Classification
Physics Abstracts
42.30.Sy — 02.40.Ft — 02.40.Pc

Geometric Shape Recognition and Classification

Marie-José Labouré, Isabelle Fillère, Michel Jourlin and Jean-Marie Becker

Laboratoire Image, Signal et Acoustique (LISA), CPE-Lyon, 43 boulevard du 11 novembre, 69616 Villeurbanne Cedex, France

Résumé. — Parmi les méthodes de reconnaissance et de classification de formes planes, nous avons étudié plus particulièrement les méthodes géométriques. Nous avons développé une technique de vectorisation de contours utilisant des points de "forte" courbure. Les formes ainsi résumées peuvent être comparées à des formes références au moyen de paramètres de forme, de coefficients de symétrie et de métriques dans un but de classification ou de reconnaissance.

Abstract. — Among pattern recognition and classification methods, geometric methods are more particularly developped. Firstly an edge vectorization technique based on "high" curvature points has been studied. Summarized shapes can then be compared to reference shapes using shape parameters, symmetry coefficients and metrics in order to establish a shape classification.

1. Introduction

In the field of image analysis the lack of mathematical tools often forbids rigorous developments. Moreover, the set of convex bodies has been thoroughly studied providing a lot of concepts and techniques whose interest for applications is undoubtful.

Binary shapes (simply connected compact set in $\mathbb{R}^2$, with a non empty interior) given by their contour are presently studied. In order to reduce the amount of information, the shape is firstly summarized by some characteristic edge points. Then the shapes are compared in order to conclude about their similarity either from metrics or their membership of a given class or shape parameters. Our researches are based on old papers from famous mathematicians (Besicovitch, Brunn, Minkowski, ...). They were adapted to the frame, developped, improved and implemented. Our own shape parameters and our own metrics were also created.

2. Edge Vectorization

Information on the shape edge is concentrated at dominant points. A recursive algorithm using convex hull procedure is used [1]. The first step consists in determining the farthest points of the convex hull of the contour (*i.e.* the edge points realizing the maximal Euclidean distance with at

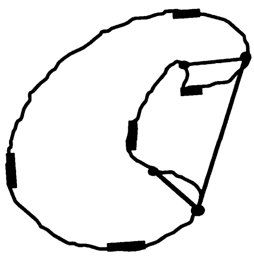


Fig. 1.

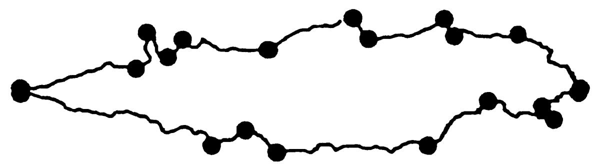


Fig. 2.

Fig. 1. — Vectorization process, dominant points are bold.

Fig. 2. — Original shape (turbulent jet) and dominant points.

least one point of the contour). These points are high curvature points [2]. Then, at each step the cavities which compose the difference between the current shape and its convex hull are studied. For each cavity the farthest points from the boarding segment and its extremities are retained as dominant points (see Figs. 1 and 2). When the difference between the cavity and its convex hull has a too small area, the algorithm stops. The dominant points are often high curvature points [2], but not always. By joining these points, a polygonal approximation is obtained, which is very close to the initial contour specially for rough contours.

The advantages of this method are: independence of scale, no curvature computation, rapidity and the fact that a small number of points is sufficient to rebuild accurately the original shape.Then later computations (shape parameters for example) are simplified.

3. Metrics

In shape recognition a notion of distance between two bodies is needed, to evaluate their relative position or overlapping, to compare their geometric shapes, ... One should think that the most interesting distance to use is the Euclidean distance, but in the space of compact sets it does not yield a distance. Thus, specific metrics have been defined and adapted to different kinds of problems: the Hausdorff distance [3] which allows to estimate for example the relative position of two shapes, the Asplund distance [4] which can gauge one shape with regard to the other, or the symmetric difference distance which allows to evaluate the degree of overlapping by measuring an area. To *determine the degree of similarity* of two shapes, these metrics must be extended to the set of compact sets equal up to a similarity: the two shapes are first best superimposed and then the distance is computed.

The metric study and the definition of their extension led us to the study of the analogy between metrics on compact sets and functional metrics [5]. It also led to the creation of new distances [6]:

- the generalization of Hausdorff metric to non circular structuring elements: this proves that using a square (or any convex body) on the grid yields a real distance,

- the radial distance which is based on the radial function adapted to star-shaped bodies (see Fig. 3 for more explanations).

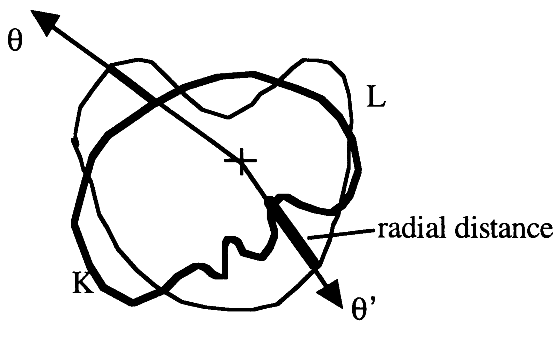


Fig. 3.

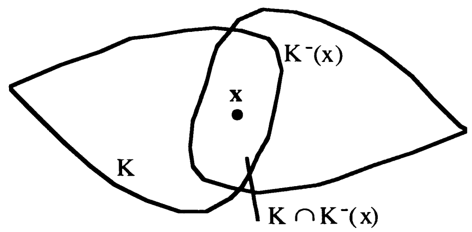


Fig. 4.

Fig. 3. — Radial distance is the greatest difference between the radial functions.

Fig. 4. — $\mathrm{K} \cap \mathrm{K}^-(x)$ is one of the central convex bodies contained in K. $b(\mathrm{K}) = \mathrm{Sup}_{x \in \mathrm{K}} \dfrac{\mu(\mathrm{K} \cap \mathrm{K}^-(x))}{\mu(\mathrm{K})}$.

4. Shape Parameters

Image processing uses shape parameters (defined in $\mathbb{R}^2$) in order to give a shape classification, or more simply, a proximity degree of the studied shape to a reference one. If the reference shape is a disk these parameters are circularity parameters and if the reference shape is a centrally symmetric convex body they are coefficients of symmetry. Let us recall that a positive real valued function defined on the set of planar shapes is a shape parameter provided it is scale invariant; moreover, this parameter applied to convex bodies is a coefficient of symmetry if its values are in [0, 1], the value 1 characterizing the centrally symmetric convex bodies.

Let us recall some coefficients of symmetry: let K be a convex body.

• Besicovitch coefficient [7] is the ratio between the area of the maximal central convex body included in K and the area of K (see Fig. 4). This maximum is obtained when K and its symmetric set with respect to a point x are best overlapped. The coefficient is then $b(\mathrm{K}) = \dfrac{\mu(\mathrm{K} \cap \mathrm{K}^-)}{\mu(\mathrm{K})}$ and varies in [2/3, 1] (μ is the area). The value 2/3 characterizes triangles.

• Winternitz coefficient [8]: if one considers the minimal ratio between the left and right areas in K which are delimitated by a cutting line through a given interior point x, the coefficient of Winternitz $w(\mathrm{K})$ is the maximum of these ratios for x varying in K (see Fig. 5). It varies in [4/5, 1], the value 4/5 characterizes triangles.

• Minkowski coefficient [8]: it presents symmetry concept as the research of the best centered point in the band delimitated by the two support lines of the convex set (see Fig. 6). The coefficient can be expressed by $m(\mathrm{K}) = \mathrm{Sup}_{x \in \mathrm{K}} \, \mathrm{Inf}_{\theta \in [0,2\pi]} \dfrac{h(x,\theta)}{h(x,\theta+\pi)}$ where h is the support function of K: the distance between x and the support line of direction θ of K. It varies in [1/2, 1], the value 1/2 characterizes triangles.

The triangle is the most asymmetric convex shape for all these coefficients and each of them leads to a unique point in K which is the best centered. These best centered points are generally different except for the triangle. In this latter case they are all equal to the centroid.

Specific parameters can also be used to evaluate other geometric properties: let P be the perimeter, R the circumradius and r the inradius of K.

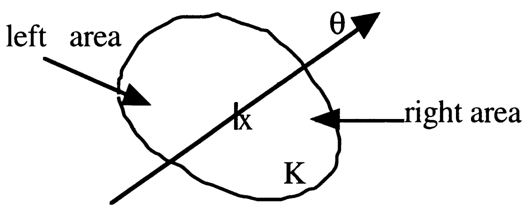


Fig. 5.

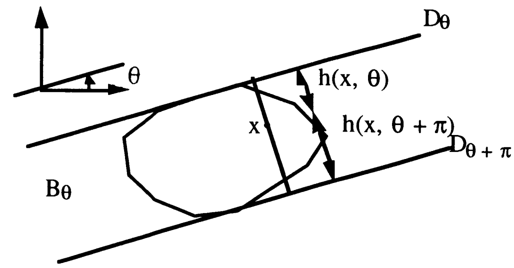


Fig. 6.

Fig. 5. — $w(\mathrm{K}) = \mathrm{Sup}_x \mathrm{Inf}_\theta$(right area/left area).

Fig. 6. — Minkowski coefficient. $m(\mathrm{K}) = \mathrm{Sup}_{x\in\mathrm{K}}\ \mathrm{Inf}_{\theta\in[0,2\pi]} \dfrac{h(x,\theta)}{h(x,\theta+\pi)}$.

As the parameters $I_0 = r/R$, $I_1 = P^2/4\pi\mu$ were not sufficient to conclude about the circularity of a contour, new shape parameters have also been created: they are issued from old or new isoperimetric inequalities [9]: $I_2 = PR/(2\mu), I_3 = Pr/(\mu + \pi r^2)$ evaluate the proximity of a convex shape K to a disk (the value 1 characterizes the disk).

Some of these shape parameters are defined in the case of n-sided convex polygons; they are issued from equalities about inradius and circumradius of regular n-sided polygons [10]; they evaluate the proximity to a regular polygon, the value 1 characterizes a regular polygon:

$$I_4 = 2\mu/(n \sin\frac{2\pi}{n}R^2) \quad I_5 = \mu/(n \tan\frac{\pi}{n}r^2) \quad I_6 = P/(2n \sin\frac{\pi}{n}R)$$

$$I_7 = P/(2n \tan\frac{\pi}{n}r) \quad I_8 = P^2/(4n \tan\frac{\pi}{n}\mu).$$

Then, one can associate to each convex body a vector of which coordinates are the parameters I_j. If the reference shape is the disk, the distance between our shape and the disk is evaluated by calculating $\Sigma|I_j - 1|$ for $j = (0), 1, \ldots, 3$ and if the reference shape is a regular polygon $\Sigma|I_j - 1|$ for $j = 4, \ldots, 8$ is calculated. Then it is possible to classify convex shapes in the order of circularity and convex polygonal shapes in the order of regularity (see examples in Tabs. I, II). In fact, when performing a granulometric analysis on a binary image, a practical issue consists in classifying the studied particles according to their size or Féret diameter. The main interest of the present approach is to supplement the classical criteria by circularity and regularity informations.

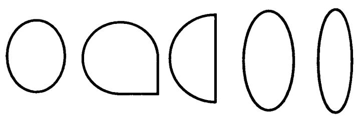
5. Conclusion

During last years old symmetry coefficients were re-used and we proposed new shape parameters in order to classify planar compact sets. For best results one has to use several parameters. The implementation is easier if the contour is polygonal, hence the interest of a good vectorization. Metrics are also useful to compare similar shapes or to evaluate the error between a contour and its polygonal approximation.

Table I. — *C: classification in circularity order.*

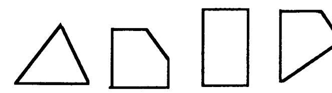
Table II. — *R: classification in regularity order.*

Tab. I.



I_1	1.04	1.07	1.13	1.17	1.59
I_2	1.20	1.16	1.21	1.52	2.26
I_3	1.00	1.03	1.06	1.02	1.06
$\Sigma(I_j-1)$	0.24	0.26	0.4	0.71	1.91
C	1	2	3	4	5

Tab. II.



I_4	0.87	0.74	0.92	0.72
I_5	1.11	1.26	1.50	1.19
I_6	0.98	0.89	0.98	0.93
I_7	1.11	1.17	1.25	1.19
I_8	1.11	1.08	1.04	1.19
$\Sigma\lvert I_j-1\rvert$	0.48	0.88	0.89	0.92
R	1	2	3	4

References

[1] Labouré MJ., Azema J. and Fournel T., Detection of dominant points on a digital closed curve, *Acta Stereol.* **11** (1992) 169-174.

[2] Labouré MJ., Hivert M., Jourlin M. and Zeboudj R., High curvature subsets of the boundary of a convex body, *Acta Stereol.* **8** (1989) 83-88.

[3] Hausdorff F., *Math. Ann.* **79** (1919) 157.

[4] Asplund E., Comparison between plane symmetric bodies and parallelogramms, *Math. Scand.* **8** (1960) 171-180.

[5] Jourlin M., Métriques de formes et applications, Congrès International sur l'Analyse en Distance, Distancia '92 (1992) pp. 279-281.

[6] Fillère I. and Labouré MJ., A different approach of distances in the set of planar bodies, *Acta Stereol.* **12** (1993) 89-94.

[7] Besicovitch AS., Measure of asymmetry of convex curves, *J. London. Math. So.* **23** (1948) 237-240.

[8] Grünbaum B., Measure of symmetry for convex sets, *Proc. Symp. Pure. Math.* **7** (1963) 233-270.

[9] Labouré MJ., Jourlin M., Fillère I., Becker JM. and Frenea F., Isoperimetric inequalities and shape parameters, *Acta Stereol.* **15** (1996) 65-70.

[10] Fejes-Toth L., Lagerungen in der Ebene auf der Kugel und im Raum (Springer Verlag Berlin-Göttingen-Heidelberg, 1953).