discriminant*.

Figure 5.2 shows the extended distance transform or discriminant of a circle. It consists of a lower cone which corresponds to the ordinary Euclidean distance transform but extends above this to form another inverted cone. It is this upper portion which has



We can imagine the extended distance transform being generated by a normal inclined at 45 degrees to the plane as it is swept around the boundary of a shape.

Figure 5.1: Intuitive explanation of the extended distance transform

the potential to interact with surfaces from other parts of the contour and thus define symmetries other than those of the SAT. In figure 5.3 we see the surface generated by a parabola. There are two significant types of "event" on this surface, *i.e* points where the surface appears "creased" and points where two smooth portions of the surface intersect. The creases correspond to the *evolute* of the curve which is the locus of centres of radii of curvature and the intersections correspond to the skeletal symmetry points.

5.2.2 Theoretical analysis

Let $\gamma(t)$ be a regular parameterization of a curve, which we can assume to be unit speed for simplicity of calculation. (That is, $\| (\gamma'(t)) \| = 1$, where the prime denotes differentiation.) The function

$$f(t, x, y) = \| \gamma(t) - (x, y) \|$$

measures the distance from the point $\gamma(t)$ of the curve to the general point (x, y) in the plane. (Compare with [27, p.33], where the *square* of f (the 'distance-squared' function) is used. The only purpose of squaring is to make the function differentiable when (x, y) lies on the curve, and for our purposes we can ignore this problem.)

We also consider the family F of functions as follows:

$$\mathbb{R} \times \mathbb{R}^2 \xrightarrow{F} \mathbb{R} \times \mathbb{R}^2$$

$$t, x, y \longmapsto \| \gamma(t) - (x, y) \|, x, y$$

The **critical set** Σ of F is the set of points t, x, y where the Jacobian matrix of F is singular. This is the same as the set of points where f has a critical point, and this occurs precisely when x, y lies on the normal to the curve at $\gamma(t)$. (See [27, p.33].) Thus

$$\Sigma = \Big\{(t, x, y) \mid \frac{\partial f}{\partial t} = 0\Big\}.$$

The **critical locus** or **discriminant** Δ of F is defined by

$$\Delta = F(\Sigma).$$

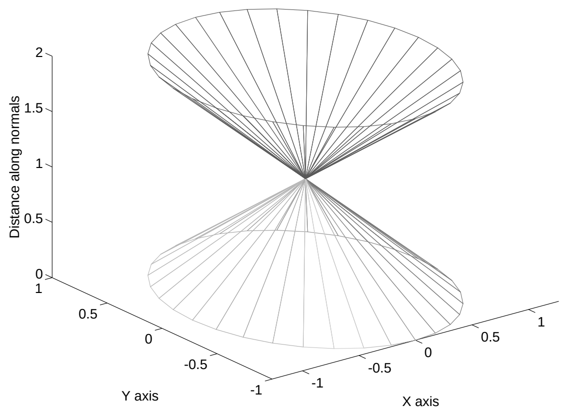
This is a subset of the target space $\mathbb{R} \times \mathbb{R}^2$.

To see how this relates to the intuitive explanation above, fix t. In Σ there is a line l of points, namely the normal at $\gamma(t)$ but raised to height t. The image $F(l)$ is a line whose projection to the (x, y)-plane is still the normal line. But $F(t, x, y)$, for $(t, x, y) \in l$, is raised above this normal line to a height equal to its distance from the point $(0, \gamma(t))$ at the foot of the normal. Hence the image $F(l)$ is a line through $(0, \gamma(t))$ at 45^o to the base plane $t = 0$.

The surface formed by the 45^o lines is a *ruled surface* and will usually have singularities. The circle (figure. 5.2) is a rather un-typical case here in that all the lines actually pass through a common point. A more typical case is that shown in figure 5.3, where the surface has cusp edges and a swallowtail point.

It is our task first to characterise the singularities of the discriminant surface and then to outline methods for detecting them.

We use singularity theory to define the type and position of singularities which occur on Δ. The **critical set** Σ is a smooth surface and so Δ can have only (i) 'multi-local' and (ii) local (or intrinsic) singularities. We proceed to describe these in turn.



The discriminant can be thought of as normals to the original curve rising from the plane at 45 degrees. For a circle the discriminant looks like two cones with a solitary singularity at their common apex.

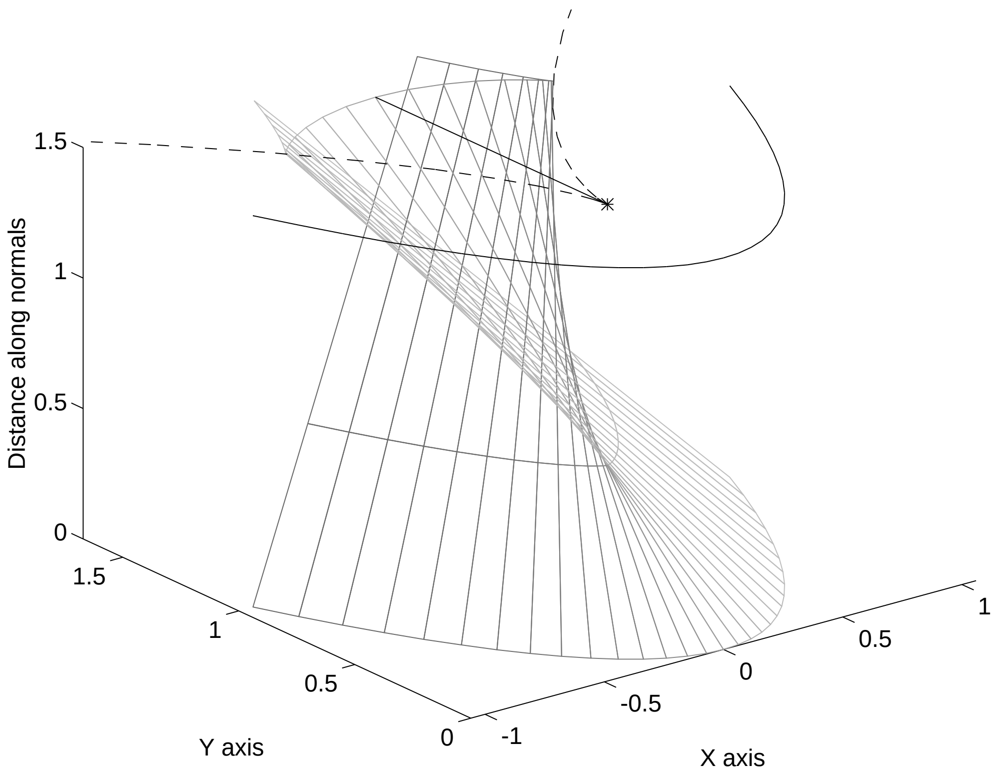
Figure 5.2: Discriminant for a circle

5.2.3 Multi-local singularities

Multi-local singularities are caused by two points of Σ mapping to the same point of Δ. That is, there exist points $(t_1, x_1, y_1), (t_2, x_2, y_2) \in \Sigma$ with $F(t_1, x_1, y_1) = F(t_2, x_2, y_2)$. This clearly requires $x_1 = x_2$ and $y_1 = y_2$, so corresponds to two normals *of the same length*, intersecting at (x_1, y_1). (The lengths are measured from the 'feet' $\gamma(t_1), \gamma(t_2)$ of the normals to the intersection point.)

5.2.4 Local singularities

The local (or intrinsic) singularities of Δ are images of points $(t, x, y) \in \Sigma$ where the map $F|\Sigma$ (that is, F restricted to the smooth surface Σ) is singular. It turns out that $F|\Sigma$ is singular when $(t, x, y) \in \Sigma$ and (x, y) is at the centre of curvature of γ at $\gamma(t)$. (The calculation for this is almost identical to that in [27, p.33].) This means that (x, y) is on the evolute which is the locus of centres of curvature (see figure 5.3). Note that $F(t, x, y)$ only belongs to the Euclidean distance transform when the distance along the normals from (x, y) to $\gamma(t)$ is the shortest among all normals through (x, y). In fact this never happens on the evolute except at cusps (corresponding to maxima and minima of curvature).



The discriminant for a parabola shows a "swallow tail" surface. Above the discriminant is drawn the parabola complete with evolute (dashed line) and symmetric axis (straight solid line). The central symmetric axis corresponds to points of multi-local symmetry on the discriminant, the evolute corresponds to intrinsic singularities which are along the cusped edges of the "swallow tail". The only intrinsic singularity to have a minimum normal distance and therefore to be on the symmetric axis is that corresponding to the cusp on the evolute (asterisk).

Figure 5.3: Discriminant for a parabola

5.3 Interpretation of the Euclidean distance transform

In the analysis we described multi-local singularities as the intersection of equal length normals. Equal length normals intersect in the centres of bi-tangent circles, that is, circles tangent to γ at two points, namely $\gamma(t_1)$ and $\gamma(t_2)$. Thus the points (x_1, y_1) generated in this way are part of the symmetry set ([42]). The point (x_1, y_1) will also be part of the SAT, if the equal normals are the shortest normals from (x, y) to the curve.

This confirms that the so called "local maxima" of the Euclidean distance transform do indeed correspond to the SAT but we can now define them more formally not as "local maxima" but as singularities. We can also see that the standard Euclidean distance transform is a subset of the discriminant; specifically it is that portion of the discriminant consisting of normals running from the boundary of the shape and terminating at the first singularity they encounter. The standard distance transform is single valued for any position x, y whereas the extended distance transform can be multi-valued. The only intrinsic singularities on the standard distance transform are points corresponding to cusps of the evolute at maxima of curvature whereas large sections of the extended distance transform have intrinsic singularities corresponding to the entire evolute.

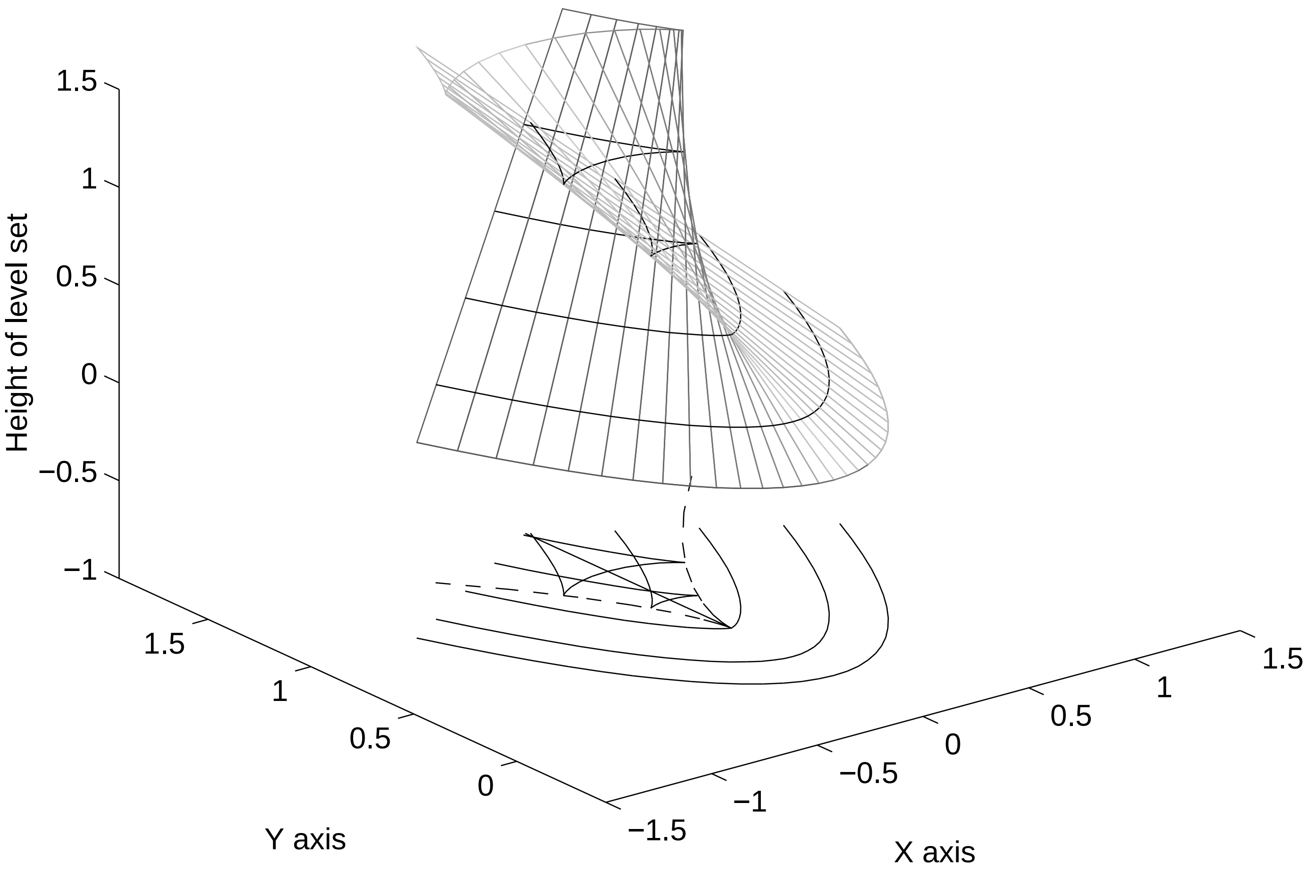
5.4 Interpretation of wave propagation algorithms

The relationship between wave based methods such as Brady and Scott's algorithm and the extended Euclidean distance transform can be described using the concept of level sets. A familiar example of level sets are the contours on a geographic map, each individual contour telling us what points on the earth's surface are at the same height above sea level. In a similar way we can construct level sets of the discriminant. Firstly a curve is defined by the intersection of a horizontal plane of height c with the discriminant and then this curve is projected down to the plane.

Formally the level set for c is the set (c, x, y) where $c = f(t, x, y)$ and (x, y) is on the normal at t. Thus (x, y) lies on the parallel of the curve at a distance c. The level set is this parallel lifted up to the discriminant surface.

In other words a wavefront which has propagated a distance c from the boundary is the parallel to the original curve at distance c and can be thought of as the projection of the level set of the discriminant at height c to the plane. The entire wave propagation process thus computes the projection to the plane of all level sets of the discriminant from $c = 0$ to some positive upper bound in ascending order.

We demonstrate this relationship for a parabola in figure 5.4. It is easy to see that collisions of wavefronts in the plane correspond to multi-local singularities of the discriminant occurring at a particular level set. It is perhaps less intuitively obvious that the wavefronts should form cusps and that as the wavefront progresses the cusps follow the evolute. These cusps correspond to local singularities of the discriminant at a particular level set.



The discriminant for a parabola is a "swallow tail" surface. Below the discriminant is drawn the parabola complete with evolute (dashed line) and symmetric axis (straight solid line). Four waves or parallels of the parabola are seen to correspond to four distinct cross sections of the discriminant. Note that once the wavefront encounters the evolute it forms cusps which propagate along the evolute.

Figure 5.4: Discriminant level sets and wavefronts for a parabola

An example of the wave propagation process is shown in figure 5.5. Figure 5.5a shows a cubic oval (outer smooth curve) and its evolute (cusped curve); the inner smooth curve is a wavefront which is propagating inwardly from the oval boundary. It is interesting to note that as the wave propagates from the shape boundary the twin cusps of a "swallow tail" are created as the wave hits an evolute cusp associated with and pointing towards a maximum of curvature on the boundary, figure 5.5b. These cusps then follow the evolute, as in figures 5.5c and d, until they are annihilated at cusps of the evolute associated with and pointing away from adjacent minima of curvature, figures 5.5e and f. If there is an inflection in the boundary due to a concavity then the cusps become asymptotic to the normal at the point of inflection and propagate to infinity. For a curve which is smooth, convex and closed, once all points of all wavefronts have propagated beyond the evolute no further collisions are possible and so the skeletal points are constrained to lie within the evolute.

Colliding wavefronts correspond to multi-local singularities of the discriminant and thus define skeletal points. Using the geometry of the discriminant, we can define two types of skeletal point: those which have been created by wavefronts which have *not* propagated beyond the evolute, as in figure 5.5b; and those which have been created by wavefronts which *have* passed beyond the evolute, as in figure 5.5d. This can be used as a principle by which to filter or rank skeletons. Indeed, most "pre-evolute" skeletons appear to be of perceptual relevance whereas many post-evolute skeletons do not. For example, large wavefronts propagate from smooth convex corners of small radius. However, it should be noted that counter examples exist such as the skeletal branch associated with the minor axis of an ellipse.

The discriminant provides an explanation of the patterns produced by the wave propagation algorithm of the last chapter for objects which are approximately rotationally symmetric. These correspond to parts of the discriminant close to intrinsic singularities, *i.e.* normals which pass close to the evolute of the curve boundary. For a perfectly circular object the evolute collapses to a stationary point at the centre of the circle. The slightest perturbation however, creates at least two maxima and two minima of curvature in the shape boundary and hence an evolute with several cusps. The evolute remains close to the centre of such a shape and doesn't move much as the shape boundary is traversed because the shape is of approximately constant curvature. This means that normals of approximately the same length but from widely spaced points on the boundary can come very close to each other near the slow moving evolute of near circular parts. The evolute is also slow moving near any curvature extremum as the evolute changes direction at that point and therefore stops, thus the normals show a similar behaviour. The wave propagation algorithm has no way of telling these instances apart from multi-local singularities and so marks them as skeletal points. So the patterns produced by the wave propagation algorithm at the centre of near circular parts and at curvature extrema correspond to the evolute and its interior.

5.5 Summary

Using standard techniques from singularity theory we have defined an *extended* Euclidean distance transform. The study of distance transformations from this point of view is a valuable tool that offers many new insights; in particular we should note the following points. Singularity theory tells us exactly what type of singularity can exist, under what conditions they occur and their local structure [27, 26, 11, 6]. It was demonstrated that two types of singularity exist on this surface, *i.e.* multi-local singularities, corresponding to the symmetry set, and intrinsic singularities corresponding to the evolute.

The standard distance transform is a subset of the extended Euclidean transform and is that portion from the curve boundary to the first singularity that any normal encounters. In the literature the artifacts on the standard Euclidean distance transform which correspond to skeletal points are commonly called "local maxima" [3]. This is a misnomer as strictly speaking they are not local maxima of any variable in the usual sense. Using singularity theory in this context these artifacts can be described simply and formally as singularities of the distance function.

Wave propagation algorithms such as that due to Brady and Scott [22] can be interpreted as the ordered computation of level sets of the discriminant projected down to the plane. The wave propagation algorithm cannot distinguish between multi-local singularities, corresponding to skeletal branches, and intrinsic singularities corresponding to the evolute. This shortcoming explains some of the additional skeletal patterns encountered in chapter 4, particularly for near circular objects.

The discriminant can be thought of as an *extended* Euclidean distance transform as it contains the standard Euclidean distance transform but also describes distances along normals which are not minimal, the description of singularities in such cases enabling us to detect smoothed local symmetry and symmetry set points. This fact is exploited to advantage in the next chapter.

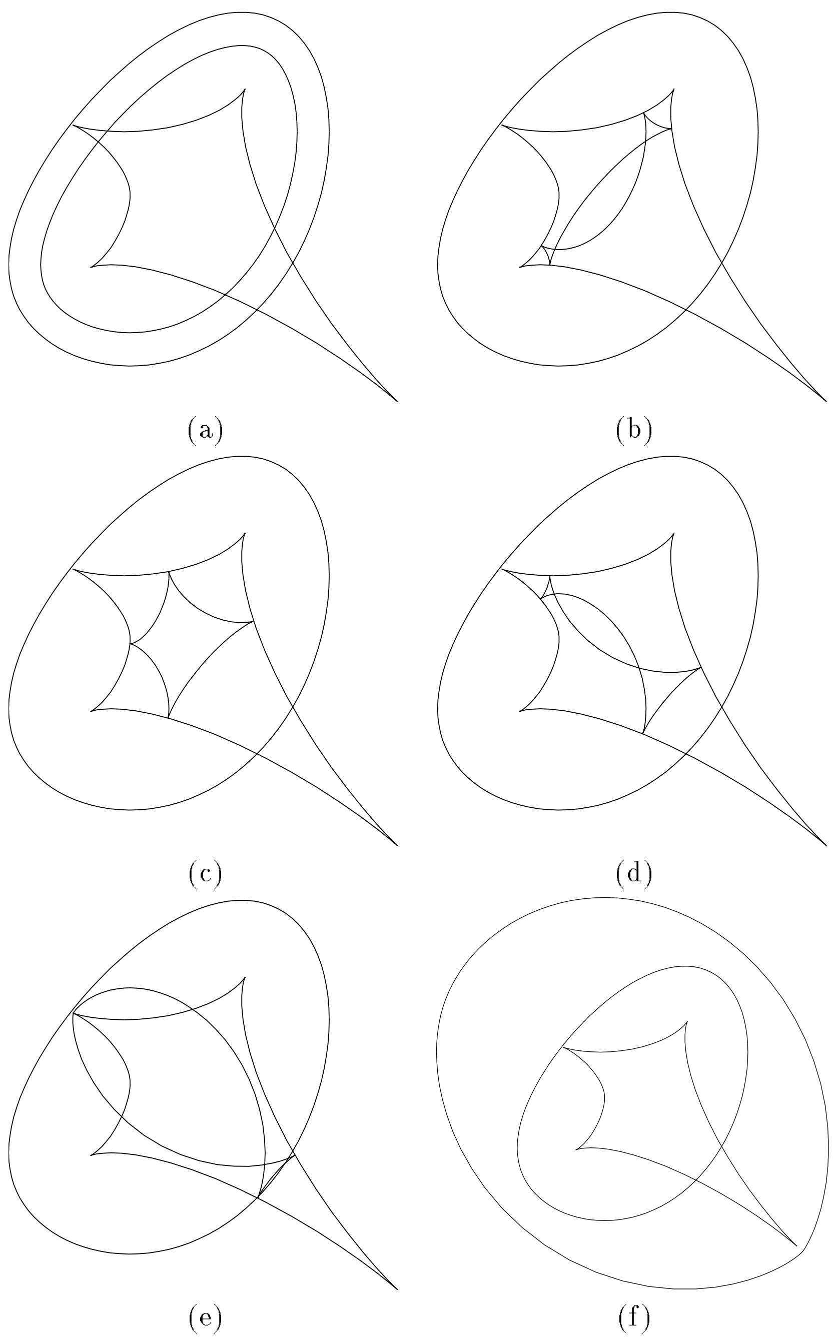


Figure 5.5a shows a cubic oval shape (outer smooth curve) and its evolute (cusped curve). A wavefront (inner smooth curve) has been propagated inwardly from the shape boundary. Figures 5.5b, c, d, e, and f show the evolution of this wavefront and its interaction with the evolute.

Figure 5.5: Wavefront evolution for a cubic oval

Chapter 6

Skeletonisation using an extended Euclidean distance transform

6.1 Introduction

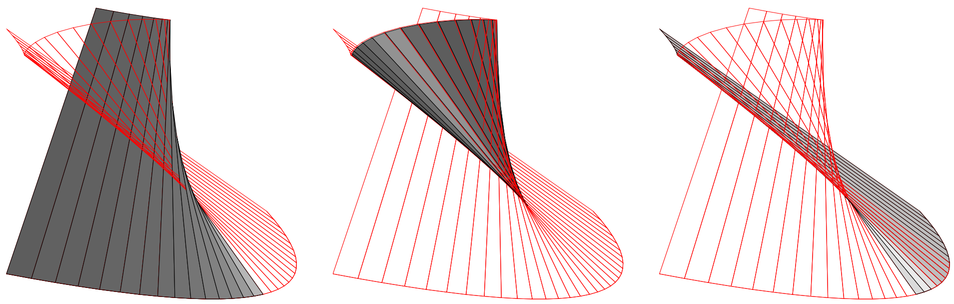
In the last chapter we introduced a novel shape representation called the extended Euclidean distance transform. We now build on that work by showing how this representation can be used as the basis for an algorithm to perform skeletonisation. In this chapter a brief description of the principle behind the algorithm is given followed by an outline of the specific implementation details, examples of output, and discussion of results. It is found that this technique gives a skeleton which is of similar quality to those produced using the standard distance transform approaches but which captures more symmetries of a shape than can be achieved using these methods.

6.2 Principle behind the algorithm

In the previous chapter it was determined that skeletal points correspond to multi-local singularities on the discriminant where smooth portions of its surface appear to "intersect" one another. The goal of the algorithm therefore is firstly to compute the discriminant and then to detect these intersections.

A natural decomposition of this problem is as follows: firstly, to dissect the discriminant into separate *single valued* sheets, [1] (a single valued sheet is one where only one point on the sheet is above any one point on the plane).

Secondly, to compute the distance from the boundary to every point on these sheets. Thirdly, to compare sheets to see if any occupy the same point over the plane and have the same distance from the boundary at that point. If this is the case then we can conclude that they intersect and their points of intersection can be marked as skeletal points. The key question in this approach is how can the discriminant be dissected into separate single valued sheets? This can be achieved by constructing sheets from curve segments bounded by two consecutive maxima of curvature and their corresponding evolutes.



(a) (b) (c)

The discriminant of a parabola can be dissected into three separate sheets. These are of two types, pre-evolute sheets as in figure 6.1a and figure 6.1c and post-evolute sheets as in figure 6.1b.

Figure 6.1: The dissected discriminant of a parabola.

Figure 6.1 shows how the discriminant can be thought to consist of separate sheets. Figure 6.1 shows a parabola which has a maximum of curvature in the middle which bisects the curve. Figure 6.1a shows one of the three separate sheets formed on the discriminant of the parabola. It consists of all inclined normals originating from the left half of the parabola and terminating at the first "crease" or local singularity they encounter on the discriminant. Figure 6.1c shows another sheet which due to the symmetry of the parabola can be considered a mirror image of the first. This consists of all

[1] This also allows us to solve the problem encountered by the wave propagation algorithm of distinguishing between multi-local and intrinsic singularities

inclined normals originating from the right half of the parabola and again terminating at the first "crease" or local singularity they encounter. It is the intersection of these two sheets projected down onto the plane which forms the skeleton of the parabola. Figure 6.1b shows the third sheet of the parabola discriminant. This consists of all normals which have passed beyond the "crease" in the discriminant. In the case of the parabola this does not intersect with any other sheet and therefore does not contribute to a skeletal branch. If a single sheet was constructed from the whole curve then the portions in figure 6.1a and figure 6.1c would together form a self intersecting sheet.

6.3 Implementation of the algorithm

The process of going from raw image to skeletal description involves three main stages: preprocessing, computation of discriminant and detection of singularities on the discriminant.

6.3.1 Preprocessing

The input to the algorithm is a greylevel image. Edge segments are found using a Canny edge detector and the edge segments from this are chained together using a simple linker. These edge chains are then approximated by cubic B-Splines. These preprocessing stages are very similar to those found in the work of Saint-Marc and Medioni [85] on symmetry detection and Chong *et al* [31] on fingerprint data compression. The output from the preprocessing stage is a list of separate spline chains. Each chain consists of the coordinates of the spline's control polygon vertices as floating point numbers and additional information such as if the curve is open or closed.

6.3.2 Computation of discriminant sheets

This part of the algorithm computes the extent of the separate discriminant sheets over the plane. A pseudo code listing is given in figure 6.2. The inputs to this stage are the separate spline chains. The output is a list of blocks of pixels which lie under particular sheets.

Firstly, the curvature κ along the splines is computed, defined by the following equation:

```
For each spline chain{
    Compute curvature at sample intervals.
    Compute maxima of curvature along chain.
    Compute evolute for spline chain.
    Break chain into segments between maxima of
    curvature and spline boundaries.
    For each segment{
        Create a polygon consisting of the spline boundary
        within the segement and its corresponding evolute.
        Call a scanline conversion algorithm to find what
        pixels lie within the polygon.
        Output segment number and list of pixel coordinates.
    }
}
```

Figure 6.2: Pseudo code for computation of discriminant sheets

$$\kappa = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{3/2}}$$

First and second derivatives of the splines required for this calculation can be obtained by differentiating the polynomials which make up the spline basis matrix to obtain new matrices. Maxima of curvature are found by a simple search and then the evolute E is computed which is defined as:

$$E : \left(x - \frac{y'(x'^2 + y'^2)}{x'y'' - x''y'}, y + \frac{x'(x'^2 + y'^2)}{x'y'' - x''y'} \right)$$

The curve is then broken into segments between maxima of curvature and spline boundaries. This ensures that no sheet can overlap itself. For each of these sheets a polygon with floating point vertices is created consisting of the spline boundary segment and its corresponding evolute. These are two open curves so the polygon is closed by defining sides between their two end points. A scanline conversion algorithm due to Hackbert [43] is then called. Scanline conversion is a standard computer graphics technique which when given an arbitrary polygon returns the pixels that lie within the polygon. This is done in order to register the sheets to the underlying pixel grid.

This means that when distances on separate sheets are compared they are sampled at precisely the same point over the plane for each sheet.

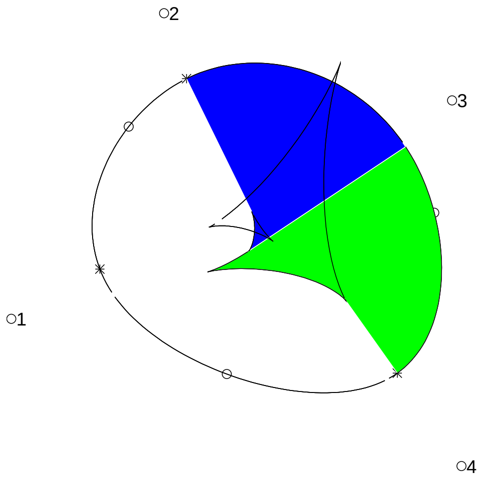


Figure 6.3: Dissection of discriminant
Simple spline object (round shape) with control polygon denoted by numbers. Within the shape we see the evolute and one section of discriminant. The discriminant section is shaded both black and grey as there is a spline boundary present. The section is bounded by positive curvature extrema (asterisks on the curve boundary) and evolute.

Figure 6.3 depicts a very simple shape to explain how the discriminant is dissected. The numbers denote the control polygon of the round spline boundary. The evolute is the cusped curve inside the shape. One sheet of the discriminant is shaded in between two consecutive positive curvature extrema (asterisks) and the evolute. The discriminant sheet comprises two shaded areas because there is a spline boundary between the extrema. There are two further partially overlapping segments to be shaded.

6.3.3 Detection of singularities on the discriminant

The preceding stage of the algorithm gave us separate smooth sheets of the discriminant. These consisted of a number of spline segments each with an associated block of pixels. We now have to find where these sheets intersect and to mark these intersections as skeletal points. The basic requirement is to compute the distance of every pixel from its associated spline segment and then to search to see if any other segment has a pixel of the same coordinates and of the same distance from that spline segment.

This part of the algorithm could be very inefficient in terms of time and memory if care were not taken in its implementation. The memory requirement varies over the image with some pixels having many sheets and others having few. The distances from pixels to spline segments are computed using a root finding algorithm due to Schieder, full details of which are given in [43]. This is rather slow and so should only be applied when necessary. Large portions of typical images only have a single sheet over them and so distance calculation is unnecessary.

It was decided that an efficient solution would be to use an insertion sort implemented with linked lists [54]. A pseudo code listing of this part of the algorithm is given in figure 6.4. Firstly an array of pointers to list records is created which is the size of the image, *i.e.* there is a pointer for each pixel. In the beginning this array is empty but as the algorithm progresses lists of arbitrary size can be hung from each pixel pointer. Each pixel for each sheet is considered in turn. A list record is created. If there is no list at this pixel location the record is inserted to start a list but the distance to the spline segment is *not* computed as it may not be needed. Only when a second record is added is the distance for the first record computed thus saving much computation time. The record is added to the list using an insertion sort based on the distances contained in each record. If adjacent records have distances within a certain threshold, typically 1 to 2 pixel widths, then the pixel is signalled as a skeletal point. This avoids the need for a further search through the lists for intersections. As large portions of typical images only have one or a low number of sheets covering them, it was thought that more efficient sorts [87] were unnecessary.

6.3.4 Experimental results

Figure 6.5 gives a comparison of the skeletons produced by the extended distance transform and the standard distance transform approach. The skeleton of a tee shape in figure 6.5a produced by the extended distance transform can be seen to make more symmetries explicit than the symmetric axis transform (SAT) in figure 6.5b derived from the standard distance transform. In particular it depicts more of the vertical axis of symmetry than the SAT and more of the diagonal symmetries emanating from corners. It also marks the symmetry between the base and the top of the tee which is entirely omitted in the SAT. However the new skeleton does not reach to the corners of the shape and there appear to be additional branches near the cross of the tee which at first do not appear to have any intuitively obvious cause. These artifacts arise because the discontinuous corners of the tee shape are represented by smooth spline sections of finite radii. At convex corners the skeleton can only approach as near to the border as the cusp on the evolute. The concave corners interact with each other and straight edge segments to produce skeletal branches which are long compared with the smooth corner sections which contribute to them.

Figure 6.6 shows the result of applying the extended distance transform to the skeletonisation of a spanner and gives comparisons with the wave propagation and standard distance transform approaches. In addition to the standard SAT branches running inside the spanner, figure 6.6c, other symmetries are made explicit. Extra branches are formed in the jaws and the end to end symmetry of the jaws is captured by a large lateral branch. The SAT is a subset of this skeleton and it is interesting to compare their local structure. Note the crossing of branches in the ends of the wrench which are a three branch junction in the SAT. The skeleton made using the extended distance transform is of a higher quality than that of the wave based approach of figure 6.6b. The wave based skeleton has disconnected branches projecting from the ends of the spanner but these are much clearer on the extended transform skeleton. We can also see that there is a skeletal branch at right angles to the main SAT skeleton in figure 6.6a. There are only a few isolated points of this branch in figure 6.6b; this is because the normals projecting from the concave jaws of the spanner spread out and miss each other.

Figure 6.7 displays the remarkably complex symmetries of the human hand. These arise partly from the multiple symmetries created between the fingers and thumb. The best example is perhaps the four "echoing" skeletal branches to the right of the thumb. This figure also highlights a limitation of the thresholding technique used to detect the intersection of the discriminant sheets. Boundary sections with normals pointing towards one another create thin skeletal branches as in the fingers. Boundary sections where normals point in a similar direction stay close for longer and produce broad skeletal branches. The arcs at the base of the fingers are examples of such boundaries. Figure 6.7a was obtained using the extended distance transform and figure 6.7b was obtained using the wave based approach. The extended distance transform skeleton is of a higher quality than the wave based skeleton. The wave based skeleton is incomplete, particularly in the region of the palm. The wave based skeleton also marks the evolute at the finger tips and this is avoided in the case of the extended distance transform.

The example of the screwdriver in figure 6.8 shows how the problem of broad skeletal branches can be alleviated to some extent. This is done by detecting a sign change or zero crossing in the subtraction of distances between sheets. Positive and negative results have been indicated by two different shades. The border between the two regions on the branches marks the zero crossing and the position of the skeleton to within one pixel. In this example the skeletal branches protruding from the handle appear to mark important geometric relationships. The branches extending laterally from the shaft seem less important and are due to minor fluctuations of the spline which are probably artifacts of the spline fitting process.

6.4 Discussion

6.4.1 Advantages and limitations

The quality of skeletons produced by the extended distance transform algorithm is good. This is partly due to the production of a dense and precise distance map made possible in part by the use of splines to approximate the boundary. The algorithm produces much more information than the SAT and it is not adversely affected by incomplete segmentation in the preprocessing stages. Comparisons with the SAT and other skeletonisation routines is difficult as most assume segmented binary data as input and this presupposes preprocessing of unspecified nature and parameters.

The use of splines does however impose limitations on the approach. The repeatable fitting of splines to data is a difficult problem. Residuals can lead the spline to oscillate and produce false skeletal branches. The detection and modelling of discontinuities is another important issue

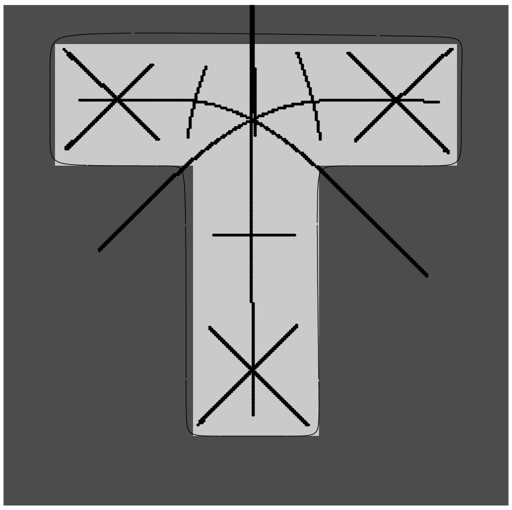
6.5 Summary

An algorithm to perform skeletonisation using an extended distance transform has been introduced based on the theoretical analysis of the previous chapter. Its fundamental principle is to compute the discriminant and then detect intersections of its sheets as these correspond to skeletal branches. Efficient implementation of the algorithm is required to make the algorithm viable in terms of memory and computation time. An insertion sort using linked lists was used as a solution to this problem. The use of splines to approximate the boundary provides a smooth curve as data input for the algorithm. However fitting of splines to noisy data and discontinuities can be problematic. It has been shown that the skeleton produced makes more symmetries of a shape explicit than the SAT derived using the standard distance transform. The use of a distance transform approach with its dense distance map provides a good quality skeleton.

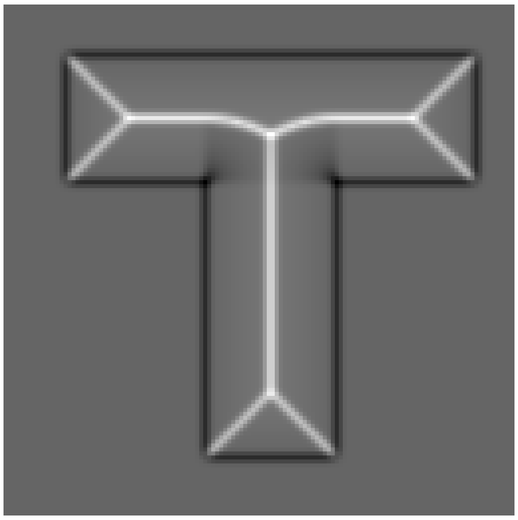
```
Allocate memory for image sized array of pointers to link lists
For each segment of spline{
    Input list of pixels within segment polygon.
    For each pixel in list{
        Perform insertion sort using linked lists.
    }
}

Insertion_sort(pixel_coordinates)
    Create a single linked list record.
    Insert pixel coordinates and spline segment into record.
    Check linked list image array at pixel coordinates.
    If(No list at this location)
      insert new record and return.
    elseif(Only one record in linked list)
        Compute distance of pixel to spline segment of existing record.
        Insert distance into existing record.
        Compute distance of pixel to spline segment of new record.
        Insert distance into new record.
        Traverse list performing insertion sort based on distance.
        If(neighbouring record distances < threshold)
            Mark Skeletal point in output image
    Else(More than one record in linked list)
        Compute distance of pixel to spline segment of new record.
        Insert distance into new record.
        Traverse list performing insertion sort based on distance.
        If(neighbouring record distances < threshold)
            Mark Skeletal point in output image
```

Figure 6.4: Pseudo code for detection of singularities on the discriminant



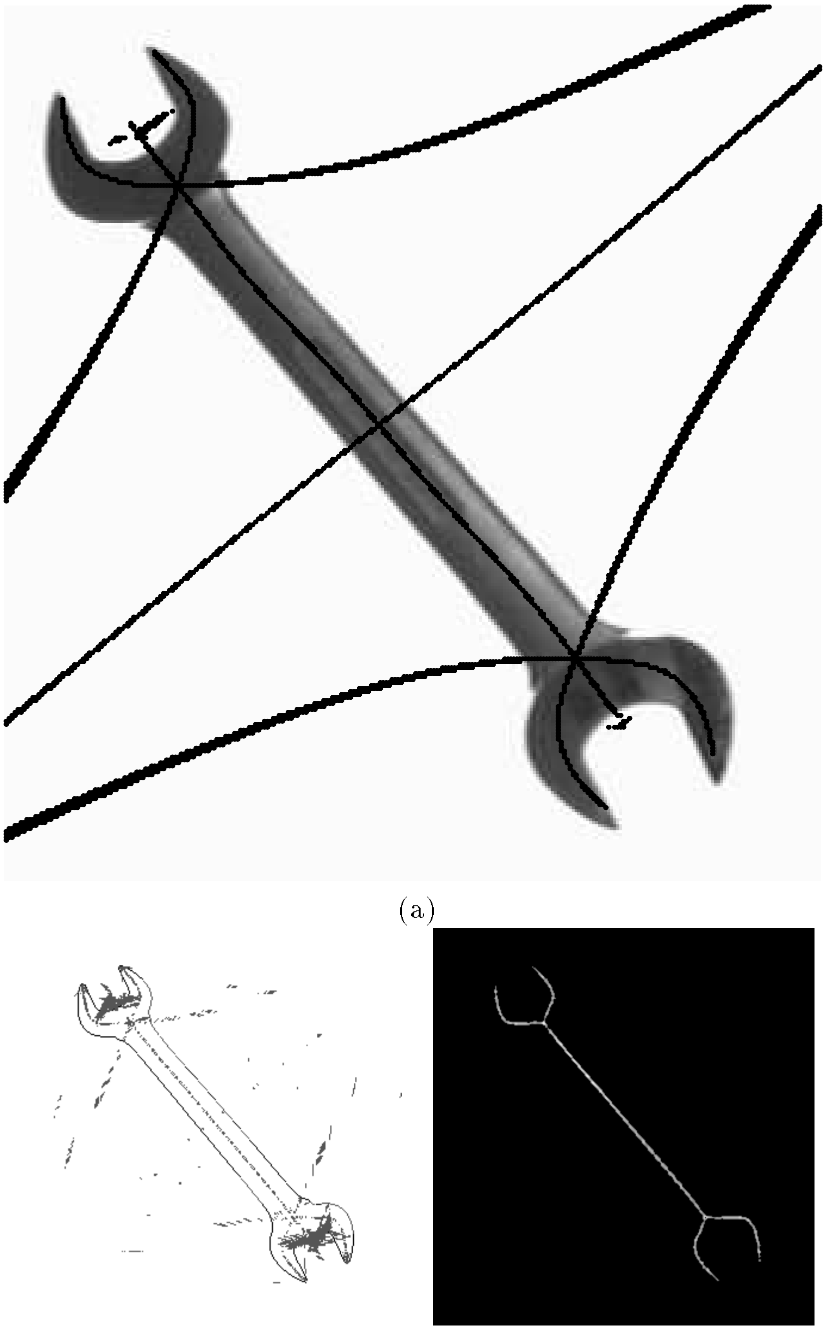
(a)



(b)

Figure 6.5a shows the skeleton for a Tee shape obtained using the extended distance transform whereas figure 6.5b shows the symmetric axis transform of the same shape obtained using the standard distance transform. We can see that many more symmetries are made explicit using the extended distance transform. Skeletal branches do not touch the shape boundaries in figure 6.5a, this is due to the fact that the discontinuous corners of the original shape are approximated by smooth splines.

Figure 6.5: The additional information aquired by the extended distance transform

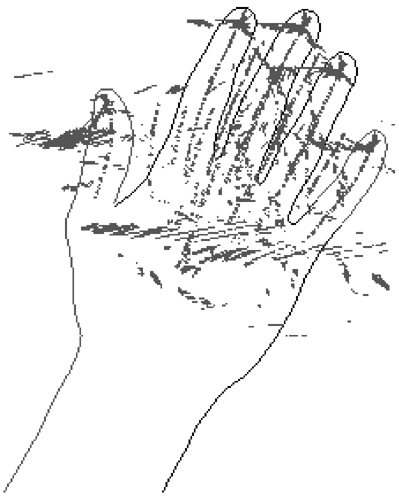


(a) (b) (c)

Figure 6.6a shows the skeleton of a spanner produced using the extended distance transform. This skeleton has the addition symmetries present in the wave based skeleton of figure 6.6b and is also of a quality similar to the skeleton produced using the distance transform and Petrou operator shown in figure 6.6c.

Figure 6.6: Skeleton of a spanner produced by the extended distance transform

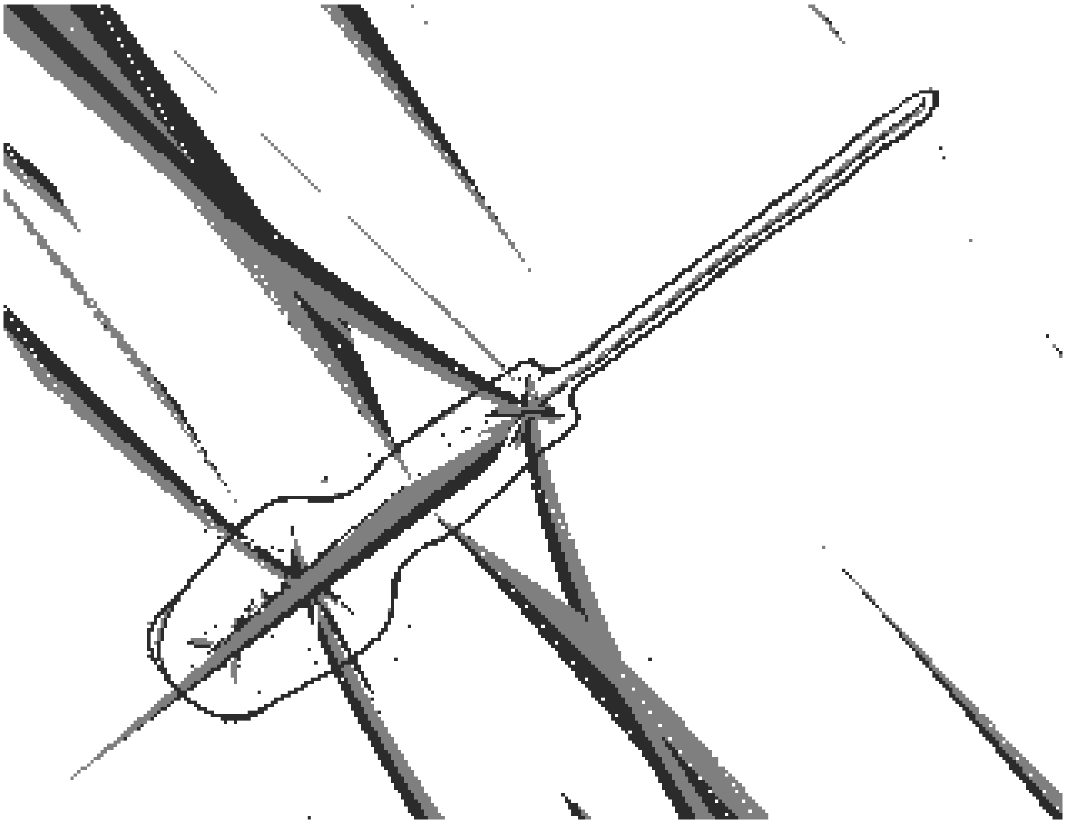
(a)



(b)

The skeleton for a hand obtained using the extended distance transform is shown in figure 6.7a. The complex symmetries of the hand are due in part to multiple symmetries between the fingers and thumb. Note the wide skeletal branches caused by intersecting sheets of the discriminant meeting at oblique angles. The skeleton obtained using the wave propagation algorithm is given in figure 6.7b. This skeleton is incomplete, particularly in the palm region. Notice the evolute markings at the finger tips which are absent for the case of the extended distance transform.

Figure 6.7: The symmetries of a hand using the extended distance transform



We can determine the centre of a broad skeletal branches by detecting the sign change in the subtraction of distances between sheets. Note that even fairly flat boundaries can produce skeletal branches some distance from the object.

Figure 6.8: The skeleton of a screwdriver using the extended distance transform

Chapter 7

Conclusions and further work

7.1 Introduction

The work in this thesis has addressed the problem of shape description in computer vision and in particular the computation of skeletal shape descriptors. This has entailed three main bodies of work. Firstly skeletonisation using the standard distance transform was studied using a new filter based approach. Secondly a wave propagation algorithm proposed by Brady and Scott was studied and the issues involved in its mapping to a MIMD parallel architecture were investigated. Thirdly an extended distance transform was introduced and it was demonstrated how this could be used to obtain rich skeletal descriptors of high quality. In this chapter a brief description of each piece of work is given along with the connections between them. Conclusions are then drawn for each piece of work individually and then as a whole. Further research directions are then presented.

7.2 Conclusions

7.2.1 Filter based approach

A standard method to perform skeletonisation is to use a distance transform. The symmetric axis transform corresponds to the ridges or "local maxima" of the distance transform. Skeletonisation can be performed by first computing the distance transform and then finding the "local maxima". It was decided to introduce a new method to find the "local maxima" by posing this task as a signal processing or filtering problem. The rationale for this approach came from the observation that the human visual system can distinguish the SAT clearly on a greylevel representation of the distance transform. This effect is called a Mach band which occurs due to lateral inhibition in the retina. A model for this process is the Marr-Hildreth operator and it was found that convolution of the distance transform with this operator picked out the SAT clearly. An attempt was then made to use a filter designed to detect a specific geometric feature on the distance transform namely a roof profile. The filter chosen was a Petrou operator and it was found that this produced much more clearly defined skeletons than the Marr-Hildreth operator.

Strong points of the technique are as follows. This approach produces skeletons of high quality. Filter parameters can be varied to alter the degree of smoothing and give a multi-scale shape description. Skeletal branches could be ranked according to the magnitude of operator output. This can be used at a single scale or to aid the problem of tracking branches through multiple spatial scales which has been identified in the literature as an important problem [76]. Subpixel interpolation of the filter output could be used to fix the position of skeletal branches more accurately. The raw filter output can be seen as a kind of *continuous* shape representation as opposed to a discrete *binary* representation which is the normal output format for skeletonisation algorithms.

However some drawbacks of the technique have also been identified. The pure roof profile is not a universal model for skeletal points on the distance transform and this fact can complicate operator output in some instances. The output of the operator can drop out at junctions in a similar way to edge detectors. As with all methods based on the Euclidean distance transform this approach has the intrinsic limitation that it can only be used to compute the SAT and not more general skeletal shape descriptors such as the SLS or the symmetry set. Nevertheless the filter based method of finding the "local maxima" on the distance transform represents a promising and novel approach to the problem of obtaining high quality skeletons.

7.2.2 Parallel wave propagation

Brady and Scott's parallel wave propagation algorithm was studied. This algorithm transcends the limitation of the standard distance transform approach in that it can compute more than just the symmetric axis transform. Brady and Scott implemented their algorithm on a simulator of the Connection Machine which is a computer with a massively parallel architecture.

The algorithm was mapped onto an array of transputers and the issues involved in this mapping were investigated. A Transputer array cannot provide the sheer computing power of the Connection Machine but it was found that an efficient implementation is possible if communication and synchronization overheads are minimised. The key to this goal was to divide the work up between processors on the basis of the distance of the wavefront from the shape boundary (which has been called *interval decomposition*) rather than to give each processor a portion of the boundary. It was found that the algorithm gives good results for thin elongated shapes but produced skeletons of poorer quality in general than the distance transform approach. The root cause of this poorer quality is the fact that the propagation process consists of separate "messages" sent down normals and does not constitute a continuous wavefront.

7.2.3 The Extended distance transform

The filtering approach using the standard distance transform produced good quality skeletons but could only compute the symmetric axis transform. The wave propagation approach produced more general skeletal descriptions than the symmetric axis transform but the quality of skeleton was lower in general. It was therefore decided that some new technique was required which would combine the strengths of both these approaches. Using standard techniques from singularity theory an *extended* Euclidean distance transform was developed. A theoretical analysis was performed that formally defined the extended Euclidean distance transform. An algorithm was proposed which used the extended Euclidean distance transform to perform skeletonisation.

In using this technique for skeletonisation the following points should be noted. The technique depends on the fitting of splines to pixel chains obtained from the Canny edge detector which can lead to a number of problems in the skeletonisation process. The spline can oscillate on straight edge segments thus giving false skeletal branches. If the error threshold of the fitting process is set too high then the spline can be a poor approximation of the shape boundary and the skeleton can drift from the expected position or have missing branches. If the original shape has sharp concave vertices then these should be detected and modeled explicitly. If not the small bends of finite radii that result produce large skeletal branches.

Care was needed to ensure that the computational costs in terms of time and memory were not too high. The choice of data structure was of key importance, and dynamic memory allocation using linked lists provided a simple solution greatly reducing memory requirements.

The fact that the extended distance transform enables more symmetries of a shape to be made explicit has been a key motivation for this work. This means that skeletal axes are less disrupted due to occlusion or imperfect segmentation than the SAT. However, there is a potentially negative side to the extraction of these additional symmetries in that along with additional useful and perceptually relevant skeletal axes many irrelevant axes can be produced. These axes come from two main sources. Firstly, they are generated by sheets of the extended distance transform which have passed beyond the standard distance transform. Secondly, when input is taken from greylevel images there will be sheets originating from any lines formed by the preprocessing algorithm such as from other objects, surface markings and shadows, *etc.* It is clear that the SAT is inadequate as a skeletal shape descriptor but it is not clear that the full generality of the symmetry set is required either. This puts even more emphasis on the ability to filter and impose structure on the output from skeletal shape descriptors. The new extended distance transform provides a new criterion for this which is whether the intersecting sheets of the discriminant which create a skeletal branch are pre-evolute or post-evolute. Most post-evolute branches appear to be perceptually relevant and many post evolute branches do not. However, there are exceptions such as the skeletal branch associated with the minor axis of an ellipse.

Given the above reservations, it is concluded that the extended Euclidean distance transform succeeds in its primary goal of producing skeletons which capture more of the symmetries of a shape than the standard distance transform approach but which are of a similar quality. In addition the extended Euclidean distance transform provides an elegant unification of previous work on skeletonisation. The relationship between the SAT, SLS, PISA, SS, standard distance transform and wave propagation methods can be established in terms of the extended Euclidean distance transform. It was found that the ridges on the standard distance transform are not "local maxima" as is suggested in the literature but are better described as singularities of the distance function.

7.3 Further work

In this section ways in which the work in this thesis could be improved or extended are discussed. This is done in turn for the three main areas of study.

7.3.1 Filter based approach

The Petrou operator detects instances of a specific geometric entity, *i.e.* a roof profile. The assumption is that the so called "local maxima" of the standard Euclidean distance transform are roof profiles. This is a reasonable assumption but they are only perfect roof profiles for thin parallel sided objects. It would be interesting to try to characterise what type of profile exists and their effect on the output of the operator.

This work has established that finding the "local maxima" can be viewed as a signal processing problem. It would be interesting to study the performance of other types of filter, for example a matched filter or morphological filter, and see how these relate to the Petrou operator.

7.3.2 Parallel wave propagation

It should be possible to improve the performance of this technique in various ways. More smoothing of the boundary would help, indeed the smoothed local symmetries approach includes a multi-scale analysis of the boundary called the curvature primal sketch. It would be possible to combine this with the region based skeleton detection thus providing a multi-scale skeleton as had been achieved for the SAT.

The voting scheme still uses discrete messages passing down normals as in the original Brady and Scott algorithm. For concave curved sections of shape boundaries the normals spread out and so many pixels can be missed. The algorithm could interpolate distances at these pixels and therefore not miss them out or spread votes over many pixels with some kind of weighting function. It might be better to use a continuous function such as the B-splines rather than the output from the Canny edge detector to approximate the boundary.

7.3.3 The Extended distance transform

Efficient discriminant computation

The method used here for computing the discriminant is accurate but is still rather inefficient. The main concern to date has been to prove the principle of the algorithm, however if this technique is to be viable a more efficient method for deriving the discriminant surfaces and detecting their intersection is needed. In this section we discuss two possible alternatives for these tasks.

One method would be to compute each smooth surface of the discriminant using separate *constrained distance transforms*. Constrained distance transforms are used in the literature for tasks such as path planning. These algorithms tend to be more complex than ordinary distance transforms particularly when the distance is propagated from many source pixels. This problem may be circumvented if we take advantage of certain constraints which apply in our case. If an evolute has an inflexion then the corresponding involute (the shape boundary in this case) will have a cusp. By the converse argument if the involute is smooth (which is one of our assumptions) then the evolute will never have an inflexion. So for any section of evolute running between two cusps the curvature of the evolute will have constant sign. This means we can use an ordinary distance transform if we use the following computational sleight of hand. We can initialise the pixels on the evolute with their corresponding radius of curvature. This would mean that the evolute would act as a kind of "wave guide" for the distance transform as it propagated from the source pixels.

For the second method, if we sample the boundary finely enough using closely spaced normals, the discriminant appears to be made up of many thin rectangular strips. Discriminant values within each strip can be interpolated from radii of curvature calculated at their vertices. This kind of process is common in graphics applications such as shading calculations. Indeed many workstations have dedicated hardware to perform this operation in real-time for thousands of quadrilaterals. In principle this would open the possibility for real-time computation of the discriminant for a moving or deforming object.

Detection of discriminant intersection

Currently detection of discriminant intersection points is due to a simple thresholding method. We have demonstrated that this can lead to broad skeletal branches in some cases. The detection of zero crossings in the subtraction of distances between sheets does offer a solution but there may be better techniques. Leymarie and Levine [57] use active contours to detect the SAT on the standard Euclidean distance transform. It may be possible to use this type of technique on the extended distance transform thus producing a continuous skeleton.

Appendix A

Singularities of distance functions

Consider the family of functions

$$\mathbb{R} \times \mathbb{R}^2 \xrightarrow{F} \mathbb{R} \times \mathbb{R}^2$$

$$t, x, y \longmapsto \| \gamma(t) - (x, y) \|^2, x, y$$

where $\gamma(t)$ is a unit speed parameterization of the curve, in particular we assume that γ is smooth *i.e* that $\gamma'(t)$ is never zero. A geometric interpretation of this mapping is shown in figure A.1. At every point (x, y) off the curve $\gamma(t)$ it plots the function $f(t, x, y) \; \| \; \gamma(t) - (x, y) \; \|^2$, thus building up a distanced squared function from the curve. (The only purpose of the <u>squared</u> distance is to make it easier to differentiate. So long as (x, y) is not on the curve, all results about the singularities of F apply equally well to $\| \gamma(t) - (x, y) \|, x, y$.)

The critical set

Write

$$f(t, x, y) = \| \gamma(t) - (x, y) \|^2 .$$

The **critical set** Σ of F is the set of points t, x, y with a singular Jacobian

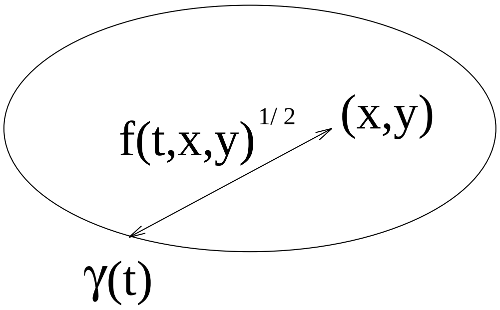


Figure A.1: distance squared function to a parameterized curve

$$\begin{vmatrix} \frac{\partial f}{\partial t} & \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0$$

i.e. $\frac{\partial f}{\partial t} = 0$:

$$\frac{\partial f}{\partial t} = 0 \qquad \underline{x} = (x, y)$$

If $T(t)$ is the unit tangent vector to the curve and . stands for scalar product,

$$\frac{\partial f}{\partial t} = 2(\gamma(t) - \underline{x}).T(t) \tag{A.1a}$$

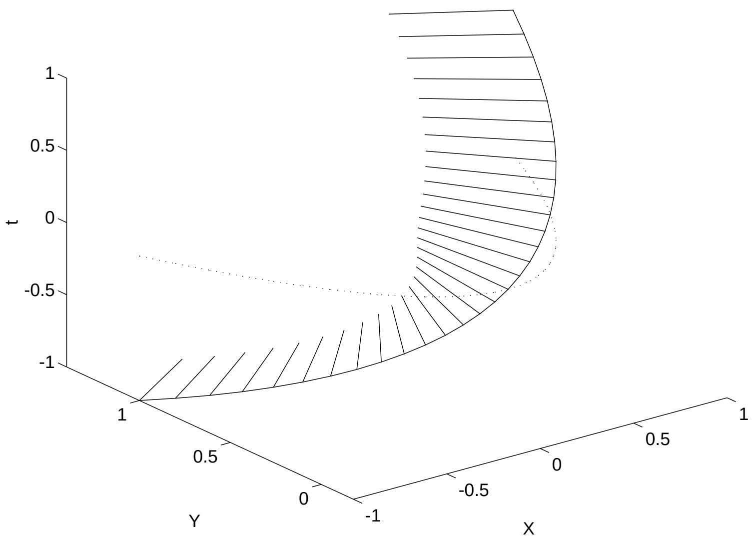
$$= 0 \tag{A.1b}$$

$$\Leftrightarrow \gamma - \underline{x} \quad \text{along normal} \tag{A.1c}$$

So the **critical set** Σ is the set of points (t, x, y) where (x, y) lies on the normal of the curve at t. These normals form a kind of ruled surface on the xy plane but we can separate them by varying t as a third dimension. Such a configuration is called a normal bundle and one for a parabola is shown in figure A.2.

Now

$$\Sigma = \left\{(t, x, y) \mid \frac{\partial f}{\partial t} = 0\right\}$$



The normals of the parabola have been separated by plotting the parameter t as a third dimension. The original parabola is displayed as a dotted line in the xy plane at $t = 0$ as an aid to interpretation.

Figure A.2: Normal bundle for a parabola

This will be smooth provided that at least one of

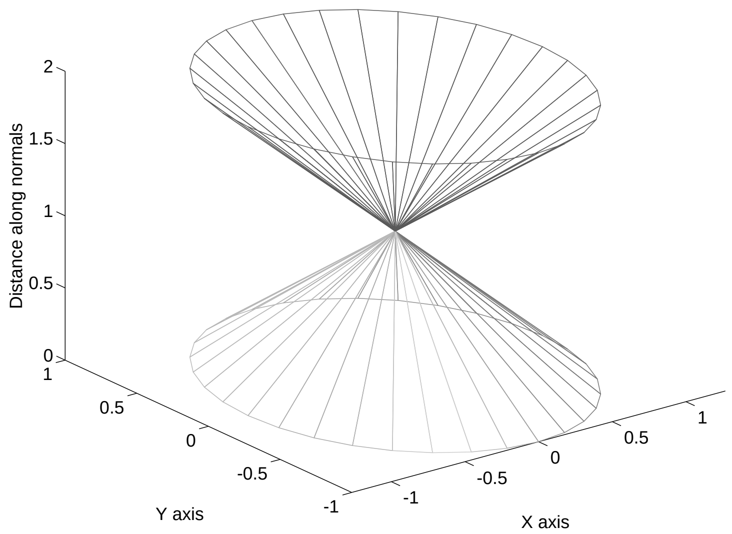
$$\frac{\partial^2 f}{\partial t^2}, \frac{\partial^2 f}{\partial x \partial t}, \frac{\partial^2 f}{\partial y \partial t}$$

is nonzero at each point of Σ.

The discriminant

The **critical locus** or **discriminant** Δ of F is $F(\Sigma)$ which is a subset of the target space $\mathbb{R} \times \mathbb{R}^2$. This contains the Euclidean distance transform if we use $\| \quad \|$ rather than $\| \quad \|^2$. The discriminant can be thought of as constructed from normals to the original curve rising at 45 degrees to the plain. To see this fix t; in Σ there is a line of points, namely the normal at $\gamma(t)$ but raised to height t. The image of this is a line at 45^o to the base.

Singularities occur when these normals cross one another, *e.g.* if the curve is a circle then the normals would form a right circular cone with the normals crossing at a single singularity at the apex (see figure A.3). A crucial point is to realise that the normals continue to rise and, in this case form another inverted cone above the first. In general the normals may pass through several singularities, the first singularities that each normal meets as it rises from the plane contribute to the Symmetric axis transform, other higher singularities they meet contribute to smoothed local symmetries and the symmetry set.



The discriminant can be thought of as normals to the original curve rising from the plane at 45 degrees. For a circle the discriminant looks like two cones with a solitary singularity at their common apex.

Figure A.3: Discriminant for a circle

So we have defined the **critical set** Σ and its **discriminant** Δ where :

$$\Sigma = \left\{(t, x, y) \mid \frac{\partial f}{\partial t} = 0\right\}$$

$$\Sigma = \left\{(t, x, y) \mid (\underline{x} - \gamma) \perp \gamma'(t)\right\}$$

$$\Delta = F(\Sigma)$$

We can now use singularity theory to define the type and position of singularities which occur on this surface. The **critical set** Σ is a smooth surface and so there can only be two types of singularity namely multi-local and local (or intrinsic) singularities. Multi-local singularities are caused by two points of Σ mapping to the same point of Δ. Local (or intrinsic) singularities of Δ are images of points $(t, x, y) \in \Sigma$ where $F|\Sigma$ has rank < 2 . This follows from the inverse function theorem.

Multi-local singularities

There exist points $(t_1, x_1, y_1), (t_2, x_2, y_2) \in \Sigma$ (distinct) with $F(t_1, x_1, y_1) = F(t_2, x_2, y_2)$

This means that:

$$\| \gamma(t_1) - (x_1, y_1) \| = \| \gamma(t_2) - (x_2, y_2) \|$$

and $x_1 = x_2$ and $y_1 = y_2$

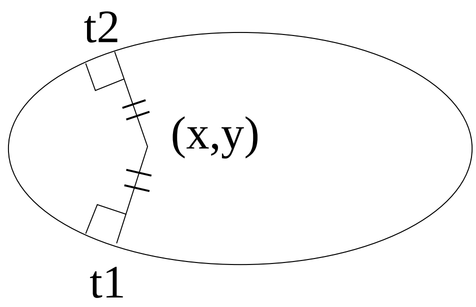


Figure A.4: Point (x,y) for which normal distances to the curve are equal

In figure A.4, obviously these points correspond to the centres of bi-tangent circles and thus are part of the symmetry set and also the symmetric axis transform (if these are the shortest normals to (x_1, y_1)).

Local singularities

To find these we need to determine when $F|\Sigma$ is singular (*i.e.* has a Jacobian of rank < 2).

$$\gamma(t) = (X(t), Y(t))$$

$$f(t,x,y) = (X-x)^2 + (Y-y)^2$$

$$\frac{\partial f}{\partial x} = -2(X-x)$$

$$\frac{\partial^2 f}{\partial x \partial t} = -2X' \qquad \frac{\partial^2 f}{\partial y \partial t} = -2Y'$$

The implicit function theorem tells us that we can use t, x or t, y as local coordinates on Σ since X', Y' are not both zero for any t.

Use t, x (so assume $f_{ty} \neq 0$):

$$t, x \xrightarrow{F|\Sigma} \Big(f(t,x,\eta(t,x)), x, \eta(t,x)\Big)$$

where $\eta(t,x)$ is a function of t, x from solving $\frac{\partial f}{\partial t} = 0$, so if $\frac{\partial f}{\partial t} = 0$ then

$$\frac{\partial f}{\partial t}(t,x,\eta(t,x)) \equiv 0$$

$$f_{tt} + f_{ty}\eta_t = 0 \qquad \text{chain rule}$$

$$f_{tx} + f_{ty}\eta_x = 0 \qquad \text{chain rule}$$

where $f_{ty} \neq 0$.

The Jacobian matrix of $F|\Sigma$ is

$$\begin{pmatrix} f_t + f_y\eta_t & f_x + f_y\eta_x \\ 0 & 1 \\ \eta_t & \eta_x \end{pmatrix}$$

f_t is zero because the point is on Σ and so the Jacobian is

$$\begin{aligned} \text{Rank2} \quad &\Leftrightarrow \quad \eta_t \neq 0 && \text{(A.2a)} \\ &\Leftrightarrow \quad f_{tt} \neq 0 && \text{(A.2b)} \end{aligned}$$

Now the Serret-Frenet formulae state that for a unit speed plane curve

$$T' = \kappa N$$

and

$$N' = -\kappa T$$

$$\begin{aligned} f &= (\gamma(t) - \underline{x})^2 && \text{(A.3a)} \\ f_t &= 2(\gamma(t) - \underline{x}).T && \text{(A.3b)} \\ f_{tt} &= 2(\gamma(t) - \underline{x}).\kappa N + 2(T.T) && \text{(A.3c)} \end{aligned}$$

$$f_t = f_{tt} = 0 \Leftrightarrow$$

$$\gamma - \underline{x} = \lambda N \qquad (f_t = 0)$$

$$2.\lambda N.\kappa N + 2 = 0 \qquad \lambda = \frac{-1}{\kappa}$$

$$\underline{x} = \gamma + \frac{1}{\kappa} N$$

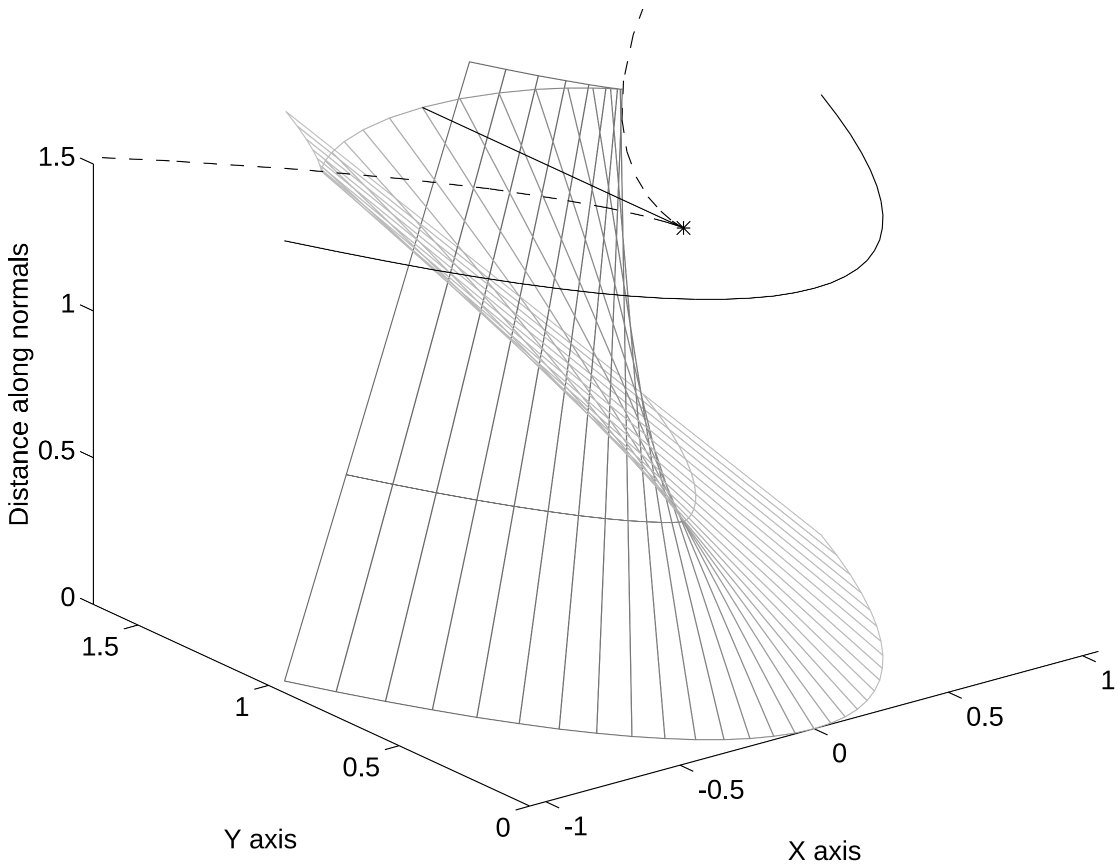
Hence $F(\Sigma)$ is singular when $(t, x, y) \epsilon \Sigma$ and (x, y) is at the centre of curvature of γ at t. This corresponds to the evolute which is the locus of centres of curvature (see figure A.5).

$F(t, x, y)$ only belongs to the Euclidean distance transform when the distance along the normals from (x, y) to $\gamma(t)$ is the shortest among all normals through (x, y). This never happens at the evolute except at cusps. To see that this is so consider the distance function from a fixed point (x, y) to a curve $\gamma(t)$ where (x, y) is on the evolute:

$$f(t) = \parallel \gamma(t) - \underline{x} \parallel^2$$

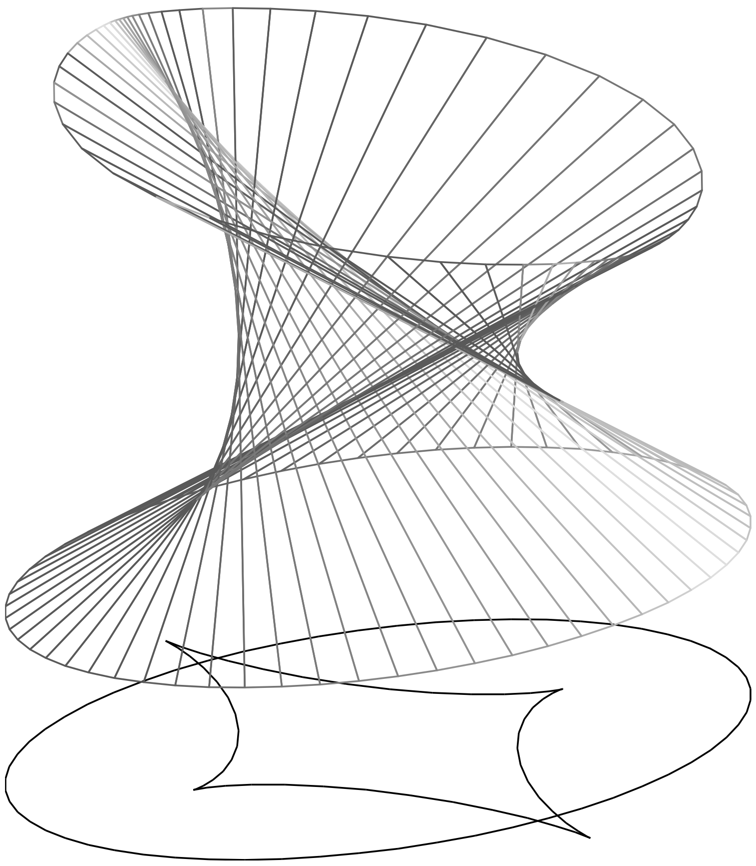
At the evolute $f' = f'' = 0$ so this is a point of inflection and hence there will always be a lower minimum somewhere else *unless* $f''' = 0$ which occurs at a vertex of the curve or cusp of the evolute.

Figure A.5 shows the discriminant of a parabola and its relationship to the symmetric axis and evolute. The original parabola has been drawn at the top of the discriminant along with the evolute (dashed line) and symmetric axis (straight solid line). The discriminant can be seen to undergo a "swallow tail" transition. Most of the length of the symmetric axis corresponds to multi-local singularities of the discriminant except for the end point which is a local singularity at the cusp of the evolute. The evolute itself corresponds to local singularities on the discriminant which form the cusps of the "swallow tail". Figure A.6 shows the discriminant of an ellipse which exibits even more complexity.



The discriminant for a parabola shows a "swallow tail" surface. Above the discriminant is drawn the parabola complete with evolute (dashed line) and symmetric axis (straight solid line). The central symmetric axis corresponds to points of multi-local symmetry on the discriminant, the evolute corresponds to intrinsic singularities which are the cusps of the "swallow tail". The only intrinsic singularity to have a minimum normal distance and therefore to be on the symmetric axis is that corresponding to the cusp on the evolute (asterisk).

Figure A.5: Discriminant for a parabola



The discriminant for an ellipse shows even more complexity. Below the discriminant is drawn the ellipse with its star shaped evolute.

Figure A.6: Discriminant for an ellipse

Bibliography

[1] J. Aloimonos and A. Bandyopadhyay. Active vision. In *First International Conference on Computer Vision*, pages 35–54, 1987.

[2] C. Arcelli, L. P. Cordella, and S. Levialdi. From local maxima to connected skeletons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3:134–143, 1981.

[3] C. Arcelli and G. S. di Baja. Finding local maxima in a pseudo-Euclidean distance transform. *Computer Graphics Vision and Image Processing*, 43:361–367, 1988.

[4] C. Arcelli and G. Sanniti di Baja. A width-independent fast thinning algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(4):463–474, July 1985.

[5] C. Arcelli and G. Sanniti di Baja. A one-pass two-operation process to detect the skeletal pixels on the 4-distance transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(4):411–414, April 1989.

[6] V.I. Arnol'd. Wave front evolution and the equivalent morse lemma. *Communications on pure and applied mathematics*, 29:557–582, 1976.

[7] H. Asada and M. Brady. The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):2–14, January 1986.

[8] B. B. Baker and E. T. Copson. *Mathematical theory of Huygens's principle*. Oxford, second edition, 1950.

[9] D. H. Ballard. Animate vision. *Artificial Intelligence Journal*, 48:57–86, 1991.

[10] D. H. Ballard and J. Sklansky. Tumour detection in chest radiograms. *Computers in Biomedical Research*, 6:299–321, August 1973.

[11] T.F. Banchoff and P.J. Giblin. Global theorems for symmetry sets of smooth curves and polygons in the plane. *Proceedings of the Royal Society of Edinburgh*, 106A:221–231, 1987.

[12] T. O. Binford. Visual perception by computer. In *IEEE conference on systems and cybernetics*, Miami, December 1971.

[13] A. Blake, M. Taylor, and A. Cox. Grasping visual symmetry. In *Fourth International conference on computer vision*, pages 724–733, 1993.

[14] A. Blake and A. Yuille. *Active Vision*. MIT Press, 1992.

[15] H. Blum. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form*. M.I.T Press, 1964.

[16] H. Blum. Biological shape and visual science (part 1). *Journal of Theoretical Biology*, 38(2):205–287, February 1973.

[17] H. Blum and R. N. Nagel. Shape description using weighted symmetric features. *Pattern Recognition*, 10:167–180, 1978.

[18] R. A. Boie and I. J. Cox. An analysis of camera noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(6):671–674, June 1992.

[19] G. Borgefors. Distance transforms in digital images. *Computer Graphics Vision and Image Processing*, 34:344–371, 1986.

[20] M. Brady. *Human and Machine Vision*, chapter Criteria for Representations of Shape. Academic Press, 1983.

[21] M. Brady and H. Asada. Smoothed local symmetries and their implementation. A.I.Memo 757, MIT, February 1984.

[22] M. Brady and G. Scott. Parallel algorithms for shape representation. In I.Page, editor, *Parallel Architecture and Computer vision*. Oxford university press, 1988.

[23] R. Brooks. Model-based three-dimensional interpretations of two-dimensional images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):140–149, March 1983.

[24] R. A. Brooks. Symbolic reasoning among 3-D and 2-D images. *Artificial Intelligence*, 17:285–348, 1981.

[25] J.W. Bruce. Wavefronts and parallels in Euclidean space. *Mathematical proceedings of the Cambridge philosophical society*, 93:323–333, 1983.

[26] J.W. Bruce and P.J. Giblin. Growth, motion and 1-parameter families of symmetry sets. *Proceedings of the Royal society of Edinburgh*, 104A:179–204, 1986.

[27] J.W Bruce and P.J. Giblin. *Curves and singularities.* Cambridge University Press, second edition, 1992.

[28] J.W. Bruce, P.J. Giblin, and C.G. Gibson. Symmetry sets. *Proceedings of the Royal society of Edinburgh*, 101A:163–186, 1985.

[29] J. F. Canny. Finding edges in images. Master's thesis, M.I.T, 1983.

[30] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, November 1986.

[31] M. M. S. Chong, R. K. L. Gay, H. N. Tan, and J. Liu. Automatic representation of fingerprints for data compression by B-spline functions. *Pattern Recognition*, 25(10):1199–1210, 1992.

[32] P. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.

[33] J.G. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20:847–856, 1980.

[34] L. S. Davis. A survey of edge detection techniques. *Computer Graphics and Image Processing*, 4:248–270, 1975.

[35] R. Deriche. Optimal edge detection using recursive filtering. In *First International Conference on Computer Vision*, 1987.

[36] A. R. Dill, M. D. Levine, and P. B. Noble. Multiple resolution skeletons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):495–504, July 1987.

[37] M. D. Edwards. *A review of MIMD Architectures for Image Processing*, chapter 5. Research Studies Press, 1986.

[38] A. Fiorentini, G. Baumgartner, S. Magnussen, P. H. Schiller, and J. P. Thomas. The perception of brightness and darkness relations to neuronal receptive fields. In L. Spillmann and J. S. Werner, editors, *Visual perception the neurophysiological foundations*, chapter 7, pages 129–161. Academic Press, 1990.

[39] A. Fiorentini and L. Maffei. Contrast in night vision. *Vision Research*, 13:73–80, 1973.

[40] M. M. Fleck. Local rotational symmetries. In *Computer Vision and Pattern Recognition*, pages 332–338, Miami, Florida, 1986.

[41] J. M. Gauch. *Multiresolution image shape description*. Springer-Verlag, 1993. ISBN 354097682-5.

[42] P.J. Giblin and the late S.A. Brassett. Local symmetry of plane curves. *American mathematical monthly*, 92(10):689–707, 1985.

[43] Andrew. S. Glassner. *Graphics Gems*. Academic Press, 1990. ISBN 0-12-286165-5.

[44] Khoros Group. *Khoros Programmers Manual*. University of New Mexico, 1992.

[45] G. Healey and R. Kondepudy. CCD camera calibration and noise estimation. In *International conference on computer vision and pattern recognition*, pages 90–95, Illinois, June 1992. IEEE.

[46] C. J. Hilditch. Linear skeletons from square cupboards. In B.Meltzer and D.Michie, editors, *Machine Intelligence IV*, chapter 22, pages 403–420. Elsevier,New York, 1969.

[47] W. D. Hillis. The Connection Machine. *Scientific American*, 256:108–115, June 1987.

[48] Hitachi UK Ltd. *Hitachi CCD camera model KP-180.*

[49] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand-eye coordination. In J. Illingworth, editor, *British Machine Vision Conference*, volume 2, pages 389–398. BMVC Press, 1993.

[50] D. H. Hubel. *Eye, Brain and Vision.* Scientific American library. W. H. Freeman, New York, 1987. ISBN 0-7167-5020-1.

[51] J. Illingworth and J. Kittler. A survey of the Hough transform. *CVGIP*, 44:87–116, 1988.

[52] INMOS. *Transputer Reference Manual*, 1988.

[53] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. In *First International conference on computer vision*, pages 259–268, 1987.

[54] J. F. Korsh and L. J. Garrett. *Data structures, algorithms and program style using C.* PWS-Kent, 1992. ISBN 0-87150-099-X.

[55] L. Lam, S. Lee, and C.Y. Suen. Thinning methodologies - a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(9):869–885, September 1992.

[56] F. Leymarie. Tracking and describing deformable objects using active contour models. Technical Report TR-CIM-90-9, McGill research centre for intelligent machines, February 1990.

[57] F. Leymarie and M.D.Levine. Simulating the grassfire transform using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(1):56–75, January 1992.

[58] M. Leyton. Symmetry-curvature duality. *Computer Graphics Vision and Image Processing*, 38:327–341, 1987.

[59] M. Leyton. A process-grammar for shape. *Artificial Intelligence*, 34:213–247, 1988.

[60] P. A. Lynn. *An introduction to the analysis and processing of signals*, chapter 10. Macmillan, 1990.

[61] M. Markus and B. Hess. Isotropic cellular automaton for modelling excitable media. *Nature*, 347:56–58, September 1990.

[62] D. Marr. *Vision: A computational investigation into the human representation and processing of visual information.* Freeman, 1982.

[63] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal society, London*, B270:187–217, 1980.

[64] D. Marr and H. K. Nishihara. Representation and recognition of the spatial organisation of three-dimensional shapes. *Proceedings of the Royal society London*, B200:269–294, 1978.

[65] S. Marshall. Review of shape coding techniques. *Image and Vision Computing*, 7(4):281–294, November 1989.

[66] U. Montanari. Continuous skeletons from digitized images. *Journal of the Association for Computing Machinery*, 16(4):534–549, October 1969.

[67] R. J. Morris. *Symmetry of curves and the geometry of surfaces.* PhD thesis, Department of Pure Mathematics, University of Liverpool, 1990.

[68] L. R. Nackman and S. M. Pizer. Three-dimensional shape description using the symmetric axis transform 1: Theory. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2):187–202, march 1985.

[69] T. Pavlidis. A review of algorithms for shape analysis. *Computer Graphics Vision and Image Processing*, 7:243–258, 1978.

[70] T. Pavlidis. Applications of splines to shape description. In *Visual Form analysis and recognition*, pages 431–442. Plenum, 1991.

[71] T. Pavlidis. Why progress in machine vision is so slow. *Pattern Recognition Letters*, 13:221–225, April 1992.

[72] M. Petrou. Detecting edges with intrinsic length-scale. In *Proceedings of the 6th Scandinavian conference on image analysis*, 1989.

[73] M. Petrou. Optimal convolution filters and an algorithm for the detection of linear features. Technical Report VSSP-TR-21/90, Department of Electronic and Electrical Engineering University of Surrey, Guilford, GU2 5XH, UK, 1990.

[74] M. Petrou and J. Kittler. On the optimal edge detector. In *Proceedings of the 4th Alvey vision conference*, 1988.

[75] M. Petrou and J. Kittler. Optimal edge detectors for ramp edges. Technical Report VSSP-TR-2/90, Department of Electronic and Electrical Engineering University of Surrey, Guilford, GU2 5XH, UK, 1990.

[76] S. M. Pizer, W. R. Oliver, and S. H. Bloomberg. Hierarchical shape description via the mutilresolution symmetric axis transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):505–511, July 1987.

[77] J. Ponce. On characterizing ribbons and finding skewed symmetries. *CVGIP*, 52:328–340, 1990.

[78] I. Ragnemalm. *The Euclidean distance transform.* PhD thesis, Department of Electrical Engineering, Linkoping University, 1993. Linkoping studies in science and technology. Dissertations No 304.

[79] S. E. Reichenbach, S. K. Park, and R. Narayanswamy. Characterizing digital image acquisition devices. *Optical Engineering*, 30(2):170–177, February 1991.

[80] R.W. Rodieck. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5:583–601, 1965.

[81] F. Rolland, J-M Chassery, and A. Montanvert. 3d medial surfaces and 3d skeletons. In *International Workshop on Visual Form*, Capri, Italy, May 1991.

[82] A. Rosenfeld. Axial representations of shape. *Computer Graphics Vision and Image Processing*, 33:156–173, 1986.

[83] A. Rosenfeld and J. L. Pfaltz. Sequential operations in digital picture processing. *Journal of the Association for Computing Machinery*, 1966.

[84] J. Ross and J.R. Johnstone. High frequency limitations on Mach bands. *Vision Research*, 21:1165–1167, 1981.

[85] P. Saint-Marc and G. Medioni. B-spline contour representation and symmetry detection. Technical Report IRIS#262, Institute for robotics and intelligent syutems, University of Southern California, Los Angeles, California 90089-0273, 1990.

[86] G. L. Scott, S. C. Turner, and A. Zisserman. Using a mixed wave/diffusion process to elicit the symmetry set. *IVC*, 7(1):63–70, February 1989.

[87] R. Sedgewick. *Algorithms in C.* Addison Wesley, 1990. ISBN 0-201-51425-7.

[88] J. Serra. *Image analysis and mathematical morphology.* Academic Press, 1982.

[89] J. Serra. *Image analysis and mathematical morphology: Volume 2.* Academic Press, 1988.

[90] L. A. Spacek. Edge detection and motion detection. *Image and Vision Computing*, 4(1):43–56, February 1986.

[91] D. Terzopoulos. Visual modelling. In *The second British machine vision conference*, Glasgow, September 1991.

[92] D. Vernon. Two-dimensional object recognition using partial contours. *IVC*, 1987.

[93] R. J. Watt. Space-scale analysis in the human primal sketch. In *Alvey Vision Conference*, pages 163–168, Cambridge, september 1987.

[94] M. W. Wright. Computation of smoothed local symmetries on a MIMD architecture. In *British Machine Vision Conference*, 1991.

[95] M. W. Wright. Computation of smoothed local symmetries on a MIMD architecture. Technical Report CUED/F-InfEng/Tr.68, Cambridge University Engineering Department, Cambridge, CB2 1PZ, 1991.

[96] M. W. Wright. Skeletonistion as model based feature detection. In *4th International Conference on Image Processing and its Applications*, Maastricht, Holland, April 1992. IEE.

[97] M. W. Wright, R. Cipolla, and P. J. Giblin. Skeletonisation using an extended Euclidean distance transform. In *British Machine Vision Conference*, pages 559–568, 1994.

[98] M. W. Wright, R. Cipolla, and P. J. Giblin. Skeletonisation using an extended Euclidean distance transform. *Image and Vision Computing*, 1995. In press.

[99] M. W. Wright and F. Fallside. Skeletonisation as model based feature detection. *IEE Proceedings-I*, 140(1), February 1993.

[100] Y. Xia. Skeletonization via the realization of the fire front's propagation and extinction in digital binary shapes. *PAMI*, 1989.

[101] Y. Yomdin. On the local structure of a generic central set. *Compositio Mathematica*, 43:225–238, 1981.

[102] S. K. Yuen. Shape from contour using symmetries. Cognitive Science Research Report 141, University of Sussex School of Cognitive Science, May 1989.

[103] A. Yuille and M. Leyton. 3d symmetry-curvature duality theorems. *Computer Graphics Vision and Image Processing*, 52:124–140, 1990.