AFFINE INVARIANT MATHEMATICAL MORPHOLOGY APPLIED TO A GENERIC SHAPE RECOGNITION ALGORITHM

J.L. LISANI*, L. MOISAN*, P. MONASSE* and J.M. MOREL*
CMLA ENS-Cachan, 61 av du Pdt Wilson, 94235 Cachan Cedex, France.

Abstract. We design a generic contrast and affine invariant planar shape recognition algorithm. By generic, we mean an algorithm which delivers a list of all shapes two digital images have in common, up to any affine transform or contrast change. We define as "shape elements" all pieces of level lines of the image. Their number can be drastically reduced by using affine and contrast invariant smoothing Matheron operators, which we describe as alternate affine erosions-dilations. We then discuss an efficient local encoding of the shape elements. We finally show experiments. Applications aimed at include image registration, image indexing, optical flow.

Key words: shape recognition, contrast and affine invariant, partial occlusion

1. Introduction

Recently, various strategies to rigorously define distances between shapes have been proposed[25]. This distance method allows large nonparametric deformations. In this communication, we shall restrict ourselves to the case where perturbations boil down to contrast changes, planar affine transforms and occlusions. This restrictive framework is just sufficient to recognize an image which has undergone a xerocopy or a photograph (if it is a painting) and is thereafter subject to contrast changes and an arbitrary framing (occlusion on the boundary). The affine invariant framework is a well acknowledged topic[3, 4, 12, 13].

The restrictions we are taking are not arbitrary, but result from a hopefully rigorous invariance analysis. We first argue that the local contrast invariant information of an image is completely contained in its level lines ([5, 6]), which turn out to be Jordan curves. In order to overcome the occlusion phenomena, we wish to have an encoding as local as possible. The locality is obtained by segmenting each level line into its smallest meaningful parts which must finally be described by small codes. The curve segmentation-encoding process must therefore be itself invariant.

Moreover, the description of the curves must involve some smoothing since level lines are influenced by the quantization process. Thus, smoothing must be performed in order to get rid of this influence. Another reason to smooth shapes, is given by the "scale space ideology"[24]. Indeed, many of the fine scale oscillations of the shapes may be parts of the shape; the analysis of the

* {lisani,moisan,monasse,morel}@cmla.ens-cachan.fr

shape would be lost in those details.

Following [1], the only contrast invariant, local, smoothing and affine invariant scale space leads to a single PDE,

$$\frac{\partial u}{\partial t} = |Du|\mathrm{curv}(\mathrm{u})^{\frac{1}{3}}, \tag{1}$$

where Du is the gradient of the image, curv(u) the curvature of the level line and t the scale parameter. This equation is equivalent to the "affine curve shortening" ([22])

$$\frac{\partial x}{\partial t} = |\mathrm{Curv}(\mathrm{x})|^{\frac{1}{3}}\vec{\mathrm{n}}, \tag{2}$$

where x denotes a point of a level line, Curv(x) its curvature and $\vec{n}$ the signed normal to the curve, always pointing towards the concavity.

This equation is the only possible smoothing under the invariance requirements mentionned above. This gives a helpless bottleneck to the local shape recognition problem, since it is easily checked ([1]) that no further invariance requirement is possible. Despite some interesting attempts [10], there is no way to define a projective invariant local smoothing. The use of curvature-based smoothing for shape analysis is not new[2, 14, 9].

The contrast invariance requirement leads us to describe the shapes in terms of mathematical morphology[23]. In [7], connected components of level sets are proven to be invariant under contrast changes and [6] proposed to take as basic elements of an image the boundaries of the level sets (the so called **level lines**), a complete representation of the image which they call **topographic map**. A fast algorithm for the decomposition of an image into connected components of level lines is described in [20] and its application to a semi-local scale-space representation in [21]. Each one of these connected components is a closed Jordan curve and in many cases, we shall identify the term "shape" with these Jordan curves.

In Section 2, a fast algorithm to perform equation (2) is derived by going back to the mathematical morphology formalism ([23, 16]) and defining first an affine distance and then affine erosions and dilations. This leads us to an axiomatic justification for a fast algorithm introduced by Moisan ([17, 18]). This presentation follows the general line of a book in preparation [11].

In Section 3, we explain how to segment the smoothed curves into affine invariant parts and how these pieces of level lines can be encoded in an efficient way for matching. Section 4 gives a first account of what can be done with the generic algorithm.

2. Affine invariant mathematical morphology and PDE's

We first define an "affine invariant distance" which will be a substitute to the classical euclidean one. We consider shapes X, subsets of $\mathbb{R}^2$. Let $x \in \mathbb{R}^2$ and Δ an arbitrary straight line passing by x. We consider all connected components of $\mathbb{R}^2 \setminus (X \cup \Delta)$. If $x \notin \bar{X}$, exactly two of them contain x

in their boundary. We denote them by $\mathrm{CA}_1(\mathrm{x},\Delta,\mathrm{X})$, $\mathrm{CA}_2(\mathrm{x},\Delta,\mathrm{X})$ and call them the "chord-arc sets" defined by x, Δ and X, and we order them so that $\mathrm{area}(\mathrm{CA}_1(\mathrm{x},\Delta,\mathrm{X})) \leq \mathrm{area}(\mathrm{CA}_2(\mathrm{x},\Delta,\mathrm{X}))$.

Definition 1 *Let X be a "shape" and $x \in \mathbb{R}^2, x \notin \bar{X}$. We call affine distance of x to X the (maybe infinite) number $\delta(x,X) = \inf_\Delta \mathrm{area}(\mathrm{CA}_1(\mathrm{x},\Delta,\mathrm{X}))^{1/2}$, $\delta(x,X) = 0$ if $x \in X$.*

Definition 2 *For $X \subset \mathbb{R}^2$. We call affine a-dilate of X the set $\tilde{D}_a X = \{x, \delta(x,X) \leq a^{1/2}\}$. We call affine a-eroded of X the set $\tilde{E}_a X = \{x, \delta(x,X^c) > a^{1/2}\} = (\tilde{D}_a X^c)^c$.*

Proposition 1 *$\tilde{E}_a$ and $\tilde{D}_a$ are special affine invariant (ie they commute with area preserving affine maps) and monotone operators.*

Proof 1 *It is easily seen that if $X \subset Y$, then for every x, $\delta(x,X) \geq \delta(x,Y)$. From this, we deduce that $X \subset Y \Rightarrow \tilde{D}_a X \subset \tilde{D}_a Y$. The monotonicity of $\tilde{E}_a$ follows by the duality relation $\tilde{E}_a X = (\tilde{D}_a X^c)^c$. The special affine invariance of $\tilde{D}_a$ and $\tilde{E}_a$ follows from the fact that if* $\det \mathrm{A} = 1$, *then* $\mathrm{area}(\mathrm{X}) = \mathrm{area}(\mathrm{AX})$.

Remark 1 *One can show that $\tilde{E}_a$ and $\tilde{D}_a$ are affine invariant in the sense of Definition 14.19, in [11] that is, for every linear map A with $\det A > 0$, $A\tilde{E}_{(\det A)^{1/2}a} = \tilde{E}_a A$.*

We shall now use Matheron Theorem (Theorem 6.2 in [11]) in order to give a standard form to $\tilde{E}_a$ and $\tilde{D}_a$.

Definition 3 *We say that B is an affine structuring element if 0 is in interior of B, and if there is some $b > 1$ such that for every line Δ passing by 0, both connected components of $B \setminus \Delta$ containing 0 in their boundary have an area larger or equal to b. We denote the set of affine structuring elements by $\mathbb{B}_{\mathrm{aff}}$.*

Proposition 2 *For every set X,*

$$\tilde{E}_a X = \bigcup_{B \in \mathbb{B}_{\mathrm{aff}}} \bigcap_{y \in a^{1/2}B} X - y = \{x, \exists B \in \mathbb{B}_{\mathrm{aff}}, x + a^{1/2}B \subset X\}$$

Proof 2 *We simply apply Matheron theorem. The set of structuring elements associated with $\tilde{E}_a$ is $\mathbb{B} = \{X, \tilde{E}_a X \ni 0\}$. Now,*

$$\tilde{E}_a X \ni 0 \Leftrightarrow \delta(0,X^c) > a^{1/2} \Leftrightarrow \inf_\Delta \mathrm{area}(\mathrm{CA}_1(0,\Delta,X))^{1/2} > a^{1/2}$$

This means that for every Δ, both connected components of $X \setminus \Delta$ containing 0 have area larger than some number $b > a$. Thus, X belongs to $a^{1/2}\mathbb{B}_{\mathrm{aff}}$ by definition of $\mathbb{B}_{\mathrm{aff}}$.

By Proposition 2, x belongs to $\tilde{E}_a X$ if and only if for every straight line Δ, chord-arc sets containing x have an area strictly larger than a. Conversely we can state :

Corollary 1 *$\tilde{E}_a X$ is obtained from X by removing, for every straight line Δ, all chord-arc sets contained in X which have an area smaller or equal than a.*

2.1. Application to curve affine erosion/dilation schemes

Let c_0 be a Jordan curve, boundary of a simply connected set X. Iterating affine erosions and dilations on X gives a numerical scheme that computes the affine shortening c_T of c_0 at a given scale T. In general, the affine erosion of X is not simple to compute, because it can be strongly non local. However, if X is convex, then it has been shown in [18] that it can be exactly computed in linear time. In practice, c will be a polygon and the exact affine erosion of X —whose boundary is made of straight segments and pieces of hyperbolae— is not really needed; numerically, a good approximation of it by a new polygon is enough. Now the point is that we can approximate the combination of an affine erosion plus an affine dilation of X by computing the affine erosion of each *convex component* of c, provided that the erosion/dilation area is small enough. The algorithm consists in the iteration of a four-steps process:

1. **Break the curve into convex components.**
2. **Sample each component.**
3. **Apply discrete affine erosion to each component.**
4. **Concatenate the pieces of curves obtained at step 3.**

• **Discrete affine erosion**. This is the main step of the algorithm: compute quickly an approximation of the affine erosion of scale σ of the whole curve. The first step consists in the calculus of the "area" A_j of each convex component $\mathcal{C}^j = P_0^j P_1^j ... P_{n-1}^j$, given by $A_j = \sum_{i=1}^{n-2} \left[P_0^j P_i^j, P_0^j P_{i+1}^j \right] /2$. Then, the effective area used to compute the affine erosion is $\sigma_e = \max \{\sigma/8, \min_j A_j\}$. We restrain the erosion area to σ_e because the simplified algorithm for affine erosion may give a bad estimate of the continuous affine erosion+dilation when the area of one component is less than the erosion parameter. The term $\sigma/8$ is rather arbitrary and guarantees an upper bound to the number of iterations required to achieve the final scale. The discrete erosion of each component is defined as the succession of each middle point of each segment $[AB]$ such that

1. A and B lie on the polygonal curve
2. A or B is a vertex of the polygonal curve
3. the area enclosed by $[AB]$ and the polygonal curve is equal to σ_e

• **Iteration of the process**. To iterate the process, we use the fact that if E_σ denotes the affine erosion plus dilation operator of area σ, and $h = (h_i)$ is a subdivision of the interval $[0, H]$ with $H = T/\omega$ and $\omega = \frac{1}{2} \left(\frac{3}{2}\right)^{2/3}$, then

$$E_{(h_1-h_0)^{3/2}} \circ E_{(h_2-h_1)^{3/2}} \circ ... \circ E_{(h_n-h_{n-1})^{3/2}} \left(c_0\right) \longrightarrow c_T$$

as $|h| = \max_i h_{i+1} - h_i \to 0$, where c_T is the affine shortening of c_0 described above by (2).

The algorithm has linear complexity in time and memory, and its stability is ensured by the fact that each new curve is obtained as the set of the middle points of some chords of the initial curve, defined themselves by an integration process (an area computation). Hence, no derivation or curvature computation appears in the algorithm.

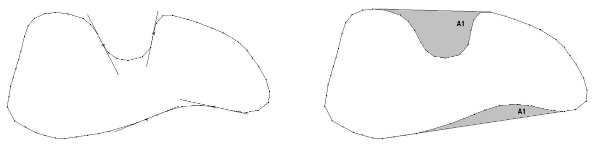

Fig. 1. Inflexion points (marked with small triangles) and bitangents of a closed curve. The area defined by each bitangent and the original curve is marked ($A1$).

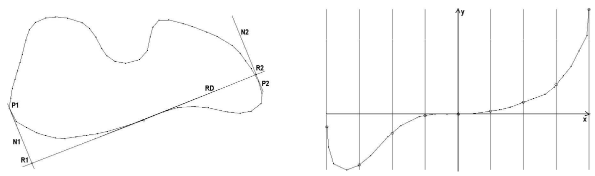

Fig. 2. Left: Local reference system for similarity invariant normalization: reference direction (RD), normal directions (N1, N2) and reference points (R1, R2). The portion of the curve normalized with this reference system starts at P1 and ends at P2, passing through the inflexion point. Right: Similarity invariant normalization. The y-ordinate of the marked points is used to encode the piece of curve.

3. Algorithms for the description of the shapes in an image.

3.1. Similarity invariant description of curves

In the search for an invariant description of a curve, the starting point for the sampling must be invariant, and so must be the sampling mesh. Typically, inflexion points have been chosen because they are affine invariant. Now, since the curve is almost straight at inflexion points, their position is not robust, but the direction of the tangent to the curve passing through them is. Another affine invariant robust semilocal descriptor is given by the lines which are bitangent to the curve (see Fig. 1).

Our reference system is formed by such a line, and the next and previous tangents to the curve which are orthogonal to it (see Fig. 2). The intersections of each one of these lines with the reference line provide two reliable points independent of the discretization of the curve. The portion of the curve to be normalized is limited by these points. Normalization consists in a similarity transform that maps the reference line to the x-axis and that sets the distance between the two reference points to 1. We discretize each one of the normalized portions of the curve with a fixed number n of points, and we store, for each discretized point, its y coordinate (see Fig. 2). This set of n values is used to compare portions of curves.

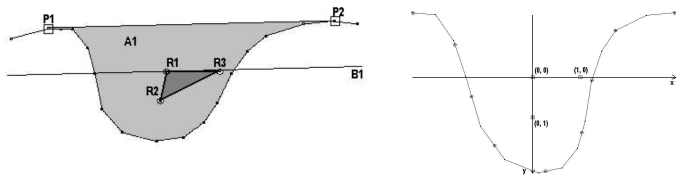

Fig. 3. Left: Local reference system for affine invariant normalization: reference points (R1, R2, R3). The portion of the curve to be encoded has endpoints P1 and P2. Right: Affine invariant normalization. The length of the normalized piece of curve together with the x and y coordinates of the marked points are used to locally encode the curve.

3.2. Affine invariant description of curves

If we look at Fig. 1, we can observe that the portion of the curve between the points defining the bitangent, together with the bitangent itself, define an area (*A1*), from which further invariant features can be computed. In particular, we can compute the barycenter of this area, an affine invariant reference point. We compute then the line *B1* parallel to the bitangent and passing through the barycenter. *B1* divides the initial area into two parts and we compute the barycenter of the part which does not contain the bitangent (see Fig. 3). This second barycenter is a second reference point. Finally a point in line *B1* such that the area of the triangle formed by this point and the two preceding barycenters is a fixed fraction of the initial area *A1* is a third reference point (see Fig. 3). We therefore obtain three nonaligned points, that is an affine reference system. This strategy is related to [8]. The discretization points are taken at uniform intervals of length on the normalized curve. The total length of the normalized curve is also used in the code. This set of $2n + 1$ values is what we use to compare portions of curves.

4. Experimental results

Figure 4 displays a picture of a man and the same picture after an occlusion of the face with his forearm and their level lines after smoothing with the iterative scheme described in section 2. Clearly some level lines have suffered a significant occlusion, and, even if some parts of the level line remain unchanged, registration methods based on global matching would fail in detecting those lines. In Figure 5, we show the result of the matching of several pieces of an occluded level line with other pieces of level lines in the second image.

References

1. L. Alvarez, F. Guichard, P.L. Lions, J.M. Morel, *Axioms and Fundamental Equations on Image Processing*, Technical Report 9231, CEREMADE, 1992 and Arch. for Rat. Mech. Anal. 16(9), 200-257, 1993.

Fig. 4. Up: Original images (from the film 'Analyze This' (Warner Bros)). Down: their smooth level lines (smoothing method of section 2).

2. H. Asada and M. Brady. *The curvature primal sketch.* IEEE Trans. Pattern Analysis and Machine Intelligence, 8(1), 2-14, 1986.
3. K. Astrom. *Affine and projective normalization of planar curves and regions*, ECCV 94, pp. B :439-448, 1994.
4. S.K. Bose, K.K. Biswas and S.K. Gupta. Model-based object recognition : the role of affine invariants. AIEng, 10(3) : 227-234, 1996.
5. V. Caselles, B. Coll, and J.M. Morel. *A Kanisza programme.* Progress in Nonlinear Differential Equations and their Applications, 25, 1996.
6. V. Caselles and B. Coll and J.M. Morel, *Topographic maps*, preprint CEREMADE, 1997. To appear in I.J.C.V.
7. V. Caselles and J.L. Lisani and G. Sapiro and J.M. Morel, *Shape preserving histogram modification*, IEEE Trans. on Image Processing, February 1999.
8. Cohignac, T. and Lopez, C. and Morel, J.M., *Integral and Local Affine Invariant Parameter and Application to Shape Recognition*, ICPR pp. A:164-168, 1994.
9. G. Dudek and J.K. Tsotsos. *Shape representation and recognition from multiscale curvature*, CVIU, 68(2), pp. 170-189, 1997.
10. O. Faugeras and R. Keriven, *Some recent results on the projective evolution of 2D curves.* In Proc. IEEE International Conference on Image Processing, vol. 3, pp. 13-16, Washington, October 1995.
11. F. Guichard and J.M. Morel, *Image iterative filtering and PDE's*, Preprint, 1999. Book in preparation.
12. R.A. Hummel, H.J. Wolfson. *Affine invariant matching.* DARPA88, pp. 351-364, 1988.
13. P. Kempenaers, L. Van Gool, and A. Oosterlinck. *Shape recognition under affine distortions.* VF91, pp. 323-332, 1991.
14. A. Mackworth and F. Mokhtarian, *A theory of multiscale, curvature-based shape representation for planar curves*, IEEE Trans. Pattern Analysis and Machine Intelligence, 14 : 789-805, 1992.

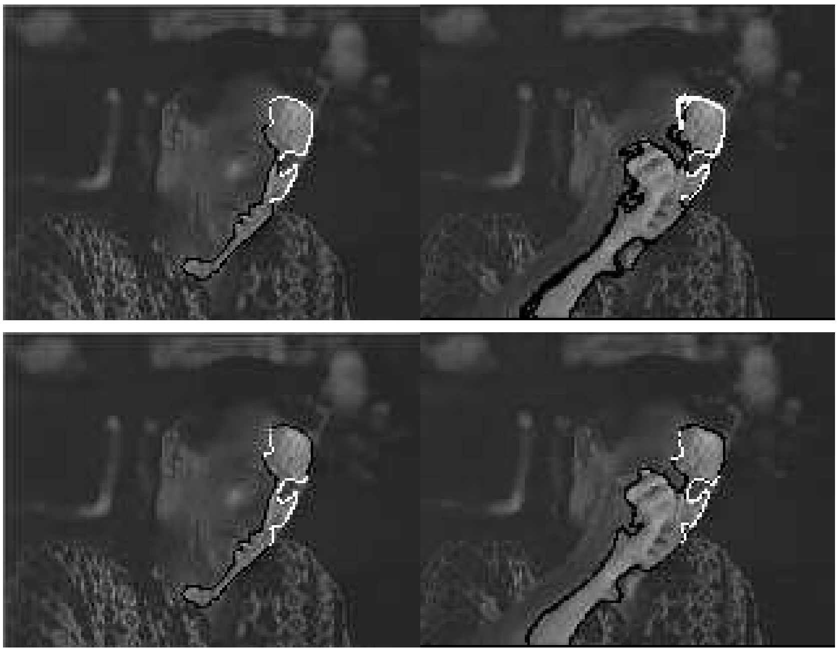

Fig. 5. Correspondences (marked in white, the remaining parts of the curves in black) between some pieces of a level line in the first image (left) and some pieces of other level lines in the second image (right). Up: By using the similarity invariant registration method based on inflexion points. Down: By using the affine invariant registration method based on bitangents.

15. G. Koepfler, L. Moisan, *Geometric Multiscale Representation of Numerical Images*, Proc. of the Second Int. Conf. on Scale-Space Theories in Computer Vision, in Springer Lecture Notes in Computer Science, vol. 1682, pp. 339-350, 1999.
16. G. Matheron, *Random Sets and Integral Geometry*, John Wiley, N.Y., 1975.
17. L. Moisan, *Traitement numérique d'images et de films: équations aux dérivées partielles préservant forme et relief*, PhD dissertation, Université Paris-Dauphine, France, 1997.
18. L.Moisan, *Affine Plane Curve Evolution : a Fully Consistent Scheme*, IEEE Transactions On Image Processing, vol. 7:3, pp. 411-420, 1998.
19. P. Monasse, *Contrast Invariant Image Registration*, Proc. of I. V. Conf. on Accoustics, Speech and Signal Processing, Phoenix, Arizona, vol 6, 1999, pp. 3221-3224.
20. P. Monasse and F. Guichard, *Fast Computation of a Contrast-Invariant Image Representation*, to appear in IEEE Transactions on Image Processing, 1998, preprint CMLA 9815, available at *http://www.cmla.ens-cachan.fr*
21. P. Monasse and F. Guichard, *Scale-Space from a Level Lines Tree*, Proc. of 2nd Int. Conf. on Scale-Space Theories in Computer Vision, Corfu, Greece, 1999, pp. 175-186.
22. G. Sapiro and A. Tannenbaum, *Affine Invariant Scale Space*, International Journal of Computer Vision, 11 (1), 25-44, 1993.
23. J.Serra, *Image Analysis and Mathematical Morphology*, Academic Press, 1982.
24. A.P. Witkin, *Scale space filtering*, Proc. IJCAI, 1019-1023, 1983.
25. L. Younes *Computable elastic distances between shapes*, SIAM J. of Applied Mathematics, 1998.