Shape Description by Bending Invariant Moments

Paul L. Rosin

School of Computer Science & Informatics, Cardiff University, Cardiff, UK
Paul.Rosin@cs.cf.ac.uk

Abstract. A simple scheme is presented for modifying geometric moments to use geodesic distances. This provides a set of global shape descriptors that are invariant to bending as well as rotation, translation and scale.

Keywords: moments, transformation invariance, bending, articulation.

1 Introduction

In the literature there is a huge range of techniques for shape description, not only for computer vision [14,20,23], but also in particle science [22], medical imaging [3], geography [4], art [1], etc. Shape descriptors can be categorised into global or local methods. Global methods have the advantage of simplicity, and are efficient to both store and match. Local methods provide a richer representation, and so are sometimes more capable of performing object discrimination. However, computing the descriptor tends to require some parameters (e.g. a scale parameter), and matching such descriptors can be computationally expensive.

A desirable property for shape descriptors is that they are invariant to certain transformations of the shape in the image. Invariance to translation, rotation, scale, and possibly skew have become standard practice. However, for shapes that undergo articulation or bending deformation, although there exist methods for matching local descriptors [5,11,15], there is little work on invariant global descriptors [19]. In computer vision textbooks [21] the only instance given is the Euler number, which is generally not sufficiently discriminative.

Moments are widely used as shape descriptors [9,17,18]. Central moments provide translation invariance, and further moment invariants were developed to provide additional invariance to rotation and scale [13] and skew [8]. The contribution of this paper is to describe a simple scheme to enable shapes to be described by geometric moments that are invariant to bending in addition to translation, rotation and scale.

2 Bending Invariant Moments

Our starting point is radial or rotational moments that are generally defined in terms of polar coordinates (r, θ) over the unit disk as

A. Berciano et al. (Eds.): CAIP 2011, LNCS 6854, pp. 253–260, 2011.

$$R_{pq} = \int_0^{2\pi} \int_0^1 r^p e^{iq\theta} f(r\cos\theta, r\sin\theta) r dr d\theta.$$

Since the polar angles of points will vary under bending, we discount these values and just use distances over the 2D image plane. Given a discretised two dimensional shape S, the bending invariant moments are

$$A_p = \sum_{\mathbf{x}_i \in S} d_g(\mathbf{x}_i, \mathbf{x}_c)^p$$

where $\mathbf{x}_c$ is the centre of S. To provide invariance to bending, the Euclidean radial distances have been replaced by geodesic distances. The geodesic distance $d_g(\mathbf{x}_1, \mathbf{x}_2)$ between two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in S is the length of the shortest path between $\mathbf{x}_1$ and $\mathbf{x}_2$ such that the path lies within the S. This path will consist of linear segments and possibly sections of the boundary of S.

The centre also needs to be chosen such that it is invariant to bending. For the Euclidean metric, the centroid is the point minimising the sum of squared distances, while the geometric median is the point that minimises the sum of distances (for which there is no direct closed-form expression). In a similar manner, we choose

$$\mathbf{x}_c = \arg\min_{\mathbf{y} \in S} \sum_{\mathbf{x}_i \in S} d_g(\mathbf{x}_i, \mathbf{y}).$$

Since geodesic distances have been used, the above guarantees that A_p is invariant to translation, rotation and bending. In addition, a normalisation is applied to provide invariance to scale:

$$\alpha_p = \frac{A_p}{A_0^{\frac{p+2}{2}}}$$

where A_0 is the area of S.

Standard geometric image moments of a 2D shape are defined as

$$m_{pq} = \sum_{(x_i, y_i) \in S} x_i^p y_i^q$$

where the distances from the origin are measured along the Cartesian axes. We modify the moments A_p to loosely follow the concept of geometric moments, using something like a set of curvilinear axes. The first axis is the geodesic path to the centre $\mathbf{x}_c$ (used for A_p). The second axis at $\mathbf{x}_i$ is determined locally as the shortest geodesic path from $\mathbf{x}_i$ to the boundary of S. Thus, the second bending and scale invariant moment is

$$\beta pq = \frac{B_{pq}}{B_0^{1+\frac{p+q}{2}}}$$

where

$$B_{pq} = \sum_{\mathbf{x}_i \in S} d_g(\mathbf{x}_i, \mathbf{x}_c)^p d_g(\mathbf{x}_i, E)^q$$

and E is the exterior of S. This provides a more general version of α_p since $\beta_{p0} = \alpha_p$.

To ensure that all points in S are reachable from the centre by geodesic paths we require that S is a single connected component (although holes are allowed). The geodesic centre is not unique – a simple counterexample is given by an annulus. However, this has not proven to be a problem in our experiments.

The only other global bending invariant descriptors of shapes in the computer vision literature that we are aware of is the recent work by Rabin *et al.* [19]. They also use geodesic distances, but they are computed between pairs of points. To reduce computational complexity they perform uniform furthest point sampling of the points in S, and compute geodesics between these points and the full point set in S. Three quartiles of these distances are taken as global descriptors.

Other related work is by Gorelick *et al.* [12]. Instead of the distance transform or geodesic paths they used the Poisson equation which generates expected distances travelled by particles following a random walk. Gorelick *et al.* also computed shape descriptors using moments, but they were not concerned with invariance to bending. Rather than directly computing moments from the random walk distances they used several features derived from the distances. Since several of these features involved the local orientation of the distance field, and combinations of local orientations, the features (and their moments) were not bending invariant. Moreover, the shape was centred at the (standard) centroid which is bending invariant.

3 Implementation Details

To compute the geodesic distances there are several approaches. Long and Jacobs [15] created a graph whose vertices were sample points. Edges were included in the graph if the straight line between a pair of sample points (corresponding to the vertices) was completely included within the shape. They only used sample points on the boundary, and the cost of both graph construction and determination of the shortest paths between all pairs of points is $O(n^3)$ for n boundary points.

However, in order to reduce the effects of noise, we wish to use geodesic distances between interior points as well as boundary points. This will ensure that the centre will be relatively insensitive to boundary noise. An extreme case is shown in figure 1, in which half the circle has been perturbed. The point minimising the summed geodesic distance to the boundary is significantly affected, unlike the summed geodesic distance to all points in the shape.

We compute geodesic distances using a distance transform capable of operating in non-convex domains. Algorithms for computing the geodesic distance transform using ordered propagation are available with computation complexity linear in the number of pixels [6][1]. Finding the centres of shapes is the most computationally expensive step ($O(n^2)$ for a shape containing n points). Therefore

[1] For simplicity, a geodesic distance transform was used. It performs ordered propagation with a heap based priority queue, and produces approximate geodesic distances.

(a) (b)

Fig. 1. Centres determined by minimising the geodesic distances to (a) the shape boundary, (b) the shape interior. The geodesic distances to the centres are displayed inside the shapes.

a multiresolution strategy is used. The centre is first found at a low resolution after the image has been downsampled so that the shape contains about 5000 pixels[2]. The centre position is then refined at full resolution within a $3F \times 3F$ window, where F is the ratio of high to low resolution. To perform downsampling, each pixel in the low resolution image is set to the foreground value if any of its parent pixels in the high resolution image are foreground. This ensures that the downsampled shape remains a single connected component (even though the topology may change, i.e. holes can be removed).

4 Experiments

We start by showing the effectiveness of the geodesic centre. Despite the shape in figure 2 undergoing articulation and deformation the centre of the shape is captured reliably.

Next, we show classification results for Ling and Jacob's dataset of articulated shapes [15] which consists of 40 shapes made up of 8 classes each containing 5 examples (see figure 3). As a baseline, comparison is made with two methods that are not invariant to bending: the standard Hu moments invariants and Belongie *et al.*'s shape context [2]. In addition, comparison is made with two bending invariant methods: Ling and Jacob's inner-distance (geodesic) extension of shape context [15], and Rabin *et al.*'s work [19] which extracts a 4D description from geodesic quartiles and uses stochastic gradient descent to compute the Wasserstein distance between distributions.

For our classifier we compute the following moments: all A_p of order 1 to 9, all B_{pq} of order 1 to 4 (i.e. 14 moments), and the first 7 Hu moment invariants. For each type of moment we choose the combination of up to 5 moments that

[2] Choosing 5000 pixels for the low resolution version of the shape provides a reasonable compromise between maximising accuracy and computational efficiency. The value was experimentally determined as suitable for all the data tested in this paper.

Fig. 2. Despite undergoing articulation and deformation the shape's centre is correctly determined

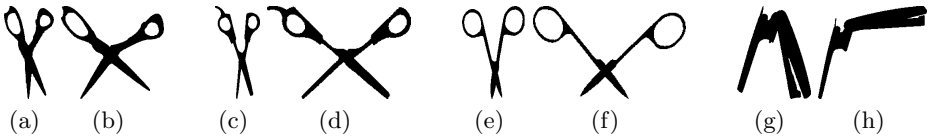

Fig. 3. Examples of pairs of shapes from four classes from Ling and Jacob's dataset of articulated shapes

gives the best classification. Nearest neighbour classification is performed using Mahalanobis distances and leave-one-out cross validation.

The bending moment invariants perform very well, as do the geodesic quartiles, and are only marginally outperformed by the inner distance shape context; see table 1.

Table 1. Classification accuracies for Ling and Jacob's dataset of articulated shapes

Method	**Accuracy (%)**
Hu moment invariants	70
shape context	50
inner distance shape context	100
4D geodesic quartiles	97
α_p	97
β_{pq}	97

Classification performance is evaluated on a second dataset of articulated shapes provided by Gopalan *et al.* [11]. However, since their images contain substantial segmentation errors producing many small holes and components, we have pre-processed the images to extract only the single largest foreground object – examples are shown in figure 4. Gopalan *et al.* report recognition rates

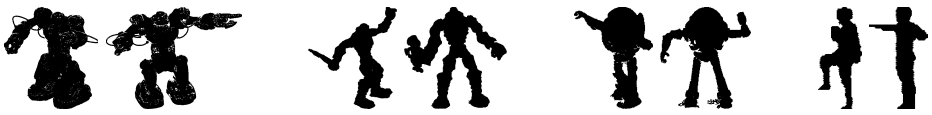

Fig. 4. Examples of pairs of shapes from four classes from Gopalan *et al.*'s dataset of articulated shapes. Only the single largest foreground object from each original image has been retained.

Table 2. Classification accuracies for Gopalan *et al.*'s dataset of articulated shapes

Method	**Accuracy (%)**
Hu moment invariants	72
shape context	82
inner distance shape context	90
α_p	92
β_{pq}	100

Table 3. Bulls-eye test scores for MPEG-7 CE-1 database

Method	**Bulls-eye scores**
Hu moment invariants	46
shape context	76
inner distance shape context	85
ASC & LCDP	96
4D geodesic quartiles	60
α_p	38
β_{pq}	61

of 58% for the inner distance shape context and 80% for their own method, although our experiments (on the cleaned data) using the inner distance shape context gave a much higher accuracy (90%). Again, as table 2 shows, the bending moment invariants perform very well.

Our final evaluation was performed on the MPEG-7 CE-1 database 1400 shapes. As table 3 shows, the invariance to bending does allow β_{pq} to demonstrate clear improvements on the Hu moment descriptors which are only invariant to similarity transformations. However, the bending moment invariants do not perform as well as many other reported methods. The reason is that for some of the object classes there is considerable variation in shape which cannot be captured easily by global descriptors. Note that Rabin *et al.*'s geodesic quartiles give a similar accuracy to β_{pq}. Approaches based on local matching such as shape context can generate excellent accuracies. Currently, the best performance for the MPEG-7 CE-1 database has been achieved by Lin *et al.*'s aspect shape context (ASC) [16] which was combined with locally constrained diffusion

process (LCDP) to achieve a bulls-eye score of 96%. However, such methods are computationally expensive.

5 Conclusions

This paper describes an approach to generate global shape descriptors that are invariant to bending, rotation, translation and scale. The shape centre is estimated as the point that minimises the summed geodesic distances to other points, and then moments are computed on the distances from the centre. The benefit of such global shape descriptors is that they tend to be much more efficient to match compared to local shape descriptors. For instance, matching by shape context requires dynamic programming, while both [19] and [5] use iterative matching schemes.

Several approaches in the literature could be adapted to the proposed framework. For instance, for describing 3D shapes Gal *et al.* [10] use a histogram of centricity, which is the average of the geodesic distance from a mesh vertex to all other vertices. The proposed moment descriptors could also be generated from centricity values instead of distances from the centre. However, since computing pairwise distances is computationally expensive this would have the disadvantage of requiring all the pairwise distances to be computed at full resolution.

An alternative to directly using the geodesic distances to modify the moments, is to use the geodesic distances along with multidimensional scaling to construct a bending invariant form of the shape [7,15]. Moments subsequently computed would therefore be bending invariant. Future work will investigate the effectiveness of this approach.

Finally, extending the proposed method to 3D shapes would be straightforward. The geodesic distance transform using ordered propagation can be readily and efficiently applied in higher dimensions, with run-time proportional to the number of elements (e.g. voxels).

References

1. Arnheim, R.: Art and Visual Perception: A Psychology of the Creative Eye. University of California Press, Berkeley (1974)
2. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. on Patt. Anal. and Mach. Intell. 24(4), 509–522 (2002)
3. Bookstein, F.L.: Shape and the information in medical images: A decade of the morphometric synthesis. Computer Vision and Image Understanding 66(2), 97–118 (1997)
4. Boyce, R.B., Clark, W.A.V.: The concept of shape in geography. Geographical Review 54, 561–572 (1964)
5. Bronstein, A.M., Bronstein, M.M., Bruckstein, A.M., Kimmel, R.: Analysis of two-dimensional non-rigid shapes. Int. J. of Computer Vision 78(1), 67–88 (2008)
6. Cárdenes, R., Alberola-López, C., Ruiz-Alzola, J.: Fast and accurate geodesic distance transform by ordered propagation. Image Vision Comput. 28(3), 307–316 (2010)

7. Elad, A., Kimmel, R.: On bending invariant signatures for surfaces. IEEE Trans. on Patt. Anal. and Mach. Intell. 25(10), 1285–1295 (2003)
8. Flusser, J., Suk, T.: Pattern recognition by affine moment invariants. Pattern Recognition 26, 167–174 (1993)
9. Flusser, J., Zitova, B., Suk, T.: Moments and Moment Invariants in Pattern Recognition. Wiley Publishing, Chichester (2009)
10. Gal, R., Shamir, A., Cohen-Or, D.: Pose-oblivious shape signature. IEEE Trans. Vis. Comput. Graph. 13(2), 261–271 (2007)
11. Gopalan, R., Turaga, P.K., Chellappa, R.: Articulation-invariant representation of non-planar shapes. In: Europ. Conf. Computer Vision, vol. 3, pp. 286–299 (2010)
12. Gorelick, L., Galun, M., Sharon, E., Basri, R., Brandt, A.: Shape representation and classification using the poisson equation. IEEE Trans. on Patt. Anal. and Mach. Intell. 28(12), 1991–2005 (2006)
13. Hu, M.: Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory 8(2), 179–187 (1962)
14. Kindratenko, V.V.: On using functions to describe the shape. J. Math. Imaging Vis. 18(3), 225–245 (2003)
15. Ling, H., Jacobs, D.W.: Shape classification using the inner-distance. IEEE Trans. on Patt. Anal. and Mach. Intell. 29(2), 286–299 (2007)
16. Ling, H., Yang, X., Latecki, L.J.: Balancing deformability and discriminability for shape matching. In: Europ. Conf. Computer Vision, vol. 3, pp. 411–424 (2010)
17. Mukundan, R., Ramakrishnan, K.R.: Moment Functions in Image Analysis – Theory and Applications. World Scientific, Singapore (1998)
18. Prokop, R.J., Reeves, A.P.: A survey of moment-based techniques for unoccluded object representation and recognition. CVGIP: Graphical Models and Image Processing 54(5), 438–460 (1992)
19. Rabin, J., Peyré, G., Cohen, L.D.: Geodesic shape retrieval via optimal mass transport. In: Europ. Conf. Computer Vision, vol. 5, pp. 771–784 (2010)
20. Rosin, P.L.: Computing global shape measures. In: Chen, C.H., Wang, P.S.-P. (eds.) Handbook of Pattern Recognition and Computer Vision, 3rd edn., pp. 177–196. World Scientific, Singapore (2005)
21. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision. Thomson-Engineering (2007)
22. Taylor, M.A.: Quantitative measures for shape and size of particles. Powder Technology 124(1-2), 94–100 (2002)
23. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recognition 37(1), 1–19 (2004)