Acta Numerica (1994), *pp.* 1–59

Formalization and computational aspects of image analysis

Luis Alvarez
Departamento de Informatica y Sistemas
Universidad de Las Palmas
Campus de Tafira, 35017 Las Palmas, Spain

Jean Michel Morel
C.E.R.E.M.A.D.E.
Université Paris IX Dauphine
75775 Paris cedex 16, France

In this article we shall present a unified and axiomatized view of several theories and algorithms of image multiscale analysis (and low level vision) which have been developed in the past twenty years. We shall show that under reasonable invariance and assumptions, all image (and shape) analyses can be reduced to a single partial differential equation. In the same way, movie analysis leads to a single parabolic differential equation. We discuss some applications to image segmentation and movie restoration. The experiments show how accurate and invariant the numerical schemes must be and we compare several (old and new) algorithms by discussing how well they match the axiomatic invariance requirements.

CONTENTS

1 Introduction 2
2 Image multiscale analysis 6
3 Axiomatization of image multiscale analyses and classification of the main models 12
4 Shape multiscale analyses 16
5 Relation between image and shape multiscale analyses 20
6 Multiscale segmentation 22
7 An example: texture discrimination 26
8 Movies multiscale analysis 30

9 Invariance and stability requirements for numerical schemes of the fundamental equation. 32
10 Finite difference schemes for the AMSS model 36
11 Morphological (set evolution) schemes 42
12 Conclusions 50
Appendix A. The 'fundamental theorem' of image analysis 50
Appendix B. Proof of the scale normalization lemma 51
Appendix C. Classification of shape multiscale analyses 53
References 55

1. Introduction

1.1. What will be done, what not, and why?

Before starting with what will be the main object of this survey – image multiscale analysis – we intend to give a very brief account of what image processing is and the choices we have made about what should be developed here and what should be omitted. Image processing may be viewed as a long list of techniques for capturing, transmitting, and extracting information from digital images, in close relation with what is assumed to be relevant to human perception. Here is, accordingly, the list of subjects treated in a classical manual of image processing: *Visual perception, Digitization, Compression, Enhancement, Restoration, Reconstruction, Matching, Segmentation, (Semantic) Representation* (Rosenfeld and Kak, 1982). This defines image processing as a somewhat abstract theory. Other manuals focus on practical applications to perception-based control of robots (Horn, 1986) vision theory (Marr, 1982), while many other monographs treat a single technological application: radar vision, microscopy, satellite imaging, compression standards, character recognition, etc. These involve specific mathematical techniques which will not be presented here. Indeed, whenever some *a priori* (statistical, structural) knowledge about the processed image is at hand, the process must be adapted accordingly and (for instance) the use of stochastic filtering techniques, is justified, but specific.

The IEEE monographs and proceedings give a good account of what is being done in image processing and one can get a rather complete view of the image analysis subject by reading the proceedings of the biannual ICCV (International Conference on Computer Vision). Now, whatever the envisaged applications are, the nine items from Kak and Rosenfeld are a reliable common denominator. They represent what everybody should know before starting any application dealing with images or a formalized theory of vision. From the nine terms quoted above, only the second (digitization)

and the last (semantic representation) fall outside our field because the first relates to the engineering of captors and the last to artificial intelligence and structured programming. Starting with the mathematical classification of the subjects, let us say that compression and reconstruction rely on sharp mathematical techniques related to harmonic analysis. Indeed, the main step of compression–reconstruction devices is a decomposition of the image on a well chosen functional orthonormal basis which can be classical (Fourier, Haar, Hadamard) or new: Wavelets (Meyer, Mallat, Daubechies, etc.), Wavelet packages, etc. All these theories, from the mathematical as well as from the numerical viewpoint, are well explained in several recent books (see e.g. Meyer (1992)), accessible to both mathematicians and engineers, and we simply choose not to present them here. However, there is another reason for this: we must distinguish between techniques for *analysing* images (which therefore strongly rely on the geometry of images) and techniques for *storing* them, where there is no need for geometrical invariance in the numerical representation. When one wants to compress, 'tous les coups sont permis'. In addition to this intuitive difference there is a corresponding strong difference in mathematical techniques. As we shall see, image-analysis geometry-preserving techniques must be fully nonlinear (probably one of the first to understand this and draw the mathematical consequences was Matheron (1975)). Therefore, from the initial nine subjects from the Kak–Rosenfeld classification we shall keep only four: *visual perception, enhancement, restoration and segmentation.* (We have also omitted matching because this is a secondary task, only effectuated after some of the four preceding processes have been applied.) Because the four mentioned subjects obey the same geometric requirements, we shall see that they can be treated by a common theory which we shall call *multiscale analysis.* We shall therefore not treat them as primary subjects and they will simply appear as natural consequences or applications of the theory of geometric multiscale analysis.

1.2. Geometric multiscale analysis

This section is devoted to a short overview of what will be covered.

A numerical image can be modelled as a real function $u_0(x)$ defined in $\mathbb{R}^N$ (in practice, $N = 2$ or 3). The main concept of vision theory and image analysis is *multiscale analysis* (or '*scale space*'). Multiscale analysis associates with $u(0) = u_0$ a sequence of simplified (smoothed) images $u(t, x)$ which depend upon an abstract parameter $t > 0$, the *scale.* The image $u(t, x)$ is called *analysis of the image* u_0 *at scale* t. The formalization of *scale space* has received much attention in the past ten years; more than a dozen of theories for image, shape or 'texture' multiscale analysis have been proposed and recent mathematical work has permitted a formalization

of the whole field. We shall see that several formal principles (or axioms) are sufficient to characterize and unify these theories and algorithms and to show that some of them simply are equivalent. These principles are *causality* (a concept in vision theory which can be led back to a maximum principle), *Euclidean (and/or affine) invariance*, which means that image analysis does not depend upon the distance and orientation in space of the analysed image, and *morphological invariance* which means that image analysis does not depend upon a contrast change.

The characterization and classification of the numerous theories of image and shape analysis will be obtained by identifying the underlying partial differential equations (which have been more or less implicit in many theories!). The axiomatic characterization leads, as we shall see, to a significant improvement in most proposed algorithms as well as to new ones with more invariance properties. Among the theories which will be axiomatically or numerically tested here, we shall mention

- the *Raw Primal Sketch* by Hildreth and Marr,
- the *Scale Space* by Witkin, Koenderink, etc.,
- the *Intrinsic Heat Equation* by Gage, Hamilton, Grayson, Angenent, etc.,
- the *Motion by Mean Curvature* (Osher, Sethian, Evans, Spruck, Giga, Goto, Barles, Souganidis, etc.),
- the *Entropy Scale Space* by Kimia, Tannenbaum and Zucker,
- the *Texton* theory by Julesz,
- the *Dynamic Shape* by Koenderink and Van Doorn,
- the *Curvature Primal Sketch* by Mackworth and Mocktarian, Asada and Brady, etc.,
- the *Morphologie Mathématique* by Matheron, Serra and the 'Fontainebleau school',
- the *Anisotropic Diffusion* by Perona and Malik,
- the *Affine Scale Space of Curves* by Sapiro and Tannenbaum,
- the *Affine Morphological Scale Space* of images by Alvarez, Guichard, Lions and Morel,
- the *Affine Morphological Galilean Scale Space* of movies by the same authors.

The classification of these multiscale theories will lead us to focus on the only one of them which simultaneously matches all invariance and stability requirements partially satisfied by the others: the Affine Morphological Scale Space (AMSS). This multiscale analysis can be defined by a simple Partial Differential Equation (PDE),

$$\frac{\partial u}{\partial t} = |\mathrm{D}u|(t \cdot \mathrm{div}(\mathrm{D}u/|\mathrm{D}u|))^{1/3}, \quad u(0,x) = u_0(x), \tag{1.1}$$

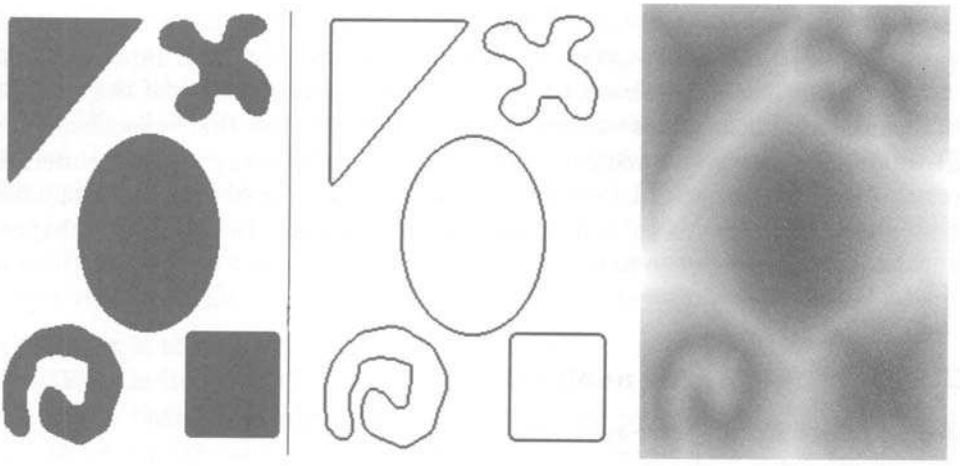

Fig. 1. Three different representation of a 200×300 pixels binary image (from left to right): (a) the classical discrete representation, (b) the level curve representation, (c) the distance function to the level curve. See Section 9 for more details.

where $u(t,x)$ denotes the image analysed at scale t and point x. (This parabolic equation admits a unique 'viscosity solution' in the Crandall–Ishii–Lions (1991) sense.) As we shall see, the equation of the AMSS handles independently all *level sets* of the analysed image and is therefore compatible with the Morphologie Mathématique (which asks for contrast invariance). In addition, contrast invariance means that the boundary of every level set of the image is analysed as a shape and we get a common multiscale analysis for shapes and images.

A multiscale formalization of *image segmentation* can be developed with analogous principles and leads to multiscale segmentation algorithms. Multiscale segmentation associates an initial image u_0 with a sequence $u(t,x)$, $K(t,x)$, where $u(t,x)$ is the image simplified at scale t and $K(t,x)$ the set of boundaries of the homogeneous regions of u_0 at scale t. Thanks to the formalization, many segmentation algorithms can be reduced to one.

As a first application of the axiomatic method, we shall show how both scale space and image segmentation theories lead to texture segmentation algorithms as well as to a rigorous discussion of Julesz' axiomatic theory of texture discrimination. The experimental result of this discussion is unexpected.

We finally devote some pages to the above mentioned Affine Morphological Galilean analysis of movies, with Guichard's (1993) remarkable experimental results in movie denoising. The underlying equation,

$$\frac{\partial u}{\partial t} = (|\nabla u| \operatorname{curv}^{1/3}(u))^{1-q} \left((|\nabla u| \operatorname{sgn}(\operatorname{curv}(u)) \operatorname{accel}(u))^{+}\right)^{q}, \quad (1.2)$$

is also a parabolic equation, where curv denotes the nonlinear differential operator computing the curvature of the level lines, and accel represents the 'apparent acceleration' observed at a given space–time point of the movie.

Roughly speaking, this survey has two parts: in the first, (Sections 1 to 8) we develop the above mentioned theories and give comparative numerical results on test images. The second is devoted to a long discussion of old and new algorithms for Mean Curvature Motion (applied to images or shapes) and the AMSS model (equation (1.1)).

2. Image multiscale analysis

2.1. A short story of the subject

Computer vision deals with a philosophical, psychological, physiological and technical question which can be stated in a few words: how can the local brightness information arriving at the retina of some individual (or any optical sensor) be transformed into a global percept of the objects surrounding him, including their distance, colour and shape? In the 1960s, this question was translated into a very practical framework with the new possibilities for experimentation offered by digital pictures with computers. The new technology has enabled accurate measurements of human visual performance on digital pictures and the first experiments in 'computer vision'. The joint developments in pschophysics and computer vision have led to a new doctrine the existence of *low level vision*. The story of the doctrine is well explained in David Marr's book *Vision* and we shall just give a few hints of how this doctrine developed.

On the other hand, several psychophysical experiments due to Bela Julesz and his school proved that the reconstruction of the spatial environment from binocular information was an automated, reflex process, independent of any learning. Julesz also studied the 'preattentive' perception of textures and proved the existence of a process for discriminating textures independently of any *a priori* knowledge. The discrimination process is fast, parallel and Julesz and his school discussed it in *mathematical terms* from statistics and geometry. These experiments, as well as the neurobiological experiments of Hubel and Wiesel, gave proof of the existence, in the first milliseconds of the perception process, of a series of parallel, fast and irreversible operations applied to the retina information and already yielding very rich and useful information to further understanding of the 'image'.

2.2. The visual pyramid as an algorithm

We shall call this series of operations the 'visual pyramid'. In mathematical terms, it may be thought of as an algorithm but not in the Turing sense; it

the more general sense where we define an algorithm as a black box transforming its input into an output in a deterministic way by a physical process: in any case, this machine is assumed to be physically implemented in the brain. We must distinguish the problem of how this machine works in the case of the brain and what it really does as an information processor (what Chomsky called performance versus competence). Indeed, the second question is simply a mathematical question, while the first one is very relevant in neurobiology. We shall now focus on the mathematical question and treat it in a rather rough way by answering the three questions:

(a) What is the input of the visual pyramid?
(b) What is its output?
(c) What basic principles must obey the visual pyramid if it is considered to be a physical system?

2.3. What is the input of the visual pyramid?

A simple model to discuss image processing is to define an 'image' as a 'brightness' function $u_0(x)$ at each point x of a domain of the plane. This domain, which may be the plane itself, is a model of the retina or any other photosensitive surface. In what follows, we shall take the plane for simplicity. For commodity, in the discussion, we shall always assume that $u_0(x)$ is in the space $\mathcal{F}$ of all continuous real functions $u(x)$ on $\mathbb{R}^N$ such that $\|(1+|x|)^{-N}u(x)\| \leq C$ for some N and C. Of course, the datum of $u_0(x)$ is not absolute in perception theory, but can be considered as the element of an equivalence class. If y is a vector of the plane, the shifted datum $u_0(x-y)$, which is the image shifted by y, is an equivalent datum. In the same way, the change of $u_0(x)$ into $u_0(Rx)$ where R is an isometry of the plane should not change the visual analysis. Finally, we can think of u_0 as belonging to a projective class, that is, as a representative of the class $u_0(Ax)$ where A is any projective map of the plane. Indeed, a plane image can be viewed by an observer from any distance and orientation in space (think of a painting in a gallery). Therefore, the input $u_0(x)$ is assumed to be equivalent to any of its anamorphoses $u_0(Ax)$. We shall assume, in the following, A to be only any affine map, which makes sense when the ratio between the size of the observed objects and the distance to the sensor is small. Last but not least, the observation of $u_0(x)$ does not generally give any reliable information on the number of photons sent by any visible place to the optical sensor. Therefore, the equivalence class in consideration will be $g(u_0(x))$, where g stands for any (unknown) contrast function depending on the sensor. This last assumption, that only isophotes matter, is associated with the 'mathematical morphology' school. So we shall call it the 'morphological' assumption.

To summarize, an image is an equivalence class of functions $u_0(Ax)$ where

A is a translation or an isometry in most classical geometrical models, A is any affine map in the simplified projective model and it is a class of functions $g(u_0(x))$ where g is any continuous nondecreasing function in the morphological model. We can combine these models and consider a morphological projective model, that is, an equivalence class under the action of all gs and As: $g(u_0(Ax))$.

2.4. What is the output *of the visual pyramid?*

Starting from the local brightness information, each layer of the visual pyramid is assumed to yield more and more global 'low level' information about the image. This information is assumed to be usable for the geometric reconstruction (stereovision) as well as for 'high level vision', that is, the interpretation of the scene. Whatever its use might be, most models define the basic output as either

- a smoothed image (from which reliable 'features' can be extracted by local and therefore differential operators); or
- a segmentation, that is, either a decomposition of the image domain into homogeneous regions ('strong segmentation'), with boundaries or a set of boundary points or 'edge map'.

In both cases, the output depends on two variables: a variable x which denotes the centre of a spatial neighbourhood and a variable t which can be identified or correlated with

- the 'height' in the visual pyramid (or distance from the first layer: the retina). This distance corresponds to the biological time between the 'arrival' at the retina and the first arrival at a given layer; and
- the degree of globality of the local information in the considered layer, that is, the size of the neighbourhood in the retina which influences what happens at x.

To summarize, the output of the vision pyramid is:

- either a multiscale image $u(t, x)$; or
- or a multiscale 'edge map' $K(t, x)$, where t is a parameter which can be identified with a time of analysis or with a measure of the spatial globality of the information provided by $u(t, x)$. The bottom of the output is the original image $u(0, x) = u_0(x)$.

2.5. What basic principles must the visual pyramid obey?

The causality These principles come first from the 'preattentive' assumption that no feedback is allowed, the visual information being processed in parallel through a sequence of filters. This means that what happens at

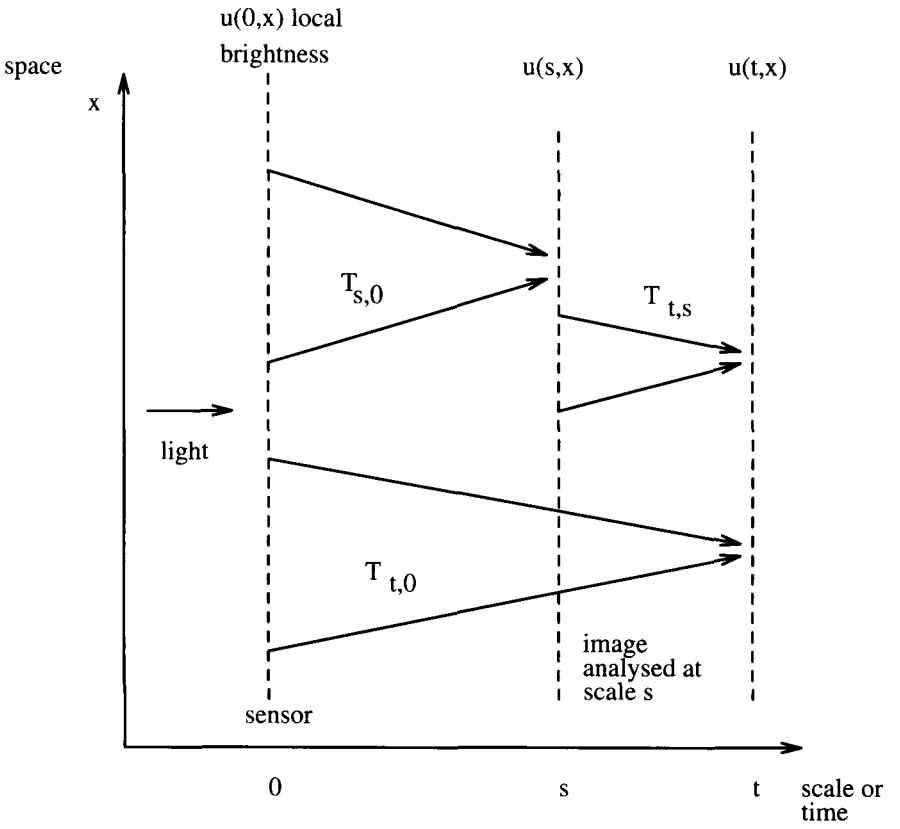

Fig. 2. The scale-space visual pyramid.

higher scales cannot influence what happens at lower scales: the pyramid acts 'from fine to coarse'. Furthermore, there is no time for taking into account at scale t what happens at a significantly smaller scale s. So we assume that the output at scale t can be computed from the output at a scale $t-h$ for very small h. (Take into account the fact that the visual pyramid is a series of filters through which new visual information is constantly being processed. So, to look at what happens at a smaller scale means to 'look into the future', at newly arriving perceptual images.) To formalize this relation, we call $T_t : \mathcal{F} \to \mathcal{F}$ the map which associates an image u_0 with its 'smoothed image' at the scale t, $T_t u_0$. (In the same way, we denote by T_t the map associating a set K of boundaries with a set of boundaries $T_t K$ simplified at scale t.) This mapping is obtained by constructing 'transition filters' which we call $T_{t+h,t} : \mathcal{F} \to \mathcal{F}$ and hence we have the

Pyramidal Structure (Causality 1) $T_{t+h} = T_{t+h,t} T_t$, $T_0 = \mathrm{Id}$. Furthermore, the operator $T_{t+h,t}$ will always be assumed to act 'locally', that is, to

look at a small part of the processed image. In other terms, $(T_{t+h,t}u_0)(x)$ must essentially depend upon the values of $u_0(y)$ when y lies in a small neighbourhood of x.

We shall give two formal versions of this 'locality assumption'. Let us now just give its 'physical' interpretation: if the basic elements of the pyramid are assumed to be 'neurons', this only means that a neuron is primarily influenced by its neighbours. A clear argument for this is time: only neurons which are close can have an influence without transmission delay. Let us finish with an intuitive requirement which is called in image-processing 'causality'. Since the visual pyramid is assumed to yield more and more global information about the image and its features, it is clear that when the scale increases, no new feature should be created by the multiscale analysis: the image and the boundaries at scale $t' > t$ must be simpler than the boundaries at scale t. The causality assumption must of course be formalized. Its formalization has been discussed by Hummel (1986), Koenderink (1984, 1990a), Yuille (1988), Witkin (1983), Perona and Malik (1987) in the framework of image processing, by Kimia *et al.* (1992) in the framework of shape analysis and by Muerle and Allen (1968), Brice and Fennema (1970), Horowitz and Pavlidis (1974) in early works on image segmentation.

The result of the discussion in the case of image processing is that causality must be formalized as *pyramidality* plus a *local comparison principle*: if an image u is locally brighter than another v, then this order must be conserved some time by the analysis (prevalence of local behaviour on global behaviour). In formal terms, it can be expressed as the

Local Comparison Principle (Causality 2) If $u(y) > v(y)$ for y in a neighbourhood of x and $y \neq x$, then for h small enough,

$$(T_{t+h,t}u)(x) \geq (T_{t+h,t}v)(x).$$

In the case of edge detection, there are several formalizations, but the simplest states that no new boundary is created when the scale increases, that is, $T_{t'}K$ is contained in T_tK if $t' > t$.

We finally need some assumption stating that a very smooth image must evolve in a smooth way with the multiscale analysis. Somehow, this belongs to the 'causality' galaxy, but we prefer to call it regularity and it clearly corresponds to the assumption of the existence of an infinitesimal generator for the multiscale analysis.

Regularity Let $u(y) = \frac{1}{2}(A(x-y), x-y) + (p, x-y) + c$ be a quadratic form of $\mathbb{R}^N$. There exists a function $F(A, p, x, c, t)$, continuous with respect to A, such that

$$\frac{(T_{t+h,t}u - u)(x)}{h} \to F(A, p, x, c, t), \quad \text{when} \quad h \to 0.$$

Morphological and affine invariance In addition to the causality requirement, we must keep in mind that the pyramid acts on equivalence classes of images of the form $g(u_0(Ax))$, where g is any nondecreasing continuous function and A any isometry (or any affine map) of the plane. Therefore, the output should not depend upon u_0 but on the equivalence class. So the transition operators $T_{t+h,t}$ must somehow commute with the perturbations g and A. In the case of a change in contrast g, this is easily translated into the

Morphological invariance $gT_{t+h,t} = T_{t+h,t}g$, which means that change of contrast and multiscale analysis can be applied in any order. If A is an isometry, the same kind of relation must be true. Denote by Au the function $Au(x) = u(Ax)$. Then we state the

Euclidean invariance $AT_{t+h,t} = T_{t+h,t}A$.

Let us now examine the case of an arbitrary linear map A. The commutation relation cannot be so simple because A can reduce or enlarge the image. (Think of the case where A is a zoom defined by $Au(x) = u(\lambda x)$ for some positive constant λ.) Since the zoom has changed the scale of the image, we can just impose a weak commutation property:

Affine invariance For any A and $t \geq 0$, there exists a C^1 function $t'(t, A) \geq 0$ such that

$$AT_{t'(t,A),t'(s,A)} = T_{t,s}A.$$

Moreover, the function $\phi(t) = (\partial t'/\partial \Lambda)(t, \lambda \,\mathrm{Id})$ is positive for $t > 0$.

This relation means that the result of the multiscale analysis T_t is independent of the size and position in space of the analysed features: an affine map corresponds to the anamorphosis of a plane image when it is presented to the eye at any distance large enough with respect to its size and with an arbitrary orientation in space. (The general visual invariance should be projective, but for small objects at some distance, we shall be contented with the affine invariance.) (See Forsyth *et al.* (1991), Lamdan *et al.* (1988).) The assumption on t', $\phi(t) = (\partial t'/\partial \lambda)(t, \lambda \,\mathrm{Id}) > 0$, can be interpreted by looking at the relation $\lambda(\mathrm{Id})T_{t'} = T_t(\lambda \,\mathrm{Id})$ when λ increases, i.e. when the image is shrunk before analysis by T_t. Then, the corresponding analysis time before shrinking is increased. In more informal terms we can say that the analysis scale increases with the size of the picture.

Let us point out the fact that the affine invariance must be stated in such a general framework because we have, up until now, made no attempt to fix the relation between the abstract 'scale' parameter and the concrete scale understood as having some relation with the size of objects. As for all future results, they will be true whatever the change in abstract scale $T_t \to T_{\sigma(t)}$ provided σ is a smooth increasing function: $\mathbb{R}^+ \to \mathbb{R}^+$. Now, the

next lemma will permit a full normalization of the scale, thanks to affine invariance.

Lemma 1 (Normalization of scale.) Assume that $t \to T_t$ is a one-to-one family of operators satisfying pyramidality and affine invariance. Then the function $t'(t, B)$ only depends on t and $|\det B|$: $t'(t, B) = t'(t, |\det B|^{1/2})$ and is increasing with respect to t. Moreover, there exists an increasing differentiable rescaling function $\sigma : [0, \infty] \to [0, \infty]$, such that $t'(t, B) = \sigma^{-1}(\sigma(t)|\det B|^{1/2})$ and if we set $S_t = T_{\sigma^{-1}(t)}$ we have $t'(t, B) = t|\det B|^{1/2}$ for the rescaled analysis.

We shall give a proof of this lemma in Appendix B, because it is of particular relevance in image processing.

To summarize, the multiscale analysis T_t must (or may) satisfy:

- **Causality** $T_{t+h} = T_{t+h,t}T_t$, $T_{t,t} = T_0 = \mathrm{Id}$, and $(T_{t+h,t}u)(x) > (T_{t+h,t}v)(x)$ if $u(y) > v(y)$ for y in a neighbourhood of x and $y \neq x$. In the case of boundary multiscale analysis, this last assumption is replaced by $T_tK \subset T_sK$ if $t > s$.
- **Regularity** Let $u(y) = \frac{1}{2}(A(x-y), x-y) + (p, x-y) + c$ be a quadratic form of $\mathbb{R}^N$. There exists a function $F(A, p, x, c, t)$, continuous with respect to A, such that

$$\frac{(T_{t+h,t}u - u)(x)}{h} \to F(A, p, x, c, t) \quad \text{when} \quad h \to 0.$$

- **Morphological invariance** $gT_{t+h,t} = T_{t+h,t}g$ for any change of contrast g.
- **Euclidean invariance** $AT_{t+h,t} = T_{t+h,t}A$ for any isometry A of $\mathbb{R}^N$.
- **Affine invariance** $AT_{t'+h',t'} = T_{t+h,t}A$ with $t' = |\det A|^{1/2}t$ and $h' = |\det A|^{1/2}h$. Notice that the affine invariance implies the Euclidean invariance.
- **(Optional) linearity** $T_{t+h,t}(au + bv) = aT_{t+h,t}(u) + bT_{t+h,t}(v)$. We add this property because it has been very much in use in computer vision.

There are therefore five main axioms and we shall see that they allow us to classify and characterize the theories of multiscale image and shape processing completely, to unify and to improve several of them.

3. Axiomatization of image multiscale analysis and classification of the main models

Theorem 1 (Koenderink, 1990a; Hummel, 1986; Yuille and Poggio, 1986.)

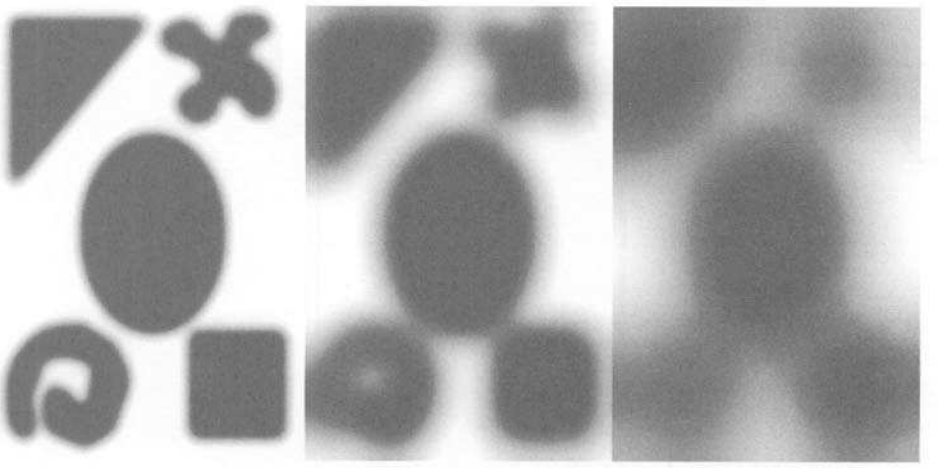

Fig. 3. Linear multiscale analysis (heat equation). Evolution of the original image given in Figure (1a) (from left to right): (a) $t = 10$; (b) $t = 70$; (c) $t = 300$.

If a multiscale analysis is causal, Euclidean invariant and linear, then it obeys (up to a rescaling $t \to \sigma(t)$) the heat equation

$$\frac{\partial u}{\partial t} = \Delta u,$$

(that is, $u(t,x) = (T_t u)(x)$ obeys the heat equation).

This model of multiscale analysis (the 'raw primal sketch') is due (among others) to Hildreth and Marr (see Marr (1982)) and Witkin (the *scale space*) (1983). See also Koenderink (1984) who was the first to state the heat equation explicitly. More recently Lindeberg (1990) studied the associated discrete scale space and Florack *et al.* (1992) showed how to use the heat equation to find corners, T-junctions, etc. by simple differential operators.

What happens if we remove the linearity axiom? As noted in Perona and Malik (1988) in their nonlinear theory of scale space, 'anisotropic diffusion', we can get nonlinear heat equations

$$\frac{\partial u}{\partial t} = F(\mathrm{D}^2 u, \mathrm{D}u, u, x, t). \tag{3.1}$$

The converse implication and a complete study of nonlinear models is given in Alvarez *et al.* (1992a):

Theorem 2 (Fundamental theorem.) If an image multiscale analysis T_t is causal and regular then $u(t,x) = (T_t u)(x)$ is a viscosity solution of (3.1), where the function F, defined in the regularity axiom, is nondecreasing with

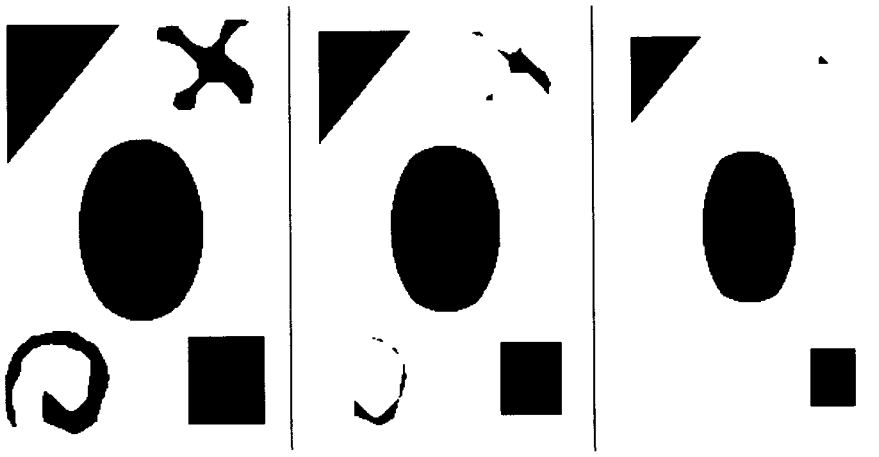

Fig. 4. 'Dilation' multiscale analysis. Evolution of the original image given in Figure (1a) (from left to right): (a) $t = 5$; (b) $t = 10$; (c) $t = 15$.

respect to its first argument D^2u. Conversely, if u_0 is a bounded uniformly continuous image, then equation (3.1) has a unique viscosity solution.

(For a quick proof of this theorem, see Appendix A.)

The particular case of the heat equation corresponds to $F(A, p, c, x, t) = \text{trace}(A)$. As a consequence of this theorem, all multiscale models can be classified and new, more invariant models can be proposed:

Theorem 3 Let $N = 2$. If a multiscale analysis is causal, regular, Euclidean invariant and morphological, then it obeys an equation of the form

$$\frac{\partial u}{\partial t} = |Du| F(\text{div}(Du/|Du|, t), \tag{3.2}$$

where $\text{div}(Du/|Du|)(x)$ can be interpreted as the curvature of the level line of the image $u(t, x)$ passing by x and $F(s, t)$ is nondecreasing with respect to the real variable s.

An important particular case is when F is a constant function: if $F = +1$ or $F = -1$, then the equation becomes $\partial u/\partial t = |Du|$ (resp. $\partial u/\partial t = -|Du|$), which corresponds to the so-called morphological erosion when the sign is '$-$' and to a morphological dilation when the sign is '$+$' (see Brockett and Maragos (1992)). *Dilation* and *erosion* are the basic operators of the Morphologie Mathématique, founded by Matheron (1975) and his 'Fontainebleau School'. The dilation at scale t is defined by

$$D_t u_0(x) = \sup_{y \in B(x,t)} u_0(y),$$

where $B(x, t)$ is a set centred at x: the 'structuring element' which is gen-

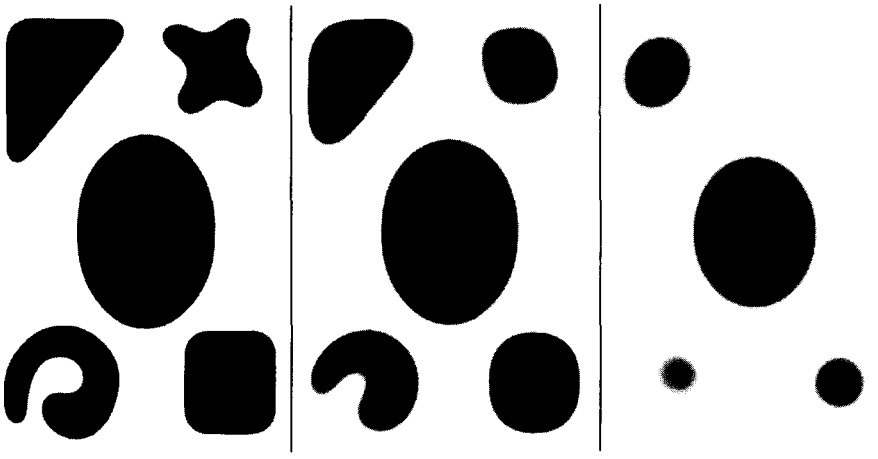

Fig. 5. 'Mean curvature motion' multiscale analysis. Evolution of the original image given in Figure (1a) by using scheme (10.4) (from left to right): (a) $t = 3$; (b) $t = 9$; (c) $t = 14$.

erally a ball with radius t. Assume, for instance, that u_0 is the characteristic function of a set X, then $\mathrm{D}_t u_0$ is the characteristic function of the t-neighbourhood of X. For the erosion, one simply replaces 'sup' by 'inf'. As noted by Michel Rascle, another relevant example satisfying the multiscale morphological axioms and therefore a parabolic PDE is the family of zooms with ratio t,

$$(T_t u_0)(x) = u(t, x) = u_0(tx).$$

Indeed, the preceding theorem applies and it is easily seen that the underlying equation is

$$\frac{\partial u}{\partial t} = \frac{1}{t}(\mathrm{D}u \cdot x).$$

Note that this formulation may be useful, the zooming operators on a digital picture being in no way easy to implement. Now, the preceding examples have been very particular instances of the equation and there are many other possibilities for morphological multiscale filtering! If we set $F(s, t) = s$, we obtain the 'mean curvature equation' (MCM)

$$\frac{\partial u}{\partial t} = t \cdot |\mathrm{D}u| \mathrm{div}(\mathrm{D}u/|\mathrm{D}u|). \tag{3.3}$$

This equation comes from a reformulation by Osher and Sethian (1988) of a differential geometry model studied by Grayson (1987) and Gage and Hamilton (1986). It is also very close to the 'anisotropic diffusion' of Perona and Malik (1987) and to an image restoration equation due to Rudin *et al.*

(1992a):

$$\frac{\partial u}{\partial t} = \operatorname{div}(Du/|Du|).$$

Now, the most invariant model is new: it is proved in Alvarez *et al.* (1992a) that

Theorem 4 (AMSS Model.) Let $\mathbb{N} = 2$. There is a single causal, regular, morphological and affine invariant multiscale analysis. Its equation is

$$\frac{\partial u}{\partial t} = |Du|(t \cdot \operatorname{div}(Du/|Du|))^{1/3}. \tag{3.4}$$

We shall better understand this equation in the framework of shape analysis: Sapiro and Tannenbaum (1992a,b), who independently discovered the model as a *shape scale space* have given it a remarquable geometric interpretation in this framework. Of course, fundamental Theorem 2 can be applied in any dimension. Let us just state a last example of scale space in dimension 3 (of particular relevance for medical solid images) (see again Alvarez *et al.* (1992a) and Caselles *et al.* (1993)).

Theorem 5 Let $N = 3$. There is (up to a rescaling) a single causal, regular, affine invariant and morphological multiscale analysis, associated with the equation

$$\frac{\partial u}{\partial t} = |Du|(G(u)^+)^{1/4}. \tag{3.5}$$

By $G(u)$ we denote the Gaussian curvature, that is the determinant of $D^2(u)$ restricted to $Du^{\perp}$. We shall expand more on this subject when looking for movie analysis equations.

4. Shape multiscale analyses

We could deduce the shape analysis statements from image analysis statements. However, since in this case the axiomatics is particularly simple and intuitive, we shall list well-adapted principles, which are, however, equivalent to the general image analysis principles. For more details, see Kimia *et al.* (1992), Lopez and Morel (1992), Mackworth and Mockhtarian (1992).

We define a shape or ('silhouette') as a closed set X whose boundary is a Jordan curve of $\mathbb{R}^2$. We denote by $T_t(X)$ the *shape analysed at scale t*. X is identified with its characteristic function $X(x) = 1$ if $x \in X$ and 0 else. We call multiscale analysis any family of operators $(T_t)_{t \geq 0}$ acting on shapes and we set $X(t) = T_t(X)$. As before, we shall state *causality* principles, the first of which remains unchanged:

Pyramidal Structure (Causality 1) $T_{t+h} = T_{t+h,t}T_t$, $T_0 = \mathrm{Id}$. Instead of the local comparison principle, we shall give a very intuitive statement: the shape local inclusion principle.

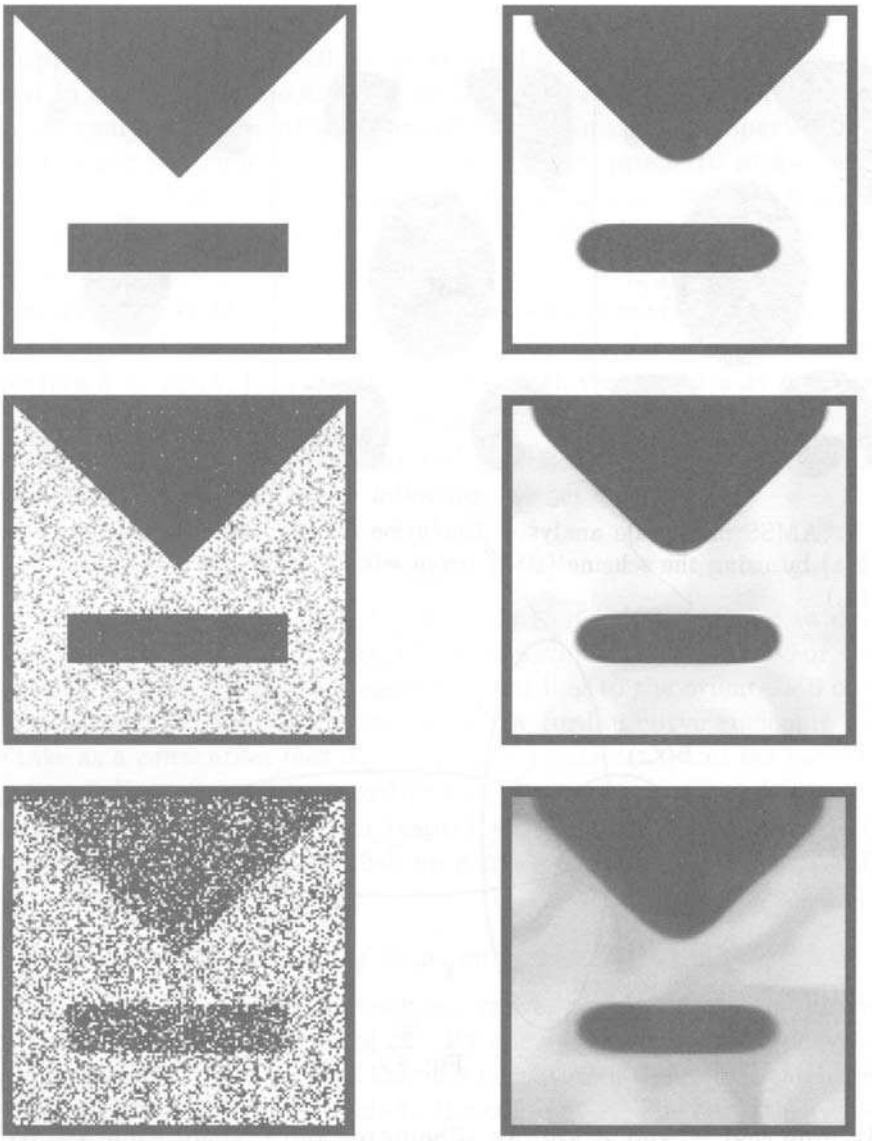

Fig. 6. Affine Invariant Morphological Multiscale Analysis. Authors: L. Alvarez–F. Guichard. A multiscale analysis T_t associates an image $u_0(x)$ with more regular images $u(x,t) = (T_t u_0)(x)$, where t is the scale of analysis. In this experiment, $u(t,x)$ is computed by the AMSS, $\partial u/\partial t = |Du|(t \cdot \text{curv}(u))^{1/3}$ where Du is the gradient of u, $\text{curv}(u)(x) = \text{div}(Du/|Du|)(x)$ the curvature of the level line of u passing by x. In order to illustrate the use of the AMSS model as a way of keeping only reliable information in image analysis, we display on the left-hand column two increasingly distorted (noisy) versions of an original synthetic image (up-left). The left-down image is obtained from the left-up by giving 70% of its pixels a random value. In the right-hand column, we display the respective analyses at the *same* scale t of these images. Two phenomena are illustrated by comparing the right and left columns: first, the *causality* (a smoothing effect), selective loss of information from fine to coarse; second, the *morphologically invariant* behaviour. Unless the level sets of the triangle and rectangle have quite different mean brightness in the noisy images, their shape is handled by the analysis in the same way. Size of the images: 128 × 128. CPU time: 21 s.

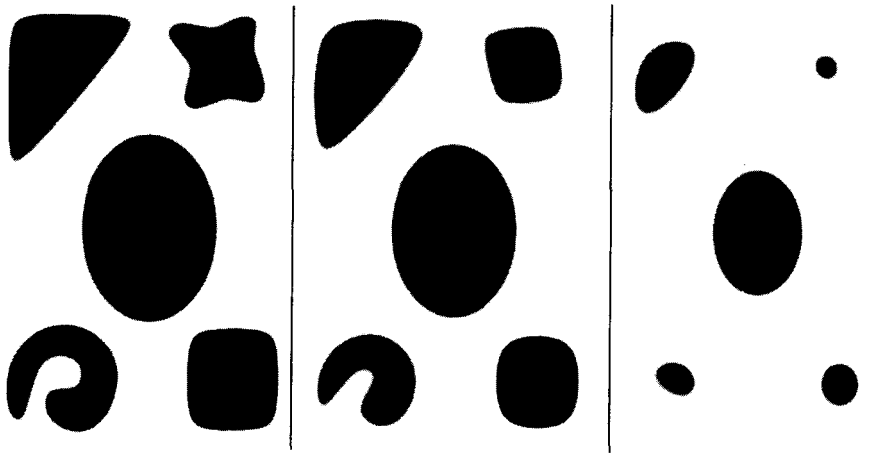

Fig. 7. AMSS multiscale analysis. Evolution of the original image given in Figure 1(a) by using the scheme (10.8) (from left to right): (a) $t = 3$; (b) $t = 9$; (c) $t = 14$.)

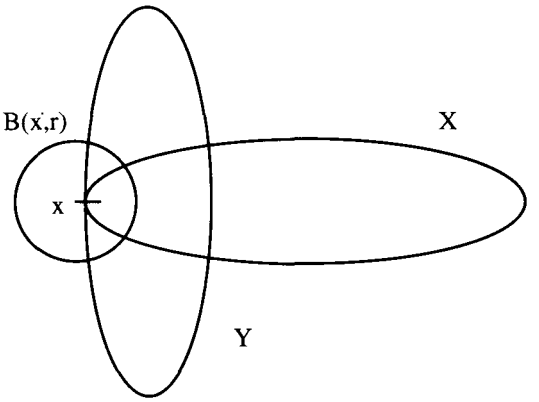

Fig. 8.

Assume that X and Y are two silhouettes and that for some $x \in \partial Y$ and some $r > 0$, one has $X \cap B(x, 2r) \subset Y \cap B(x, 2r)$. Assume further that the inclusion is strict in the sense that ∂X and ∂Y only meet possibly at x. Then we shall say that *the shape X is included in shape Y around x.*

Shape local inclusion If X is included in Y around x, then for h small enough, $T_{t+h,t}(X) \cap B(x, r) \subset T_{t+h,t}(Y) \cap B(x, r)$.

This last axiom implies that the value of $T_{t+h,t}(X)$ for h small, at any point x, is determined by the behaviour of X near x. We are allowed to take r infinite. Therefore the shape local inclusion principle also implies that if a shape is globally contained in another, this order is preserved for every scale (Mackworth and Mokhtarian, 1992).

Both preceding principles allow, as we shall see, the shape analysis to be localized in space and time and we therefore only need to state what the multiscale analysis makes of very simple shapes in order to specify it. So we add a 'basic principle' which will state what happens to disks. As we shall see, disks are somehow a 'basis' for shape analysis because whenever we know how they are analysed, we know what will happen to every other shape.

Basic principle Let $D = D(x, 1/r)$ be a disk with curvature $1/r$ and centre x. Then $T_{t+h,t}(D)$ is a disk with radius $\rho(t, h, 1/r)$ and centre x. Moreover, the function $h \to \rho(t, h, 1/r)$ is differentiable with respect to h at $h = 0$ and the differential is continuous with respect to $1/r$.

The 'basic principle' implies that the multiscale analysis behaves in a smooth and isotropic way. In the following, we set

$$g(t, 1/r) = \frac{\partial \rho}{\partial h}(t, 0, 1/r). \tag{4.1}$$

Note that $g(t, s)$ is defined for $t \geq 0$ and $s \in \mathbb{R}$. Now in order to define $g(t, 0)$ we must assume that $\lim\ g(t, 1/r)$ exists when r tends to $+\infty$ or $-\infty$. The radius r may be positive or negative, according to the orientation of the normal $\vec{n}(x)$. In the case where the curve is a Jordan curve enclosing a set X, we take as a *convention that* $\vec{n}(x)$ is *pointing outside* X and *the curvature is negative if* X *is convex at* x, *positive else.* It may seem natural to assume therefore that $g(t, \kappa)$ is odd with respect to κ. Indeed, this corresponds to the assumption that a black disk on grey background and a white disk behave in the same way.

4.1. The fundamental equation of shape analysis

When a point x belongs to an evolving curve, we denote by $\dot{x}$ the time derivative of x, which is a vector of $\mathbb{R}^2$. By $\mathrm{curv}(x)$ we denote the curvature of a curve which is C^2 at x. Recall that the curvature is defined as the inverse of the radius of the osculatory circle to the curve at x. The curvature is zero if the radius is infinite.

Theorem 6

1 Under the three principles (pyramidal, local shape inclusion, 'basic'), the multiscale analysis of shapes is governed by the curvature motion equation

$$\dot{x} = g(t, \mathrm{curv}(x))\vec{n}(x) \tag{4.2}$$

where g is defined by (4.1).

2 If the analysis is affine invariant, then the equation of the multiscale analysis is, up to rescaling,

$$\dot{x} = \gamma(t \cdot \mathrm{curv}(x))\vec{n}(x) \tag{4.3}$$

where γ is defined by $\gamma(x) = a \cdot x^{1/3}$ if $x \geq 0$ and $\gamma(x) = b \cdot x^{1/3}$ if $x \leq 0$ and a, b are two nonnegative values.

3 If we add that $T_t(X^c) = T_t(X)^c$ (reverse contrast invariance) then the function g in (i) is odd and we get

$$\dot{x} = (t \cdot \mathrm{curv}(x))^{1/3} \vec{n}(x). \tag{4.4}$$

We prove this theorem in Appendix C. By the expression 'governed by', we mean that the equation must be satisfied at any point (t, x) of $\partial X(t)$ where the boundary of the silhouette is smooth enough to give a classical sense to both terms of the equation.

Remark It has been proved (Gage and Hamilton, 1986; Grayson, 1987; Angenent, 1989) that equation (4.2) with $g(t, s) = s$ has smooth solutions. This equation has been introduced in picture processing by Kimia *et al.* (1992a), Mackworth and Mockhtarian (1992) and Alvarez *et al.* (1992a) in different contexts. An early version of an algorithm leading to equation (4.2) is given in Koenderink and Van Doorn (1986). See also Yuille (1988). Equation (4.4) has been introduced and axiomatically justified (with a more complicated axiomatic however) in Alvarez *et al.* (1992a). It has also been proposed in Sapiro and Tannenbaum (1992a). The existence and regularity of the solution are proved in Sapiro and Tannebaum (1992b). The axiomatic presentation adopted here follows Cohignac *et al.* (1993a).

5. Relation between image and shape multiscale analyses

In this section, we shall show how the AMSS model (1.1) can be deduced from the shape evolution (4.4). We give in Appendix C proofs of the axiomatic deduction of this last equation, so that our exposition will be rather complete. In addition, the shape multiscale analysis equation has an easy geometric interpretation as an 'intrinsic diffusion' which we shall explain at the end of this section. Since the multiscale analysis satisfies the obvious inclusion principle that

$$\text{If} \quad A \subset B \quad \text{then} \quad T_{t+h,t}(A) \subset T_{t+h,t}(B),$$

we can, as well known in the 'mathematical morphology school' (Matheron 1975, Maragos 1987), associate with a picture u the set of its level sets

$$X_a u = \{(x, y), u(x, y) \geq a\}.$$

Then, assuming that each level set is a union of silhouettes, we can simply define $T_t(u)$ from the multiscale analysis $T_t(X)$ of silhouettes by

Morphological Principle For any a, t, h and

$$u : X_a T_{t+h,t}(u) = T_{t+h,t}(X_a u).$$

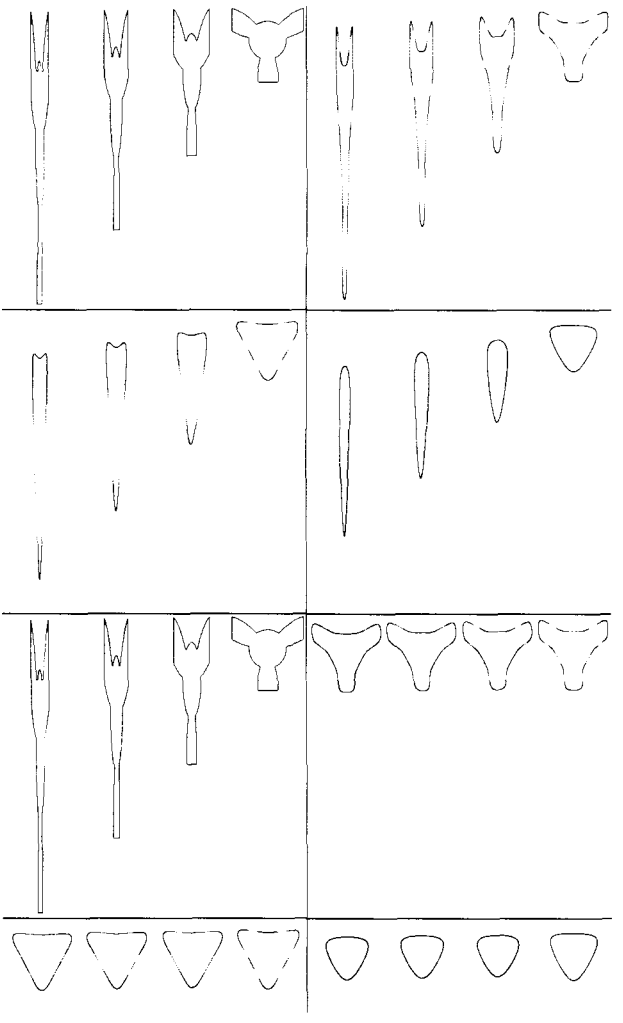

Fig. 9. Affine invariant shape recognition made possible thanks to the AMSS model. Authors: Cohignac–Eve–Guichard–Lopez–Morel. Which numerical implementation can be done using the AMSS model? We wish, after many iterations of the discrete algorithm, to be able to recognize shapes. So the numerical algorithm should strictly have the same invariance properties as the continuous model. In the case of affine invariance, this happens to be a rather new problem in numerical analysis. We display here the results of an algorithm which will be discussed later on. These results illustrate the affine invariance of the analysis. A shape is distorted by affine anamorphoses with respective eigenvalues (1/2,2), (1/3,3) et (1/4,4) (first quadrant, up-left). The quadrants 5, 6, 7, 8 display the multiscale analyses of these shapes at four successive scales. If we apply the inverse anamorphoses to the shapes after analysis, they must be the same (if the algorithm is correct!). This can be checked in quadrants 2, 3, 4.

Theorem 7 Assume that a multiscale image analysis T_t satisfies the morphological principle and that each of the level sets is governed by equation (4.2). Then $T_t u$ satisfies

$$\frac{\partial u}{\partial t} = g(t, \mathrm{curv}(u))|\mathrm{D}u|. \tag{5.1}$$

Proof. Let us prove (3.2) at any (x, a) such that $u(t,x) = a$ and $u(t,y) = a$ implies $\mathrm{D}u(t,y) \neq 0$ and u is C^2 at (t,y). The first condition implies by the implicit function theorem that x belongs to a Jordan curve Γ enclosing two regions Γ^+ and Γ^-, with $u(t,x) < a$ on Γ^- , $u(t,x) > a$ on Γ^+ and $u(t,x) = a$ on Γ. Set $u(t) = u(t,.)$. The relation $X_a T_{t+h,t} u(t) = T_{t+h,t} X_a u(t)$ implies that $T_{t+h,t}\Gamma^+$ is equal to a connected component of $X_a u(t+h)$. Therefore, the point $x(t+h)$ defined by equation (4.2) with initial value $x(t) = x$ belongs to the boundary of $X_a u(t+h)$. So we obtain $u(t+h, x(t+h)) - u(t, x(t)) = 0$. Dividing by h and passing to the limit yields

$$\frac{\partial u}{\partial t}(t,x) + \mathrm{D}u(t,x)\dot{x}(t) = 0$$

and using equation (4.2) and $\vec{n}(x) = -\mathrm{D}u/|\mathrm{D}u|$ we obtain equation (3.2).

The Affine Scale Space (ASS) model (4.4) of Sapiro and Tannenbaum yields a simple geometric interpretation of the AMSS model (1.1). Let us consider two ways of parametrizing a smooth Jordan curve $x(s)$:

- either in an Euclidean-invariant way by imposing $|x_s| = 1$; or
- or in an affine-invariant way by setting $|\det(x_s, x_{ss})| = 1$.

In the second case, we say that the curve has been parametrized by its 'affine length parameter' s. This parametrization is affine covariant because $|\det(Ax_s, Ax_{ss})| = |\det A||\det(x_s, x_{ss})|$ for any affine map A. Sapiro and Tannenbaum (1992b) proved (it is an easy computation) that the ASS model is equivalent to the following 'intrinsic heat equation ':

$$\frac{\partial x(s,t)}{\partial t} = t^{1/3}\frac{\partial^2 x(s,t)}{\partial s^2}, \quad x(0,s) = x_0(s).$$

So the application of the AMSS model to an image u can be interpreted as the affine invariant diffusion of all the level lines (isophotes) of u. Therefore we obtain a nonlinear generalization of the linear classical scale spaces.

6. Multiscale segmentation

Segmentation is acknowledged as the main tool in image interpretation. As we shall see, the segmentation problem is quite well understood in the framework of multiscale analysis. To define this problem in two lines, let us say that the objective of segmentation is to find the homogeneous regions of an image as well as their boundaries. However, the term 'homogeneous' is

extremely vague and in order to state what homogeneity is, one has to rely either on perceptual experiments or on axiomatic definitions. In any case, homogeneity may concern clues as different as brightness, colour and texture. In this section, we shall not discuss what these clues are (this will be axiomatically introduced in the next section for textures). We shall assume that the image datum is composed of a finite set of k 'channels', each channel being itself a real image. In the simplest case, there is a single channel, the brightness. In the colour image case, the actual technology (partly based on perceptual criteria) yields three channels, i.e. three images of the same size (red, green, blue). In the case of other clues, like texture elements, the number of channels is unlimited and experiments showed in the next section involve up to forty channels computed from an initial grey level image and with the same size. The segmentation problem assumes that an initial multichannel image u_0, with $u_0(x) \in \mathbb{R}^k$, and an initial boundary map K_0, where K_0 is a subset of the image domain with finite Hausdorff length, are given. The initial boundaries can simply be the boundaries of all pixels. The segmentation process computes a multiscale sequence $(K(t))$ of segmentations, as well as a multiscale sequence of piecewise homogeneous images $u(t)$. $K(t)$ is assumed to be the set of the boundaries of the homogeneous regions of $u(t)$.

We shall give simple multiscale principles which will closely determine the segmentation process. Setting, as usual, $T_t K_0 = K(t)$, we impose

- **Causality** $K(t') \subset K(t)$ if $t < t'$ and $T_{t,s} T_s = T_t, \quad T_0 = \mathrm{Id}$.
- **Euclidean invariance** $T_{t+h,t} R = R T_{t+h,t}$ for every isometry R.
- **Invariance by zooming** Let Z_λ denote the zoom with ration λ. Then

$$T_t Z_\lambda = Z_\lambda T_{\lambda^{1/2} t}.$$

Finally we need an axiom fixing the relations between the boundaries $K(t)$ at scale t and the image $u(t)$. Our choice (see Morel and Solimini (1993)) is to take the simplest principle, which Mumford and Shah (1988) called the 'cartoon principle'.

Cartoon Principle $u(t)$ is locally constant in $\mathbb{R}^2 \setminus K(t)$ and equal to the mean value of u_0 on each connected component of $\mathbb{R}^2 \setminus K(t)$.

The causality and cartoon principles nearly fix the type of segmentation algorithm to be used: it is a *region growing* algorithm. Let us define *region* as every connected component of $\mathbb{R}^2 \setminus K(t)$. Since, for

$$t' > t, \quad K(t') \subset K(t),$$

we deduce that the regions at scale t' are unions of regions at scale t (which is another way of stating the causality axiom). Thus, in order to completely fix the multiscale analysis, we only need a criterion for region 'merging'. Mumford and Shah (1988) proposed the following Euclidean and zoom-invariant

criterion: every segmentation is associated with an energy (which somehow measures its complexity),

$$E(u(t), K(t)) = \int_{\mathbb{R}^2 \setminus K(t)} (u(t,x) - u_0(x))^2 \, dx + t \cdot \text{length}(K(t)).$$

Then two regions of the segmentation will be merged at scale t if and only if the energy of the resulting segmentation decreases. The associated 'recursive merging' algorithm is extremely simple. We start with an initial trivial segmentation of the image at scale $t = 0$. In this case, the image simply is (e.g.) divided into 'pixels' (small squares in the actual technology).

Region Growing Variational Algorithm

- Fix u_0, K_0 as the initial trivial segmentation, where K_0 is the union of boundaries of all pixels.
- A scale t being fixed: for every pair of regions, check whether their merging decreases the Mumford–Shah energy and, if so, merge them. Then the new K is obtained by removing the common boundary of both regions and u takes the mean value of u_0 on the union of both regions as the new value.
- Increment the scale t and go back to the preceding step.

Mumford and Shah (1989) proved that a segmentation which is minimal for their energy has a finite number of regions with smooth boundaries. It is, however, impossible to find a minimal segmentation, because the energy is highly not convex. Therefore, it is useful to get information about the segmentations obtained by a concrete algorithm computing local minima. Here is such a theorem, which justifies the use of region growing associated with the Mumford–Shah energy.

Theorem 8 (Koepfler *et al.*, 1994.) Let us say that a segmentation of a bounded vectorial image is *2-normal at scale t* if no pair of regions can be merged without increasing the Mumford–Shah energy. Then the set of 2-normal segmentations is compact for the Hausdorff distance and there is a bound, only depending on t, for the number of regions of a 2-normal segmentation.

We display in the following several examples of segmentations (Figures 10–11). This section has been kept short, as more than a thousand papers have been written on segmentation algorithms. Now, Morel and Solimini (1993), following and updating the terminology of Zucker (1976), have discussed more than ten classes of algorithms in image processing and have shown that they are merely variants of the Mumford–Shah energy minimizing process. Among the many theories which lead to the Mumford–Shah formalization (or variants), let us mention the 'snakes' (Kass *et al.*, 1987), the survey by Haralick and Shapiro (1985), the stochastic segmentation of

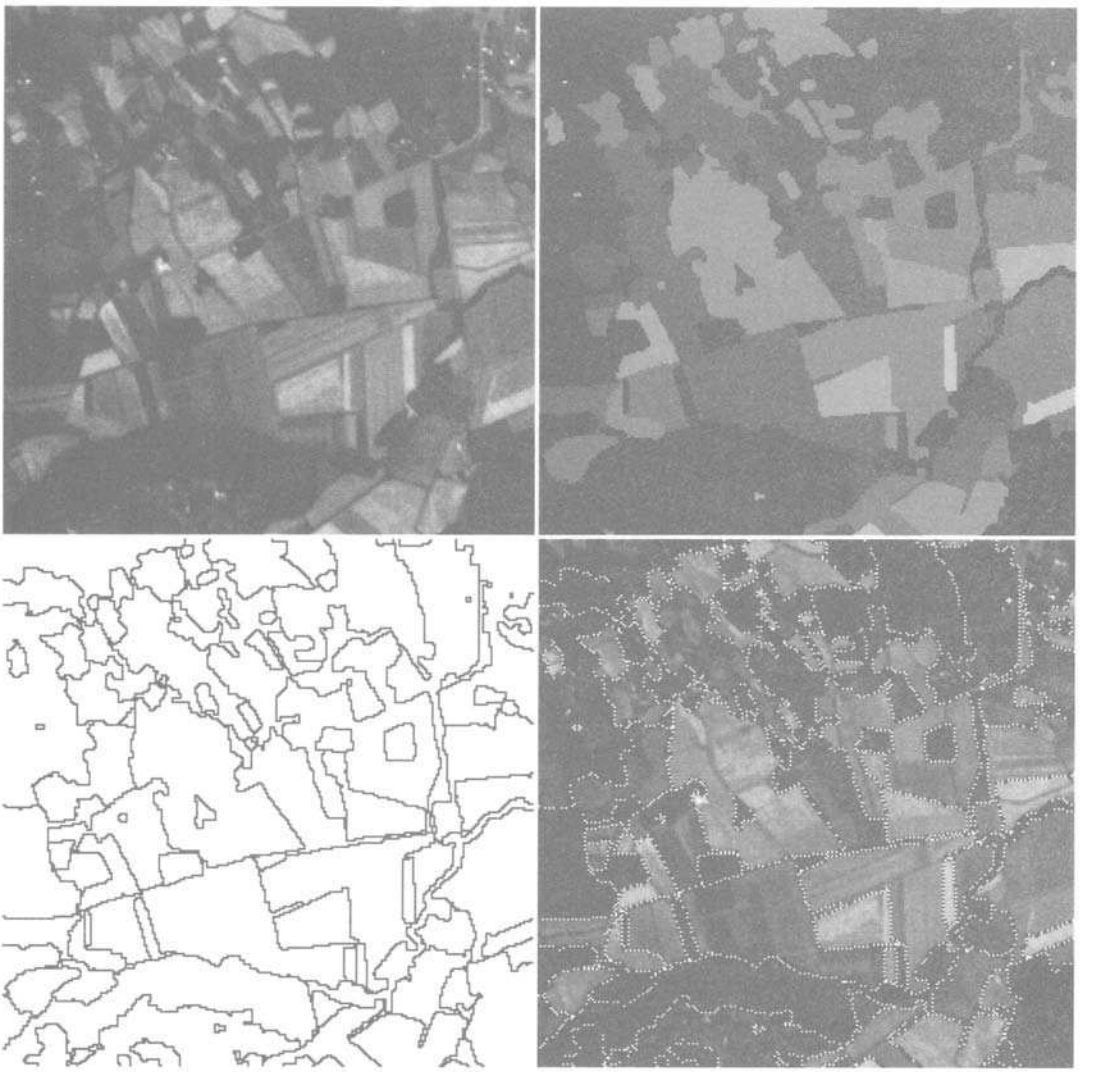

Fig. 10. Image segmentation by the Munford–Shah criterion. Authors: Koepfler–Morel. Up-left to right-down: an original satellite image (one single channel), the boundary map when it has been decided to keep 100 regions, the 'cartoon' (i.e. the associated piecewise constant image) and finally the superposition of the original image and the boundary map.

Geman and Geman (1984), the Blake and Zisserman 'weak membrane' model (1987), and the region growing algorithms of Brice and Fennema (1970), Muerle and Allen (1968), Pavlidis (1972), Horowitz and Pavlidis (1974). An affine invariant version of the Mumford–Shah theory has been proposed by Ballester and Gonzalez (1993).

7. An example: texture discrimination

Some theories of low level vision happen to be axiomatized and these can therefore be matched with the mathematical axiomatics which we developed above. This is the case for the Julesz Texton Theory (Julesz, 1981; 1986; Julesz and Kroese, 1988), which attempts to give a formal account of how human perception can discriminate textures in a few milliseconds. In order to understand the intuitive concept of *texture*, it is sufficient to look at the texture pair in Figure 11. The discrimination of both textures is easy for human perception, therefore whatever the computation involved in it is should also be. In order to confirm his texton theory experimentally, Julesz created pairs of simple different shapes, which were taken as building elements for creating two different but undiscriminable textures. This showed that not all *different* texture pairs are discriminable. Look at Figure 12 where a region of '10s' is surrounded by '5s'. Unless these shapes are individually quite distinguishable, they become quite equivalent when they are building elements of a 'texture'.

The axioms of shape analysis can be matched with the (not completely formal) axiomatics proposed by Julesz for his theory of preattentive texture discrimination. Therefore the texton theory can be numerically tested, and the result is quite different from the previous attempts and surprising. Indeed, previous attempts at formalization were based on linear image multiscale filtering followed by nonlinear mechanisms compatible with physiological data. One of the most conclusive in this way is due to Malik and Perona (1991), whose algorithm experimentally matches the human performance in texture discrimination. Now, as we shall see, a faithful implementation of the texton theory yields hyperdiscrimination! Julesz created pairs of textures which are undiscriminable for the preattentive vision but not for the algorithm deduced from his axioms.

In order to give a brief account of the texton theory, let us say that Julesz assumes that in perception only the local means of the curvatures and orientations of the shapes (or 'textons') of level sets are retained. These local means are called *texton densities*. Thus, every image u_0, supposedly containing textures, is associated with a multichannel image $u_i(t,x)$, $i = 1,2,\ldots,k$ where each image $u_i(t,x)$ stands for the density of the texton of type i at scale t. Then a segmentation process, computing 'texture gradients' is effectuated and yields the boundaries of the textures.

Therefore, the Julesz theory can be immediately translated into the multiscale analysis framework developed above. The translation yields a texture segmentation algorithm authorized to check how valid the texton theory is.

Texture Segmentation Algorithm

1 Apply the fundamental equation (AMSS) to the image: it is the only multiscale analysis allowing an invariant computation of multiscale curvatures and orientations.

2 Texton densities can be deduced as local means of curvatures and orientations at several scales.

3 Assuming that the analysed image only contains two texture regions, apply the region growing Mumford–Shah algorithm and stop the segmentation process when only two regions remain. If the regions are those predicted by the theory (and if the contrast of densities between them is strong), there is discrimination.

Let us briefly state how we organize the numerical experimentation of Julesz theory by using the above discussed scale-space algorithms.

7.1. Formalization and computation of textons densities

Let $u_0(x)$ be an image and $u(t,x)$ its multiscale analysis by equation (1.1). Then the curvature and orientation at scale t and location x are given respectively by $\mathrm{curv}(u)(x)$ and $Du/|Du|(x)$ which are both well defined and computed by equation (1.1).

Notice that textons can be white on black or white on black, so we define for any $\theta \in [0,\pi]$ and $t > 0$ the 'density on a ball $B(x,\Delta)$ of textons at scale t and orientation θ' by

$$G_{B_\Delta}(x) * (1_{\Theta_{\theta,\delta\theta}} \mathrm{curv}^+(u)(t,x))$$

$$G_{B_\Delta}(x) * (1_{\Theta_{\theta,\delta\theta}} \mathrm{curv}^-(u)(t,x))$$

where $\Theta_{\theta,\delta\theta}$ is the set of points where $Du/|Du| = (\cos\phi, \sin\phi)$ satisfies $|\theta - \phi| \leq \delta\theta$ and $G_{B_\Delta}(x)$ is the isotropic Gaussian with variance Δ. By a^+ we mean $\max(a,0)$ and by a^-, $\max(-a,0)$.

Then we get a formal definition of texton densities at each scale and orientation. Based on the obtained texton density channels, a segmentation of the image can be achieved. If the Julesz theory is correct, all the pairs of textures which we discriminate should be clearly discriminated by at least one of the texton densities.

Conversely, any pair of undiscriminable textures should be undiscriminated by all channels.

The basic facts which can be discovered by experimentation (Figures 11 and 12) are following. The textures proposed by Julesz, and most textures proposed in the literature as discriminable are easily discriminated by texton densities even rougher than those described above: Indeed, it is sufficient to compute global curvature densities, that is, we take $\delta\theta = \pi$ and do not

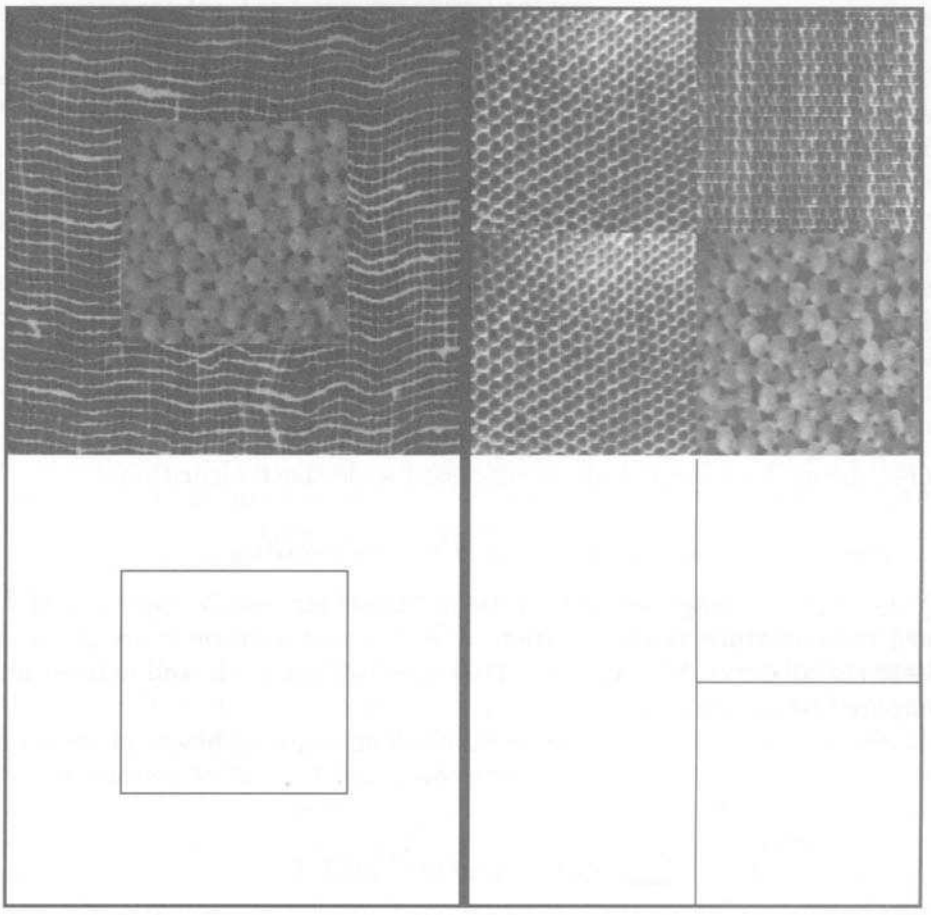

Fig. 11. Segmentation of natural textures. Authors: Koepfler–Lopez–Morel. On the left: two natural textures extracted from the Brodatz collection. On the right, three different Brodatz textures. Below: boundaries computed by the multichannel segmentation algorithm based on the Mumford–Shah functional. The 18 used channels are curvatures at different scales. Note how the same texture is repeated twice on the left of the right-hand image. The algorithm, which only takes into account local mean values of curvature, yields the right segmentation.

take into account orientation. The same process is followed for the Julesz undiscriminable textures!

The computational theory of texture discrimination which we have presented is the most complete attempt at computationally discussing the psychophysical theory of textons. It is, however, not the first successful attempt. The first one, already mentioned, is due to Malik and Perona (1991), who modelled texton density computations by a combination of linear oriented filters and nonlinear mechanisms. The Malik and Perona algorithm

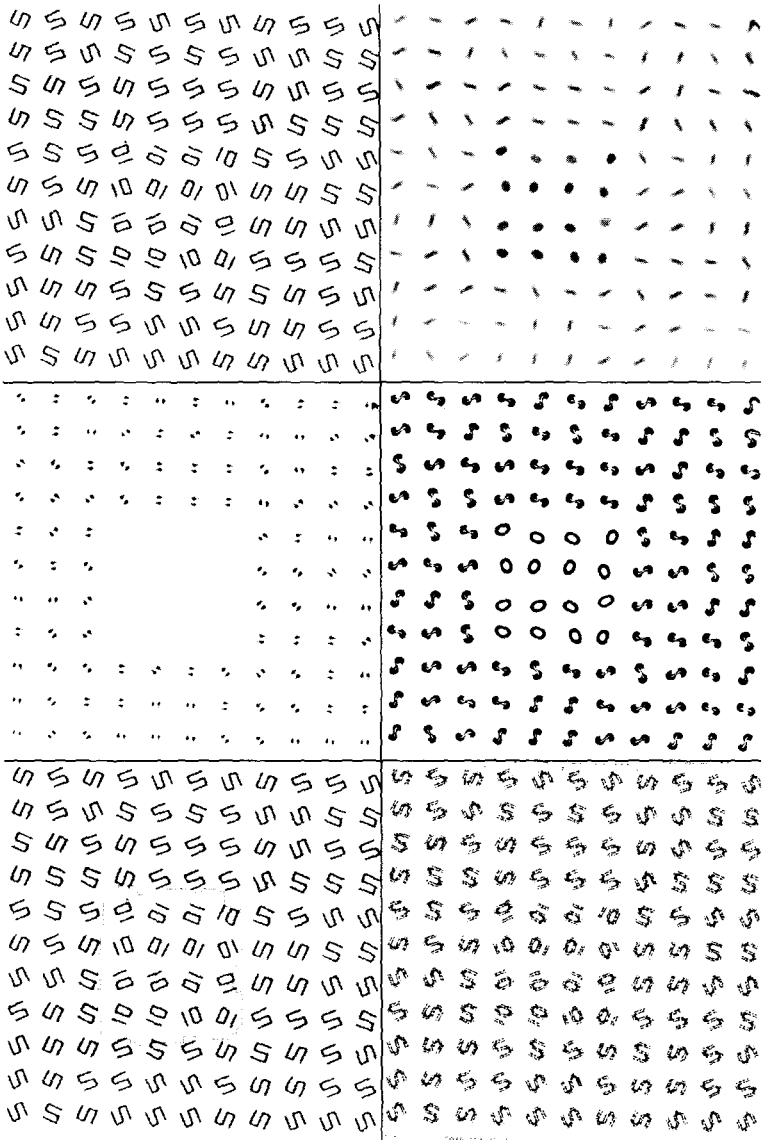

Fig. 12. Verification of Julesz' theory of textons which yields hyperdiscrimination algorithms. Experiment: Lopez–Morel. Julesz created textures which are undiscriminable for the human preattentive vision, as apparently predicted by the texton theory. However, the AMSS model, whose requirements match the axioms of the texton theory, leads to discriminate–undiscriminable texture pairs. Up–down and left–right: (a) a preattentively undiscriminable texture pair, u_0; (b) $u(t) = T_t u_0$: image obtained after application of the multiscale analysis T_t for $t = 7$ pixel units; (c) negative part of the curvature, $\mathrm{curv}(u(t))^-$; (d) positive part $\mathrm{curv}(u(t))^+$; (e) segmentation obtained by the Mumford–Shah piecewise constant model applied to the two-channel image of the curvatures; (f) negative curvature $\mathrm{curv}(u_0)^-$ of u_0. This last image shows that curvature-based discrimination was not possible on the original image. In order to explain the hyperdiscrimination of the AMSS model, the texton theory must be made more precise (Lopez and Morel, 1992).

matches the human preattentive vision and does not exceed it, as the above presented algorithm does. It is, however, easily seen by looking at their algorithm that it essentially computed curvature-based features, with less and less accuracy as the scale increased. The same article presents a very precise account of the previous attempts to make texton theory computational. (See, in particular, Treisman (1985), Voorhees and Poggio (1987), Enns (1986).)

8. Movie multiscale analysis

In this section, we return to scale space theory and look for its adaptation to movies. It is rather easy, since we still stay in the framework of the Fundamental Theorem 2. Indeed, a movie can be modelled as a 3D image, $u(x,\theta)$, where $x \in \mathbb{R}^2$ and $\theta \in \mathbb{R}$ is the time parameter. The statement and justification of the causality axioms remain unchanged, as well as the morphological invariance. Of course, the Euclidean and affine invariance make less sense in 3D, and we shall keep their 2D versions, the 2D-Euclidean invariance and the 2D-affine invariance, where isometries (respectively affine maps) are restricted to the image plane. We, however, add two specific axioms related to motion.

- **Time affine invariance** For any affine time rescaling $A_{a,b} : u(x,\theta) \to u(x, a\theta + b)$, there exist $t'(a,t)$ such that $T_t A_{a,b} = A_{a,b} T_{t'}$.
- **Galilean invariance** Denote by B_v any Galilean motion operator defined by $B_v u(x,\theta) = u(x - v\theta, \theta)$. Then $B_v(T_t u) = T_t(B_v u)$.

The Galilean invariance means that the analysis is invariant under 'travelling', that is motion of the whole picture at constant velocity v does not alter the analysis. In the following, we distinguish the 'spatial gradient' $\nabla u = (u_x, u_y)$, and the space–time gradient, $\mathrm{D}u = (u_x, u_y, u_\theta)$.

Theorem 9 If a multiscale analysis of movies is causal, regular, time-translation invariant, space-Euclidean invariant and morphological, then it is governed by the equation

$$\frac{\partial u}{\partial t} = |\nabla u| \, F(\mathrm{curv}(u), \mathrm{accel}(u), t). \tag{8.1}$$

If, in addition, we assume the analysis to be time-affine and space-affine invariant, the equation is (up to a rescaling)

$$\frac{\partial u}{\partial t} = (|\nabla u| \, \mathrm{curv}^{1/3}(u))^{1-q} \, ((|\nabla u| \mathrm{sgn}(\mathrm{curv}(u)) \, \mathrm{accel}(u))^+)^q, \tag{8.2}$$

for some $q \in [0,1]$.

(See Alvarez *et al.* 1992a,b.) In the above formulae, we use the convention that the power preserves the sign, that is $a^q = |a|^q \mathrm{sgn}(a)$. Hence, when

$q = \frac{1}{4}$, we obtain the equation

$$\frac{\partial u}{\partial t} = |\mathrm{D}u|(G(u)^+)^{1/4}. \tag{8.3}$$

This equation was mentioned in Section 2 as the only affine invariant morphological scale space in $\mathbb{R}^3$. Of course, this full affine invariance has no meaning for classical movies: what is the meaning of a rotation involving spatial and time variables? Now, in the field of relativity theory, such invariance makes sense because the Lorentz transform is nothing but a spatial-time rotation. In other words, when $q = \frac{1}{4}$, we have an equation which is both Galilean invariant and relativist invariant. We did not explain what the differential operator accel(u) is. (As for curv(u), it is simply the curvature of the spatial level curves, as in Section 2.) We could give the explicit formula for accel(u) in terms of the first and second derivatives of u, but this would prove disastrous, since the formula takes several lines. We shall use two ways to characterize accel and justify its name of 'apparent acceleration'. Let us first explain a classical notion in motion analysis: the apparent velocity. Assuming that a movie displays moving objects, let us call $x(\theta)$ the trajectory of a point of one of these objects. If we make the assumption that the object is *Lambertian*, which means that the light that it is sending to the camera is constant, then we can ensure that $u(x(\theta), \theta) = C$ for some constant C. Differentiating with respect to θ yields $\langle \nabla u, \dot{x}(\theta) \rangle + u_\theta = 0$. So we see that the component of the velocity in the direction of the spatial gradient can be recovered from the partial derivatives of u:

Definition We call *apparent velocity* of a movie $u(x, \theta)$ at point (x, θ) the scalar

$$v_1 = -u_\theta / |\nabla u|.$$

Now we are in a position to justify (and define) accel(u):

Lemma 2 (Interpretation of accel.) Consider a picture in translation motion $u(x, \theta) = w(x - \int_0^\theta \vec{v}(\tau)\, \mathrm{d}\tau)$, where $\vec{v}(\theta)$ is the instantaneous real velocity vector. Then the apparent velocity v_1 is equal to the true velocity in the direction of the gradient, $\langle \vec{v}, \nabla u / |\nabla u| \rangle$. Let $V = (\vec{v}, 1)$ be the real space–time velocity. Then

$$\mathrm{accel}(u) = -\langle \mathrm{D}v_1, V \rangle. \tag{8.4}$$

The first result is easy to check. The second formula can be taken as a definition of accel and shows that in the case of objects in translation motion, accel(u) is the derivative of the apparent velocity in the direction of $-\nabla u$. This is why we call it 'apparent acceleration'. For a proof, see Alvarez *et al.* (1992a,b) or Guichard (1993).

The second way of explaining what accel is consists in using Guichard's (1993) numerical approximant of it, which has proved essential when dealing with actual movies. Normal movie display series of frames where objects may jump more than 3 (and up to 50...) pixels from one frame to the next. Indeed, quick motions may let an object jump from one side of the screen to the other side in a very few frames. So, real movies are dramatically undersampled in time, which does seem to affect the trained zapper, but makes classical (local) numerical schemes impossible. So we define a nonlocal search space for the 'possible velocity vectors',

$$\mathcal{W} = \{\mathrm{V} = (\alpha, \beta) \text{ for all } \alpha \text{ and } \beta \text{ in } \mathbb{R}\}.$$

Theorem 10

$$\begin{aligned} & |\nabla u|(\mathrm{sgn}(\mathrm{curv}(u))\,\mathrm{accel}(u))^{+} \qquad (8.5) \\ = & \min_{w_1, w_2 \in \mathcal{W}} \frac{1}{\Delta\theta^2}(|u(x - w_1, \theta - \Delta\theta) - u(x, \theta)| \\ & + |u(x + w_2, \theta + \Delta\theta) - u(x, \theta)| + |\langle \nabla u, w_1 - w_2 \rangle|) + o(1). \end{aligned}$$

Interpretation Of course, for numerical experiments, we shall not compute the minimum for all vectors in $\mathcal{W}$, but only for the vectors on the grid. We have two differents parts in the second term: the first part is the variation of the grey level value of the point x, for candidate velocity vectors: w_1 between $\theta - \Delta\theta$ and θ, and w_2 between θ and $\theta + \Delta\theta$. This variation must be as small as possible, because a point is not supposed to change its grey level value during its motion. The second part is nothing but an 'apparent acceleration', that is, the difference between w_1 and w_2 in the direction of the spatial gradient $|\nabla u|$.

Since movie analysis theory is very recent and computationally heavy (but not desperate), it has not yet received technological applications. The most promising applications at hand are detection of hidden trajectories and denoising of movies. Both are done in the same way, by simply applying the equation to a movie at some small scale. As can be deduced from the equation itself, a trajectory will be easily eliminated if either its acceleration is high or the moving object is small (and therefore with high curvature). The multiscale analysis acts like a sieve, by removing first all the most erratic trajectories, and leaving as $t \to +\infty$ only the Galilean trajectories. (Which, by the Galilean invariance principle, remain unaltered.) This phenomenon is illustrated in Figures 13 and 14.

9. Invariance and stability requirements for numerical schemes of the AMSS model

In order to compare schemes, we shall fix in this section a list of formal requirements which derive from the axiomatic analysis developed above. We

Fig. 13. Eliminating parasites in a real movie (Author: F. Guichard). We present (up) a sequence of three images of a real movie where a man is talking and moving his head. Notice that in the second image we have added two black rectangles. In the middle and down, we present two states of the multiscale analysis of the above sequence.

first have the properties which imply convergence of the schemes towards the equations, and second the invariance properties of the equations themselves. All schemes will be defined as the iteration of some nonlinear discrete filter, which we call T. We set $u^n = T^n u_0$, where u_0 is the (discretized) initial image. All schemes T will depend on a scale step Δt and a space step Δx.

We shall consider different discrete representations of u^n (see Figure 1):

- The classical discrete representation on a grid $u^n_{i,j} = u^n(i,j)$, with $1 \leq i \leq N$, $1 \leq j \leq N$. The image is the union of the squares centred

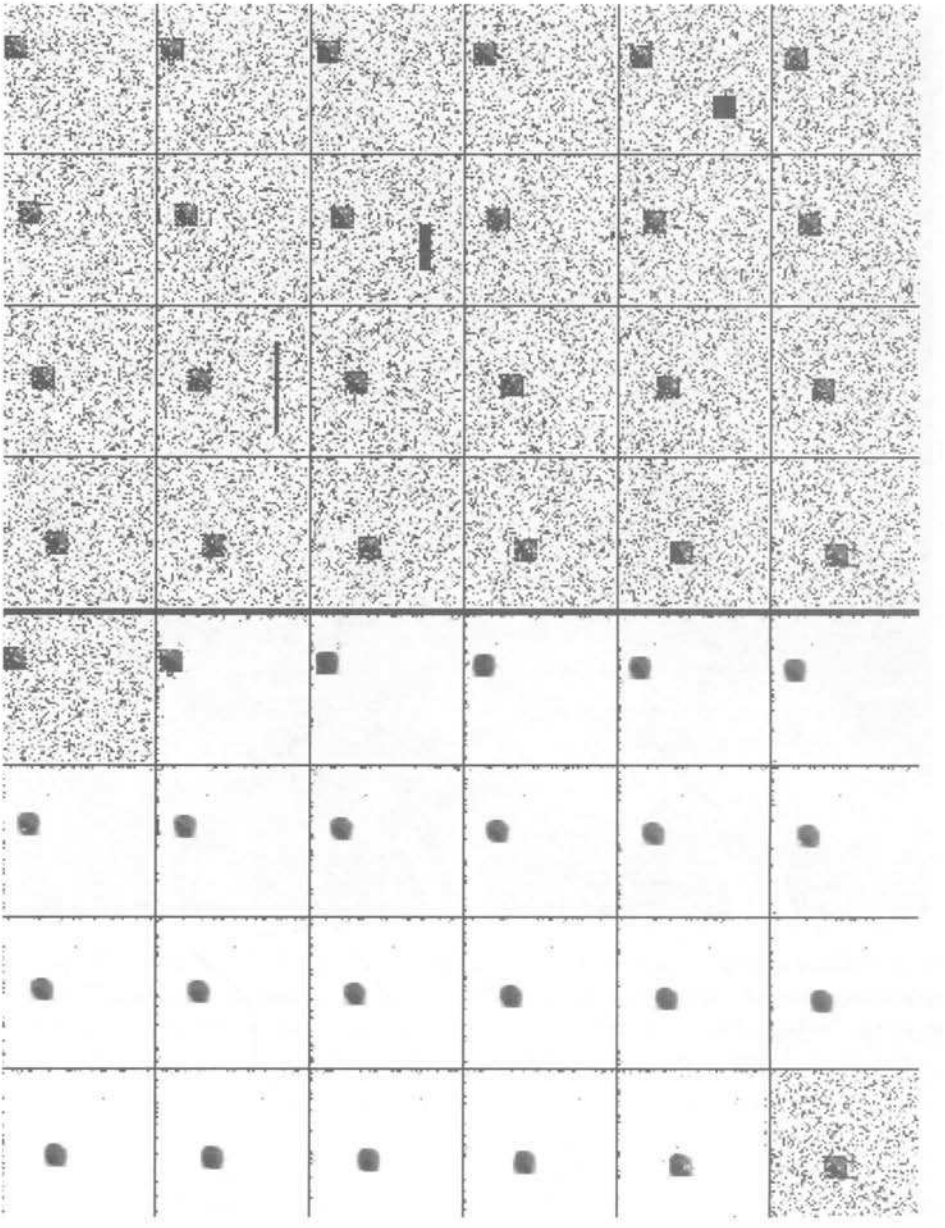

Fig. 14. Movie example. (Author: F. Guichard). Up: original noisy movie of a square and a point in uniform motion, 24 images 64×64. Down: the same movie analysed at a small scale, $q = 0.3$. Spurious trajectories due to noise are eliminated. Only the 'true' trajectories remain: the square and the point.

at the points (i, j), and the brightness in each square is constant: $u(i, j)$. Each one of the squares is called *pixel* (for 'picture element').

- The representation of u as a collection of all its Jordan level curves. In order to define these curves, one considers the level sets $X_a = \{u(x) \geq a\}$, where a admits, for technological reasons, 255 values. Thus each X_a is a finite union of squares in the initial image u_0. By elementary planar topology, its boundary ∂X_a is a finite union of Jordan curves, which we call 'level curves'. Conversely, in this discrete framework, u_0 can easily be reconstructed from all its Jordan level curves. So we call this numerical representation the 'level curve' representation. By the previous sections, we know that it is equivalent to applying AMSS to the initial image or to running ASS on each level curve. But of course, these are very different ways of handling an image in practice.
- We have now to say how we represent each level curve $C(s)$. A first solution is to discretize it as a polygon $C(i), i = 1, \ldots, N$ where the $C(i)$ have real (floating point) values. Then the description of the image will be very precise, with an 'underpixel' accuracy.
- Another solution to represent the level curves is to associate them with their distance function, which is a Lipschitz image. Of course, this representation is heavy since an image can have many level curves (more than pixels) and each becomes associated with a new image. However, the interest in this representation is that, as the distance function is Lipschitz, classical finite difference codes are easy to implement and we avoid the direct programming of a finite element method for curve evolution (Osher and Sethian, 1988). To visualize this representation on the screen, we use the following distance function (see Figure 1).

$$u^n_{i,j} = 128 + \operatorname{dist}((i, j), C) \quad \text{if } (i, j) \text{ is not surrounded by } C(s)$$

and

$$u^n_{i,j} = 128 - \operatorname{dist}((i, j), C) \quad \text{if } (i, j) \text{ is surrounded by } C(s).$$

Let H_λ, R, and A the operators defined by $H_\lambda u(x) = u(\lambda x)$, $Ru(x) = u(Rx)$ where R is an isometry of $\mathbb{R}^2$, $Au(x) = u(Ax)$ where A is a linear application of $\mathbb{R}^2$. Since the schemes to be considered will be very different in structure, we cannot state all following properties in a completely formal way. What they mean in each particular case will be clear in context.

- [Consistency] We shall say that a scheme is consistent if the discrete operator T (which only depends upon Δx and Δt) tends to the differential one (the second member of AMSS or ASS, MCM), when the steps Δx and Δt tend to 0 in a suitable way.
- [Convergence] $\forall t\ T^n u \to T_t u$ a.e. when Δt, $1/n$ and Δx tend to 0 in a suitable way. (T_t denotes the continuous multiscale analysis under consideration.)

- [L^∞ stability] Let c, d be real numbers. If $c \leq u(x) \leq d$ then $c \leq Tu(x) \leq d$.
- [Order Preserving] If $u(x) \geq v(x)$ for all x in $\mathbb{R}^2$, then $(Tu)(x) \geq (Tv)(x)$.
- [Morphology] For all nondecreasing functions h: $Th(u) = h(Tu)$. This means that *the level sets* are handled independently by the discrete operator.
- [Isom. invariance] $TR_\theta = R_\theta T$.
- [Affine invariance] $TA = AT$ for any affine map with determinant equal to 1.

In the following sections, we define schemes and discuss which of the above properties they satisfy. As must be clear from the above list of requirements, as well as from the axiomatic analysis of the first sections, consistency is by no way a sufficient criterion for a good scheme in image analysis. The algorithms we look for must satisfy as much as possible all causality and invariance properties, and consistency, though necessary, should be a consequence rather than a primary requirement.

We shall not only discuss the AMSS model, but also the Mean Curvature Motion (MCM) equation. As we have shown, the MCM model has all the desirable properties except for full affine invariance. (It is, however, Euclidean and zoom invariant.) Therefore, it can prove quite useful in image processing and shape recognition. In addition, as we already have seen, it has been very much used by the image processing community and there is a rich list of attempts to run MCM (or algorithms which *a posteriori* prove to be equivalent to MCM). This will help us when we come to discuss schemes for the AMSS model. Of course, this last model is still much more puzzling, because unless it is written as a PDE, its behaviour is, because of the affine invariance, highly nonlocal (a circle is equivalent to an ellipse of any eccentricity!).

10. Finite difference schemes for the AMSS model

10.1. Mean curvature motion

We start with the 'MCM' equation given by

$$u_t = |\nabla u| \operatorname{curv}(u) = \frac{u_y^2 u_{xx} - 2u_x u_y u_{xy} + u_x^2 u_{yy}}{u_x^2 + u_y^2}.$$

In order to discretize this equation by finite differences we shall introduce an explicit scheme which uses a fixed stencil of $3 * 3$ points to discretize the differential operators. For simplicity, we assume that the spatial increment Δx is the same in the x- and y-axes. We approach the first derivatives u_x

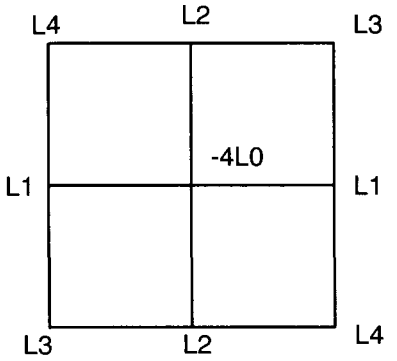

Fig. 15. A 3 × 3 stencil.

and u_y at a point (i,j) of the lattice by using the following linear scheme:

$$\begin{aligned}(u_x)_{i,j} &= \frac{2(u_{i+1,j}-u_{i-1,j})+u_{i+1,j+1}-u_{i-1,j+1}+u_{i+1,j-1}-u_{i-1,j-1}}{4\Delta x} \\ &\quad +\mathcal{O}(\Delta x^2) \\ (u_y)_{i,j} &= \frac{2(u_{i,j+1}-u_{i,j-1})+u_{i+1,j+1}-u_{i+1,j-1}+u_{i-1,j+1}-u_{i-1,j-1}}{4\Delta x} \\ &\quad +\mathcal{O}(\Delta x^2).\end{aligned}$$

Denoting by ξ the direction orthogonal to the gradient of u, one easily sees that $|\nabla u|\,\mathrm{curv}(u)$ is equal to $u_{\xi\xi}$. We have

$$\xi = (-\sin\theta, \cos\theta) = \left(\frac{-u_y}{\sqrt{u_x^2+u_y^2}}, \frac{u_x}{\sqrt{u_x^2+u_y^2}}\right),$$

and

$$u_{\xi\xi} = \sin^2\theta\, u_{xx} - 2\sin\theta\cos\theta\, u_{xy} + \cos^2\theta\, u_{yy}. \tag{10.1}$$

We want to write $u_{\xi\xi}$ as a linear combination of the values of u on the fixed stencil 3*3. Of course, the coefficients of the linear combination may depend on ξ. Because the direction of the gradient (and then ξ) is defined modulo π, by symmetry we must assume that the coefficients of points symmetrical with respect to the central point are the same (see Figure 15).

In order to have consistency, we must find $\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4$, such that

$$\begin{aligned}(u_{\xi\xi})_{i,j} &= \frac{1}{\Delta x^2}(-4\lambda_0 u_{i,j} + \lambda_1(u_{i+1,j}+u_{i-1,j}) + \lambda_2(u_{i,j+1}+u_{i,j-1}) \\ &\quad +\lambda_3(u_{i-1,j-1}+u_{i+1,j+1}) + \lambda_4(u_{i-1,j+1}+u_{i+1,j-1})) \\ &\quad +\mathcal{O}(\Delta x^2).\end{aligned} \tag{10.2}$$

We write

$$u_{i+1,j} = u_{i,j} + \Delta x(u_x)_{i,j} + \tfrac{1}{2}\Delta x^2(u_{xx})_{i,j} + \mathcal{O}(\Delta x^3),$$

and the same relation for the other points of the stencil. By feeding (10.2) with these relations and by using relation (10.1), we obtain four relations between our five coefficients

$$\begin{cases} \lambda_1(\theta) = 2\lambda_0(\theta) - \sin^2\theta, \\ \lambda_2(\theta) = 2\lambda_0(\theta) - \cos^2\theta, \\ \lambda_3(\theta) = -\lambda_0(\theta) + 0.5(\sin\theta\cos\theta + 1), \\ \lambda_4(\theta) = -\lambda_0(\theta) + 0.5(-\sin\theta\cos\theta + 1). \end{cases} \tag{10.3}$$

There remains one degree of freedom for our coefficients given by the choice of $\lambda_0(\theta)$. We shall choose $\lambda_0(\theta)$ following the stability and geometric invariance criteria. Denoting by $u_{i,j}^n$ an approximation of $u(i\Delta x, j\Delta x, n\Delta t)$ we can write our explicit scheme as

$$u_{i,j}^{n+1} = u_{i,j}^n + \Delta t(u_{\xi\xi}^n)_{i,j}. \tag{10.4}$$

Note that this scheme can be rewritten as

$$u_{i,j}^{n+1} = \sum_{k,l=-1}^{1} \alpha_{k,l} u_{i+k,j+l}^n$$

where $\alpha_{k,l}$ satisfy $\sum_{k,l=-1}^{1} \alpha_{k,l} = 1$.

The following obvious lemma shows a general condition for having [L^∞ stability] in this type of scheme:

Lemma 3 Let a finite difference scheme be given by

$$T(u)_{i,j} = \sum_{k,l=-1}^{1} \alpha_{k,l} u_{i+k,j+l},$$

where $\alpha_{k,l}$ satisfy

$$\sum_{k,l=-1}^{1} \alpha_{k,l} = 1.$$

Then the scheme satisfies [L^∞ stability] if and only if $\alpha_{k,l} \geq 0$ for any k, l.

Proof. If $\alpha_{k,l} \geq 0$ for any k, l, set $\min = \inf_{i,j}\{u_{i,j}\}$, $\max = \sup_{i,j}\{u_{i,j}\}$ and take a point (i, j). Then [L^∞ stability] follows from the inequality:

$$\min = \sum_{k,l=-1}^{1} \alpha_{k,l} \min \leq \sum_{k,l=-1}^{1} \alpha_{k,l} u_{i+k,j+l} = (Tu)_{i,j} \leq \sum_{k,l=-1}^{1} \alpha_{k,l} \max = \max.$$

On the other hand, if there exists $\alpha_{k_0,l_0} < 0$ then choosing u and (i, j) such

that $u_{i+k_0,j+l_0} = \min$ and $u_{i+k,j+l} = \max$ for any other k, l, we obtain

$$(Tu)_{i,j} = \sum_{k\neq k_0, l\neq l_0}^{1} \alpha_{k,l} \max + \alpha_{k_0,l_0} \min = \max + \alpha_{k_0,l_0}(\min - \max) > \max$$

and therefore [L^∞ stability] is violated.

Following this lemma, in order to have [L^∞ stability] in the scheme (10.4) we must seek for λ_0 such that $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$ and $(1 - 4\lambda_0/\Delta x^2) \geq 0$. Unfortunately, because of the relations between our coefficients, it is impossible to obtain these relations, except for particular values of $\theta = (0, \frac{1}{4}\pi, \frac{1}{2}\pi, \ldots)$. Indeed, We remark that for θ in $[0, \frac{1}{4}\pi]$,

$$\lambda_1 \geq \lambda_2 \quad \text{and} \quad \lambda_3 \geq \lambda_4.$$

But

$$\lambda_2(\theta) \geq 0 \Rightarrow \lambda_0(\theta) \geq \tfrac{1}{2}\cos^2(\theta)$$

$$\lambda_4(\theta) \geq 0 \Rightarrow \lambda_0(\theta) \leq \tfrac{1}{2}(1 - \sin\theta\cos\theta).$$

So, we cannot find $\lambda_0(\theta)$ satisfying both inequalities, since

$$\tfrac{1}{2}\cos^2\theta \geq \tfrac{1}{2}(1 - \sin\theta\cos\theta).$$

Then, if we choose $\lambda_0(\theta) \geq \frac{1}{2}\cos^2\theta$ we have $\lambda_4(\theta)$ very negative. If we take $\lambda_0(\theta) \leq \frac{1}{2}(1 - \sin\theta\cos\theta)$ we obtain $\lambda_2(\theta)$ very negative. We prefer to choose λ_0 between both functions, and then to have λ_2 and λ_4 negative, but slightly (see Figure 14).

On the other hand, if we impose on λ_0 the following geometrical requirements:

(i) Invariance by rotation of angle $\frac{1}{2}\pi$

$$\lambda_0(\theta + \tfrac{1}{2}\pi) = \lambda_0(\theta).$$

(ii) Pure diffusion in the case $\theta = 0, \frac{1}{2}\pi, \ldots$

$$\lambda_0(0) = 0.5.$$

This condition implies that $\lambda_2(0) = \lambda_3(0) = \lambda_4(0) = 0$.

(iii) Pure diffusion in the case $\theta = \frac{1}{4}\pi, \frac{3}{4}\pi, \ldots$

$$\lambda_0(\tfrac{1}{4}\pi) = 0.25.$$

This condition implies that $\lambda_1(\frac{1}{4}\pi) = \lambda_2(\frac{1}{4}\pi) = \lambda_4(\frac{1}{4}\pi) = 0$.

(iv) Symmetry with respect to the axes $i + j$ and $i - j$,

$$\lambda_0(\tfrac{1}{2}\pi - \theta) = \lambda_0(\theta).$$

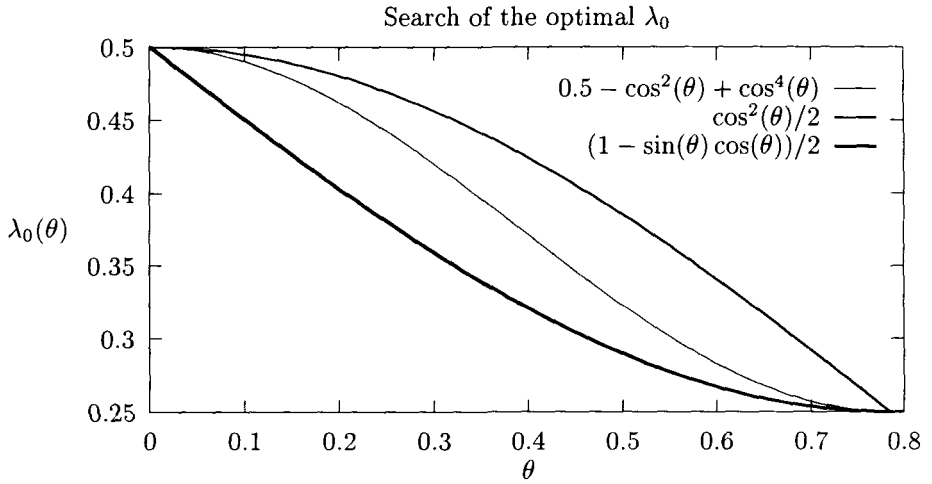

Fig. 16.

We remark that, by the above conditions, it is sufficient to define the function $\lambda_0(\theta)$ in the interval $[0, \frac{1}{4}\pi]$ because it can be extended by periodicity elsewhere.

We have tested two choices for the function $\lambda_0(\theta)$ using the trigonometric polynomials as basis. The first one corresponds to an average of the boundary functions:

$$\lambda_0(\theta) = \tfrac{1}{4}(\cos^2\theta + 1 - \sin\theta\cos\theta). \tag{10.5}$$

As we shall see this choice is well adapted to the 'affine curvature motion' equation. However, if we extend this function by periodicity, the extended function is not smooth at $\frac{1}{4}\pi$. If we seek for a smooth function for $\lambda_0(\theta)$, we must impose $\lambda_0'(0) = \lambda_0'(\frac{1}{4}\pi) = 0$. The simplest polynomial, of degree as small as possible, satisfying the above conditions, and between both boundary functions is

$$\lambda_0(\theta)) = 0.5 - \cos^2\theta\sin^2\theta. \tag{10.6}$$

We deduce the values of the other λs using (10.3). For instance with the above choice of $\lambda_0(\theta)$ we have

$$\begin{cases} \lambda_1(\theta) = \cos^2\theta(\cos^2\theta - \sin^2\theta), \\ \lambda_2(\theta) = \sin^2\theta(\sin^2\theta - \cos^2\theta), \\ \lambda_3(\theta) = \cos^2\theta\sin^2\theta + 0.5\sin\theta\cos\theta, \\ \lambda_4(\theta) = \cos^2\theta\sin^2\theta - 0.5\sin\theta\cos\theta. \end{cases}$$

We have tested this scheme in a workstation and we have noticed that if we impose a 'natural' stability condition such as

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2},$$

then the algorithm has good behaviour and remains stable in the sense that there exists experimentally a (small with respect to 255) $\epsilon > 0$ such that for any $n \in \mathbb{N}$ and (i,j),

$$-\epsilon + \inf_{i,j}\{u^0_{i,j}\} \leq u^n_{i,j} \leq \sup_{i,j}\{u^0_{i,j}\} + \epsilon.$$

10.2. The AMSS model

We will use the ideas developed in the above section. We rewrite the AMSS equation (1.1) as

$$u_t = (|\nabla u|^3 \operatorname{curv}(u))^{1/3} = (u_y^2 u_{xx} - 2u_x u_y u_{xy} + u_x^2 u_{yy})^{1/3}. \tag{10.7}$$

We remark that

$$|\nabla u|^3 \operatorname{curv}(u) = |\nabla u|^2 u_{\xi\xi}$$

where ξ corresponds to the direction orthogonal to the gradient. Therefore, in order to discretize this operator, it is enough to multiply the discretization of $u_{\xi\xi}$ presented in the above section by $|\nabla u|^2$. We choose $\lambda_0(\theta)$ given by (10.5) because it corresponds to a trigonometric polynomial of degree two and then multiplying it by $|\nabla u|^2$ the coefficients

$$\eta_i = |\nabla u|^2 \lambda_i, i = 0,1,2,3,4,$$

are polynomials of degree two with respect to u_x and u_y. Indeed, we obtain for $\theta \in [0, \frac{1}{4}\pi]$

$$\begin{aligned}(|\nabla u|^2 u_{\xi\xi})_{i,j} \;=\; & \frac{1}{\Delta x^2}(-4\eta_0 u_{i,j} + \eta_1(u_{i+1,j} + u_{i-1,j}) + \eta_2(u_{i,j+1} + u_{i,j-1}) \\ & +\eta_3(u_{i-1,j-1} + u_{i+1,j+1}) + \eta_4(u_{i-1,j+1} + u_{i+1,j-1})) \\ & +\mathcal{O}(\Delta x^2),\end{aligned}$$

where $\eta_0, \eta_1, \eta_2, \eta_3, \eta_4$ are given by

$$\begin{cases} \eta_0 = 0.25(2u_x^2 + u_y^2 - u_x u_y), \\ \eta_1 = 0.5(2u_x^2 - u_y^2 - u_x u_y), \\ \eta_2 = 0.5(u_y^2 - u_x u_y), \\ \eta_3 = 0.25(u_y^2 + 3u_x u_y), \\ \eta_4 = 0.25(u_y^2 - u_x u_y). \end{cases}$$

Finally, the finite difference scheme for the AMSS equation is

$$u^{n+1}_{i,j} = u^n_{i,j} + \Delta t(|\nabla u^n|^2 u^n_{\xi\xi})^{1/3}_{i,j}. \tag{10.8}$$

We have tested this algorithm and we have noticed that in this case the condition for the experimental stability (in the sense presented in the above subsection) is

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{10}.$$

Remark The finite difference schemes that are presented satisfy [Consistency] and we conjecture [Convergence]. Morphological invariances are obtained asymptotically by taking a little time step Δt. The experimental results presented in Figures 5 and 7 have been obtained by using these schemes with $\Delta x = 1$ and $\Delta t = 0.1$ in the case of MCM and $\Delta t = 0.01$ in the case of affine curvature motion. One has to take Δt that small because unless experimental stability is achieved with $\Delta t \leq 0.1$, the experimental affine invariance needs $\Delta t < 0.05$.

11. Morphological (set evolution) schemes

11.1. A theoretical scheme

In this subsection, we discuss theoretical (only discretized in time) schemes inspired by the Mathematical Morphology School. These schemes will be, in contrast to the above presented finite difference schemes, fully geometrically invariant and stable. Now, as we shall see in a latter subsection, their implementation on a grid is problematic and they will only be practically implementable by working with the 'Jordan level curve' representation of the images (Section 11.3). Denote by C a set of convex sets which is stable either by isometries or by linear maps. For example, we can take for C the set of disks, ellipses or triangles,.... For all $t \geq 0$, let us set

$$C_t(x) = \{B \in C,\ \text{area}(B) = t^{3/2},\ \text{ and } x \text{ is the barycentre of } B\}.$$

Then we define two operators

$$\begin{aligned} IS_t(u)(x) &= \inf_{B \in C_t(x)} (\sup_{y \in B}(u(y)), \\ SI_t(u)(x) &= \sup_{B \in C_t(x)} (\inf_{y \in B}(u(y)). \end{aligned}$$

For example, if we take for C the set of disks, then IS_t is by definition the 'dilation operator' with radius t, and SI_t is the 'erosion' with radius t, and these are the basic operators of the mathematical morphology discussed in Section 3 (see Serra (1982), Maragos (1987)). If we impose that C is stable by all affine maps (for instance C can be the set of all ellipses), we obtain 'special' morphological operators which clearly are affine invariant, because C_t is unchanged under any affine map with determinant 1.

Consider the schemes

$$\text{(i)} \quad u^{n+1} = IS_{\Delta t}(u^n),$$

$$\text{(ii)} \quad u^{n+1} = SI_{\Delta t}(u^n).$$

It is proved in Guichard *et al.* (1993) that scheme (i) is consistent with :he equation

$$\frac{\partial u}{\partial t} = |\mathrm{D}u|(\mathrm{curv}^+(u))^{1/3}$$

vhile scheme (ii) is consistent with the equation

$$\frac{\partial u}{\partial t} = |\mathrm{D}u|(\mathrm{curv}^-(u))^{1/3}.$$

In order to approximate equation (1), one can alternate affine closing and ffine opening and one is led to the (consistent with AMSS) scheme (first nnounced in Cohignac *et al.* (1993b)):

$$\text{(iii)} \quad \begin{cases} u^{n+1/2}(x) &= \inf\limits_{B \in C_t(x)} (\sup\limits_{y \in B}(u^n(y))), \\ u^{n+1}(x) &= \sup\limits_{B \in C_t(x)} (\inf\limits_{y \in B}(u^{n+\frac{1}{2}}(y))). \end{cases} \tag{11.1}$$

Let us end with a variant of the schemes (i), (ii) and (iii) which is consstent with MCM. Catté *et al.* (1993) proved that if one takes for C the :t of all segments in the plane, then scheme (iii) converges towards the ICM model. We shall explain in the next sections why this type of scheme, iough fully invariant in theory, can hardly be implemented on a fixed grid. owever, it can, as we shall later see, be well adapted to the 'Jordan level irve' representation of images.

1.2. Iterated median filters and curvature motion

orphological schemes for MCM have been proposed by Koenderink and an Doorn (1987) and Merriman *et al.* (1992). They define a weighted edian filter which weighs the contribution of points y close to a point x :cording to a Gaussian law of distance. So the algorithm can be rewritten vhen applied to a characteristic function $\tilde{u}^0$ of a level set).

Veighted median filter

Let $\tilde{u}^0$ be a binary image:

$$\begin{cases} \tilde{u}^0(x) = 1, & \text{if } x \text{ belongs to the level set } E \\ \tilde{u}^0(x) = 0, & \text{else.} \end{cases}$$

We solve the heat equation, with initial datum $\tilde{u}^n$, for a small time Δt, by a convolution with a Gauss function. We obtain a new function v^n. We set (median filter)

$$\begin{cases} \tilde{u}^{n+1}(x) = 1, & \text{if } v^n(x) \geq \frac{1}{2} \\ \tilde{u}^{n+1}(x) = 0, & \text{else .} \end{cases}$$

4 We turn back to step 2.

Barles and Georgelin (1992), proved that this theoretical scheme is consistent with and convergent to the MCM evolution of the boundary of E. Earlier works by Yuille (1988) and Mascarenhas (1992) have also proved the consistency. However, we shall see that its implementation on a fixed grid is problematic. A simpler and obvious discrete implementation of the preceding weighted iterated filter on a grid was proposed long ago by the Mathematical Morphology School. If A is a finite set of real numbers, we call $\text{med}(A) = a$ any real number such that

$$\text{Card}\{r \in A, r \geq a\} = \text{Card}\{r \in A, r \leq a\}.$$

Of course, the possible as make an interval and if we want to specify a, we take for a the middle point of this interval. Then a classical morphological filter, with well-known 'denoising' properties is the original 'iterated median filter' (Matheron, 1975; Serra, 1982; Maragos, 1987)

$$\text{(iv)} \qquad u^{n+1}(x) = \text{med}_{y \in B}(u^n(y)).$$

This filter (where B is a circular fixed stencil around x) can be viewed as a simplification of the Gaussian-weighted median filter, where the Gauss function has been replaced by the characteristic function of a disk.

Let us now pass to discrete versions on a grid. Among the schemes discussed in this article, the morphological schemes (i)–(iv) discretized on a fixed grid are the only ones to be both morphological and order preserving (whatever the discretization of sup, inf, med on the grid is.) They are also consistent. Now, they do not have all geometrical properties. They only satisfy the geometrical invariances of the grid and are by no way scale invariant or affine invariant. So in practice they prove to be useless schemes when discretized on a fixed grid: because of the lack of rotational invariance, they make corners appear in the shapes with sides parallel to the principal directions of the stencil. However, since these schemes are the only ones to be fully morphological and monotonic, it is very desirable to have them implemented in one way or the other. There is, however, no cheap solution: either one uses grid refinements or one adapts these theoretical schemes to the Jordan level curve representation (Section 11.3).

Let us say a little more about the consistency of these schemes when discretized on a grid. Unless consistency is proved, in practice they are not: $\text{curv}(u)$ is incorrectly calculated. If we use a fixed stencil, it will simply be computed as equal to zero. Indeed, it is well known in computer graphics that the curvature of a discretized circle is impossible to compute accurately on a fixed window unless this window is very large: if e.g. the radius is equal to 20, the window should have a width of at least 12 pixels to yield a roughly correct guess of the curvature. If the window is, say, 5 pixels wide,

the curvature will simply be implicitly computed as equal to 0 by a median filter and so we shall have $u^{n+1} = u^n$, which stops the evolution. In other words, when Δt tends to 0, the scheme tends to an equation

$$\partial u/\partial t = |\mathrm{D}u| F(\mathrm{curv}(u))$$

where F is defined by

$$F(s) = \begin{cases} s, & \text{if } s \geq \mathrm{C}\,, \\ 0, & \text{if } -\mathrm{C} \leq s \leq \mathrm{C}\,, \\ s, & \text{if } s \leq -\mathrm{C}\,, \end{cases}$$

and C is a constant depending on the grid step. So consistency is in practice not true; there is a thresholding effect on the curvature. A striking example of the lack of consistency can be seen in Figure 17 the iterated weighted median filter (with an $8 * 8$ stencil) is applied to both a grey level image (left down) and a binary image (left up) which is one of its level sets. The right-hand images display what happens when $n \to \infty$: the displayed images remain steady under the median filter because the curvature of all level sets is implicitly computed as equal to 0.

11.3. Geometrical curve evolution schemes

Curve sampling problems We have seen in Section 5 MCM and the AMSS can be reformulated as intrinsic heat equations,

$$\dot{C}(t,s) = \partial^2 C(t,s)/\partial s^2$$

where s stands for the length parameter along the curve in the first case and for the 'affine length' parameter in the second. This formulation suggests a computationally cheap implementation of MCM, which was developed by Mackworth and Mocktharian (1992).

Intrinsic heat equation

- Discretize C_0 as a polygon with vertices $C(i), i = 1, \ldots, N$. Parametrize this polygon with length s.
- C_n being given, together with its length parametrization, convolve it with a discretized Gaussian filter with (small) variance δt: $\tilde{C}_{n+1}(i) = (G_{\delta t} * C_n)(i)$.
- Reparametrize $\tilde{C}_{n+1}$ with length, which yields $C_{n+1}(i)$ and go back to the previous step.

This type of algorithm is very accurately scale invariant and isotropic and satisfies, in practice, the *local shape inclusion principle* when δt is chosen small enough (close to the pixel size). The consistency of such a scheme is easily proved but the convergence is, as far as we know, unproved. One could be willing to extend this kind of scheme to the ASS model: it suffices

to compute the affine length instead of the Euclidean one in the last step. This is, however, *impossible*. Indeed, it is easily seen that any C^2 curve (for which affine length is well defined) can be approached in C^1 by C^1 curves with affine length tending to zero. This is due to the fact that any straight line has zero affine length. Conversely, every polygon can be approximated by C^2 curves whose affine length tends to zero. So the computation of the affine length makes no computational sense in the irregular context of shapes and images. There has, however, been some attempt to base schemes for the ASS model on discrete versions of affine length: Bruckstein *et al.* (1992) define the affine length of a polygon as the number of vertices and iterate a convolution kernel on it. Therefore, they obtain a Mackworth–Mockhtarian scheme, but the evolution is highly dependent on the sampling of the polygon.

A geometrical scheme In order to avoid the above discussed problems, we shall define an easy to implement version of the alternate scheme 11.1.

Let C^0 be a Jordan curve, X^0 the set enclosed by this Jordan curve, and x a point of $C^0 = \partial X^0$. Let K_t be an affine invariant and translation invariant set of convex sets (e.g. all ellipses,...) with area equal to $t^{3/2}$. Then, we define the *t-affine opening* $O_t X^0$ as the set of the barycentres of the elements of K_t contained in X^0. In the same way, we define the *affine closing* $F_t X^0$ as the complementary of the opening of X^c. And we consider the alternate scheme

$$X^{n+1/2} = O_t X^n,$$

$$X^{n+1} = F_t X^{n+1/2}.$$

This is the set evolution scheme associated with the convergent image evolution scheme 11.1. Indeed, it is easy to check that u^n obeys 11.1 if and only if its level sets obey the preceding scheme.

We shall now define a practical scheme associated with this theoretical scheme. We notice that, in order to define the opening, we can restrict the sets in K_t to be *chord sets*, that is, sets bounded on the one side by the boundary of the set and on the other by a chord. This leads to a computationally easier variant.

If X is a set enclosed by a Jordan curve $C(s)$, let $(C(s_0), C(s_1))$ be a 'chord', that is, a segment joining two points of C with $s_0 < s_1$. We define the Δt-opening as follows. For every s_0, we consider the chords $(C(s_0), C(s))$ and we call $A(s)$ the area of X between the chord and the curve $C(\sigma)$ for $s_0 \leq \sigma \leq s$. The function $A(s)$ is nondecreasing and we take $s_1 = \sup\{s, A(s) \leq \Delta t^{3/2}\}$. We call $CH(s_0)$ the chord set associated with the chord $(C(s_0), C(s_1))$. It has an area less than or equal to $\Delta t^{3/2}$.

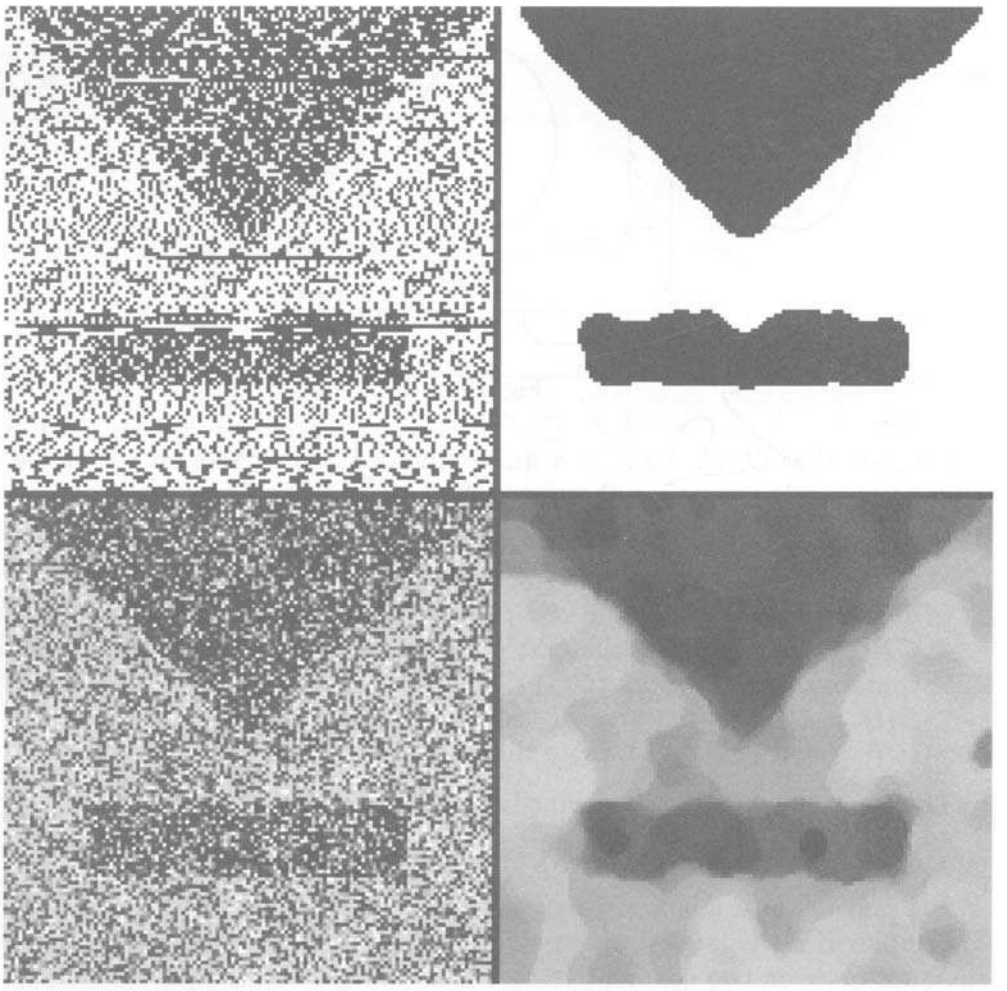

Fig. 17. The iterated weighted median filter (with an $8*8$ stencil) is applied to both a grey level image (left down) and a binary image (left up) which is one of its level sets. The right-hand images display what happens when $n \to \infty$: the displayed images remain steady under the median filter because the curvature of all level sets is implicitly computed as equal to 0.

Then the Δt-opening of X is defined by

$$O_{\Delta t}X = X \setminus (\cup_s CH(s)).$$

In the same way, we can define the closing $F_{\Delta t}X$, by $F_{\Delta t}X = (O_{\Delta t}X^c)^c$. Of course, $O_{\Delta t}X$ may be no Jordan set. Now, if we discretize X and C, we

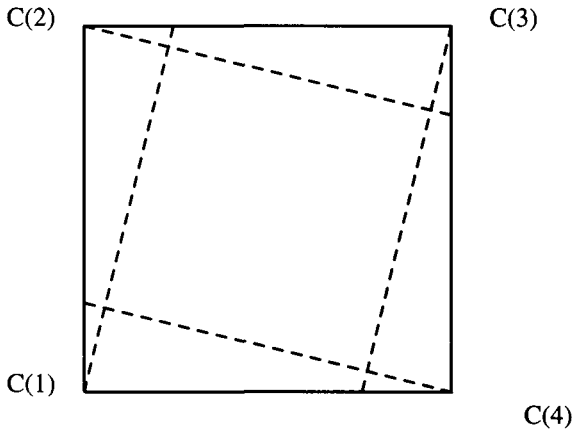

Fig. 18.

can ensure that $O_{\Delta t}X$ remains a finite union of Jordan sets on which the algorithm can be iterated. We now define a practical algorithm based on the same principle.

1 $C(i)$, $i = 1 \cdots N$, is a polygon curve enclosing X.

2 For every i, we find the last j such that the area of the chord set $CH(C(i), C(j))$ is less than $\Delta t^{3/2}$. Then, we define a new vertex $\tilde{C}(i)$, contained in the vertex $(C(j), C(j+1))$. $\tilde{C}(i)$ is chosen so that either $\tilde{C}(i) = C(j)$ or the area of the chord set associated with the chord $(C(i), \tilde{C}(i))$ is $\Delta t^{3/2}$. Denote this chord set by $CH(C(i), \tilde{C}(i))$.

The new polygon $O_{\Delta t}C$ is defined as

$$O_{\Delta t}X = X \setminus (\cup_i CH(C(i), \tilde{C}(i)).$$

It is a polygon or a union of polygons. The algorithm for computing $O_{\Delta t}X$ is complex if Δt is large, but very simple if Δt is of the same order as the distance between two consecutive vertices.

We notice that such a method is 'self-sampling' since the number of vertices cannot increase. The final algorithm consists in alternating $O_{\Delta t}$ and $F_{\Delta t}$ as explained at the beginning of this section.

11.4. Comparison and cross-validation of the schemes

To summarize the above presentation, we have rejected several existing schemes and essentially proposed two for the AMSS model: first a finite difference scheme and second a curve evolution scheme (which can in theory be used for handling images but, because of high computational cost, this has not yet been done). In any case, we can compare both schemes on images which are characteristic functions. The match, however, would not

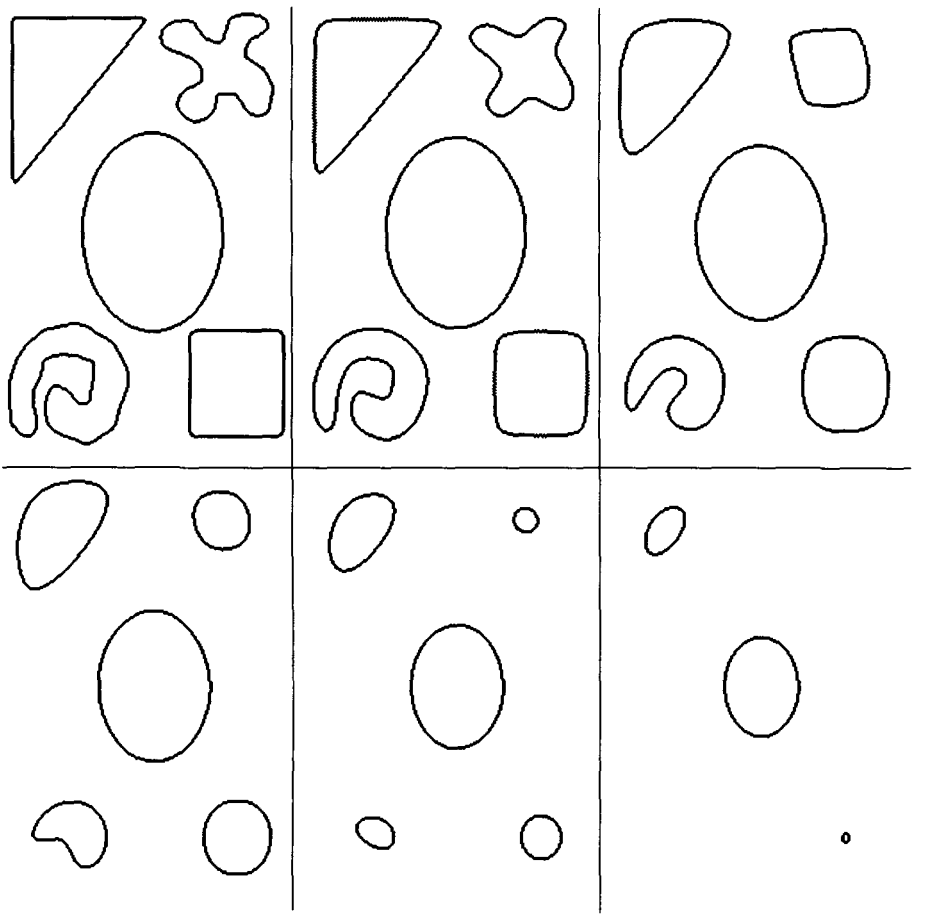

Fig. 19. Curve evolution. From left to right and from up to down: (a) $t = 0$; (b) $t = 9$, $t = 15$, $t = 21$, $t = 25$.

be fair, were the finite difference scheme to be applied to an image which presents abrupt fronts. So, following Cohignac *et al.* (1993a) we used the 'morphologization' of the numerical scheme for comparison. We know that the morphological invariance can be restored by using the 'morphological school' idea of running the algorithm separately on each level set of the image and then reconstructing it. Moreover, a very efficient numerical idea of Osher and Sethian (1988) is to run the algorithm not on a set, but on some Lipschitz function having this set as a level set (see Section 9), therefore allowing good behaviour of numerical schemes by avoiding explicit front tracking (Figure 1(c)). When done in this framework, Cohignac *et al.* (1993a) proved a full cross-validation of the above selected finite difference scheme and curve evolution scheme: complicated objects such as a spiral

evolve in exactly the same way when handled with schemes with so different an implementation (see Cohignac *et al.* (1993a) for more details).

12. Conclusions

In this review, we have described with the unified formalism of *multiscale analysis* more than 20 acknowledged theories of image, shape and texture analysis. Following recent mathematical work, we have shown how axiomatic analysis reduced these theories to not more than one model for standard multiscale analysis, the AMSS model, while a large number of segmentation multiscale devices were reduced to a single variational algorithm. The same analysis applied to movies also led to a single model, with applications to motion analysis and movie denoising. As an application of the unified theory, we have shown how it permitted a rigorous axiomatic and experimental discussion of a psychophysical theory such as Julesz texture discrimination theory.

The second part of our review deals with algorithms implementing a highly geometrically invariant equation, the AMSS model, as well as its Euclidean version, MCM. We have discussed several algorithms directly proposed for these equations, as well as procedures which, although not directed towards solving any PDE, happen to be discrete versions of the MCM.

Unless we have considered a wide range of theories and algorithms, the conclusion of the study considerably narrows the idea of what has been aimed and attained by the image analysis research. Indeed, under the variety of methods lies essentially one aim for the multiscale analysis: the computation of *multiscale curvature, multiscale orientation, multiscale affine curvature, and multiscale apparent velocity.* **And nothing else!** This is due to the obvious invariance requirements, which do not leave space for any other differential operator. In the same way, the invariance requirements practically specify the forms the algorithms must take. In any case, image processing has contributed to mathematics by proposing new variational problems and new methods for solving them (for the segmentation problem). In the case of smoothing multiscale analyses, it has brought a new equation, the AMSS model, as well as new variants of schemes for the MC equation.

Appendix A. The 'fundamental theorem' of image analysis

Fundamental Theorem 2 If an image multiscale analysis T_t is causal and regular then $u(t,x) = (T_t u)(x)$ is a viscosity solution of (3.1), where the function F, defined in the regularity axiom, is nondecreasing with respect to its first argument $\mathrm{D}^2 u$. Conversely, if u_0 is a bounded uniformly continuous image, then equation (3.1) has a unique viscosity solution.

Proof. Assume for simplicity that $u(t,x)$ is C^2 in the neighbourhood of (t,x). Then, we have

$$u(t,y) = u(t,x) + (\mathrm{D}u, y-x) + \tfrac{1}{2}\mathrm{D}^2u(y-x,y-x) + o(|y-x|^2).$$

Let $\epsilon > 0$ and Q_ϵ a quadratic form given by

$$Q_\epsilon(y) = u(t,x) + (\mathrm{D}u, y-x) + \tfrac{1}{2}\mathrm{D}^2u(y-x,y-x) + \epsilon \cdot |y-x|^2.$$

Then, in a neighbourhood of (t,x)

$$Q_\epsilon(y) < u(t,y) < Q_\epsilon(y) \qquad \text{for } y \neq x,$$

and by using the causality principle we obtain

$$(T_{t+h,t}Q_{-\epsilon})(x) \leq (T_{t+h,t}u(t))(x) \leq (T_{t+h,t}Q_\epsilon)(x).$$

On the other hand, we also have

$$Q_{-\epsilon}(x) = Q_\epsilon(x) = u(t,x) = (T_{t,t}Q_{-\epsilon})(x) = (T_{t,t}Q_\epsilon)(x).$$

Therefore we deduce from the above relations

$$\begin{aligned}\frac{\partial(T_{t+h,t}Q_{-\epsilon})}{\partial h}(x) &\leq \liminf \frac{T_{t+h,t}u(,x) - T_{t,t}u(x)}{h} \\ &\leq \limsup \frac{T_{t+h,t}u(,x) - T_{t,t}u(x)}{h} \leq \frac{\partial(T_{t+h,t}Q_\epsilon)}{\partial h}(x).\end{aligned}$$

By using the regularity principle and the continuity of the function F, and taking $\epsilon \to 0$ we obtain that $u(t,x)$ satisfies equation (3.1). Finally, in order to obtain that $F(A,p,c,x,t)$ is nondecreasing with respect to A, we notice that if $A \leq B$ then the quadratic forms

$$\begin{aligned}Q_A(y) &= \tfrac{1}{2}(A(y-x),(y-x)) + (p,y-x) + c, \\ Q_B(y) &= \tfrac{1}{2}(B(y-x),(y-x)) + (p,y-x) + c,\end{aligned}$$

satisfy $Q_A(y) \leq Q_B(y)$ and by using an obvious adaptation of the above proof, we obtain $F(A,p,c,x,t) \leq F(B,p,c,x,t)$ if $A \geq B$.

To simplify the exposition, we have showed that equation (3.1) is true in the case where u is a C^2 function. By using the same ideas in the framework of viscosity solutions (see Crandall *et al.* (1991)), it is possible to show that equation (3.1) is true in the sense of viscosity solutions for any $u(t,x)$ uniformly continuous satisfying the causality and regularity principles. The fact that if $u_0(x)$ is a bounded uniformly continuous function, equation (3.1) has a unique viscosity solution is proved in Chen *et al.* (1991), Crandall *et al.* (1991) and Evans and Spruck (1991).

Appendix B. Proof of the scale normalization lemma

Normalization Lemma (Normalization of scale.) Assume that $t \to T_t$ is a one-to-one family of operators satisfying [affine invariance]. Then the

function $t'(t,B)$ only depends on t and $|\det B|$: $t'(t,B) = t'(t,|\det B|^{1/2})$ and increases with respect to t. Moreover, there exists an increasing differentiable rescaling function σ: $[0,\infty] \to [0,\infty]$, such that

$$t'(t,B) = \sigma^{-1}(\sigma(t)|\det B|^{1/2})$$

and if we set $S_t = T_{\sigma^{-1}(t)}$ we have $t'(t,B) = t|\det B|^{1/2}$ for the rescaled analysis.

Proof. First we notice that for any linear transforms B and C and any t one has the semigroup property

$$\text{(i)} \qquad t'(t,BC) = t'(t'(t,B),C).$$

Indeed, we have $BCT_{t'(t,BC)} = T_tBC = BT_{t'(t,B)}C = BCT_{t'(t'(t,B),C)}$. The map which associates T_t with t being one to one, this implies the stated relation.

Next, we show that

$$\text{(ii)} \qquad t'(t,A) \text{ increases with respect to } t.$$

Let us prove that $t'(t,A)$ is one to one with respect to t for any A. Indeed, if not, there would be some A and some (s,t) such that $t'(t,A) = t'(s,A)$. Thus $T_tA = AT_{t'(t,A)} = AT_{t'(s,A)} = T_sA$ and therefore $t = s$ because T_t is one to one. Notice that this implies, in particular, that $t'(0,A) = 0$. Therefore, since $t'(t,A)$ is nonnegative (by definition), one to one and continuous with respect to t, we can deduce that it is increasing with respect to t.

Moreover $t'(t,A)$ satisfies

$$\text{(iii)} \qquad t'(t,R) = t \quad \text{for any orthogonal transform } R.$$

Indeed, let R be an orthogonal transform. Then iterating the formula of (i) we have

$$t'(t'(t'(t'(\dots t'(t,R)\dots,R),R),R) = t'(t,R^n).$$

Remark that there is a subsequence of R^n tending to Id. (Indeed, there is a subsequence R^{n_k} which converges to some H, orthogonal, because the orthogonal group is compact. Therefore, the subsequence $R^{n_{k+1}-n_k}$ converges to Id.) Since there exists a subsequence of R^n tending to Id and since t' is continuous we have for this subsequence $\lim t'(t,R^n) = t'(t,\mathrm{Id}) = t$. Assume by contradiction that $t'(t,R) = t''$ with $t'' < t$ then $t'(t'(t,R),R) = t'(t'',R) \le t'(t,R) = t''$ and by recursion,

$$t'(t,R^n) = t'(t'(t'(t'(\dots t'(t,R)\dots,R),R),R) \le t'' < t.$$

This is a contradiction. Thus $t'(t,R) \ge t$. We prove the converse inequality in the same way and we obtain $t'(t,R) = t$.

We note that any linear transform B of $\mathbb{R}^2$ can be obtained as a product of orthogonal transforms and of linear transforms of the kind $A(\lambda)$: $(x,y) \to$

$(\lambda x, y)$ where λ is nonnegative. We only need to make a singular value decomposition of B: $B = R_1 \mathrm{D} R_2$, where R_1 and R_2 are both orthogonal transforms and D is a transform of the kind $(x, y) \to (\lambda_1 x, \lambda_2 y)$ where λ_i are non negative. Now, it is clear that D can be decomposed as $\mathrm{D} = A(\lambda_1) R A(\lambda_2) R^{-1}$ where R is the orthogonal transform: $(x, y) \to (-y, x)$. Using (i), $t'(t, R_i) = t$, the singular value decomposition and $A(\lambda_1)A(\lambda_2) = A(\lambda_1\lambda_2)$, we obtain

$$t'(t, B) = t'(t, \lambda_1\lambda_2) = t'(t, |\det B|^{1/2}).$$

Using (i) and (ii), we have

$$\text{(iv)} \qquad t'(t, \lambda\mu) = t'(t'(t, \mu)\lambda)$$

for any positive λ and μ. Differentiating this relation with respect to μ at $\mu = 1$ yields

$$\lambda \frac{\partial t'}{\partial \lambda}(t, \lambda) = \frac{\partial t'}{\partial \lambda}(t, 1) \frac{\partial t'}{\partial t}(t, \lambda). \tag{B.1}$$

Choose σ such that $\phi\sigma' = \sigma$ and set

$$t'(t, \lambda) = G(t, \sigma(t)\lambda),$$

where

$$\sigma(t) = \exp\left(\int_1^t \mathrm{d}s/\phi(s) \right).$$

Then the preceding relation (B.1) yields $\partial G/\partial x(x, y) = 0$. Thus $G(x, y) = \beta(y)$ for some differentiable nondecreasing function β. We obtain that $t'(t, \lambda) = \beta(\sigma(t)\lambda)$. Returning to the definition of $\phi(t)$, we have

$$\phi(t) = \partial t'/\partial\lambda(t, 1) = \partial\beta(\sigma(t)\lambda)/\partial\lambda(t, 1)$$

and

$$\phi(t) = \sigma(t)\beta'(\lambda\sigma(t)) = \phi(t)\sigma'(t)\beta'(\sigma(t)\lambda).$$

Thus the derivative of $\beta(\sigma(t))$ is 1 and integrating this last relation between 0 and t yields $\beta(\sigma(t)) = t + \beta(\sigma(0))$. Using the fact that $t'(0, \lambda) = 0$ (which derives from the injectivity of the T_t), we obtain $\beta(\sigma(0)) = 0$ and therefore $t'(t, \lambda) = \sigma^{-1}(\lambda\sigma(t))$. To finish the proof, we set $S_t = T_{\sigma^{-1}(t)}$ and we prove that the affine invariance is true for S_t with $t'(t, \lambda) = \lambda t$. $S_t B = T_{\sigma^{-1}(t)} B = B T_{t'(\sigma^{-1}(t), \lambda)} = B T_{\sigma^{-1}(\lambda\sigma(\sigma^{-1}(t)))} = B T_{\sigma^{-1}(\lambda t)} = B S_{\lambda t}$.

Appendix C. Classification of shape multiscale analyses

Theorem

(i) Under the three principles (pyramidal, local shape inclusion, 'basic'),

the multiscale analysis of shapes is governed by the curvature motion equation

$$\dot{x} = g(t, \mathrm{curv}(x))\vec{n}(x), \tag{C.1}$$

where g is defined by (4.1).

(ii) If the analysis is affine invariant, then the equation of the multiscale analysis is, up to rescaling,

$$\dot{x} = \gamma(t \cdot \mathrm{curv}(x))\vec{n}(x), \tag{C.2}$$

where $\gamma(x) = a \cdot x^{1/3}$ if $x \geq 0$ and $\gamma(x) = b \cdot x^{1/3}$ if $x \leq 0$ and a, b are two nonnegative values.

(iii) If we add that $T_t(X^c) = T_t(X)^c$ [Reverse contrast invariance] then the function g in (i) is odd and we get

$$\dot{x} = (t \cdot \mathrm{curv}(x))^{1/3}\vec{n}(x). \tag{C.3}$$

Proof. (i) Let X be a silhouette and assume that $T_t(X)$ has a boundary which is a C^2 manifold in a neighbourhood of a point x of ∂X. Then it has a curvature κ at point x and we consider a subosculatory and a surosculatory disk, that is, a disk D with curvature $\kappa - \epsilon$ and a disk D′ with curvature $\kappa + \epsilon$, both tangent to the silhouette at x. Applying the same two principles as in the lemma, we see that

$$T_{t+h,t}(\mathrm{D}) \cap B(x,r) \subset T_{t+h,t}(X) \cap B(x,r) \subset T_{t+h,t}(\mathrm{D}') \cap B(x,r).$$

Thus, denoting by $x(t+h)$ the point of $\partial T_{t+h}(X)$ such that $x(t+h) - x(t)$ is parallel to $\vec{n}(x)$, we obtain

$$\rho(t,h,\kappa-\epsilon) - \rho(t,0,\kappa-\epsilon) \leq (x(t+h)-x(t))\cdot\vec{n}(x) \leq \rho(t,h,\kappa+\epsilon) - \rho(t,0,\kappa+\epsilon).$$

Dividing by h and passing to the limit when h tends to 0 yields

$$\frac{\partial \rho}{\partial h}(t,0,\kappa-\epsilon) \leq \liminf \frac{x(t+h)-x(t)}{h} \cdot \vec{n}(x),$$

$$\limsup \frac{x(t+h)-x(t)}{h} \cdot \vec{n}(x) \leq \frac{\partial \rho}{\partial h}(t,0,\kappa+\epsilon).$$

We obtain equation (C.2) by passing to the limit when ϵ tends to 0 and using the fact that $\kappa \to \partial\rho/\partial h(t,0,\kappa)$ is continuous.

(ii) After renormalization, we can use the identity

$$T_{t+h,t}\mathrm{D}_\lambda = \mathrm{D}_\lambda T_{(t+h)\lambda,t\lambda},$$

(where $\mathrm{D}_\lambda = \lambda\,\mathrm{Id}$) and so, we can deduce that the function ρ of the basic principle must satisfy $\rho(t,h,\lambda/r) = \lambda^{-1}\rho(\lambda t, \lambda h, 1/r)$ Therefore, we obtain the relation (after differentiation with respect to h at 0)

$$g(t,\lambda s) = g(\lambda t, s),$$

for any $t > 0$, $\lambda > 0$ and $s \in \mathbb{R}$ Changing t in t/λ and taking $\lambda = 1/t$ we get $g(t,s) = g(1,ts) = \beta(ts)$ for the function β defined as $\beta(x) = g(1,x)$.

On the other hand, we can use the identity $T_{t+h,t}A = AT_{t+h,t}$, where A is the linear transform whose determinant is one,

$$(x,y) \rightarrow (\lambda x, (1/\lambda)y), \quad \lambda > 0.$$

Let us apply this identity to the unit disk Δ. Look at the point $x_0 = (1,0)$ on the boundary of Δ. Then the velocity of x_0 is $\beta(-t)$, and this velocity is transformed into $\lambda\beta(-t)$. Now, look at $A\Delta$. Since $A\Delta$ is an ellipse with curvature $-\lambda^3$ at point Ax_0, the velocity of Ax_0 is $\beta(-t \cdot \lambda^3)$. Using the first identity, we obtain $\beta(-t.\lambda^3) = \lambda\beta(-t)$. Taking $t = 1$, we get $\beta(x) = b \cdot x^{1/3}$ for $x < 0$ ($b = \beta(-1)$). Now, apply the same technique to Δ^c and we get the result $\beta(x) = a \cdot x^{1/3}$ for $x > 0$ ($a = \beta(1)$).

(iii) With the same technique as above we obtain that the function β is odd. □

REFERENCES

L. Alvarez, F. Guichard, P.L. Lions and J.M. Morel (1992a), 'Axioms and fundamental equations of image processing', Report 9216, C.E.R.E.M.A.D.E.,, Université Paris Dauphine, *Arch. Rat. Mech.*, to appear.

L. Alvarez, F. Guichard, P.L. Lions and J.M. Morel (1992b), 'Axiomatisation et nouveaux opérateurs de la morphologie mathematique', *C.R. Acad. Sci. Paris* **315**, 265–268.

L. Alvarez, P.L. Lions and J.M. Morel (1992c), 'Image selective smoothing and edge detection by nonlinear diffusion (II)', *SIAM J. Numer. Anal.* **29**, 845–866.

L. Alvarez and L. Mazorra (1992), 'Signal and image restoration by using shock filters and anisotropic diffusion', Preprint, Dep. de Inf. U.L.P.G.C. ref:0192, *SIAM J. Numer. Anal.*, to appear.

S. Angenent (1989), 'Parabolic equations for curves on surfaces I, II', University of Wisconsin-Madison Technical Summary Reports, 19, 24.

H. Asada and M. Brady (1986), 'The curvature primal sketch', *IEEE Trans. Patt. Anal. Machine Intell.* **8**(1).

C. Ballester and M. Gonzalez (1993), 'Affine invariant multiscale segmentation by variational method', *Proc. Eighth Workshop on Image and Multidimensional Signal Processing (8–10 September, Cannes)*, IEEE (New York), 220–221.

G. Barles (1985), 'Remarks on a flame propagation model', Technical Report No 464, INRIA Rapports de Recherche.

G. Barles and C. Georgelin (1992), 'A simple proof of convergence for an approximation scheme for computing motions by mean curvature', Preprint.

G. Barles and P.E. Souganidis (1993), 'Convergence of approximation schemes for fully nonlinear second order equation', *Asymp. Anal.*, to appear.

C. Brice and C. Fennema (1970), 'Scene analysis using regions', *Artificial Intelligence* **1**, 205–226.

W. Brockett and P. Maragos (1992), 'Evolution equations for continuous-scale morphology', *ICASSP, San Francisco* 23–26.

A.M. Bruckstein, G. Sapiro and D. Shaked (1992), 'Affine-invariant evolutions of planar polygons', Preprint.

V. Caselles, F. Catté, T. Coll and F. Dibos (1992), 'A geometric model for active contours in image processing', Report 9210, C.E.R.E.M.A.D.E., Université Paris Dauphine (Paris).

F. Catté, F. Dibos and G. Koepfler (1993), 'A morphological approach of mean curvature motion', Report 9310, C.E.R.E.M.A.D.E., Université Paris Dauphine (Paris).

Y-G. Chen, Y. Giga and S. Goto (1989), 'Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations', Preprint, Hokkaido University.

T. Cohignac, F. Eve, F. Guichard, C. Lopez and J. M. Morel (1993a), *Numerical Analysis of the Fundamental Equation of Image Processing*, to appear.

T. Cohignac, C. Lopez and J. M. Morel (1993b), 'Multiscale analysis of shapes, images and textures', *Proc. Eighth Workshop on Image and Multidimensional Signal Processing (8–10 September, Cannes)*, IEEE (New York) 142–143.

M. G. Crandall, H. Ishii and P.L. Lions (1991), 'User's guide to viscosity solution of second order partial differential equation', C.E.R.E.M.A.D.E., Preprint (Paris).

G. Dal Maso, J.M. Morel and S. Solimini (1989), 'Une approche variationnelle en traitement d'images: résultats d'existence et d'approximation', *C.R. Acad. Sci. Paris* **308**, 549–554.

I. Daubechies (1992), 'Ten lectures on wavelets', SIAM (Philadelphia).

F. Dibos and G. Koepfler (1991), 'Propriété de régularité des contours d'une image segmentée', *C.R. Acad. Sci. Paris* **313**, 573–578.

J. Enns (1986), 'Seeing textons in context', *Perception and Psychophysics* **39**(2), 143–147.

L.C. Evans and J. Spruck (1992), 'Motion of level sets by mean curvature I', Preprint.

O. Faugeras (1993), 'A few steps toward a projective scale space analysis', *C.R. Acad. Sci. Paris* to appear.

L. Florack, B. ter Haar Romeny, J.J. Koenderink and M. Viergever (1991), 'General intensity transformations and second order invariants', *Proc. 7th Scandinavian Conference on Image Analysis (Aalborg)* 13–16.

L. Florack, B. ter Haar Romeny, J.J. Koenderink and M. Viergever (1992), 'Scale and the differential structure of images', *Image Vision Computing* **10**.

D. Forsyth, J.L. Mundy and A. Zisserman (1991), 'Invariant descriptors for 3-D object recognition and Pose', *IEEE Trans. Patt. Anal. Machine Intell.* **13**, No.10.

M. Gage and R.S. Hamilton (1986), 'The heat equation shrinking convex plane curves', *J. Diff. Geom.* **23**, 69–96.

S. Geman and D. Geman (1984), 'Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images', *IEEE Patt. Anal. Machine Intell.* **6**.

F. Guichard (1993), 'Multiscale analysis of movies', *Proc. Eighth Workshop on Image and Multidimensional Signal Processing (8–10 September, Cannes)*, IEEE (New York), 236–237.

F. Guichard, J.M. Lasry and J.M. Morel (1993), 'A monotone consistent theoretical scheme for the fundamental equation of image processing', Preprint.

M. Grayson (1987), 'The heat equation shrinks embedded plane curves to round points', *J. Diff. Geom.* **26**, 285–314.

R.M. Haralick and L.G. Shapiro (1985), 'Image segmentation techniques', *Comput. Vision Graph. Image Process.* **29**, 100–132.

B. Horn (1986), *Robot Vision*, MIT (Cambridge, MA).

S. L. Horowitz and T. Pavlidis (1974), 'Picture segmentation by a directed split-and-merge procedure', *Proc. Second Int. Joint Conf. Pattern Recognition* 424-433.

R. Hummel (1986), 'Representations based on zero-crossing in scale-space', *Proc. IEEE Computer Vision and Pattern Recognition Conf.*, 204–209.

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

B. Julesz (1986), 'Texton gradients: the texton theory revisited', *Biol. Cybern.* **54**, 245–251.

B. Julesz and J.R. Bergen (1983), 'Textons, the fundamental elements in preattentive vision and perception of textures', *Bell System Tech. J.* **62**(6), 1619–1645.

B. Julesz and B. Kroese (1988), 'Features and spatial filters', *Nature* **333**, 302–303.

M. Kass, A. Witkin and D. Terzopoulos (1987), 'Snakes: active contour models', *1st Int. Comput. Vis. Conf. IEEE 777.*

B.B. Kimia (1990), 'Toward a computational theory of shape', PhD Dissertation Department of Electrical Engineering, McGill University, Montreal, Canada.

B.B. Kimia, A. Tannenbaum and S.W. Zucker (1992), 'On the evolution of curves via a function of curvature, 1: the classical case', *J. Math. Anal. Appl.* **163**(2).

G. Koepfler, J.M. Morel and S. Solimini (1991), 'Segmentation by minimizing a functional and the "merging" methods', *Proc. 'GRETSI Colloque' (Juan-les-Pins, France).*

G. Koepfler, C. Lopez and J.M. Morel (1994), 'A multiscale algorithm for image segmentation by variational method', *SIAM J. Numer. Anal.* **31**, to appear.

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

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

J.J. Koenderink (1990b), 'The brain, a geometry engine', *Psychol. Res.* **52**, 122–127.

J.J. Koenderink and A.J. van Doorn (1986), 'Dynamic shape', *Biol. Cybern.* **53**, 383–396.

J.J. Koenderink and A.J. van Doorn (1987), 'Representation of local geometry in the visual system', *Biol. Cybern.* **55**, 367–375.

Y. Lamdan, J.T. Schwartz and H.J. Wolfson (1988), 'Object recognition by affine invariant matching', in *Proc. CVPR 88.*

T. Lindeberg (1990), 'Scale-space for discrete signal', *IEEE Trans. Patt. Anal. Machine Intell.* **12**, 234–254.

C. Lopez and J.M. Morel (1992), 'Axiomatisation of shape analysis and application to texture hyperdiscrimination', *Proc. Trento Conf. on Surface Tension and Movement by Mean Curvature*, De Gruyter (Berlin).

A. Mackworth and F. Mokhtarian (1986), 'Scale-based description and recognition of planar curves and two-dimensional shapes', *IEEE Trans. Patt. Anal. Machine Intell.* **8**(1).

A. Mackworth and F. Mokhtarian (1992), 'A theory of multiscale, curvature-based shape representation for planar curves', *IEEE Trans. Patt. Anal. Machine Intell.* **14**, 789–805.

J. Malik and P. Perona (1991), 'Preattentive texture discrimination with early vision mechanisms', *J. Opt. Soc. Am.* A **7**(5), 923–932.

P. Maragos (1987), 'Tutorial on advances in morphological image processing and analysis', *Opt. Engrg* **26**(7).

D. Marr (1976), 'Analyzing natural images: a computational theory of texture vision', *Cold Spring Harbor Symp. on Quantitative Biology, XL* 647–662.

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

P. Mascarenhas (1992), 'Diffusion generated motion by mean curvature', Preprint.

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

B. Merriman, J. Bence and S. Osher (1992), 'Diffusion generated motion by mean curvature', CAM Report 92-18, Department of Mathematics, University of California (Los Angeles CA 90024.1555, USA).

Y. Meyer (1992), *Ondelettes et Algorithmes Concurrents*, Hermann (Paris).

J.M. Morel and S. Solimini (1988a), 'Segmentation of images by variational methods: a constructive approach', *Rev. Matematica de la Universidad Complutense de Madrid* Vol. 1 **1,2,3**, 169–182.

J.M. Morel and S. Solimini (1988b), 'Segmentation d'images par méthode variationnelle: une preuve constructive d'existence', *C. R. Acad. Sci. Paris.*

J.M. Morel and S. Solimini (1993), *Variational Methods in Image Segmentation*, Birkhauser (Boston) to appear.

J.L. Muerle and D. C. Allen (1968), 'Experimental evaluation of techniques for automatic segmentation of objects in a complex scene', in *Pictorial Pattern Recognition* (G. C. Cheng *et al.*, eds), Thompson (Washington), 3–13.

D. Mumford and J. Shah (1988), 'Boundary detection by minimizing functionals', *Image Understanding* (S. Ullman and W. Richards, eds).

D. Mumford and J. Shah (1989), 'Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems', *Commun. Pure Appl. Math.* XLII **4**.

S. Osher and J. Sethian (1988), 'Fronts propagating with curvature dependent speed: algorithms based on the Hamilton–Jacobi formulation', *J. Comput. Phys.* **79**, 12–49.

T. Pavlidis (1972), 'Segmentation of pictures and maps through functional approximation', *Comput. Graph. Image Process.* **1**, 360–372.

T. Pavlidis and Y.T. Liow (1988), 'Integrating region growing and edge detection', *Proc. IEEE Conf. on Comput. Vision Patt. Recognition.*

P. Perona and J. Malik (1987), 'A scale space and edge detection using anisotropic diffusion', *Proc. IEEE Computer Soc. Workshop on Computer Vision.*

L. Rudin, S. Osher and E. Fatemi (1992), 'Nonlinear total variation based noise removal algorithms', *Proc. Modélisations Matématiques pour le Traitement d'Images*, INRIA 149–179.

A. Rosenfeld and A. Kak (1982), *Digital Picture Processing* Vol. 1, Academic (New York).

G. Sapiro and A. Tannenbaum (1992a), 'On affine plane curve evolution', EE Pub 821, Department of Electrical Engineering, Technion Israel Institute of Technology, Haifa, Israel.

G. Sapiro and A. Tannenbaum (1992b), 'Affine shortening of non-convex plane curves', EE Pub 845, Department of Electrical Engineering, Technion Israel Institute of Technology, Haifa, Israel.

J. Serra (1982), *Image Analysis and Mathematical Morphology* Vol. 1, Academic (New York).

A. Treisman (1985), 'Preattentive processing in vision', *Comput. Vision, Graph. Image Process.* **31**, 156–177.

H. Voorhees and T. Poggio (1987), 'Detecting textons and texture boundaries in natural images', *Proc. Int. Conf. Computer Vision*, IEEE (New York), 250–258.

A. P. Witkin (1983), 'Scale-space filtering', *Proc. IJCAI (Karlsruhe)* 1019–1021.

A. Yuille (1988), 'The creation of structure in dynamic shape', *Proc. Second International Conference on Computer Vision (Tampa)* 685–689.

A. Yuille and T. Poggio (1986), 'Scaling theorems for zero crossings', *IEEE Trans. Patt. Anal. Machine Intell.* **8**.

S. W. Zucker (1976), 'Region growing: childhood and adolescence (survey)', *Comput. Graph. Image Process.* **5**, 382–399.