A Distance for Elastic Matching in Object Recognition

Robert Azencott, François Coldefy, Laurent Younes

Abstract

We define distances between geometric curves by the square root of the minimal energy required to transform one curve into the other. The energy is formally defined from a left invariant Riemannian distance on an infinite dimensional group acting on the curves, which can be explicitely computed. The obtained distance boils down to a variational problem for which an optimal matching between the curves has two be computed. An analysis of the distance when the curves are polygonal leads to a numerical procedure for the solution of the variational problem, which can efficiently be implemented, as illustrated by experiments.

Acknowledgements: This work comes as a part of a project on object recognition, supported by the D.R.E.T, helded at SUDIMAGE research laboratory under the direction of R. Azencott. Publication is courtesy of the D.R.E.T.

1. Introduction

1.1. Generalities

The problem of matching two objects together is very important in computer vision and shape recognition. In many cases, recognition is based on shapes (outlines), with the help of some suitably designed distance. A general principle is to associate to any pair (C_1, C_2) of objects to be compared a measure of discrepancy $d(C_1, C_2)$. The recognition of an observed object C may be done by finding, from a dictionary of "templates", the previouly recorded object C_{temp} for which $d(C, C_{temp})$ is minimal. Clearly, the definition of the distance is the crucial step of the method, and many researches have been done in this direction. See, for example [12] for a review of the huge litterature existing on the subject.

Instead of basing recognition on a finite collection of points of interest (primitives) taken from the outline of an object (corners, inflexion points, ...), which is a popular way of handling the problem, our purpose is to base the comparison on the whole outline, considered as a plane curve. The distance we shall define incorporates some deformation energy between the curves. The approach, as we will see, comes out to be intrinsic, and robust to usual Euclidean transformations.

The method is related to the wide litterature on "snakes" ([10], [4],...), in the way that our distance corresponds to some continuous process of deformation of one curve into another. It is also related to papers on elastic matching such as [5]; however, we provide an elastic matching algorithm which is based on a true distance between intrinsic properties of the shapes, taking into account possible invariance to scaling or Euclidean transformations in the case they are required. From this point of view, our results are indebted to the seminal work of Grenander on group theory applied to pattern recognition (cf. [7], and [8], in particular).

1.2. Principles of the approach

Denote by $\mathcal{C}$ our object space, which will be the set of arc-length parametrized plane curves. Assuming that each object in $\mathcal{C}$ can in some way be deformed into another, our purpose is to define a distance measuring the amount of deformation which is necessary for this operation. The deformations are represented by a *group action*

$$\begin{aligned} G \times \mathcal{C} &\longrightarrow \mathcal{C} \\ (a, C) &\longrightarrow a.C \end{aligned}$$

on $\mathcal{C}$, where G is a group (the fact that we have a group action means that, for all $a, b \in G$, and for all $C \in \mathcal{C}$, $a.(b.C) = (ab).C$ and $e.C = C$, e being the identity element of G). The fact that each object can be deformed into another means that the group action is transitive, that is, we assume that, for all C_1, C_2 in $\mathcal{C}$, there exists a in G such that $a.C_1 = C_2$.

Assume that we have some way of measuring the cost of the transformation $C \rightarrow a.C$, and denote this cost by $\Gamma(a, C)$. To compare two objects, we put

$$d(C_1, C_2) = \inf\{\Gamma(a, C_1), C_2 = a.C_1\} \qquad (1)$$

which is the smallest cost required to deform C_1 into C_2. If G acts transitively on $\mathcal{C}$, it can be easily proved that, if E is

*This work comes as a part of a project on object recognition, supported by the D.R.E.T, helded at SUDIMAGE research laboratory. Publication is courtesy of the D.R.E.T. R.Azencott and L. Younes: Groupe DIAM, CMLA (URA 1611), ENS de Cachan, 61 av. du President Wilson, 94235 Cachan CEDEX, France, and SUDIMAGE, 47 av Carnot, 94230 Cachan, France.
F. Coldefy: SUDIMAGE.
E-mail : Laurent.Younes@ens-cachan.fr

a function defined on $G \times \mathcal{C}$, taking values in $[0, +\infty[$, such that : i) For all $C \in \mathcal{C}$, $\Gamma(e, C) = 0$ — ii) For all $a \in G$, $C \in \mathcal{C}$, $\Gamma(a, C) = \Gamma(a^{-1}, a.C)$ — iii) For all $a, b \in G$, $C \in \mathcal{C}$, $\Gamma(ab, C) \leq \Gamma(b, C) + \Gamma(a, bC)$; then d defined by (1) is symmetric, satisfies the triangle inequality and is such that $d(C, C) = 0$ for all $C \in \mathcal{C}$.

Moreover, when G acts transitively on $\mathcal{C}$, we know, from elementary group theory, that $\mathcal{C}$ can be identified (at least as a set) to a coset space on G. Indeed, fixing some reference element $C_0 \in \mathcal{C}$, and letting $H_0 = \{h \in G, h.C_0 = C_0\}$, $\mathcal{C}$ can be identified to G/H_0 through the well-defined correspondance $a.H_0 \leftrightarrow a.C_0$. Using this identification, one can define a cost function Γ provided that G itself is equipped with a left-invariant distance d_G (such that for some function $\gamma_0 : G \to \mathbb{R}$ with $\gamma_0(h) = 1$ if $h \in H_0$, the identity $d_G(ca, cb) = \gamma(c) d_G(a, b)$ is true on G) ; for this, define for $C \in \mathcal{C}$, with $C = b.C_0$ $\Gamma_0(a, C) = d_G(e, a^{-1})/\gamma(b)$. Then Γ_0 satisfies properties i), ii) and iii) above. The obtained distance is, if $C = b.C_0$,

$$d_0(C, C') = \gamma(b)^{-1} \inf\{d(e, a), aC' = C\} .$$

To define a suitable left invariance distance on G, we base our intuition on standard constructions of differential geometry, thinking of G as a Lie group acting on $\mathcal{O}$, on which we define an invariant metric characterized by its trace on the Lie algebra of G. However, in order to make an elementary description, it suffices to remark that, to define a geodesic distance on G, we only need to know how to compute the lengths of smooth paths in G. If $\mathbf{a} = (\mathbf{a}(t), t \in [0, 1])$ is such a path, subject to suitable regularity conditions which will depend on the context, we must define the length $L(\mathbf{a})$ and then set

$$d_G(a_0, a_1) = \inf\{L(\mathbf{a}), \mathbf{a}(0) = a_0, \mathbf{a}(1) = a_1\}$$

the infimum being computed over some set of *admissible paths*.

We would like to define the length of a path $t \to \mathbf{a}(t)$ by the formula

$$\int_0^1 ||\dot{\mathbf{a}}_t(t)|| dt , \tag{2}$$

for some norm. Thinking of $\dot{\mathbf{a}}_t(t) dt$ as a way to represent the portion of path between $\mathbf{a}(t)$ and $\mathbf{a}(t + dt)$, defining the norm corresponds to defining the cost of a small variation of $\mathbf{a}(t)$. Note that we must have

$$d_G(\mathbf{a}(t), \mathbf{a}(t + dt)) = \gamma(\mathbf{a}(t)) d_G(e, \mathbf{a}(t)^{-1} \mathbf{a}(t + dt)) ,$$

so that the problem is to define γ and the cost of a small variation from the identity.

For this purpose, we have considered some small deformation of a plane curve C into some plane curve $C + \delta C$, for which we define, in some natural way, its cost Γ. We then represent $C \to C + \delta C$ as a transformation $C \to a.C$ for some group action on $\mathcal{C}$ and identify the cost Γ under the form $\Gamma = \tilde{\gamma}(C)^{-1} d_G(e, a^{-1})$ at least for a close to identity. Now, having chosen a reference curve C_0, and expressed $C = b.C_0$, we see that we have to set $\gamma(b) = \tilde{\gamma}(C)$.

2. Comparison of plane curves

2.1. Infinitesimal deformations

Consider a plane curve in parametric form

$$C = \{m(s) = (x(s), y(s)),\ s \in [0, l]\} .$$

We assume that the parametrization is done by arc-length, that is (denoting by f_s the derivative of a function f with respect to s): $\dot{x}_s^2 + \dot{y}_s^2 \equiv 1$, so that l is the Euclidean length of the curve C.

Assume that, by some deformation, each point of C is moved, the displacement being given by some vector field carried by C (ie. a function $V(s) = (u(s), v(s))$, considered as a vector starting at the point $m(s)$). A new curve is obtained, which is

$$\tilde{C} = \{\tilde{m}(s) = (x(s) + u(s), y(s) + v(s))\} .$$

The field V (and its derivatives) is infinitely small.

We define the *energy* of this deformation by

$$\delta E(V) = \int_0^l ||\dot{V}_s(s)||^2 ds ,$$

and its cost by the square root of this energy.

Denote by g^* the function, $g^* : [0, l] \to [0, \tilde{l}]$, which associates to s the arc length $\tilde{s}$ in $\tilde{C}$ of the point $m(s) + V(s)$. Denote also by $\theta^*(s)$ the angle between the tangent to C at point $m(s)$ and the axis $y = 0$. Let $\tilde{\theta}^*(\tilde{s})$ be the similar function defined for $\tilde{C}$. One can show that, at first order

$$\delta E = \int_0^l (\dot{g}_s^* - 1)^2 ds + \int_0^l (\tilde{\theta}^* \circ g^*(s) - \theta^*(s))^2 ds ,$$

We separate l from the other quantities by letting $g(s) = g^*(ls)/\tilde{l}$ (which is defined on $[0, 1]$ and takes values in $[0, 1]$), $\lambda = \tilde{l}/l$, $\theta(s) = \theta^*(ls)$ and $\tilde{\theta}(\tilde{s}) = \tilde{\theta}^*(\tilde{l}\tilde{s})$ (which also are defined on $[0, 1]$). Still at order one, we have

$$\delta E(\lambda, g, D, l) = l(\lambda - 1)^2 + l \int_0^1 \left[(\dot{g}_s - 1)^2 + D^2\right] ds ,$$

with $D(s) = \tilde{\theta} \circ g(s) - \theta(s)$.

We now can identify a group action. We characterize a curve C up to translations by the pair $(l, \zeta(.))$ in which

$l \in]0, +\infty[$ is its length and $\zeta = e^{i\theta}$ where $\theta(.)$ is its normalized angle function, defined on $[0, 1]$. Now, define the action

$$(\lambda, g, r).(l, \zeta) = (l/\lambda, r.\zeta \circ g) \tag{3}$$

where $\lambda > 0$, g is a diffeomorphism of $[0, 1]$ and r is a measurable function, defined on $[0, 1]$, with complex values and $|r| \equiv 1$. Let G be the set composed with these 3-uples. The product

$$(\tilde{\lambda}, \tilde{g}, \tilde{r}).(\lambda, g, r) = (\lambda\tilde{\lambda}, g \circ \tilde{g}, \tilde{r}.r \circ \tilde{g}) \tag{4}$$

is a group operation, with identity $e_G = (1, \mathrm{Id}, \mathbf{1})$ (where $\mathrm{Id}(s) = s$ and $\mathbf{1}(s) = 1$), and inverse $(\lambda, g, r)^{-1} = (1/\lambda, g^{-1}, \bar{r} \circ g^{-1})$ where $\bar{r}$ is the complex conjugate of r ($\bar{r}.r \equiv 1$), and (3) is a group action. Now, letting $r = e^{-iD}$, the relation (3) may be rewritten

$$(l, \theta) \to (\lambda, g, r)^{-1}.(l, \theta)\,. \tag{5}$$

and our evaluation of small deformations implies that, if (λ, g, r) is close to $(1, \mathrm{Id}, \mathbf{1})$, the effect of the action on (l, θ) is

$$\delta E(\lambda, g, r, l) = l(\lambda - 1)^2 + l \int_0^1 \left[(\dot{g}_s - 1)^2 + |r - 1|^2\right] ds\,, \tag{6}$$

with the first order approximation $|e^{-iD} - 1| \simeq |D|$.

Letting $a = (\lambda, g, r)$, we thus have evaluated the cost of a small the deformation $C \to a^{-1}.C$ by

$$\Gamma(a^{-1}, C)^2 = l(\lambda - 1)^2 + l \int_0^1 \left[(\dot{g}_s - 1)^2 + |r - 1|^2\right] ds\,.$$

Let us choose a reference curve C_0 as the horizontal plane segment of length 1, that is $C_0 = (1, \mathbf{1})$. The curve $C = (l, \zeta)$ is equal to $b.C_0$ if and only if $b = (\lambda, g, r)$ with $\lambda = 1/l$ and $r = \zeta$. This above expression can be written, for a close to e,

$$\Gamma(a^{-1}, C)^2 = d_G(e, a)^2 / \gamma(b)^2$$

with

$$\gamma(b) = 1/\sqrt{l} \tag{7}$$

and

$$d_G(e, a)^2 = (\lambda - 1)^2 + \int_0^1 \left[(\dot{g}_s - 1)^2 + |r - 1|^2\right] ds\,. \tag{8}$$

2.2. Definition of a distance

We now are in position to apply the paradigm of the introduction : translating this distance from the neighbourhood of e to the neighbourhood of any $a \in G$, defining the length of a path in G by the sum (the integral) of the lengths of infinitely small portions of the path, and obtaining the geodesic distance d_G between two elements a and b of G as the minimal length of the paths which join them. The surprising fact is that this distance can be explicitely computed. In [15], it is shown that, for $a = (\lambda, g, e^{i\Delta})$, $b = (\mu, h, e^{i\tilde{\Delta}})$,

$$[d_G(a, b)]^2 = 2(\lambda + \mu - 2\sqrt{\lambda\mu} \int_0^1 \sqrt{\dot{g}_s \dot{h}_s}\,|\cos(\frac{\Delta - \tilde{\Delta}}{2})|ds \tag{9}$$

As a consequence, one defines a distance between two plane curves $C = (l, e^{i\theta})$ and $\tilde{C} = (\tilde{l}, e^{i\tilde{\theta}})$,

$$[d(C, \tilde{C})]^2 = l + \tilde{l} - 2\sqrt{l\tilde{l}} \sup_g \int_0^1 \sqrt{\dot{g}_s(s)} \left|\cos \frac{\tilde{\theta} \circ g(s) - \theta(s)}{2}\right| ds\,. \tag{10}$$

the supremum being taken over functions g which are increasing diffeomorphisms of $[0, 1]$.

2.3. Definition of distances with invariance restrictions

We now modify the previous distance by requiring some additional Euclidean invariance properties.

For example, the same analysis provides a scale invariant distance by putting

$$d^{(1)}(C, \tilde{C}) = 2 \inf_g \arccos \int_0^1 \sqrt{\dot{g}_s(s)} \left|\cos \frac{\tilde{\theta} \circ g(s) - \theta(s)}{2}\right| ds\,. \tag{11}$$

the infimum being taken over functions g which are strictly increasing diffeomorphisms of $[0, 1]$.

Finally, a rotation invariance distance is obtained by

$$d^{(0)}(C, \tilde{C}) = 2 \inf_g \min_{c \in]-\pi, \pi]} \arccos \int_0^1 \sqrt{\dot{g}_s(s)} \left|\cos \frac{\tilde{\theta} \circ g(s) - \theta(s) - c}{2}\right| ds\,. \tag{12}$$

3. Numerical implementation

3.1. Case of polygonal curves

Given two functions θ and $\tilde{\theta}$, the core of the numerical problem is to compute $\sup_g Q(g)$ when

$$Q(g) = \int_0^1 \sqrt{\dot{g}_s(s)} \left|\cos \frac{\tilde{\theta} \circ g(s) - \theta(s)}{2}\right| ds\,.$$

This is not trivial, since the functional to maximize is not concave, and even not differentiable because of the absolute value. The approach we have used is to approximate our curves by polygons, for which some explicit computation may be carried on. Assume that both curves are piecewise linear, ie. that θ and $\tilde{\theta}$ are piecewise constant. Then, there exists $0 = s_0 < s_1 < \cdots < s_m = 1$ (resp. $0 = \tilde{s}_0 < \tilde{s}_1 < \cdots < \tilde{s}_n = 1$) and constants $\theta_1, \ldots, \theta_m$ (resp. $\tilde{\theta}_1, \ldots, \tilde{\theta}_n$) such that $\theta(s) \equiv \theta_i$ on $[s_{i-1}, s_i[$ (resp. $\tilde{\theta}(s) \equiv \tilde{\theta}_j$ on $[\tilde{s}_{j-1}, \tilde{s}_j[$). Let g be an increasing diffeomorphism of $[0,1]$, and let $\tau_i = g^{-1}(s_i)$, $i = 1, \ldots, m$. One can compute the best function g when the τ_i are fixed. This gives an expression of $Q(g)$ as a function of τ which is

$$Q(g) = \sum_{i=1}^{m} \sqrt{(s_{i+1} - s_i)Q_i}\,. \tag{13}$$

with

$$\begin{aligned} Q_i \;=\; & \cos^2 \frac{\tilde{\theta}_{j(i)+1} - \theta_{i+1}}{2}(\tilde{s}_{j(i)+1} - \tau_i) \\ & + \sum_{j=j(i)+1}^{j(i+1)-1} \cos^2 \frac{\tilde{\theta}_{j+1} - \theta_{i+1}}{2}(\tilde{s}_{j+1} - \tilde{s}_j) \\ & + \cos^2 \frac{\tilde{\theta}_{j(i+1)} - \theta_{i+1}}{2}(\tau_{i+1} - s_{j(i+1)}) \end{aligned}$$

Where $j(i)$ is the index j such that $\tilde{s}_j \le \tau_i < \tilde{s}_{j+1}$. We see that there exists a combinatorial part in the maximization of Q, which is due to the $j(i), i = 1, \ldots, m$. Each $j(i)$ may take any value between 1 and n, with the constraint that $j(1) \le j(2) \le \cdots \le j(m)$. If the $j(i)$ are fixed, the τ_i may be obtained by the maximization of a smooth function, with the constraint that, for all i

$$\max(s_{j(i)}, \tau_{i-1}) \le \tau_i < \min(s_{j(i)+1}, \tau_{i+1})\,. \tag{14}$$

Now, $Q(g)$, as given in equation (13), may very quickly be maximized by linear programming, when the number of edges in the polygonal curve is not too large. When the number of edges is large, a suboptimal steepest-descent procedure may be used.

3.2. General curves

When one deals with general differentiable curves, each of them may be replaced by a polygonal approximation. We generally use a multi-scale approach, starting with a rough polygonal approximation for which dynamic programming can be used, and then refine the result for enhanced approximations by steepest-descent.

To estimate the rotation parameter c in the expression of $d^{(0)}$, we start with an initial value c_0 and find the optimal g with this fixed c_0, and then compute the best c given g. The procedure can be iterated a few times.

4. Experiments

We present examples from a small database composed with 8 outlines of planes, for 4 types of planes. The shapes have been extracted from 3D synthesis images, under two slightly different view angles for each plane. We have applied some smooth stochastic noise to the outlines in order to obtain variants of the same shape which look more realistic. The outlines are presented in figure 1. The lengths of the curves have been computed after smoothing (using a cubic-spline representation). The complete matrix of distances has been computed on this database, and is given in table 1. We see that the distance between a plane and the other one from the same class is always smaller than between any plane in another class. The computed distance is $d^{(0)}$ (insensitive to rotations), since the orientations of the planes vary.

	sailplan-1	sailplan-2	aero-1	aero-2	x29-1	x29-2	bomb-1	bomb-2
sailplan-1	0	0.25	0.43	0.46	0.79	0.73	0.9	0.81
sailplan-2	0.25	0	0.47	0.48	0.71	0.69	0.77	0.82
aero-1	0.43	0.47	0	0.28	0.76	0.8	0.77	0.81
aero-2	0.46	0.48	0.28	0	0.79	0.77	0.78	0.76
x29-1	0.79	0.71	0.76	0.79	0	0.38	0.84	0.81
x29-2	0.73	0.69	0.8	0.77	0.38	0	0.82	0.8
bomber-1	0.9	0.77	0.77	0.78	0.84	0.82	0	0.29
bomber-2	0.81	0.82	0.81	0.76	0.81	0.8	0.29	0

Table 1. Matrix of distances (in radian) within the plane database

References

[1] Y. Amit, U. Grenander, M. Piccioni : Structural image restoration through deformable templates. Preprint (1989).

[2] V.I. Arnol'd : *Méthodes mathématiques de la mé-*

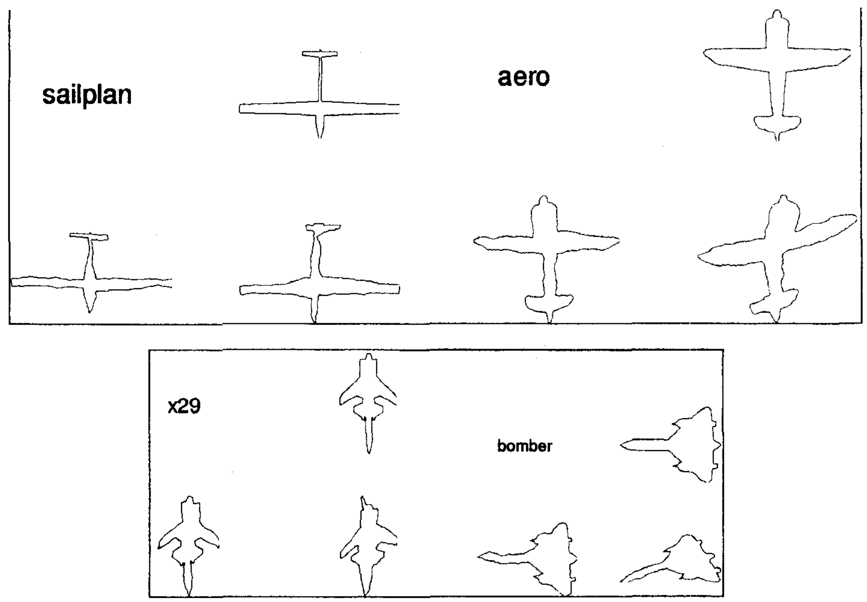

Figure 1. Outlines of planes from 4 classes. For each type of plane: upper-right original view (from above); lower-left : view from above degraded by smooth noise ; lower-right : slight variation of the angle of view and noise. The compared outlines are lower-left and lower-right

canique classique. Mir. Moscow (1976).

[3] R. Azencott : Deterministic and random deformations ; applications to shape recognition. Conference at HSSS workshop in Cortona, Italy (1994).

[4] L. D. Cohen and I. Cohen : Finite element methods for active contour models and balloons for 2D and 3D images. *IEEE PAMI* 15, (1993).

[5] I. Cohen, N. Ayache, P. Sulger : Tracking points on deformable objects using curvature information, *in Computer Vision - ECCV'92* G. Sandini (ed.) pp 458-466 (1992).

[6] M. P. Do Carmo : *Riemannian geometry*. Birkaüser.

[7] U. Grenander : *Lectures in Pattern theory*, Springer Verlag, Berlin (1976, 1978, 1981).

[8] U. Grenander and D. M. Keenan : Towards automated image understanding : *J. Appl. Stat.*, vol 16, No 2, 207-221 (1989).

[9] U. Grenander and D. M. Keenan : On the shape of plane images : Siam J. Appl. Math. vol 53, No 4 1072-1094 (1991).

[10] M. Kass, A. Witkin and D. Terzopoulos : Snakes: active contour models. *Int. J. of Comp. Vision*, 1, 321-331 (1988).

[11] D. G. Kendall : Shape manifolds, procrustean metrics and complex projective spaces. *Bull. London Math. Soc., 16, 81-121* (1984).

[12] C. G. Perrot and L. G. C. Hamey : Object recognition: a survey of the litterature. Macquarie computing reports No 91-0065C (1991).

[13] A. Trouvé : Infinite Dimensional Group Action and Pattern Recognition. Preprint (1995).

[14] A. L. Yuille : Generalized deformable models, Statistical Physics, and matching problems. *Neural computations*, 1, 1-24 (1990).

[15] L. Younes : Computable elastic distances between shapes, To appear in *SIAM J. of Applied Math.*