On the complexity of curves and the representation of visual information

Benoit Dubuc

Department of Electrical Engineering
McGill University, Montréal

October 1995

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

ISBN 0-612-12358-8

ABSTRACT

Measures, dimensionality, and complexity are coupled notions, and the relationships between them are important practically as well as theoretically. Measures on curves might include their length, the number of components, or the area covered (such as the one for space-filling curves). In differential geometry, curves are characterized as mappings from an interval to the plane. In topology, curves are characterized as a Hausdorff space with certain countability properties. Neither of these definitions captures the role that curves play in vision, however, in which curves can denote simple objects (such as a straight line), or complicated objects (such as a jumble of string). The difference between these situations is in part a measure of their complexity, and in part a measure of their dimensionality. This thesis lays the groundwork for a formal theory of curves appropriate for computational vision in general, and for addressing problems such as separating straight lines from jumbles in particular.

Specifically, after presenting a parameterisation-free intermediate representation, the *Besicovitch tangent set*, we build a measure of complexity on it, the *complexity map*. This naturally decomposes into two distinct components, the *tangential and the normal complexity indexes*, which separate (0-dimensional) dust patterns, from (1-dimensional) contours, from (2-dimensional) textures. One main consequence of this analysis is that an intermediate representation, such as the *discrete tangent map*, is necessary for properly separating curve-like patterns that fill areas from those that extend mainly along their length. Most importantly, it provides the basis of a classification scheme for curve-like sets such as those encountered in edge/line detection.

RÉSUMÉ

Mesure, dimension et complexité sont des notions liées et leur apparentement s'avère important tant au niveau pratique que théorique. Les mesures pour les courbes du plan peuvent inclure leur longueur, le nombre de composantes, ou l'aire couverte. En géométrie différentielle, les courbes planes sont caractérisées comme étant une application de l'intervalle au plan. En topologie, elles sont caractérisées comme un espace de Hausdorff avec certaines propriétés de dénombrement. Aucune de celles-ci ne capture vraiment l'essence de ces courbes en vision cependant, c'est-à-dire qu'elles puissent représenter des objets simples, comme un segment de droite, ou des objets compliqués, comme un paquet de lignes tracées au hasard. La différence entre ces deux situations dépend à la fois d'une mesure de leur complexité, et de leur dimension. Cette thèse jette les bases pour une théorie formelle des courbes appropriée pour la vision par ordinateur en général, et pour des problèmes tels la séparation des ensembles de segments simples de ceux formés d'un paquet de lignes.

Après avoir présenté une représentation locale exempte de paramétrisation, l'*ensemble des tangentes de Besicovitch*, nous lui associons une mesure de complexité: la *carte de complexité*. Celle-ci, qui se divise en deux composantes distinctes, les *index de complexité normal et tangentiel*, est liée à une notion abstraite de la dimension, puisque c'est elle qui marque la distinction entre les poussières de courbes (0-dimensionnelles), les contours simples (unidimensionnels) et les textures (bidimensionnelles). La conséquence première de l'analyse présentée est qu'une représentation intermédiaire telle la *carte des tangentes* est nécessaire pour pouvoir distinguer correctement les ensembles qui couvrent une aire de ceux qui s'étendent sur leur longueur. De plus, elle procure les bases d'un procédé de classification pour les ensembles à structure de courbes tels ceux que nous rencontrons lors de la détection de contours.

ACKNOWLEDGEMENTS

This document is the result of many years of hard work, but also to an incredible supportive environment. During this long adventure, many people have contributed to making this piece of work what it is, and I am very grateful to all of them.

First, I want to thank my supervisor, Steve Zucker, for his endless support, for giving me the freedom to express my ideas, and for the drive to question what sometimes seems evident at first glance. His passion for computer vision was truly inspiring, and his input at all stages was critical to making the document what it is.

Secondly, I'd like to thank my dearest friend, Juan Stéphan, who has been supporting me 110% in this project. Not only has he helped me with many of the figures, but also took care of all the burden at home, so I could concentrate on my work as much as possible (I guess that it will now be time for me to participate in household activities again!). Sometimes, research leads to the greatest excitement, but sometimes it takes you to the worst moods. Juan never complained about that and was always there to encourage me to continue on and persevere: I will never be grateful enough for his support. I'd like to thank Carolyn Heltne, a special neighbor from whom I like the description of my thesis work to be the "ultimate in deconstructivism!"

Then, there are all the people from the lab: John Lloyd for his thesis advice and style discussions; Dominik Hartmann and Robert Lucyshyn for Mathematica advice; Peter Whaite for endless TEX advice, Lee Iverson for the invaluable tangent maps and L/L code, and for proofreading parts of the document; Pierre Breton for his inspiration and day-to-day advice; Mike Langer, for reading my original incomprehensible ideas, guiding some of this work to make it clearer and easier to understand; Yves Bérubé for thoroughly reading the introduction; Thomas Jelonek, James Elder, and Jonas August for all the involved discussions. I must also thank all the support staff at the Centre for Intelligent Machines, in particular: Jennifer Quinn for her day-to-day support, always there to resolve a new crisis; and der Mouse for his always sound technical advice. In addition, I cannot forget two good friends of mine who were around at the beginning but left for better worlds: Frédéric Leymarie and Frank DiGiuseppe. Thanks guys!

Three special people have contributed significantly to my confidence in the different mathematical aspects developed in this thesis. A special thanks to Allan Tannenbaum, for actually understanding my incomprehensible babble at the beginning. Thanks also to Jal Choksi from the McGill Math Department and to Peter Caines from Electrical Engineering.

I need to thank National Sciences and Engineering Research Council of Canada (NSERC), the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR), and the Stewart-Reid Foundation for their financial support. Thanks to Zilon for the great graffiti in the introduction; Jerome Holmes and Maurice Angers at McGill; Rashad Tawashi, at Université de Montréal, for his endless support, pushing me everytime we would meet to finish my thesis from very early on; Nicole from Saint-Henri and Doug Oram, a very good friend, who helped me sorts things out when all this stuff was in a highly confused state.

Finally, I would like to thank my parents for their support at all time. Both have always had faith in me, and did everything they could to support me at all time. I'd like to pay a special tribute to my father, Serge Dubuc, whom I admire very much, and who will always be an inspiration for me.

CONTRIBUTIONS

How should edge elements be organized so they can be grouped into curves? How can curves be segmented from texture flows? How can "dust-like" patterns be segmented from textures? What is the proper representation for early visual information? How do representations change as a function of scale? These questions are all examined by a novel approach based on tangential and normal dilations of a set. In this thesis we will be concerned with a generalization of geometric measure theory that is relevant to computer vision. It both provides a formal connection between previously unrelated problems, and an algorithm for solving them. In particular, the following points have been addressed:

- reflexion on the problems inherent to integration of local information;
- introduction of the tools from geometric measure theory to computational vision in the context of curve detection;
- characterization of the structure of the tangent maps through the tangent separation theorems;
- definition of a new intermediate representation, the complexity map, which splits both the normal and tangential components of complexity in a curve-like set;
- classification scheme for the type of patterns and a rule on the choice of representation subtending integration;
- justification for an oriented local representation at an early stage in curve detection.

TABLE OF CONTENTS

CHAPTER 1

Introduction and motivation

"How complex or simple a structure is depends critically upon the way we describe it. Most of the complex structures found in the world are enormously redundant, and we can use this redundancy to simplify their description. But to use it, to achieve the simplification, we must find the right representation" HERBERT A. SIMON [1968]

A classical problem in the design of a general purpose artificial vision system is the localization and description of image curves (edges or bars). For instance, imagine a dark cube against a white background. The task of early vision is to abstract a description of this cube sufficiently rich to enable its recognition, while segmenting it as a figure from the background. Such a description must certainly involve the bounding contour around this cube, and it is the task of boundary detection to recover this contour[1]. Complexity and dimension issues arise immediately. From an intuitive geometrical point of view, surfaces are boundaries of solids, lines are boundaries of surfaces and points are boundaries of lines, as was pointed out by the French mathematician Poincaré (1926). Because the cube may subtend a large visual angle covering an enormous number of pixels, "re-presenting" it by its edges reduces the amount of information tremendously while keeping the essence of the information about the object. Line drawings are another example of abstraction of information in which the essence of the scene is kept and reduced to its minimal expression. Look for instance at the graffiti by Zilon in Figs. 1.1 and 1.2: his drawings clearly illustrate

[1]Throughout this work, when referring to curve , edge or boundary detection, we imply the process just described here.

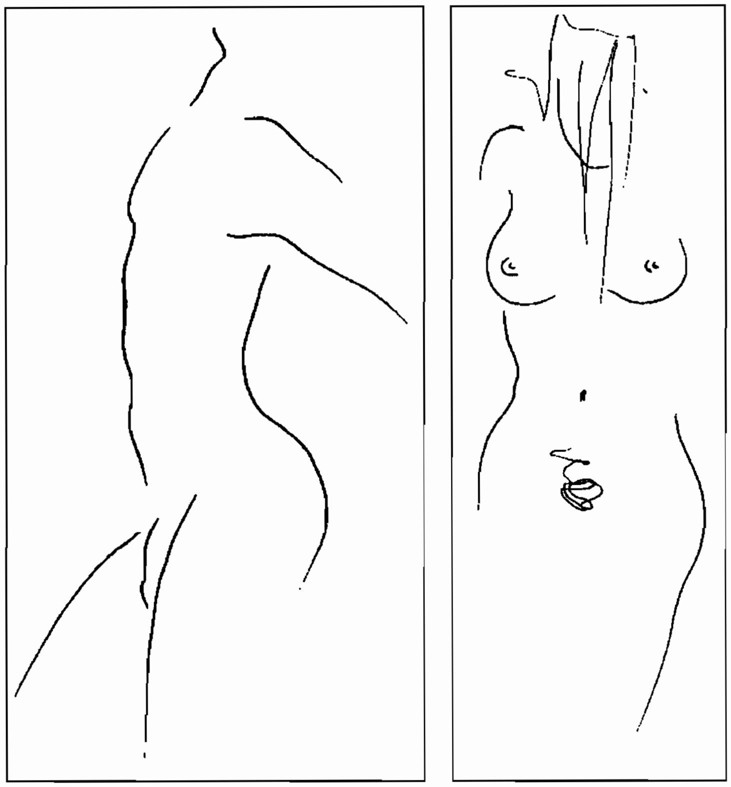


FIGURE 1.1. A man and a woman. Two graffiti by Zilon (one of Montréal's most prolific graffitist) illustrating the point that just a few lines can capture the essence of a scene.

the point that just a few lines are sufficient to provide the percept of a fairly complex scene. Thus, it made intuitive sense thirty years ago to begin to build computer vision systems by developing algorithms that would extract these edges and segment images automatically. At the same time, neurophysiology was providing important conceptual support for these techniques. The result is that now, among the stages for the processing of visual information, edge detection is one of the best understood.

1. Structure detection in early vision

Edge detection implies however a basic problem in perceptual grouping: once the local structure is established, the transition to global ones must be effected. To illustrate, imagine standing on an edge element in an unknown image, as in Fig. 1.4a or Fig. 1.4b. Is this edge element part of a curve, or perhaps part of a texture? If the former, which is the next element along the curve? If the pattern is a texture, is it a hair pattern (in which nearby elements are oriented similarly) or a spaghetti pattern

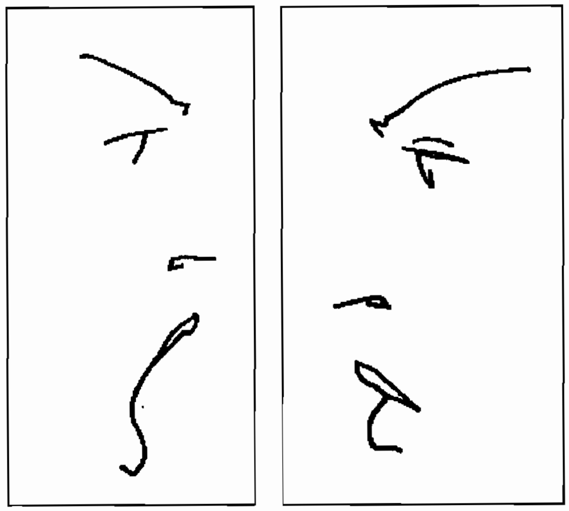
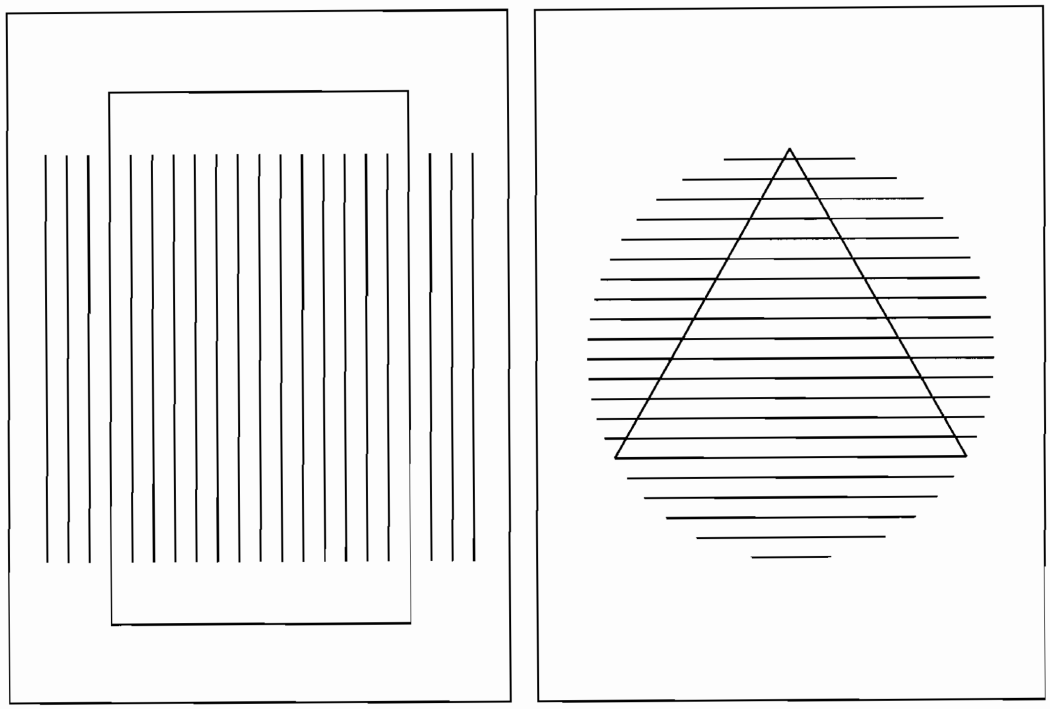


FIGURE 1.2. Shouting out! Two more graffiti by Zilon again illustrating the fact that just a few lines can capture the essence of a scene. This time the curve intersections are more severe. The number of lines stays small nevertheless inducing an elaborate percept.

(in which they are not)? These questions are in part about complexity since curves are "simpler" than textures, and in part about dimensionality, since some discontinuities are 0-D, curves are 1-D, and textures are 2-D. In this thesis a complexity measure that seeks to address these questions will be proposed. The ultimate goals are to show, in the context of curve detection, that the choice of representation and support for the grouping process is an important issue, and to provide a means of making an appropriate decision regarding the choice of representations.

Measures, dimensionality, and complexity are coupled concepts, and the relationships between them are important practically as well as theoretically. Measures for curves might include their length, the number of components (cardinality), or the area covered. However, the situation is more subtle than this, as is illustrated in several examples. The first example is taken from a classical demonstration by the Italian psychologist Kanizsa (1979). Figure. 1.3a, in which a pinstriped surface appears to be occluding a rectangle, demonstrates that curves, or sets of curves, can actually connote either the outline of objects (as in the rectangle) or surfaces (the pinstripes). Closer examination reveals that the rectangle is actually continuous through the surface, suggesting that visual inferences somehow group the pinstripes together and ignore the fact that the rectangle is the longest curve in the image. Neither the

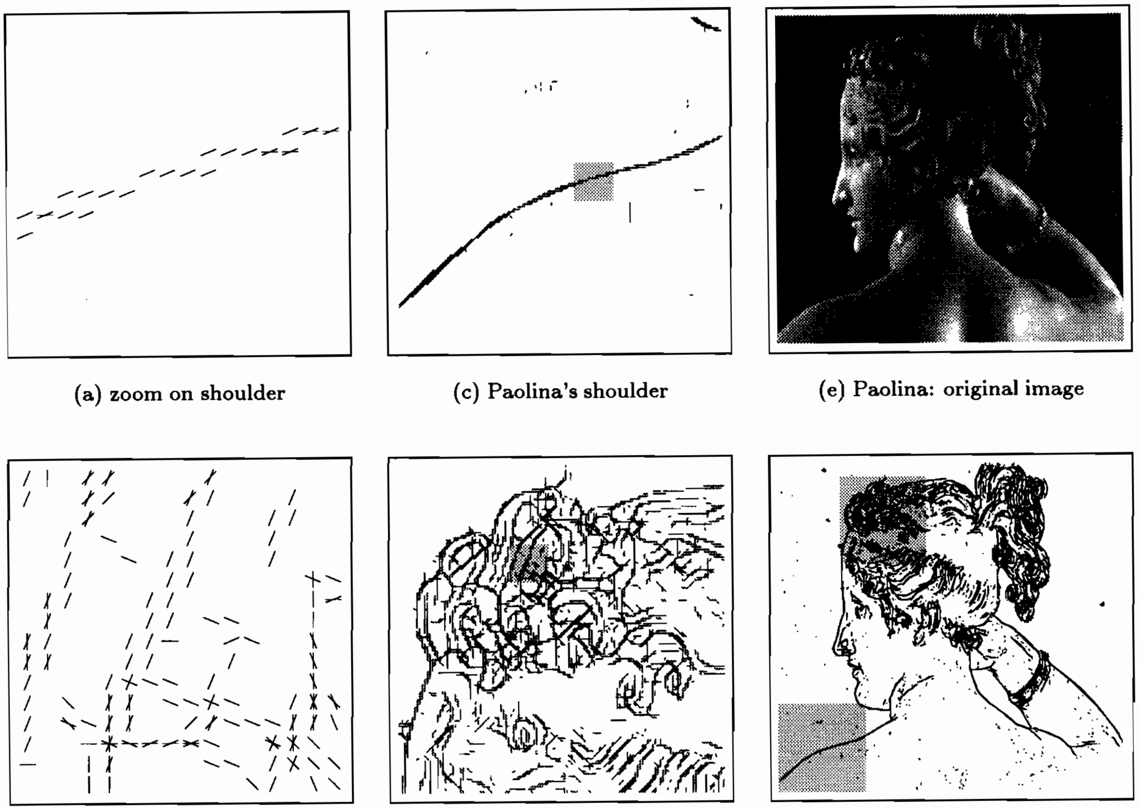


(a) (b)

FIGURE 1.3. Two Kanizsa patterns. (a) a pattern due to Kanizsa (1979) (b) circle/triangle pattern due to Galli & Zama (1931). Both these examples illustrate the need of using different representations for integration. Why does the texture "absorb" portions of the rectangle in (a) and of the triangle in (b)? In both cases the grating seems to predominate over the perception of the closed curve.

length of the curve, defining the rectangle, nor the number of components, defining the pinstripes, is dominant. Figure 1.3b, a triangle, is similarly camouflaged within a horizontal grating, and once more illustrates what Kanizsa (1979) called the "social conformity of a line".

Our second example is an image of a statue (which Pietro Perona refers to as "Paolina", see Fig. 1.4e). The result of an edge operator [Iverson & Zucker 1995] at a given scale is shown in Fig. 1.4f, and the problem of integrating local information raises the following observations. For the shoulder region (Fig. 1.4c), the underlying object is simple and a curve representation seems appropriate to group the edge elements. If we examine instead regions subtending part of the hair structure (Fig. 1.4d), then choosing a curve representation and walking along a hair would lead very quickly



(a) zoom on shoulder (c) Paolina's shoulder (e) Paolina: original image

(b) zoom on hair (d) Paolina's hair (f) Paolina: edge map

FIGURE 1.4. The subtlety of "walking through" a tangent map: (a & c) curves, (b & d) texture. Moving from right to left, the gray shaded areas are expanded to show the need for different representations to support the grouping of local edge elements. Integrating the responses of local edge detectors in the hair region is problematic. By what principles should the tangents be grouped?

to confusion, since it will be difficult to know on which part of the curve one is. A texture representation in this case seems more appropriate.

2. The problem

The leading question of this thesis is: given the output of edge/line detectors at a given scale and for a given resolution, how can these be grouped together? This clearly involves a local-to-global transition which has been described as "collecting individual edge points together to form continuous curves" [Cox et al. 1993]. However, it assumes that edge points should be grouped into curves; but local edges can arise from other image structures, such as a texture (hair, or field of grass, for instance).

We therefore question this assumption that edges should be grouped only into curves, and rather seek to determine which representation should be chosen for the grouping, together with the dimension of its support.

A second view of grouping is that it is a noise problem [Zucker 1993]. Since there are spurious responses from the local detectors, a global estimation procedure is necessary to eliminate them (Kalman filtering, for instance). Another is that it is simply an image-domain phenomenon linked to scale. Since larger operators have more image support, they should be less susceptible to such local variations. However, they are also more likely to average across features belonging to different objects.

This thesis questions all these assumptions and will try to shine light on the grouping and representation problems through arguments of complexity. The starting point of our investigation will be the search for edges, positive or negative contrast lines. The local structure will be given by the output of an edge/line detector [Iverson & Zucker 1995] sometimes followed by a few iterations of relaxation labeling [Zucker et al. 1977, Iverson 1993]. The process that will be described in this thesis in some cases would decide where those of Cox et al. (1993), David & Zucker (1990) or Mumford (1992) could be used, i.e. where a curve support is appropriate and over which extent. In the case when the support indicates a surface (textures for instance), then approaches such as [Zucker 1985] and [Kass & Witkin 1987] for integral curves, or [Malik & Perona 1990] for oriented texture characterization should rather be used.

3. Three approaches to complexity

But what exactly is complexity? Until now we have been rather vague in defining the concept and relied on its colloquial meaning. The Webster dictionary states that complexity is *the quality or state of being hard to separate, analyze or solve.* But how can one quantify complexity? How can it be defined as a quantitative measure assigned to a physical system or computation? How can we measure complexity for those patterns encountered in curve detection and not only determine that a certain representation is no longer appropriate, but also indicate a suitable one?

The following will open three parentheses about three different lines of thought on complexity. We will therefore quickly introduce: (i) the notion of computable number and algorithmic complexity; (ii) the notion of saliency map in edge grouping; (iii) the notion of confusion and the link between fractal dimension and complexity. These are in no way intended to be complete presentations, as they each portray informally a large body of research, but should rather provide the 'toile de fond' to what will follow.

3.1. Computable numbers and algorithmic complexity. Alan Turing's work on computable numbers [Turing 1936] is a key step toward the definition of algorithmic complexity (sometimes called Kolmogorov complexity). According to his theory, it is possible to characterize different numbers by the length of the program that is required to compute them. A number will be *computable* if there is a simple algorithm that gives the number even if the number is infinitely long [Pagels 1988]. A number will be called *non-computable* if the only algorithm we know is to explicitly specify the number itself within the program (the number is then said to be *incompressible*). For instance, let us look at the following sequences:

(i) 0.0101010101010101010101010101010101010101

(ii) 0.4285714285714285714285714285714285714285

(iii) 0.1234567891011121314151617181920212223242

(iv) 0.4142135623730950488016887242096980785696

(v) 0.8120961395426294503517263064533891042827

The first sequence is easy to describe in that it consists of 20 sets of 01 after the decimal point. The second sequence, although it looks more complicated shows periodicity very early on: it is simply the decimal expansion of 3 divided by 7. Therefore a program saying "Divide 3 by 7 and print the result" would be sufficient to describe the sequence. The third sequence is Champernowne's number C which is constructed by writing out the integers in order. Again a simple number since it could be generated by the program: "Print the integers in order after the decimal point". The fourth example displays the first digits in the numerical estimation of $\sqrt{2}-1$.

Thus far, there was a simple algorithm to give the number even if it was very long. The last example shows the output of a random number generator simulating the flipping of a 10-sided die. For truly random numbers there would be no better program for such a sequence than simply saying "Print the following: 0.8120961395426...".

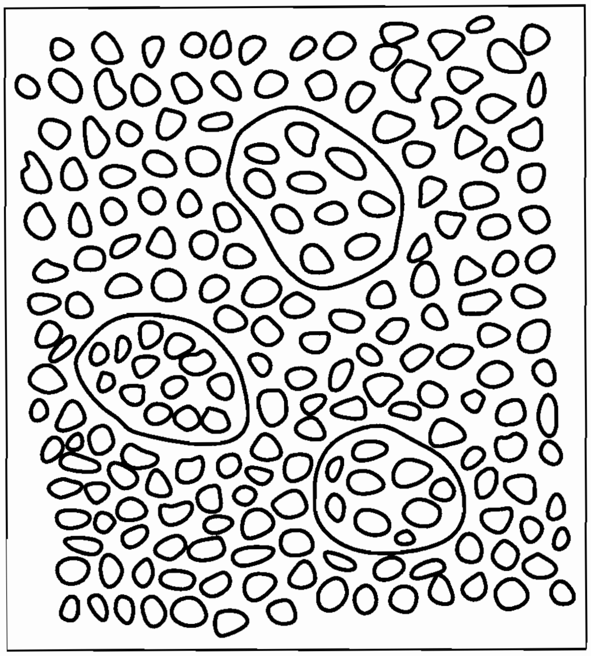
Thus the descriptive complexity of a truly random sequence is as long as the sequence itself.

Algorithmic complexity [Kolmogorov 1965, Kolmogorov 1987], is based on the definition of a *minimal program*: the shortest one following encryption as a string of integers. The algorithmic complexity of a number will therefore be the length of the minimal program required to compute it. This gives a quantitative measure to every number in the continuum, and results in the algorithmic definition of complexity. Algorithmic complexity can also be intuitively thought of as follows [Cover & Thomas 1991]: "if one person can describe a sequence to another person in such a manner as to lead unambiguously to a computation of that sequence in a finite amount of time, then the number of bits in that communication is an upper bound on the algorithmic complexity". Although the link to vision is not immediate, we will propose one in the next section.

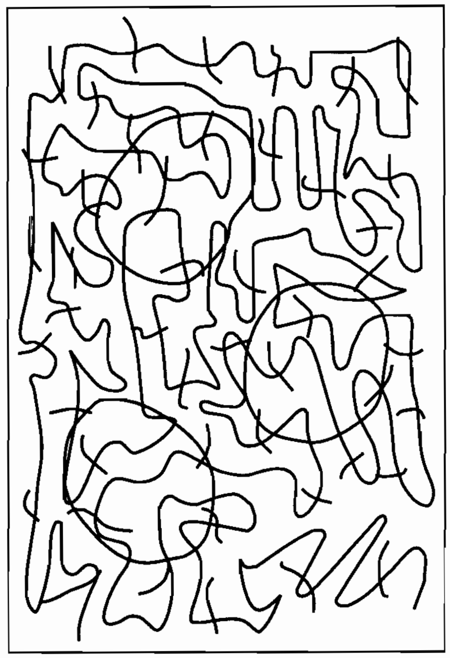
3.2. The saliency map for edge grouping. In an attempt to use the notion of complexity in the grouping of edge elements, Ullman defined a *saliency measure* and a *saliency map* to be able to represent what he calls 'globally conspicuous locations' (i.e. the structural saliency of structures as a whole) in the image [Ullman 1990, Sha'ashua & Ullman 1988]. He states that context is important when it comes to integrating information. Objects are rarely seen in isolation, so the question of how they can be segmented or selected from their environment arises. Using the edge map, he suggests building a saliency measure for each position in the image. This measure will favor image contours that are long and smooth; i.e., those that maximize length, but minimize curvature and curvature variation. Long curves, that are as straight as possible, and have the least number of gaps, will then be preferred. The two global properties considered in this construction are therefore length and shape (curvature and curvature variation).

As a demonstration, let us recall two of his figures [Ullman 1990]. These, which are reproduced in Fig. 1.5, show three large closed blob-like curves drawn within two different contexts. We notice that when the object is different from its context, it induces a "pop-out" effect [Treisman 1985][2]. This, in fact, was also present in

[2] In Chapter 6 we will come back to the fact that the circles in (b) tend to blend in more into their environment.



(a) Ullman pop-out



(b) Ullman hidden

FIGURE 1.5. Curve complexity and context. In these two figures taken from Ullman (1990), notice how in (a) the big circles pop-out as opposed to (b) where they tend to blend more into the scene.

the Kanizsa examples, but is illustrated at its best in Fig. 1.5a. When applied to the patterns, Ullman's saliency measure will be higher for the three blob-like closed curves in both cases: in (a) because these are the longest with the smallest average curvature, and in (b) since they have small curvature variation (as opposed to the long, wiggly curve) and are relatively long (as opposed to the distractors). Based on the saliency map, three blob-like figures would then be the detected salient structures on which subsequent processes such as segmentation and recognition can focus [Ullman 1990].

In this thesis, we will build a measure of complexity that in part resembles the saliency index described by Ullman. We will use it to define which representation is adequate for the grouping of edge elements, but will also argue that a measure that only considers length and shape is not sufficient. In order to show this, reconsider the Kanisza example. According to Ullman's theory, the most salient curves in Fig. 1.3a should be the two long sides of the rectangle, yet these are camouflaged by the grating. Furthermore, the Ullman saliency index says nothing about the whether the global structure is a curve or a texture. This index, preceptor to perceptual grouping, needs more than what we will call the "tangential component", it needs also to consider the context in a direction normal to the local edge elements.

Our belief is that, although global entities such as length and shape play an important factor for the grouping of local elements, complexity in the normal direction plays a role in determining which objects will "pop out" within the scene. While Fig. 1.5 illustrated the fact that objects that differ from their context pop out, the Kanizsa example brought a counterexample to Ullman's original saliency measure. Different representations are needed prior to the integration process.

3.3. Complexity, entropy and fractal dimension. In a series of articles, Mendès-France has developed mathematical tools to characterize the complexity of curves linked to models of physical systems. His contributions led to notions such as

- the temperature of a curve [DuPain et al. 1986]: the entropy of a plane curve is defined in terms of intersection points with a random line [Santaló 1976]. The Gibbs distribution which maximizes the entropy can be used to define the "temperature of a curve". At zero temperature, the curve reduces to a line segment, the more complicated the curve is, the higher its temperature;
- the confusion index [Mendès-France 1991*b*]: which is a measure of the uncertainty in the location of a point on the curve. As seen earlier in the hair texture, when the branches of a curve are tightly packed, it becomes impossible to decide which branch of a curve contains a given point;
- the Planck constant of a curve [Mendès-France 1991*b*]: calculating this resulted in saying that if the entropy is not zero, it is impossible to define with infinite precision both the location of a point on a curve and its tangent.

Much of Mendès-France's contributions in the area of curve complexity has been motivated by the work of the mathematician Benoit Mandelbrot [1982], who argues that many physical processes and structures are best modeled by a class of non-differentiable functions called *fractals*. The resulting theory, *fractal geometry*, has popularized the notion of fractal dimension which has roots in the formalization of general topology and geometric measure theory developed at the beginning of this century. In his classical paper, entitled "How long is the coast of Britain?", Mandelbrot

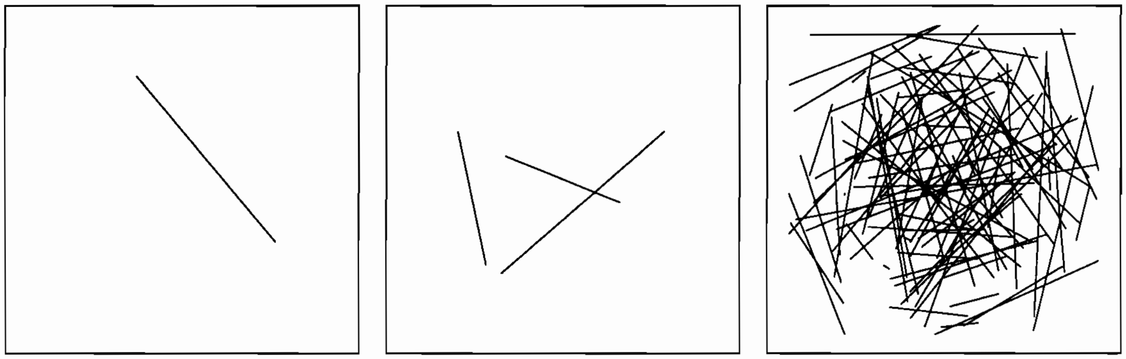
(1970) discussed the link between measure and dimension. His theory suggests that notions such as length, surface area, slope, and surface orientation ought to be abandoned entirely in favor of global measures of the processes' behavior over scale (rate of growth of measured arc length, for instance). In perception, Pentland [1984, 1985] has shown how fractals and the notion of fractal dimension can capture striking regularities in highly complex visual structures. Fractal dimension and complexity therefore seem related and Mendès-France (1991*a*) showed that indeed there is a relationship between the dimension of a curve and its entropy, a quantity known to correlate with the complexity of an object. One is therefore justified to draw a parallel between fractal dimension and complexity more deeply than by the enumeration of examples.

Although we will not use Mendès-France constructions directly, his work supports the nomenclature we will use in this thesis. In this thesis, we will work with curves as they occur in computational vision, and will thus choose definitions of dimension suited to the problem. Our measure of complexity will be linked to the types of calculations done in fractal analysis, but will differ in many respects. We will borrow from Mendès-France's analysis to justify its name: the *complexity map*. More formally, dimensional analysis classically takes place in the limit as scale tends to zero. Our analysis, relevant to computer vision, must take place at a finite scale. The phrase *complexity map* denotes this difference.

4. Complexity in early vision?

Not considering complexity could be disastrous in the design of a general computer vision system. Unfortunately, formally proving that complexity is necessary turns out to be a very difficult task. Most researchers assume that the only examples are "pathological" and would not occur in practice anyway. Our goal will be to show that confusion arises even with the simplest combination of building blocks. The following examples are chosen to highlight the issues involved and to show the interrelationships between the concepts of complexity and representation.

Example 1.1 (The loop and the spaghetti). Consider two images, one of a circle, and the other of a plate of spaghetti. Two tasks can be envisioned: to draw the curves and the other, to follow them. In order to draw the circle, the algorithm can



(a) one single line (b) three lines (c) one hundred lines

FIGURE 1.6. Line segments or texture? Three examples of sets composed of line segments. Notice how difficult it is to follow every path in (c).

be very simple: repeat n times: take 1 pixel step, rotate $2\pi/n$. In the case of the spaghetti pattern: repeat n times: take a long step, rotate random amount. Now, suppose one would like to reproduce the two patterns. In the first case the complexity is bounded, while in the second case we need remember the various random rotations.

Another issue is the one of following the path. In the case of the circle, the task is very easy. In the case of the spaghetti pattern, the task is much more complicated and confusing. The fact that there are many branching points makes the path not unique: there are many different paths. In fact it leads to a combinatorial explosion of possibilities. The "curve" representation then fails to be efficient. Following the pattern leads to integration but the representation that needs to be used varies from one case to the next.

Example 1.2 (Pick-up sticks). Suppose we are only working with line segments and we want to find out when the segment representation is no longer appropriate. Three instances are shown in Fig. 1.6. Now, let us make a parallel with the game called "Pick-Up Sticks"[3]. Complexity will be related to the difficulty of picking up a stick without moving the other sticks. In Fig. 1.6a, the task is trivial since there is only one segment. In Fig. 1.6b, it is a little more difficult, but still very easy. In Fig. 1.6c, the task is hard. Contrary to intuition, the task is not simply a matter of the number of line segments. Even with only two segments, the task can be hard if

[3]Actually in this example the length of the sticks is random.

the sticks are parallel and close one to another, or if they cross at a small angle. This is due in part to the width of our fingers. With tweezers, the class of "manageable configurations" becomes larger. Even other factors come into play: we could imagine sticks with variable lengths and widths. A stick that is too short, would be hard to pick up.

Example 1.3 (Frequency). An example taken from the representation of numbers states that, at first glance, it is very unlikely that somebody would be able to distinguish between the numbers 99999999999999 and 9999999999999 without serially scanning each digit. The same point can be made with grating patterns, suggesting that our organization of understanding for visual patterns is richer than the simple enumeration of the individual curves. In Fig. 1.7, we reconsider the Kanizsa pattern to determine whether the percept changes when a pinstripe is added or removed[4]. For instance, compare Fig. 1.3a and Fig. 1.7a: which has 20 and which has 21 lines? The distinction is very hard to make, but the one between Fig. 1.7a and Fig. 1.7b is immediate. Rounding the corners would be noticeable, however, even though these events are isolated.

Taken together, Examples 1.1 and 1.2 show how algorithmic complexity relates to computational vision. The first example stressed the complexity of communication of a pattern and of a visual task through the pattern following example. The second example introduced some of the central principles we will develop in this thesis. Most importantly, it showed how the class of all patterns could be partitioned through a complexity measure into equivalence classes of equally hard tasks. There are many games of Pick-Up Sticks that are equivalent in terms of difficulty – i.e., in terms of complexity – even though the particular arrangement of the sticks may differ enormously.

One goal of this thesis is to make clear that such a partitioning is necessary for the transition from local to global representations in computer vision. Complexity analysis is not, however, sufficient for all aspects of the local-global transition. Consider another local-global problem: connectedness. Fig. 1.8 clearly illustrates this

[4]Note that the placement of the figures is a deliberate strategy we adopted to make the task harder: having to flip from one page to another, helps to make our point unambiguous.

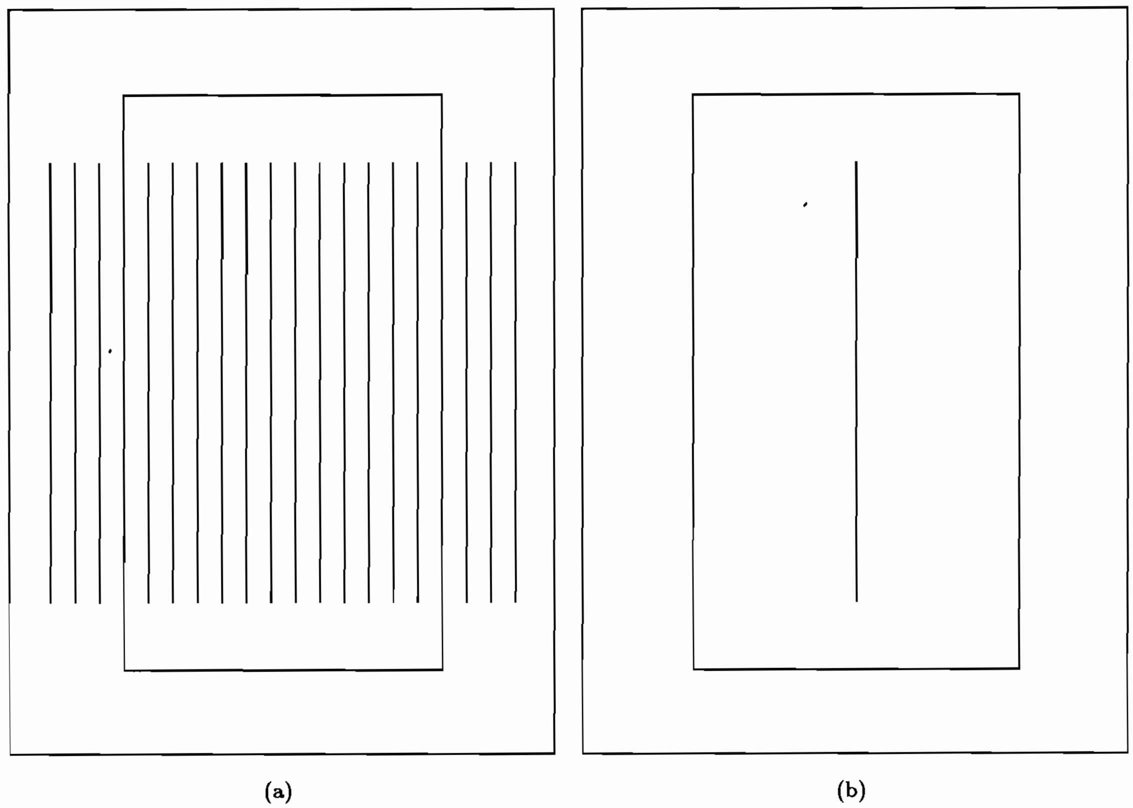
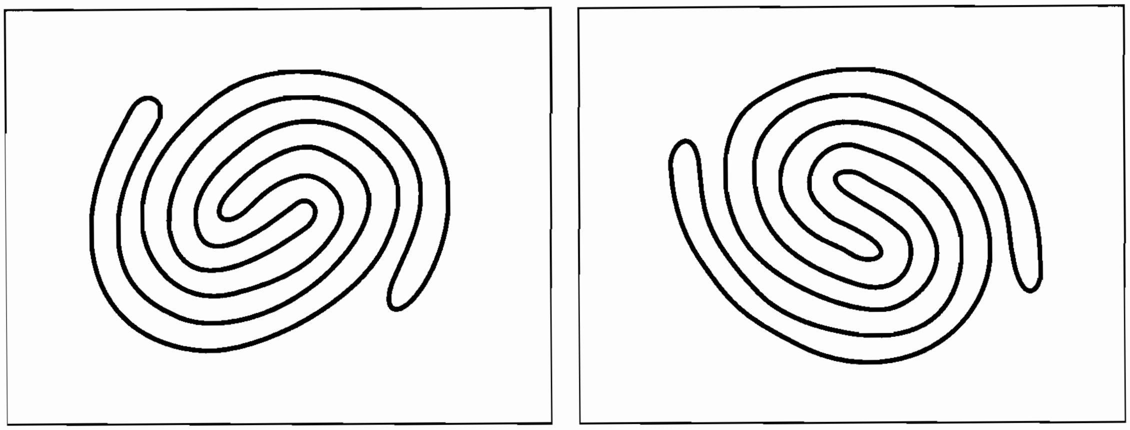


FIGURE 1.7. Kanizsa patterns again but this time with different grating frequencies. Notice how easy it is to tell these two apart but how difficult it is to differentiate between (a) and Fig. 1.3a.

point: it is very difficult to tell which of the two figures is connected and which is not. The explanation we propose for this, returning to the analogy of our Pick-Up Sticks example, is that the two figures belong to the same equivalence class; they constitute two equally hard tasks, and are thus very difficult to differentiate from each other. The further question – which is connected and which is not – will require other techniques for solving.

How can one relate algorithmic complexity to vision? Recalling the intuitive description of algorithmic complexity presented earlier, let us build the following experiment. Take three persons: Robert, Bruno, and Veronica. Bruno has a pattern that he shows to Veronica who then describes it to Robert. From this description, Robert reproduces the pattern. He then shows his reproduction to Veronica. If Veronica cannot distinguish the reproduced pattern from the original in a bounded amount



(a) (b)

FIGURE 1.8. Connected or not? This figure, taken from the classical work of Minsky & Papert (1972) illustrates well the local-global problem. The problem is to determine which of the two figures is connected and which is not. Indeed, one of these two figures is composed of only one curve, while the other has two.

of time, then we will say that the *representational complexity* is bounded by the length of the description. Representational complexity would thus build equivalence classes of patterns. Relating back to our Pick-Up Sticks example, within each equivalence class, the visual tasks are equally hard. The question now is how to build a measure of complexity that will define these equivalence classes of patterns, and then how to assess it.

Returning to our examples, the last one showed that, for some patterns, approximation is sufficient. For instance, in the case of grating patterns, the exact number of lines and their exact location might not be relevant. There is an obvious difference between a pattern with one segment and one with two segments provided the lines are (i) sufficiently long and (ii) sufficiently apart one from one another. In the grating part of the Kanizsa pattern, for instance, the difference between n and $n+1$ lines is irrelevant for our percept if n is large enough and if the lines are reasonably distributed. Compare the gratings between Fig. 1.3a and Fig. 1.7 and try to say at a quick glance the difference between the two. The difference between 20 and 200 lines would however be noticeable (provided again that they are reasonably positioned), again reinforcing the suggestion of equivalence classes of patterns that would be indexed by complexity measures.

The take-home message from all these examples is that, within a single representation and within a general setup, integration from local to global representations is intractable. On the other hand, integration is a key step in moving from an image-centered representation to an object-centered one. It has been done successfully in controlled environments where one knows the complexity of the scene *a priori*, or assumes the complexity to be within some bounds. The *blocks world* is an example of such a constrained environment. How does one build a general theory of integration for edge detection?

The core of this thesis will show that different actions in the integration stage should be taken depending on the context. First, we need to choose and define both an *intermediate representation* and a *complexity measure.* Given this representation, the local information can be integrated only if the underlying object is simple enough. Knowing this can make feasible the transition from early to intermediate levels of vision. If the complexity of an object at a certain scale and for a particular representation exceeds some value, then two choices could be made:

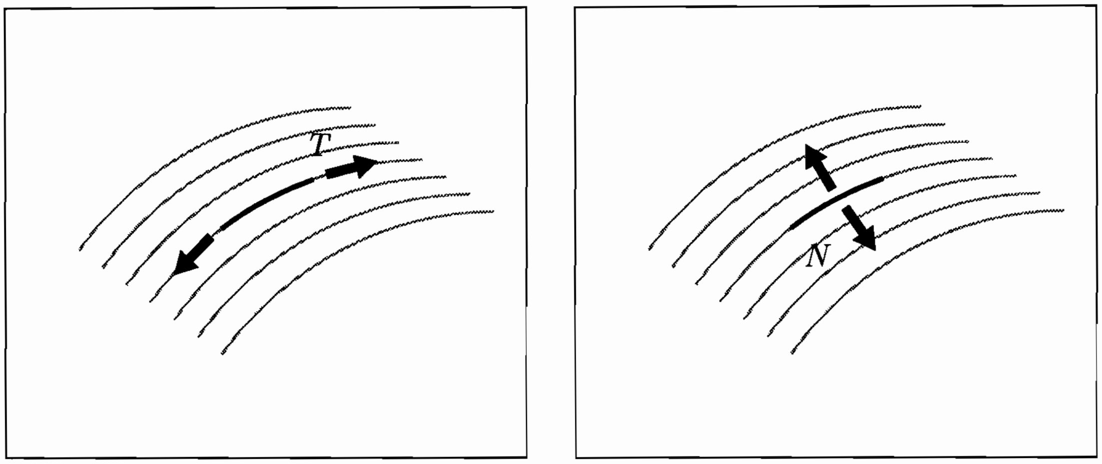
(i) keep the same scale but adopt another representation, or

(ii) change scale.

Keeping *status quo*, i.e. keeping the sole curve representation under the current scale, is bound to failure.

5. Mapping complexity and indexing representation

As curve detection is central to vision, what is required is a measure of the complexity of curves, and our specific goal in this thesis is to propose one. We will show how it successfully handles the Kanizsa and the Paolina examples, among others. It is based on an intermediate representation—the *discrete tangent map*, or a discretized tangent field—and a consequence of our analysis is that such intermediate representations are necessary for a proper segregation of curve-like patterns that fill areas, from curve-like patterns that extend mainly along their length and also from dust patterns (discontinuities, for instance). These representational differences capture the first stages of segmentation; but via complexity analysis not pixel grouping.



(a) *tangential* for *continuity*

(b) *normal* for *density*

FIGURE 1.9. The main idea: examine the rate of growth of oriented dilations, in the normal direction N to test density ("space-fillingness") and in the tangential direction T to test continuity.

The complexity measure we derive will be tailored to discrete "curve-like" sets such as those we seek in edge detection. The basic idea will be to look in two directions: in the tangential direction to assess continuity and in the normal direction to assess density of the object within a local extent (Fig. 1.9). This will lead to two complexity indexes, that we call the *normal* and *tangential complexity indexes*, and constitute the basis for our *complexity map.* Although the tangential complexity captures the same line of thought as previous researchers such as Ullman and Mumford, it is the normal complexity that provides some further insight into segregating textures from curve patterns. Both must be used together.

6. Outline of the thesis

This thesis will explore a broad range of concepts. Chapter 2 briefly reviews the detection of structure in early vision: i.e. edge detection, texture analysis, early perceptual grouping and scale-space analysis. The reader well aware of the material can therefore skip this chapter. Chapter 3 is the most "mathematically concentrated". Classical techniques of geometric measure theory are introduced and placed in perspective to articulate their relevance for computer vision. The important notion of *curve-like set* is brought to the forefront. These sets form a richer class of objects than

Jordan curves and are independent of any specific functional form or parametrization. Curve-like sets are appropriate for vision in that they are not unlike the set of pixels through which a curve might pass (at a finite scale). From this we introduce a parametrization-free intermediate representation, the *Besicovitch tangent set*, which will be the basis for the construction of our complexity map. Tangent sets are related to the discrete domain as discrete tangent maps and these are obtained from image operations. Finally, the rectifiability property of curve-like sets will allow us to study the properties of the local representation. This will lead to our main results in the continuous domain, namely what we call the *tangent separation theorems.*

Chapter 4 is the heart of the thesis. It will concentrate on performing a finer characterization of patterns. This will be done by dilating sets and looking at the Minkowski functional locally, both in space and in scale, using the fact that measures, dimensionality, and complexity are coupled. Measures for curves such as length, number of components, or area covered, are all captured by the Minkowski functional. The major novelty of our approach is to perform oriented dilations to define both the normal and tangential complexity indexes. These will be the building blocks of the *complexity map.* The complexity indexes are linked to an abstract notion of dimension, since it is dimension that separates (0-dimensional) dust patterns from (1-dimensional) contours and from (2-dimensional) texture flows. In the Kanizsa example (Fig. 1.7a), for instance, the occluding surface is 2-D, while the protruding top and bottom portions of the rectangle are 1-D and some of the discontinuities (the corners) are 0-D. The end of Chapter 4 shows how to build the complexity map on a discrete grid. It illustrates the different steps involved before computing the complexity indexes and highlights the parameters that need to be set.

Chapter 5, will be devoted to the choice of representation and parameter estimation. It starts by using the complexity map to perform a segmentation on the image. The idea will be to partition the complexity space and to lift the partition off the image. This will provide us with an early perceptual grouping that is based on the type of representation appropriate for integration. The segmentation scheme and a series of experiments on simple patterns will enable us to obtain a strategy to set the parameters involved in the analysis. Finally, we will show in Chapter 6 that, when applied to our test images, the complexity map agrees with our intuition and with the preset requirements.

While much of the thesis is mathematical, we stress that it is not a goal to only develop a mathematical structure. Rather, we believe that complexity analysis of the sort proposed is essential for vision systems. Our goal is therefore to lay out the structure of an appropriate complexity theory, and to illustrate its properties by computational experiments as well as mathematical analysis. Much is formally incomplete and we hope that our conjectures will be taken as incentives for future research both in mathematics and in computer vision.

7. A note on the experiments

A series of examples will be carried over throughout this thesis. These will serve to illustrate the points we are trying to make and clarify different concepts related to the algorithms involved. We will constantly refer to these as

(i) the Kanizsa pattern: Fig 1.3a;

(ii) the Perceptron spirals: one curve (Fig 1.8a) and two curves (Fig. 1.8b);

(iii) the Ullman patterns: pop-out (Fig. 1.5a) and hidden (Fig. 1.5b);

(iv) the Paolina image: Fig. 1.4e.

These examples are not all equivalent. Some are sets that can be described easily, others are images in which the underlying curve-like sets must be inferred. For the Kanizsa pattern and the Perceptron spirals, the structure was known *a priori*. The tangent map will be discretized from the continuous one only for the Kanizsa pattern however. The discrete tangent map of the Perceptron spirals and the Ullman patterns will be obtained by analyzing the corresponding images and provide structural properties. The Paolina image is our example of a full grey level image for which we will infer edges and interpret the inference process. Note the difference between both the Ullman patterns and Perceptron spirals, in which we will try to detect *negative contrast lines* (i.e. dark lines on a light background), and the Paolina image, for which we are seeking *edges* (i.e. boundaries between light/dark regions).

In all cases we will project the set or the image into the unit square, so our numbers (scale for grouping and resolution) will be relative to the unit square. Two different scales will be considered: (i) the *scale of the operator* σ, expressed in pixels and (ii) the *scale for complexity analysis* δ. δ is linked with the *spatial extent* Ω over which grouping should be considered. *Resolution* is the inverse of the size of smallest element of the digitized grid. Finally, whenever a variable is "hatted", $\hat{\omega}$, for instance, it means it is expressed in image coordinates (pixels).

CHAPTER 2

Edge detection and early vision

Vision is the process which interprets the patterns of light projected on a photosensitive device that allows an organism to interact with its environment. There are two key observations one can make about biological vision: (i) its immediacy, and (ii) its complexity. Although just the action of opening our eyes allows us to see, a complete understanding of vision still remains a challenging problem, and a general artificial vision system remains elusive. One question then with hard problems is "is this problem solvable"? Evolution has surely found an elegant solution, since we are equipped with a highly sophisticated visual system: we have two eyes, and use the information for stereopsis, we have different types of photoreceptors (the rods and the cones) to allow the perception of color and patterns with low contrast, we have a fovea to focus attention, but most importantly, almost half of our brain with its 10^{15} synapses is devoted to visual processing of one sort or another. The resulting system, which is highly adaptable and flexible, provides us with the necessary visual information processing to ensure our survival.

The sophistication of our visual system and its versatility has driven man to try to understand and even duplicate it. With the advent of the general computer, the advances in neurophysiology and psychology, the question that arose was: can we make a machine see[1]? A computer that could process visual inputs would allow the automation or control of tasks, which until now have required human supervision. This is especially important when tasks are dangerous, repetitive or time-consuming. The challenge has therefore triggered the interest of scientists from various disciplines. As

[1] Actually, the question was even more general: can we make a machine think? And this was the basis for *artificial intelligence*.

neurophysiologists and psychophysicists pursued their effort to understand biological vision and its intricate behavior, mathematicians, engineers, and computer scientists, among others, tried to develop the necessary paradigms that would provide vision to a machine.

Early and later vision. The human visual system being highly organized, it is natural to model vision as proceeding in stages, each stage producing increasingly more useful descriptions of the images, and increasingly abstracting toward scene properties. At the end of the last century von Helmholtz (1867) advocated a bimodal hierarchy of vision: early and later vision (sometimes referred to as low-level and high-level vision). Early vision is typically based on estimating scene properties, while later vision relates to the recognition and meaning of objects (an excellent presentation, well beyond the scope of this chapter, can be found in Zucker (1992)). This is why low-level vision is thought to use very general knowledge, such as the physics of imaging, while high-level vision is related to domain specific knowledge. Examples of low-level vision problems are edge detection, stereovision, texture analysis, shape from X (where X can be motion, shading, texture, etc). Shape analysis (such as the decomposition of an object into parts, for instance), and object recognition are examples of high-level vision processes.

Top-down and bottom-up approaches. There are at least two types of information involved in the vision process: *data* and *knowledge*. Visual information is usually represented by images or two-dimensional arrays of numbers. The image represents the activity of photosensitive cells in the retina, or the response of CCD-elements in a camera. The data is therefore the information provided by these sensors. Knowledge can be obtained from different sources: it can be knowledge about the visual system (sensors, optical system, body position), about the environment, past experience, etc. A spectrum of approaches in the development of theories for machine vision then emerges [Breton 1994]. At one end of this spectrum is what is called the "bottom-up" (or data-driven) approach, which builds on the data and infers more structured information. At the other end is what is called the "top-down" (or model-based) approach, which uses the data to answer specific questions. Hypotheses are verified from the knowledge base and conclusions are confirmed by the data. Knowledge-based approaches, where the knowledge is rich and specific (such as in OCR, for example), make the vision problem (e.g., recognizing text) easier, but the solution may have only limited interest (for the autonomous land vehicle, for instance). On the other hand, when the knowledge is more general, the solution might be more widely applicable, but may also be more difficult to reach.

The approach advocated in this thesis is to enrich the data with more structured representations, carefully organizing the information to break complexity (recalling our opening quote). We will consider a vision system with only one single intensity image $\mathcal{I}$. We assume the surface of the photosensitive array to be a square lattice composed of square cells which all have the same response, and the value of the intensity at a given point will be written $\mathcal{I}(\hat{x})$, where $\hat{x} = (x_i, y_i)$ is one of the pixels of the square lattice. The theories that will be developed will be mostly bottom-up, following Marr's *principle of least commitment* (i.e., postpone decisions as far as possible) [Marr 1982], and will sit at the upper limit of what we described to be low-level vision. In this chapter, the first topic will be edge detection, one of the most primitive perceptual tasks. Then, issues of grouping will be addressed, and we will close the chapter with a quick survey of the basic principles underlying texture analysis and scale-space theory.

1. Edge detection...

Given an intensity image $\mathcal{I}$, we seek to detect edges (light/dark boundaries), or lines (positive contrast lines, i.e., light lines on a dark background, or negative contrast lines, i.e., dark lines on a light background). How can this be done? We now briefly sketch two lines of investigation that have coupled to define current approaches to solving this problem.

1.1. ...in neurobiology. Viewed in the large, the architecture of the visual cortex appears ideally suited as an edge/curve detection machine [Zucker 1993]. In particular, patterns of light projected onto the retina influence the activity of cells in the visual system, either in an excitatory manner, leading to an increase in that cell's firing rate, or an inhibitory manner, leading to a decrease. The resultant map of the activity of an individual cell as a function of light distribution is called a *receptive field.* Cells participating in the processing of visual information have been classified into different functional categories. *Simple cells* are those cells whose receptive fields are separated into distinct subzones. They most resemble line and edge detectors in computer vision (see Fig 2.1a); they are orientationally selective, as well as selective for a number of other properties including stimulus contrast, direction of motion and stereo disparity. They appear at a range of sizes, and are optimally selective to different spatial frequencies, or to bars of different widths. The pioneering work of Hubel & Wiesel on cats [1962*b*] and monkeys [1962*a*] was seminal and led to the development of different mathematical models for simple cells. Some models are based on Gabor functions [Jones & Palmer 1987], others on difference of Gaussians [Hochstein & Spitzer 1984] or derivatives of Gaussians [Young 1985]. In most cases, the data fit the various models. The derivative of Gaussian model is often the one adopted for computational vision because of its theoretical implications [Koenderink 1990]. The graph of such an operator is represented as grey values on Fig. 2.1b.

We will not go into the details of the models at this moment. The important point is that the visual cortex has an apparatus sensitive to local orientation; i.e., there are cells that respond to lines of a given length, width, orientation and contrast. One other point about these cells as a group however, is that, at any given time, only a small fraction of them are active. Over orientation hypercolumns, the pattern of activity is very sparse as opposed to other regions of the striate cortex where the activity is much more sustained. This observation has led to a conjecture that there exists a functional classification for regions of the visual cortex which segregates cells into those representing scalar variables from those representing geometric variables [Allman & Zucker 1990]. For our investigation, this prefigures a point developed in the next chapter: curves in an image should be of "measure zero", i.e., sparse with respect to other types of structure. This observation raises an important issue that is rarely discussed in the physiological literature. Since orientation selective cells will respond to oriented stimuli that are part of edges and part of textures (such as "hair patterns"), how can it be determined, from the responses of such cells, whether

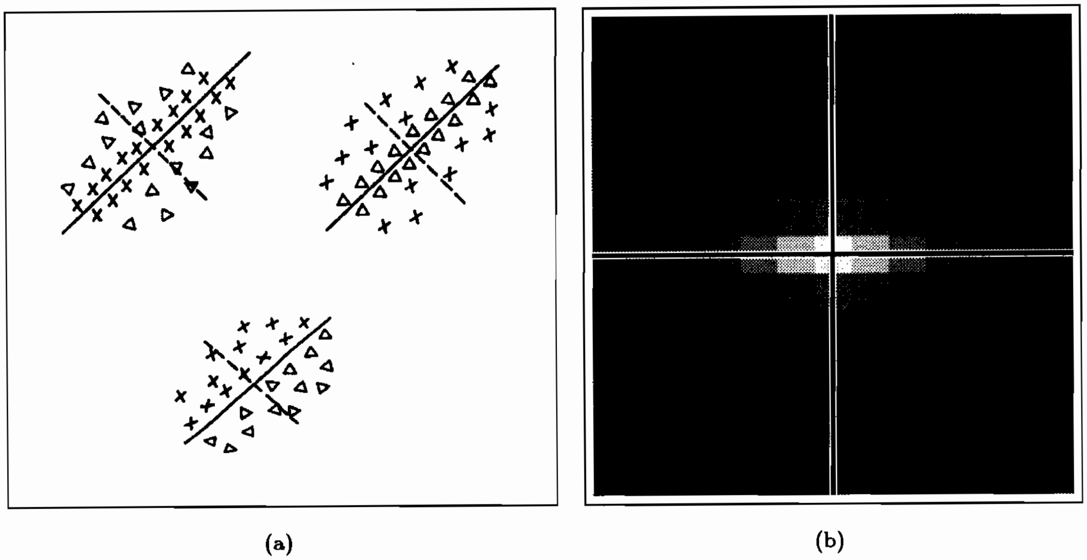


FIGURE 2.1. Edge detection: from neurophysiology to models: (a) arrangement of simple cortical receptive fields sketched by Hubel & Wiesel (1962*b*) (b) a biological model of edge detection: the second derivative of Gaussian operator. The patterns of activity for the receptive fields in (a) are noted by the crosses, areas giving excitatory responses, and by the triangles, areas giving inhibitory responses. In (b) the intensity correlates with the desired activity: the light central region is excitatory, while the dark side bands are inhibitory.

the stimulus was an edge or a texture. The model developed in this thesis will provide an answer to this question.

1.2. ...in computer vision. An edge in a picture can roughly be defined as a discontinuity or abrupt change in the grey level or color (see Fig. 2.2). That is why idealized edges have been represented at first as a singularity in the graph of a function, the intensity map, and have been conceptually linked to the process of digital differentiation. In general, we can say that local edge detection is typically accomplished in two steps [Zucker 1993]: *(i)* the convolution of an operator against the image, and *(ii)* some process of interpretation of the operator's responses.

An early significant paper on this topic was due to Roberts (1965), who employed a simple 2x2 operator, the so-called Roberts cross operator, to enhance edges in digital images of polyhedra. Because of its high sensitivity to noise, more sophisticated masks were developed such as the Sobel and Prewitt operators [Levine 1985].

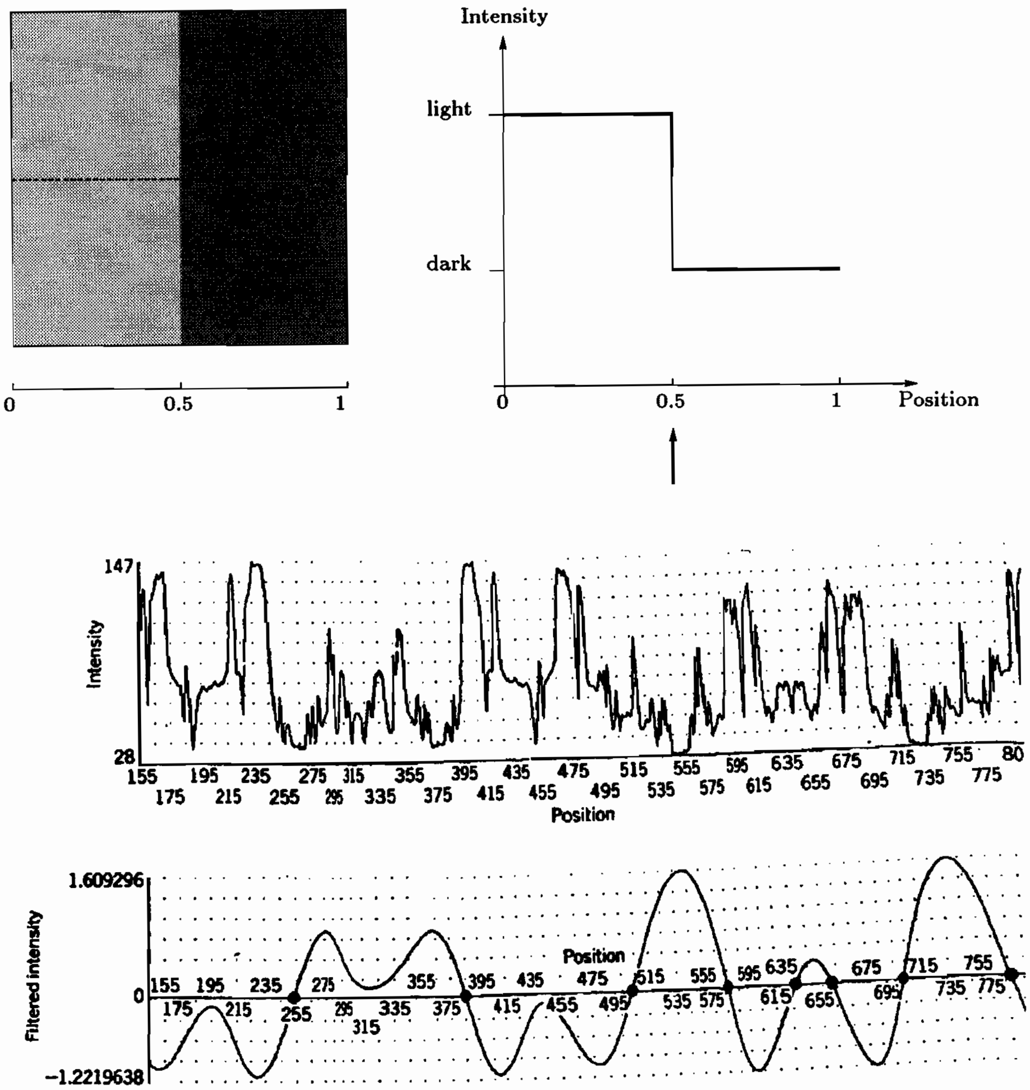


FIGURE 2.2. Edge detection in computer vision. The top row shows on the left an image of an ideal edge. On the right, we displayed the grey value along one scan line (corresponding to the dotted line on the left) of the original pattern as a function of position. The position of the edge corresponds with the discontinuity in the graph and this is what needs to be detected. (b) A more realistic 1D signal, and (c) the output obtained after applying the Gaussian second derivative operator. The large dots on the x-axis give the position of the so-called *zero-crossings* (figures (b) and (c) are reproduced from [Hildreth 1992]).

The new family of operators were still not satisfying in a general case. The idealized edge model was not very realistic (see Fig.2.2b for a more realistic 1D signal). Moreover, edges as they occur in images, have not only a position, but an orientation.

The need for multiple orientations at a position, the fact that edges could occur at different scales, and the search for an optimal operator, were all factors that drove the development of more sophisticated algorithms. A new trend of approaches arrived that tried to mimic the physiology of the visual system or to optimize the operator's response. Marr & Hildreth (1980) based their theory of edge detection on physiological findings. They used unoriented Laplacian operators and based the decision stage on the coincidence of zero-crossings across scale. Canny (1986) developed an edge operator based on optimality principles: the optimal edge operator tuning having to be a trade-off between detection and localization. He arrived at a line operator whose cross-section is similar to a Gaussian second derivative and an edge operator similar to a Gaussian derivative. His operators had the difference of being oriented but only allowed a single edge element at each position.

Although these (together with the more modern approaches of Deriche (1987), Perona & Malik (1990), Freeman & Adelson (1991)) are probably among the most popular edge detection techniques, they have two major shortcomings [Zucker 1993]. First, the assumption of linearity, which inevitably blurs nearby structures together and smoothes around corners and discontinuities. Secondly, the assumption of a single value at each position, failing to represent the places where there are orientation discontinuities such as corners and T-junctions [Guzman 1968, Waltz 1975].

A different approach to local line/edge detection has been presented by Iverson & Zucker (1995). It is based on what they called *logical/linear operators*, in which a set of non-linearities were developed to significantly improve the sensitivity of the initial operators over the optimal linear ones. These non-linearities implement a test on continuity of support along the preferred direction of the operator, and a test on variation across it. The non-linearities are formulated within a logic that accumulates consistent evidence linearly, but in which incompatible evidence provokes a nonlinear suppression. This approach is particularly well suited to detect the ends of lines and to report places where there are multiple orientations. It also makes a clear distinction between the detection of positive and negative contrast lines versus the detection of edges. It is the approach we will adopt in this thesis, since it better represents the local structure of an edge/line scene.

1.3. ...but then what? The work of Hubel and Wiesel demonstrated that there exist some cells in the brain that respond to elementary geometrical patterns: namely lines and bars of various lengths and orientations. If these constitute the building blocks for a theory of computational vision, the question to answer now is how can the local information be combined to eventually be able to interpret the scene? How can local edge detector responses be glued together into global entities such as 'contours' or 'textures'? This, in the context of edge detection, constitutes the local to global transition.

2. The local to global transition

The perceptual grouping problem is a key issue to fill the gap between low-level and high-level processes and has been addressed by various groups of researchers in computational vision. Many directions have been explored, but none seems to predominate. One of the biggest problems is that it is not clear what should be the output of the process. The following is an attempt to sketch some of the different solutions proposed to group edge information into more global representations, to illustrate the diversity of approaches, assumptions, and conceptualizations.

Fitting polynomials to edge points. For a long time, research related to curve detection has had two main thrusts: *(i)* designing operators and *(ii)* fitting global functions through their responses. In his early system, Roberts (1965) fitted long straight lines to edge detector outputs, because he was working in a world only composed of a relatively small number of blocks such as cubes. Even now, modern investigators are still fitting lines [Viéville & Faugeras 1990]. A world of lines being rather restrictive (although sufficiently rich to be intractable), approximations with higher order polynomials (spline fitting, for instance) were attempted. As a compromise, Saund (1992) suggested token grouping with arcs of circles over multiple scales to link edge/line elements to reduce the complexity of the search space. Any fitting procedure however is confronted with the problem that errors are large in both the dependent and independent variable due to the digitization process and to the inference process itself. This makes the curve fitting very unreliable. Moreover, it is not clear what the final representation should be.

Edge following and graph search. Following in the steps of Montanari (1972), much research have been made around edge following. As its name indicates, it is mainly concerned with following the paths in the edge map based on the local tangent and (sometimes) curvature information available. This originally seemed to be the right thing to do since, after all, it is the basis the *Fundamental Theorem of the Local Theory of Curves* [do Carmo 1976], which asserts that the combination of local orientation and curvature defines a plane curve uniquely modulo rigid body motion. The problem was therefore expressed as a graph search trying to optimize some cost function associated with the grouping process. Ramer's [1975] early *strokes* and *streaks* suggested an algorithm to merge oriented edge elements using a minimum cost state space. The cost function was based on the connectedness of the elements, their orientation compatibility (leading to a low overall curvature), and a unique assignment rule. Even today approaches dealing with building up decision trees to provide the grouping are advocated. Cox et al. (1993) recently enhanced the tree with multiple hypotheses allowing decisions to be taken at a later stage. They stress the fact that earlier approaches lacked the ability to handle intersections and to extrapolate over significant gaps, and they suggest that any edge grouping scheme should

(i) provide a mechanism to integrate information in the neighborhood of an edge and to avoid making irrevocable grouping decisions based on insufficient data;

(ii) have a prior model for the smoothness of curves on which to base grouping decision;

(iii) incorporate noise models for the edge detector;

(iv) be able to handle intersecting edges.

But even approaches meeting all these criteria would get lost on images like Fig. 1.4d. In fact, it is not clear that the dimensionality of the support is known in advance. It is therefore not surprising that techniques trying to blindly follow the paths lead to very deceiving results in general cases. Because they assume a curve-based representation, local-to-global grouping schemes modeled on edge following fail on highly textured images.

Lowe: organization and grouping. Lowe [Lowe 1989, Lowe & Binford 1987, Lowe & Binford 1979] stressed that, by grouping together features that are likely to have been produced by a single object, an object search can be intelligently ordered. He suggests [Lowe & Binford 1987] computing a measure of *meaningfulness*: how likely a grouping could have arisen from an underlying physical relationship rather than by accident. His solution forms all potential groupings and then tests for meaningfulness of the combination based on parallelism, collinearity and proximity. These groupings can then be used as an index for the next step in the search. The advantage of the proposed approach is the ability to detect globally significant features from locally weak information, since it groups over a wide range of sizes. The other aspect to be considered is the fact that he does not start from an idealized model, therefore there is no need to make assumptions about the amount of noise, and other domain specific assumptions. The problem, once more, is that testing linearity, parallelism and proximity is intractable when the scene is highly complex.

Mumford et al.: gaps, T-junctions, 2.1D sketch. Gaps arise as image features, but also because edge detectors are rarely perfect. In a series of papers, Mumford and his team [Nitzberg & Mumford 1990, Mumford 1992] complements the integration of contour information by a technique for filling out gaps as they occur in edge detection. The gap filling problem is modelled as a stochastic process that minimizes an energy functional. Their representation is called the 2.1D sketch [Nitzberg & Mumford 1990], where they advocate a partition of the intensity image into regions. They stress the use and detection of T-junctions (see Fig. 3.12), since they constitute such important perceptual clues but they avoid the complexity issue by assuming there is no texture. Edges must group into curves by definition. The edge recovery algorithm presented by this team [Nitzberg et al. 1993] would be complementary to the steps following our complexity analysis, as we shall present a method for validating the "curve assumption" (introduced in the next chapter).

2.1. Heitger & von der Heydt: occluding contours. A dreadful challenge to any theory of edge detection and edge grouping is the detection of subjective contours[2] [Kanizsa 1979], vivid percepts that arise at locations where discontinuities in

[2]Also called illusory contours, anomalous contours or contours without gradient.

the intensity image are totally nonexistent. Heitger & von der Heydt (1993) presented recently a technique to group *key points* responsible for occluding contours including subjective contours. These key points include T-junctions, corners and line endings. The grouping, which provides improved definition of occluding contours where contrast is weak, proceeds into four stages: (i) convolution with a set of orientation selective kernels; (ii) nonlinear pairing; (iii) enriched curve representation including the information about the potential occluding contours; (iv) final contours from local maxima of combined representations. The subjective contours could be classified into two extreme categories: *ortho*, when the grouping goes along the perpendicular direction to the key points, and *para* when it goes in the same direction (extrapolation). The pairing itself was made possible by setting rules for valid configurations. Their results successfully detected the completion not only for some of the classical patterns in the perception literature, but also on natural images. As opposed to Mumford (1992), who presented a mathematically-based model for the completion of gaps, this one is based on neural mechanisms suggested by physiological experiments. Once more however, it is not obvious how this would behave on complex scenes.

Ullman et al.: saliency maps. In Chapter 1, we outlined Ullman's contributions to the grouping of local edge elements. More generally, he observed that "segmentation should be conducted on an area of interest rather than applied to the entire image, implying that some preattentive process is required to detect prominent locations from which an area of interest is defined prior to the act of segmentation" [Ullman 1990]. Our approach is consistent with the observation. However Ullman's solution, the saliency map, is incomplete in the sense that we showed earlier. Our solution will not only overcome the problems pointed out in Chapter 1, but will also provide a unified scheme to work with curves, textures and sparse patterns.

David and Zucker: dynamic coverings. David & Zucker (1990) presented an alternate view of grouping of edge elements. Their method was based on the minimization of an energy functional which would provide a covering of the curves in the image. This process was done in parallel over the image domain by setting up small dynamic elements and a potential field built from the information provided by the tangent (edge) map. Problems with curve crossings (multiple orientations) occured however, and the representation for the global recovery of objects was not explicitly described. These questions in part motivated this thesis. In the end, we shall provide a technique for properly segmenting the tangent (edge) map so that corners and crossings can be properly handled.

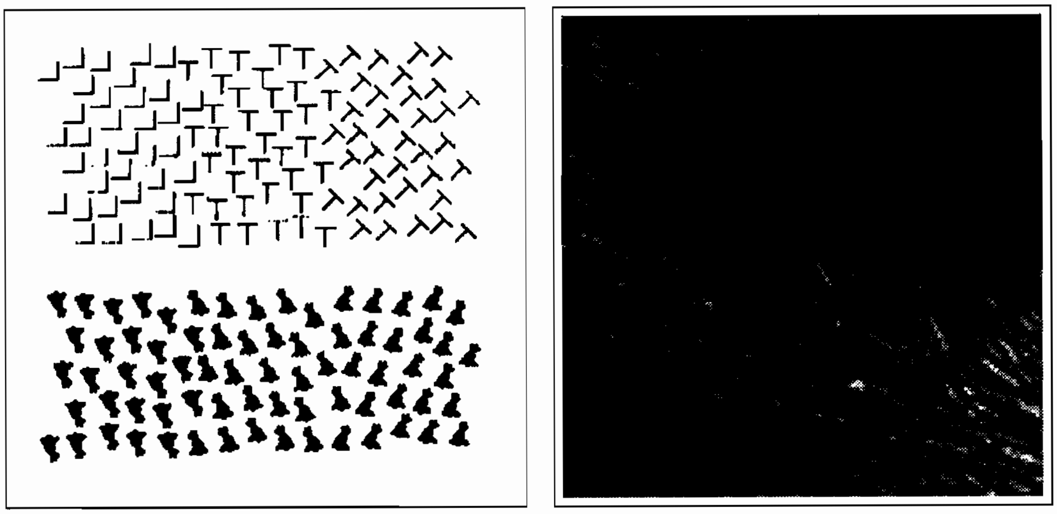
3. Textures in intensity images

Natural scenes are rich in image textures: just think of images of a brick wall, of a field of grass, of a pebble beach, or of a plaid jacket[3]. Although textures are ubiquitous in our visual environment, they admit no rigorous definition. A textbook definition states that a texture is "something composed of closely interwoven elements" [Ballard & Brown 1982]. The intuitive characteristic that applies to most cases is that image textures are underlaid by a two-dimensional support: textures, as we see them, cover an area.

The analysis and description of visual textures has quite a long history. One of the first pioneers was Gibson (1950), trying to infer shape from texture gradients. The history of texture analysis is also linked to the study of aerial photographs, trying to segment different regions in images. This was then considered as a pattern recognition problem. Julesz (1981) and Beck (1982), in psychology, studied the intricate nature of texture perception. This led to the *texton theory* [Julesz 1981] and triggered a whole line of research in the organization of texture patterns.

One view of texture is that it is based on the repetition of a pattern, called a *texture primitive* or *texel*. Take the example of a checkerboard. The basic elements would be the black and white tiles. Their careful arrangement of interwoven black and white squares in two orthogonal directions, provides the final pattern. The checkerboard texture is deterministic in nature since there is a clear (and short) set of rules to describe the placement of the tiles. These types of textures, which fall under the class of *structural models* [Ballard & Brown 1982], are very common in man-made objects.

[3] A classic collection of such image textures is presented in Brodatz' [1966] well-known book.



(a) Textons (b) TYPE II pattern

FIGURE 2.3. Textures. (a) example of a typical pattern from the texton theory (reproduced from [Levine 1985]) (b) example of a typical TYPE II pattern: fur (reproduced from [Zucker 1985]).

In nature, the organization of the basic patterns tends to be more probabilistic, leading to *statistical* texture models[4]. We therefore need to consider the *statistical distribution* of the basic patterns (such as the first moments, autocorrelation, cooccurence). Classical theories of texture [Julesz 1981, Beck 1982] attribute preattentive texture discrimination to difference in first order statistics of stimulus features (textons) such as orientation, size, color, and brightness of their constituents. These ideas were developed for black and white patterns such as those in Fig. 2.3. More recently, Malik & Perona (1990) applied it to grey-level images. The image, $\mathcal{I}$, needed first to be filtered by a bank of linear filters (spanning a range of orientations), followed by half-wave rectification and then a model of intracortical inhibition. The result, once averaged out by a second set of oriented filters with large receptive fields, provided the texture gradient, and showed an excellent degree of texture discrimination.

A complete survey of texture analysis is beyond the scope of this introduction. The important point that will be used for image textures is their 2-dimensional support. The type of textures we will be interested in are those induced by curves (hair

[4] It is interesting to see that the study of texture analysis is closely related to the one of fractals [Mandelbrot 1982], in which they talk about deterministic and statistical self-similarity. The rules for the organization of the basic patterns tend to be more complicated but the same distinction applies with respect to the type of arrangement of the primitives.

patterns, fur, fields of grass, arborizations, etc). Their underlying texture elements are curves or pieces or curves, and the grouping of these basic elements clearly has a 2-dimensional support. The study of TYPE I and TYPE II patterns [Zucker 1985] has a direct relationship to some of the ideas we will present in this thesis. TYPE I patterns were originally defined as those having 1-dimensional support (occluding contours, for instance), while TYPE II patterns were those with 2-dimensional support such as flows (fur, for instance). One of our contributions will be to extend (and alter a little) the classification, and to provide a scheme to decide which kind of pattern one is confronted with.

4. Scale-space theory and scale selection

One message that stands out from the foundations of edge detection is that raw numerical signals are insufficient for tasks requiring any sophistication (a justification for this can be found for instance in Ullman (1990)). Since one wants to have a representation as compact as possible that will correlate with meaningful events, the observation that significant changes in the image occur at multiple resolutions is of paramount importance. Consider for instance a leopard's coat. At a fine scale, one's attention will be driven by the individual hairs making the fur, while at a coarser scale, the spot patterns on the coat will emerge. Two scales, two drastically different texture patterns. Which scale should one choose?

The study of scale-space was pioneered by Witkin (1983) and Koenderink (1984). If we consider the study of 1D-signals, then meaningful events seem to correlate with extrema of the signal and of the derivatives. The derivative operator and the local extrema operator depend on the signal but also on the spatial support of the operator. In edge detection, for instance, the events sought were the zero-crossings of the second derivative of Gaussian (a technique for detecting inflection points). The multiscale representation is then obtained by embedding the signal in a one-parameter family of derived signals, the *scale-space*, where the *scale* parameter σ is the standard deviation of the Gaussian kernel

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

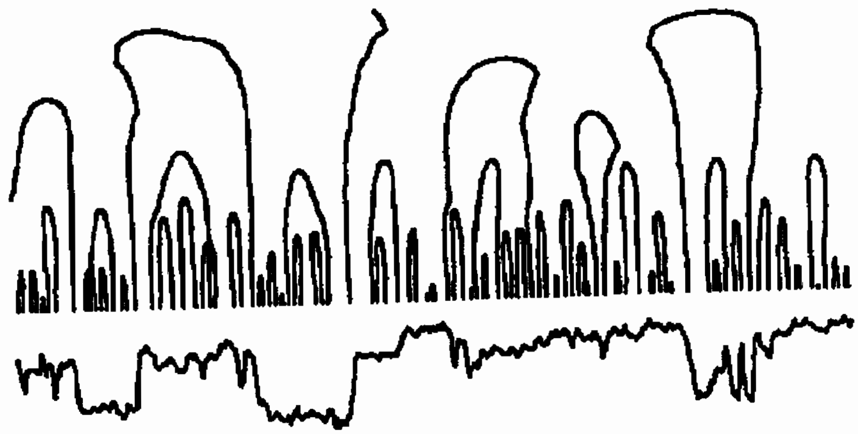


FIGURE 2.4. The scale-space theory. First, the signal on the bottom. The x-axis shows position and the y axis, signal amplitude. On the top are shown the trace of the zero-crossings of a second derivative of Gaussian operator through scale-space. The x-axis corresponds again to space, but this time the y-axis represents scale. The bottom part of the scale-space loops corresponds to fine scales while the upper part is for coarser scales. Figure reproduced from Witkin (1992).

and then tracking the zero-crossings across scales (see Fig. 2.2c for the detection and Fig. 2.4 for the tracking). One main result to justify the technique is the *causality principle* which states that events can appear at a finer scale but existing ones (the ones from coarser scales) cannot (in general) disappear, cross each other, or change direction. The resulting patterns in scale-space have the characteristic shape of embedded, non-intersecting open loops. The main advantage of such patterns was the ability to continuously track events across scale. The *scale of the event* was then defined to be the scale at which it vanishes, while the *location of the event* was the one at the finest scale. This is also what Witkin (1983) calls the *identity* and the *localization assumptions*, respectively.

Having set the foundations for a continuous theory of scale-space, efforts have been made to apply it to images (not only continuous 1D signals). A good survey of the work done in scale-space since the early 90s can be found in Lindeberg (1991).

One of the most important issues, however, is the one of scale selection (remember the leopard's coat example). Originally, the existence of a single fixed scale for analysis was advocated [Witkin 1992]. This became a very controversial subject, and continues to be even now. After presenting a technique to select interesting scales from local extrema of normalized scale invariant derivatives, Lindeberg (1993) concludes that the selection of "the best scale" without *a priori* knowledge is an impossible mathematical problem. His method instead proposes reasonable "initial hypotheses". In another line of thought, Elder & Zucker (1995) argue that there should not be one single scale for the entire image, but rather multiple local (in space) scales. Their claim was that these could be uniquely determined when using *a priori* information about the sensor noise, with the end result called the *minimum reliable scale.*

In this thesis, issues of scale will also arise. It will not be a scale for the operator, since we assume this to be fixed and given, but rather a scale for complexity analysis that we will seek. Our definition of scale will emerge in Chapter 4, and our approach to select it will be presented later, in Chapter 5.

5. Summary

Early vision encompasses a wide range of processes intending to obtain from raw numerical signals a richer representation, yet more compact description of the image structure. Different modules for early vision have been presented and edge detection is one of them. In this thesis we will not consider issues of finding "optimal" edge operators in terms of sensitivity to noise and localization, but rather focus on the choice of intermediate representation, setting desired requirements.

Texture analysis also is thought of as being a low-level vision problem. Although the definition of texture still stays ambiguous, it is confronted with this same problematic issue of scale selection, or representation of the signal at multiple scales. Texture analysis, edge detection, edge grouping, scale selection, are all intrically related concepts, as will be shown in this thesis. The coupling of these different issues will be the solution we will provide to bridge the gap from retinotopic maps to object-centered ones.

CHAPTER 3

Curve-like sets and curve detection

In the case of the world on a human scale you don't care much about problems involving infinities or infinitesimals, whereas you certainly care whether something is line-like or point-like.

JAN. J. KOENDERINK [1990]

This chapter is an attempt to answer the question "what is a curve?" in the context of curve detection. From an elementary analysis textbook [Simmons 1963], we get that *"a continuous curve is usually thought as the path of a continuous moving point and this rather vague notion is often felt to carry with it the even vaguer attribute of 'thinness' or 'one-dimensionality' "*. This definition of a curve is bound to the one of dimension. The first section of this chapter will informally address the issue of dimension to give a feel for what will follow. Then the notion of "curve" as usually presented in elementary differential geometry, namely the one of Jordan, will be presented. It will remind us of concepts such as local linear representation (the tangent) and global measure (the length). But curves, such as those encountered in early vision, are not Jordan curves and the required parametrization, among other things, is exactly what is sought after and not what is given. A generalization of the previous ideas through measure theory to maps that are not necessarily smooth needs therefore to be introduced. The resulting *curve-like sets* and their associated parametrization-free tangents will constitute a much better basis for our needs and will represent the core apparatus for our work.

When inferring curve-like objects from images, one is confronted with issues such as discretization, quantization and choice of scale. One of the main characteristics of curve-like sets is that they lie in the continuous domain. How can one make a parallel between these sets and the output of a finite set of "edge detectors"? An answer to this will lead to a definition of what we will call *discrete curve-like sets.* The underlying philosophy will be somewhat different since we are trying to infer the set from images. We start by getting the discrete tangent map and then infer the curve-like set from its local properties. The local structure of the discrete curve-like sets will be obtained through edge detection. From our previous description of the edge we will be able to present the implications of our choice of operators and decision method.

1. The A-B-C of dimension

The history of the various notions of dimension involves the greatest mathematicians of the turn of the century: Poincaré, Lebesgue, Brouwer, Cantor, Peano, Hilbert, just to name a few. That history is very closely related to the creation of space-filling curves and early fractals [Peitgen et al. 1992]. Hausdorff remarked that the problem of creating the right notion of dimension is very complicated. People had an intuitive idea about dimension: the dimension of a set, say E, is the number of independent parameters (coordinates), which are required for the unique description of its elements. This turned out to be incorrect, as the counterexamples of Cantor and Peano showed. In this section, we will review three different approaches to defining the concept of dimension. These will help in setting up the framework for our own intuitive requirements (presented last) which we call the *curve assumption.*

1.1. Poincaré's cut dimension. Poincaré's cut dimension is inductive by nature and starts with a point. A point has dimension 0. Then he observes that "...if to *divide* a continuum it suffices to consider as cuts a certain number of elements all distinguishable from one another, we say that this continuum is of *one dimension*; if on the contrary, to divide a continuum it is necessary to consider as cuts a system of elements themselves forming one or several continua, we shall say that this continuum is of *several dimensions*" [Poincaré 1926].

From this definition we get that a segment has dimension 1 since it can be split by a point (dimension 0). The same happens for a circle, since it can be disconnected using a pair of points (dimension 0). The unit square has dimension 2, since it needs a line (dimension 1) to get disconnected. Finally, the cube has dimension 3, since can be disconnected using a plane (dimension 2).

Poincaré's definition has the advantage of being very intuitive and easy to grasp. It will be used later to motivate the proofs of the *tangent separation theorems* (Section 4). This idea of dimension also formed the basis of the now accepted one developed by Menger and Urysohn [Hurewicz & Wallman 1948]; namely that

(i) the empty set has dimension -1,

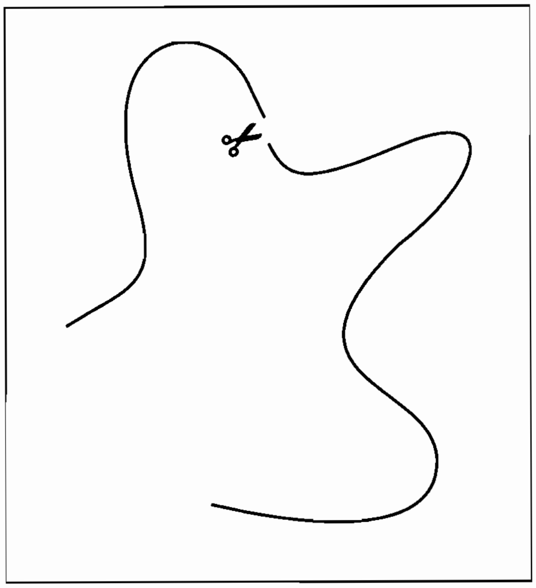
(ii) the dimension of a space is the least integer n for which every point has arbitrarily small neighborhoods whose boundaries have dimension smaller than n.

1.2. Lebesgue covering dimension. The Lebesgue covering dimension is the most frequently used in point set topology to define the notion of dimension for a topological space [Munkres 1975, Edgar 1990]. It consists in covering the set E with little disks (such as those used in point set topology) and then focusing on the maximal number of disks in the cover which have non-empty intersection. This is called the *order of the cover.*

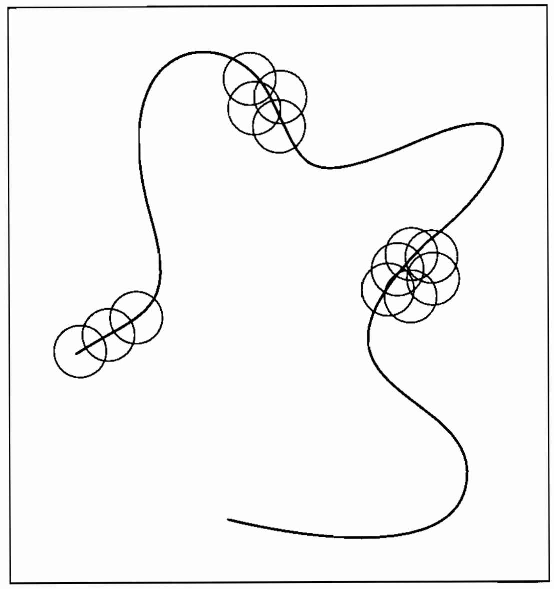
An object E has covering dimension n provided any cover admits an open refinement of order $n+1$, but not of order n. Taking the line segment as an example, it is easy to see that the order of the cover cannot exceed 2, leading to a topological dimension smaller than 1 (see Fig. 3.1).

Another equivalent definition would be that the topological dimension of a set E is the smallest integer k such that, for all $\epsilon > 0$, there exists a covering A_i of E by closed sets of diameter $\leq \epsilon$, with the following property: the intersection of any $k+2$ sets A_i is empty.

1.3. Measure and dimension. In their classical monograph, Hurewicz & Wallman (1948) presented different approaches to defining dimension. The one we will adopt here will associate the concepts of measure and dimension. An object will be called one-dimensional if it has length (one-dimensional measure), 2-dimensional if it has an area (2-dimensional measure), 3-dimensional if it has a volume (3-dimensional



(a) toward cut dimension



(b) toward covering dimension

FIGURE 3.1. Illustration of Poincaré's cut and Lebesgue's covering dimension of a line. (a) a simple line disconnected by a single point; (b) covers of orders 2, 3, and more. Since it is possible to cover the curve with a cover of order 2, its dimension is not bigger than 1 from the definition.

measure), and so on. The measure that will be used in this case is the one developed by Hausdorff. It will also allow us to study in more details the local structure of the set.

1.4. The curve assumption. We are interested in developing a definition applicable to the kinds of objects encountered in curve detection within computer vision. Three intuitive guidelines will drive the development of our theory:

(i) a notion of "thinness": curves extend along their length;

(ii) curves are "of measure zero" with respect to surfaces;

(iii) discontinuities are of "measure zero" with respect to curves.

These further relate to the fact that, typically, there is more than one curve in an image and these intersect causing discontinuities in orientation. The properties listed are necessary for a suitable definition of the type of objects perceived when doing curve detection, and will imply constraints on the representation of visual information. We will say that an object satisfying these rules verifies the *curve assumption.* Thus far, the *curve assumption* is very informal and intuitive; in Chapter 5, we will make the definition more precise.

Before introducing geometric measure theory, we will review the concepts of local linear approximation and rectifiability for Jordan simple curves in $\mathbf{R}^2$. This is most easily approached through differential geometry. We will then try to extend it to a more general class of objects that we will call *curve-like sets.* It will be shown that the concept of rectifiability allows us to derive important properties and constraints on the local structure of the sets to be studied.

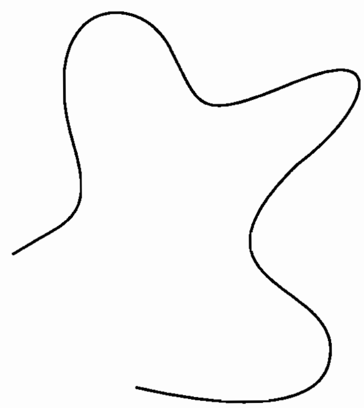
2. Elementary differential geometry: Jordan curves

The most common definition of a curve is the one of Jordan, namely that a curve Γ is the range of a continuous map γ from an interval I to Euclidean space (typically $\mathbf{R}^2$ or $\mathbf{R}^3$). In elementary differential geometry, this definition precedes two other basic notions, namely the *length* and the best local linear approximation or the *tangent* to Γ at $\gamma(t)$.

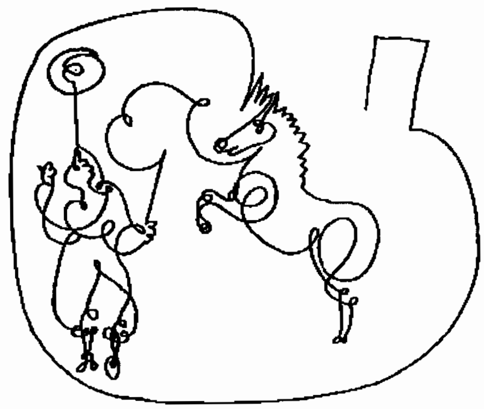
Definition 3.1 (Jordan curve [Tricot 1995]). A *curve* Γ in $\mathbf{R}^2$ is the range of a continuous function $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ defined on an interval $[a, b]$. If γ is an injection, the curve Γ is called *simple.* Its endpoints are $\gamma(a) = A$ and $\gamma(b) = B$. The mapping γ is called a *parametrization* for the curve Γ.

Remark 3.1. In the rest of this section, when using the word "curve", we mean a simple Jordan curve. Although too restrictive a definition for the type of patterns detected in edge detection, it will be useful to study the basic concepts. We will widen the definition in the next section to include the types of sets sought for in computer vision.

Two examples of Jordan curves are shown on Fig. 3.2. In (a), a simple Jordan curve, i.e. a Bézier curve with 7 control points. In (b), a Jordan curve that is not simple (since the curve cuts itself). This is a drawing due to Pablo Picasso that we adapted from [Mendès-France 1991*b*], and it illustrates the fact that the definition of "curve" is indeed very large. The following definition builds an equivalence relation between different parametrizations of a curve.



(a) a very simple curve



(b) a Jordan curve that is not simple!

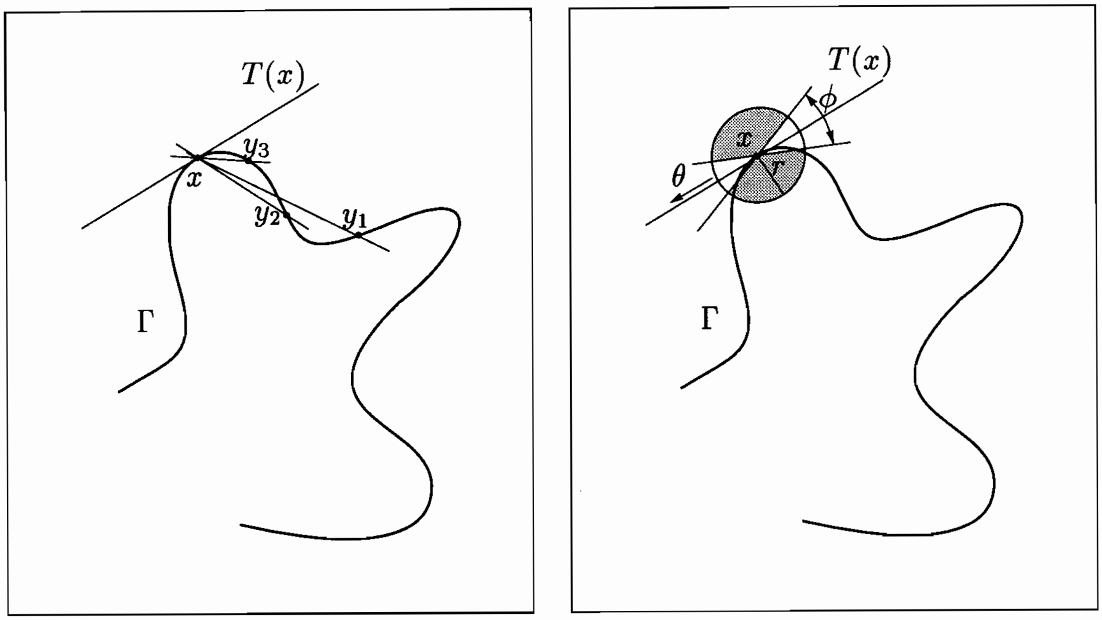
FIGURE 3.2. Examples of Jordan curves. In (b) we reproduced a drawing due to Picasso entitled "Le Jongleur" (adapted from [Mendès-France 1991*b*]).

Definition 3.2 (Fréchet equivalence [Cesari 1965]). A mapping $\Gamma : \gamma = \gamma(t), t \in I$ is said to be *Fréchet equivalent* to another mapping $\Gamma_1 : \gamma = \gamma_1(s), s \in I_1$ if for every $\epsilon > 0$ there exists a homeomorphism h_ϵ from I_1 to I such that $|\gamma(h_\epsilon(s)) - \gamma_1(s)| < \epsilon$ for all $s \in I_1$. This defines an equivalence relation between Γ and Γ_1, and then we write $\Gamma \sim \Gamma_1$.

2.1. A curve in the small: tangent. The notion of a tangent, the best linear approximation to a curve at a point, is key in the study of curves in the small. Intuitively, it is defined as follows:

Definition 3.3 (Tangent to a curve [Hilbert & Cohn-Vossen 1990]). If Γ is a simple (parametrized) curve and x is a point on Γ, the *tangent* $T(x)$ to the curve at x is the limit (if it exists) of the straight line passing through x and y when $y \in \Gamma$ and $y \to x$.

The reason why we called this an intuitive definition is the fact that the limit $x \to y$ is not always defined. Furthermore, in applications such as computer vision, for one, the parametrization is exactly what one is trying to infer. Thus a more general model is required. However, when the parametrization is given, the tangent



(a) the tangent $T(x)$ to Γ at x (b) Besicovitch tangent

FIGURE 3.3. This figure illustrates in (a) the intuitive definition of a tangent T to a curve Γ at a point $x = \gamma(t)$. Take a sequence $\{y_1, y_2, \cdots\}$ of points on the curve converging to x. Draw the lines passing through y_i and x. The "limit line" gives us the tangent $T(x)$ to the curve at x. In (b) we illustrate the parametrization-free definition of the tangent, looking at a cone that shrinks around the point x (presented later in the text).

is the first derivative of the map γ. We will see in the next section that this local notion is tightly linked to the global one of *length*. In this highly structured situation, there is a clear model of the local-to-global transition.

2.2. The length of a curve. We mentioned previously that a curve was a set extending along its length. How can we compute such a length? And even before that, does such a measure exist for a particular set? First we review what is usually used to compute the length of a parametrized simple curve and then we present a non-parametric algorithm to calculate the length of a linear set based on projections.

2.2.1. *Parametric.* Our formalisation of the intuitive definition of the length of a curve will be derived from the ancient device of inscribed polygons.

Definition 3.4 (Partition and its norm [Tricot 1995]). Let $[a, b] \subset \mathbf{R}$, then a *partition* $\mathcal{P}([a, b])$ is a set of points $\{t_0, t_1, \cdots, t_n\}$ such that

$$a = t_0 < t_1 < \ldots < t_n = b$$

The *norm* $|\mathcal{P}|$ of a partition is defined as being

$$|\mathcal{P}| = \max(t_{i+1} - t_i), i = 0, 1, 2, \ldots, n-1$$

Given Γ, a simple curve in $\mathbf{R}^2$, let $\mathcal{P}_n$ be a sequence of partitions such that $\lim_{n \to \infty} |\mathcal{P}_n| = 0$. Define

$$L(\mathcal{P}_n, \Gamma) = \sum_{i=0}^{n-1} |\gamma(t_{i+1}) - \gamma(t_i)|$$

From the triangle inequality we see that the insertion of new points of subdivision will produce an increase in $L(\mathcal{P}_n, \Gamma)$.

Definition 3.5 (Jordan length [Burkill & Burkill 1970]). If $L(\mathcal{P}, \Gamma)$ is bounded for all dissections $\mathcal{P}$ of $[a, b]$, the *length* of a curve Γ in the Jordan sense is as follows:

$$L(\Gamma) = sup_{\mathcal{P}} L(\mathcal{P}, \Gamma).$$

It is sufficient to obtain the length from a limit of a sequence of partitions:

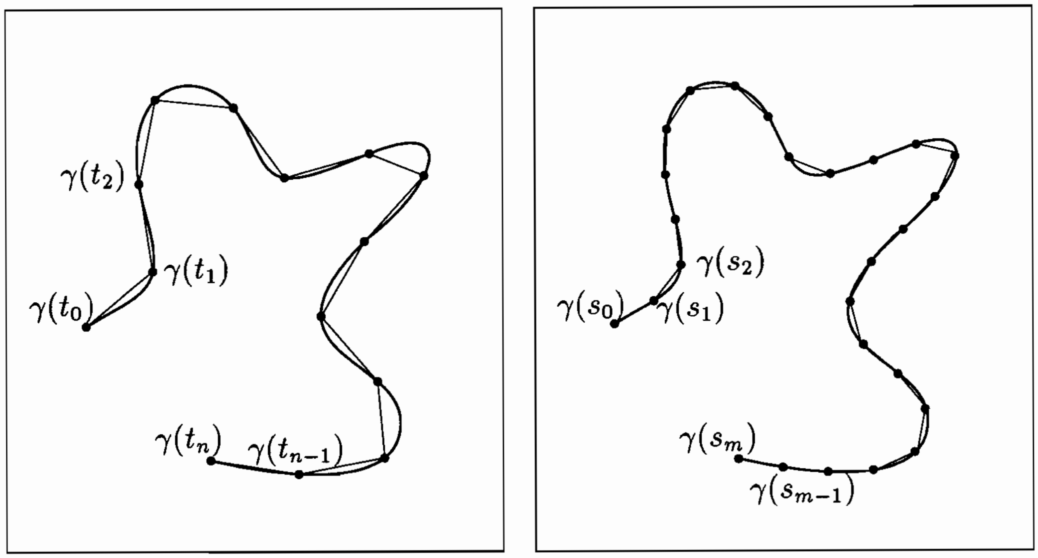
Theorem 3.1 (Length as a limit [Tricot 1995]). *If $(\mathcal{P}_n)$ is a sequence of partitions such that $\lim_{n \to \infty} |\mathcal{P}_n| = 0$ then*

$$L(\Gamma) = \lim_{n \to \infty} L(\mathcal{P}_n, \Gamma).$$

This last theorem implies that the length is independent of the choice of polygonal approximations as these get finer and finer.

Definition 3.6 (Rectifiability [Burkill & Burkill 1970]). A curve Γ is called *rectifiable* if it has finite length in the Jordan sense.

Remark 3.2. It is interesting to see that the word RECTIFIABLE derives from the Latin word *rectus* which means STRAIGHT. In French, the expression RECTIFICATION D'UNE COURBE means calculating the length of a curve as if it were a straight line segment. Not being able to unfold a curve into a straight line segment implies that the curve is not rectifiable.



(a) coarse partition **(b) fine partition**

FIGURE 3.4. Two different partitions leading to two different approximations of the length of a simple Jordan curve. If the curve is rectifiable, the finer the partition gets, the more accurate the estimate of length will be.

A few more results need to be mentioned. These provide the invariance properties one would like for the calculation of length and link a local notion (the tangent) to a global one (the length):

(i) **Length and rigid body motions:** the length of a curve is invariant under rigid body motions, i.e. translations and rotations;

(ii) **Length and arclength [Smith 1971]:** if the mapping Γ is differentiable, then

$$L(\Gamma) = \int_a^b |\Gamma'(t)|dt$$

i.e., the notion of length as just presented corresponds with the one of arclength in differential geometry;

(iii) **Length and parametrization [Cesari 1965]:** length $L(\Gamma)$ is independent of the parametrization (Fréchet-independent), i.e., $\Gamma \sim \Gamma_1$ implies $L(\Gamma) = L(\Gamma_1)$.

2.2.2. *Non-parametric.* The definition of length through polygonal approximations has the disadvantage of relying on the existence of a parametrization. It can be shown that length is independent of the parametrization, but one was nevertheless needed to perform calculations. In vision, **no** such parametrization is given. Rather, for many applications, this is what is sought. We therefore now turn to various definitions of "linear measure" that are independent of the parametrization. The *Hausdorff measure* and the *integral geometric measure* (see [Morgan 1987]) are examples of such measures. The integral geometric measure takes its roots from the middle of the 19th century, when Cauchy (1850) discovered a parametrization-free algorithm to compute the length of a curve. His idea was to project the set to be measured on lines with different orientations. We begin with this historically-important algorithm which can be considered as the first step toward what is now called *integral geometry*:

Theorem 3.2 (Projective length [Cauchy 1850]). *Let Γ be a curve of length $L(\Gamma)$. Let p be a line of direction θ passing through the origin. Let $m(\Gamma, p)$ be the measure of the projection of Γ on p counting multiplicities*[1]*. Then*

$$L(\Gamma) = \frac{1}{4}\int_{-\pi}^{\pi} m(\Gamma, p)dp$$

Example 3.1. Take Γ to be a straight line of length l with orientation θ. Since the length is invariant under rigid body motion, we can translate the original curve and rotate so it sits on the x-axis. If p is a line with orientation ϕ, then $m(\Gamma, p) = l|\cos\phi|$ and we get

$$L(\Gamma) = \frac{1}{4}\int_{-\pi}^{\pi} l|\cos\phi|d\phi = l$$

Example 3.2. Take Γ to be a circle of radius r centered at the origin. Then $m(\Gamma, p) = 4r$ (twice the radius two times since we have to count multiplicities) for all p (see Fig. 3.5 for a given p), and we get:

$$L(\Gamma) = \int_{-\pi}^{\pi} rdp = 2\pi r$$

Cauchy went further and determined what would be the bounds on the error of the estimation of the length if one were to take only a finite number of lines for the projection.

[1] Counting multiplicities means that if a curve projects on a segment more than once, say n times, then the length of the segment is counted n times. See Ex. 3.2.

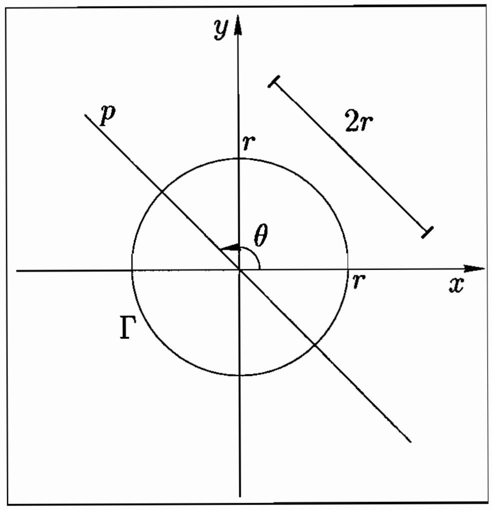


FIGURE 3.5. Cauchy projective length for a circle

Theorem 3.3 (Projective length approximation [Cauchy 1850]). *Let Γ be a curve of length $L(\Gamma)$. Let $\{p_1, p_2, \cdots, p_n\}$ be a collection of lines passing through the origin with orientation θ_i, such that $(\theta_i - \theta_{i-1}) = 2\pi/n$ (i.e., equispaced in orientation). If $m(\Gamma, p_i)$ is the measure of the projection of Γ on p_i counting multiplicities, then if $n > 2$*

$$\left| L(\Gamma) - \frac{\pi}{2n} \sum_{i=1}^{n} m(\Gamma, p_i) \right| \leq \frac{\pi}{2n^3} \sum_{i=1}^{n} m(\Gamma, p_i) \tag{3.1}$$

Remark 3.3. This theorem shows us the consequences of discretizing orientation space in terms of the computation of the length of a "linear" set (a curve in this case). To demonstrate this, let us take a simple example: the approximation of the perimeter of a circle. For every orientation defined by p_i, we have $m(\Gamma, p_i) = 2r$. Therefore the error is bounded by $\frac{L}{2n^2}$, where L is the perimeter. Choosing $n > 8$ in Eq. 3.1 ensures us to get an error smaller than 1%. Later in this document we will be interested in knowing what are the consequences of digitizing space coordinates.

We end this section by stating another version of Theorem 3.2 as presented in Steinhaus (1954) and do Carmo (1976). A line in the plane being uniquely determined by an orientation $\theta \in [-\pi, \pi]$, and a distance to the origin $\rho > 0$, we can consider the "area" of a set in this strip. Given a line defined by (p, θ), let $m(\Gamma, \theta, \rho)$ be the number of its intersection points with the curve Γ (that we call *multiplicity*). The

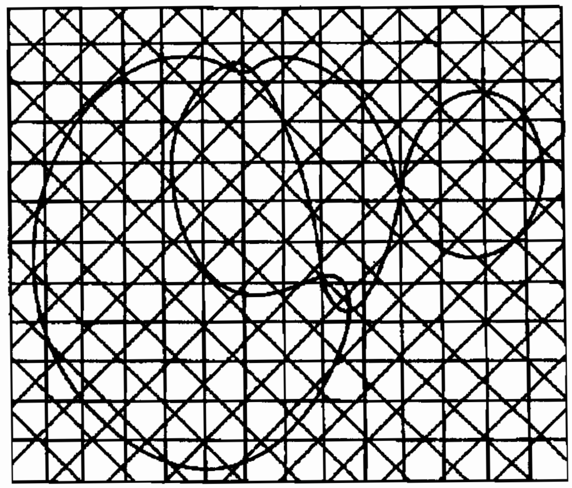


FIGURE 3.6. Estimating length from projections. A grating with parallel lines is overlaid on the set. Then count the number of intersections repeat by rotating the grating. Integrating the result provides an estimate of the length of the set. Figure reproduced from [do Carmo 1976].

measure of a set of straight lines (counted with multiplicities) which meet with Γ is related to $L(\Gamma)$:

Theorem 3.4 (Cauchy-Crofton formula [do Carmo 1976]). *Let Γ be a curve with length $L(\Gamma)$. The measure of the set of straight lines (counted with multiplicities) which meet Γ is equal to $2L(\Gamma)$, i.e.*

$$L(\Gamma) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{0}^{\pi} m(\Gamma,\theta,\rho)d\theta d\rho. \tag{3.2}$$

The Cauchy-Crofton formula allows us to dismiss the condition of studying only a simple Jordan curve. In fact Eq. 3.2 can define the *length* of any plane set E in all cases for which the function $m(\Gamma,\theta,\rho)$ is integrable. Thus, we have arrived at a definition of length that does not rely on any *a priori* parametrization. It is also independent of the notion of a tangent, of a derivative and of the approximation of the set by inscribed polygons. This definition covers more than rectifiable arcs: a finite or an enumerable set of such arcs gets a length in this new sense, provided the sum of lengths of its components is finite [Steinhaus 1954]. In the next section, we will elaborate on the definition of such sets; they will be called *1-sets*.

Remark 3.4. The Cauchy-Crofton theorem can be used to develop an algorithm for estimating the length of curves. Imagine having a curve Γ drawn on a piece of paper for which an estimate of length is needed. Take a transparent sheet on which is drawn family of parallel straight lines equispaced with distance d (we will call it a *grating*). Overlay this grating on the original drawing and count the number of intersections with the curve Γ. Rotate the grating and repeat the counting process for k orientations π/k. If m is the total number of intersections of a curve Γ (over all orientations), then

$$\hat{L}(\Gamma) = \frac{1}{2} m d \frac{\pi}{k} \tag{3.3}$$

is a good approximation to Eq. 3.2 [do Carmo 1976]. An example of this measurement process is shown on Fig. 3.6. The number of intersection points found is 153, the separation between the lines is $d = 7$ millimeters (mm) and the number of orientation is $k = 4$, therefore the estimate

$$\frac{1}{2} m d \frac{\pi}{k} = \frac{1}{2}(153)(7)\frac{\pi}{4} \approx 421 \ mm$$

while the actual value is 413.4 mm (adapted from [do Carmo 1976, p.46]).

3. Curve-like sets in geometric measure theory

As we mentioned before, we need a wider class of objects as an underlying model for curve recovery. Simple curves are too restrictive since they do not allow multiple curves and various kinds of discontinuities that are key in our description and understanding of the visual world. Even some of the simplest patterns could not be expressed by the Jordan definition (see Figs. 1.1 and 1.2). Mathematicians have however described a wider class of objects which would be more suited for our needs and these are called 'curve-like sets' (regular sets with finite positive Hausdorff measure). Instead of considering a mapping, we will rather consider sets. The notion of length will be kept implicit: it will be one of the basic requirements for these sets to exist. This section is a brief introduction to curve-like sets. It starts with the definition of the Hausdorff measure, and leads to the one of 1-sets, or those with finite length. Then we present density properties which will provide a hierarchy for one-dimensional sets. It is within this hierarchy that the type of sets to be considered

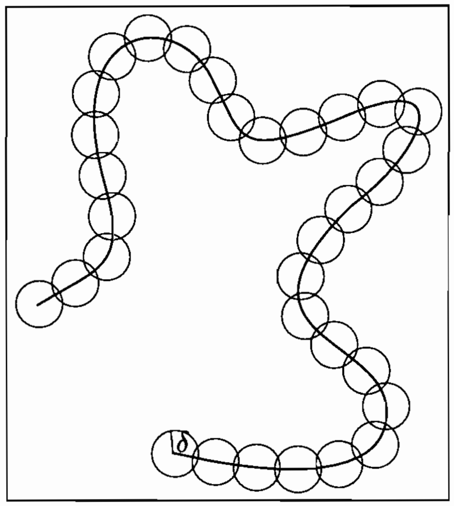


FIGURE 3.7. Parametrization-free approximation of the length of a set by a covering with δ-balls: a first step to the calculation of the Hausdorff measure.

throughout this thesis arise: the *curve-like sets.* A fine study of the local structure of curve-like sets will lead in the next section to constraints on the distribution of tangents and discontinuities.

3.1. Hausdorff measure. One way to compute the length, area or volume of an object is to use the Hausdorff s-dimensional measure $\mathcal{H}^s$, where, in the case of a smooth rectifiable curve, $s = 1$, in the case of a surface, $s = 2$ (classical references for this are [Rogers 1970, Federer 1985], but one can also look at [Falconer 1987], a more readable presentation). Consider the problem of defining the length $\mathcal{H}^1$ of a set E in the plane. Hausdorff's idea was to cover the set with small circles and to take the sum of the diameters (Fig. 3.7). If the balls are restricted to be smaller than some given value $\delta > 0$, and if the 'most economical' covering is chosen, we get an approximation of the length of the set at resolution δ. Allowing arbitrary covers, instead of covers by balls, gives us an outer measure, and for $\delta > 0$ we write

$$\mathcal{H}^1_\delta(E) = \inf \sum_i |U_i|$$

where $|U|$ is the *diameter* of U, (i.e., $|U| = \sup\{|x - y| : x, y \in U\}$) and $\{U_i\}$ is any sequence of sets of diameter less than δ covering E. The infimum here is taken over all (countable) δ-cover $\{U_i\}$ of E. It can be shown that $\mathcal{H}_\delta(E)$ increases as δ decreases, therefore:

Definition 3.7 (Hausdorff measure [Falconer 1987]). The *one dimensional Hausdorff measure* of E is given by

$$\mathcal{H}^1(E) = \lim_{\delta \to 0} \mathcal{H}^1_\delta(E) = \sup_{\delta > 0} \mathcal{H}^1_\delta(E).$$

Since no confusion will arise in this thesis, we will write $\mathcal{H}$ for $\mathcal{H}^1$.

Remark 3.5. One can show that the Hausdorff measure is in fact a *measure* in the measure theoretic sense [Falconer 1987]. It then can be used to define the notion of "almost". In this thesis we will use the terms "for almost all x in E" and "almost everywhere" (sometimes denoted a.e.). This means that the property applies for all $x \in E$, except maybe on a (very small) set G with $\mathcal{H}(G) = 0$. When writing $\mathcal{H}$-almost everywhere, or $\mathcal{H}$-a.e., we want to emphasize that this is with respect to the Hausdorff measure $\mathcal{H}$, and not with respect to another measure (the Lebesgue measure, $|\cdot|_1$, for instance). The term "almost nowhere" (which is used only once in this document) means that the property holds at most on a set of measure 0.

One might wonder if the Hausdorff measure coincides with the Jordan length for simple Jordan curves, or with the Lebesgue one-dimensional measure for Lebesgue-measurable subsets of the real line.

Theorem 3.5 (Hausdorff, Lebesgue and Jordan measures [Falconer 1987]). *If Γ is a curve, and E a Lebesgue measurable subset of* $\mathbf{R}$ *then*

(i) *the Jordan length L and the Hausdorff measure coincide, i.e. $\mathcal{H}(\Gamma) = L(\Gamma)$.*

(ii) *the Lebesgue measure $|\cdot|_1$ and the Hausdorff measure coincide, i.e. $\mathcal{H}(E) = |E|_1$.*

Hausdorff measure permits a classification of sets. One of its most popular uses is as the basis for *Hausdorff dimension*, an important construction which was used as an abstract formulation for the concept of dimension. Among other results, it led to the definition of a class of sets called 'fractals' [Mandelbrot 1982]: sets with non-integer Hausdorff dimension. In this chapter, we are not considering non-rectifiable curves and rather concentrate on those with positive but finite measure:

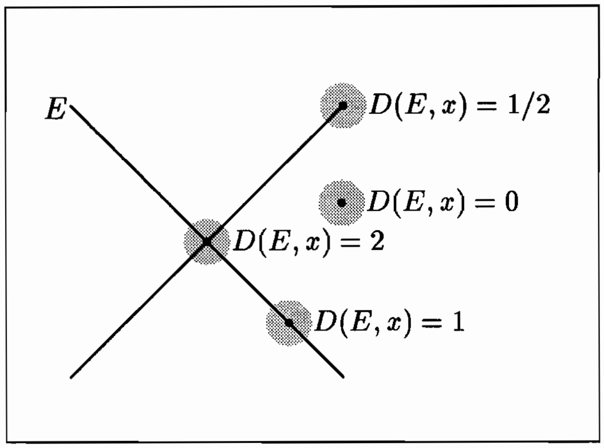


FIGURE 3.8. Example of an set where the density is not always 1: the truncated cone in $\mathbf{R}^2$. The cross is the set studied. The grey regions are places where we wanted to focus attention for the density.

Definition 3.8 (1-set [Falconer 1987]). An $\mathcal{H}$-measurable set E with $0 < \mathcal{H}(E) < \infty$, will be called a *1-set* (originally called 'linearly measurable set' by Besicovitch (1928)).

3.2. Basic density properties. The notion of densities for sets will be used in the definition of the local approximation of a set (tangent). Intuitively, densities indicate of the local measure of a set compared with the expected measure [Falconer 1987]. The definition is as follows:

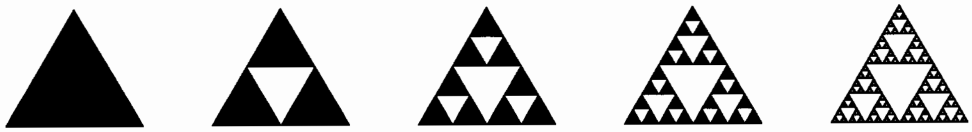
Definition 3.9 (Density [Falconer 1987]). Let $B_r(x)$ denote the closed ball of centre x and radius r. The *upper* and *lower densities* of a 1-set E at a point $x \in \mathbf{R}^2$ are defined as

$$D_u(E,x) = \limsup_{r \to 0} \frac{\mathcal{H}(E \cap B_r(x))}{2r}$$

and

$$D_l(E,x) = \liminf_{r \to 0} \frac{\mathcal{H}(E \cap B_r(x))}{2r}$$

respectively. If $D_u(E,x) = D_l(E,x)$, we say that the *density* of E at x exists and write $D(E,x)$ for the common value.



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

FIGURE 3.9. Example of an irregular 1-set in $\mathbf{R}^2$: the Sierpinski triangle.

Example 3.3. To better understand the previous definition, let us consider the following subset of $\mathbf{R}^2$

$$E = \{(u,v) \in B_r((0,0)) : u^2 = v^2\}$$

has density

$$D(E,x) = \begin{cases} 1 & \text{for } x \in E \setminus \{(0,0), \text{end points}\}, \\ 0 & \text{for } x \notin E, \\ \frac{1}{2} & \text{for } x \in \{\text{end points}\}, \\ 2 & \text{for } x = (0,0) \end{cases}$$

Note that in the last example the density is zero when outside the set and non-zero otherwise. In fact, one of the most interesting results about densities is that the density is zero almost everywhere outside the set:

Proposition 3.1 (Falconer (1987)). *If E is a 1-set in $\mathbf{R}^2$, then*

(i) $D(E,x) = 0$ *at $\mathcal{H}$-almost all x outside E, and*

(ii) $2^{-s} \leq D_u(E,x) \leq 1$ *at almost all $x \in E$.*

The last proposition is used mainly in the structure study of one-dimensional sets. Requiring $D_u(E,x)$ to be greater than zero insures that we are almost surely on the set E.

Definition 3.10 (Regular and irregular sets [Falconer 1987]). A point $x \in E$ at which $D_u(E,x) = D_l(E,x) = 1$ is called a *regular* point of E; otherwise it is called an *irregular* point. A 1-set is said to be *regular* if $\mathcal{H}$-almost all of its points are regular, and *irregular* if $\mathcal{H}$-almost all of its points are irregular.

Remark 3.6. Examples of irregular 1-sets in $\mathbf{R}^2$ include constructions similar to the one of the Cantor set [Falconer 1987]. An example taken from Morgan (1987) defines $E \subset \mathbf{R}^2$ by starting with an equilateral triangle and removing triangles at different scales. Start with E_0, a closed equilateral triangular region of side 1 (Fig. 3.9a). Let E_1 be the three equilateral triangular regions of side 1/3 in the corners of E_0 (Fig. 3.9b). In general let E_{j+1} be the triangular regions, a third the size, in the corners of the triangles of E_j. Finally, let $E = \bigcap E_j$ (an approximation is shown in Fig. 3.9e).

E is a 1-set since the projection of each E_j onto the x-axis is the unit segment, therefore the projection of $E = \bigcap E_j$ is also the unit segment which gives that $\mathcal{H}(E) \geq 1$. As for the other inequality we have that E_j is covered by 3^j equilateral triangles of side $\left(\frac{1}{3}\right)^j$, therefore $\mathcal{H}(E) \leq 1$, confirming that E is a 1-set. The proof that E is irregular, i.e. that its density is different than 1 $\mathcal{H}$-a.e. on E, can be found in Tricot (1991).

1-sets and regularity lead to the cornerstone definition for our work. They constitute the basic types of objects to be studied within the rest of this thesis. Originally, they were called *Y-sets* by Besicovitch (1928) (and this nomenclature still persists in some books [Tricot 1995, Falconer 1987]), but we decided to adopt the nomenclature used by Falconer (1990), since it fits more closely to intuition:

Definition 3.11 (Curve-like set [Falconer 1990]). A 1-set contained in a countable union of rectifiable curves will be called a *curve-like set.*

This definition is more general than the one of Jordan. It allows for multiple curves and these can intersect. It does not require a parametrization, since it is rather based on the notion of a set. Moreover

Theorem 3.6 (Regularity [Falconer 1990]). *A curve-like set is a regular 1-set.*

This therefore assures us that curve-like sets are free of the potential curve-free structures that the more general class of 1-sets could contain. Curve-like sets will suit our needs for edge detection, where we know that some kind of curve structure

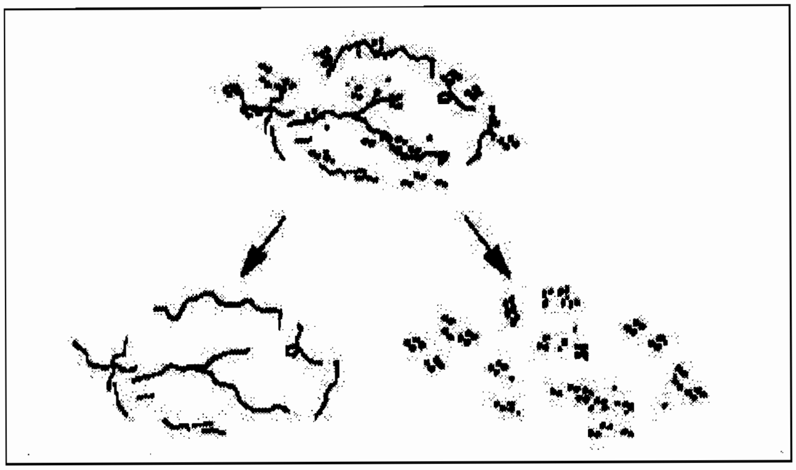
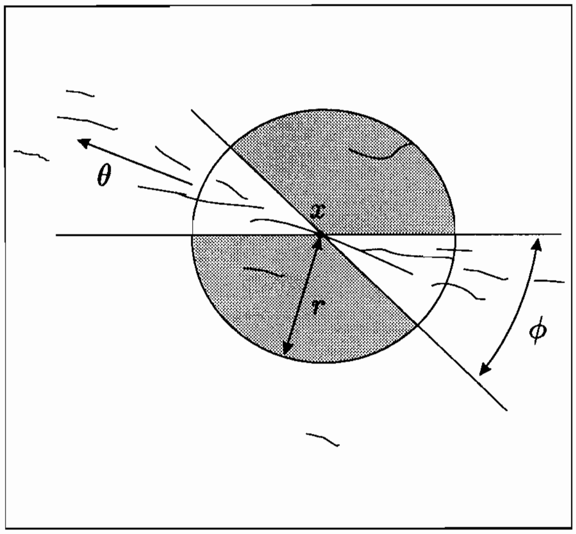


FIGURE 3.10. Decomposition of a 1-set. This figure, reproduced from Falconer (1990), illustrates the concept of decomposing a rectifiable one-dimensional set into a regular "curve-like" part and an irregular "curve-free" part.

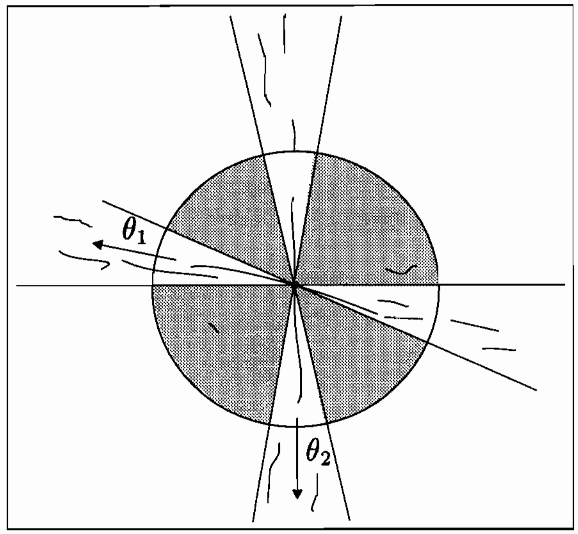
is present, since the set to inferred is provided by the output of oriented line/edge operators.

3.2.1. *Hierarchy for one-dimensional sets.* In the continuous domain, measures and densities have allowed mathematicians to partition the space of one-dimensional sets (not only the 1-sets), and to build a hierarchy for them. The first distinction, a rather crude one, is between those that have finite measure, the 1-sets, and those that have infinite length. Among the 1-sets, a finer subdivision provides regular (curve-like) and irregular (curve-free) sets. One nice result, called the *decomposition theorem* [Falconer 1990], enables a split of 1-sets into a regular and an irregular part, as shown in Fig. 3.10. It can be shown that each part from the set can be analyzed separately and then recombined without affecting density properties. The spirit of this decomposition is similar to what we will do with the discrete tangent map obtained though edge/curve detection. While our decomposition scheme will be different than the one presented here, the underlying idea is very similar. For vision applications, of course, further types of structures will be important. This will be discussed in detail in Chapter 5. Until then, the reader should keep in mind what was shown in Fig. 3.10.

3.3. Local structure of curve-like sets. Before discussing the existence of tangents for curve-like sets, we will present an alternate definition of a tangent that



(a) Besicovitch tangent $T(x)$



(b) Tangent set $\Theta(x)$

FIGURE 3.11. Illustrating the parametrization-free definition of tangents (Besicovitch tangent) at a point. Such definitions require that a significant part of E lies near x, of which a negligible amount lies outside the wedges. In (a) we illustrate the Besicovitch tangent at one scale, adapted from Falconer (1990), and in (b) the tangent set (multiple tangents) at one point.

does not rely on a parametrization of the set. This definition is due to Besicovitch (1928):

Definition 3.12 (Tangent: Besicovitch [Falconer 1987]). A curve-like set E has a *tangent* $T_B(x)$ at x in the direction $\pm\theta$ if

(i) $D_u(E, x) > 0$ and

(ii) for every angle $\phi > 0$,

$$\lim_{r \to 0} \frac{\mathcal{H}(E \cap (B_r(x) \backslash S_r(x, \theta, \phi) \backslash S_r(x, -\theta, \phi)))}{r} = 0 \tag{3.4}$$

where $B_r(x)$ is the ball of radius r centered at x, $S_r(x, \theta, \phi)$ is the sector of radius r at angle θ with opening ϕ, and $S_r(x, -\theta, \phi)$ is the sector in the opposite direction (it could have been written $S_r(x, \theta + \pi, \phi)$).

Suppose $x \in E$, then this definition means that at x the set E is locally concentrated around the line $T_B(x)$ with orientation θ passing through x. The first condition in this definition ensures that x is indeed on the set. The second condition ensures that the concentration is around the tangent line only. Fig. 3.11a illustrates the definition, namely that the second condition consists in looking at the rate of growth of what is found outside the local angular sector centered at x. If this rate of growth is much faster than r, the curve is ensured (from the first condition) to be concentrated around the line with orientation θ at x.

How does this definition of tangent relate to the usual definition of a tangent to a parametrized curve? In his seminal work, Besicovitch (1928) showed that this parametrization-free definition was indeed equivalent to the classical definition we presented in Section 2.

Theorem 3.7 (Besicovitch and classical tangent [Besicovitch 1928]). *Let E be a parametrized simple curve. If $x \in E$, and if both $T(x)$ and $T_B(x)$ exist, then the Besicovitch and the usual definition of the tangent at x correspond, i.e., $T(x) = T_B(x)$.*

PROOF. The original proof can be found in Besicovitch (1928), but a modern presentation can be found in Tricot (1995). □

Remark 3.7. Since we know now that the Besicovitch tangent and the usual tangent to a parametrized curve coincide, we will denote the tangent to a set E at x by $T(x)$ and always imply the Besicovitch construction.

One can now wonder if the Besicovitch tangent is solving some of the problems encountered with the classical definition for representing data obtained from edge detection. We will focus here on line endings and intersections. For this, let us recall Example 3.3 in which one of the lines was at an angle $\theta_1 = \pi/4$, while the other was at $\theta_2 = 3\pi/4$. Suppose we are at one of the end points of the line with orientation θ_1. The density $D(E,x) = 1/2 > 0$, and the rate of growth outside a sector is zero since

$$E \cap (B_r(x) \backslash S_r(x, \theta_1, \phi) \backslash S_r(x, -\theta_1, \phi)) = \emptyset$$

for all $0 < \phi < \pi/8$ and $r > 0$. A tangent at the end points is thus defined with the same orientation as for the rest of the line. At the intersection, i.e. for $x = (0,0)$, the density $D(E,x) = 2 > 0$, but

$$\frac{\mathcal{H}(E \cap (B_r(x) \backslash S_r(x, \theta_1, \phi) \backslash S_r(x, -\theta_1, \phi)))}{r} = 1,$$

for all $r < 1$ and $0 < \phi < \pi/8$, therefore the Besicovitch tangent at $x = (0,0)$ fails to exist. Since these events (curve intersections) are of paramount importance for the description of an edge map, we will present in the next section a wider representation for the local structure of a set than the Besicovitch tangent.

The study of the distribution of tangents for curve-like sets will be based on this critical result about the existence on tangents:

Theorem 3.8 (Tangent a.e [Falconer 1987]). *A curve-like set E has a tangent at almost all its points.*

SKETCH OF PROOF. The proof has several steps and is the subject of [Falconer 1987, chapter 3]. First, one proves that a rectifiable curve Γ has a tangent at almost all its points. This can be done using the following

Lemma 3.1 (Falconer (1987)). *If $\phi > 0$ and E is the set of points on a rectifiable curve Γ that belong to pairs of arbitrarily small subarcs of Γ subtending chords that make an angle of more than 2ϕ with each other, then $\mathcal{H}(E) = 0$.*

which characterizes the distribution of tangents for a rectifiable curve. It says that on a rectifiable curve, the chords defined by triples of points that are sufficiently close should almost never make a large angle between them. The existence of a tangent almost everywhere for a single rectifiable curve then follows from the continuity of the mapping. Once one knows that a rectifiable curve has a tangent almost everywhere, properties of densities, together with Thm 3.6 provide with the final result. □

Although for this thesis the central theorem will be Thm 3.8, we end our review of the classical results from geometric measure theory by the *structure theorem*, which constitutes a very deep result about the structure of arbitrary subsets of $\mathbf{R}^n$ [Morgan 1987]. We will cite here its one-dimensional version and a corollary that partly follows from the previous theorem. Recalling that a *continuum* is a compact connected set, we have

Theorem 3.9 (Structure theorem [Falconer 1987]). *If E is a continuum with $\mathcal{H}(E) < \infty$, then it consists of a countable union of rectifiable curves together with a set of $\mathcal{H}$-measure zero.*

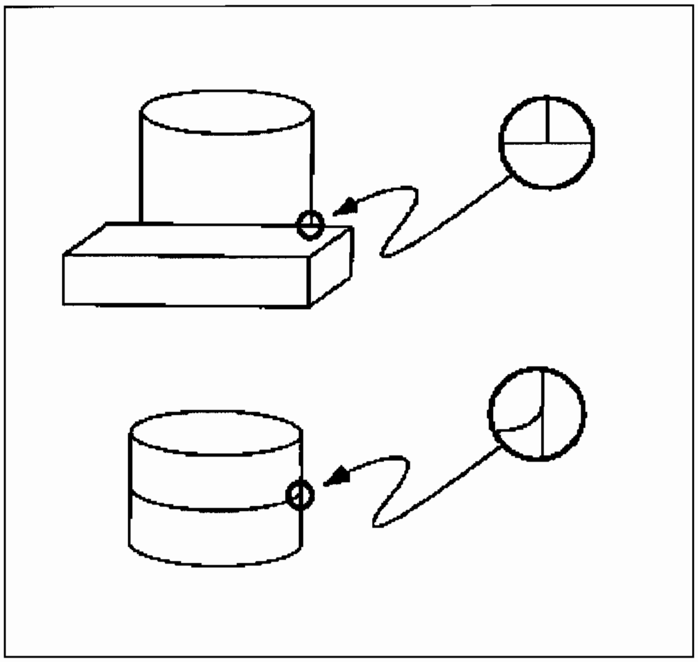


FIGURE 3.12. T-junctions. When objects occlude one another, they induce discontinuities in the edge map. T-junctions such as those shown in this figure are one type of these occlusion discontinuities. Reproduced from [Nitzberg et al. 1993].

Corollary 3.1. *If E is a continuum with $\mathcal{H}(E) < \infty$, then E is regular and has a tangent at almost all its points.*

The material we presented in this chapter so far is standard in elementary differential geometry and in geometric measure theory. We have briefly established links with relevant issues in computer vision, and have presented the mathematical apparatus needed for the development of our intermediate representation. The rest of this chapter begins the presentation of my original contributions.

3.4. Multiple tangents. The Besicovitch tangent must be extended for applications in computer vision. This is due to the fact that when objects occlude, they create discontinuities in bounding contours, leading to T-junctions [Guzman 1968, Waltz 1975, Nitzberg et al. 1993, Heitger & von der Heydt 1993] such as those presented in Fig. 3.12. At these points of discontinuity in orientation, "multiple tangents" must be represented [Zucker et al. 1989]. Intuitively the rationale is as follows: in the limit, as the discontinuity is approached from one side, one tangent is obtained, while from the other side, the second tangent is obtained. The following is an extension of the Besicovitch tangent to allow the representation of multiple tangents at a point:

Definition 3.13 (Multiple tangents). A curve-like set E has a *tangent set* $\Theta(x)$ at x if $D_u(E,x) > 0$ and for every angle $\phi > 0$,

$$\lim_{r\to 0} \frac{\mathcal{H}(E \cap (B_r(x) \backslash (\bigcup_{\theta\in\Theta(x)} [S_r(x,\theta,\phi) \cup S_r(x,-\theta,\phi)])))}{r} = 0 \tag{3.5}$$

but also, for each $\theta \in \Theta(x)$, $\exists\, r_0$ and ϕ_0 such that $\forall 0 < \phi < \phi_0$ and $0 < r < r_0$,

$$\limsup_{r\to 0} \frac{\mathcal{H}(E \cap (B_r(x) \cap (S_r(x,\theta,\phi) \cup S_r(x,-\theta,\phi))))}{2r} > 0 \tag{3.6}$$

As in the definition of the Besicovitch tangent, the density condition makes (almost) sure we are on the set. The condition given by Eq. 3.5 prevents things from being too crumpled around the point, while the third (Eq. 3.6) ensures that indeed there is something going on in the directions contained in the tangent set. Eq. 3.6 can be interpreted as the requirement that the conical density around each tangent direction be positive. In the case of the usual Besicovitch tangent (Defn 3.12), we know that this is true (see for instance [Morel & Solimini 1995]), therefore both definitions agree.

Remark 3.8. The value chosen for ϕ_0 will be linked to the orientation resolution of the operators when doing curve detection, while r_0 will be linked to their tangential extent (partly defining the scale of the operator).

EXISTENCE OF MULTIPLE TANGENTS. It is easy to build a set with multiple tangents. The set E from Example 3.3, for instance, has multiple tangents at $x = (0,0)$. In this case $\Theta(x) = \{\pi/4, 3\pi/4\}$. For both $\theta \in \Theta(x)$, we have for all $r > 0$ and $0 < \phi < \pi/8$

$$\mathcal{H}(E \cap (B_r(x) \cap (S_r(x,\theta,\phi) \cup S_r(x,-\theta,\phi)))) = r > 0,$$

while

$$E \cap \left(B_r(x) \backslash (\bigcup_{\theta\in\Theta(x)} [S_r(x,\theta,\phi) \cup S_r(x,-\theta,\phi)]) \right) = \emptyset$$

which ends the verification. □

Until now we did not put any constraint on the cardinality of the tangent set $\Theta(x)$ at a point. This first result addresses part of the issue:

Corollary 3.2 (Unique tangent a.e.). *If E is a curve-like set and $x \in E$, then the set of tangents $\Theta(x)$ at x is composed of a unique tangent for almost all x in E.*

PROOF. The proof follows from Thm 3.8. □

3.5. The tangent map. The definition of multiple tangents allows us to define a new structure essential for the development of our complexity measure. This structure will give the orientation of the set at each of its points:

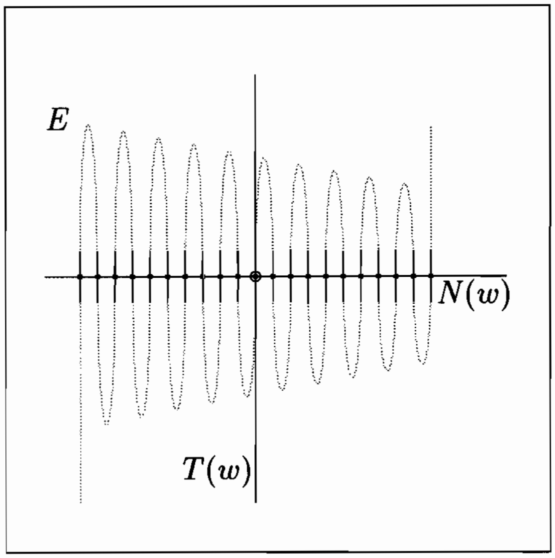
Definition 3.14 (Tangent map). Given a curve-like set E, the *tangent map* τ is

$$\tau = \bigcup_{x \in E} (x, \Theta(x)).$$

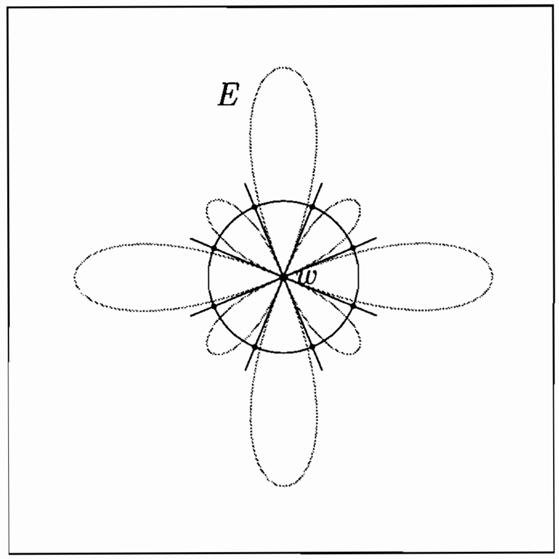
The tangent map will provide the mechanism for relating geometric structure to visual structure. The links will be provided by showing that the geometric structure is directly analogous to that obtained in computer vision. Thus we must define the tangent map in the discrete domain. This is how the structure we just defined will be linked to the output of edge detectors. The resulting intermediate representation will be the one used to characterize the complexity of the tangent (edge) map, and to provide a decision scheme for the representation underlying the grouping process. Before describing the discrete counterpart of the tangent map, we will investigate the structure of the underlying tangent map for continuous curve-like sets.

4. Tangent separation theorems

Rectifiability constrains the global and local distribution of tangents. Basically, for a one-dimensional set to be rectifiable, it cannot be too crumpled and cannot cut itself too often. The following theorems will try to capture this last statement, and will provide constraints on the underlying local approximation to a curve-like set. Recall that a set is *totally disconnected* if no two of its points lie in the same connected component. Thus, given any pair of points in the set, there is a decomposition into two disjoint closed subsets, each containing one of the points [Falconer 1987]. Now,



(a) parallel tangents separation



(b) multiple tangents separation

FIGURE 3.13. Illustration of the tangents separation theorems. The key idea is that for a curve to be rectifiable, the parallel or the multiple tangents cannot form a continuum. (a) illustrates the parallel tangents separation, where $w \in E$ is the point with the circle around it. The tangent $T(w)$ and the normal $N(w)$ are drawn and for this particular example the set $\mathcal{D}(w, E)$ is composed of all the other dots. (b) multiple tangents separation.

looking in the neighborhood of a tangent, in the normal direction for instance, we get the following:

Theorem 3.10 (Separation of parallel tangents). *Let E be a curve-like set. Consider a point w on E with tangent $T(w)$. If $N(w)$ is a line passing through w with orientation different than $T(w)$, then $\mathcal{D}(w, E) = \{y \in E \bigcap N(w) : T(w) = T(y)\}$ is a totally disconnected set.*

PROOF. We will proceed by contradiction. Suppose there exists a connected component $C \subset \mathcal{D}(w, E)$. Since $C \subset N(w)$ we have

$$L(C) = |C|_1 = \mathcal{H}(C) = \mathrm{diam}(C) > 0$$

where $|\cdot|_1$ is the usual one-dimensional Lebesgue measure and $L(\cdot)$, the Jordan length. Take now any point z inside C. For small enough ρ we get

$$\mathcal{H}(C \cap B_\rho(z)) = 2\rho$$

Therefore, if θ is the angle for $T(w)$, then we have for all ϕ sufficiently small

$$\lim_{r\to 0} \frac{\mathcal{H}(C \cap (B_r(z) \backslash S_r(z,\theta,\phi) \backslash S_r(z,-\theta,\phi)))}{r} \geq 1$$

which shows that there is no tangent at that point. This being true for all interior points of C, and since $\mathrm{diam}(C) > 0$, we have found a set of $\mathcal{H}$-measure > 0 for which there does not exist a tangent. However E being a rectifiable set, this contradicts the fact that it should have a tangent almost everywhere. □

This says that, if a curve gets squeezed into itself too much, it reaches a point where it may be difficult to know on which portion of the curve one is. The branches of the curve are tightly packed. When walking along the set, such as was tried for Paolina's hair in Chapter 1, one is confused.

Now consider the distribution of multiple tangents. More importantly, however, Theorem 3.11 will be the equivalent of Theorem 3.10 but in orientation space rather than in the spatial domain.

Theorem 3.11 (Separation of multiple tangents). *If w is a point on a curve-like set E, then $\Theta(w)$, the set of multiple tangents at w, is a totally disconnected set.*

SKETCH OF PROOF. Suppose there exists a connected component $C \subset \Theta(w)$. We can then find a circular arc in the neighborhood of w for which each point has a multiple tangent. That means we would be able to build a set of measure greater than zero with more than one tangent everywhere, which contradicts Theorem 3.2 since E is rectifiable. □

Remark 3.9. Proofs of Theorems 3.10 and 3.11 show the spirit of Poincaré's cut dimension idea [Poincaré 1926]. In both proofs (using contradiction) we showed that in order to disconnect the set, we needed a continuum; a straight line in the case of Theorem 3.10, an arc of circle in Theorem 3.11. This led to a contradiction because then the object couldn't be a rectifiable curve.

5. From curve-like sets to edge detection

One of the central features of the theory of curve-like sets was the Besicovitch tangent formulated in a parametrization-free manner. This is interesting for computer vision, because it is analogous to parametrization-free methods for estimating tangents, namely edge detection. We shall place tremendous emphasis on this analogy. In particular, just as Besicovitch sought a dense collection of points within a cone, "edge" operators seek a dense collection of pixels at a certain contrast [Marr & Hildreth 1980, Canny 1986, Hubel & Wiesel 1962*a*]. There are two important differences, however, and these differences will motivate the rest of this thesis. First is the notion of *scale*. The Besicovitch tangent was defined in the infinitesimal limit, and is related to a classification of curves as being either finite length (rectifiable) or infinite (non-rectifiable). Those with finite length were called 1-sets. The second difference is *resolution*. Orientation for the Besicovitch tangent is a real variable, as is spatial location; for any computation on a computer, these will be quantized numbers.

To develop the analogy between Besicovitch tangent sets and curve detection, we must face several subtleties. Finite scale and finite resolution have deep consequences, which we shall now attempt to illustrate. The result will both help us to frame the measure-theoretic problems that are appropriate for vision, and will lead to a statement of what we seek formally: discrete curve-like sets. To avoid the impression that all the mathematical questions are resolved, we also switch to a more informal style of presentation.

5.1. The paradox of length. Length is a discontinuous functional in the following sense [Steinhaus 1954]. In the vicinity of a rectifiable curve Γ, another curve Γ_1 can be defined whose length exceeds an arbitrary, previously defined limit, or even is infinite: this is what Steinhaus (1954) called the *paradox of length*. One aspect is a problem of numerical errors. If we measure the circumference of a circular object, we will not obtain πd, d the diameter, but rather something close to it. We know we are inaccurate, but we don't worry because, if a more accurate result is needed, we can just increase the level of precision in our measurement (as we showed in Section 2.2.2). Measurement requires units such as meters, yards: all idealized

FIGURE 3.14. An example of a set with no tangents almost everywhere. Named "la couronne flamboyante" after Dubuc (1982), this set doesn't have an integer Hausdorff dimension and constitutes an interesting example of a fractal set.

straight line segments (we saw that a rectifiable curve meant it could be "unfolded" somehow). Curved objects, such as circles, also have a definitive length that can be measured as accurately as necessary. Somehow, our experience is that objects which fit on a piece of paper have finite length, but this is a misleading intuition.

We introduced Hausdorff measure to study the local structure of sets, and to provide constraints on the distribution of tangents. However, it is more popularly known for its use in the study of 'fractals', or sets with non-integer Hausdorff dimension [Mandelbrot 1982]. Fractal curves are those in which tangents exist almost nowhere (in a measure theoretic sense). These led to very intricate figures such as Fig. 3.14. Returning to Mandelbrot's famous question: "How long is the coast of Britain?" [Mandelbrot 1970], we have that, in some cases, the length of certain objects depends on the precision of the measuring instrument. The length of the coast of Britain gets longer and longer as the resolution of the measuring instrument is increased. Could we neglect such objects with infinite lengths? The answer is 'no', since most curves encountered in nature are not rectifiable: they tend to be the rule rather than the exception. This statement is contrary to the belief that unrectifiable curves are an invention of mathematicians, and that natural curves are rectifiable: it's the opposite that is true.

Steinhaus's [1983] solution to overcome the paradox of length was to compute a length of finite order. The Cauchy-Crofton formula provided a technique to compute length by counting the number of intersections of a grating of parallel lines with the curve, and rotating this grating over a set of orientations (see Eq. 3.3). The exact formula integrates over all orientations, and the spacing between the lines is meant to be infinitesimal. It was mentioned that the length could be approximated by taking a positive spacing d, and averaging over a finite set of k orientations. To avoid the paradox of length, one can count the intersections up to some number n, and neglect subsequent intersections with the curve. Thus we get L_n, the length of order n. Steinhaus (1983) states that this concept is free from the paradox of length. If the resolution of the image is increased, the distance d between the lines in the grid is reduced, the number of orientations increased, then keeping the order fixed, the numbers L_n computed will approach more and more closely a definite limit: the ideal value of length of order n.

In computer vision we are confronted with a similar problem in the context of determining how, and when, to seek a transition from local representations to global ones. These are indeed the central questions of this thesis, and the answer is based on the observation that for curves, the computation of length is an archetypal local-to-global transition. But, as we learned from the Steinhaus paradox, the length computation is not always sensible or well-defined. Thus the central **philosophical proposal** on which this thesis is based can be summarized with the following statement:

> *The length of objects should only be considered when the result will be meaningful and of practical value.*

Curves in the world are projected on an image. At a given resolution and scale for detection, fractals and unrectifiable curves make no sense. Returning to our fractal set (Fig. 3.14), and projecting it onto an image with a given resolution, we obtain clusters of points, and few parts that extend along their length (the tips of the crown for instance, or the linking units). Now, applying edge detection (for this example,

6. Discrete equivalent to the tangent separation theorems

Theorem 3.10 has its discrete equivalent: if there is a discrete tangent at position $\hat{x}$, then there is a limit on the number of tangents that can be locally parallel to it (i.e., with the same orientation). The same applies at a single image location, where all tangents cannot exist at a point, giving a discrete equivalent to Theorem 3.11. The detection of edges or lines therefore constrains the discrete tangent map: not all tangents can be on, even if the image is of a bowl of spaghetti.

In the continuous domain, this constraint was defined as the fact that, given an orientation θ, one cannot find an *interval* in a *direction different from* θ for which one would have tangents with the same orientation θ. Two key changes need to be considered for the discrete equivalent:

- *different orientation:* Suppose the discrete orientations considered are $\theta_1, \theta_2, \cdots,$ θ_{N_θ}, and that a discrete tangent θ_i is on at a given position $\hat{x}$. An orientation θ_j will be said to be "different" if $d(\theta_j, \theta_i) \geq 2$ where
$$d(\theta_i, \theta_j) = \frac{|(\theta_i - \theta_j) \bmod_c r|}{\pi / N_\theta}$$
and where $r = \pi$ for lines, and $r = 2\pi$ for edges. Details about this distance between orientation cells can be found in [Iverson 1993];
- *interval:* The discretization induces a lateral spreading of tangents as we saw earlier. An interval here will therefore be defined as a set of M adjacent (8-connected) pixels within a given orientation. It is the value that M will take that will define the minimum size of contiguous pixels to be considered as an interval. For instance, if the resolution of the image is N, then a discrete interval is definitely smaller than N. To refine this assertion, we will need to use the scale of the operator $\sigma = (\sigma_N, \sigma_T)$.

Both these notions then lead to a conjecture about the size of the largest neighborhood over which information can spread laterally:

Conjecture 3.1 (Separation of discrete parallel tangents). Let $\mathcal{I}$ be an image with spatial resolution N. Let $\hat{\tau}$ be its discrete tangent map obtained through a bank of L/L operators for N_θ orientations with normal scale extent $\sigma_N < N$. If the response at $\hat{x}$ for orientation θ_i is positive, then there exists a constant k such that

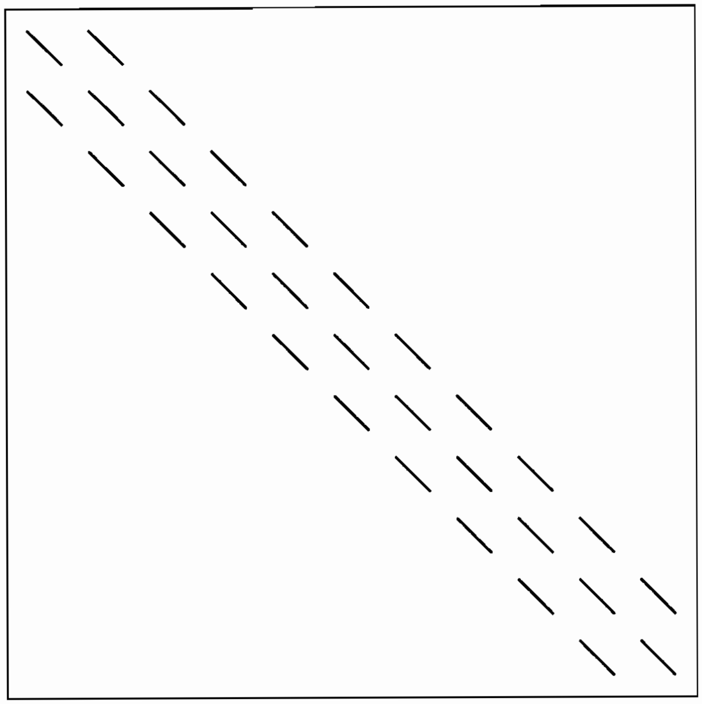


FIGURE 3.18. Discrete tangent map for a line segment at an orientation $\theta = 45°$ obtained from L/L operators. Notice the spreading of tangents over a 3 pixels connected linear neighborhood: i.e. drawing a line at 0° orientation can hit 3 contiguous tangents.

the coexistence of $M - 1$ contiguous tangents. What needs to be understood out of this conjecture is that at a given pixel, not all tangents (orientations) should have significant response at the same time.

This type of analysis of image operator's behavior is not completely new. Canny's [1986] original analysis presented something along these lines when trying to characterize x_{zc}, the mean distance between zero-crossing of f' and $x_{\max}$, the distance between adjacent maxima in the noise response to the filter f. This was used as a constraint to limit the number of peaks in the response.

7. Summary

Now the appropriate tools are gathered to attack the grouping problem and to set the ground for integration: namely choosing the appropriate support and representation. After briefly reviewing the notion of "dimension", we have seen that the Jordan curve definition was too restrictive, especially for passing from local to more global descriptions. An approach from geometric measure theory was then introduced, leading

CHAPTER 4

Characterizing complexity

In Chapter 3 we introduced curve-like sets through the Hausdorff measure and studied their local properties: densities and tangents. We also presented a discrete equivalent to curve-like sets, and interpreted this as the result of edge detection. In the same spirit as the decomposition scheme presented in Chapter 3, we will present here a finer classification of curve-like sets. This refinement will provide a means of deciding which representation should be used for a grouping process.

In order to assess the complexity of the tangent map, we will use a variant of an approach due to Minkowski. Originally Minkowski's approach consisted in covering a set with balls of radius ϵ, and computing the measure of the dilated set. The rate of growth of this measure as scale ϵ changes can be linked to the complexity of the set. Our approach differs in the way that the dilation is achieved; in particular, it will not be done isotropically. After reviewing the standard technique, we will present our variation. The end result will be a covering of the tangent map with oriented segments of "size" 2ϵ. The rate of growth of the measure of the dilated sets with respect to scale in a neighborhood of a fixed scale $\delta > 0$ will provide our normal and tangential complexity measures. This chapter is the heart of the thesis.

1. Minkowski polynomials

1.1. Isotropic dilations. Minkowski dilations are routinely used in mathematical morphology [Matheron 1975, Serra 1982], and for the estimation of fractal dimension [Falconer 1990, Tricot 1995]. The approach consists in creating a new set, which is the Minkowski sum of the original set with a dilating kernel, sometimes called a *structuring element*. The dilation is done isotropically over the set. More formally, given a set $E \subset \mathbf{R}^2$, we have

Definition 4.1 (Dilation [Serra 1982]). The *dilation*[1] of a subset E of $\mathbf{R}^2$ by another subset F of $\mathbf{R}^2$, called *structuring element*, is obtained through their Minkowski sum

$$E \oplus F = \{a + b, a \in E, b \in F\}$$

where $+$ here denotes a vector sum[2].

An example of the isotropic dilation of a curve with a ball is shown in Fig. 4.1. In fractal analysis, the case of a dilation with a ball is often called the *Minkowski sausage*. Dilations with other shapes (squares, segments, etc.) lead to *generalized Minkowski sausages* [Tricot et al. 1988].

Remark 4.1. In the case where the dilation is done with a ball as structuring element, we will write $E(\epsilon)$ for $E \oplus \epsilon B$. The resulting set can also be thought of as the set of points that are at a distance from E that is smaller than ϵ:

$$E(\epsilon) = \{y \in \mathbf{R}^2 : (\exists x \in E)\ |x - y| < \epsilon\}.$$

1.2. The Minkowski functional. When dilating with balls B of radius ϵ, the n-volume of the dilated object (for sets $E \subset \mathbf{R}^2$, the studied measure will be the area and will be denoted by $|\cdot|_2$) can be computed, and its rate of growth with respect to ϵ is related to the complexity of the set.

Example 4.1. To illustrate the concept we will take three elementary examples: a point, a segment of length l, and a circle of radius r. Dilating each set with a ball of radius ϵ gives us three different Minkowski sausages: a ball in the case of the point and the disk, a "wiener-shaped" object in the case of the segment (see Fig. 4.2). The corresponding areas for the dilated sets are:

[1] Matheron (1975) actually calls this 'dilatation', but the term 'dilation' is more common in mathematical morphology.

[2] Formally, the definition needs another twist. Considering $F_x = F \oplus \{x\}$, then the set $\{z : E \cap F_z \neq \emptyset\}$ of the points z such that E hits the translate F_z, is called the *dilation* of E by F. We deliberately omitted this step, since in this thesis our structuring elements will always be balls which are symmetric.

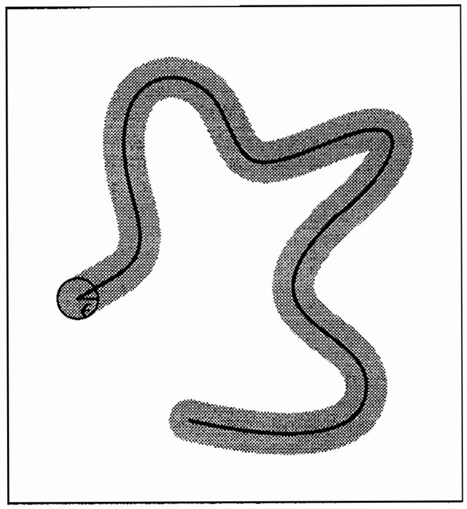


FIGURE 4.1. Illustration of a Minkowski dilation with a ball of radius ϵ. In black is the original set, in grey is the resulting dilated set. The resulting set is often called a *Minkowski sausage.*

- for the point

$$|E(\epsilon)|_2 = \pi\epsilon^2 \tag{4.1}$$

- for the segment of length l

$$|E(\epsilon)|_2 = 2l\epsilon + \pi\epsilon^2 \tag{4.2}$$

- for the disk of radius r

$$\begin{aligned} |E(\epsilon)|_2 &= \pi(\epsilon + r)^2 \\ &= \pi r^2 + 2\pi r\epsilon + \pi\epsilon^2. \end{aligned} \tag{4.3}$$

This simple example just illustrates that for some subsets E of $\mathbf{R}^2$,

$$|E(\epsilon)|_2 = \phi(E)\epsilon^0 + \psi(E)\epsilon^1 + \gamma(E)\epsilon^2, \tag{4.4}$$

where the functions ϕ, ψ, and γ are independent of the scale ϵ. Equations such as Eq. 4.4 are called *Minkowski functionals* [Matheron 1975]. In the case of image curves, we would like to detect which term dominates as ϵ is changed. Intuitively we have

(i) terms in ϵ^0: intersections and "dense" (space-filling) regions,

(ii) terms in ϵ^1: parts that extend along their length,

(iii) terms in ϵ^2: line endings, points of high curvature, sharp tangent discontinuities.

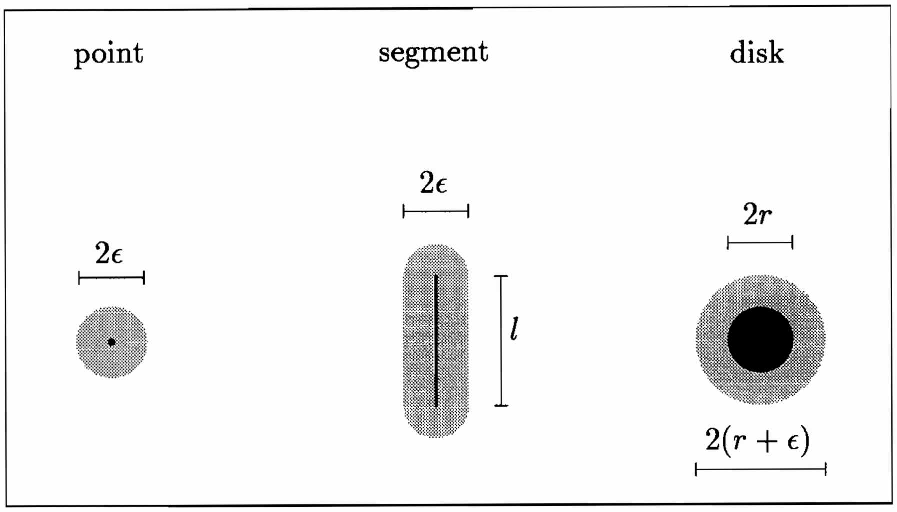


FIGURE 4.2. Dilating simple sets with a ball. Three elementary sets are shown with their corresponding dilations: (left) a point, (middle) a segment of length l, (right) a disk of radius r.

Suppose we were able to detect these regions. Then, for places in which surface information dominates (terms in ϵ^0), any attempt to integrate the information from the tangent map using one-dimensional support would not end up in an efficient representation. A field or density representation would be better to describe the image than counting the number of parts or following the path. To be more specific, let us recall the Perceptron example from Chapter 1. The interpretation with respect to what we just laid out could be as follows: when seen as a whole, the dominating term in the Perceptron spirals is the area covered, not the number of components, not the length. The areas of the dilated sets being more or less equivalent, the patterns are almost indistinguishable. Yet, the distinction between the two is the number of components: one has a single curve, while the other has two. Since the term corresponding to the number of components (term in ϵ^2) is not dominating, the differentiation between the two is hard.

When the dominating term is the one with respect to linear measure (length or perimeter), then grouping under a curve representation (i.e. with a one-dimensional support) results in a useful and efficient description of the scene. This is the case for global curve detection.

1.3. The Steiner formula. In some cases, the Minkowski functional can be linked to metric properties of the set. For instance, if the set E is a non-empty compact convex set, and if the dilation is done with a ball B (Minkowski sausage), then the area $|E(\epsilon)|_2$ is determined by the *Steiner formula* [Matheron 1975]:

$$\begin{aligned} |E(\epsilon)|_2 &= \sum_{k=0}^{2} \begin{pmatrix} 2 \\ k \end{pmatrix} W_k(E)\epsilon^k \\ &= W_0(E)\epsilon^0 + 2W_1(E)\epsilon^1 + W_2(E)\epsilon^2 \end{aligned}$$

where the W_k are called *the Minkowski coefficients.* For $n = 2$, we have that W_0 is the *area* of the set E, $2W_1$ is the *perimeter* and W_2/π is the *number of components* [Matheron 1975]. Recalling the examples for the point, segment and disk in the previous section, the last statements can be easily verified.

Although the Steiner formula looks very appealing, it cannot be applied in our case because the sets studied are non-convex. However, we would like to keep in mind the "idea" behind this formula, that different coefficients in the polynomial relate to metric properties of the set (the area, the perimeter and the number of components). We might then seek the functional for more complicated but still simple examples such as: (i) a pair of parallel lines; (ii) a pair of crossing lines; or (iii) a set of n equidistant parallel lines. We shall study these elementary sets to understand our complexity proposal, and to infer the various parameters needed for the analysis. This is the subject of Chapter 5.

1.4. Rate of growth and complexity. Our complexity measure will reflect the dominating term at a scale ϵ for $|E(\epsilon)|_2$. For instance, in the case of the segment of length l, in the neighborhood of 0, the term in ϵ^1 dominates, while in the neighborhood of ∞, the dominating term is the one in ϵ^2. At a given scale ϵ, if the functional grows[3] like ϵ^α, we let the complexity be $2 - \alpha$. In the small (around 0) the segment looks like a "line" ($2 - \alpha = 1$), in the large, i.e. seen from very far away (around ∞), the segment looks like a point ($2 - \alpha = 0$).

[3] The reader interested in a more complete treatment of rate of growths should consult [Tricot 1995]

1.4.1. *Log-log domain.* Given a set E, and its Minkowski sausage $E(\epsilon)$, the rate of growth around zero gives the well-known Minkowski-Bouligand[4] (fractal) dimension $\Delta(E)$. For $E \subset \mathbf{R}^2$, this is formally written as follows

$$\Delta(E) = \limsup_{\epsilon \to 0} \left(2 - \frac{\log |E(\epsilon)|_2}{\log \epsilon} \right) \tag{4.5}$$

Applying this formula to the point, segment, and disk examples, gives dimensions of 0, 1, and 2 respectively.

In applications however (i.e., when working with measured data), instead of estimating this limit, a standard practice is to go into the "log-log domain" and determine the rate of growth from the slope of a straight line fit to the data [Dubuc et al. 1989]. For the Minkowski-Bouligand dimension, the log-log data used are

$$\left(\log\left(\frac{1}{\epsilon}\right),\ \log\left(\frac{1}{\epsilon^2}|E(\epsilon)|_2\right) \right). \tag{4.6}$$

In practice, a finite list of scales $\{\epsilon_k\}$, $k = 1, 2, \cdots, K$, is used and the estimated fractal dimension $\hat{\Delta}(E)$ is obtained by fitting a straight line to the data points

$$\left(\log\left(\frac{1}{\epsilon_k}\right),\ \log\left(\frac{1}{\epsilon_k^2}|E(\epsilon_k)|_2\right) \right) \quad k = 1, 2, \cdots, K. \tag{4.7}$$

Since the fractal dimension is defined in a neighborhood of zero, the ϵ_k are chosen to be small. The ϵ and ϵ^2 in the denominators of Eqs. 4.6 and 4.7 are just a trick to obtain the dimension directly from the fit and not having to do a small arithmetic adjustment. Similar formulas will also be used for our measures of complexity.

Before jumping to our complexity measures, we will further illustrate the approach by providing log-log plots for the point, segment and disk examples (see Fig. 4.3). The data for these three sets (that can be calculated from Eqs. 4.1-4.3) are displayed on the same plot, together with their straight line fits. The slopes of these fits provide the estimate of the estimated fractal dimension $\hat{\Delta}(E)$.

Two different ranges of scale were selected, leading to two sets of estimates. The results are shown in Table 4.1. The first column gives the estimates for the scale range $2^{-15} \leq \epsilon \leq 2^{-10}$ (the points in Fig. 4.3 are those corresponding to this scale range). The results are clearly as expected. The ones for the coarser scale range, $2^{-5} \leq \epsilon \leq 2^{0}$, are more intriguing, since they differ from the expected values (i.e.,

[4] Also called the Cantor-Minkowski dimension.

E	$\Delta(E)$ $2^{-15} \leq \epsilon \leq 2^{-10}$	$\Delta(E)$ $2^{-5} \leq \epsilon \leq 2^{0}$
point	0.00	0.00
segment	1.00	0.75
disk	2.00	1.64

TABLE 4.1. Estimating the fractal dimension of simple sets in $\mathbf{R}^2$. This table gives the estimates obtained while fitting a straight line to the respective log-log data (Eq. 4.6). The difference between the two columns is the range of scales chosen. In both cases, 10 equidistant points in the log-log domain were computed.

the topological dimensions: a point is 0-dimensional, a segment is 1-dimensional, and a disk is 2-dimensional). One reason for this discrepancy is that, over this range of scales, the terms in ϵ^2 are coming into play (if this is not intuitive, the reader should go back to Eqs. 4.1-4.3 and 4.5), inducing errors in the estimation of the dimension. This behavior has been observed in the past, and was the main justification for the development of a new technique for estimating the fractal dimension of graphs of functions: the *variation method* [Tricot et al. 1988]. It was noticed that, in the case of the Minkowski sausage, when the scale was too large, it induced "rolls" on the dilated set, which themselves induced a concavity in the log-log plot, affecting the estimates of the dimension [Dubuc et al. 1989]. This is one of our main motivations for splitting the dilations into their normal and tangential components.

2. Oriented dilations

One of the key differences between our approach and the standard fractal analysis techniques is that the dilations will not be done isotropically, but will adapt to the local structure of the set. We call them *oriented dilations*, and we will show that they are necessary for separating sets of different complexity. Let E be a line segment. Fig. 4.4a shows the dilation with a ball, while Figs. 4.4b,c illustrate the normal and tangential dilations. The normal and tangential dilations become possible, of course, because we have a notion of (Besicovitch) tangent at each point.

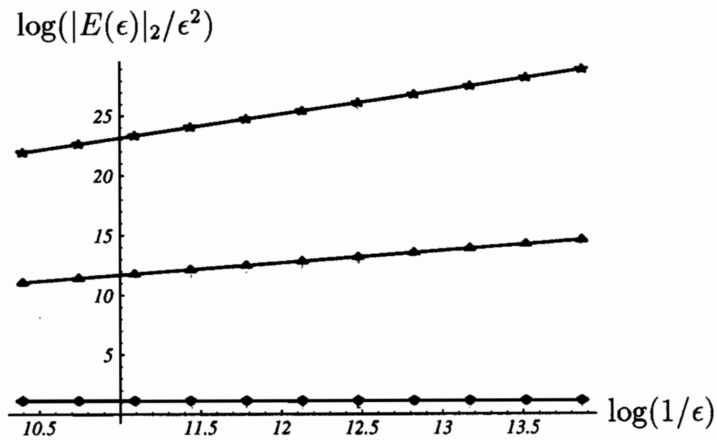


FIGURE 4.3. Obtaining the complexity from log-log plots. In this example we show the Minkowski-Bouligand log-log plots (Eq. 4.6) for a point (stars), a segment of length $l = 1$ (triangles), and a disk of radius $r = 1$ (diamonds). The rate of growth can be obtained from the slope of the straight line fit to the log-log data.

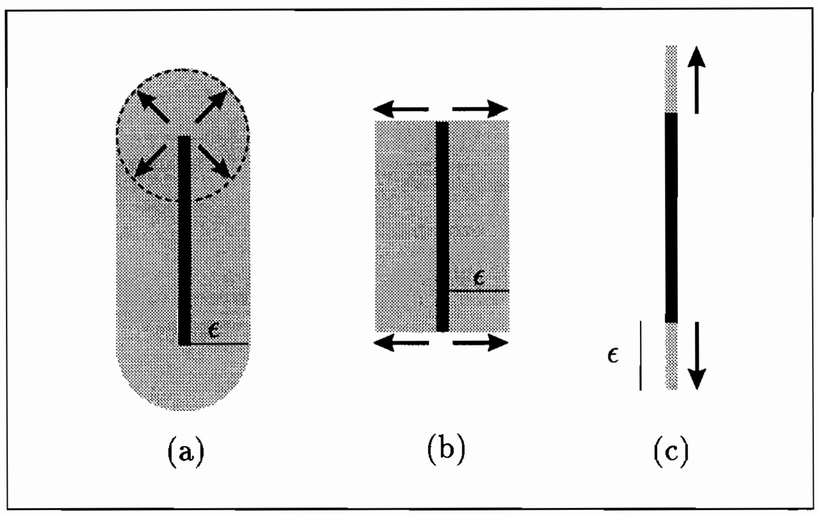


FIGURE 4.4. Isotropic and oriented dilations. (a) isotropic dilation with a ball of radius ϵ. (b) normal dilation (c) tangential dilation. Oriented dilations are possible because of the intermediate representation provided by the Besicovitch tangent sets.

Definition 4.2 (Normal and tangential dilations). Let E be a curve-like set, and τ its tangent map. The *normal dilation* $E_N(\epsilon)$ of E at a scale ϵ is the dilation of the set E with the segment $(-\epsilon, \epsilon)$ in the direction normal to the tangents $\theta \in \Theta(x)$ at x (Fig. 4.4b). The *tangential dilation* $E_T(\epsilon)$ is obtained by dilating E with the segment $(-\epsilon, \epsilon)$ in the direction of the tangents (Fig. 4.4c).

This departure from the standard Minkowski dilation approach will be essential for our analysis, since it will segregate the classification of curves from textures (using normal dilation), from the one of dust from curves (using tangential dilation).

2.1. Applying oriented dilations to test data. One way of building intuitions about the implications of oriented dilations and discrete complexity maps is through computational experiments. In the following examples, the dilation will be done at two scales, ϵ_1 and ϵ_2, with $\epsilon_1 < \epsilon_2$, that were carefully chosen to highlight various issues. The dilated sets will be displayed with two different greys: the darker grey for the smaller scale; and the union of the darker and the lighter grey for the dilation at the larger scale. The case of the Kanizsa pattern is shown in Fig. 4.5, where on bottom is displayed the result of the tangential dilation, and on the top is shown the normal dilation. A close-up of the process is shown in the middle and right columns of Fig. 4.5. Taking a closer look, we see first that in the neighborhood of the largest scale ϵ_2:

(i) the growth of the normal dilation within the grating patch has stopped because it is saturated (to verify this, just look at the top left panel). For the top and bottom parts of the rectangle, the growth can continue well beyond ϵ_2.

(ii) for the tangential dilation, the growth is saturated everywhere in the pattern, except at the four corners, and at the tips of the pinstripes.

A technical point now arises. The local approximations, the discrete tangents, have a length and a width that we call *normal* and *tangential* extents (denoted $\hat{\omega}_N$ and $\hat{\omega}_T$ respectively). These stay fixed when doing the dilations, but then have initial values that are set beforehand. In the top right panel, one can see the effect of setting the tangential extent smaller than the size of a pixel (it looks like a dotted pattern rather than a continuous line). In the bottom right panel, we see the effect of the normal extent: the width of the dilation. The chosen values for these two parameters will help overcome the effects of digitization as shown in the next chapter.

The second example shows a closeup of the tangential dilation for one of the Ullman discrete tangent maps. The results are shown in Fig. 4.6. What can be seen here (Fig. 4.6b) are the "star-shaped" patterns obtained on the small blobs, while, for the larger blob, the dilated set more or less resembles the original set (except for its

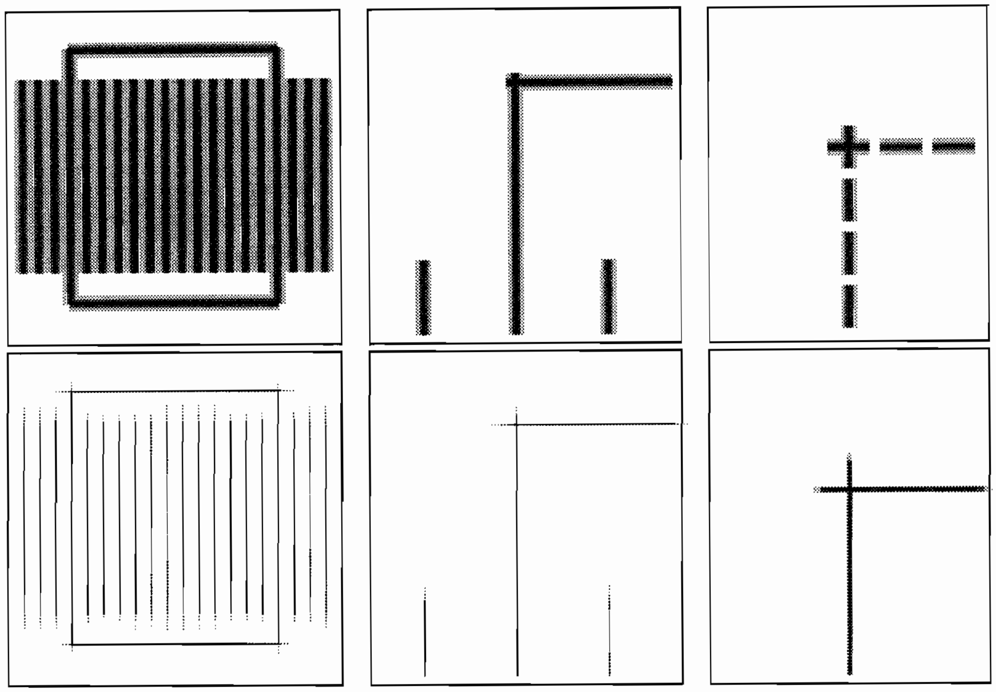
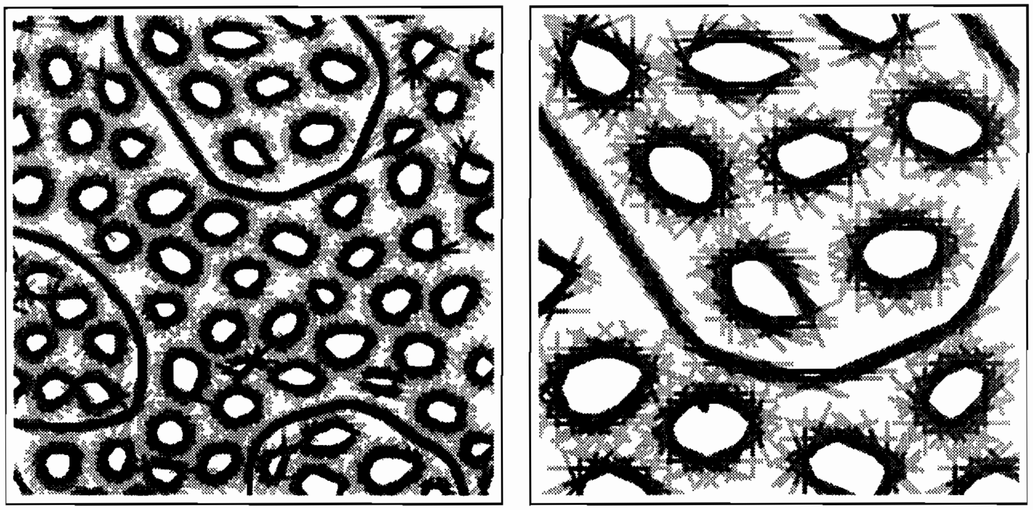


FIGURE 4.5. Oriented dilated sets for the Kanizsa pattern (Fig. 1.7a). Top row gives the normal dilation, while the bottom row is for the tangential dilation. Going from left to right shows increasingly large close-ups at the corner of the rectangle discrete tangent map dilations.

thickness). This is due to the difference in the rate of growth for these two types of objects, and illustrates the point that tangential dilations are sensitive to curvature, making them useful to detect regions of high curvature.

Fig. 4.7 displays the tangential and normal dilations on the hair and shoulder regions for the discrete tangent map of the Paolina image. In the shoulder region (Fig. 4.7a,b), the normal dilated set seems to grow linearly, while the tangential dilations saturate very early on. In the hair region (Fig. 4.7c,d), the normal dilation stops growing at an early stage, leading to a nil (zero) rate of growth. Being able to capture the rate of growth and select the appropriate scale, will provide anchors for perceptual grouping. This grouping process could then select regions that extend along their length (such as the shoulder), but could avoid textures (such as those encountered in the hair region). Although this is still an intuitive statement, it will be clearly laid out in the next chapter.



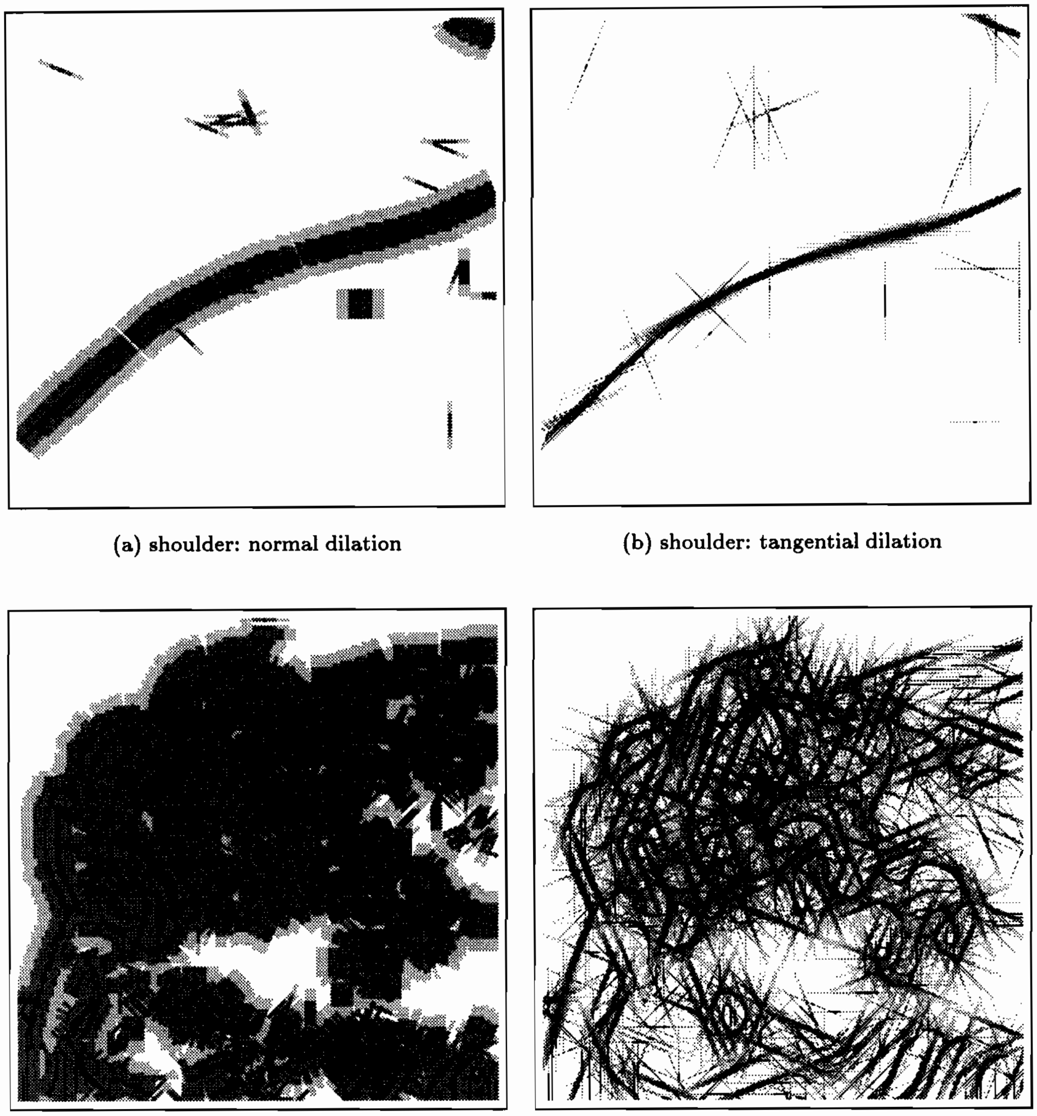
(a) tangential dilation (b) tangential dilation: zoom

FIGURE 4.6. Tangential dilation for the Ullman pop-out figure (Fig. 1.5a).

Density vs continuity. From the examples in the previous sections, we can provide a second justification for oriented dilations (the first being given in Section 1.4.1). The normal dilations provide a way of testing the *density* or *"space-fillingness"* of a set, while the tangential dilations will test for *continuity* (look back at Fig. 1.9). These concepts will be captured by the normal and tangential complexity indexes induced by the oriented dilations, in the neighborhood of a fixed scale $\delta > 0$.

3. A measure of complexity

3.1. The normal complexity. The local information contained in the tangent map τ can be used to calculate what we will call the *normal complexity* $C_N(\delta)$ of a curve-like set at a given scale δ. The main idea is to look at the rate of growth of $|E_N(\epsilon)|_2$ in the neighborhood dictated by a scale $\delta > 0$. In fractal analysis [Tricot 1995], the rate of growth of the measure of isotropically dilated sets, $E(\epsilon)$, is studied in the neighborhood of zero. If the area $|E(\epsilon)|_2$, is of order α, the fractal dimension is $2 - \alpha$. Here we rather consider the left derivative of the rate of growth, evaluated at $\delta > 0$, and call this the *normal complexity* $C_N(\delta)$ for E. We stress that, since our measures are evaluated at a finite scale, we refer to them as *complexities* rather than *fractal dimensions*, even though some of the formulas are analogous.



(a) shoulder: normal dilation

(b) shoulder: tangential dilation

(c) hair: normal dilation

(d) hair: tangential dilation

FIGURE 4.7. Oriented dilations for the Paolina subregions.

Definition 4.3 (Normal complexity). Let E be a curve-like set and τ its tangent map. Let $E_N(\epsilon)$ be the normal dilation of E at scale ϵ, and let $|\cdot|_2$ denote its area. Then the *normal complexity log-log plot* is defined as follows

$$\left(\log\left(\frac{1}{\epsilon}\right), \log\left(\frac{1}{\epsilon^2}|E_N(\epsilon)|_2\right)\right). \tag{4.8}$$

The *normal complexity*, $C_N(\delta)$ at scale $\delta > 0$, will be the left derivative of the normal complexity log-log plot evaluated at $\epsilon = \delta$.

Remark 4.2. Two aspects of this definition require clarification. First, the "trick" in the denominators is to ensure that the resulting number corresponds directly to the expected topological dimension for simple sets. We could have also considered the log-log plot

$$(\log \epsilon,\ \log |E_N(\epsilon)|_2)\,, \tag{4.9}$$

calculated the left derivative at δ, say α, but then we would have had to set $C_N(\delta) = 2-\alpha$. Secondly, (if one forgets for a moment about the *left* derivative), this definition can be interpreted as finding the tangent of the log-log plot at a given scale δ

$$\alpha = \left.\frac{d\log |E_N(\epsilon)|_2}{d\log \epsilon}\right|_{\epsilon=\delta}$$

To ensure existence however, we need to consider the left derivative.

Returning to Paolina (Fig. 4.7): in the hair region, the rate of growth α is approximatively 0 at the chosen scale δ, thus the complexity at this scale is $2-\alpha \approx 2$ (indicating a texture). For the shoulder, the rate of growth is linear ($\alpha \approx 1$), leading to a normal complexity of $2-\alpha \approx 1$ (indicating a curve). To develop a more solid intuition about normal complexity we next show how to calculate it for several basic patterns. These calculations are based on the following lemma:

Lemma 4.1. *Suppose* $|E_N(\epsilon)|_2 = a\epsilon^0 + b\epsilon^1 + c\epsilon^2$. *Then the normal complexity at scale* δ *will be*

$$C_N(\delta) = \frac{2a + b\delta}{a + b\delta + c\delta^2} \tag{4.10}$$

PROOF. From $|E_N(\epsilon)|_2$ we obtain the following normal complexity log-log plot

$$\left(\log\left(\frac{1}{\epsilon}\right), \log\left(\frac{a+b\epsilon+c\epsilon^2}{\epsilon^2}\right)\right)$$

Setting $x = \log(1/\epsilon)$ (equivalently $\epsilon = e^{-x}$), we can rewrite the previous equation as follows

$$(x, f(x)) = \big(x, \log(ae^{2x} + be^x + c)\big)$$

and

$$\begin{aligned}\frac{df}{dx} &= \frac{2ae^{2x} + be^{x}}{ae^{2x} + be^{x} + c} \\ &= \frac{2a + be^{-x}}{a + be^{-x} + ce^{-2x}}\end{aligned}$$

which leads to the desired result. □

Example 4.2 (A set of n equidistant line segments). If E is the set composed of n equidistant line segments of length l, where $\delta_0 > 0$ is the spacing between the lines, then

$$|E_N(\epsilon)|_2 = \begin{cases} 2nl\epsilon & \text{if } \epsilon < \delta_0/2 \\ 2l\epsilon + (n-1)l\delta_0 & \text{otherwise} \end{cases} \tag{4.11}$$

and, from the last lemma,

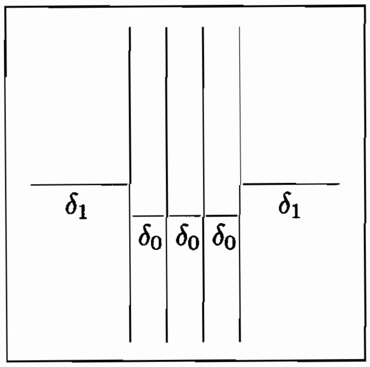
$$C_N(\delta) = \begin{cases} 1 & \text{if } \delta < \delta_0/2 \\ \dfrac{2\delta + 2(n-1)\delta_0}{2\delta + (n-1)\delta_0} & \text{otherwise.} \end{cases} \tag{4.12}$$

The result for a set E composed of four lines of unit length is shown in Fig. 4.8. Notice the discontinuity at $\delta_0/2$ and the decreasing behavior of the complexity after this point. We can show that $C_N(\delta)$ will be decreasing for $\delta \geq \delta_0/2$, therefore the normal complexity attains a maximum for $\delta = \delta_0/2$, for which

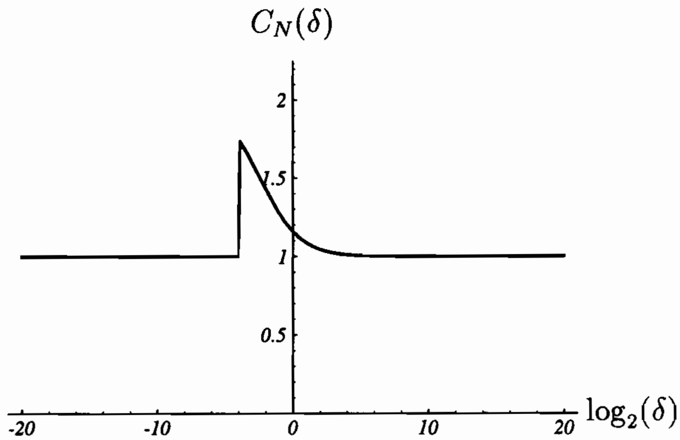
$$C_N(\delta_0/2) = 2 - \frac{1}{n}.$$

In the case of two parallel lines, the normal complexity will therefore be smaller than 1.5, and, for a given δ, we now know that, if the computed normal complexity is bigger than 1.5, the object is more "complex" than 2 parallel lines. In Chapter 5 we will generalize this example to arbitrary finite sets of lines with the same orientation.

The discontinuities in the last example are key to appreciate the behavior of the normal complexity (see Fig. 4.8b for instance). Remember, our definition involved taking the left derivative at a fixed scale instead of looking at the rate of growth in a



(a)



(b)

FIGURE 4.8. Plotting the normal complexity of a set of equidistant parallel lines. (a) the pattern to be studied: n equidistant lines with an inter-line spacing of $\delta_0 = 2^{-3}$ (b) its normal complexity: a discontinuity occurs at $\delta = \delta_0/2$ with $C_N(\delta_0/2) = 2 - \frac{1}{n}$.

neighborhood around zero or at infinity. What can insure that this definition is well formed? What values should it take and do these correlate with the complexity of the set?

From the structure of curve-like sets, we would like show that the normal complexity is indeed well-defined. For this we could first show that $|E(\epsilon)|_2$ is a concave function of ϵ and then, using a standard result from analysis [Valiron 1966], we would get that the left and right derivatives of the complexity log-log plot exist everywhere. A proof for the concavity of $|E(\epsilon)|_2$ would most probably use Theorems 3.10 and 3.11. This will provide the connection between Minkowski dilations, normal complexity and the Hausdorff measure of the curve-like set E. Our main conjecture is therefore

Conjecture 4.1 (Existence of normal complexity). If E is a curve-like set, then the corresponding normal complexity $C_N(\delta)$ exists for all $\delta > 0$.

Although we do not have a complete proof for this result, we will present here several results in support of it. First we recall the following Lemma from [Dubuc & Dubuc in press]:

Lemma 4.2. *If A is a bounded subset of the real line, then the Lebesgue measure of the Minkowski sum of A with $(-t, t)$, where $t > 0$, is a concave function of t.*

PROOF. Let $A(t) = \{a + x : a \in A \text{ and } |x| < t\}$. For all t, $A(t)$ is a bounded open subset of the real line, therefore it consists of a disjoint union of intervals, each of which is at least of length $2t$. If ϵ is positive, then $A(\epsilon)$ is a finite union of open intervals (a_i, b_i), $i = 1, 2, \cdots, N$. We can suppose that the a_i and b_i are such that $a_1 < b_1 < a_2 < b_2 < \cdots < a_N < b_N$. If $t > \epsilon$, we can compute the length of $A + (-t, t)$ as a sum of $N + 1$ functions f_i, where

- $f_0(t)$ is the measure of $A(t) \cap (-\infty, a_1)$,
- $f_i(t)$ is the measure of $A(t) \cap (a_i, a_{i+1})$, for $1 \leq i \leq N - 1$,
- $f_N(t)$ is the measure of $A(t) \cap (a_N, \infty)$,

with

$$
\begin{aligned}
f_0(t) \quad &= \quad t - \epsilon \\
f_i(t) \quad &= \quad \begin{cases} 2(t-\epsilon) + (b_i - a_i) & \text{if } t - \epsilon \leq (a_{i+1} - b_i)/2 \\ a_{i+1} - a_i & \text{if } t - \epsilon > (a_{i+1} - b_i)/2 \end{cases} \\
& 1 \leq i \leq N - 1 \\
f_N(t) \quad &= \quad (b_N - a_N) + (t - \epsilon).
\end{aligned}
$$

Since each of the f_i are concave, this is also true for the sum. This being true for all $\epsilon > 0$, we get the final result. □

Lemma 4.3. *If E is a curve-like subset of $[0, 1]^2$ containing only lines with the same orientation, then $|E_N(\epsilon)|_2$ is a concave function of ϵ.*

PROOF. Since E is composed only of curves with one orientation, there is only one normal that can be chosen. Let N_y be the line oriented normal to the pattern at y. Without loss of generality, we will assume that N_y is parallel to the x-axis. Since E is rectifiable, then from Theorem 3.10 we have that $E \cap N_y$ is a totally disconnected set and we can write:

$$E \cap N_y = \{\cdots, a_{-1}, a_0, a_1, \cdots\}.$$

From the last lemma we know that $|(E \cap N_y) \oplus (-\epsilon, \epsilon)|_2$ is concave. Since

$$|E_N(\epsilon)|_2 = \int_0^1 |(E \cap N_y) \oplus (-\epsilon, \epsilon)|_2 dy,$$

we obtain the final result. □

If E is a curve-like set for which $|E_N(\epsilon)|_2$ is concave, then the complexity log-log plot corresponding to normal dilation will be concave. A standard result from analysis (for instance, see [Valiron 1966]), that says that for a concave function the left and right derivatives exist everywhere could then be applied. However, the previous material does not provide a proof in the general case. Constraining the set to a bounded region can still be kept since it makes sense for computer vision applications. One would have however to allow different orientations (not only one orientation). The argument might then use the result with respect to the multiple tangents separation theorem. This time, given a point x subtending non empty tangent set, the intersection of a circle centered at x with the tangent map, will provide a totally disconnected set. We suggest using this to prove the concavity of the area of the normal dilation.

With this background, and accepting the existence of normal complexity, we now calculate bounds on its value. It is instructive, once again, to consider the case of finite linear sets. This will predict the values taken by our algorithm developed at the end of the chapter.

Theorem 4.1 (Normal complexity bounds for finite linear sets). *If E is a curve-like set composed of a finite number m of line segments, then the normal complexity takes values between 1.0 and 2.0.*

PROOF. Let the segments be $s_1, s_2, \cdots, s_m$, each of finite length $L(s_i)$. For the upper bound, consider α, the left derivative for $(\log\epsilon, \log|E_N(\epsilon)|_2)$ at $\epsilon = \delta$. Then $C_N(\delta) = 2 - \alpha$. Since $|E_N(\epsilon)|_2$ is a monotonic (increasing) function, then $\alpha \geq 0$, giving the upper bound 2.0.

For the second inequality, notice that, for any ϵ, the area of the dilated set can be split into two parts, a linear part (for the region where growth still occurs) and a constant part (for the saturated regions). Therefore the area can be written as $|E_N(\epsilon) = a + b\epsilon$. Since the number of segments is finite, there will be only a finite number of scales for which the parameters for the area will change. Then using Lemma 4.1, we obtain that

$$C_N(\delta) = 1 + \frac{a}{a + b\delta} \geq 1$$

since $a, b, \delta > 0$. $\square$

3.2. The tangential complexity . We just defined a complexity measure that was based on normal dilations. Its main feature was to determine whether a curve was dense (i.e., space-filling) or not. But the normal complexity alone is not sufficient (in the same spirit as length and shape were not sufficient for Ullman's saliency measure): we also need to test for continuity. This will be done through tangential dilations, which will enable discrimination between "curves" and isolated events (which we shall call "dust").

An intuitive illustration of why we need to consider the distribution of the tangents in the tangential direction (actually, it could be any direction, provided it is different from the normal direction) is as follows. Take a vertical, unit length, line segment centered in the unit square. Divide this segment into n subsegments of equal size, then shift each subsegment horizontally a random distance, but such that it still stays within the unit square. The resulting set has the same normal complexity as the original segment, but its structure is very different. A test on continuity would detect the difference, and this is one of the characteristics of the tangential complexity. Looking only in the normal direction (and then only using normal complexity) is therefore not enough since we do not know the structure of the set *a priori*. This was not the case in previous work [Dubuc et al. 1989], where the objects to be studied were graphs of *continuous non-constant functions*. Taking only the normal component was then sufficient to characterize the complexity of the curve.

Another important technical detail needs to be discussed. In the case of normal complexity, we considered the rate of growth of the *area* of the dilated set. For tangential dilations, the area of the dilated set is not the appropriate measure to take, since it is zero for some curve-like sets (the area of the tangential dilation of a line, for instance, is zero). To obtain a definition analogous to what we had for the normal complexity, we will take a measure of length of the dilated set. Since, in some cases, the dilated set can cover an area, we specifically consider the *perimeter* [Santaló 1976], i.e. the length of the boundary of the dilated set, $\partial E_T(\epsilon)$, denoted $|\partial E_T(\epsilon)|_1$.

Definition 4.4 (Tangential complexity). Let E be a curve-like set, and τ its tangent map. If $E_T(\epsilon)$ is the tangential dilation of E at scale ϵ, and if $|\partial E_T(\epsilon)|_1$ denotes its perimeter, then the *tangential complexity log-log plot* is defined as follows

$$\left(\log\left(\frac{1}{\epsilon}\right), \log\left(\frac{1}{\epsilon}|\partial E_T(\epsilon)|_1\right)\right). \tag{4.13}$$

The *tangential complexity* $C_T(\delta)$ at scale $\delta > 0$ will be the left derivative of the tangential complexity log-log plot evaluated at $\epsilon = \delta$.

To illustrate, let E be a set of lines with the same orientation. Dilate the set tangentially and look at the rate of growth of the perimeter for the dilated set. For instance, a line of length l becomes a line of length $l + 2\epsilon$ when dilated tangentially. Therefore we have

$$|\partial E_T(\epsilon)|_1 = 2(l + 2\epsilon)$$

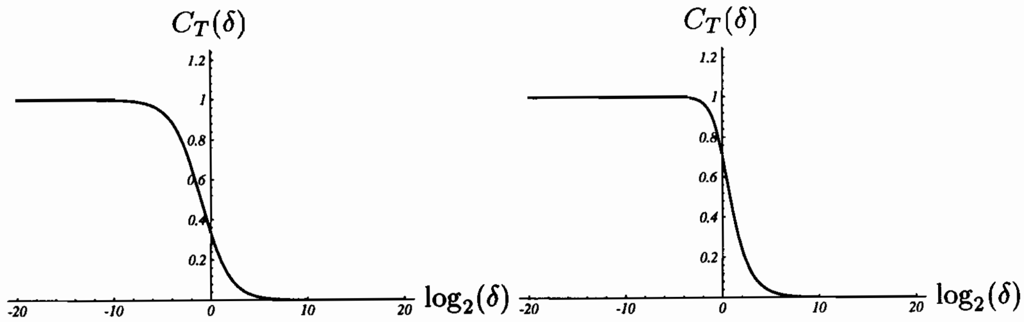
The "measure' chosen being the *perimeter*, we expect the rate of growth, α, to be between 0 and 1 (consider the example of the line segment to convince yourself). We defined $C_T(\delta)$, the tangential complexity, to be $1 - \alpha$. In our example, when l is large with respect to ϵ, the rate of growth is small and α is "close" to zero, and E has a "curve" structure: continuity and smoothness are verified. Otherwise, if l is small with respect to ϵ, then α is closer to 1 and the object would be better described as having a "dust-like" or curve-free structure. For large δ, i.e. from far away, a line is seen as a point. Thus it is necessary to consider the scale δ in the definition as well.

The following lemma is the tangential equivalent to what we had in Lemma 4.1. This will then provide answers to elementary calculations, and will confirm quite a number of things in the case of sets with a finite number orientations. Then, the case of a circle, where there is a continuum of orientations, will be studied. So we first start with

Lemma 4.4. *Suppose $|\partial E_T(\epsilon)|_1 = a\epsilon^0 + b\epsilon^1$. Then the tangential complexity at scale δ will be*

$$C_T(\delta) = \frac{a}{a + b\delta} \tag{4.14}$$

PROOF. The proof is similar to the one of Lemma 4.1 and was omitted here. $\square$



(a) tangential complexity: line segment

(b) tangential complexity: circle

FIGURE 4.9. The tangential complexity for (a) a single line of length $l = 1$, (b) a circle with radius $r = 1$.

Example 4.3 (Single line, equispaced lines). If E is a line of length l, then we have

$$|\partial E_T(\epsilon)|_1 = 2(l + 2\epsilon) \tag{4.15}$$

and from the last lemma we get

$$C_T(\delta) = \frac{l}{l + 2\delta} \tag{4.16}$$

A plot of $C_T(\delta)$ with respect to scale δ for a line of length $l = 1$ is shown in Fig. 4.9a. Notice the sigmoid shape which leads to the following observations:

(i) for a given line of length l to have a tangential complexity $C_T(\delta)$ bigger than some value $0 < \Delta_T \leq 1$, we must have

$$l > \frac{2\delta\Delta_T}{1 - \Delta_T},$$

therefore making it easy to predict conditions in which a set will be characterized as being represented as curves (as opposed to dust), an issue that will be further discussed in Chapter 5;

(ii) since l and ϵ (also δ) are coupled within this definition, we have that, for $l/\delta = c$,

$$C_T(\delta) = \frac{c}{c + 2},$$

or in other words, a line of length l has the same tangential complexity at scale δ as a line of length $2l$ at scale 2δ, for instance;

(iii) the tangential dilation is related to the idea that, as the scale becomes bigger, the half-perimeter of the dilated object diverges from the length of the curve.

Example 4.4 (Circle of radius r). Let E be a circle of radius r. Try to imagine the tangential dilated set: it will be an annulus exterior from the circle of width ζ (which itself can be calculated from ϵ and r). Moreover, we get that

$$|\partial E_T(\epsilon)|_1 = 2\pi \left(r + \frac{\epsilon}{\sin(\arctan(\epsilon/r))} \right)$$

leading to the following tangential complexity

$$C_T(\delta) = 1 - \frac{\delta^2}{r^2 + r\sqrt{r^2+\delta^2} + \delta^2}. \tag{4.17}$$

A graph of the tangential complexity for a circle of radius $r = 1$ as a function of scale is shown in Fig. 4.9b. Notice that the shape is similar to what we had for the line. Changing the radius shifts the function. We therefore conclude that the tangential complexity of a circle is close to 1 when the radius is large with respect to the scale. The opposite situation, namely a small circle considered at a large scale, looks like a point and its tangential complexity will be close to 0; exactly as we would expect.

Theorem 4.2 (Tangential complexity bounds for finite linear sets). *If E is a curve-like set composed of a finite number m of line segments, then the tangential complexity take values between 0.0 and 1.0.*

PROOF. Let the segments be $s_1, s_2, \cdots, s_m$, each of finite length $L(s_i)$. The upper bound follows from the monotonicity of $|\partial E_T(\epsilon)|_1$. The rate of growth is bounded above by growth of all the segment line endings. In the worse case we therefore have

$$|\partial E_T(\epsilon)|_1 = 2(2m\epsilon + \sum_{i=1}^{m} L(s_i))$$

and, from Lemma 4.4, we get the desired inequality

$$C_T(\delta) = \frac{\sum_{i=1}^{m} L(s_i)}{2m\delta + \sum_{i=1}^{m} L(s_i)} \geq 0.$$

□

In the next two chapters we will present results using a discrete implementation of Definition 4.4 and show that the tangential complexity will be sensitive to ends of lines, corners, and points of high curvature. The result saying that our definition is well-formed stays conjectural. We presented here the intuition behind the definition and showed that it made sense on elementary sets.

4. Mapping complexity

The complexity measures presented in the last sections depended on scale. This is a major departure from the standard 'fractal analysis' approach in which the rate of growth is studied around zero [Tricot 1995]. The other key difference is to make the complexity measure local in space, by computing it over a given compact region $\Omega(x)$ centered at x. The normal complexity at x can be obtained by first restricting the dilation to the region $\Omega(x)$:

$$E_N(x,\epsilon) = \{E_N^{\Omega}(\epsilon)\} \cap \Omega(x). \tag{4.18}$$

where $E^{\Omega} = E \cap \Omega(x)$. This means that the tangent sets considered from the tangent map are only those within $\Omega(x)$ and the dilations are limited to this region. The case of the tangential dilation is slightly more complex and consists in not dilating any further than the region $\Omega(x)$, therefore leading to $E_T(x,\epsilon)$.

Definition 4.5 (Complexity indexes). The *normal and tangential complexity indexes* at x over a region $\Omega(x)$ are denoted $C_N(x,\delta)$ and $C_T(x,\delta)$ and are obtained by looking at the rate of growth of their local oriented dilations. More formally they are obtained as the left derivatives of their corresponding log-log plots

$$\left(\log\left(\frac{1}{\epsilon}\right), \log\left(\frac{1}{\epsilon^2}|E_N(x,\epsilon)|_2\right)\right) \tag{4.19}$$

for $C_N(x,\delta)$, and

$$\left(\log\left(\frac{1}{\epsilon}\right), \log\left(\frac{1}{\epsilon}|\partial E_T(x,\epsilon)|_1\right)\right) \tag{4.20}$$

for $C_T(x,\delta)$.

Remark 4.3. In some cases, we will consider the entire set E within a bounding region Ω (in Chapter 5, for instance) . We then write $C_N(\cdot,\delta)$ and $C_T(\cdot,\delta)$ for the normal and tangential complexities restricted to the region Ω.

In a similar fashion as we did in the case of the tangent map, we can bundle all the complexity indexes and build our main tool, the *complexity map*:

Definition 4.6 (Complexity map). Given a curve-like set E and its tangent map τ, we define the *complexity map* C to be

$$C = \bigcup_{x\in E} (x, C_N(x,\delta), C_T(x,\delta)).$$

5. Computing discrete complexity maps

Chapter 3 introduced curve-like sets, and we just developed a tool to get a finer classification of these sets at a given scale through the complexity map. Issues now arise of how to calculate or approximate the tangential and normal complexity of discrete curve-like sets. We end this chapter by describing, step by step, the algorithm used to numerically estimate the maps on the discrete tangent maps obtained from the edge/line detection operators. Once the algorithm is clearly laid out, we discuss, in the next chapter, decisions that need to be made about proper parameter choices.

The algorithm consists of several steps. The dilations are done digitally, therefore the tangents must be projected onto an image. Once the tangent map is projected, the oriented dilations need to be performed and their measures need to be estimated. From the measures of the discretely dilated sets, the rate of growth is estimated through a best fit line in much the same way as presented in Section 1.4.1. The result of the fit provides the two complexity indexes, at each non-empty entry of the tangent map. Bundling these gives the discrete complexity map.

5.1. Step 1: Projecting the discrete tangent map. The line/edge operators as presented in Chapter 3 provide the discrete tangent map at a given scale σ. The first step toward the discrete complexity map consists in projecting the discrete tangents and generating the *tangent map image.* Each tangent within the discrete tangent map is attached to a pixel from the original image. We further subdivide each pixel into an array of $p_x \times p_y$ pixels and project each individual tangent from the discrete tangent map. Each discrete tangent can be thought of a "needle". It has an orientation $\hat{\theta}$, a width $\hat{\omega}_N$ (the *normal extent*) and a length $\hat{\omega}_T$ (the *tangential extent*). The resulting image is what we call the *tangent map image.*

As we will see in the next section, isotropic dilations could be done directly on the tangent map image but, for oriented dilations, the algorithm we used will link together the projection and the dilation into a single step.

5.2. Step 2: Dilating the projected discrete tangent map. In order to compute the complexity we need to look at the rate of growth of the area of a dilated set around a fixed scale $\delta > 0$. Many dilations will be done in the neighborhood of δ, so let $\{\delta_k\}$ be the sequence of scales in this neighborhood that will be used to obtain the straight-line fit of the corresponding log-log plot. In the following subsections, the case of isotropic dilations will be presented first, since it is standard; then the technique to approximate the oriented dilations will be addressed.

5.2.1. *Isotropic dilations using distance transforms.* The information in the discrete tangent map will be used to generate the oriented dilations. If one only needed isotropic dilations, the dilated sets could be obtained by applying a dilation on the tangent map image as is currently done in mathematical morphology [Serra 1982]. We would then choose the appropriate structuring element and dilate the tangent map with respect to it in the neighborhood of a scale δ. For instance if the structuring element was a disk (leading to a Minkowski sausage), then we would take a range of scales $\{\delta_k\}_k$ in the neighborhood of δ, and dilate with disks of radius δ_k by iterating through the list.

Although mathematical morphology operations are appealing theoretically, the problem can also be solved in practice by applying distance transforms [Borgefors 1984]. A distance transform is an operation on an input binary image (object – non-object points) over a domain S, which returns a real value, the minimum distance from the point to the object, given a particular metric. The first step then consists in choosing the desired metric [Dubuc 1988, Borgefors 1986], and then in applying a distance transform to the image. For instance, in the case of the Minkowski sausage, there exist simple algorithms that can give a good approximation to the Euclidean metric [Leymarie & Levine 1992]. Using one of these, one could first project the unit tangents on a discrete lattice and consider the result of the projection as the object points. The Euclidean metric distance transform of this binary image could then be used to assess the complexity of the object once the areas have been extracted.

Isotropic dilations however are not appropriate to compute the complexity map, but will nevertheless be used in Chapter 5 to obtain a rough estimate for the local extent over which to do the analysis. The algorithm just described in this section will therefore be recalled later.

5.2.2. *Approximating oriented dilations.* Section 2.1 stressed the need for oriented dilations. One suggestion would be to develop oriented distance transforms. But due to our choice of intermediate representation (the discrete tangent map), we can do it much more easily in a very straightforward manner. Given a PostScript[5] interpreter, it is possible to use the features of PostScript to generate the approximated dilated sets. The end result is not efficient but has the advantage of being fast to implement and easy to debug, therefore ideal to generate a prototype. The main idea is to repetitively draw the tangent map with the appropriate unit tangent properties (normal extent, tangential extent, color). The generation of the oriented dilated set will therefore incorporate the projection step as well (Step 1). We now describe our approach in greater detail.

Most drawing operations in PostScript are done through a "virtual pen" which has various properties: a width, a color, a tip shape, etc. Taking a wider pen generates wider lines. A red pen draws red lines, and so on. To simulate normal dilations, we will change the width of the drawing pen. To simulate tangential dilations, we will change the length of the unit tangents. Scale will therefore be linked to the width of the unit tangent (in the case of normal dilation), or to its length (in the case of tangential dilation). If for each scale δ_k we draw the tangent map by changing the pen parameters (its size and color). Then there will be a direct correlation between the resulting "colored" image and the oriented dilation of the set.

[5]PostScript is a device-independent page description language developed by Adobe Systems Incorporated which has become the industry standard for printing high-quality integrated text and graphics.

5.2.3. *Local extent.* Another issue that needs to be discussed is the one of local extent. The complexity indexes within the complexity map are computed over a bounded region $\Omega(x)$. Since we are working on a square lattice, we will take $\Omega(x)$ to be a square with side $\Omega < 1$ centered at x, therefore resulting in a square window of $\hat{\Omega} \times \hat{\Omega}$ pixels in image coordinates (where $\hat{\Omega} = \lfloor \Omega * N \rfloor$). For a given x in the image domain subtending a non-empty tangent set in the discrete tangent map, we will restrain the computations within a bounded window $\Omega(x)$ (an $\hat{\Omega} \times \hat{\Omega}$ window in image coordinates). Sliding the window over the image domain will eventually provide the discrete complexity map.

5.3. Step 3: Estimating the measures and rate of growth. Once the dilated sets are generated, calculating their measures is very straightforward. A simple histogram on the data provides the result. Integrating through the histogram gives estimates of the measures at the various scales δ_k in a neighborhood of δ. If $H_N(\delta_k)$ is the number of pixels covered for the (local) normal dilation and $H_T(\delta_k)$ is the one for the tangential dilation at scale δ_k, then the discrete normal and tangential complexity indexes, denoted $\hat{C}_N(\hat{x}, \delta)$ and $\hat{C}_T(\hat{x}, \delta)$ respectively, are given by a straight line fit through the data:

$$\left(\log\left(\frac{1}{\delta_k}\right), \log\left(\frac{1}{\delta_k^2} H_N(\delta_k)\right)\right), \quad k = 1, 2, \cdots, K_N \tag{4.21}$$

for the normal complexity and

$$\left(\log\left(\frac{1}{\delta_k}\right), \log\left(\frac{1}{\delta_k} H_T(\delta_k)\right)\right), \quad k = 1, 2, \cdots, K_T \tag{4.22}$$

for the tangential complexity. Bundling the indexes gives us the desired discrete complexity map:

Definition 4.7 (Discrete complexity map). Given an image $\mathcal{I}$, and $\hat{\tau}$ its discrete tangent map, we define the *discrete complexity map* $\hat{C}$ to be

$$\hat{C} = \bigcup_{\hat{x} \in \mathrm{Dom}(\hat{\tau})} (\hat{x}, \hat{C}_N(\hat{x}, \delta), \hat{C}_T(\hat{x}, \delta)).$$

where $\hat{x}$ denotes a pixel in the image domain.

Remark 4.4. In our computations, dilations are defined as a set of "pixels". The distinction between area and perimeter did not apply, and we kept the same measure in both cases: counting the number of pixels of the dilated set. Experiments in Appendix B show that this approach leads to good approximations. The calculation of the perimeter would be possible but more computationally expensive.

Corollary 4.1 (Bounds for discrete complexity indexes). *Given a discrete tangent map $\hat{\tau}$, the discrete normal complexity index takes values between 1.0 and 2.0, while the discrete tangential complexity index takes values between 0.0 and 1.0.*

PROOF. Both complexity maps being generated by local linear approximations of the set by a finite number of line segments, we can apply two previous theorems (Thm 4.1 and Thm 4.2) to obtain the result. □

6. The parameters involved

The algorithm to compute the complexity map of the discrete tangent map being clearly laid out, we can now list the various parameters that need to be determined. Starting from the original image, with resolution N, the discrete tangent map was obtained at a scale σ with N_θ different orientations, and the algorithm for the discrete complexity maps needed the following parameters (see also Table 4.2):

- the local extent $\Omega(x)$ for the calculation of the indexes: here it will be set to a square $\Omega \times \Omega$;
- the scale for the analysis $\delta > 0$;
- the range of scales over which to dilate $\{\delta_k\}$, $k = 1, 2, \ldots, K$ or stated differently, given δ, choose $\{\delta_k\}$, its neighborhood, at which the dilations are done.
- the subresolution for the projection of the tangents given by p_x and p_y;
- the extent of the projected tangents $\hat{\omega}_T$ (tangential extent: original length) and $\hat{\omega}_N$ (normal extent: original width);

Two indexes need to be calculated for each position where there is at least one tangent, i.e. the tangential and the normal indexes. The above mentioned set of parameters will be different for each of these two. This is why in Table 4.2 we indicate by a subscript (N or T), the process to which it belongs. As far as the extent for the

Effect	**Parameter**	**Symbol**
Structure/	local extent	Ω_N, Ω_T
geometry issues	scale of the analysis	δ_N, δ_T
	neighborhood of δ_T	$\delta_{T,1}, \cdots, \delta_{T,K_T}$
	neighborhood of δ_N	$\delta_{N,1}, \cdots, \delta_{N,K_N}$
Numerical/	extents	$\hat{\omega}_T$, $\hat{\omega}_N$
quantization issues	precision (subresolution)	(p_x, p_y)

TABLE 4.2. Parameters to be set for the calculation of the complexity indexes. The subscripts reflect the fact that the values could be different for the tangential and the normal analysis.

projected tangents are concerned, only one of the two is important in any case: the tangential extent, in the case of the normal dilations, and the normal extent in the case of tangential dilations. This is why the order in Table 4.2 is reversed for that entry. In the next chapter we discuss methods for automatically setting these parameters.

7. Summary

This chapter presented a technique that will be used to eventually partition the ensemble of (discrete) curve-like sets. It is based on a complexity analysis, done through normal and tangential dilations. The rate of growth of the area of the dilated sets provided the normal and tangential complexities which, when restricted to a bounded region in space, form the basis of our complexity measure. The different steps of the algorithm to obtain discrete estimates of both the normal and tangential complexity indexes were laid out, and the parameters to be set were clearly presented. We have shown that the discrete normal complexity index was a number between 1.0 and 2.0, and the tangential complexity index a number between 0.0 and 1.0. How should this new layer of organization be used to group the edge elements and prepare the ground for grouping? This is the question we shall address in the next chapter.

CHAPTER 5

From complexity to decision

Let us return to questions of computer vision, and ask which types of geometric objects are natural for early representations. For instance, looking back at Paolina's discrete tangent map, we notice the emergence of different types of curve-like substructures (see Fig. 5.1). Curves are clearly among them, as are textures. But is texture a distinct class; is there a difference between texture flows, as they arise from hair patterns that are well combed, and turbulent patterns such as wind-blown hair? What sort of object is the edge of the Kanizsa pinstripes? The complexity map provides an answer, by classifying discrete curve-like sets into:

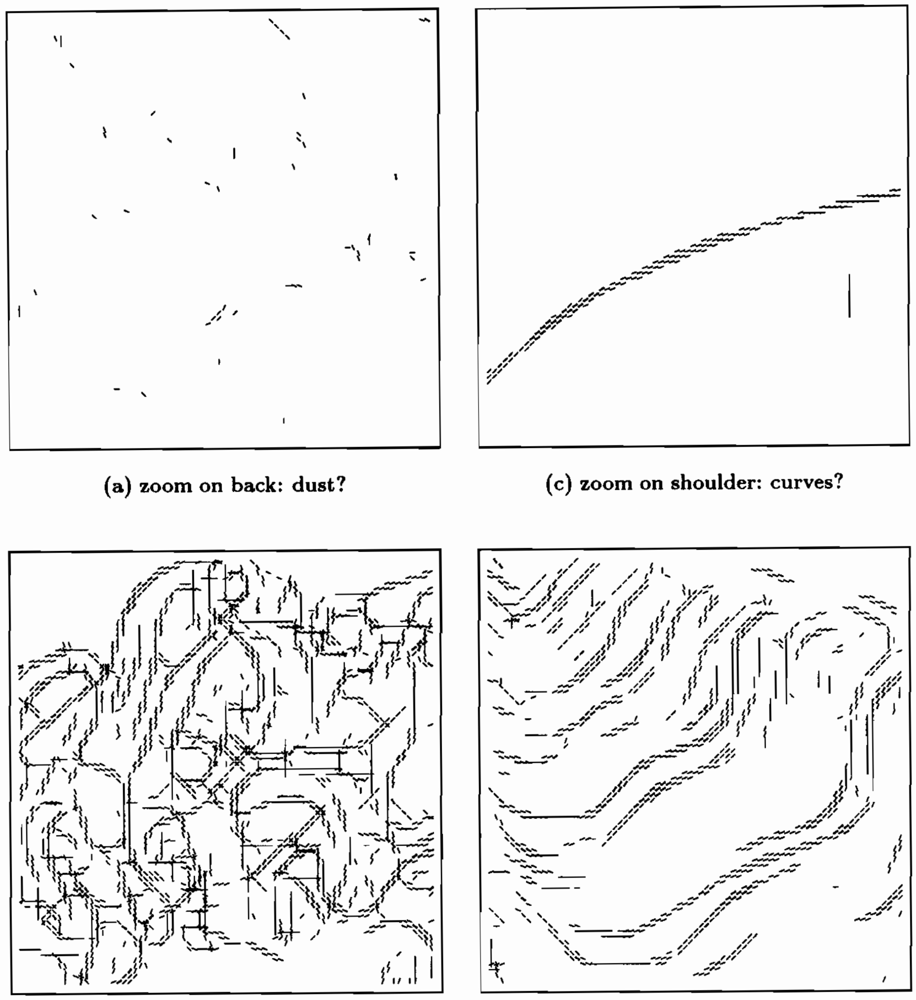
dust-like: sets in which the tangent map is sparse, the object almost nowhere extends along its length locally;

curve-like: discrete tangent maps for which a curve representation is completely adequate. Objects extend along their length, like Paolina's shoulder, and the density of other tangents is low almost everywhere along it in a local neighborhood;

turbulence-like: tangent maps that are characterized by objects that do not extend along their length but are dense in the normal direction; e.g., Paolina's uncombed hair or the edge of a grating;

texture-flow-like: tangent maps for which the objects extend along their length and are also dense in the normal direction; e.g. Paolina's combed hair or the grating part of the Kanizsa pattern.

In this chapter we show how to use the complexity map to partition the tangent map into these classes. In the process, we show how to set the various parameters that



(a) zoom on back: dust?

(c) zoom on shoulder: curves?

(b) zoom on hair: turbulent?

(d) zoom on hair: flow?

FIGURE 5.1. Curve-like sets substructures. Zooming in different parts of Paolina's discrete tangent map highlights different types of curve-like substructures: dust (spurious responses in the back), curves (the edge of the shoulder), flow (well-combed hair), and turbulence (wind-blown hair).

emerged from the last chapter. From a purely mathematical perspective, these parameters are arbitrary. However, from the perspective of computational vision, there are two principles that dictate basic relationships between them. We describe these principles in Section 2, and then derive constraints on the parameters by studying the behavior of our measure on simple sets. The key idea is that image segmentation will be lifted onto the tangent map and effected through notions of complexity.

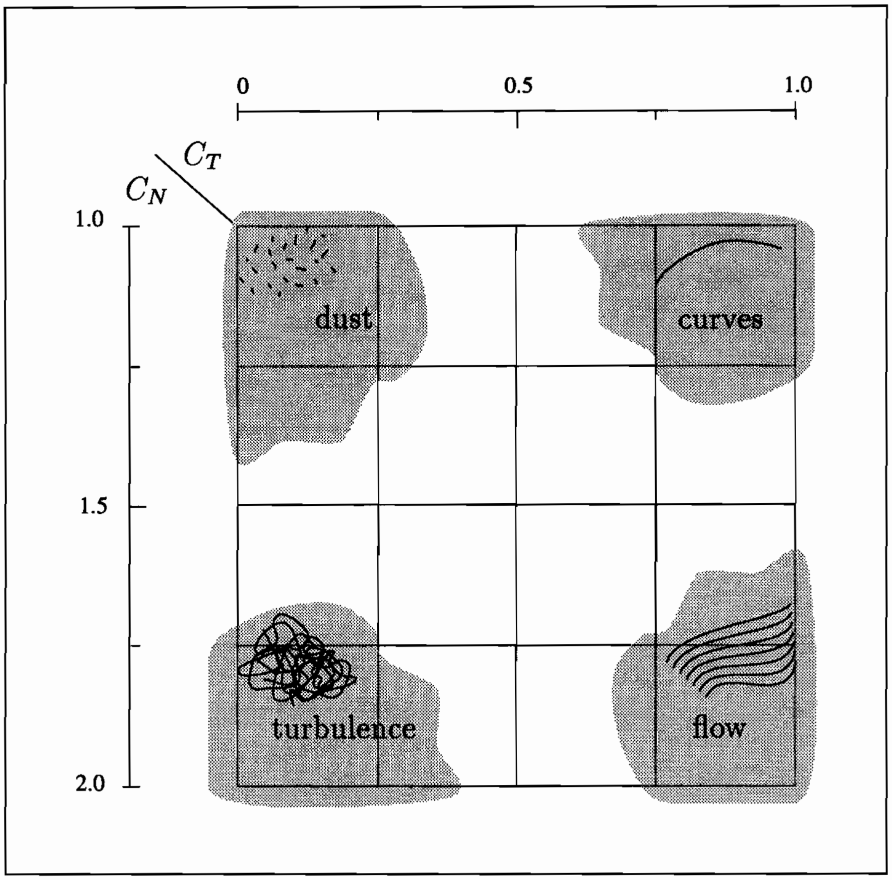


FIGURE 5.2. THE COMPLEXITY SPACE. The tiled square represents the space of complexity indexes pairs that can be encountered. The normal complexity varies between 1.0 and 2.0 while the tangential complexity is between 0.0 and 1.0. Each of the corners of the complexity space corresponds to a different type of curve-like substructure. Partitioning the complexity space will result in an image segmentation scheme bound to the structure of the objects in the visual scene.

1. Indexing representations through the complexity map

The four patterns listed in the introduction arise as the extrema in normal and tangential complexity space. In general, the space of valid tangential/normal complexity pairs is a subset of $\mathbf{R}^2$, namely $[0.0, 1.0]$x$[1.0, 2.0]$. This is because the tangential complexity is a number between 0.0 and 1.0, and the normal complexity a number between 1.0 and 2.0. Partitioning this space allows us to organize the space of possible patterns into equivalence classes (see Fig. 5.2). In this thesis the partition will consist in setting two values $1.0 < \Delta_N < 2.0$ and $0.0 < \Delta_T < 1.0$ and then looking at the four regions that they induce. This provides us the following nomenclature:

dust: low normal complexity: $1.0 \leq C_N(x,\delta) < \Delta_N$
low tangential complexity: $0.0 \leq C_T(x,\delta) < \Delta_T$

curves: low normal complexity: $1.0 \leq C_N(x,\delta) < \Delta_N$
high tangential complexity: $\Delta_T \leq C_T(x,\delta) < 1.0$

turbulence: high normal complexity: $\Delta_N \leq C_N(x,\delta) < 2.0$
low tangential complexity: $0.0 \leq C_T(x,\delta) < \Delta_T$

flow: high normal complexity: $\Delta_N \leq C_N(x,\delta) < 2.0$
high tangential complexity: $\Delta_T \leq C_T(x,\delta) < 1.0$

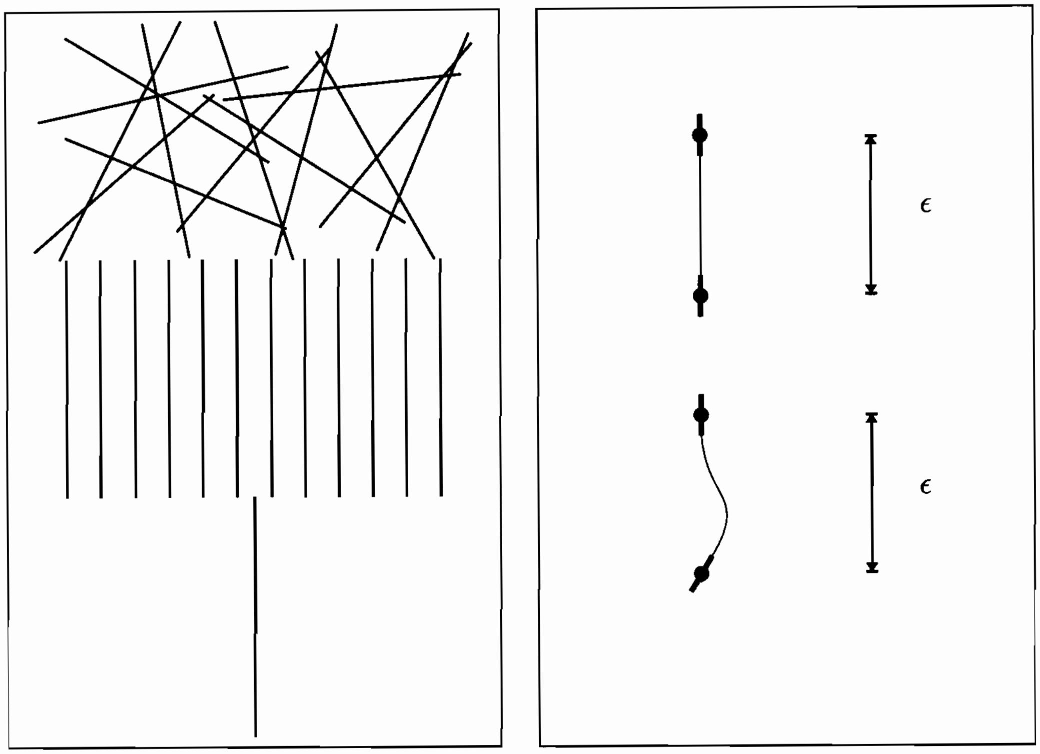
We will show shortly that under an appropriate choice of parameters this partitioning scheme can successfully segment an image in terms of these various kinds of curve-like substructures.

2. The anchor problem: verifying the curve assumption

Our approach will be to focus on the top right quadrant in the complexity space (see Fig. 5.2), i.e. **curves**[1] and verifying the guiding *curve assumption* from Chapter 3, namely that *(i)* a curve must extend along its length; *(ii)* it should not be space-filling; and *(iii)* the number of discontinuities must be negligible.

We claim that a set satisfying these conditions should belong to the upper right corner of the complexity space. Detecting regions in an image for which this is true constitutes our anchor problem, i.e. *verifying the curve assumption.* Recalling our philosophical proposal from Chapter 3, we claim that curves are the objects for which the notion of length is meaningful and of practical value. Now, if there are curves in the image, what is their maximal local extent? How can we make the local to global transition? Curves in the shoulder region of Paolina have a large local extent while those in the hair region have a small one. The sole fact that the discrete tangent map was obtained from edge-detection operators puts a lower bound on the size of the local extent to verify the curve assumption. But it could be verified on larger extents. For normal complexity, we will seek the largest extent Ω for which the curve

[1] From now on in this chapter, when refering to "curves", we imply objects from the upper right corner of the complexity space.



(a) normal complexity tree (b) neighboring tangents

FIGURE 5.3. Relative complexity and wiggles. (a) normal complexity tree: a line is simpler than a grating which is simpler than a jumble of random lines (b) compatibility between pairs of tangents: a line is simpler than a wiggle.

assumption is verified over at least one point in the discrete tangent map, and use this local extent to compute the complexity. This provides the following notion of *relative complexity*: a line is simpler than a grating, which itself is simpler than a bunch of oriented random lines, as illustrated in Fig. 5.3a; we call this the *normal complexity tree*. The lower in the tree an object is, the simpler the object is, but always relative to the rest. On the right are shown two samples of pairs of local orientations with their centers both displaced ϵ apart. This illustrates tangential complexity, for which a line is simpler than a wiggle. The first pair is aligned, while for the second pair, the orientation of the bottom tangent is random. How can we pass from one to another with the shortest path? If the length of the link correlates with the complexity of the interpolated piece of curve, a line is simpler than a wiggle. Both the tangential and normal complexity can capture the last statements and we will set the parameters

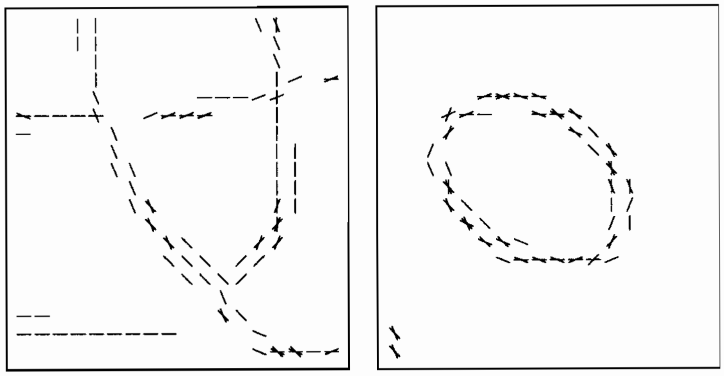


FIGURE 5.4. Transversality and quantization revisited.

for this to be so. But this process needs also to take into account the constraints imposed by the discretization, namely that the detection of tangents tends to spread both spatially and in orientation. Since the operators were originally seeking pieces of curves at a given scale σ, we can therefore now put forward two principles to guide the local to global transition. These two principles will constrain the guiding curve assumption from Chapter 3. To verify it, one needs to select from

(i) the **quantization principle:** *the scale* δ as small as possible, but large enough to overcome the effects of digitization and to detect curve intersections;

(ii) the **simplicity principle:** *the local extent* Ω as small as possible, but large enough to segregate curves from textures.

The first point is illustrated in Fig. 5.4. As can be seen, the tangents do not align perfectly and multiple tangents occur in three cases: (i) when the orientation cannot be represented due to the discretization, (ii) when the curvature is high, (iii) at curve intersections. The complexity analysis should be able to discriminate between these cases, since this is necessary for effecting the local-to-global transition.

The second point was illustrated by the Kanizsa and Paolina examples. As mentioned in Chapter 2, events occur at different scales. Even at a fixed scale for the edge detection process, different types of structures may emerge. Since we are doing curve detection, if there were curves in this image, where would they be, and what should be the scale and extent for grouping? One can realize that given a local extent Ω and a discrete tangent map, not all scales are interesting. Scales smaller than the pixel size provide little information, while scales on the order of the chosen local extent yield saturated growth. The interesting range therefore lies in between these two extremes.

The parameters that need to be set to obtain the complexity map were given in Chapter 4. Some are bound to geometric/structural properties of the object and the scene, while others are used to obtain reliable numerical estimates and to overcome the effects of digitization. The discretization scheme and the parameters of the operator are known *a priori* for our analysis. We chose a square lattice and used L/L operators for detection of lines and edges. Using this knowledge will help selecting the parameters of the complexity analysis in such a way as to overcome the effects of digitization.

Our decision strategy parallels the notions of *integration scale* and *local scale* described by Lindeberg (1993). The integration scale in our case will be the *local extent* Ω, while the local scale is the *scale* δ at which the normal and tangential complexity are estimated. It will be large enough to overcome the effect of digitization, but then its value for the normal complexity will be directly correlated with the frequency allowed within the spatial extent. A large scale allows only for very few distinct parallel lines for instance. We will therefore be biased toward Koenderink's assessment that counting beyond 3 is a rare capability in vision and has to be subsumed under the heading of combinatorics [Koenderink 1990, p.64]. However, we stress the difference that Lindeberg's scales were applied to image operations, and ours are applied to the tangent map.

3. Some simple curve-like sets

Normal complexity for a line segment on a discrete grid is always greater than or equal to one. Tangential complexity of a line is a number between 0 and 1 depending on the length of the line, the scale of analysis, and the local extent. According to the simplicity and tangent quantization principles, parameters must be set so this is clear. To obtain the proper values, we therefore begin with a detailed analysis of a segment and circle discrete tangent map and then proceed with patterns containing multiple curves.

3.1. A single line segment and a circle. Given spatial and orientation resolutions, how can we choose scale and local extent for the complexity? Length, scale, spatial extent and complexity are coupled notions. The first cases analyzed here will focus on tangential complexity, having in mind the detection of line endings and points of high curvature while palliating the effects of digitization. The choices will assume for tangential complexity a fixed spatial extent for all images using a given operator. The choices will therefore be a function of the spatial and orientation resolutions and of the operator. Various constraints will be set up. Most constraints will be bound to the operator and expressed in image coordinates. The derivation of the curvature and orientation constraints is done by studying the discrete tangent maps (obtained through L/L) of classes of test patterns. All test images (see Fig. 5.5) had a resolution of $N = 51$, which was used to calculate the values with respect to the image coordinates; i.e. in pixel units.

3.1.1. *Continuity.* The tangential complexity will provide a means for detecting line endings. Suppose $\Omega(x)$ is centered on a line ending that extends on the other side beyond the border of the bounding region (see Fig. B.1g in Appendix B). For small enough δ, i.e. $\delta < \Omega/2$, we have (using Lemma 4.4)

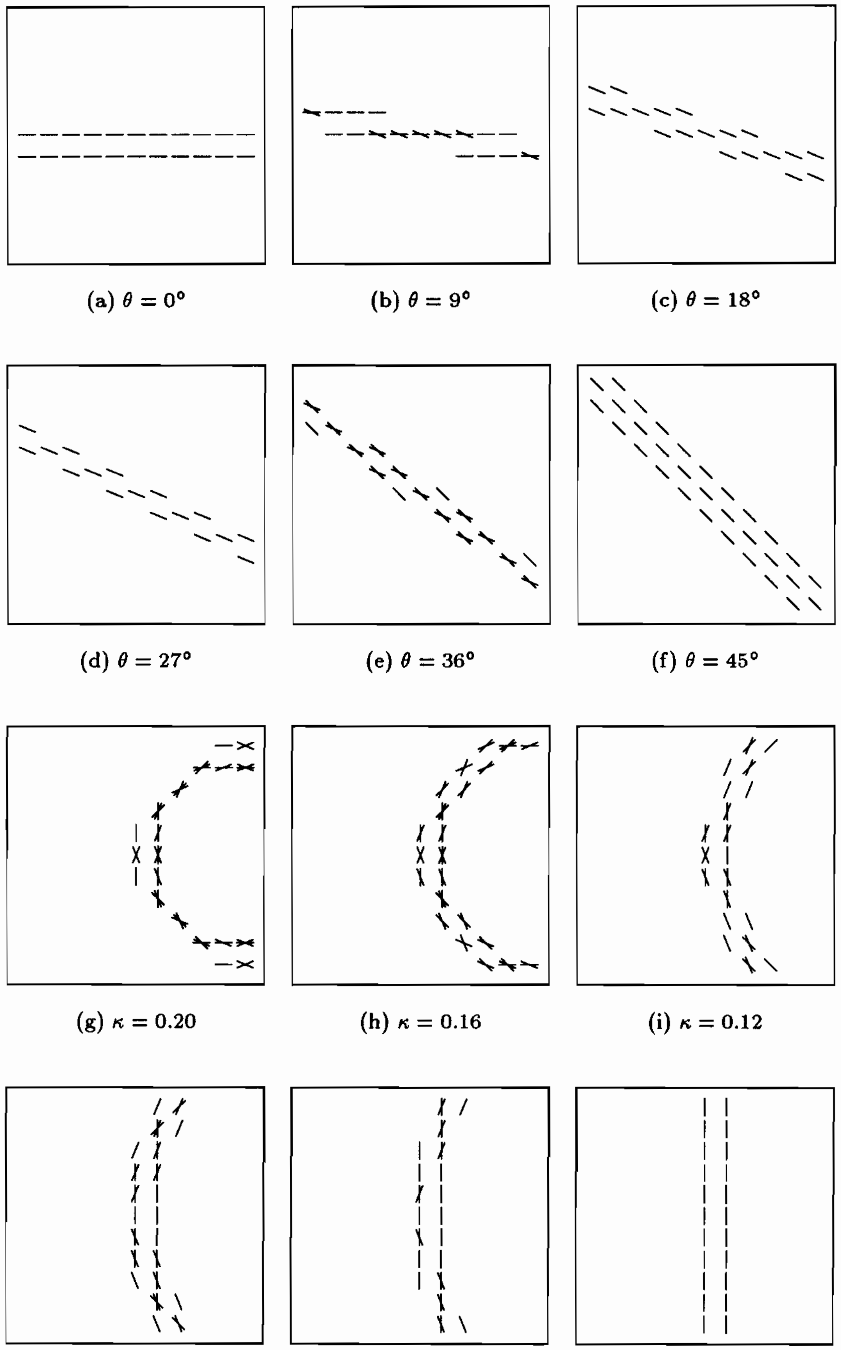
$$C_T(x, \delta) = \frac{l}{l + \delta},$$

where l is the length of the line within the spatial extent. This can be used to build the following constraint on length for detecting line endings:

Constraint # 1 (Continuity). *Given the local extent Ω and $0 < \Delta_T < 1$, a line segment of length $\sqrt{2}\Omega/2$ with line ending at x, will have $C_T(x, \delta) < \Delta_T$ if*

$$\frac{\sqrt{2}\Omega/2(1 - \Delta_T)}{\Delta_T} < \delta < \Omega/2$$

3.1.2. *Curvature.* It was shown in the previous chapter that tangential complexity can be used for constraining curvature (see Eq. 4.17). The tangential complexity is therefore sensitive to the distribution of orientation differences along the curve. Curvature of a curve as the rate of change of the angle of the tangent vector field was originally due to Euler. We will use here the tangential complexity to control the



(a) $\theta = 0°$ (b) $\theta = 9°$ (c) $\theta = 18°$

(d) $\theta = 27°$ (e) $\theta = 36°$ (f) $\theta = 45°$

(g) $\kappa = 0.20$ (h) $\kappa = 0.16$ (i) $\kappa = 0.12$

(j) $\kappa = 0.10$ (k) $\kappa = 0.06$ (l) $\kappa = 0.02$

FIGURE 5.5. Discrete tangent maps for segments and arcs obtained from L/L operators. (a)-(f) for a single line segment with varying orientation θ; (g)-(l) arcs of circles with varying radii r, in pixels ($\kappa = 1/r$).

bending of a curve by restricting the range of valid "curves" to some given so-called "curvature classes". These were first described by Parent & Zucker (1989), and used subsequently in Iverson (1993) to partition the space of curvatures encountered in the local analysis of curves (see Fig. 5.5g-l for discrete tangent maps of arcs with various curvatures). We will therefore set up the parameters such that curves with low curvature will be treated as such. In his thesis, Iverson (1993) chose 5 curvature classes which were bounded by digital radii of 5 pixels, since this "is close to the minimum radius of a circle which can be reliably distinguished from a blob and simultaneously categorized into either a line or edge-like discontinuity" [Iverson 1993, p.122]. The decision we took is to keep all curves having a curvature radius bigger than 10 pixels, i.e. from straight to slightly curving lines:

Constraint # 2 (Curvature). *Given a local extent $\hat{\Omega} = 8$ pixels, and $\Delta_T = 0.73$, objects with curvature radius higher than 10 pixels will have tangential complexity $C_T(x, \hat{\Omega}) \geq \Delta_T$ whenever*

$$2^{-5.3}N < \delta < 2^{-4.5}N$$

A plot of the tangential complexity as a function of scale (Fig.5.6), provides a justification for the choice of scale for tangential dilations. The limiting circle radius was chosen to be 10 pixels ($\kappa = 0.10$) and Δ_T was set to 0.73. Searching through spatial and normal extents, we found that $\hat{\Omega} = 8$ pixels for the spatial extent, and $\hat{\omega}_N = 1.1$ pixels for the normal extent provided a good split for curvature. With these chosen, we get that in order to satisfy the condition, the scale needs to be between 1.29 and 2.25 pixels as can be derived from Fig. 5.6 and from the fact that the resolution was $N = 51$ for this experiment.

3.1.3. *Discretization.* A straight line of a given length intersects a different number of pixels of a digital grid depending on its orientation. The scale δ and the normal extent $\hat{\omega}_N$ will therefore be set to overcome the quantization and digitization (see Fig. 5.5a-f), and ensure that a line is seen with tangential complexity greater than some $\Delta_T < 1$, no matter how it is oriented with respect to the digitizing grid.

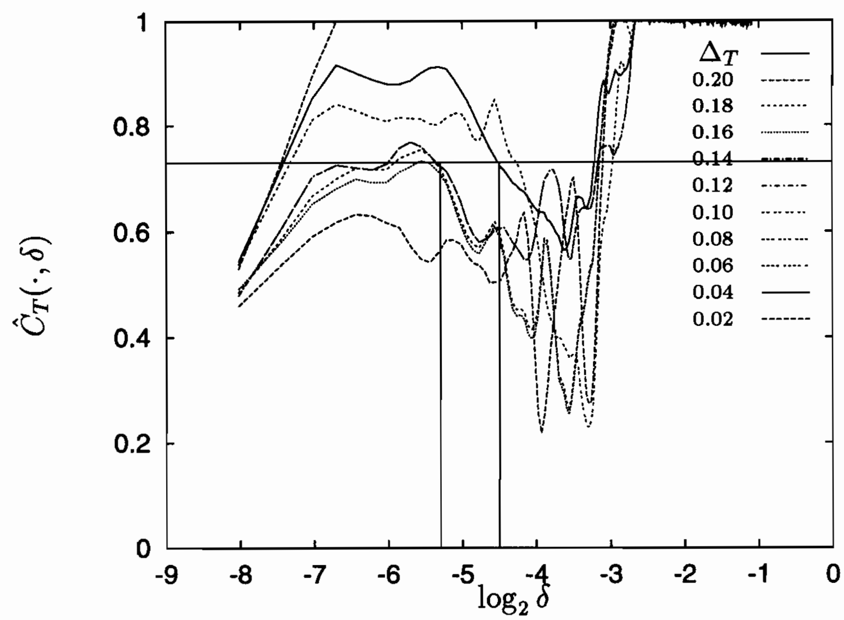


FIGURE 5.6. Finding the scale for tangential complexity to select curves with low curvature and reject those with high curvature. This plot shows the estimation of the tangential complexity across all scales of arcs of circles. The local extent was 8 pixels in all cases and the normal extent was 1.1 pixels. The legend gives the curvature κ for each plot. The goal is to be able to segment the two bundles: high curvature from low curvature.

A constraint was built by considering the set of lines passing through a square lattice at various orientations and infering the discrete tangent map with L/L operators. The goal was to make sure that, over the chosen local extent ($\hat{\Omega}$) and with the normal extent ($\hat{\omega}_N$), a line would be seen with tangential complexity larger than Δ_T. To find the appropriate scale, we set $\Delta_T = 0.73$ and used the same extents as those found for curvature, i.e. $\hat{\omega}_N = 1.1$ and $\hat{\Omega} = 8$. Doing the tangential analysis at all scales using these parameters, we derived that in order to have a tangential complexity greater than $\Delta_T = 0.73$ for lines at any orientation, we must take the scale δ to be within some bounds:

Constraint # 3 (Orientation). *Given the image resolution $N = 51$, for the complexity of a line to be larger than $\Delta_T = 0.73$, we need to have*

$$2^{-5.2}N < \delta < 2^{-4.7}N$$

Looking at the complexity as a function of scale provides the possible values for the scale as shown in Fig. 5.7. Looking at the graph one sees that we need to choose

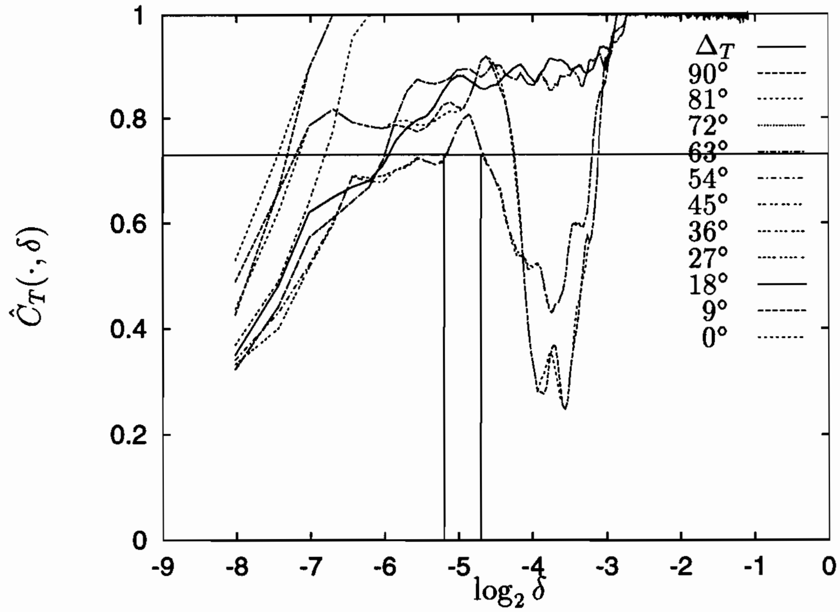


FIGURE 5.7. Finding the scale for tangential complexity such that a line is seen with high tangential complexity no matter what its orientation is. This plot shows the estimation of the tangential complexity across all scales of lines with different orientations. The local extent was 8 pixels in all cases and the normal extent was 1.1 pixels and the resolution $N = 51$.

the scale δ to be between 1.39 and 1.96 pixels to ensure the tangential complexity of a line to be larger than $\Delta_T = 0.73$ no matter what its orientation is.

3.2. Multiple lines. So far the objects that have been studied were elementary, but most importantly, there was only a single one in the image. What happens when the scene is composed of more than one line or curve? The normal complexity will eventually grow larger than 1. When is it too large, i.e. when does linear structure start to become confusing? The study of simple gratings and crossings will provide constraints that will be used for the scale selection for the normal complexity. This time, as opposed to the constraints obtained for the tangential complexity, the values will mostly be in relation to the structure of the image and expressed with respect to the unit square. The values will therefore vary from one image to the next.

3.2.1. *A set of parallel lines.* Let us first consider a set of lines with the same orientation and length, but spaced arbitrarily. The spacing between the lines will be strictly positive and finite. A general algorithm to find the normal complexity of this kind of pattern will therefore emerge from:

Theorem 5.1. *Let E be a set of n aligned and parallel vertical segments $s_1, s_2, \cdots, s_n$ of length $l < 1$, bounded by the unit square. The segments are separated by the distances $d_0, d_1, \cdots, d_n$, where d_0 is the distance between the left edge and s_1, $d_i = d(s_i, s_{i+1})$, and d_n is the distance between the last segment s_n and the right edge. Let now $b_0 = d_0$, $b_{2i-1} = b_{2i} = d_i/2$ for $i = 1, 2, \cdots, n-1$, and $b_{2n-1} = d_n$. If the b_i are reordered such that*

$$b_{\sigma(0)} \leq b_{\sigma(1)} \leq \cdots \leq b_{\sigma(2n-1)},$$

then

$$(5.1) \quad C_N(\cdot, \delta) = \begin{cases} 1 & \text{if } \delta < b_{\sigma(0)} \\ \dfrac{(2n-i-1)\delta + 2\displaystyle\sum_{k=0}^{i} b_{\sigma(k)}}{(2n-i-1)\delta + \displaystyle\sum_{k=0}^{i} b_{\sigma(k)}} & \text{if } b_{\sigma(i)} \leq \delta < b_{\sigma(i+1)}, \quad i < 2n-1 \\ 2 & \text{if } \delta \geq b_{\sigma(2n-1)} \end{cases}$$

Remark 5.1. A series of results will be presented in this and the following sections. The main idea is that the analysis will be done within a bounded region Ω (most of the time, the unit square). By writing $C_N(\cdot, \delta)$ or $C_T(\cdot, \delta)$, we assume that E and all the operations (dilations, area, perimeters) are limited by the region Ω.

PROOF. First we show that

$$(5.2) \quad |E_N(\epsilon)|_2 = \begin{cases} 2nl\epsilon & \text{if } \epsilon < b_{\sigma(0)} \\ l\sum_{k=0}^{i} b_{\sigma(k)} + (2n-i-1)\epsilon l & \text{if } b_{\sigma(i)} \leq \epsilon < b_{\sigma(i+1)}, \quad i < 2n-1 \\ l & \text{if } \epsilon \geq b_{\sigma(2n-1)} \end{cases}$$

The case when $\epsilon < b_{\sigma(0)}$ is trivial, while for $\epsilon \geq b_{\sigma(2n-1)}$, it follows from the fact that $\sum_{k=0}^{2n-1} b_{\sigma(k)}$ by construction. Suppose now that $b_{\sigma(i)} \leq \epsilon < b_{\sigma(i+1)}$. The growth occurs on both sides of the lines s_i. The lines where $j < i$ is saturated while it is not for the rest leading to

$$|E_N(\epsilon)|_2 = l\sum_{k=0}^{i} b_{\sigma(k)} + (2n-i-1)\epsilon l$$

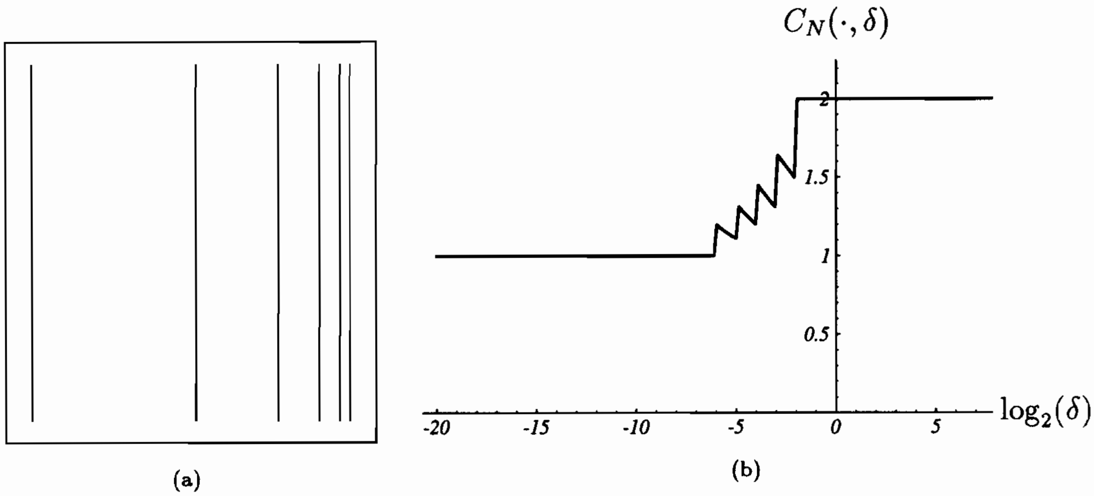


FIGURE 5.8. Plotting the normal complexity of a set of parallel lines over a bounded region (for sake of understanding the bounding box is a little bit wider to allow to see the complete pattern): (a) six parallel lines spaced according to a geometric sequence with (b) its normal complexity. Notice the discontinuities in the complexity for each of the grating separation.

Finally, applying Lemma 4.1 gives the desired result. □

An example of such a pattern is shown in Fig. 5.8a and the corresponding graph of normal complexity is shown in Fig. 5.8b. The values for the complexity can be computed from the Mathematica routine `GratingNormalLocalC` provided in Appendix B. Notice in Fig. 5.8b the peaks occurring at the half spacing distance of each of the lines. This suggests a tentative answer to the local extent selection problem for the normal complexity: calculate the complexity on the whole image at all scales in order to obtain a pattern like the one shown in Fig. 5.8b, then choose the peak in the coarsest scale to set the local extent.

Example 5.1 (Regular grating revisited). Let us now recall the example of a pattern with n equidistant lines of unit length as introduced in Chapter 4. Suppose the spacing between the lines is $0 < \delta_1 < 1$, that x is on a point underlying a non-empty tangent set closest to the center of the pattern, and the pattern is at a distance $\delta_1 < \delta_2 < 1$ to the edges of $\Omega(x)$. In this case we have

$$b_{\sigma(0)} = b_{\sigma(1)} = \cdots = b_{\sigma(2n-3)} = \delta_1/2, \quad b_{\sigma(2n-2)} = b_{\sigma(2n-1)} = \delta_2$$

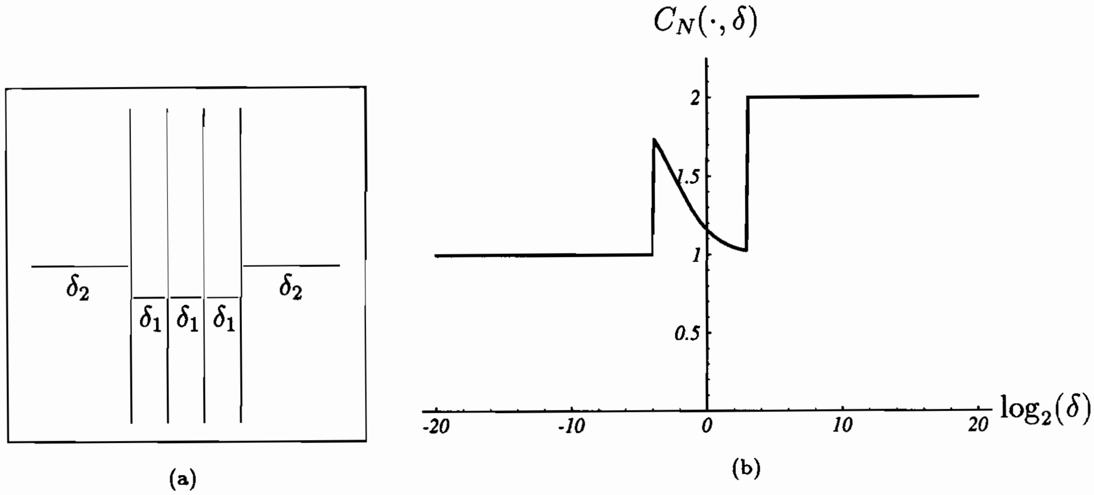


FIGURE 5.9. Plotting the normal complexity of a set of equidistant parallel lines over a bounded region. (a) the pattern to be studied: n equidistant lines with an inter-line spacing of $\delta_1 = 2^{-3}$ and with $\delta_2 = 2^3$ (b) its normal complexity: the first discontinuity occurs at $\delta = \delta_1/2$ with $C_N(\cdot, \delta_1/2) = 2 - \frac{1}{n}$.

and applying the last theorem, we obtain the normal complexity index

$$C_N(\cdot, \delta) = \begin{cases} 1 & \text{if } \delta < \delta_1/2 \\ \dfrac{2\delta + 2(n-1)\delta_1}{2\delta + (n-1)\delta_1} & \text{if } \delta_1/2 \leq \delta < \delta_2 \\ 2 & \text{if } \delta \geq \delta_2 \end{cases} \tag{5.3}$$

which corresponds to what was derived before except for the discontinuity at $\delta = \delta_2$, as can be seen by comparing Figs. 4.8 and 5.9. We now put these computations together to specify a constraint on the quantization principle.

Constraint # 4 (Lateral spreading). *If a quantization artifact leads to as many as n parallel tangents for a single line, then the scale δ to ensure a dimension smaller than Δ_N needs to be*

$$\frac{(n-1)}{2}\frac{(2-\Delta_N)}{(\Delta_N - 1)}\delta_1 \leq \delta < \frac{\Omega}{2} \tag{5.4}$$

where δ_1 is the spacing between the centers of adjacent pixels.

3.2.2. *A pair of lines.* We just saw what the normal complexity would be for parallel lines, the case of a pair of lines just being a special case. The last theorem and a result from Chapter 3 can be used to show that, provided the lines are sufficiently far from the edges of the local extent, the normal complexity attains a maximum at $\delta/2$, if δ is the spacing between the line segments, with a value of $C_N(\cdot, \delta/2) = 1.5$. Now, what happens if the lines are crossing each other at an angle θ?

Lemma 5.1. *Given two lines of length l centered in a disk $\Omega(x)$ of radius r and crossing with angle θ (see Fig. 5.10a), then the area of the local normal dilation of the set is given by*

(5.5)

$$|E_N(x,\epsilon)|_2 = \begin{cases} 4\left(r^2 \arcsin\left(\frac{\epsilon}{r}\right) + \epsilon\sqrt{r^2-\epsilon^2} - \frac{\epsilon^2}{\sin\theta}\right) & \text{if } 0 < \epsilon \leq \epsilon_0 \\ \theta r^2 + 2\left(r^2 \arcsin\left(\frac{\epsilon}{r}\right) + \epsilon\sqrt{r^2-\epsilon^2} - \epsilon^2 \tan\left(\frac{\theta}{2}\right)\right) & \text{if } \epsilon_0 < \epsilon \leq \epsilon_1 \\ \pi r^2 & \text{if } \epsilon > \epsilon_1 \end{cases}$$

where x is the center of the crossing and with $\epsilon_0 = r\sin(\theta/2)$ and $\epsilon_1 = r\cos(\theta/2)$.

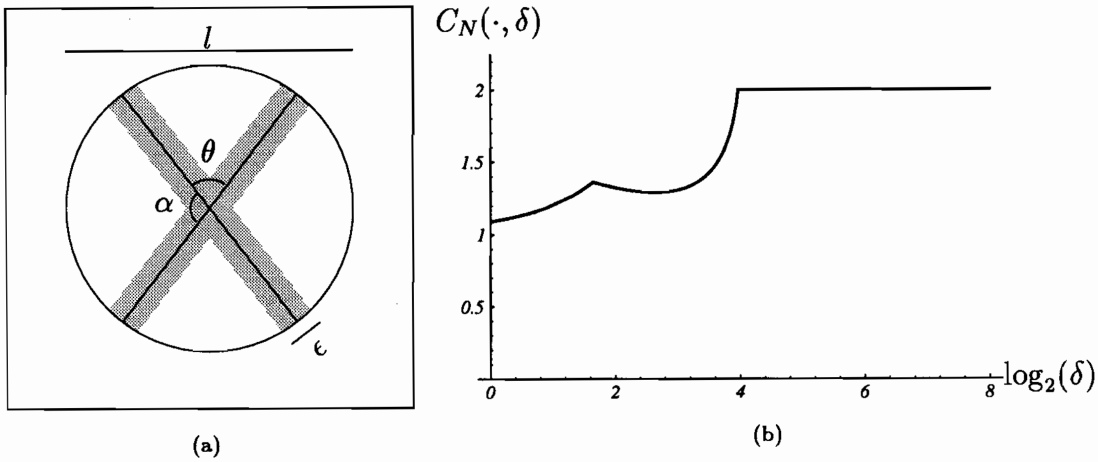
PROOF. The proof of this result is given in Appendix A. The idea is that $|E_N(x,\epsilon)|_2$ will grow smoothly with ϵ until the growth within the sectors saturates. This happens when ϵ reaches $\epsilon_0 = r\sin(\theta/2)$ and $\epsilon_1 = r\cos(\theta/2)$. □

Given the area of the dilated set, the normal complexity can be estimated. It was done by applying a straight line fit to the normal complexity log-log plot. In Fig. 5.10b, we plotted the estimated normal complexity as a function of scale. This shows that, given the local extent and the scale, you know how sensitive the normal index will be to line crossings and could use this information to detect these:

Constraint # 5 (Intersections). *Given Ω, two lines crossing with an angle $\pi/4 \leq \theta \leq \pi/2$ at x will have normal complexity $C_N(x,\delta)$ greater than $\Delta_N = 1.5$ if $\delta = \Omega_N/4$.*

The justification for this constraint is given by Fig. 5.10c, where we see that if the orientation is within some bounds, the normal complexity will be larger than 1.5.

3.2.3. *A pattern of radial lines.* In much the same way as we did for the case of parallel lines, we can look at the complexity for a set of n radial line segments of length l, which somehow can be considered as a generalization of what was shown in Section 3.2.1 for parallel lines. The result is given in Appendix A. It relies on the definition of a convenient bounding polygon. A plot for the normal complexity of



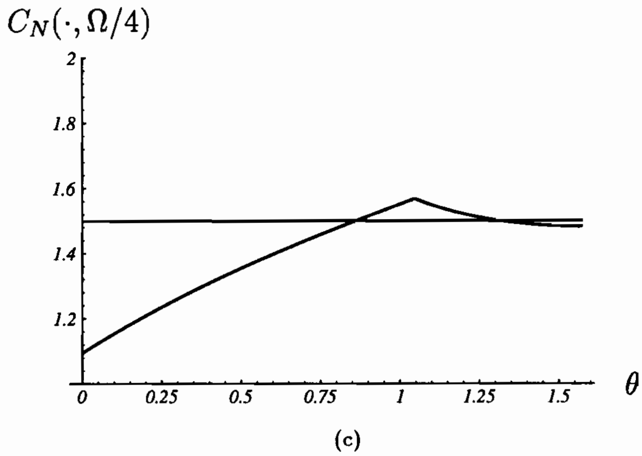


FIGURE 5.10. Plotting the normal complexity for a pair of crossing lines over a bounded region. (a) two lines of length l crossing at angle θ and the dilated set at scale ϵ shown in grey (b) its normal complexity as a function of scale. Notice the bump in the complexity. (c) the normal complexity for $\delta = \Omega/4$ as a function of the separation angle (in radians). We see that then if the angle is larger than $\pi/3$, the normal complexity is larger than $\Delta_N = 1.5$.

a radial pattern is shown on Fig. 5.11b. It corresponds to a pattern of 5 lines with separation at $\theta_i = \pi/2^i, i = 1, \cdots, 4$, and $\theta_5 = \pi - \sum_{i=1}^4 \theta_i$ (see Fig. 5.11a). Notice the same kind of peak pattern can be observed as in the case of parallel lines (Fig. 5.8b). This time the discontinuities in the complexity arise at the scales corresponding to the half-angles of the sectors.

As just mentioned, for technical reasons, the shape of the bounding polygon $\Omega(x)$ within which the analysis is carried out, has an effect on the normal complexity index of the set at x. Our algorithm is using a square window of size $\Omega \times \Omega$ (or $\hat{\Omega} \times \hat{\Omega}$ in

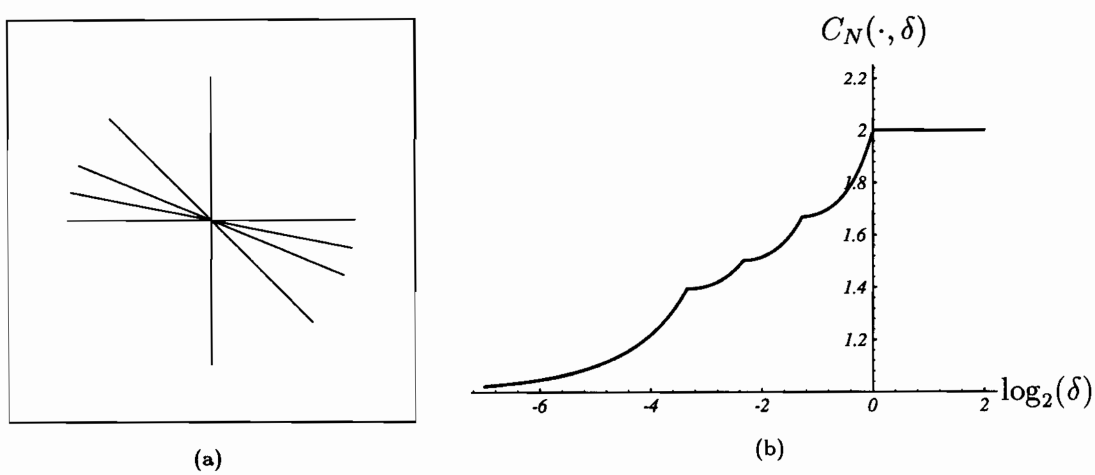


FIGURE 5.11. Plotting the normal complexity for a set of radial lines over a bounded region. (a) five lines of length l crossing at angles following a geometric sequence. (b) the normal complexity as a function of scale over the bounding polygon. Notice the bumps in the complexity for each of the angles.

image coordinates). Specifically, the complexity of a line will be affected by the shape of this region. Suppose a line is centered within the bounding square. The area of the dilated set is unaltered by the enclosing region when the line is horizontal or vertical (if the scale remains within Ω). It starts getting clipped when the line is at an angle, and the worst case occurs when the line is at an orientation $\theta/4$:

Theorem 5.2. *Let E be the diagonal to the square $\Omega(x)$ with side Ω, centered on x. Then*

$$C_N(\cdot, \delta) = \begin{cases} \dfrac{\Omega\sqrt{2}}{\Omega\sqrt{2} - \delta} & \text{if } \delta < \Omega\sqrt{2}/2 \\ 2 & \text{otherwise} \end{cases}$$

PROOF. If $\delta \geq \sqrt{2}\Omega/2$, the growth is saturated, otherwise the area of the dilated set is

$$\begin{aligned} |E_N(\epsilon)|_2 &= 2\left[2\epsilon\frac{\Omega\sqrt{2}}{2} - 2\frac{\epsilon^2}{2}\right] \\ &= 2\left[\Omega\sqrt{2}\epsilon - \epsilon^2\right] \end{aligned}$$

and Lemma 4.1 provides the final result. □

Constraint # 6 (Shape of bounding region). *Given a square bounding region $\Omega(x)$ centered on x with side Ω, its shape has an impact on the normal complexity index. From the last theorem, to ensure that the resulting normal complexity index $C_N(x,\delta)$ of a line segment centered at x be smaller than $1 < \Delta_N < 2$, we must choose*

$$0 < \delta < \frac{\Omega\sqrt{2}(\Delta_N - 1)}{\Delta_N}$$

3.2.4. *The Kanizsa pattern.* The ground is now set to calculate the normal complexity for the Kanizsa pattern. This will provide a piece of justification for our choice of local extent in the next section. The general case is developed in Appendix 1.3. To simplify matters we will suppose that $\Omega(x)$ is the bounding box for the pattern, where x is a point inside the grating somewhere around the middle. The Kanizsa pattern can be described by a set of variables: n, k, l, $\delta_2, \delta_3, \delta_5, \delta_6$ where

- n: the number of lines for the grating
- $k < (n-2)$: the number of lines inside the grating
- δ_i different key spacings:
  - δ_2: between the lines in the grating
  - δ_3: between the rectangle and the grating
  - δ_5: between the rectangle and the side of the unit square
  - δ_6: between the two side lines of the rectangle
- l the length of the lines in the grating.

Based on this definition of the pattern and on the assumption that $\delta_6/2 > \delta_3$, we get the following for the computation of the area of the normal dilation of the set:

$$\begin{aligned} |E_N(\epsilon)|_2 = \quad & (\epsilon < \delta_2/2)\ [2(n-1)l\epsilon, (n-1)l\delta_2] + \\ & (\epsilon < \delta_5)\ [4\delta_3\epsilon, 4\delta_3\delta_5] + \\ & (\epsilon < \delta_6/2)\ [4\delta_3\epsilon, 4\delta_3\delta_6/2] + \\ & (\epsilon < \delta_3, \epsilon < \delta_6/2)\ [(2\delta_6 + 4\delta_3)\epsilon - 4\epsilon^2, 2\delta_6\delta_3] \end{aligned}$$

where here, the quantities in the parentheses denote a test, and, depending on the result of this test, we take the first (if true) or the second (if false) action between the square brackets. The area can be rewritten as $|E_N(\epsilon)|_2 = a\epsilon^0 + b\epsilon + c\epsilon^2$, and from Lemma 4.1 we can explicitly compute the normal complexity.

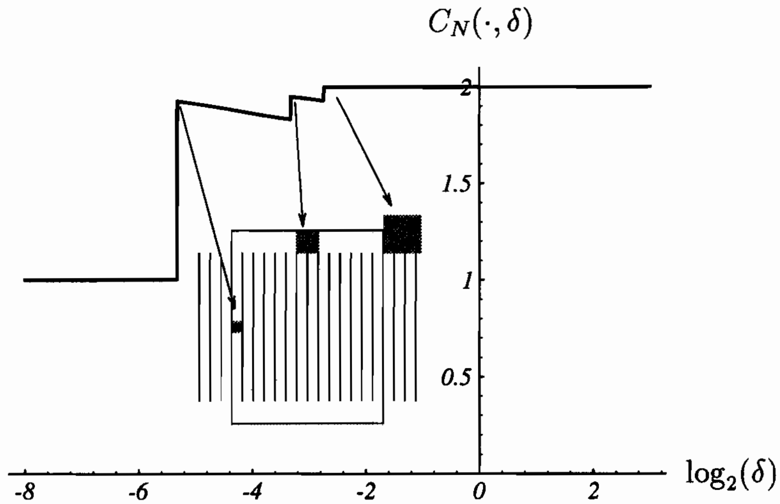


FIGURE 5.12. Emergence of natural scales. On top is the graph of the normal complexity for the Kanizsa pattern as a function of scale. The discontinuities occur at scales tied to key image structures.

A plot of the normal complexity as a function of scale is shown in Fig. 5.12. The Kanizsa pattern studied has been included as well to relate to the graph. Each of the grey squares corresponds to a discontinuity in the normal complexity graph. Those at which complexity changes rapidly are therefore of special significance, and define the *natural scales* for this pattern.

For any pattern, it would be possible to look at the values obtained for the normal complexity as a function of scale over the whole field of view. An application of this to texture analysis led to what was called the *local fractal dimension* [Peleg et al. 1984] and was used in other work [Pimienta et al. 1994] as a means to select the scale. In the case of a pattern with events occuring at a finite number of scales, you can obtain interesting results as was shown in Figs 5.8 and Section 3.2.3. The Kanizsa pattern and the Ullman images are other examples of such cases. In the case of natural images however, jumps are often blurred together and no spectral pattern emerges. This indicates that events do not cluster but are distributed across a range of scales.

4. Parameter evaluation

The approach we will adopt is as follows: given the parameters for the operator, how can we set those for both the normal and tangential complexity such that, within the local extent, a line is detected as a line? Our goal is therefore to find the conditions under which something simple will be found by applying the simplicity and tangent quantization principles. The procedure we chose will first proceed with the normal complexity parameters, and then concentrate on tangential complexity. We now write explicitly Ω_N and δ_N for the spatial extent and scale for the normal analysis, and denote by $\hat{\Omega}_T$ and $\hat{\delta}_T$ those for the tangential complexity.

4.1. ...for the normal complexity. Because the parameters are interrelated, we will first set Δ_N to be 1.5, and then find the size of the *local extent* Ω_N over which to perform the computations. This will be fixed for the whole image, and will be driven by the curves with the largest separation, i.e. we are trying to find the largest extent for which the curve assumption is satisfied. The idea in principle is to perform the computations for all spatial extents at every position x subtending a non empty tangent set with $\Omega_N(x) \subset I^2$. The spatial extent chosen (Ω_N) will be the maximum extent such that there exists a point x with non-empty tangent set with low normal complexity and high tangential complexity. This ensures that, at this position and over this extent, the curve assumption is satisfied. For small spatial extents, and if the discrete tangent map is not empty, one will always find such points from the definition of the edge detection operator. Moreover, the extent is bounded above by the total extent of the image.

However, since this procedure is very computationally expensive, we selected the spatial extent in the neighborhood of the maximal separation d_{max} between two entries in the discrete tangent map. The value for d_{max} was obtained by applying isotropic dilations to the discrete tangent map at all scales through a distance transform [Borgefors 1986], as was described in Chapter 4. This then ensures that there is a set of points F over which the separation is at least d_{max}, taking this to be a lower bound for the local extent Ω_N. We then verified that for the chosen extent, the required conditions were met.

Once the local extent Ω_N is obtained, the characteristics of the operator constrain the choice of the scale δ_N for the analysis to satisfy the tangent separation and quantization principles. For instance, δ_N has to be larger than the image resolution but smaller than $(\Omega_N\sqrt{2})/2$. What then would be an interesting choice for δ_N? This will be answered by our constraints. Given the operator, we want to test the "curve assumption" from our principles and overcome the discretization artifacts. For instance, we empirically observed that our line detector can allow up to 4 parallel tangents to occur when the line is at an awkward orientation (see Section 6 in Chapter 3), therefore

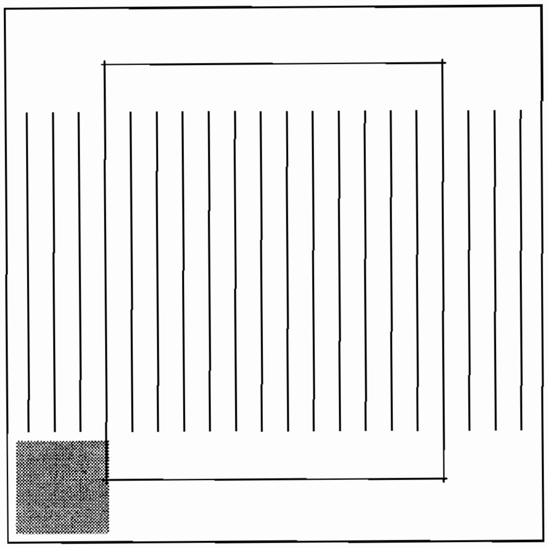
Decision rule 1 (Scale for normal complexity). *Given Ω_N, Constraints #4 (parallel spreading), #5 (intersections), and #6 (shape of bounding region) provide that, to have $C_N(x, \delta_N)$ less than $\Delta_N = 1.5$ for a line, the scale should be taken as* $\max\{1.5/N, \Omega_N/4\} \leq \delta_N \leq \Omega_N/2$. *We will therefore choose* $\delta_N = \max\{1.5/N, \Omega_N/4\}$.

Constraint #4 says that, in the case where δ_1 is the spacing between adjacent pixels and for a line detector requiring up to 4 parallel lines to represent a single line passing at a given orientation, then the scale δ_N so that the complexity is smaller than $\Delta_N = 1.5$ needs to be $\delta_N > 1.5$ pixels. Note that in all cases studied in this thesis, the dominating constraint was the one with respect to intersections.

Let us look at the results obtained for the Kanizsa pattern. The extent for the normal complexity obtained from our method is shown Fig. 5.13a by the grey square. The tangent map image being embedded in the unit square, we find $\Omega_N = 0.162$. The driving feature for the choice of local extent is the spacing between the side of the rectangle and the edge of the unit square (see Fig. 5.12). The scale for the computation is then set to be $\delta_N = \Omega_N/4 = 0.045$ in this case.

Extents for the Perceptron images and the Ullman examples are shown in Fig. 5.14. In all cases the chosen local extent $\Omega_N(x)$ will be a translated version of the grey square shown in the figures. The chosen local extent for Paolina is shown in Fig. 5.13b. The actual numbers for the extent and scale are given in Table 6.1 (at the beginning of the next chapter).

4.2. ...for the tangential complexity. As mentioned before, the approach for the tangential complexity will be slightly different. Our algorithm to estimate tangential complexity will provide useful results only if the dilations are done at small scales. The parameters will therefore be bound to the digitization and the operator. The spatial extent we choose will be fixed for all the images to $\hat{\Omega}_T = 8$ pixels. One exception is the Kanizsa pattern for which we chose $\hat{\Omega}_T$ to be 5 pixels,



(a) Discrete tangent map for the Kanizsa pattern with its local extent Ω_N for the normal complexity

(b) Discrete tangent map for the Paolina image with its local extent Ω_N

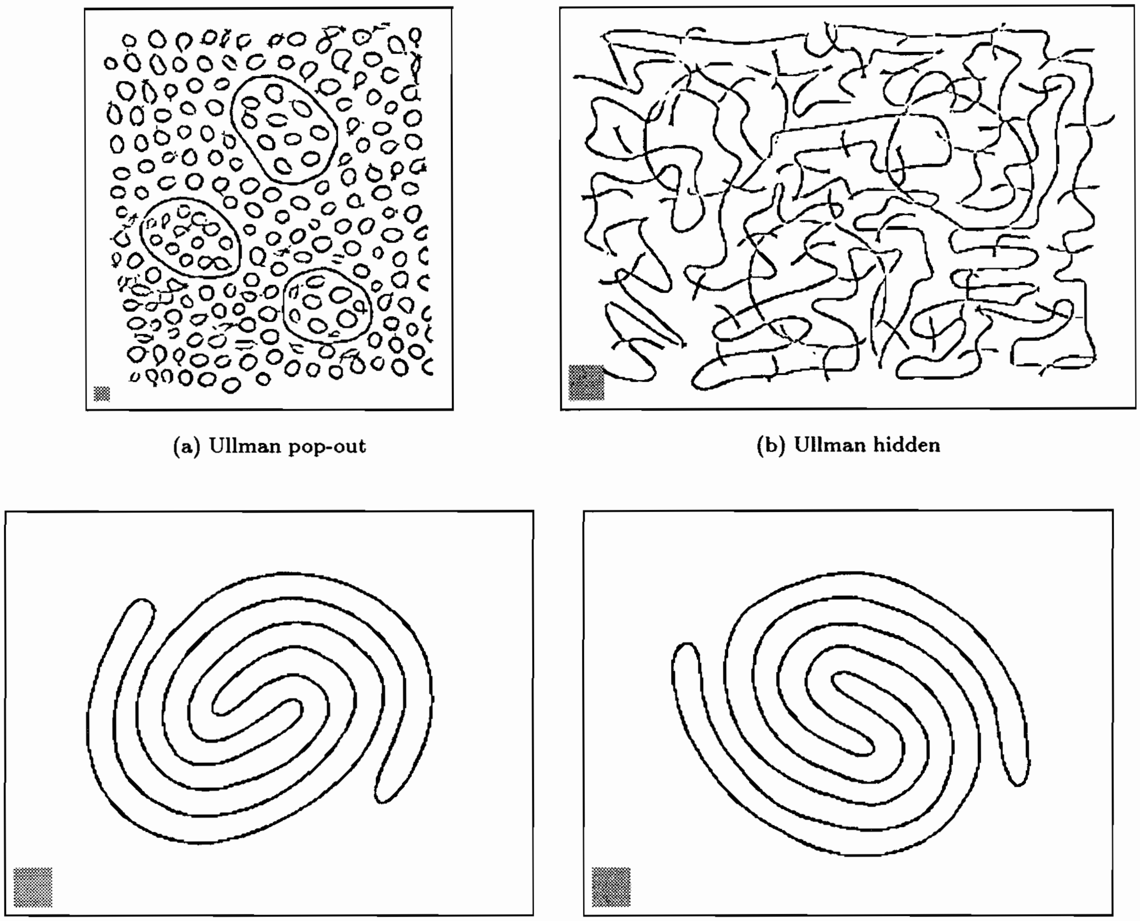
FIGURE 5.13. Local extent Ω_N for the complexity analysis. In both cases, the extent Ω_N is obtained by the size of the side of the grey square. The position of the square shows where the conditions were met.

since the discrete tangent map was derived directly from the structure of the set. As mentioned before, we chose Δ_T to be 0.73. The normal extent $\hat{\omega}_N$ was set to 1.1 pixels, therefore

Decision rule 2 (Scale for tangential complexity). *Given $\hat{\Omega}_T = 8$ pixels, $N = 51$ and $\Delta_T = 0.73$, the scale $\hat{\delta}_T = 1.59$ pixels ensures satisfaction of Constraints #1 (continuity), #2 (curvature), and #3 (orientation).*

5. Local extent and scale selection

A discussion about our choice for selecting the scale and the spatial extent is appropriate at this moment. Following our approach, the spatial extent Ω_N for a single line centered in the unit square will be the unit square itself. The scale will be $\delta_N = 0.25$ and a point on the middle of the line will be seen as a piece of line with a normal complexity of 1.0. Our choice for δ_N ensures that quantization effects will not alter the result desired. But a single line would not be the only pattern to satisfy the curve assumption. A grating of n equidistant lines, for instance, would



(a) Ullman pop-out

(b) Ullman hidden

(c) Perceptron: one curve

(d) Perceptron: two curves

FIGURE 5.14. Local extent for the normal complexity: in each subfigure, the local extent is shown as a grey square.

also be considered as such, but over a different local extent, and at a different scale. This time the local extent Ω_N will be approximately $\frac{2}{n}$ and the scale $\delta_N \approx \frac{1}{2n}$. All the tangents in the map will satisfy the curve assumption, *with respect to the (much smaller) spatial extent.* If one were to compare the grating and the line however (as in Fig 5.3a), the results would be totally different and the grating would be seen as being complex, texture-like, 2-dimensional as soon as n would be bigger than 2. This illustrates the idea of *relative complexity* that we introduced at the beginning of this chapter which is central to our approach as will be shown in the next chapter.

Another justification for our approach can be given from Paolina's discrete tangent map. The leading feature for the choice of the spatial extent was the shoulder

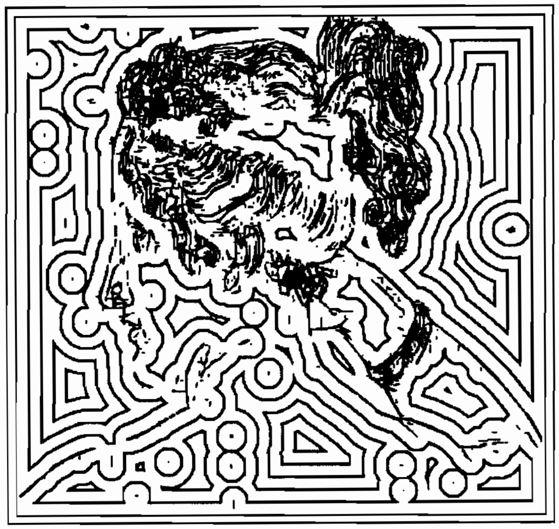


FIGURE 5.15. Camouflage on the discrete tangent map. Adding contours to camouflage the "curves" (in the complexity space sense).

Normal complexity	**Tangential complexity**
intersections	continuity
parallel spreading	discretization
bounding shape	curvature

TABLE 5.1. The constraints for the complexity analysis parameters.

region. We can force the spatial extent to be smaller by adding contours to the image (see Fig. 5.15). One notices now that the shoulder, the chin, and part of the arm structure are camouflaged. Taking the same local extent on this new image would result in an empty curve substructure. The line for the shoulder is nothing special now. It is part of another equivalence class – a texture flow – and would be detected as such.

6. Summary

This chapter presented one of the most important concepts in this thesis, namely the use of the complexity map to segment an image based on its structural properties. This was achieved by partitioning the complexity space into four distinct regions leading to a classification of objects into *dust*, *curves*, *turbulence* and *flow*. We then showed how it was possible to set the parameters to achieve such a segmentation. Experimental results are kept for the next chapter.

The partitioning scheme that we chose was elementary. Single values of the scale and spatial extent parameters were chosen for the entire image. However, more involved schemes are certainly possible, if not desirable. The partition could be refined by taking smaller regions of the complexity space, and extent parameters could vary over the image. The advantage is that the pattern discrimination scheme could then be much stronger and better suited for specific visual tasks. The price to pay will be that it will make the choices of the complexity map parameters much more complex.

CHAPTER 6

Experiments and results

The last three chapters have presented a new intermediate representation to be used for bridging the gap between image-based and object-based representations. The goal is to be able to group local elements into the appropriate structure: in our case we chose *dust* (curve-free), *curves* (satisfying the curve assumption), *turbulence*, and *flow* (oriented textures). Four examples were carried along from the first chapter: the Kanizsa pattern, the Perceptron spirals, the Ullman figures, and the Paolina image. This chapter will show the results obtained when first estimating the complexity map on these images, and then applying our grouping scheme.

For the first two types of patterns we analyzed, the geometrical structure was known ahead of time, therefore the discrete tangent map could be estimated. Originally, the discrete tangent map was defined in terms of the output of edge detectors. But for the Kanizsa pattern we simulated it by setting up a grid and discretizing position and orientation as was described in Chapter 3. It was a reasonable thing to do because the curves were not crossing (except for the corners in the Kanizsa pattern), and the spacing between distinct curves was always known ahead of time. This example can be seen as a somewhat "ideal discrete world" and will be used to enlighten the concepts presented in the previous chapters. The second section presents examples with real images for which the tangent map structure must be inferred. It starts with the analysis of the Perceptron spirals for which the discrete tangent map was obtained from the output of line detectors, and is then followed by the Ullman patterns. We end with the Paolina image.

The values computed for the various parameters are given in Table 6.1. The parameters in this table are those that are case dependent. In all cases the scale δ_N for the normal complexity was set to $\Omega_N/4$, according to the analysis in the previous chapter. For the tangential complexity, the scale $\hat{\delta}_T$ was set to be 1.59 pixels except for the Kanizsa pattern to reflect the difference in spatial extent. The tangential extent $\hat{\omega}_T$ (for the normal analysis) was set to 1.4 pixels to account for the square tessellation, while the normal extent $\hat{\omega}_N$ (for the tangential complexity) was set to 1.1 pixels for the same reasons. The values for the partition of the complexity space were $\Delta_N = 1.5$ and $\Delta_T = 0.73$.

In all cases the subresolution $p_x = p_y = 10$, i.e. each pixel from the domain of the discrete tangent map was further divided in a 10x10 subgrid for the projection of the tangents and the oriented dilations. As far as the scale neighborhood is concerned, we took in each case 8 points equidistant; i.e. $K_T = K_N = 8$. The spacing between the scale was in "subresolution" units (resolution $*$ subresolution): 2 in the case of normal complexity, 1 in the case of tangential complexity. To justify these choices, a series of experiments on groundtruth values were performed and the results are given in Appendix B.

1. When structure is known a priori...

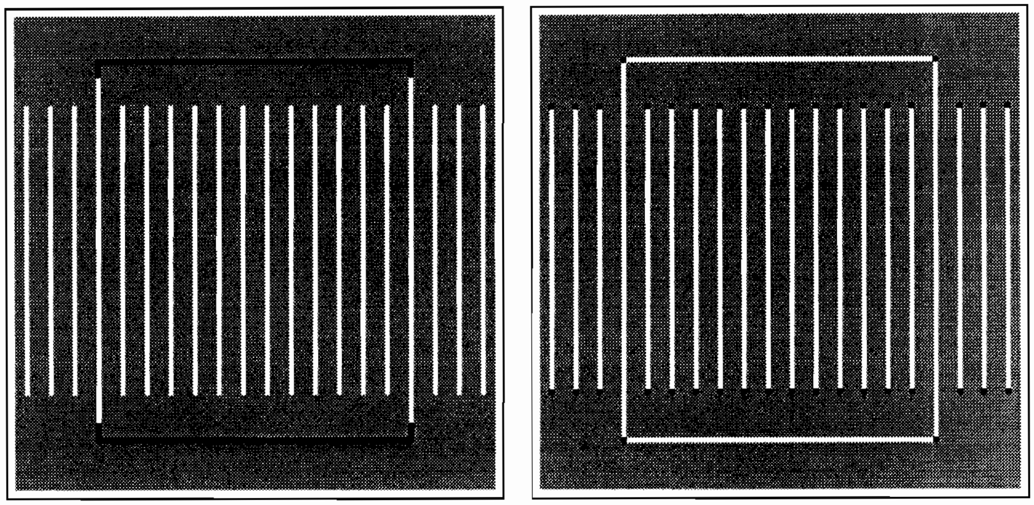
1.1. The Kanizsa pattern. The complexity map was estimated for the Kanizsa pattern (Fig. 1.7a) and the results are shown in Fig. 6.1. In (a) we display the normal complexity mapped as black and white, while in (b) we map the tangential complexity. For the normal complexity, black stands for low normal complexity while white refers to tangents embedded in a dense region. From this we see that, at the studied scale, the patch stands out as requiring a different representation from the "curves" part (the parts of the hollow rectangle in the top and in the bottom). But our classification can be finer if we add the tangential complexity component (Fig.6.1b), and would be incomplete if we did not consider it (see Section 3.2 in Chapter 4). Now black stands for low tangential complexity, and white for high tangential complexity: places where the object extends along its length. The only black spots here are the

Test image	Res	Normal complexity parameters		Tangential complexity parameters	
	N	Ω_N	δ_N	$\hat{\Omega}_T$	$\hat{\delta}_T$
Kanizsa pattern	100	0.162	0.0405	5	0.90
Perceptron spirals (one curve)	200	0.070	0.0175	8	1.59
Perceptron spirals (two curves)	200	0.070	0.0175	8	1.59
Ullman pattern (pop-out)	250	0.032	0.0080	8	1.59
Ullman pattern (hidden)	350	0.062	0.0155	8	1.59
Paolina	512	0.120	0.0300	8	1.59

TABLE 6.1. The values computed for the complexity analysis. The parameters that stayed constant throughout the analysis are discussed in the text.

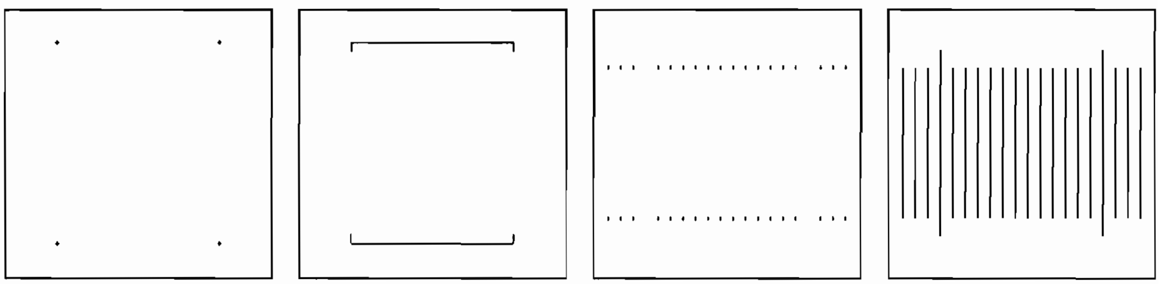
four corners and the tips of the grating pattern. The tangential complexity is therefore a good means of detecting corners and ends of lines (we will show later how it can detect points of high curvature), while the normal complexity is ideal to segregate textures from curves, when the two coexist within the scene.

The normal and tangential indexes can be collated and bundled to form the complexity map. Applying the classification scheme presented in the last section, it is possible to extract the various components: dust, curves, turbulence, and flow. This is shown in Fig. 6.2. The segmentation successfully extracts the corners, the end of lines in the grating, the "curves" part and the grating itself. We recall that the partition for the complexity space was $\Delta_N = 1.5$ and $\Delta_T = 0.73$. Most interestingly, this computational experiment reveals surprising aspects of our structural classes, e.g., the emergence of subjective borders as 1-D *turbulence* distributions (somehow similar to Heitger & von der Heydt's *ortho* groupings [Heitger & von der Heydt 1993]), and orientation discontinuities (corners) as *dust*.



(a) normal complexity (b) tangential complexity

FIGURE 6.1. Computing the normal and tangential complexity for the Kanizsa pattern. In both cases the complexity has been remapped as black or white. For the normal complexity map, the white pixels refer to tangents embedded in a complex region while the black ones represent low complexity. The "curves" regions clearly stand out in black here. In the case of tangential complexity, this time the white pixels represent regions where the set extends along its length and the black pixels, regions where the tangential complexity is low, i.e. end of lines, corners, points of high curvature.



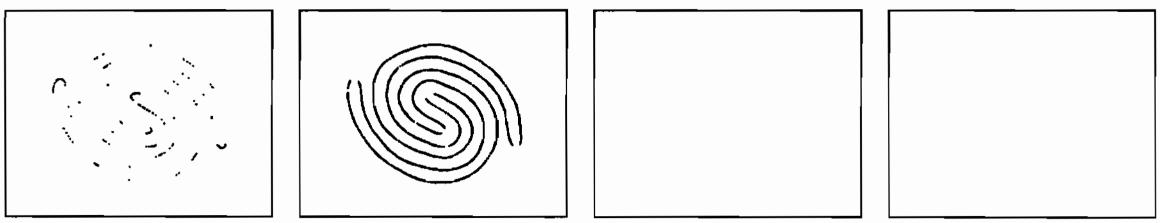
(a) dust (b) curves (c) turbulence (d) flow

FIGURE 6.2. Segmented image of the Kanizsa pattern using the complexity map and our classification rule.

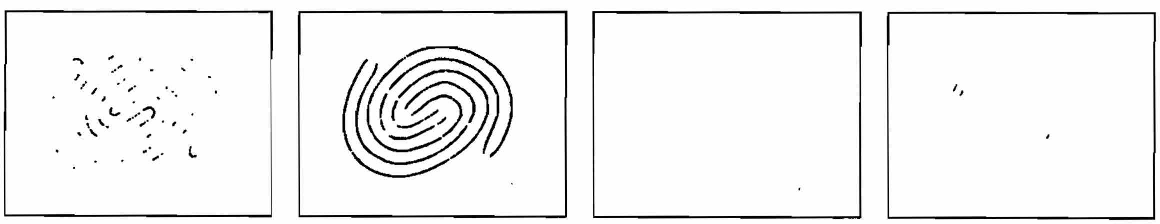
2. When structure must be inferred...

2.1. The Perceptron spirals. The case of Perceptron spirals is interesting but, once more, exceptional since it is highly degenerate: we will see that some of the subclasses will be empty. Fixing a value Ω_N^0 for the normal complexity, applying our rule provides the following: computations for the normal complexity with extents *larger* than $\Omega_N > \Omega_N^0$ leads to almost all points being categorized as being high normal (textures); computations for extents *smaller* than $\Omega_N < \Omega_N^0$ leads to almost all points being categorized as being low normal (curves). It is interesting that the singular "transition" point, taking the extent to be $\Omega_N = \Omega_N^0$, emphasizes the effects of digitization. We therefore chose $\Omega_N < \Omega_N^0$.

As shown in Fig. 6.3, the complexity maps have identical distributions both for the normal and tangential complexity. From the segmentation only, it is now almost impossible to tell which of the two sets satisfying the curve assumption belongs to the "one curve" pattern and which belongs to the "two curves" pattern. Remember that the extent over which one can safely integrate (follow the edges) is very small: $\Omega_N = 0.035$. This confirms our early prediction that the two objects belong to the same equivalence class for this scale, therefore they are very difficult to distinguish from one another; i.e., it is hard to decide which of the two is connected and which is not.



(a) dust　(b) curves　(c) turbulence　(d) flow



(e) dust　(f) curves　(g) turbulence　(h) flow

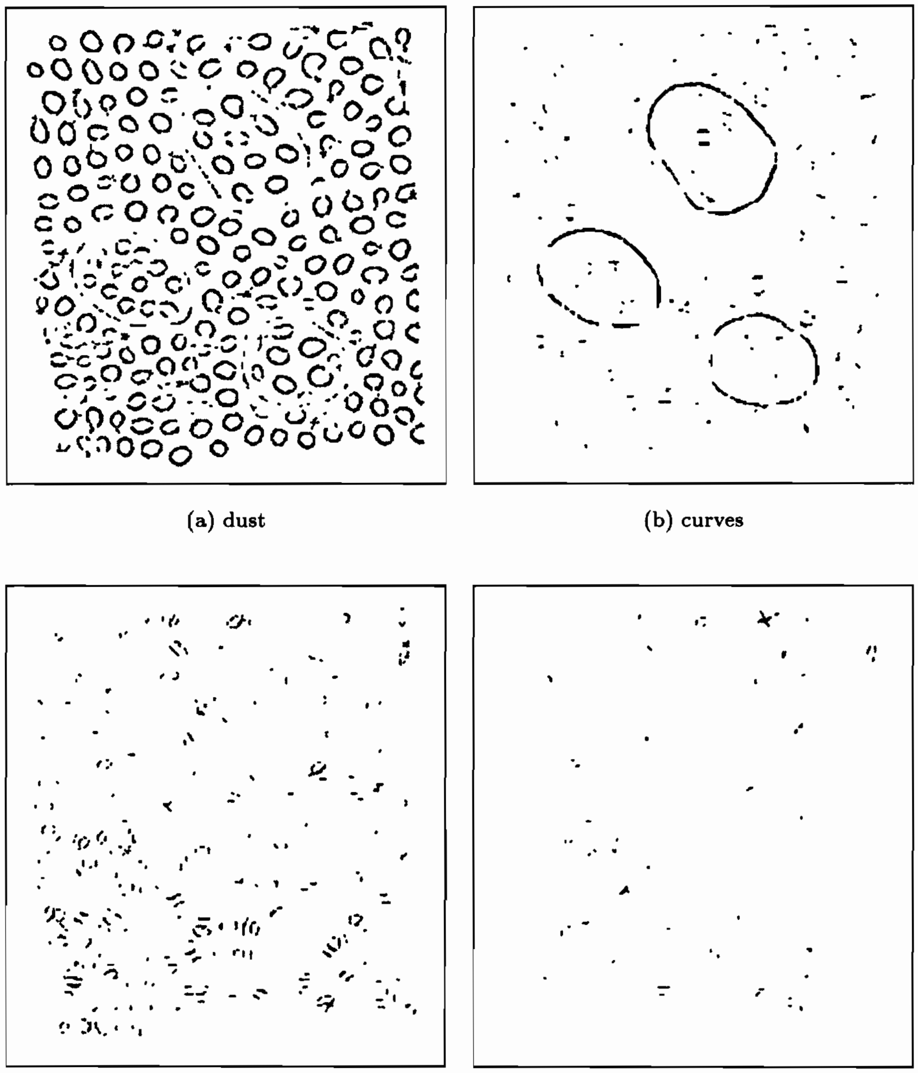
FIGURE 6.3. Segmented image of the Perceptron spirals using the complexity map and our classification rule. (a) and (e) low normal, low tangential: dust (b) and (f) low normal, high tangential: curves (c) and (g) high normal, low tangential: turbulence (d) and (h) high normal, high tangential: flow. Where is the "one curve" figure, where is the "two curves" figure?

2.2. The Ullman patterns. The algorithm was applied to the Ullman patterns (Fig. 1.5). In the case of the Ullman pop-out pattern (Fig. 1.5a), we already reported that tangential dilations should allow the segregation between the small and the large blobs by allowing only curves with small curvature to be classified in the *curves* substructure. This is shown on Fig. 6.4, where the three large blobs emerge in the "curves" section, while most of the rest is classified as *dust*, a segmentation much in the spirit of what was originally proposed by Sha'ashua & Ullman (1988).

The case of the Ullman "hidden" pattern (Fig. 1.5b) is more subtle: the tangential complexity is more uniformly distributed, not allowing anything to stand out. This is further verified in the segmented maps. As opposed to Fig. 6.4, where the large blobs were clearly poping out in the curve section, no such structure emerges in the "curves" substructure. This clearly goes against Ullman's [1990] analysis according to which the three blobs were supposed to pop-out.

Fig. 6.5 reveals a few interesting points. First, notice the points of high curvature and the line endings appearing in black in the tangential complexity image (Fig. 6.5a). These points of high curvature are known to have perceptual significance as advocated originally by Attneave (1954). Secondly, notice how the intersections show up in the flow section (Fig. 6.6d). This was to be expected from the choice of the parameters for the complexity analysis and the segmentation scheme. Since the crossings are uniformly distributed, it is hard to perceive the circles. The distractors therefore play a big role in the partial camouflage of the blobs. If the distractors are removed (as it was in the original version of this image [Mahoney 1987] shown in Fig. 6.7a), then the circles tend to be more salient. A cue to these circles is obtained in the "flow" region, as shown in Fig. 6.8c, where the crossings outline the circles. Compare the "curves" and "flow" substructures between the two figures (Figs 6.6 and 6.8).

We mentioned in the introduction two papers related to the saliency map: [Ullman 1990] and [Sha'ashua & Ullman 1988]. It is interesting to see the evolution of Fig. 1.5b. Originally [Mahoney 1987, Sha'ashua & Ullman 1988], the distractors were not present in the image (as shown in Fig. 6.7a). They were added in a subsequent paper [Ullman 1990]. I believe that, this is to support Ullman's statement about the fact that in this case "the local structure has no conspicuous local part

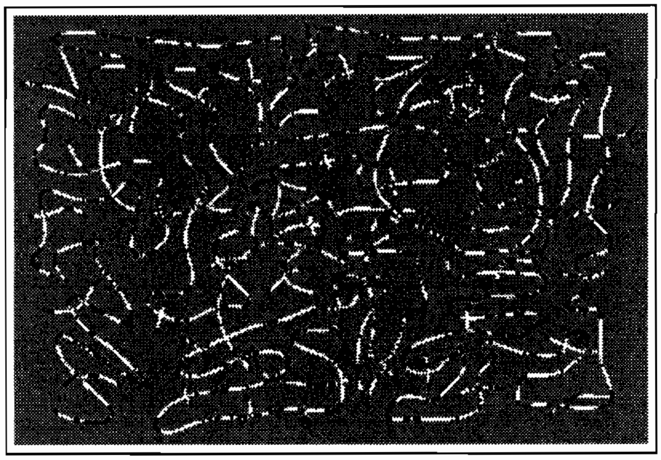


(a) dust (b) curves

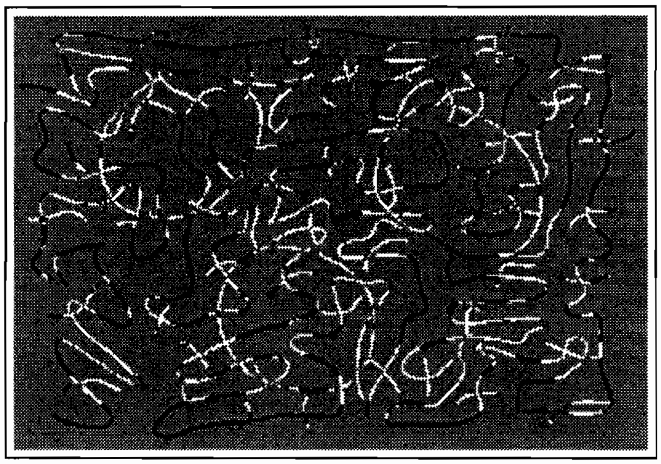
(c) turbulence (d) flow

FIGURE 6.4. Segmented image of the Ullman pop-out image using the complexity map and our classification rule. (a) low normal, low tangential: dust (b) low normal, high tangential: curves (c) high normal, low tangential: turbulence (d) high normal, high tangential: flow.

having a distinguishing local property" [Ullman 1990, p.892]. In the original image, this was not true as the curve crossings were such conspicuous local parts. Adding the distractors (Fig. 1.5b) made the statement correct, and this fact was captured by our complexity analysis and the subsequent grouping.

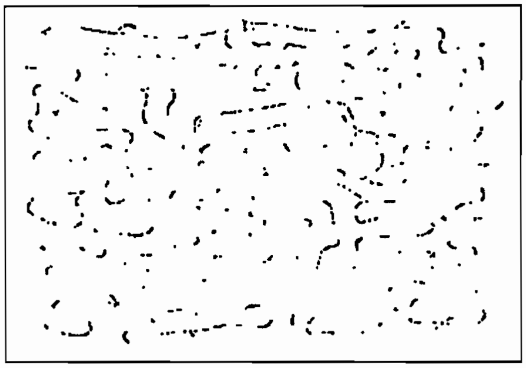


(a) tangential complexity

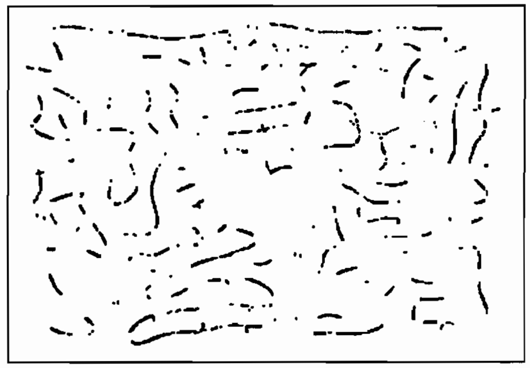


(b) normal complexity

FIGURE 6.5. Normal and tangential complexity for the Ullman patterns. Once more the complexity has been remapped as black or white. Normal complexity: white pixels refer to tangents embedded in a dense region while the black ones represent those sufficiently separated. Notice the intersections being captured in the high normal complexity. Tangential complexity: white pixels represent regions where the set extends along its length, and the black pixels, regions where the tangential complexity is low: line endings, corners, points of high curvature.

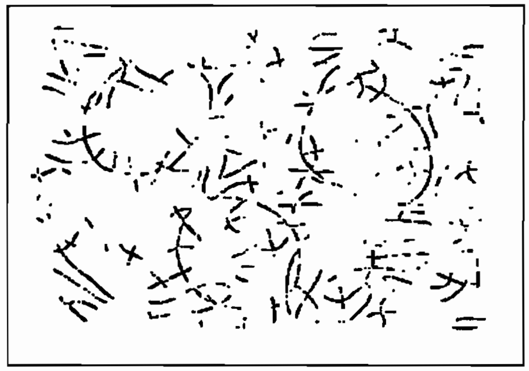


(a) dust



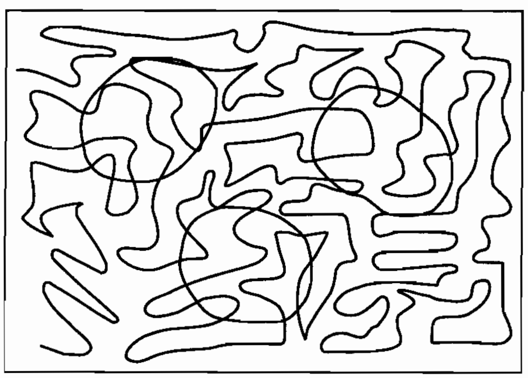
(b) curves

(c) turbulence

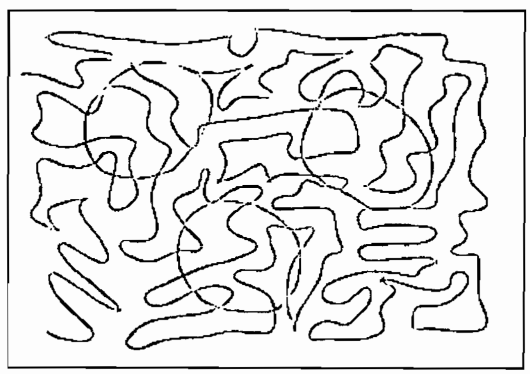


(d) flow

FIGURE 6.6. Perceptual grouping from complexity. Segmented image of the Ullman pop-in image using the complexity map and our classification rule.

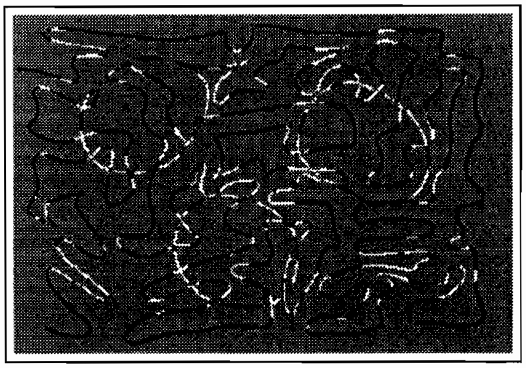


(a) original image

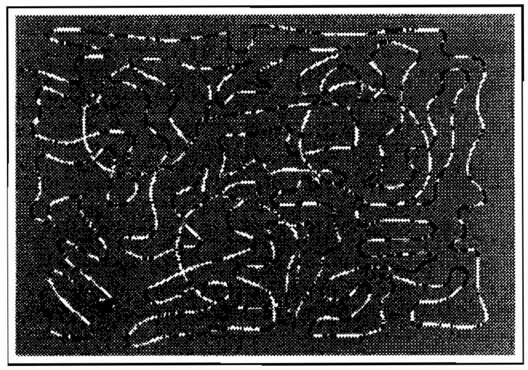


(b) discrete tangent map

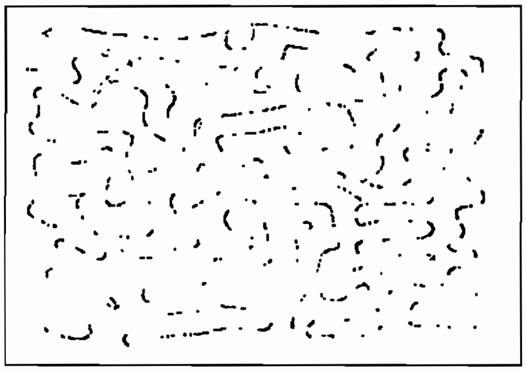
FIGURE 6.7. The original Ullman (hidden) pattern as can be seen in [Mahoney 1987]. In (a) we have the original image and in (b) the discrete tangent map obtained as the output of L/L operators. Compare this with Fig. 1.5b and decide in which of the two the three circles are the most salient.



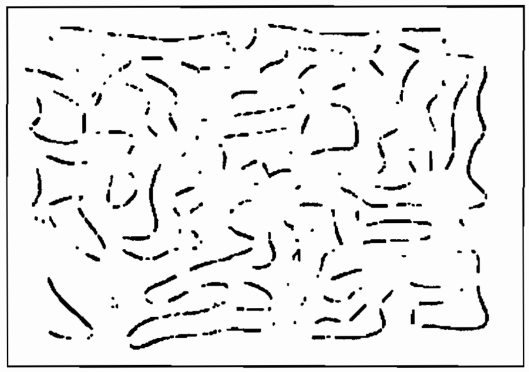
(a) normal complexity



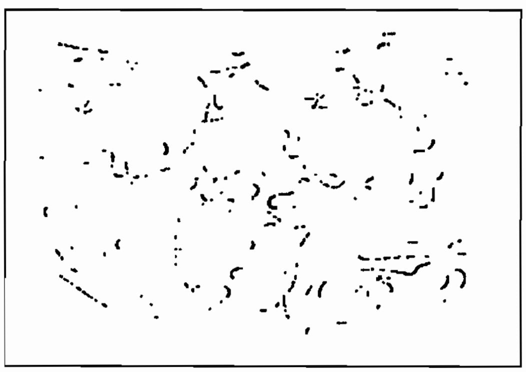
(b) tangential complexity



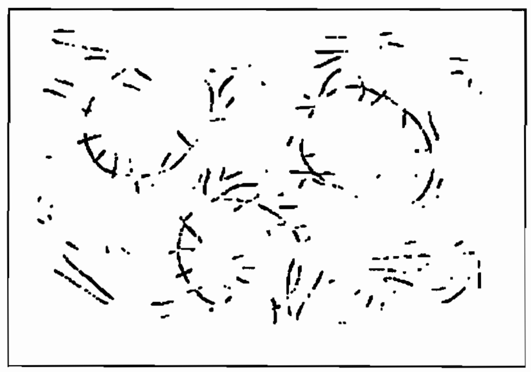
(c) dust



(d) curves



(e) turbulence



(f) flow

FIGURE 6.8. The segmented Ullman (hidden) pattern as can be seen in [Mahoney 1987]. (a) and (b) show the normal and tangential complexity maps remapped as black and white, while the usual substructures are shown in (c)-(f). Although the blobs do not really pop-out in the *curves* substructure, cues are given within the *flow* substructure from the curve intersections.

2.3. Our dearest Paolina... Now we proceed to the analysis of the Paolina image. The original image had a resolution of 512x480. The discrete tangent map (for the edges) was computed as explained in Chapter 3. The result was shown in Fig. 1.4f. The complexity maps are shown in Fig. 6.9. For the tangential complexity, we see that the spurious responses in the back get low tangential complexity. The same goes for line endings. One surprising result is the fact that the wind-blown hair has mostly high tangential complexity even if many points have high curvature. This is due to our choice of parameters. The extent being very small, and the normal extent being large, most growth is saturated within this region. In the normal complexity, the curves clearly stand out in black, namely the shoulder, chin and arm. Most of the hair structure is seen as being complex.

With the complexity map in hand, we can apply our segmentation scheme. Results of the grouping are shown in Fig. 6.10. The segmentation is as expected. The curves contain the tangents that are sufficiently separated, with no curve intersections. The flow, contains the curves that are dense and extend along their length. The dust substructure contains the spurious responses, some ends of lines, etc. It is within the "curves" substructure that the grouping into curves should go on, but spatial extent is then limited by Ω_N which was set to 0.12. This means that one can integrate safely only over that size of a region, since it is over that neighborhood that the conditions for the curve assumption were verified. Taking a neighborhood larger would not ensure that curves would be sufficiently separated, and that a curve representation would be efficient. To further illustrate, let us take back the camouflaged shoulder in Fig. 5.15. Doing the complexity analysis and our segmentation using the same local extent and scale for the normal complexity as we did for the original image will have the effect of eliminating most of the edges in the curves substructure. The shoulder at this scale would disappear. To link back with Chapter 1 and the Pick-Up Sticks example, our fingers would then be too wide to pick the (camouflaged) shoulder as opposed to the original image.

We also did the experiment with different parameters for the tangential complexity. These results were originally reported in [Dubuc & Zucker 1995]. Taking a larger

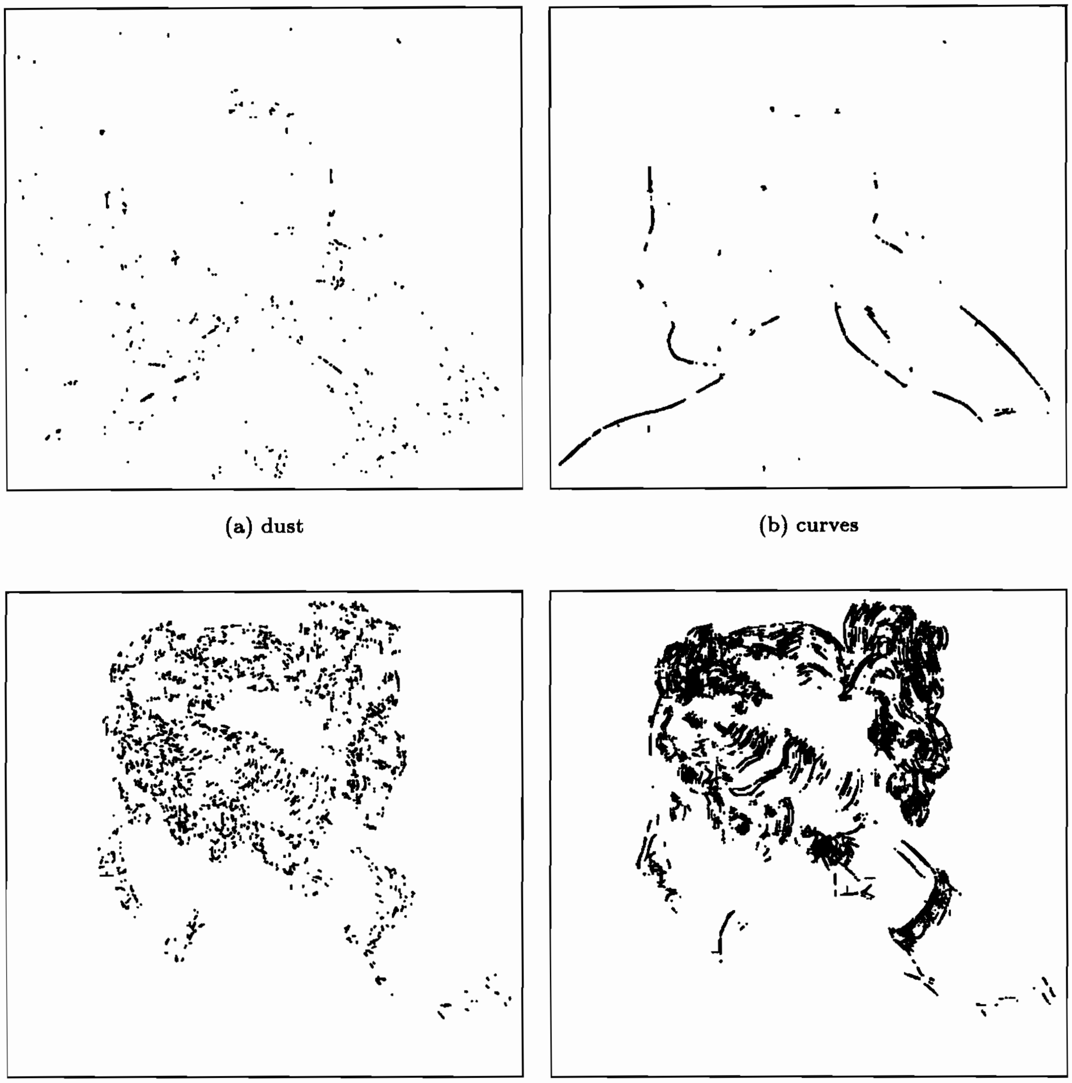
(a) tangential complexity

(b) normal complexity

FIGURE 6.9. Normal and tangential complexity for the Paolina image. As usual, the complexity (normal and tangential) has been remapped as black or white.

spatial extent for the tangential analysis provided the segmentation in Fig. 6.11. Notice the difference between this (Figs. 6.11c-d) and our previous attempt (Fig. 6.10c-d). This time the *turbulence* substructure region captures more of the windblown hair, while the *flow* consists primarily of the well-combed hair; i.e., those that extend along their length but are dense in the normal direction. This difference in tangential structure as a function of scale is analogous to what happened for normal complexity in Fig. 5.12, but much less extreme.

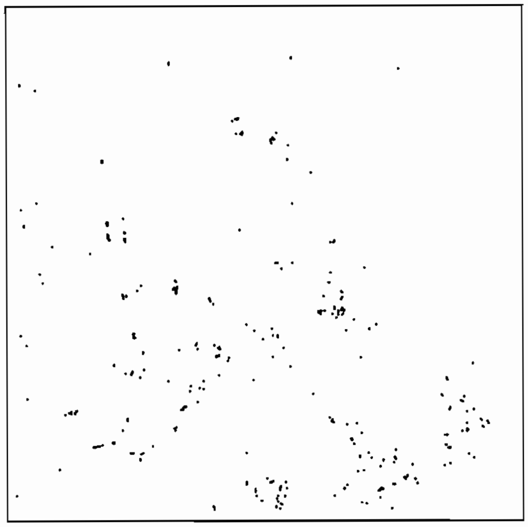
We should point out the fact that most of the analysis done was uniquely geometric. There is also a photometric aspect that should be studied [Breton 1994]: i.e. the existence of lines with various contrasts and in particular positive and negative contrast lines. These coexist with edges. In this example, we only considered edges, but the relationship between the two (edges and lines) needs to be taken care of in a complete treatment of the integration of local information.



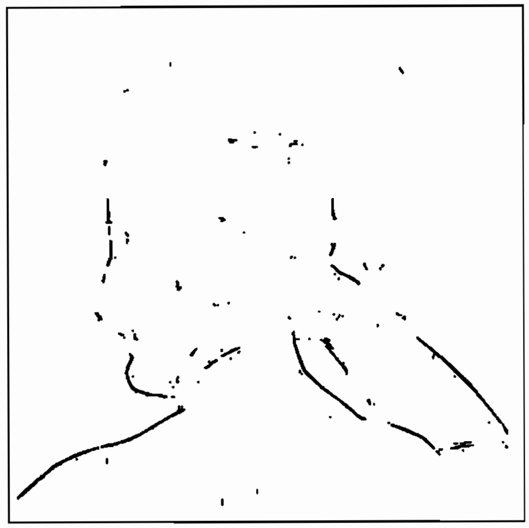
(a) dust (b) curves

(c) turbulence (d) flow

FIGURE 6.10. Segmented image of the Paolina image using the complexity map and our classification rule. The integration under a 'curve' representation should only take place within the "curves" substructure.



(a) dust



(b) curves

(c) turbulence

(d) flow

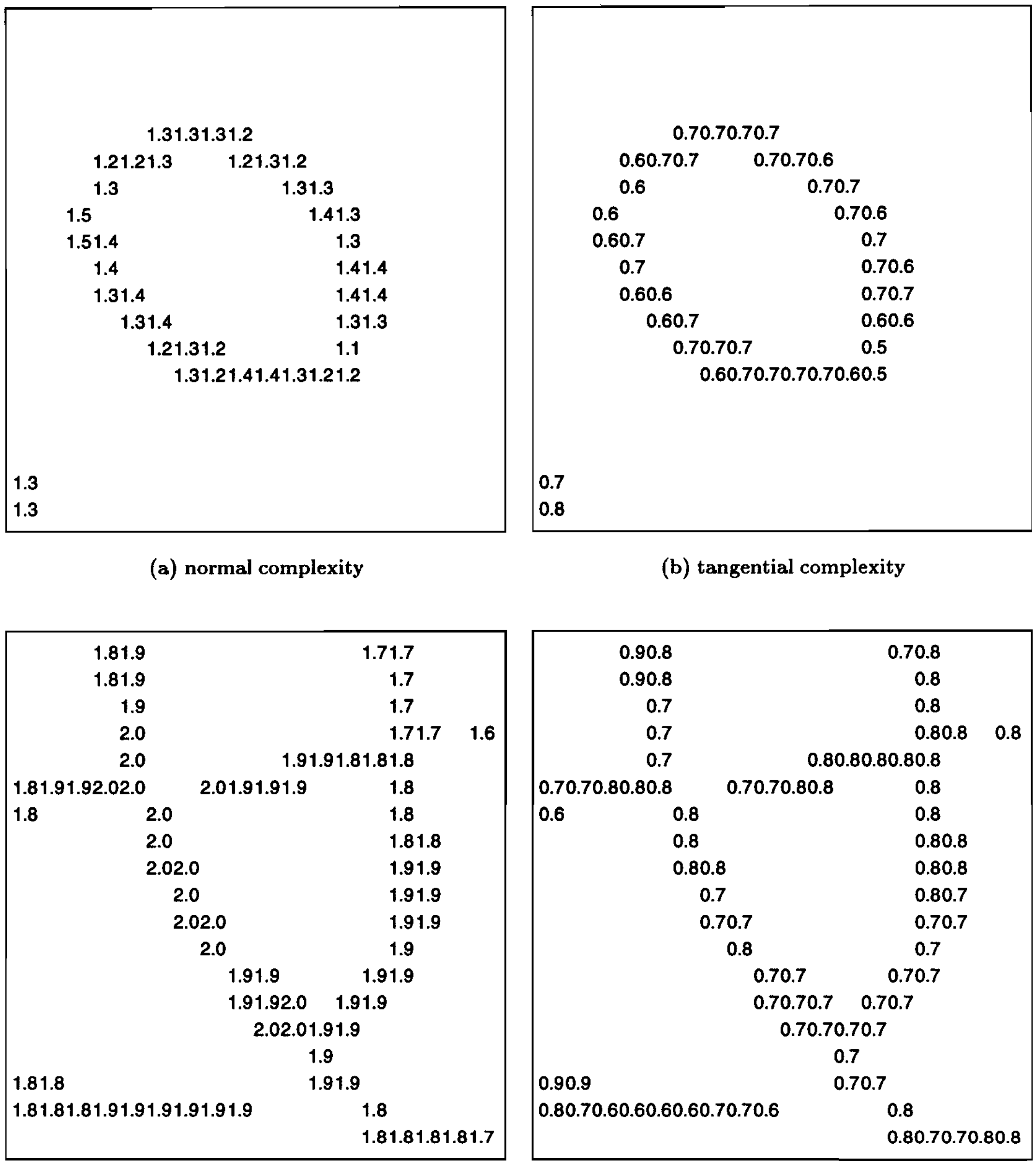
FIGURE 6.11. Segmentation using a different tangential complexity map. This time the spatial extent for the tangential analysis was taken to be larger, resulting into a larger portion of windblown hair getting classified in the "turbulence" substructure.

3. Back to the transversality/quantization problem

We are now ready to return to the problem of distinguishing between transversality and quantization, as presented in Chapter 3. It can be solved from an argument linked to the *multiple tangent separation theorem* (Thm 3.2). The idea is that, if the normal complexity is low, multiple tangents would be simply an artifact of the discretization of orientation space. From our analysis design, the hypothesis of multiple tangents at a point can only be retained in the high normal complexity substructures.

We saw earlier that the normal complexity for two lines crossing at a sufficiently large angle would be bigger than $\Delta_N = 1.5$ if the scale and spatial extent were chosen appropriately. This is how we will determine if, when two tangents are "on" at the same location, it is due to two lines crossing or to an artifact of quantization[1]. In the case of two lines crossing, the normal complexity will be higher than the value for the partition that would select the point as being simple. Our partitioning scheme will therefore ensure us not to allow crossings *per se* to occur in the low normal complexity substructures. One can then use this information to take the appropriate decision, as shown in Fig. 6.12, where we show the values for the normal and tangential complexity indexes obtained for the parts of the Ullman patterns discrete tangent maps shown in Fig. 5.4. A grouping scheme for the tangent map of simple curve-like sets was proposed by David & Zucker (1990). In their paper, they mentioned the use of branching potentials (or what they describe as *"a split of the potential distribution into layers"*) for places where curves would intersect. Our technique provides now a way of detecting where the branching could occur: when integrating through the 'curves' substructure, if one loses track at some point, in the neighborhood of an intersection, and if in this neighborhood one finds multiple tangents in the 'flow' substructure, then one has found a candidate for branching potentials. This is exactly what is meant by a local to global transition for curves.

[1] Note that two lines crossing do not necessary imply that a point in the discrete tangent map will have a tangent set with more than one tangent, as is pointed out in [Nitzberg et al. 1993] and shown in Fig. 5.4.



(a) normal complexity

(b) tangential complexity

(c) normal complexity

(d) tangential complexity

FIGURE 6.12. The complexity indexes for the transversality/quantization problem.

4. An application: clusters in axonal arbors

As our final example, we will present an application of the technique to axonal reconstruction data [Dubuc et al. 1994]. In neuroscience, a central issue in characterizing neuronal growth patterns is whether their arbors form clusters [Antonini & Stryker 1993]. Formal definitions of clusters have been elusive, although intuitively

5. Summary and discussion

The analysis of discrete tangent maps through their complexity provided a grouping scheme prior to integration. Clearly it split the sets into perceptually meaningful classes. Later processes can come now into play. In the case of curves, one can use edge following [Ramer 1975], dynamic coverings [David & Zucker 1990], token grouping [Saund 1992], or spline fitting (see for instance [Nitzberg et al. 1993]). In the case of flows, one could use the approaches of [Kass & Witkin 1987, Zucker 1985] while for turbulence, one might opt for a statistical characterization [Haralick 1979, Gagalowicz 1980] or approaches such as [Malik & Perona 1990].

How do these results relate to other edge grouping approaches? We have tried to make clear in this thesis that all aspects of the distribution of local information were important before performing the grouping. The "curves" substructure is where the curve grouping should happen. Hazardous positions in the image, as well as issues of usefulness of the representation, are captured by the spatial extent chosen and the other substructures. Is this approach following the criteria of Cox et al. (1993) presented in Chapter 2? It (i) provided a mechanism to group edges from the segmentation of the complexity space; (ii) has a model of smoothness through the tangential complexity; (iii) included the noise model at a prior stage, but digitization artifacts were dealt with for both types of complexity analysis; (iv) set up explicitly a process to detect curve intersections. But it went further in that it did not presuppose a representation *a priori* allowing for curves and textures to coexist and to be treated as such.

CHAPTER 7

Conclusion

The transition from local representations to global ones is a key problem in computer vision. In this thesis, we have presented an intermediate representation scheme for the local structure of a curve-like set, the *discrete tangent map*, and addressed the problems that are encountered when attempting to blindly follow local information. The example with the Kanizsa pattern was to that extent very informative. Within a single scale, one observes different types of objects coexisting within the scene:

(i) *discontinuities*: these were of two kinds
  (a) the four corners,
  (b) the subjective edge generated by the line endings of the grating patch.
(ii) *curves*: the top and bottom of a hollow rectangle,
(iii) *texture*: the grating patch

The *complexity map*, computed from the discrete tangent map, clearly supported what was observed. We have therefore demonstrated that the complexity of the tangent map, obtained through normal and tangential dilations, was key to determine the representation underlying the grouping of edge elements and the dimensionality of its support. Stated differently, we showed how the idea of early or primitive image segmentation could be lifted onto the tangent map, and effected through notions of complexity.

While curves and textures have been part of the descriptive repertoire for early vision, they are traditionally viewed as having no relationship. Curves are normally conceptualized via differential geometry, and textures via statistics. We have presented a unifying theory of early visual structure from the more abstract perspective of geometric measure theory. The theory is built upon the notion of a tangent map, as this derives from "edge detection", and results in a measure of structural complexity that requires separate normal and tangential components. Thus the descriptive repertoire is defined not just by the tangents, or the salient length property of curves, but by the full spatial context in which tangents are distributed. Natural extrema in the complexity map defined semantically-meaningful structural categories, including *dust*, *curves*, *turbulence*, and *flows*.

This thesis concentrated first on the careful definition of the type of objects to be studied, and stressed the need for a discrete tangent map, while providing constraints on its structure. It then introduced oriented dilations to allow the quantification of density (from the normal dilations), and continuity (from the tangential dilations), which eventually led to the complexity map. Our approach differs from the classical ones in fractal analysis [Tricot 1995], in the sense that it studies sets dilated with respect to their local structure. Another key difference is the use of the rate of growth at a particular scale rather than around zero. The separation from normal and tangential dilations is crucial to be able to classify the types of objects, therefore justifying the use of the discrete tangent map as an intermediate representation: only the tangent could provide the local orientation of the set at a particular scale.

1. Future directions

We have presented in this thesis a technique to characterize the complexity of the edges prior to grouping. As satisfying as the work done so far is, however, I now realize that this is only the beginning. Many readers have probably found within our arguments places for further investigation and are curious about how the theory presented relates to other areas of visual information processing. The following is an attempt to place pointers to key issues.

1.1. In mathematics. I have tried to introduce concepts from geometric measure theory to computer vision, and hope it will inspire researchers involved in edge detection and grouping. There are many technical aspects that remain to be resolved in the theory presented. Amongst the most important are the verification of the existence conjectures for the normal and tangential complexity. Moreover, a careful analysis is needed to predict under which conditions one is ensured to obtain reliable results. Some of this work as already been started for a small class of functions [Dubuc & Dubuc in press], a first step toward a general solution. In this thesis, we showed the validity and relevance of the analysis by applying it to a range of practical examples, and by studying its behavior on simple sets.

The local extent selection procedure, in Chapter 5, also needs to be refined. We based the selection on isotropic dilations of the edge map, but a formal scheme for selection entirely related to the work presented here is missing. One possibility would be to select the spatial extent with respect to the scale of the operator (as we did for the tangential complexity) and keep this fixed for all images. The results would not be as relevant to the choice of a final representation however. The same remark can be made about the extent for tangential complexity. Most was related to the scale of the operator. It would be nice to investigate avenues that would rather consider the structure of the image.

1.2. In psychology. The local to global problem has been extensively studied in psychology. The work presented in this thesis could be particularly relevant to the following:

(i) *Theory of attention.* How do different objects grab one's attention within a scene? Treisman (1985) has done extensive work on all aspects involved in attention. How can one, in a general setting, decide on which objects in the scene should be of interest? In this thesis we leaned toward the idea that complexity is driving attention, that objects different from their context pop out, that representations leading to simple descriptions are the ones that win. Psychological experiments along these lines would provide better evidence for our approach.

(ii) *Gestalt principles.* The Gestalt psychologists [Koffka 1935] postulated a series of grouping principles, such as the principle of good continuation for curves, which suggests why we perceive a figure "8" as a single, non-simple curve that

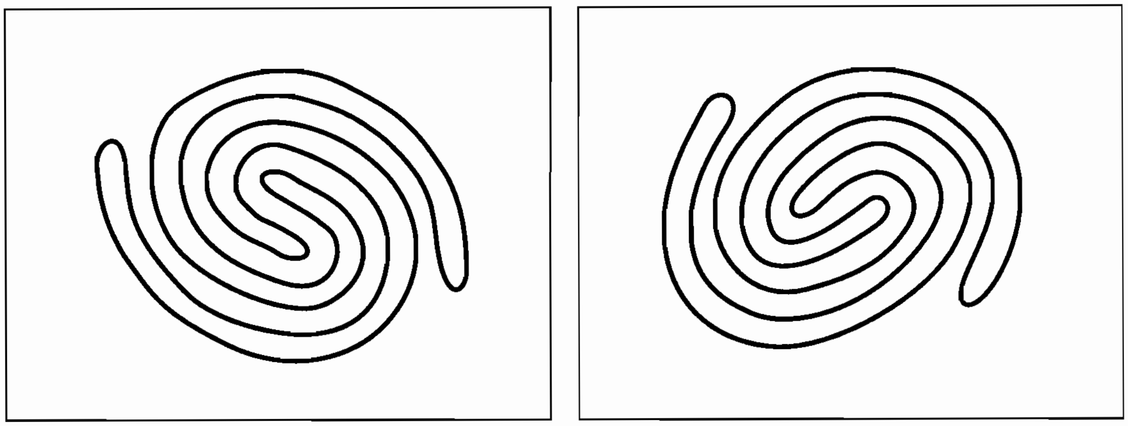


FIGURE 7.1. Connected or not? Again the Perceptron spirals leave us with a tremendous challenge: why is is so difficult to tell the two panels apart?

crosses itself, rather than two "circles", one on top of the other. Other principles include proximity, closure, symmetry and familiarity. One interpretation of the Gestalt theory is that people seem to perceive the simplest interpretation of any given data. A lack of ability to quantify this notion of "simplicity" made the theory unsatisfactory. How do Gestalt theories relate to our technique?

1.3. In theory of computing. The complexity map constitutes the first building block of an emerging representational complexity theory for curve-like sets, which should take its root from the theory of computing. In the vein of Turing computable numbers, it should now be possible to define the notion of *representable curve*. Representational complexity theory would generalize the notions presented in this thesis and would dictate the action that could be taken before integrating information. The goal would be to predict what should be the salient features in the scene and justify unambiguously the choices of possible final representations. One main result that one should get is that two patterns falling within the same equivalence class ought to be difficult to tell apart. This takes us back to the Perceptron spirals (see Fig. 7.1): why are the two patterns so hard to tell apart (besides the obvious rotation of the pattern)?

The connection to theoretical computer science is necessary to generalize the formal basis from geometric measure theory to include algorithmic complexity. In particular, as algorithmic complexity is based on the notion of a minimal program, the parallel between the time needed to perform a visual task and a minimal length algorithm is clearly central. The resulting theory would provide a complexity measure that will enable the feasibility of particular algorithms and representations to be evaluated against task requirements. Such measures have been completely lacking, making the approach to many vision tasks *ad hoc* and opportunistic. Such opportunism is of course frustrating, as it is impossible to ascribe cause to failure. In computer vision, such frustration is widespread, but representational complexity would provide an elegant solution.

2. Back to scale-space

In Chapter 2, we briefly discussed the selection of scale within the scale-space representation. This remains a controversial subject. Even if it seemed clear to many researchers that a unique fixed scale is an inappropriate solution, it is always discussed even in the latest studies (see for instance [Lindeberg 1993] or [Elder & Zucker 1995]).

The classical scale-space approach advocated this detection-localization scheme where coarse scales were used to detect features, and fine scales localized them. Recently, Elder & Zucker (1995) have questioned this approach and suggested a different use of the scale-space representation by selecting a "reliable scale" at each point in the image (which can vary from one point to the next). Did we advocate something different? One important point in my opinion is that one cannot get everything. For instance, using the finest scale to represent all the features in the image would be very inefficient. Fine scales are for small regions, large scales for larger regions: keep the *scale size/spatial extent* ratio constant. Once information is needed at fine scale, one needs to focus in that region, but then the information from coarser scales is lost. This notion of sacrifice seems to be unnatural, but I think is needed to provide any sophistication in the processing of visual information.

3. Le mot de la fin...

Because of the connection between vision and geometry, this research was based on the mathematics of geometric measure theory. It now forms part of a curve detection system, and could be considered for applications in areas as diverse as image coding (for telecommunications), optical character recognition, and mobile robotics.

To keep this theory concrete and industrially relevant, we have already begun applying the use of complexity measures to characterize the roughness of surfaces in pharmaceutical sciences [Pimienta et al. 1994] and the study of neuronal arborizations in neuroscience [Dubuc et al. 1994]. Its extension to higher dimensions is so natural that one could consider using it for other information areas. The resultant theory should provide a key to break complexity and to organize information efficiently: an essential asset these days where information is becoming one of the most valuable resources.

REFERENCES

Allman, J. & Zucker, S. W. (1990), Cytochrome oxidase and functional coding in primate striate cortex: A hypothesis, *in* C. S. H. L. Press, ed., 'Cold Spring Harbor Symposia on Quantitative Biology', pp. 979–981.

Antonini, A. & Stryker, M. P. (1993), 'Development of individual geniculocortical arbors in cat striate cortex and effects of binocular impulse blockade', *J. Neuroscience* **13**(8), 3549–3573.

Attneave, F. (1954), 'Some informational aspects of visual perception', *Psychological Review* **61**, 183–193.

Ballard, D. H. & Brown, C. M. (1982), *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ.

Beck, J. (1982), Textural segmentation, *in* 'Organization and Representation in Perception', Lawrence Erlbaum Associates, Hillsdale, NJ.

Besicovitch, A. S. (1928), 'On the fundamental geometrical properties of linearly measurable plane sets of points', *Mathematische Annalen* **98**, 422–464.

Borgefors, G. (1984), 'Distance transformations in arbitrary dimensions', *Computer Vision, Graphics and Image Processing* **27**, 321–345.

Borgefors, G. (1986), 'Distance transformations in digital images', *Computer Vision, Graphics and Image Processing* **34**, 344–371.

Breton, P. (1994), Recovering shape from shading and contour, PhD thesis, McGill University, Dept. of Electrical Engineering, Montréal.

Brodatz, P. (1966), *Textures: a Photographic Album for Artists and Designers*, Dover Publishing Co.

Burkill, J. C. & Burkill, H. (1970), *A Second Course in Mathematical Analysis*, Cambridge University Press.

Canny, J. F. (1986), 'A computational approach to edge detection', *IEEE Trans. Pattern Anal. Machine Intell.* **8**, 679–698.

Cauchy, A. S. (1850), 'Mémoire sur la rectification des courbes et et la quadrature des surfaces courbes', *Oeuvres de Cauchy* **1**, 167–177.

Cesari, L. (1965), *Surface Area*, Princeton University Press, Princeton, NJ.

Cover, T. M. & Thomas, J. A. (1991), *Elements of Information Theory*, Wiley, New York.

Cox, I. J., Rehg, J. M. & Hingorani, S. (1993), 'A Bayesian multiple-hypothesis approach to edge grouping and contour segmentation', *International Journal of Computer Vision* **11**(1), 5–24.

David, C. & Zucker, S. W. (1990), 'Potentials, valleys, and dynamic global coverings', *International Journal of Computer Vision* **5**(3), 219–238.

Deriche, R. (1987), 'Using Canny's criteria to derive a recursively implemented optimal edge detector', *International Journal of Computer Vision* **1**(2), 167–187.

do Carmo, M. (1976), *Differential Geometry of Curves and Surfaces*, Prentice Hall.

Dubuc, B. (1988), On estimating fractal dimension, Master's thesis, McGill University, Dept of Electrical Engineering, Montréal.

Dubuc, B., Quiniou, J.-F., Roques-Carmes, C., Tricot, C. & Zucker, S. W. (1989), 'Evaluating the fractal dimension of profiles', *Phys. Rev. A* **39**(3), 1500–1512.

Dubuc, B. & Zucker, S. W. (1995), Indexing visual representations through the complexity map, *in* IEEE Comp. Society, ed., 'Proc. of the 5th ICCV', Cambridge, MA, pp. 142–149.

Dubuc, B., Zucker, S. W. & Stryker, M. P. (1994), 'Complexity maps reveal clusters in neuronal arborizations', *Fractals* **2**(2), 297–301.

Dubuc, S. (1982), Une foire de courbes sans tangentes, *in* 'Actualités Mathématiques: Actes du VI^e Congrès du Regroupement des Mathématiciens d'Expression Latine, Luxembourg, September 7–12 1981', pp. 99–123.

Dubuc, S. & Dubuc, B. (in press), 'Error bounds on the estimation of fractal dimension', *SIAM Num. Anal.* .

DuPain, Y., Kamae, T. & Mendès-France, M. (1986), 'Can one measure the temperature of a curve?', *Arch. of Rat. Mech. and Anal.* pp. 155–163.

Edgar, G. A. (1990), *Measure, Topology, and Fractal Geometry*, Undergraduate Texts in Mathematics, Springler-Verlag, New York.

Elder, J. & Zucker, S. W. (1995), Scale-space surfaces and blur estimation, *in* 'Proc., Vision Interface 95', Québec City, Canada, pp. 140–147.

Falconer, K. J. (1987), *The Geometry of Fractal Sets*, Vol. 85 of *Cambridge Tracts in Mathematics*, Cambridge University Press, Cambridge, England.

Falconer, K. J. (1990), *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons Ltd, Chichester, England.

Federer, H. (1985), *Geometric Measure Theory*, Springler-Verlag, Berlin-Heidelberg, Germany.

Freeman, W. T. & Adelson, E. H. (1991), 'The design and use of steerable filters', *IEEE Trans. Pattern Anal. Machine Intell.* **13**(9), 891–906.

Gagalowicz, A. (1980), Visual discrimination of stochastic texture fields based upon their second order statistics, *in* 'Proc. of the 5th Int. Conf. on Pattern Recogn.', pp. 786–788.

Galli, A. & Zama, A. (1931), 'Untersuchungen über die wahrnehmung ebener geometrischen figuren die ganz oder teilweise von anderen geometrischen figuren verdeckt sind', *Zeitschrift für Psychologie* **123**, 308–348.

Gibson, J. J. (1950), *The Perception of the Visual World*, Riverside Press, Cambridge, MA.

Guzman, A. (1968), 'Decomposition of a visual scene into three-dimensional bodies', *AFIPS Conf. Proc.* **33**, 291–304.

Haralick, R. M. (1979), Statistical and structural approaches to texture, *in* 'Proc. of the IEEE', Vol. 67, pp. 786–804.

Heitger, F. & von der Heydt, R. (1993), A computational model of neural contour processing: figure-ground segregation and illusory contours, *in* IEEE Comp. Society, ed., 'Proc. of the Fourth ICCV', pp. 32–40.

Hilbert, D. & Cohn-Vossen, S. (1990), *Geometry and the Imagination*, Chelsea Publishing Company, New York.

Hildreth, H. (1992), Edge detection, *in* S. C. Shapiro, ed., 'Encyclopedia of Artificial Intelligence', second edn, Vol. 1, Wiley Interscience, New York, pp. 257–267.

Hochstein, S. & Spitzer, H. (1984), 'Zero-crossing detectors in primary visual cortex?', *Biological Cybernetics* **51**, 195–199.

Hubel, D. H. & Wiesel, T. N. (1962*a*), 'Receptive fields and functional architecture of monkey striate cortex', *J. Physiology (London)* **195**, 215–243.

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

Hurewicz, W. & Wallman, H. (1948), *Dimension Theory*, Princeton University Press, Princeton, NJ.

Iverson, L. A. (1993), Toward discrete geometric models for early vision, PhD thesis, McGill University, Dept. of Electrical Engineering, Montréal.

Iverson, L. A. & Zucker, S. W. (1995), 'Logical/linear operators for image curves', *IEEE Trans. Pattern Anal. Machine Intell.* **17**(10), 982–996.

Jones, J. P. & Palmer, L. A. (1987), 'An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex', *J. Neurophys.* **58(b)**, 1233–1258.

Julesz, B. (1981), 'Textons, the elements of texture perception, and their interactions', *Nature* **290**, 91–97.

Kanizsa, G. (1979), *Organisation in Vision*, Praeger, New York.

Kass, M. & Witkin, A. (1987), 'Analysing oriented patterns', *Computer Vision, Graphics and Image Processing* **37**, 362–385.

Koenderink, J. (1990), *Solid Shape*, MIT Press, Cambridge, MA.

Koenderink, J. J. (1984), 'The structure of images', *Biological Cybernetics* **50**, 363–370.

Koffka, K. (1935), *Principles of Gestalt psychology*, Hartcourt, Brace & World, New York.

Kolmogorov, A. N. (1965), 'Three approaches to the quantitative definition of information', *Problems Inform. Transmission* **1**(1), 1–7.

Kolmogorov, A. N. (1987), Information theory and theory of algorithms, *in* Y. V. Prokhorov & A. N. Shiryaev, eds, 'Selected Works', Vol. 3, Nauka. (in Russian).

Levine, M. D. (1985), *Vision in Man and Machine*, McGraw-Hill, New York.

Leymarie, F. & Levine, M. D. (1992), 'Fast raster scan distance propagation on the discrete rectangular lattice', *CVGIP: Image Understanding* **55**, 84–94.

Lindeberg, T. (1991), Discrete scale-space theory and the scale-space primal sketch, PhD thesis, Royal Institute of Technology, Stockholm, Sweden.

Lindeberg, T. (1993), On scale selection for differential operators, *in* 'Proc. of the 8th Scand. Conf. on Im. Anal.', Tromsø, Norway, pp. 857–866.

Lowe, D. G. (1989), 'Organisation of smooth image curves at multiple scales', *International Journal of Computer Vision* **3**, 119–130.

Lowe, D. G. & Binford, T. O. (1979), The interpretation of three-dimensional structure from image curves, *in* 'Proc. IJCAI-7', Vancouver, Canada, pp. 613–618.

Lowe, D. G. & Binford, T. O. (1987), Segmentation and aggregation: an approach to figure-ground phenomena, *in* M. A. Fishler & O. Firschein, eds, 'Readings in Computer Vision: Issues, Problems, Principles and Paradigms', Morgan Kaufmann Publishers, Inc., Los Altos, CA, pp. 282–292.

Mahoney, J. V. (1987), Image chunking: defining spatial building blocks for scene analysis, M.I.T. Technical Report 980, MIT Artificial Intelligence Laboratory, Cambridge, MA.

Malik, J. & Perona, P. (1990), 'Preattentive texture discrimination with early vision mechanisms', *J. of the Opt. Soc. of America A* **7**, 923–932.

Mandelbrot, B. B. (1970), 'How long is the coast of Britain?', *Physica Scripta* **32**, 257–260.

Mandelbrot, B. B. (1982), *The Fractal Geometry of Nature*, Freeman, San Francisco.

Marr, D. (1982), *Vision*, W.H. Freeman, San Francisco.

Marr, D. & Hildreth, E. (1980), 'Theory of edge detection', *Proc. Royal Society Lond.* **207**, 187–217.

Matheron, G. (1975), *Random Sets and Integral Geometry*, John Wiley & Sons, New York.

Mendès-France, M. (1991*a*), Dimension and entropy of regular curves, *in* G. Cherbit, ed., 'Fractals: Non-integral Dimensions and Applications', John Wiley & Sons Ltd, Baffins Lane, Chichester, pp. 222–230.

Mendès-France, M. (1991*b*), The Planck constant of a curve, *in* S. Dubuc & J. Bélair, eds, 'Fractal Geometry and Analysis', Kluwer Academic Publishers, pp. 325–366.

Minsky, M. & Papert, S. (1972), *Perceptrons*, M.I.T. Press, Cambridge, Massachussetts.

Montanari, U. (1972), 'On the optimal detection of curves in noisy images', *Commun. ACM* **15**(5), 335–345.

Morel, J.-M. & Solimini, S. (1995), *Variational Methods in Image Segmentation*, Vol. 14 of *Progress in Nonlinear Differential Equations and Their Applications*, Birkhaüser, Boston.

Morgan, F. (1987), *Geometric Measure Theory*, Academic Press, Inc.

Mumford, D. (1992), Elastica and computer vision, *in* 'Algebraic Geometry and Applications', Springler-Verlag, Heidelberg.

Munkres, J. R. (1975), *Topology: A First Course*, Prentice-Hall.

Nitzberg, M. & Mumford, D. (1990), The 2.1D sketch, *in* IEEE Comp. Society, ed., 'Proc. of the 3rd ICCV', Osaka, Japan, pp. 138–144.

Nitzberg, M., Mumford, D. & Shiota, T. (1993), *Filtering, Segmentation and Depth*, Vol. 662 of *Lecture Notes in Computer Science*, Springler-Verlag.

Pagels, H. R. (1988), *The Dreams of Reason: the Computer and the Rise of the Sciences of Complexity*, Simon and Schuster, New York.

Parent, P. & Zucker, S. W. (1989), 'Trace inference, curvature consistency, and curve detection', *IEEE Trans. Pattern Anal. Machine Intell.* **11**(8), 823–839.

Peitgen, H.-O., Jürgens, H. & Saupe, D. (1992), *Chaos and Fractals: New Frontiers of Science*, Springer-Verlag, New York.

Peleg, S., Naor, J., Hartley, R. & Avnir, D. (1984), 'Multiple resolution texture analysis and classification', *IEEE Trans. Pattern Anal. Machine Intell.* **6**(4), 518–523.

Pentland, A. P. (1984), 'Fractal-based description of natural scenes', *IEEE Trans. Pattern Anal. Machine Intell.* **6**(6), 661–674.

Pentland, A. P. (1985), 'On describing complex surface shapes', *Image and Vision Computing* **3**(4), 153–162.

Perona, P. & Malik, J. (1990), Detecting and localizing edges composed of steps, peaks and roofs, *in* IEEE Comp. Society, ed., 'Proc. of the 3rd ICCV', pp. 52–57.

Pimienta, C., Dubuc, B. & Tawashi, R. (1994), 'Surface fractal dimension and the quantification of roughness of titanium implant material', *Cells and Materials* **4**(4), 379–386.

Poincaré, H. (1926), Pourquoi l'espace a trois dimensions, *in* 'Dernières Pensées', Flammarion.

Ramer, U. (1975), 'Extraction of line structures from photographs of curved objects', *Computer Vision, Graphics and Image Processing* **4**, 81–103.

Roberts, L. G. (1965), Machine perception of three-dimensional solids, *in* 'Optical and Electro-Optical Information Processing', MIT Press, Cambridge, Mass., pp. 159–197.

Rogers, C. A. (1970), *Hausdorff Measures*, Cambridge University Press, London.

Santaló, L. A. (1976), *Integral Geometry and Geometric Probability*, Vol. 1 of *Encyclopedia of Mathematics and its Applications*, Addison-Wesley Publishing Company, Reading, MA.

Saund, E. (1992), Labeling of curvilinear structure accross scales by token grouping, *in* IEEE Comp. Society, ed., 'Proc. of the IEEE CVPR', pp. 257–263.

Serra, J. (1982), *Image Analysis and Mathematical Morphology*, Academic Press, London.

Sha'ashua, A. & Ullman, S. (1988), Structural saliency: the detection of globally salient structures using a locally connected network, *in* IEEE Comp. Society, ed., 'Proc. of the 2nd ICCV', pp. 321–327.

Simmons, G. F. (1963), *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York.

Simon, H. A. (1968), *The Sciences of the Artificial*, MIT Press.

Smith, K. T. (1971), *Primer of Modern Analysis*, Bogden and Quigdley, Inc.

Steinhaus, H. (1954), 'Length, shape and area', *Colloquium Mathematicum* **III**, 1–13.

Steinhaus, H. (1983), *Mathematical Snapshots*, Oxford University Press, New York. (third American edition).

Treisman, A. (1985), 'Preattentive processes in vision', *Computer Vision, Graphics and Image Processing* **31**, 156–177.

Tricot, C. (1991), Rectifiable and fractal sets, *in* S. Dubuc & J. Bélair, eds, 'Fractal Geometry and Analysis', Kluwer Academic Publishers, pp. 367–403.

Tricot, C. (1995), *Curves and Fractal Dimension*, Springler-Verlag, New York.

Tricot, C., Quiniou, J.-F., Wehbi, D., Roques-Carmes, C. & Dubuc, B. (1988), 'Evaluation de la dimension fractale d'un graphe', *Revue Phys. Appl.* **23**, 111–124.

Turing, A. (1936), 'On computable numbers with an application to the entscheidungsproblem', *Proc. London Math. Soc.* **42**, 230–265. Correction, Ibid, 43:544–546, 1937.

Ullman, S. (1990), Three-dimensional object recognition, *in* 'Cold Spring Harbor Symposia on Quantitative Biology', Vol. LV, Cold Spring Harbor Laboratory Press, pp. 889–898.

Valiron, G. (1966), *Théorie des Fonctions*, Masson et Cie, Paris.

Viéville, T. & Faugeras, O. (1990), Feed-forward recovery of motion and structure from a sequence of 2-d line matches, *in* IEEE Comp. Society, ed., 'Proc. of the 3rd ICCV', Osaka, Japan, pp. 517–521.

von Helmholtz, H. (1867), *Handbuch der physiologischen Optik*, Voss, Mineola, NY. J. P. C. Southall (ed.).

Waltz, D. (1975), Understanding line drawings of scenes with shadows, *in* P. Winston, ed., 'The Psychology of Computer Vision', McGraw-Hill, New York, pp. 19–91.

Witkin, A. (1992), Scale-space methods, *in* S. C. Shapiro, ed., 'Encyclopedia of Artificial Intelligence', second edn, Vol. 1, Wiley Interscience, New York, pp. 973–980.

Witkin, A. P. (1983), Scale-space filtering, *in* 'Proc. of the 8th Int. Joint Conf. Art. Intell.', Kalsruhe, West Germany, pp. 1019–1022.

Young, R. A. (1985), The gaussian derivative theory of spatial vision: analysis of cortical receptive field line-weighting profiles, Technical Report TR-GMR 4920, General Motors Research.

Zucker, S. W. (1985), 'Early orientation selection: tangent fields and the dimensionality of their support', *Computer Vision, Graphics and Image Processing* **32**, 74–103.

Zucker, S. W. (1992), Early vision, *in* S. C. Shapiro, ed., 'Encyclopedia of Artificial Intelligence', second edn, Vol. 1, Wiley Interscience, New York, pp. 394–420.

Zucker, S. W. (1993), 'Endowing AI with Vision: A biological and computational perspective', *Current Science* **64**(6), 407–418.

Zucker, S. W., Dobbins, A. & Iverson, L. A. (1989), 'Two stages of curve detection suggest two styles of visual computation', *Neural Computation* **1**, 68–81.

Zucker, S. W., Hummel, R. A. & Rosenfeld, A. (1977), 'An application of relaxation labelling to line and curve enhancement', *IEEE Transactions on Computers* **26**, 394–403.

APPENDIX A

The complexity of simple sets

1. Normal complexity

1.1. For a pair of crossing segments. To calculate the complexity of a pair of crossing segments at a given scale δ, we need first to calculate $|E_N(\epsilon)|_2$. In this case we have

Proposition A.1. *Let E be a pair of lines of length $2r$ crossing at an angle $0 < \theta \leq \pi/2$ and centered in a disk of radius r. Then*

(A.1)

$$|E_N(\epsilon)|_2 = \begin{cases} 4\left(r^2 \arcsin\left(\frac{\epsilon}{r}\right) + \epsilon\sqrt{r^2-\epsilon^2} - \dfrac{\epsilon^2}{\sin\theta}\right) & \text{if } 0 < \epsilon \leq \epsilon_0 \\ \theta r^2 + 2\left(r^2 \arcsin\left(\frac{\epsilon}{r}\right) + \epsilon\sqrt{r^2-\epsilon^2} - \epsilon^2 \tan\left(\dfrac{\theta}{2}\right)\right) & \text{if } \epsilon_0 < \epsilon \leq \epsilon_1 \\ \pi r^2 & \text{if } \epsilon > \epsilon_1 \end{cases}$$

where $\epsilon_0 = r\sin(\theta/2)$ and $\epsilon_1 = r\cos(\theta/2)$.

PROOF. The case when $\epsilon > \epsilon_1$ is trivial. The proof therefore splits into two parts:

First part: $0 < \epsilon \leq \epsilon_0$ First let us look at Fig. A.1a and notice that, for $\epsilon < \epsilon_0$, we have

$$|E_N(\epsilon)|_2 = 4(Y + W)$$

where Y is the area of an element with respect to angle α (i.e. the dark grey regions in Fig. A.1a) and W is the one with respect to angle θ (the light grey regions in Fig. A.1a), where $\alpha = \pi - \theta$. Analyzing Fig. A.1b, we get that W is the sum of the

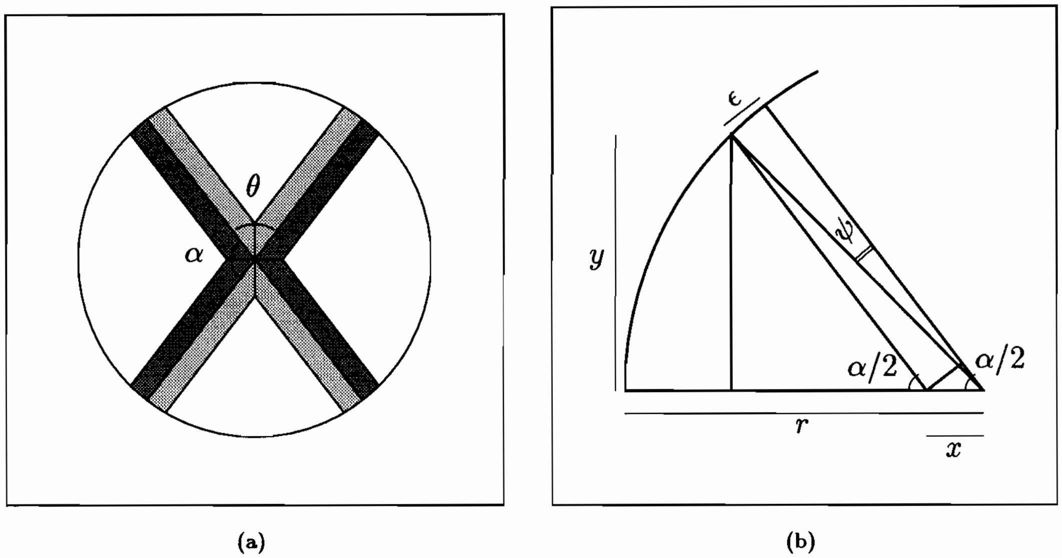


FIGURE A.1. Calculating the area of the dilated crossing within a circle or radius $r = l/2$. In (a) we have the elementary region for the calculation of the area. In (b) a close-up at one of these to calculate ψ.

areas of *(i)* a sector of angle ψ and *(ii)* a triangle:

$$W = \frac{r^2\psi}{2} + \frac{xy}{2} = \frac{r^2\psi}{2} + \frac{r\epsilon\sin(\theta/2-\psi)}{2\sin(\theta/2)}$$

since

$$x = \frac{\epsilon}{\sin(\theta/2)} \quad \text{and} \quad y = r\sin\left(\frac{\theta}{2} - \psi\right).$$

We get the same for Y but this time we replace θ by α and from the relationship between θ and α we finally obtain

$$|E_N(\epsilon)|_2 = 4\left[r^2\psi + \frac{r\epsilon}{2}\left(\frac{\sin(\pi/2-\theta/2-\psi)}{\sin(\pi/2-\theta/2)} + \frac{\sin(\theta/2-\psi)}{\sin(\theta/2)}\right)\right]$$

which after simplification gives us

$$|E_N(\epsilon)|_2 = 4\left[r^2\psi + r\epsilon\left(\cos\psi - \frac{\sin\psi}{\sin\theta}\right)\right].$$

Now we just need to find ψ, which from the sine law is given to be

$$\psi = \arcsin(\epsilon/r)$$

and then substituting we get

$$\begin{aligned}
|E_N(\epsilon)|_2 &= 4\left[r^2\arcsin(\epsilon/r) + r\epsilon\left(\cos(\arcsin(\epsilon/r)) - \frac{\sin(\arcsin(\epsilon/r))}{\sin\theta}\right)\right] \\
&= 4\left[r^2\arcsin(\epsilon/r) + r\epsilon\left(\frac{\sqrt{r^2-\epsilon^2}}{r} - \frac{\epsilon}{r\sin\theta}\right)\right] \\
&= 4\left(r^2\arcsin(\epsilon/r) + \epsilon\sqrt{r^2-\epsilon^2} - \frac{\epsilon^2}{\sin\theta}\right)
\end{aligned}$$

Second part: $\epsilon_0 \leq \epsilon \leq \epsilon_1$ Having $\theta < \alpha$, then

$$|E_N(x,\epsilon)|_2 = \theta r^2 + 4Y$$

the area of the filled sectors and the rest. But from before we know that

$$\begin{aligned}
Y &= \frac{r^2\psi}{2} + \frac{r\epsilon\sin(\alpha/2-\psi)}{2\sin(\alpha/2)} \\
&= \frac{r^2\psi}{2} + \frac{r\epsilon}{2}\left(\frac{\sin(\pi/2-\theta/2-\psi)}{\sin(\pi/2-\theta/2)}\right) \\
&= \frac{r^2\psi}{2} + \frac{r\epsilon}{2}\left(\frac{\cos(\theta/2+\psi)}{\cos(\theta/2)}\right) \\
&= \frac{r^2\psi}{2} + \frac{r\epsilon}{2}\left(\cos\psi - \sin\psi\tan(\theta/2)\right)
\end{aligned}$$

Substituting $\psi = \arcsin(\epsilon/r)$ and simplifying gives the final result:

$$|E_N(x,\epsilon)|_2 = \theta r^2 + 2\left(r^2\arcsin\frac{\epsilon}{r} + \epsilon\sqrt{r^2-\epsilon^2} - \epsilon^2\tan\frac{\theta}{2}\right).$$

□

1.2. For a radial pattern. In much the same way as was done for the set of parallel lines, we can characterize the complexity of a set of radial lines. The result will be similar to what we had if we use the following trick: we will bound the set by a convex polygon Ω. This polygon will define the spatial extent onto which we will do the analysis. This trick will simplify the result, compared with what we had for a circle of radius r.

Proposition A.2. *Let E be a set of $n > 1$ lines of length l intersecting at their midpoints. Let x be the point in $\mathbf{R}^2$ where they intersect and Ω be the polygon obtained* *by joining the perpendiculars to the ends of lines. The lines induce* n *cones with angles* $\alpha_1, \cdots, \alpha_n$*, and* $\alpha_j > 0$*. If the* α_i *are reordered such that*

$$\alpha_{\sigma(1)} \leq \alpha_{\sigma(1)} \leq \cdots \leq \alpha_{\sigma(n)},$$

then

$$C_N(x,\delta) = \begin{cases} 1 + \dfrac{\displaystyle\sum_{j=1}^{n} \cot(\alpha_j/2)}{2nr - \displaystyle\sum_{j=1}^{n} \cot(\alpha_j/2)} & \textit{if } \epsilon < \epsilon_1 \\[2em] \dfrac{2r^2 \displaystyle\sum_{j=1}^{i} \tan(\alpha_{\sigma(j)}/2) + 2\epsilon r(n-i)}{r^2 \displaystyle\sum_{j=1}^{i} \tan(\alpha_{\sigma(j)}/2) + 2\epsilon r(n-i) - \epsilon^2 \displaystyle\sum_{j=i+1}^{n} \cot(\alpha_{\sigma(j)}/2)} & \textit{if } \epsilon_i \leq \epsilon < \epsilon_{i+1} \\[2em] 2 & \textit{if } \epsilon \geq \epsilon_i \end{cases}$$

where $\epsilon_j = r\tan(\alpha_{\sigma(j)}/2)$.

PROOF. Before starting the proof, there are a few things that need to be stated: (i) Ω is a closed, convex polygon circumscribed to the circle of radius r centered on x (the intersection point of the lines); (ii) the intersections occur on the bisectors of the cones. This can be proven easily using elementary geometry.

The normal complexity will be obtained from the rate of growth of the area of the dilated sets. Suppose that ϵ is sufficiently small. Then the area of the dilated set inside one cone is comprised of four elements of same area (see Fig. A.2). Let A be one of the elements; it can be split into a rectangle and a triangle and we have

$$|A|_2 = \epsilon x + \frac{\epsilon y}{2} = \epsilon r - \frac{\epsilon^2 \cot(\alpha_j/2)}{2}$$

since

$$y = \epsilon \cot(\alpha_j/2) \quad \text{and} \quad x = r - y.$$

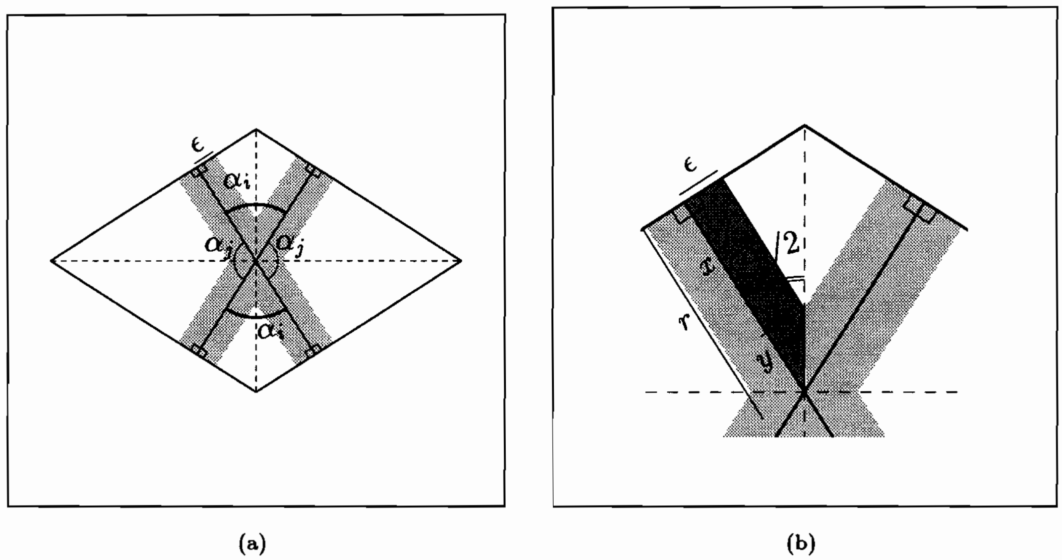


FIGURE A.2. Calculating the area of the dilated radial pattern within a polygon Ω. In (a) we have the bounding diamond shaped polygon for the calculation of the area of the bounded dilated set (in grey). In (b) a close-up to calculate an element of area (in dark grey).

The area of the dilated region inside the cone is therefore $4|A|_2$. Summing over all angles (therefore taking the union of the dilated cone interiors) yields

$$\begin{aligned}
|E_N(x,\epsilon)|_2 &= \sum_{j=1}^{n}(4\epsilon r - 2\epsilon^2 \cot(\alpha_j/2)) \\
&= 2\left(2nr\epsilon - \epsilon^2 \sum_{j=1}^{n} \cot(\alpha_j/2)\right)
\end{aligned}$$

As ϵ gets larger and larger, some of the dilated branches collapse and the growth within them stops. That is where we will use the permutation of the angles into increasing order

$$\alpha_{\sigma(1)} \leq \alpha_{\sigma(1)} \leq \cdots \leq \alpha_{\sigma(n)}.$$

The area of the dilated set for $\epsilon_i \leq \epsilon < \epsilon_{i+1}$ will then be the sum of the area of the "saturated" cones plus the "unsaturated" ones:

$$\begin{aligned}|E_N(x,\epsilon)|_2 &= 4\left[\sum_{j=1}^{i}\frac{\epsilon_j r}{2} + (n-i)\epsilon r - \frac{\epsilon^2}{2}\sum_{j=i+1}^{n}\cot(\alpha_{\sigma(j)}/2)\right] \\ &= 2\left[\sum_{j=1}^{i} r^2\tan(\alpha_{\sigma(j)}/2) + 2r(n-i)\epsilon - \epsilon^2\sum_{j=i+1}^{n}\cot(\alpha_{\sigma(j)}/2)\right]\end{aligned}$$

Putting all the pieces together, we obtain the following:

$$|E_N(x,\epsilon)|_2 = \begin{cases} 2\left(2nr\epsilon - \epsilon^2\sum_{j=1}^{n}\cot(\alpha_j/2)\right) & \text{if } \epsilon < \epsilon_1 \\ 2\left[r^2\sum_{j=1}^{i}\tan(\alpha_{\sigma(j)}/2) + 2r(n-i)\epsilon - \epsilon^2\sum_{j=i+1}^{n}\cot(\alpha_{\sigma(j)}/2)\right] & \text{if } \epsilon_i \leq \epsilon < \epsilon_{i+1} \\ |\Omega|_2 & \text{if } \epsilon \geq \epsilon_i \end{cases}$$

where $|\Omega|_2$ is the area of the bounding polygon. Finally, from a previous lemma (Lemma 4.1), we obtain the desired result. □

1.3. For the Kanizsa pattern. The Kanizsa pattern can be described by a set of variables: n, k, l, $\delta_1, \cdots, \delta_6$. We assume that the pattern is embedded and centered in the unit square. We then denote by

- n: the number of lines for the grating
- $k < (n-2)$: the number of lines inside the grating
- δ_i different key spacings:
  - δ_1: between the borders of the unit square and the first line in the grating
  - δ_2: between the lines in the grating
  - δ_3: between the rectangle and the grating
  - δ_4: between the rectangle and the top (bottom) of the unit square
  - δ_5: between the rectangle and the side of the unit square
  - δ_6: between the two side lines of the rectangle
- l the length of the lines in the grating

Based on this definition of the pattern and on the assumption that $\delta_6/2 > \delta_3$, we get the following equation for the area of the normal dilation:

$$
\begin{aligned}
|E_N(\epsilon)|_2 = \quad & (\epsilon < \delta_1)\ [2l\epsilon, 2l\delta_1] + \\
& (\epsilon < \delta_2/2)\ [2(n-1)l\epsilon, (n-1)l\delta_2] + \\
& (\epsilon < \delta_4)\ [2\delta_6\epsilon, 2\delta_6\delta_4] + \\
& (\epsilon < \delta_5)\ [4\delta_3\epsilon, 4\delta_3\delta_5] + \\
& (\epsilon < \delta_6/2)\ [4\delta_3\epsilon, 4\delta_3\delta_6/2] + \\
& (\epsilon < \delta_3, \epsilon < \delta_6/2)\ [(2\delta_6 + 4\delta_3)\epsilon - 4\epsilon^2, 2\delta_6\delta_3]
\end{aligned}
$$

where here, the expressions in the parentheses denotes a test, and, depending the result of this test, we then take the first (if true) or the second (if false) action between the square brackets. The final result is that we have $|E_N(\epsilon)|_2 = a\epsilon^0 + b\epsilon + c\epsilon^2$ and therefore from Lemma 4.1

$$C_N(x, \delta) = \frac{2a + b\epsilon}{a + b\epsilon + c\epsilon^2}$$

The following algorithm provides us with the normal complexity of the Kanizsa pattern as just described:

```
KanizsaNormalCIndex[l_,n_,d1_,d2_,d3_,d4_,d5_,d6_,e_] :=
        Block[ {a, b, c},
                a = 0; b = 0; c = 0;

                If [e < d1, b+= (2 l), a += (2 l d1)];
                If [e < (d2/2), b += (2 (n-1) l), a += ((n-1) l d2)];
                If [e < d4, b += (2 d6), a += (2 d6 d4)];
                If [e < d5, b += (4 d3), a += (4 d3 d5)];
                If [((e < d3) && (e < (d6/2))),
                        b += ((2 d6) + (4 d3)); c -= 4,
                        a += (2 d6 d3)];
                Return[(2a + b e)/(a + b e + c e^2)];
        ]
```

APPENDIX B

Troubleshooting the algorithm

The algorithm presented in this thesis is divided into several steps, and one can make mistakes at various places. We present here a *check list* and a series experiments that can be performed to make sure that the results are valid. Since it is possible to compute analytically the complexity of some simple patterns, we can use this knowledge to verify the output of the algorithm. This is what we did. Various patterns were used and these are shown in Fig. B.1. The results of the analysis are shown later.

(i) discretization: in x and in y (resolution)
(ii) region setting
(iii) tangent projection (normal and tangential extents)
(iv) scales produced: from region to scale
(v) dilations: both normal and tangential
(vi) regression
(vii) normal complexity on simple patterns
(viii) tangential complexity on simple patterns

For the normal complexity, the test patterns we chose were: (i) n vertical lines aligned with the grid (Fig. B.1a); (ii) a diagonal line (Fig. B.1b); (iii) a pair of diagonal lines crossing in the middle of the unit square (Fig. B.1c); (iv) a pair of lines crossing and aligned with grid (Fig. B.1d); (v) a pair of lines centered horizontally (Fig. B.1e); (vi) a pair of lines off centered (Fig. B.1f). These will capture different aspects of the complexity analysis. We applied the algorithm to estimate $C_N(\cdot, \delta)$ and calculated the exact values from our derivations in Chapter 5. In Table B we list the Mathematica

```
GratingNormalLocalC[l_,seplist_,e_] :=
      Block[ {doubleseparray, doubleseplist, sortedseplist,
              i, j, n, a, b, c, sumb},

      n = Length[seplist];
      Array[doubleseparray,2n-2];
      doubleseparray[1] = First[seplist];
      doubleseparray[2n-2] = Last[seplist];

      For [j=2, j<=n-1, j++,
         doubleseparray[2(j-1)] = seplist[[j]]/2;
         doubleseparray[2(j-1)+1] = seplist[[j]]/2];

      doubleseplist = Table[doubleseparray[i],{i,2n-2}];
      sortedseplist = Sort[doubleseplist];

      i = 0;
      While[((i<2n-2) && (e > sortedseplist[[i+1]])), i++];

      For [j=1; sumb=0, j<=i, j++, sumb += sortedseplist[[j]]];
      a = l sumb;
      b = l (2n - 2  - i);
      c = 0;
      Return[(2a + b e)/(a + b e + c e^2)];
]
```

TABLE B.1. Mathematica routine to calculate the normal complexity of patterns of vertical grating with arbitrary positive spacing.

routine GratingNormalLocalC used to implement the result given in Thm 5.1. This was used as groundtruth for the normal complexity of the vertical lines, centered and shifted pairs of lines.

```
TangentialLocalC[llist_,seplist_,e_] :=
      Block[ {sortedseplist, i, j, n, m,  a, b},
         n = Length[seplist];
         m = Length[llist];
         sortedseplist = Sort[seplist];
         Append[sortedseplist,0];

         i = 0;
         While[((i<n) && (e > sortedseplist[[i+1]])), i++];

         For [j=1; a=0, j<=m, j++, a += llist[[j]]];
         For [j=1,  j<=i, j++,
                  a += sortedseplist[[j]]];
         b =  (n - i);
         Return[a/(a + b e)];
]
```

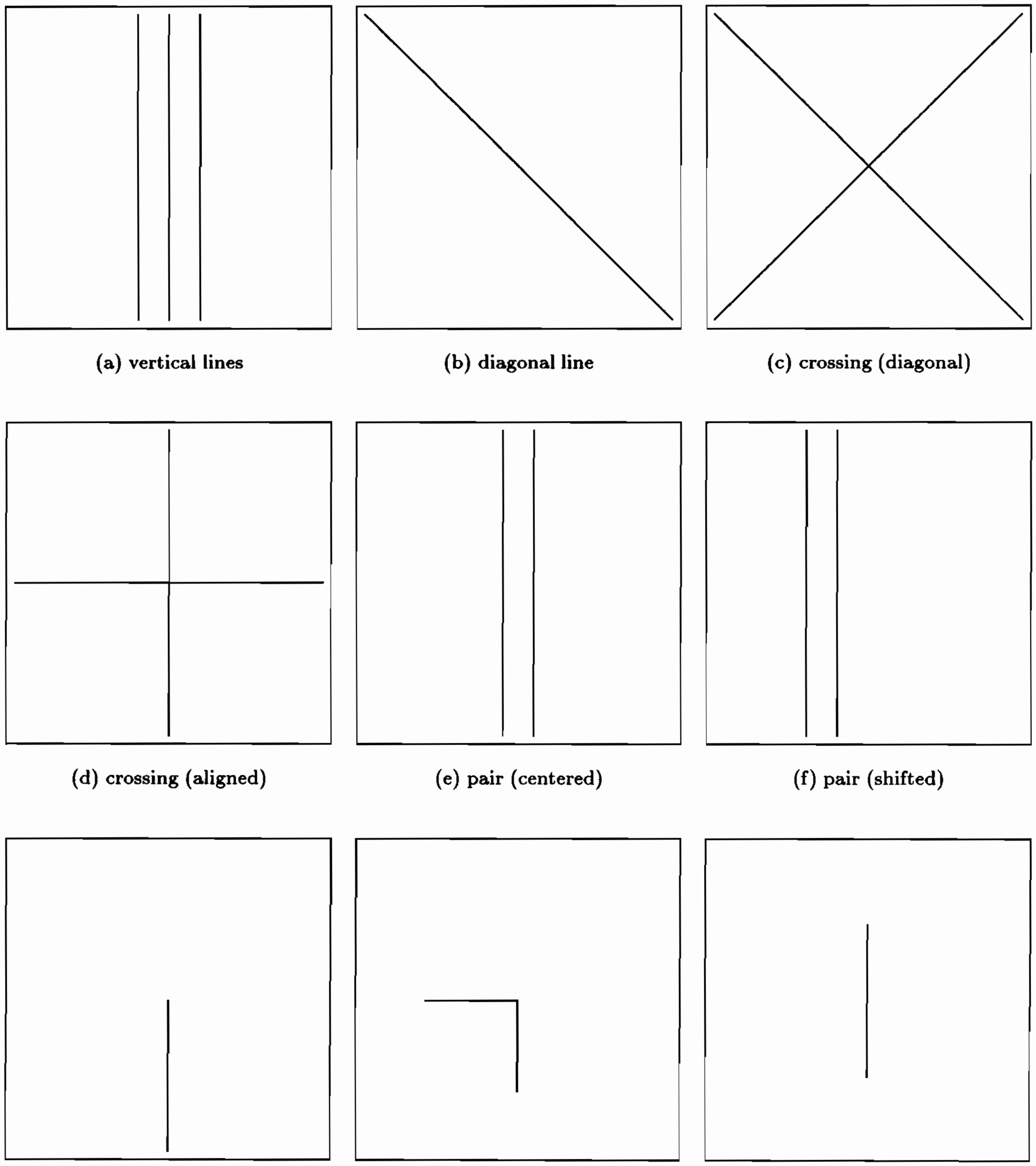
TABLE B.2. Mathematica routine to calculate the tangential complexity of simple sets.

We therefore used this with the estimated values from our algorithm. The results are shown in Fig. B.2 where the solid line gives the analytical result, and the dots are the estimated values. The results clearly agree with the expected behavior. Observe discrepancies at the discontinuities in the complexity values. Taking a tighter neighborhood improves the precision of the estimated values around discontinuities. Satisfied with this result we chose the subresolution to be $p_x = p_y = 10$ for the experiments.

For the tangential complexity, the test patterns chosen were: (i) a line ending (Fig. B.1g); (ii) a corner/wedge (Fig. B.1h); (iii) a line segment (Fig. B.1i).

The tangential complexity of all the patterns was calculated using the Mathematica routine `TangentialLocalC` given in Table B. By providing the appropriate separation list, we could obtain groundtruth result for the tangential complexity of our simple patterns.

Once more the output of our algorithm was compared to the calculated values, and the results agreed beautifully (see Fig. B.3). Optimally, it would have been nice to be able to have groundtruth for curvature, but this amounts to much more complicated calculations, so we left the results out for this analysis.

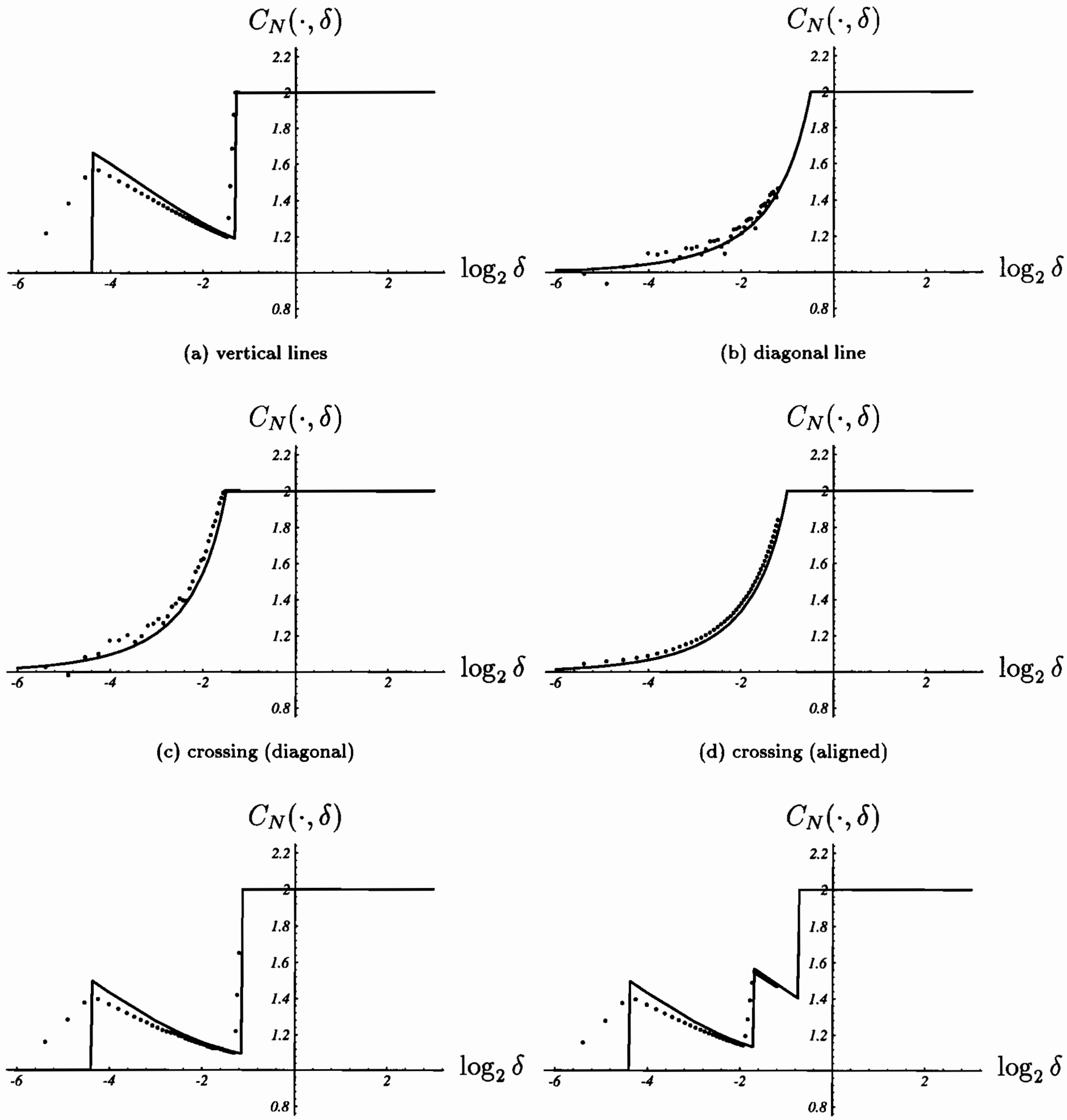


(a) vertical lines (b) diagonal line (c) crossing (diagonal)

(d) crossing (aligned) (e) pair (centered) (f) pair (shifted)

(g) line ending (h) corner (i) segment

FIGURE B.1. Patterns to test the algorithm. (a)-(f) for the normal complexity (g)-(i) for the tangential complexity.



(a) vertical lines

(b) diagonal line

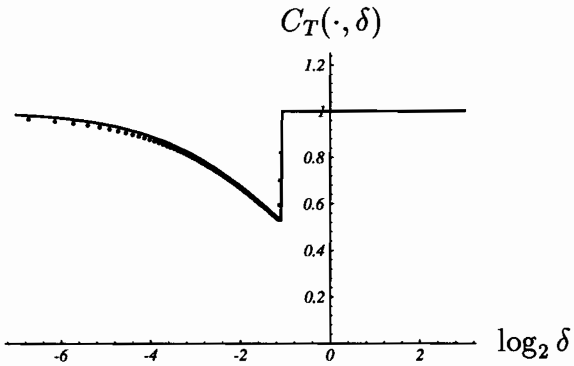
(c) crossing (diagonal)

(d) crossing (aligned)

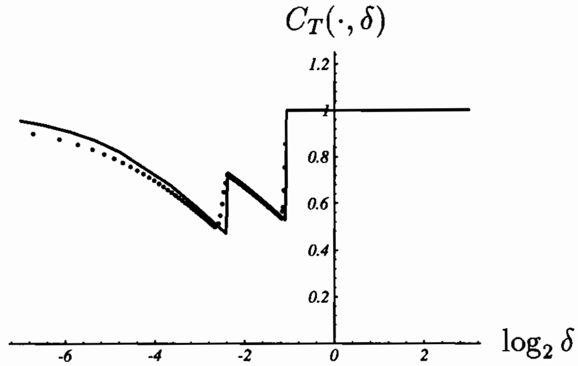
(e) pair (centered)

(f) pair (shifted)

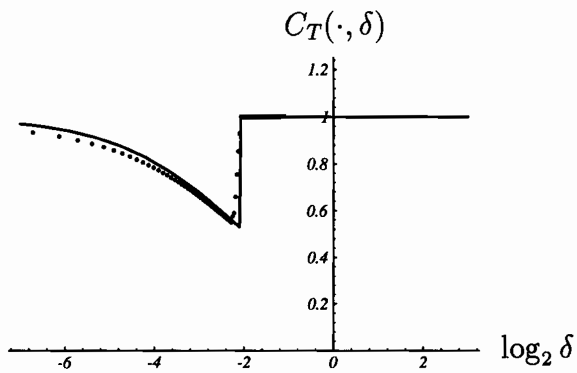
FIGURE B.2. Plotting the normal complexity across all scales for the test patterns. The solid line is the reference. The points are the estimated values.



(a) line end



(b) corner



(c) segment

FIGURE B.3. Plotting the tangential complexity accross all scales for the test patterns. The solid line is the reference. The points are the estimated values.