Chapter 1

Modelling Piecewise Continuity

This is a book about the problem of vision. How is it that a torrent of data from a television camera, or from biological visual receptors, can be reduced to perceptions - the recognition of familiar objects and the concise description of unfamiliar ones? There is of course an immense literature in psychophysics[1], neurophysiology and neuroanatomy that provides some answers in the case of biological systems (see Uttal (1981) for a taxonomy). For instance, the functioning of light-sensitive cells in mammalian vision is understood in some detail (Marks et al. 1964); and the elegant, orderly, spatial correspondence of feature detectors in the brain with the array of cells in the retina, is well known (Hubel and Wiesel 1968). There has also been much dialogue between psychophysics and neurophysiology/neuroanatomy. Examples are the discovery of spatial bandpass channels (Campbell and Robson 1968, Braddick et al. 1978), and understanding the perception of coloured light (Livingstone and Hubel 1984, Jameson and Hurvich 1961) and surface colour (Land 1983, Zeki 1983). These instances are but parts of a very large body of knowledge of biological vision.

Over the last two decades, computers have introduced a new strand into the study of vision. The earliest work (Roberts, 1965) produced systems able to recognise simple objects and manipulate them in a controlled way (Ambler et al. 1975). These systems were, of course, vastly inferior to the biological systems studied by the psychophysicists, neuroanatomists

[1]Psychophysics is the application of physical methods to the study of psychological properties. Visual psychophysics typically probes the mechanisms of human vision by noting a subject's perception of specially designed patterns, under controlled experimental conditions.

and neurophysiologists. They were rather slow, and very brittle. Nonetheless, the availability of computers affects the study of vision in three, very important ways:

1. It provides a rich and precise language in which to express visual problems and processes. Marr distinguishes three levels at which this is done (Marr 1982). At the top "computational theory" level, subtasks are described in terms of their function in processing information. At the next level a subtask, once specified, can be carried out by an appropriately designed "algorithm" - a mathematical recipe. Finally, at the implementation level, any given algorithm might be "realised physically" on any of a great variety of machines, which may be quite dissimilar in their internal architecture, and of vastly differing computing power.

2. Discussion of vision problems can be isolated from design of computing hardware. The beauty of the enriched language for specifying subtasks is that a subtask can be discussed in isolation from the structure of the machine that is to perform it, whether biological or electronic. It can be specified with mathematical precision, and the consequences of the specification can be made inescapably plain by logical predictions. Ullman (1979b), expanding Marr's philosophy (Marr 1976a), puts it like this:

> Underlying the computational theory of visual perception is the notion that the human visual system can be viewed as a symbol-manipulating system. The computation it supports is, at least in part, *the construction of useful descriptions of the visible environment.* An immediate consequence of this view is the distinction that can be drawn between the physical embodiment of the symbols manipulated by the system on the one hand, and the meaning of these symbols on the other. One can study, in other words, the *computation* performed by the system almost independently of the physical *mechanisms* supporting the computation.

Furthermore, task specifications can be tested in practice by executing an algorithm that implements them, on a computer. All this has led to considerable enrichment of studies of human vision (e.g. Marr and Poggio 1979, Mayhew 1982, Hildreth 1984, Ullman 1979b, Koenderinck and van Doorn 1976).

3. Complete, though simple, vision systems can be built and tested. The restriction to study the visual systems that nature has kindly provided is removed. It is possible to construct a system to test a particular issue and to reach a theoretical understanding of the system's behaviour. One of the issues studied in this book is how different ways of modelling the *continuity* of surfaces might affect the stability of their perception as the viewer moves. This is as important to vision by machines as to human vision - it is a generic problem in vision. Furthermore, computer vision systems are now gaining maturity. They appear at last to be approaching widespread practicability in industrial automation and robotics.

This book deals with vision as a computational problem. Little further mention will be made of psychophysics or neurophysiology. But we hope and believe that the new ways of modelling continuity presented here could eventually have a bearing, not only on computer vision, but on biological vision too.

1.1 What is Visual Reconstruction?

Visual Reconstruction will be defined as the process of

reducing visual data to stable descriptions.

"Visual data" comes in various forms, including:

- Raw intensity data direct from photoreceptors, in the form of an array of numbers
- "Optic flow" - measures of velocities of points in an image, obtained perhaps from a suitable spatio-temporal filter (e.g. Buxton and Buxton 1983).
- A depth map, consisting of points embedded, usually sparsely, in the viewer's coordinate-frame. At each point, depth (distance from the viewer) is known. Depth maps may be produced by stereopsis, in which images obtained from two slightly different viewpoints (e.g. two eyes) are compared and matched (Marr and Poggio 1979, Mayhew and Frisby 1981, Baker 1981, Grimson 1981); triangulation is then used to compute the depths. Alternatively depths may be obtained by appropriate processing of optic flow (Bruss 1983) or, artificially, from an optical rangefinder.
- Sets of discrete points making up curves in a 2D image, or in 3D ("space-curves").

In each case, data must be reduced in quantity, with minimal loss of meaningful content, if a concise, symbolic representation is to be attained. It is not enough merely to achieve compression - for example by "run-length encoding", in which an array

$$\{0, 0, 0, 0, 0, 4, 4, 4, 4, 4, 7, 7, 7, 7, 0, 0, 0\}$$

is represented more briefly as

$$\{0 \times 5, 4 \times 5, 7 \times 4, 0 \times 3\}.$$

Rather, in any vision system that is to perform in a consistent manner, it is necessary that the compressed form should be *stable.* This means that it should be *invariant* to (undisturbed by) certain distortions or variations that are likely to be encountered in the image-formation process. These include:

- sampling grain, varying in density due to perspective effects or to inhomogeneity of receptor spacing, as in the eye.
- optical blurring
- optical distortion and sensor noise
- rotation and translation in the image plane
- rotation in 3D (not including, at this point, occlusion effects in which one surface obscures another)
- perspective distortions
- variation in photometric conditions (principally in illumination of the visible scene)

Raw intensity data is affected by all these factors. Ultimately, invariance to all of them must be achieved to produce descriptions of visible surfaces that are, as far as possible, independent of imaging effects. For example, the description of a particular surface patch should not change dramatically if the image is gradually blurred by defocussing; rather it should "degrade gracefully" (Marr 1982). Those blurred snapshots of the baby still look more like a baby than, say, a table. As for 3D rotation invariance, it is required for any system that works in real-time, so that viewed surfaces appear stable as the viewer moves. Less obviously, for analysis of static images, it is still necessary to achieve descriptions that are relatively independent of viewpoint. All the factors mentioned above are relevant to the particular reconstruction processes dealt with in this book.

A prominent theme in following chapters will be *continuity*. In order to reconstruct descriptions that are not only invariant, but also relatively unambiguous, it is necessary to make simplifying assumptions about the world. Assumptions of continuity underlie visual processes of different kinds. Stereopsis is facilitated by constraints on the continuity of surfaces (Marr 1982) and, in particular, by figural continuity - continuity along curves and surface features (Mayhew and Frisby 1981). Analysis of optical flow also appears to require assumptions of continuity, either in regions (Horn and Schunk 1981, Longuet-Higgins 1984) or along curves (Hildreth 1984). Computation of lightness, the perceptual correlate of surface reflectivity (i.e. surface colour), needs constraints on continuity both of the reflectivity itself, and of the incident illumination (Land 1983).

It is clearly unreasonable, in each of these cases, to assume unremitting, global continuity. Depth, optical flow and surface colour all undergo some sudden changes across a scene. It is natural to think of them as continuous in patches. Marr (1982) used the term "continuous almost everywhere". This is not the same as "piecewise continuous" in the mathematical sense, for there is the additional expectation that "the fewer pieces the better". To put it another way, simple descriptions are best, and fewer pieces make simpler descriptions. The challenge, then, is to reach a satisfactory formalisation of "continuity almost everywhere". We do that here by borrowing the idea of a "weak constraint" - a constraint that can be broken occasionally - from Hinton (Hinton 1978). With an appropriate class of continuous surface patches, this leads to "weak continuity constraints" (Blake 1983b) - preferring continuity, but grudgingly allowing occasional discontinuities if that makes for a simpler overall description.

Another important theme emerges later in the book - *cooperativity*. Whereas the "weak continuity constraint" belongs at Marr's "computational theory" level, cooperativity is an algorithmic property. A cooperative process is a computation performed in parallel by a network of independent processing cells. Each cell is connected to just a few of its neighbours, and continually computes some function of its own state, its own input signal, and signals received from its neighbours. The attraction is that rapid computation is achievable, not only by using fast cells, but by using many cells in a large network, all sharing the computational load. The remarkable property of cooperative processes, well known in mathematics and in physical modelling, is this: despite the purely local connectivity of the cells, the network can perform global computations. It is clear that messages could pass between successive neighbours and so propagate across the network. What is more surprising is that propagation can be coordinated, unhindered by collisions between messages, to achieve a useful effect.

Networks of this kind have received much attention in theories of Psychophysics (e.g. Julesz 1971, Marr 1976b) Cognitive Science (e.g. Hinton and Sejnowski 1983, Hopfield 1984), Pattern Recognition (e.g. Rosenfeld et al. 1976) and Computer Science (e.g. Brookes et al. 1984, Milner 1980). In vision, there have been cooperative algorithms for optical flow computation (Horn and Schunk 1981), analysis of shading (Woodham 1977, Ikeuchi and Horn 1981), analysis of motion (Ullman 1979a), computation of lightness (Horn 1974, Blake 1985c) and reconstruction of stereoscopically viewed surfaces (Grimson 1981, Terzopoulos 1983). The implementation of weak continuity constraints can be achieved very naturally too, we shall see, by cooperative networks.

1.2 Continuity and cooperativity

1.2.1 Cooperativity in physical models

Two physical examples will help to provide a more concrete insight into basic properties of cooperative computations.

The first (figure 1.1) is an elastic sheet - a soap film for instance -

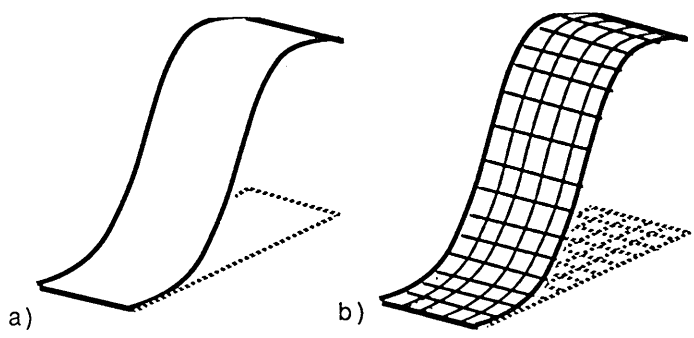

Figure 1.1: A physical example of cooperative computation. A wire frame (a) is covered by an elastic sheet (b). The shape that the sheet assumes can be calculated by an array of locally connected cells.

stretched over a wire frame. The sheet takes up a minimum energy configuration, which happens to be a solution of Laplace's equation[2]. This configuration can be computed in a local-parallel fashion, in what is called

[2]Strictly, the solution of Laplace's equation approximates to the minimum energy configuration.

a "relaxation algorithm". The shape taken up by the sheet is represented by its height at each point on a rectangular grid. Initially some rough estimate of those heights is made. The following local computation is then done, repeatedly, at each grid point: its height value is replaced by the average of the values at the four neighbouring positions. While this is going on, the heights of points on the wire frame itself remain fixed (a "boundary condition" for the cooperative process). Imagine simple computational cells, whose sole function is to accept signals from four neighbours and output their average. They repeat this perpetually. The result is that the influence of the wire frame propagates inwards on the sheet, until finally the sheet comes to rest at its true equilibrium position.

Several general properties of cooperativity are illustrated here:

Propagation - in this case from the boundary to the interior. Propagation can also occur, in certain systems, over shorter ranges, more like pressing on a mattress to produce a dent in the region of the hand. The extent of the dent depends on how elastic the mattress is. Truly global propagation (as on the soap film) occurs in visual processes - the computation of lightness is an example. Propagation over a restricted range (as on the mattress) is what occurs when weak continuity constraints are in force.

Local interaction: cells communicate only with their immediate neighbours.

Parallelism: the cells compute continuously, and independently except for the exchange of signals with neighbours.

Energy minimisation: the relaxation algorithm progressively reduces the elastic energy of the sheet, until equilibrium is reached.

The second physical example is one proposed by Julesz (Julesz 1971) as a model for stereoscopic vision, and is known in physics as an Ising model. Magnetic dipoles arranged on pivots (figure 1.2) interact with one another in such a way that they prefer to align with their neighbours. Springs on the magnets tend to return them to their natural orientations. The angles of the magnets take the place, here, of the heights in the example of figure 1.1. In just the same way, the stable states of the system of magnets can be computed cooperatively. But there is an important difference. There are not one, but many stable states. Whereas the elastic sheet always returns, after a deflection, to the same equilibrium position, the system of magnets can flip from one stable state to another. A stable state will usually consist

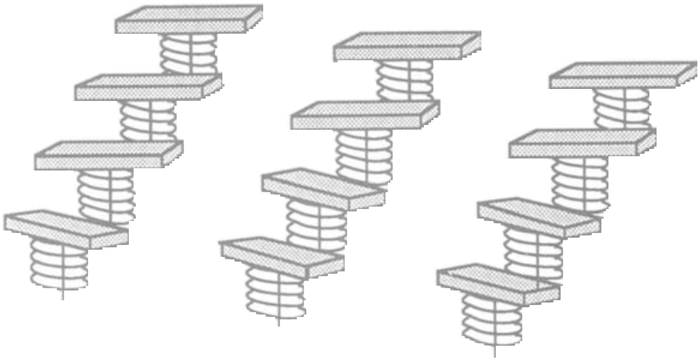

Figure 1.2: A more complex example of cooperativity. Bar magnets, arranged on a grid, tend to align with one another, but are also subject to restoring forces from the springs around their pivots.

of a patchwork of regions ("domains") each containing magnets of similar orientation. Orientation changes abruptly across domain boundaries. The size of the domains is determined by the strength of the magnetic interaction, compared with the strength of the springs: the stronger the magnetic force, the larger the domains tend to be. And all this is very much how a system behaves under weak continuity constraints - regions of continuous variation, with abrupt changes at boundaries.

1.2.2 Regression

Visual reconstruction processes of the sort discussed in this book are founded on least-squares regression. In its simplest form, regression can be used to choose the "best" straight line through a set of points on a graph. More complex curves may be fitted - quadratic, cubic or higher order polynomials. More versatile still are splines (de Boor 1978) - sequences of polynomials joined smoothly together. There is an interesting connection between cubic splines and elastic systems like the sheet in figure 1.1 (Poggio et al. 1984, Terzopoulos 1986). A flexible rod, such as draughtsmen commonly use to draw smooth curves is an elastic system. If it is loaded or clamped at several points, it takes up a shape - its minimum energy configuration - which is in fact a cubic spline[3] (figure 1.3a). Each load-point forms a "knot" in

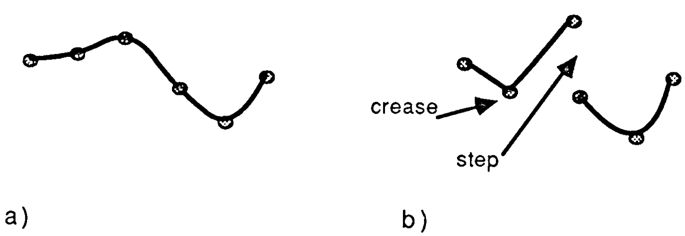

Figure 1.3: A flexible rod, under load, forms a spline. (a) A continuous spline. (b) A spline with crease and step discontinuities, controlled by multiple knots.

the spline, where one cubic polynomial is smoothly joined to the next. So spline fitting can be thought of in terms of minimising an elastic energy - the energy of a flexible rod.

Yet a further generalisation of regression, and the most important one for visual reconstruction, is to include discontinuities in the fitted curve. In spline jargon, these are "multiple knots", generating kinks ("crease" discontinuities) or cutting the curve altogether ("step" discontinuities) as in figure 1.3b. Incorporation of multiple knots, *if* it is known exactly *where* along the curve the discontinuities are, is standard spline technology. A more interesting problem is one in which the positions of discontinuities are not known in advance. In that case, positioning of multiple knots must be done automatically. An algorithm to do that might consist of constructing an initial spline fit, and then adding knots until the regression error measure reached an acceptably small value (Plass and Stone 1983). This would ensure a spline that closely fitted the data points.

For visual reconstruction that is not enough. The requirement for *stability* has already been discussed, which means that the multiple knots must occur in "canonical" (natural) positions, robust to small perturbations of the data and to the distortions and transformation listed earlier. Only then are they truly and reliably descriptive of the data.

The stability requirement is met by imposing weak continuity constraints on an elastic system like the rod. Leaving the details to later chapters, it is sufficient for now to draw on the magnetic dipole system as an analogy. Typically, it has many locally stable states with groups of dipoles of various sizes, aligned in various directions. Among these states,

[3]Again, this is an approximation.

there is a ground state, the state of lowest energy. As energy is reduced, the system is liable to stick in a locally stable state, before the ground state is reached. Similarly, an elastic material under weak continuity constraints has a ground state - its favourite configuration - which is usually very stable. A spline, for example, may have its knots arranged so as to reach its ground state, forming (by definition!) the best, stable description of the data.

Finding the ground state is a problem. Procedures for direct improvement of knot positions (Pavlidis 1977) are prone to be caught in a state other than the ground-state. But provided the system can be jostled or drawn into the ground state, the positions of discontinuities will be stable in the required manner. And this is precisely what is achieved by certain statistical algorithms (Kirkpatrick et al. 1982, Geman and Geman 1984), and the deterministic "Graduated Non-convexity" (GNC) algorithm, proposed in this book. Some examples of the operation of the GNC algorithm, reconstructing various kinds of visual data, are shown in the next chapter. A definition of the algorithm itself, however, must be delayed until chapter 3.

1.2.3 Cooperative networks that make decisions

Visual reconstruction must be more than linear filtering if it is to generate usable features for subsequent visual processes. At some point there must be an element of commitment; decisions must be made - either a feature is present or it is not. In particular, in visual reconstruction, it is necessary repeatedly to decide whether or not a discontinuity is present in a particular location. An example should clarify the distinction between mere linear filtering and feature detection. Consider the task of locating a thin, bright bar in an image. A suitable linear filter could be found which transforms an image into a new image, in which such bars, or their edges, stand out even more brightly. This is not enough. A vision system must make a decision at some point - either there is a bar (in a certain location) or there isn't. Rather than being simply "enhanced", bars must be "labelled". An elegant example due to Poggio and Reichardt (1976) illustrated a similar point. They showed that even so simple a function as detecting the direction of local motion cannot be achieved by any linear system[4].

So purely linear systems are inadequate for visual reconstruction. There must be some non-linearity, even if it is just a thresholding operation. This is what occurs in the Perceptron (Rosenblatt 1962, Minsky and Papert 1969), a simple, neuron-like switching element that computes a weighted

[4]A linear system is one that simply outputs a weighted sum of its inputs.

sum of its inputs, and produces the output 1 or 0, according to whether the sum exceeds some threshold. Similarly, in the computation of lightness (Land 1983), thresholding (used to detect edges in the conventional manner) is an adequate form of non-linearity.

Generally, *any* network that makes decisions cannot be entirely linear. Suppose the network acts to minimise an energy $F(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ is a vector of inputs to the network, and $\mathbf{y}$ is the vector of outputs. Then the output is defined (not necessarily uniquely) as that vector $\mathbf{y}$ which minimises $F(\mathbf{x}, \mathbf{y})$ - for a given, fixed $\mathbf{x}$. If F were a quadratic polynomial in the variables $\mathbf{x}$, $\mathbf{y}$ then $\mathbf{y}$ would be a linear function of $\mathbf{x}$ - the solution of the linear system

$$\partial F / \partial \mathbf{y} = 0.$$

It has already been said that a linear system cannot make decisions[5]. In fact it cannot make decisions as long as F is both "strictly convex" and smooth (differentiable in the variables $\mathbf{x}$, $\mathbf{y}$). In that case, the minimum always exists, and every input/output pair $\mathbf{x}$, $\mathbf{y}$ is a "Morse point" of the function F (Poston and Stewart 1978) which means that $\mathbf{y}$ varies continuously with $\mathbf{x}$. There is no discontinuous or sudden or catastrophic switching behaviour.

We know now that the energy function F for any system under weak continuity constraints must be either undifferentiable or non-convex or both. This is illustrated in figure 1.4.

1.2.4 Local interaction in models of continuity

Geman and Geman (1984) have forged an elegant link, via statistical mechanics, between mechanical systems like the soap film or splines, and probability theory. They have shown, in effect, that signal estimation by least squares fitting of splines is exactly the right way to behave if you have certain *a priori* probabilistic beliefs about the world in which the signal originated. Specifically, the beliefs are: that the signal - the one that is being estimated - is sampled from a "Markov Random Field" (MRF) and that Gaussian noise was added, in the process of generating the data.

What exactly is an MRF? It is a probabilistic process in which all interaction is local; the probability that a cell is in a given state is entirely determined by probabilities for states of neighbouring cells. An example based on one given by Besag (1974) illustrates this. Imagine a field full of cabbages, planted by a very methodical farmer on a precise, square grid. (A hexagonal grid would, of course, have given better packing density, but his ageing tractor runs best in straight lines.) Unfortunately, an outbreak

[5]This assumes that the system is unconstrained.

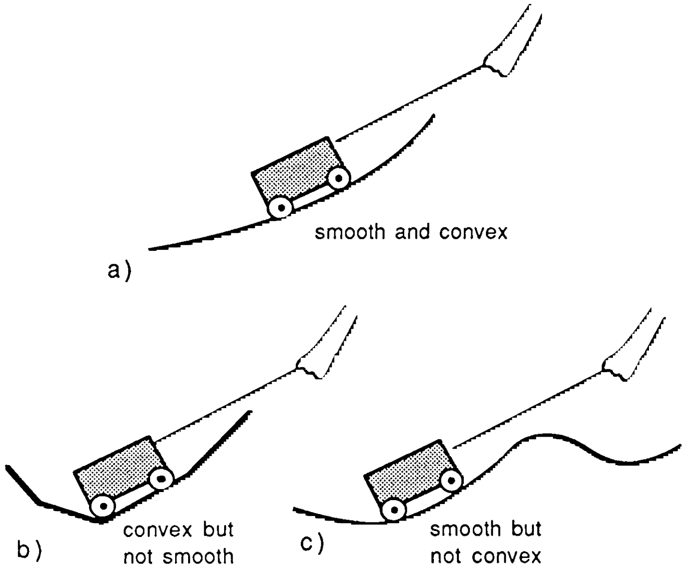

Figure 1.4: A smooth, convex energy function cannot cause discontinuous behaviour. The cart's position is a continuous function of the hand's position in (a), but jerky motion occurs in (b) and (c).

of CMV (Cabbage Mosaic Virus), which is particularly virulent when cabbages are arranged in a regular tesselation, has afflicted his crop. At a certain stage in the progress of the disease, its spread can be characterised as follows. The probability that any given cabbage has the disease depends entirely on the probability of disease of its four immediate neighbours. This is because the disease passes, with a certain probability, from neighbour to neighbour.

Qualitatively, the spread of the disease has much in common with the soap film example given earlier. In both cases, *direct* interaction occurs only between immediate neighbours. But global effects can still occur as a result of propagation. Just as the position of the wire frame influences the position of the interior of the soap film, so the introduction of disease at the edge of the field can spread, from neighbour to neighbour, towards the middle.

Formally, what Geman and Geman show is that elastic systems can also be considered from a probabilistic point of view. The link between spline energy E and probability Π is that

$$\Pi \propto e^{-E/T} \tag{1.1}$$

(T is a constant). The lower the energy of a particular signal (that was generated by a particular MRF), the more likely it is to occur. Highly deformed elastic sheets have high energy and are intrinsically "unlikely" to occur. What is more, weak continuity constraints can also be understood in probabilistic terms: they are consistent with the belief that there is a "line-process", also an MRF but not directly observable in the data, determining the positions of discontinuities.

It comes as something of a shock, when happily using splines as a very natural, mechanical model for smooth, physical surfaces, to find that this is inescapably equivalent to making certain probabilistic assumptions! The most disturbing thing is that one is *forced* to accept that the surface model is a probabilistic one, and therefore includes an element of randomness. This may be appropriate for modelling texture (Derin and Cole 1986), but in a model of smooth surfaces it has rather counter-intuitive consequences, illustrated in figure 1.5. A "1st order" MRF[6], for instance, ranks a noisy but horizontal plane more probable than a smooth inclined one. This is because the 1st order MRF is sensitive only to gradients. Later in the book, this "gradient limit" problem is discussed in some detail. It can be cured by moving to 2nd order, but then it just recurs in a different form, as

[6] 1st order, here, means that direct interaction occurs only between immediate neighbours; 2nd order means that there is direct interaction between neighbours separated by 2 steps.

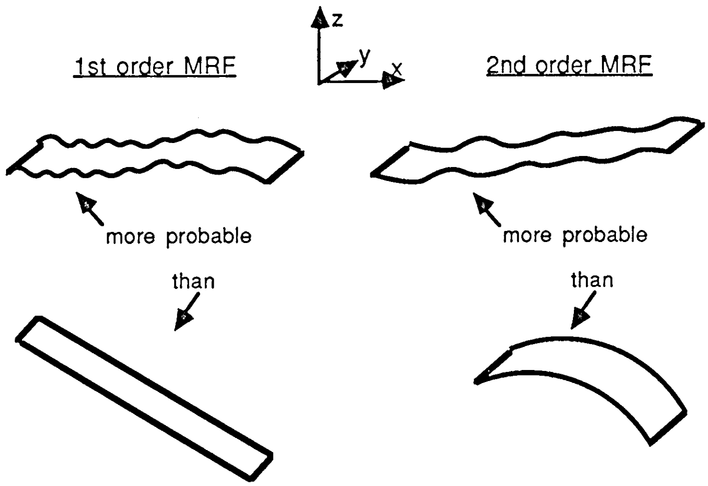

Figure 1.5: MRF models of surfaces can be somewhat counter-intuitive. A smooth but inclined or curved surface may have a *lower* MRF probability than a rough, noisy one.

in the figure. It is not clear what order of MRF would be sufficiently high to avoid the problem, if any. In any case, the higher the order, the greater the range of interaction between cells, and the more intractable the problem of signal estimation becomes. In practice, anything above 1st order is more or less computationally infeasible, as later chapters will show. What the probabilistic viewpoint makes quite clear, therefore, is that a spline under weak continuity constraints (or the equivalent MRF) is not quite the right model. But it is the best that is available at the moment.

As for choosing between mechanical and probabilistic points of analogies, we are of the opinion that the mechanical one is the more natural for representation of *a priori* knowledge about visible surfaces, or about distributions of visual quantities such as intensity, reflectance, optic flow and curve orientation. The justification of this claim must be left, however, until the concluding chapter. In the meantime, this book pursues Visual Reconstruction from the mechanical viewpoint.

1.3 Organisation of the book

Throughout the book, even in later chapters which are more technical, our aim has been to avoid obstructing the text with undue mathematical detail. Longer mathematical arguments are delayed until the appendix.

Chapter 2. Examples are given, with copious illustrations, of the applications of weak continuity constraints in Visual Reconstruction. Problems discussed include edge detection (analysis of variations in image intensity), stereoscopic vision, passive rangefinding and describing curves. This chapter is free of mathematical discussion; it should be easily accessible to most readers.

Chapter 3. The simplest possible discontinuity detection scheme is described - detecting step discontinuities in 1D data, using a "weak string". The idea of a weak continuity constraint is expanded. A simple algorithm, using Graduated Non-convexity (GNC), is described. There is some mathematics in this chapter, but nothing too difficult.

Chapter 4. The theoretical properties of the weak string and its 2D analogue, the "weak membrane", are discussed in some detail. Application of variational calculus enables exact solutions to be obtained for certain data - for example step edges and ramps. These solutions, in turn, enable the two parameters in the weak string/membrane energy to be interpreted. Far from being arbitrary, in need of unprincipled tweaking, they have clear roles in determining scale, sensitivity and resistance to noise. Moreover, it is shown that, under weak continuity constraints, the positions of discontinuities are localised with impressive accuracy. In 2D, the geometry and topology (connectivity) of discontinuities is faithfully preserved - something that, it seems, cannot be achieved by more conventional means.

Chapter 5. The "weak rod" and the "weak plate" are even more powerful means of detecting discontinuities. ("Creases" can be detected, as well as "steps".) Analytical results can again be obtained for certain cases and, as before, lead to an interpretation of parameters in the energy.

Chapter 6. So far, energy minimisation has been treated as a variational problem. For computational purposes it must be made discrete. This is done using "finite elements", together with "line-variables" to handle discontinuities. Existing minimisation algorithms are reviewed.

Chapter 7. The effectiveness of the GNC algorithm is explained. For a substantial class of signals (step discontinuities in noise), it is shown that GNC produces precisely the correct optimal solution.

Some details of designing GNC algorithms for weak membrane and plate are given. In particular, it is necessary to approximate a non-convex energy by a convex function. We explain how this is done. Both serial and parallel algorithms are dealt with, together with a full discussion of convergence properties.

Chapter 8. Some conclusions and open questions.

Appendix. The appendix contains a substantial body of work supporting, in particular, variational analysis (appendix A,B,C) and analysis of the GNC algorithm (appendix D,E).