Chapter 3
Robust Shape Directions

Abstract This chapter deals with shape affine normalization. This method associates with all shapes deduced from each other by an affine distortion a single normalized shape. A crucial ingredient for normalization is the computation of a small affine covariant set of robust straight lines associated with a shape. The set of all tangent lines to a shape has this covariance property, but it is too large. A very successful idea is to use bitangent lines, that is, lines tangent to a shape at two different points. If the shape has a finite number of inflexion points it also has a finite number of bitangent lines. In Sect. 3.3 a well-established curve affine invariant smoothing algorithm will be briefly described. This smoothing permits a drastic reduction of the number of bitangent lines. Yet, not all shapes can be encoded by using bitangents. Convex shapes have no bitangents and simple shapes have only a few. This explains why shape recognition algorithms compute other robust straight lines associated with the shape. Flat parts of curves are informally defined as intervals of the curve along which the direction of the tangent line does not vary too much. For instance, large enough polygons show as many reliable flat parts as sides. This chapter will present a simple parameterless definition of flat parts, based again on the Helmholtz principle.

3.1 Flat Parts of Level Lines

Flatness of a part of curve will be measured by comparing its direction at each point with the direction of the underlying chord (see Fig. 3.1).

Although flatness may look like a rather intuitive geometric concept, it is in fact quite complex. Our aim is to define a unique measurement, the flatness for very diverse phenomena: A long very oscillating curve may look flat seen at a distance. In another way, a short and very smooth curve can look locally very flat. One can therefore figure out that at least two parameters are involved in a flatness measurement. One measures the length of the flat part and the other gives the amplitude of the oscillations. Thus, the flatness definition problem can be viewed as the question

F. Cao et al., *A Theory of Shape Identification*. Lecture Notes in Mathematics 1948.

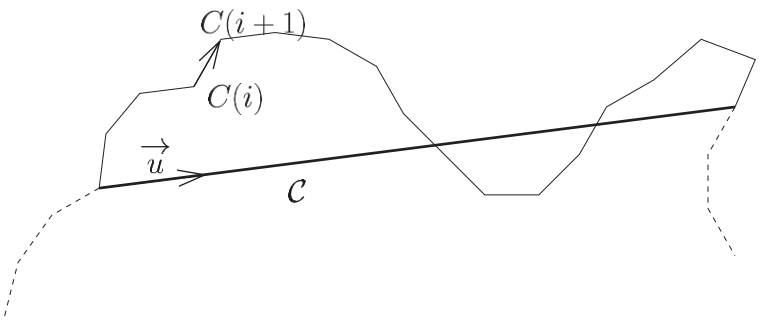


Fig. 3.1 A piece of discrete curve with the underlying chord $\mathcal{C}$ (thick segment line)

of reducing two parameters to a more abstract one, the flatness. The detection of flat parts of a curve should meet the following requirements:

- It should not detect just points around which the curve is flat, but the precise straight intervals on the curve.
- Long flat parts should be allowed to move farther from their underlying chord than short ones.
- The detection should be intrinsic to the curve, and not depend on other curves in the image.
- Detected flat parts should not overlap.
- Since detecting flat parts is generally the first step of a recognition algorithm, it deals with a huge amount of information. Therefore, computational complexity should be low.

3.1.1 Flat Parts Detection Algorithm

Consider a chord from a given curve C: its endpoints delimitate a piece of curve of length l (measured in pixels). Since one would like to measure how much the piece turns with respect to the direction $\overrightarrow{u}$ given by the chord, let us define

$$\alpha = \max_{i \in \{0 \ldots n-1\}} \left\{ \left| \text{angle}(\overrightarrow{C(s_i)C(s_{i+1})}, \overrightarrow{u}) \right| \right\},$$

where the discrete piece of curve is made of the n consecutive points $C(s_i)$.

Suppose that α is below some fixed threshold α^*. Following the discussion on independence in Sect. 2.5, consider that points at a geodesic distance (along the curve) larger than 2 are statistically independent. Thus, there are $l/2$ statistically independent segments of the type $(C(s_i), C(s_{i+1}))$ along a curve with length l. The probability of the event that $l/2$ statistically independent points on a piece of curve show a tangent line which makes an angle lower than α among all the pieces of

curve for which $\alpha < \alpha^*$ is given by:

$$p(\alpha, l) = \left(\frac{\alpha}{\alpha^*}\right)^{l/2}.$$

Of course, the lower $p(\alpha, l)$ the flatter the piece of curve.

This straightforward computation is valid under the assumption that among all the pieces of curves such that $\alpha < \alpha^*$, α is uniformly distributed over $[0, \alpha^*]$, and that the tangents are independent at Nyquist distance 2. Flat parts are now defined as rare events with regard to this *a contrario* model.

For each piece of the curve for which $\alpha < \alpha^*$, the probability $p(\alpha, l)$ is computed. Only pieces such that $p(\alpha, l)$ is under a predetermined threshold p^* are kept (these parts are called candidates). Such pieces can of course overlap. So some of them must be selected to be the flat parts of the curves. A greedy algorithm will be used: the piece of curve with the lowest p is marked as a flat part, then all candidates that share a common part with this best flat part are eliminated. The process is iterated with the remaining candidates.

3.1.2 Reduction to a Parameterless Method

The computation of α clearly depends on the discretization. The curves which the proposed algorithm deals with are level lines of images. Their natural discretization is the pixel.

The whole algorithm involves two thresholds. The first one, α^*, is not critical. Indeed, since one is interested in detecting flat parts, it is natural to *a priori* reject all pieces of curve where α is above a large threshold. We set $\alpha^* = 1$ radian once for all, which is not a strong constraint. More specifically, a change of α^* multiplies all probabilities $p(\alpha, l)$ by a constant factor. Thus, the flatness measurement is just scaled and the ordering maintained. Moreover, changing α^* also multiplies the threshold p^* by the same constant. Thus, there are not two parameters here, but just one, namely p^*. This last parameter will be eliminated by Helmholtz principle. It can be fixed in such a way that almost no flat part occurs in the level curves of a white noise.

Experimental evidence shows that $p^* = 10^{-3}$ is the maximum value for which only a few detections (on average one) occur on level lines extracted from a white noise image containing the same amount of level lines as a standard natural image. So with this value for p^* the proposed algorithm satisfies the Helmholtz principle in that there is almost no detection of flat parts in a white noise image.

3.1.3 The Algorithm

Consider a Jordan curve on which flat parts are searched for.

Part I: Candidate identification.

For each chord of the curve with length 10, 20, 30, ..., 180, 200, and then an exponential progression[1]:

1. Compute the maximum angle α between the chord and the piece of curve delimited by both ends of the chord. If n denotes the number of independent points $C(s_i)$ on this piece of discrete curve:
$$\alpha = \max_{i \in \{1 \ldots n-1\}} \left\{ \left| \text{angle}(\overrightarrow{C(s_i)C(s_{i+1})}, \overrightarrow{u}) \right| \right\}.$$
2. If $\alpha > 1$ rad, *a priori* reject the piece; else compute $p(\alpha, l) = \left(\frac{\alpha}{\alpha^*}\right)^{l/2} = \alpha^{l/2}$, where l is the length of the considered piece of curve.
3. If $p(\alpha, l) > p^* = 10^{-3}$, reject the piece.

Part II: Greedy algorithm

1. Keep the candidate for which $\alpha^{l/2}$ is minimal, mark it as *flat part*, and discard it from the list of candidates.
2. Reject all candidates that meet this best candidate.
3. Iterate until no candidate is available anymore.

3.1.4 Some Properties of the Detected Flat Parts

The condition defining the candidates ($\alpha^{l/2} < p^*$) is not a real constraint for long curves. For example, if $p^* = 10^{-3}$ and $l = 200$, all curve parts such that $\alpha < 0.97$ are accepted as candidates. Nevertheless, long pieces of curves often show subparts with a lower probability and a greedy algorithm will therefore prefer them. In the case of circles, however, this does not occur. Let us compute the arcs of circle which will be marked as flat parts. Figure 3.2 illustrates the following computations.

Proposition 3. *A circle of radius R has flat parts if and only if $R \geqslant -e \log(p^*)$. In such a case, the length of the detected flat parts is $L = 2R \sin(1/e)$.*

[1] There is a complexity issue here. All chords are not tested, but only a subsample of them so that the algorithm does not waste too much time for long curves. The only consequence of this discretization procedure is that long straight lines (in practice, lines whose length is larger than 100 pixels) can be split into two pieces (see Fig. 3.14 for an example). This is not an important drawback since the goal is to use flat parts as robust directions.

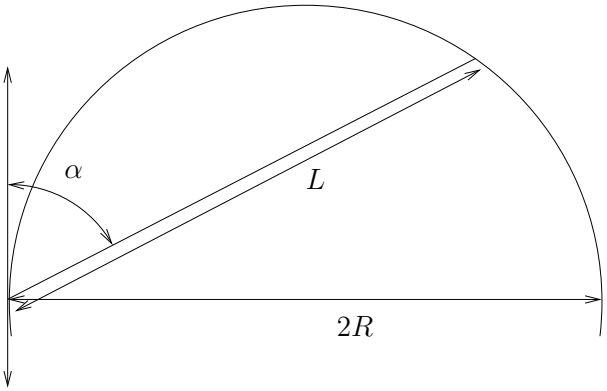


Fig. 3.2 Illustration of the flat parts computation on a circle

Proof. A circle of radius R being given, let us consider a chord of length L defining a maximum angle α with the corresponding piece of curve ($0 \leqslant \alpha \leqslant \pi/2$). The values of α and L are related by $L = 2R\sin(\alpha)$. The probability defined earlier is $p(\alpha, L) = \alpha^{R\alpha}$ (expressed as a function of L, it writes down $p(\alpha, L) = \arcsin(L/2R)^{R\arcsin(L/2R)}$). The function $\alpha \mapsto \alpha^{R\alpha}$ attains a minimum for the value $\alpha = 1/e$. Consequently, $\forall\, \alpha,\ \alpha^{R\alpha} \geqslant e^{-R/e}$.

Thus, if the probability threshold is set to p^*, and if $R < -e\log(p^*)$, then the circles of length R will show no flat part. On the contrary, if $R \geqslant -e\log(p^*)$, the detected flat parts (after the greedy step) in circles of radius R will always show a maximum angle $\alpha = 1/e$ (that is to say 21 degrees, corresponding to an arc of $1/9$ of the total circle), and their length will be $L = 2R\sin(1/e)$. □

Notice that p^* only controls the minimum radius under which no flat part will be detected: $-e\log(p^*)$. It appears only through its logarithm and small variations of it will not influence the final result. Although for symmetry reasons no piece of circle should be favored by the algorithm, the position of the detected flat parts over a circle strongly depends on the starting point of the discrete curve describing this circle. This makes flat parts of circular curves unreliable in position. However, this will not hinder the recognition of circles, as a a circle matches well with itself, up to any rotation.

3.2 Experiments

3.2.1 Experimental Validation of the Flat Part Algorithm

Experimental results are shown in Figs. 3.4 to 3.9 (original images can be seen on Fig. 3.3). For each image, the computation time is less than 10 seconds, for a 2GHz

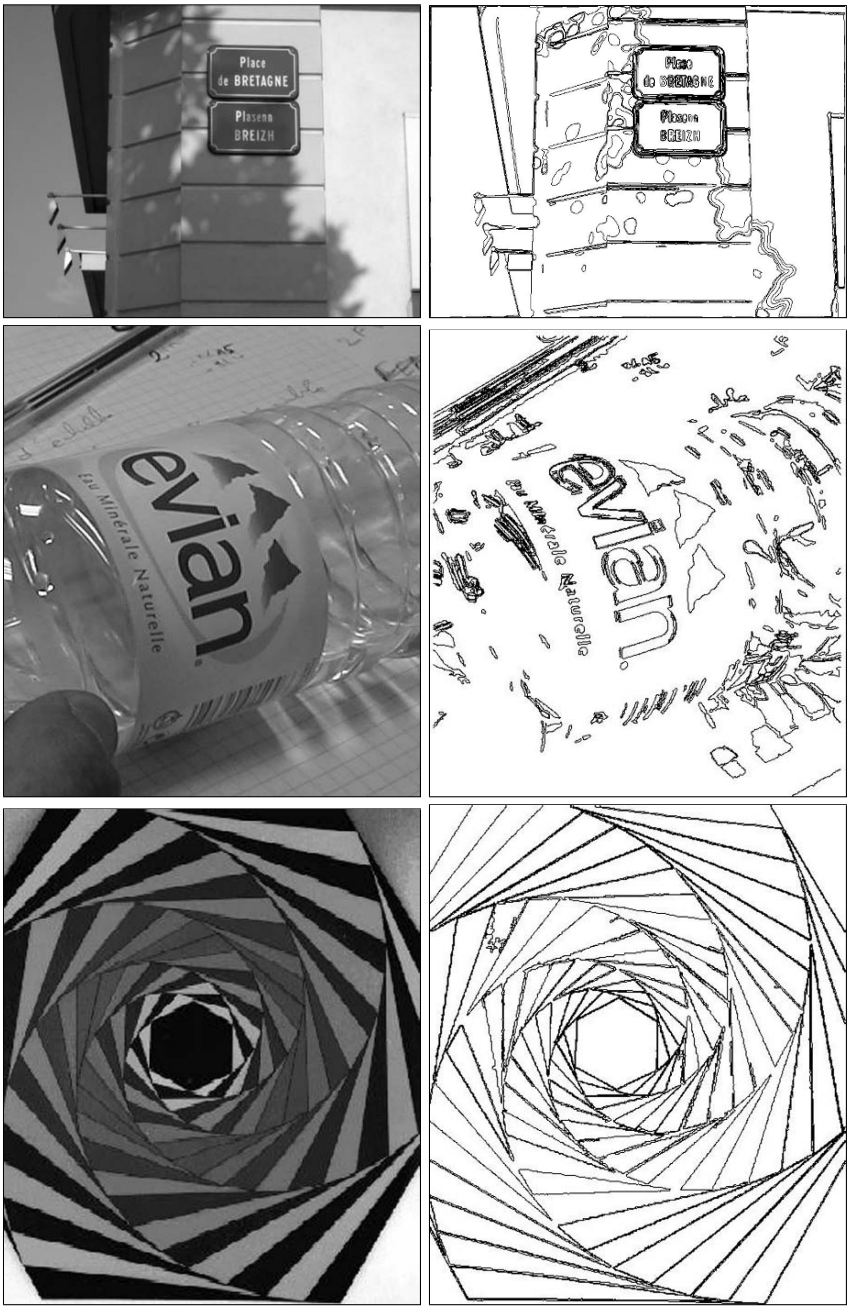


Fig. 3.3 Left column: original images. Right column: meaningful level lines detected with the method described in Chap. 2 (right). Top: *Bretagne*, 413 level lines. Middle: *Evian*, 481 level lines. Bottom: *Vasarely*, 172 level lines

standard PC. When images do not show long level lines, the computation time is less than a second.

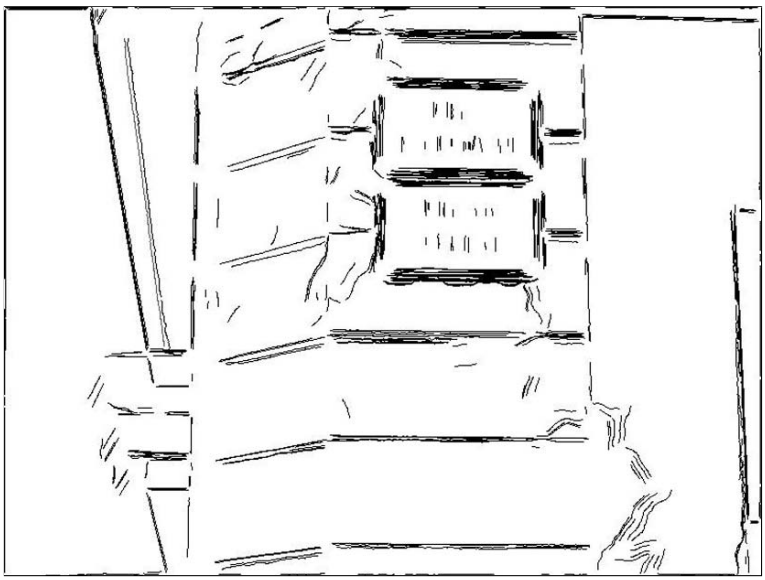


Fig. 3.4 Flat parts detection: *Bretagne*. 1004 detections. Flat parts as small as the ones in the letters of the name of the street are detected (about 10 pixels high). Flat parts in the boundaries of the shadows can be eliminated by dropping the probability threshold, as can be seen on Fig. 3.5. Nevertheless these detections actually correspond to small flat parts

3.2.2 Flat Parts Correspond to Salient Features

Figures 3.10 and 3.11 show the result of the proposed flat parts detector over all level lines in an image. By all, we mean that all level lines at all levels with quantization step equal to 1 have been extracted. This allows for an exact reconstruction of the original image from the level lines and their corresponding gray levels [135]. Some segments are detected over level lines corresponding to quantization noise (*i.e.* not contrasted level lines over perceptually uniform areas), but these segments actually correspond to small pieces of straight lines. They are no longer detected when the probability threshold p^* is set to 10^{-10} instead of the standard value (10^{-3}). Flat parts are concentrated along edges. This experiment confirms that segment lines are actually salient image features.

Comparing Fig. 3.3 to Figs. 3.4 to 3.7 shows that almost all detected flat parts belong to maximal meaningful boundaries.

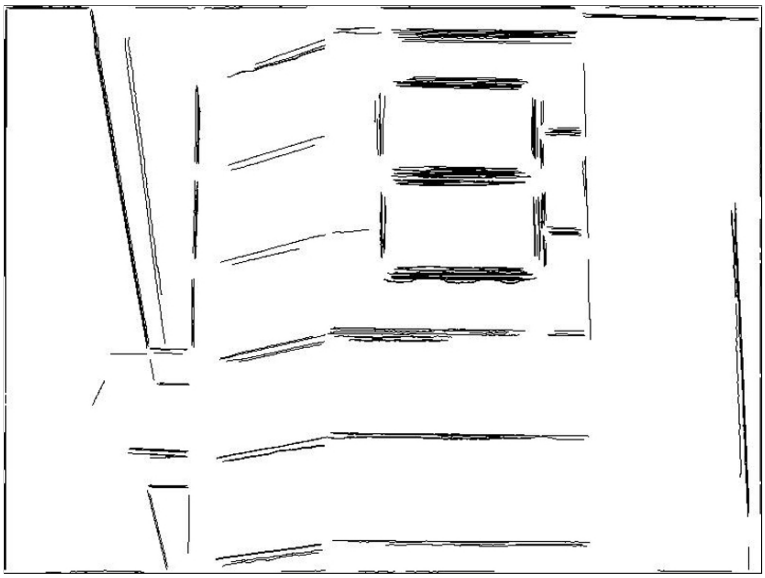


Fig. 3.5 Flat parts detection: *Bretagne*, with $p^* = 10^{-10}$, 417 detections. Letters are too small to be detected but the remaining flat parts are very accurate

Fig. 3.6 Flat parts detection: *Evian*. 448 detections

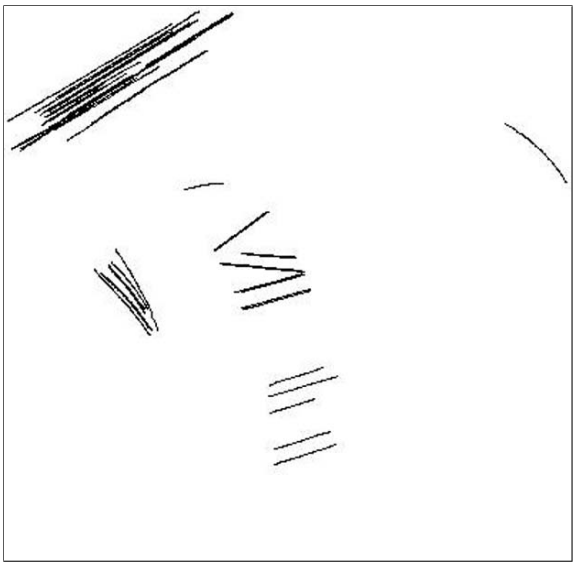


Fig. 3.7 Flat parts detection: *Evian*, with $p^* = 10^{-10}$, 64 detections

In his PhD thesis, Lisani [108] used a flat points detector to build robust semi-local normalization. Figures 3.12 to 3.15 show a comparison between the flat parts proposed in this chapter and flat points in the sense of Lisani. See captions for details.

3.3 Curve Smoothing and the Reduction of the Number of Bitangent Lines

Level lines may be subject to noise, and can have details that are too fine in relation to the essential shape information. Hence, a good shape representation requires a previous smoothing. Is this smoothing necessary? Quite, from the technological viewpoint, as otherwise there would be too many bitangent lines to level lines and therefore too many geometric codes to a level line. The general framework by which an image or a shape is smoothed at several scales in order to eliminate spurious or textural details and extract its main features is called Scale Space. The main developments of Scale Space theory in the past ten years involve invariance arguments. Indeed, a scale space will be useful for shape recognition only if it is invariant. Let us summarize a series of arguments given in [5]. A scale space computing contrast invariant information must in fact deal directly with the image level lines; in order to be local (not dependent upon occlusions), it must be in fact a partial differential equation (PDE). In order to be a smoothing, this PDE must be parabolic. The affine

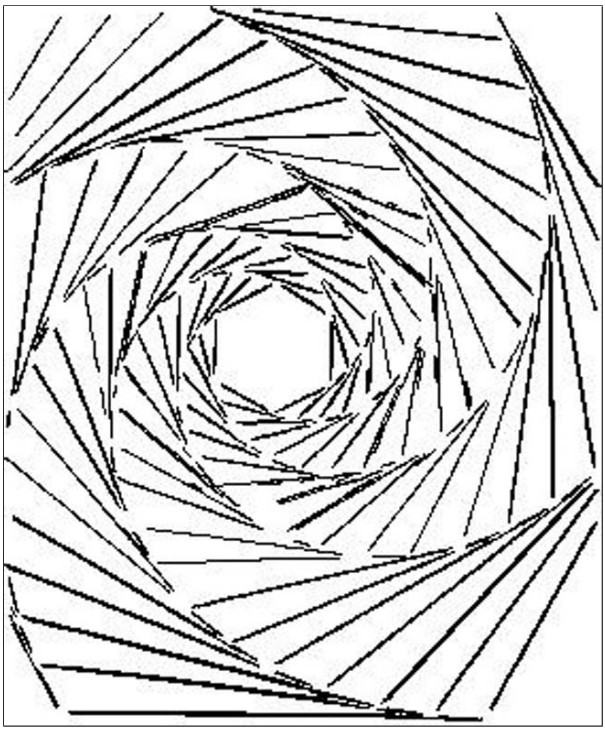


Fig. 3.8 Flat parts detection: *Vasarely*, 774 detections. Each triangle side is correctly detected as a single flat part

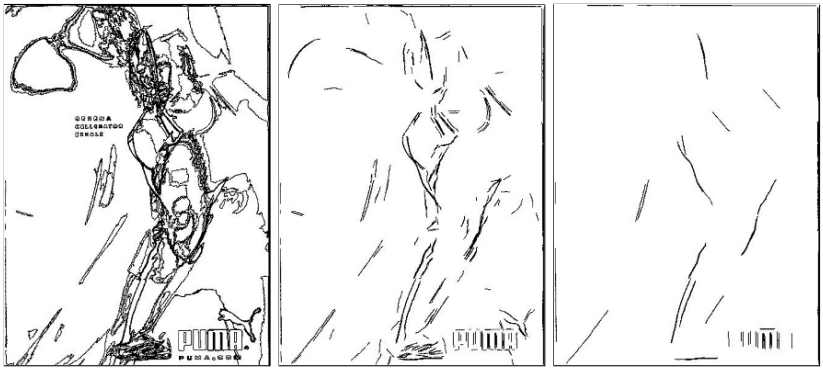


Fig. 3.9 Flat parts detection: *Serena Williams & Puma* (original image shown in Fig. 2.6). Left: Original level lines (425 lines). Middle: $p^* = 10^{-3}$ (675 detections). Right: $p^* = 10^{-10}$ (156 detections). Flat parts on letters are correctly extracted

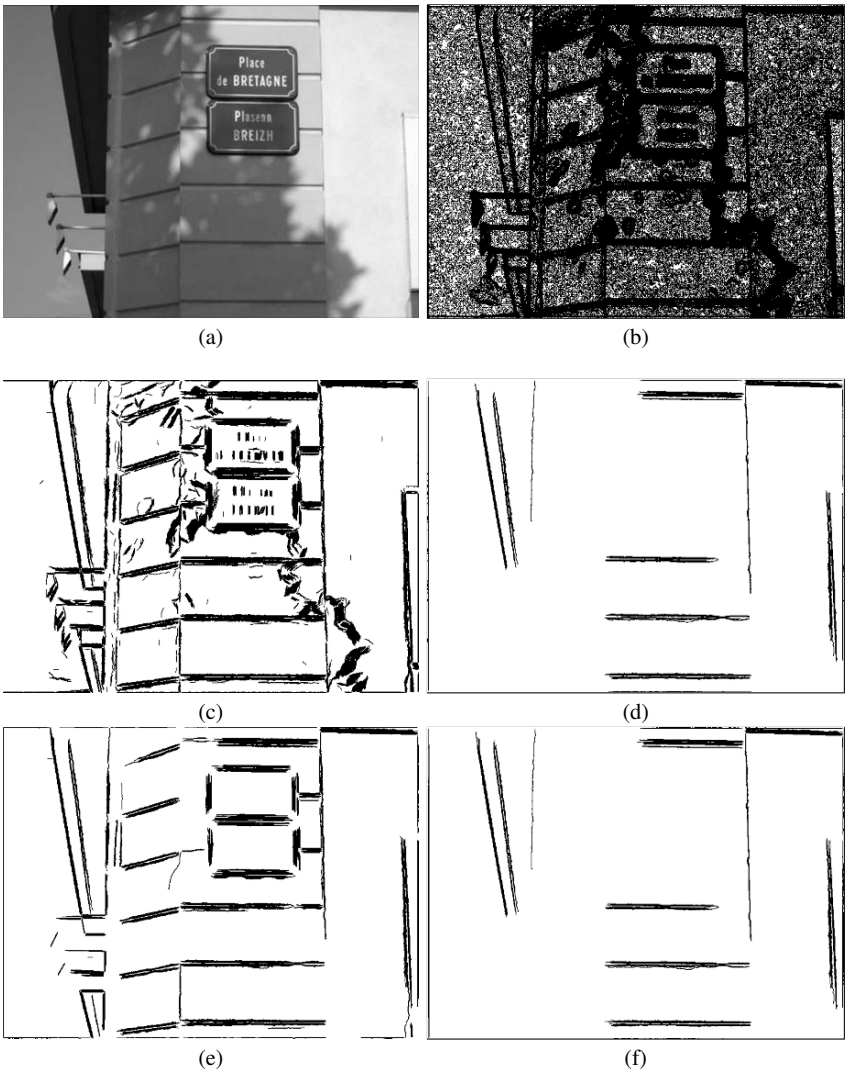


Fig. 3.10 Flat parts detection. (a) original image (size: 512×384); (b) 25,755 level lines (quantization step: 1 gray level). They cover the whole image. (c) 20,065 flat parts detected over these level lines (probability threshold p^* has here its standard value: 10^{-3}); (d) flat parts of length larger than 100 pixels among the previous ones; (e) 6,233 flat parts detected over these level lines, when the probability threshold p^* is set to 10^{-10}; (f) flat parts of length larger than 100 pixels among the previous ones. Flat parts appear to be concentrated along edges in thick bundles

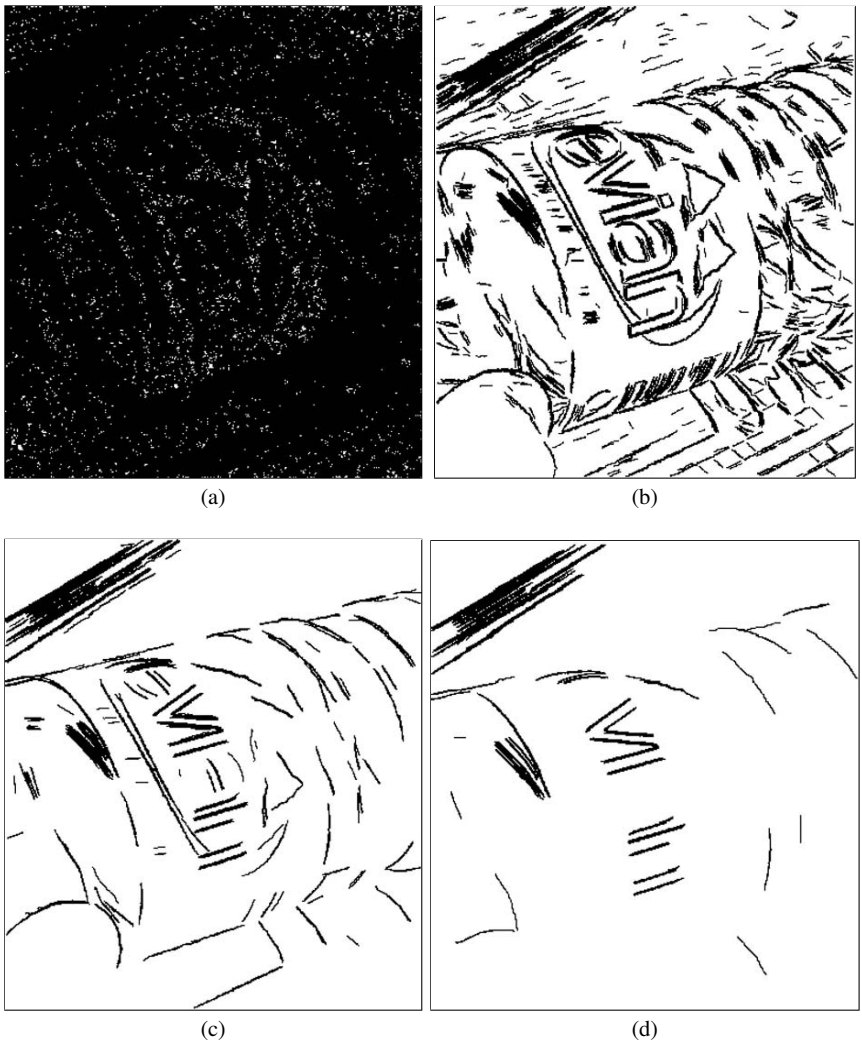


Fig. 3.11 Flat parts detection. (a) 90,078 level lines from Evian image (quantization step: 1 gray level); (b) flat parts detections over these level lines (16,533 detections); (c) flat parts detection with $p^* = 10^{-6}$ (4,659 detections); and (d) flat parts detection with $p^* = 10^{-10}$ (2,041 detections). Flat parts are concentrated along edges

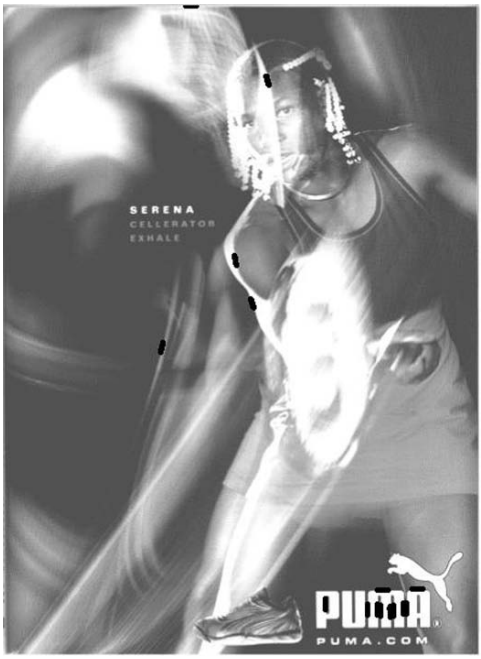


Fig. 3.12 Lisani's flat points: *Serena Williams & Puma*. Only 15 flat points (in black) are detected. To be compared to the results in Fig. 3.9

invariance requirement and the invariance with respect to reverse contrast lead to a single PDE [5]. This PDE, characterizing the unique contrast, contrast reversal and special affine invariant scale space is

$$\begin{cases} \frac{\partial u}{\partial t} = |Du|(\operatorname{curv} u)^{1/3}, \\ u(x,t) = u_0(x). \end{cases} \tag{3.1}$$

It is called *Affine Morphological Scale-Space* (AMSS). Here $u(t,0) = u_0$ is the initial image, $u(t,x)$ is the image smoothed at scale t and $\operatorname{curv}(u)(x) = \operatorname{div}(\frac{Du}{|Du|})$ denotes the signed curvature of the level line passing by x. This equation is equivalent to the affine curve shortening [155] of all of the level lines of the image, given by the equation

$$\frac{\partial x}{\partial t} = |\operatorname{Curv}(x)|^{\frac{1}{3}}\mathbf{n}, \tag{3.2}$$

where x denotes a point of a level line, $\operatorname{Curv}(x)$ its curvature and $\mathbf{n}$ the signed normal to the curve, always pointing towards the concavity.
Moisan [127] found a fast algorithm for this curvature motion. For more details on this scheme, refer to [127, 99] and to the book [29]. The invariants mentioned mean that the evolution of a shape does not depend upon any affine distortion of the plane. This corresponds to an invariance to all orthographic projections of a planar shape.

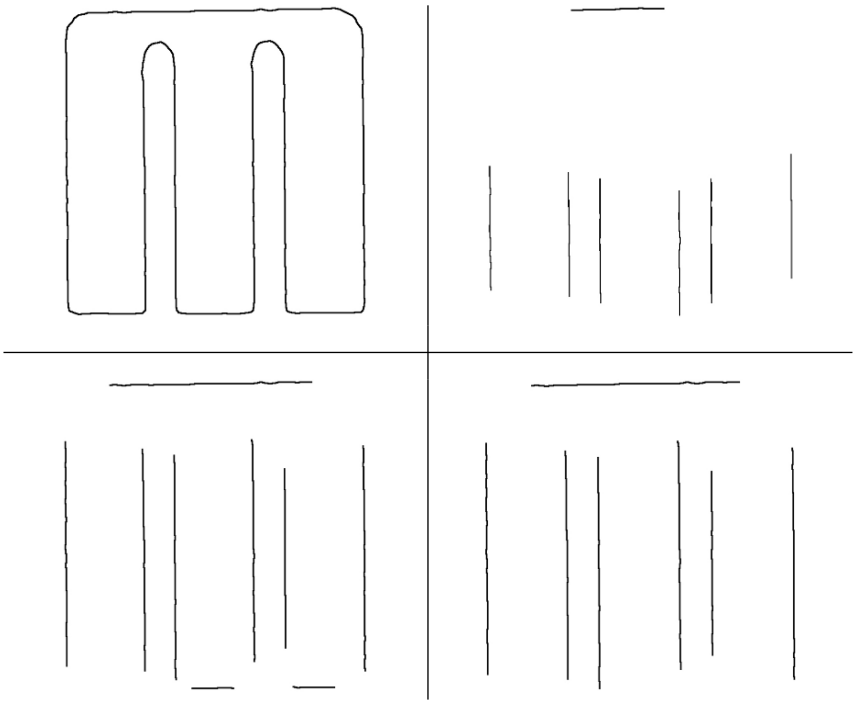


Fig. 3.13 Flat points *vs* flat parts: *Serena Williams & Puma*. From left to right and from top to bottom: considered level line, flat points (7 detections), flat parts with $p^* = 10^{-3}$ (9 detections), flat parts with $p^* = 10^{-10}$ (7 detections). One of the flat parts in the legs of the character M is not detected since these curve pieces are too small and pose a sampling problem. Since not *all* chords are tested but a subset of them, endpoints may sometimes be not conveniently distributed

Figure 3.16 shows that a slight smoothing by the affine scale space eliminates the sampling effects of a digital image and reduces drastically the number of inflexion points of a shape without altering its overall aspect. Numerically, the smoothing is slight and stops at the scale $t = 0.5$ at which a circle with radius 0.5 collapses. So the smoothing roughly eliminates details of 1 pixel size.

3.4 Bibliographic Notes

3.4.1 Detecting Flat Parts in Curves

In their seminal paper [65], Fischler and Bowles argue that any curve partitioning technique must satisfy two general principles: stability of the description, and a complete and concise explanation. Smooth sections of curves play a major role

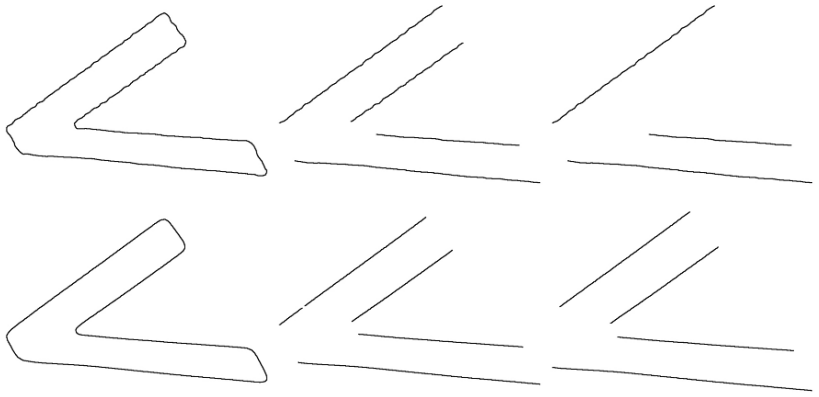


Fig. 3.14 Flat points *vs* flat parts: character V in *Evian*. Top: no smoothing. From left to right: original level line, flat parts with $p^* = 10^{-3}$ (4 detections) and with $p^* = 10^{-10}$ (3 detections). The flat points algorithm does not provide any detection. Bottom: after smoothing. From left to right: original level line, flat parts with $p^* = 10^{-3}$ (5 detections) and flat parts with $p^* = 10^{-10}$ (4 detections). With $p^* = 10^{-3}$, one of the segments is split because of the discretization procedure in the multi-scale test of chords. Again here the Lisani flat points algorithm misses the segments

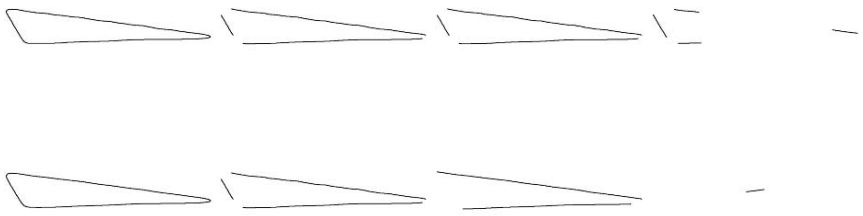


Fig. 3.15 Flat points *vs* flat parts: a triangle in *Vasarely*. Top: no smoothing. From left to right: original level line, flat parts with $p^* = 10^{-3}$ (3 detections) and flat parts with $p^* = 10^{-10}$ (3 detections), and flat points (4 detections). Bottom: after smoothing (see Sect. 3.3). From left to right: original level line, flat parts with $p^* = 10^{-3}$ (5 detections) and flat parts with $p^* = 10^{-10}$ (2 detections), and flat points (1 detection)

because they fit both principles. For instance, Guy and Medioni [78] consider segment lines as *salient* features in images. Flat part multiscale detection has been used for the more general problem of polygonal approximation of digitized curves (see [164]).

Segment or straight line detection is one of the cornerstones of computer vision. Indeed, it is often a preprocessing step of shape recognition, shape tracking [48], vanishing point detection [2], convex shape detection [92], *etc*. Most of the time,

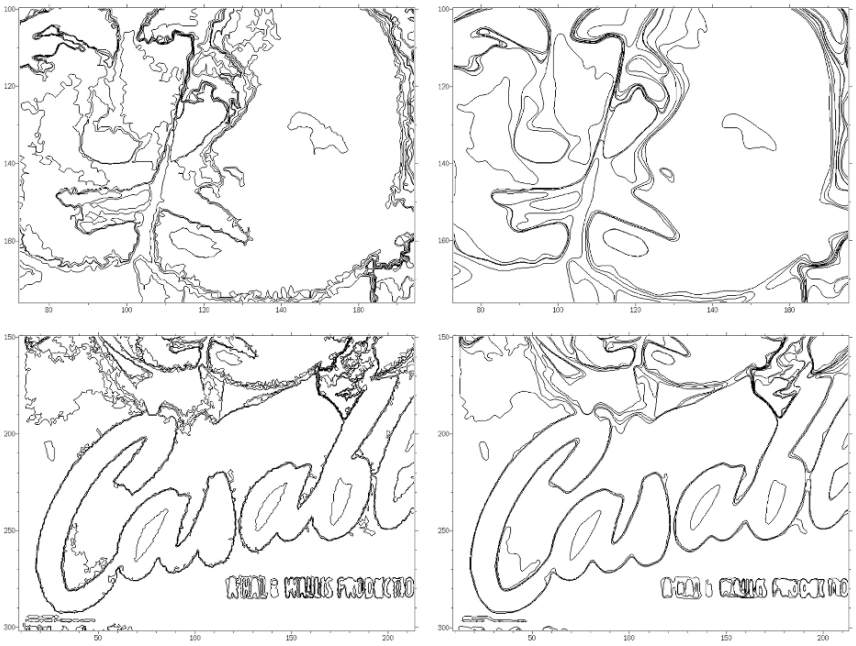


Fig. 3.16 Some level lines of a gray level image. Quantization effects and noise are seen. After a slight smoothing these effects disappear (right)

straight lines in images are conceived as contiguous edges. Many line detection algorithms therefore require a previous local edge extraction step, such as a Canny's filtering [28]. Hough Transform [85] and algorithms derived from it [91] have been widely studied for this purpose. The goal of these methods is to identify clusters in a particular space (the parameter space of a straight line, either (ρ, θ) with ρ the distance of the line to the origin, and θ the angle between a vector normal to the line and a fixed direction, or (a, b) where a is the slope and b the ordinate of the intersection between the straight line and the ordinate axis). The Hough transform is a voting procedure: every pixel votes for the parameters of the straight line going through it. Another method consists in first chaining the local edges by taking into account connectivity (see for an example [70]), and then in identifying segments among the discrete curves [107]. The main drawbacks of these methods are the number of thresholds (edge detection needs at least a gradient threshold, and the Hough Transform needs a quantization step for the parameter space discretization and a threshold for the voting procedure) and their computational burden and instability (due to local edges chaining). A fuzzy segment concept was proposed in [45]. In this method the primary detection is still based on a set of points derived from a local edge detector.

The method presented in this chapter can be viewed as an adaptation to the level lines of Desolneux et al. [50], who proposed an *a contrario* method detecting meaningful alignments in images. A meaningful alignment is a segment where a large enough proportion of points have their gradient orthogonal to the segment. More precisely a length l segment is ε-meaningful in a $N \times N$ image if it contains at least $k(l)$ points having their direction aligned with the one of the segment, where:

- $k(l)$ is given by: $k(l) = \min\{k \in \mathbb{N}, \Pr(S_l \geqslant k) \leqslant \varepsilon/N^4\}$, and
- $\Pr(S(l) \geqslant k)$ is the probability that, in at least k points in a straight segment of length l, the gradient of the image is orthogonal to the segment, up to a predetermined precision.

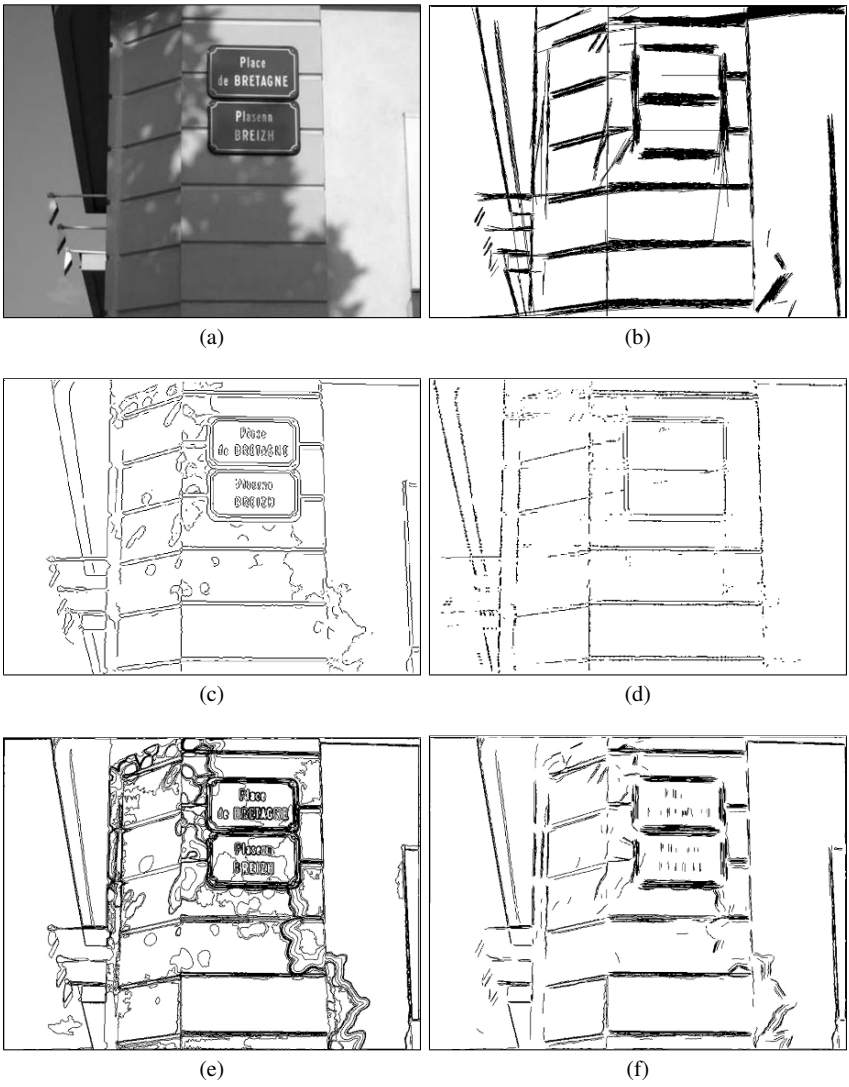
Estimating the probability that k points among l have a tangent with the same direction as the chord is not relevant to detect flat parts. In such a model, consecutive alignments are indeed not favored. They are instead crucial for shape normalization.

In his PhD thesis [108], Lisani defined flat points on curves by using two arbitrary parameters. A flat point is the center of a curve segment for which the sum of the angle variations of tangents is small enough (less than 0.2 radian) over a large enough piece of curve (larger than 15 pixels). This algorithm misses many flat points, and does not really detect segments, as several experiments have shown clearly.

Figure 3.17 shows the results for some of the algorithms which were just discussed. As far as flat parts detection is concerned, Desolneux's alignments are suitable neither for detecting accurate segment directions nor for detecting segment lengths. The naive segment detector based on Hough transform which illustrates the discussion is certainly not the best that can be done using Hough techniques. Nevertheless even a more clever algorithm would face the same problem as this one. It involves numerous critical parameters (different parameters would drastically change the results). Some isolated points are detected as segments because they fall by chance on the same straight line as another more distant segment and therefore collect its votes. Both algorithms (alignments and the Hough transform-based algorithm) are not local enough: that is why segments over the characters in the test image are not detected. Canny's edge detector is well known to suffer from lack of accuracy at edge junctions (where the gradient is badly estimated). Here, this would not be a real issue, since segment lines are searched for between junctions, where edges are more accurately detected. Nevertheless those edge detectors need several critical thresholds.

3.4.2 Scale-Space and Curve Smoothing

Since the seminal work of Lamdan *et al.* [101], bitangent lines are well-known to be of high interest to build up semi-local invariant curve descriptions. The reduction of the number of bitangent lines is linked to curve smoothing, or curve scale space. The modern concept of scale space comes from Witkin [181] and is mainly related to the



(a) (b)

(c) (d)

(e) (f)

Fig. 3.17 Segment detection. (a) original image; (b) maximal meaningful alignments [50]; (c) Canny's edge detector; (d) Points that correspond to an edge and that lie at the same time on a direction detected by voting in the Hough space; (e) local maximal meaningful level lines; (f) result of the proposed algorithm. See text for discussion.

Gaussian scale space, given by the heat equation [98]. An interesting shape recognition method using the mean curvature motion was discovered by Mocktharian and Mackworth [132]. The use of curvature-based smoothing for shape analysis is by now well established. The seminal papers are [10], [132] and [62]. These authors define a multi-scale curvature which is similarity invariant, but not affine invariant. Abbasi *et al.* [1] used the mean curvature motion and an affine length parameterization of the boundary of the solid shapes in order to get an approximately affine shape encoding. Sapiro and Tannenbaum [155] and Alvarez, Guichard, Lions and Morel [5] independently discovered the affine scale space with different approaches. Alvarez *et al.* proved existence of viscosity solutions to the affine scale space. An existence and regularity theorem was later proved by Angenent, Sapiro and Tannenbaum [7] from which it can be derived that the number of inflexion points decreases under the affine scale space. This result is crucial for shape encoding. Moisan [127] found a fast and fully affine invariant scheme implementing the affine scale space. He also proved the uniform consistency, which by a Barles and Souganidis [16] result is sufficient for convergence. The numerical scheme of Moisan was later extended by Cao and Moisan [34] to more general motions by curvature. Very recently the affine erosion scheme was used by Niethammer *et al.* [142] to compute an affine invariant skeleton of plane curves.