The Representative Capacity of Parameters Derived from the Radial Signature

Christoph Rasche

Laboratorul de Analiza si Prelucrarea Imaginilor
Universitatea Politehnica din Bucuresti
Bucuresti 061071, RO
email: rasche15@gmail.com
fax: +4021 402 4821, phone: +4021 402 4645

Abstract

A method for the boundary representation of 'simple' shapes is presented. It is based on the radial signature and exploits its extrema information to arrive at a low-dimensional geometric description (ca. 10 dimensions). This short description can represent shapes well, which is demonstrated on the Corel and MPEG7 collection. Its key advantage is the short computation duration. If this description is extended by further radial-based parameters and combined with Fourier descriptors, it leads to almost the same retrieval performance as the best-performing signature approach, which also uses Fourier descriptors. In a classification task, the radial-based descriptors clearly outperform the Fourier descriptors.

Keywords: shape boundary, low-dimensional, retrieval, Fourier descriptors, set B of MPEG7

1. Introduction

A parametric description of 'simple', planar, closed or near-closed shapes is introduced. Here, simple is understood as circular or loosely cyclic, whereby we define the latter as simple (non self-intersecting) polygons or as star shapes (star-shaped, non self-intersecting polygons). If the shape is fragmented, its gap sizes should be small only, such that the shape appears as nearly closed. In this study, only shapes with a single gap are considered, that is closed or near-closed curves. Such curves occur in abundance in gray-scale images of real-world scenes, especially at a local scale, and are frequently returned by edge or contour (elevation) detection algorithms. If such curves are described appropriately, they may improve scene classification or at least the description of some scene categories.

The shape representation is based on the radial signature, which describes the distances of the pixels to the centroid (pole) of the shape, also called centroid-distance function or radial-distance signature in other studies. This or similar distance signatures are often used as a first step in building shape descriptions using the Fourier Transform, see [1, 2] for reviews. But here, the radial signature itself is analyzed and is parameterized based on its extrema and other geometric information. Some of the parameters are simple shape descriptions as used previously [3]. It is shown that this information per se is very representative.

In order to compare the representative capacity of our new description to other shape descriptions - that are based on distance signatures and Fourier descriptors -, we perform a shape retrieval task with the MPEG7 image set as it was done for [1, 2]. It will be demonstrated that if one uses the geometric parameters of the radial signature only, one already obtains results comparable to some other methods which employ much more processing to obtain their shape description. In a simple classification task, the geometric information clearly outperforms the Fourier descriptors.

2. Theory

The radial signature is generated only for curves that are closed or nearly closed. In case of the latter, the gap size is required to be smaller than a tolerance, which is dependent on the curve's total arc length. This therefore includes L features with a limited acute angle. If a gap is present, it is closed by a straight line, which is a crude but simple choice to arrive at a more complete description of the shape.

For a given curve, its pole is determined by taking the mean of all curve points. A radial signature $R(v)$ is formed, whereby v is the arc length variable; the signature is normalized by the average radius. The signature's extrema are determined and assigned to lists $R^{\max}$ and $R^{\min}$(maxima and minima respectively), with $\mathsf{n}_{\max}$ the number of maxima. The directional angles $\gamma(i)$ of the maxima are taken and their included angles $\alpha(i)$ between them determined and ordered by decreasing value ($i = 1..\mathsf{n}_{\max}$).

Five structural biases were created, which represent to what degree the shape corresponds to a specific (global) form; $\breve{s}_{\frown}$ represents curved shapes and may correspond to the bending energy [3]; $\breve{s}_2$ coarsely represents the degree of elongation (or eccentricity in [3]), an analogue measure to the aspect ratio; $\breve{s}_3$ represents triangle or deltoid shapes; $\breve{s}_4$ stands for a quadrilateral or astroid shape; $\breve{s}_5$ is for a pentagon or any star shape with five peaks. Thus, the bias number corresponds to the number of 'outer' corners (vertices) with angles smaller than π and we refer to those shapes sometimes as n-corner shapes.

For notational simplicity we use the following angles: supplementary angles α^c as a π-complement to the included angles α; interior angles $\beta^3 = 2\pi/3$, $\beta^4 = \pi/2$, $\beta^5 = 2\pi/5$ for expressing the degree of structural bias - they correspond to the interior angles for an equilateral triangle, a square and a regular pentagon.

- **Curved**, $\breve{s}_\frown$: the bias allows to express whether a shape is of type bean, bicorn, D or crescent and is only larger zero if the largest included angle $\alpha(1)$ is a reflex angle (larger than $\pi/2$). The bias corresponds to the difference above π:

$$\breve{s}_\frown = \begin{cases} \alpha(1) - \pi & , \alpha(1) > \pi \\ 0 & , \text{else.} \end{cases} \tag{1}$$

- **Two-corner**, $\breve{s}_2$: this bias is suitable to express shapes such as ovals, ellipses or U-turns. The bias is taken only if two or more maxima are present; its value depends on the complementary angles for the included angles, and is proportional to the elongation η and weight w:

$$\breve{s}_2 = \begin{cases} \frac{\pi - \alpha^c(1) - \alpha^c(2)}{\pi} w\eta & , \mathsf{n}_{\max} \geq 2 \\ 0 & , \text{else.} \end{cases} \tag{2}$$

The weight (or strength) equals the range of values for the 2nd derivative of the (normalized) radial signature: $w = \mathrm{rng}(R'')$. The weight so corresponds to the degree of 'peakness' of the shape; the elongation is a parameter by itself and is defined later.

- **Three-corner**, $\breve{s}_3$: expresses triangles or star shapes with three peaks. The bias decreases with inreasing asymmetry from an equilateral triangle (or deltoid or equivalent star shape):

$$\breve{s}_3 = \begin{cases} \frac{\beta^3 - \sum_{i=1}^{3}(\beta^3 - \alpha(i))}{\beta^3} w & , \mathsf{n}_{\max} \geq 3 \\ 0 & , \text{else.} \end{cases} \tag{3}$$

- **Four-corner**, $\breve{s}_4$: the bias corresponds to the ratio of the fourth included angle and the interior angle for a square:

$$\breve{s}_4 = \begin{cases} \frac{\alpha(4)}{\beta^4} w & , \mathsf{n}_{\max} \geq 4 \\ 0 & , \text{else.} \end{cases} \tag{4}$$

- **Five-corner**, $\breve{s}_5$: the bias is analogous to the one for the four-corner polygon:

$$\breve{s}_5 = \begin{cases} \frac{\alpha(5)}{\beta^5} w & , \mathsf{n}_{\max} \geq 5 \\ 0 & , \text{else.} \end{cases} \tag{5}$$

In summary, the corner biases have a large value for cyclic polygons (equilateral triangle, square, rectangle, regular pentagon,...). The bias decreases the more the shape deviates from the interior angle of the cyclic polygon. The bias also decreases the less 'acute' the interior angles are such as in a squircle (four-cornered wheel). Conversely, the bias grows very large if the corners become more acute such as in a hypocycloid (deltoid, astroid,...) or any star shape - due to the use of the derivative-dependent weight value w. A circle is expressed by all structural biases being 0.

The biases are not mutually exclusive. There are many shapes (in gray-scale images) that cannot be assigned clearly to a single bias only. The strength of the description lies in allowing multiple biases and avoids a 'feature classification' (see also [4] for explanations).

- **Elongation**, η: measures the spatial extent of the shape and is 0 for symmetric shapes such as circles, squares and pentagons; it is proportional to the range (rng) of radii otherwise:

$$\eta = \begin{cases} \text{rng}(R(v)) & , \breve{s}_2 > 0 \\ \max(\text{rng}(R^{\max}), \text{rng}(R^{\min})) & , \mathsf{n}_{\max} \geq 3 \\ 0 & , \text{else.} \end{cases} \tag{6}$$

For polygons with three or more corners, the range of the signature values is not a sufficient elongation measure, because an equiangular polygon (or in particular a hypocycloid) shows a certain range due to its corners but it has no true elongation. The elongation is therefore determined with the lists of maximum and minimum radii.

- **Symmetry**, y: is determined by two measures. One is the signature's irregularity, ι (iota), which is the integrated difference between the two signature halves whereby one half is inverted, whereby the point of havening is the total maximum of the radii $R^{\max}$. The other measure is the minimum of the (absolute) derivative for the (ordered) maximum radii:

$$y = 1 - \iota - \min(\Delta R^{\max}). \tag{7}$$

The symmetry is largest for any shape, that has at least two equal angles such as an isosceles triangle, a kite shape and so on.

The gap size - if present - is included as dimensions as well by the spatial and angular separation between the endpoints, ω_d and $\omega_\angle$. The angular separation may be zero but there may exist a spatial separation, in which case the curve corresponds to the beginning of a spiral shape.

The following vector is then formed, whose performance is later reported as 'RS geo simple':

$$\mathbf{s}(o, r, \eta, y, \omega_d, \omega_\angle, \breve{s}_\frown, \breve{s}_2, \breve{s}_3, \breve{s}_4, \breve{s}_5),$$

whereby o is the angular orientation of the global maximum's directional angle; r is the average of the (unnormalized) signature, normalized here by the image dimensions.

A longer, 28-dimensional vector is formed to improve the performance for the image retrieval task, whose performance is later reported as 'RS geo complex'. The 20 additional parameters are the degree of concavity, which is determined by the fraction of the signature whose course is in reverse direction to the dominant direction; the above-mentioned weight value w; the standard deviation of the angular directions of the corners; the angular gap, which is above zero in case of a horse-shoe shape; the mean, minimum, maximum and standard-deviation value for the maxima and minima radii ($R^{\max}$ and $R^{\min}$); and the mean, minimum, maximum and standard-deviation value for the maxima and minima in the second derivative. The creation of more parameters did not lead to better performance. For the task of shape retrieval with the MPEG7 collection, 3 dimensions of the above shape vector will be dropped, see Result section for details.

3. Methods

Radial Signatures (RS). Figure 1 demonstrates an example of the signature. The shape's signature was low-pass filtered by a Gaussian filter, whose size was proportional to curve length.

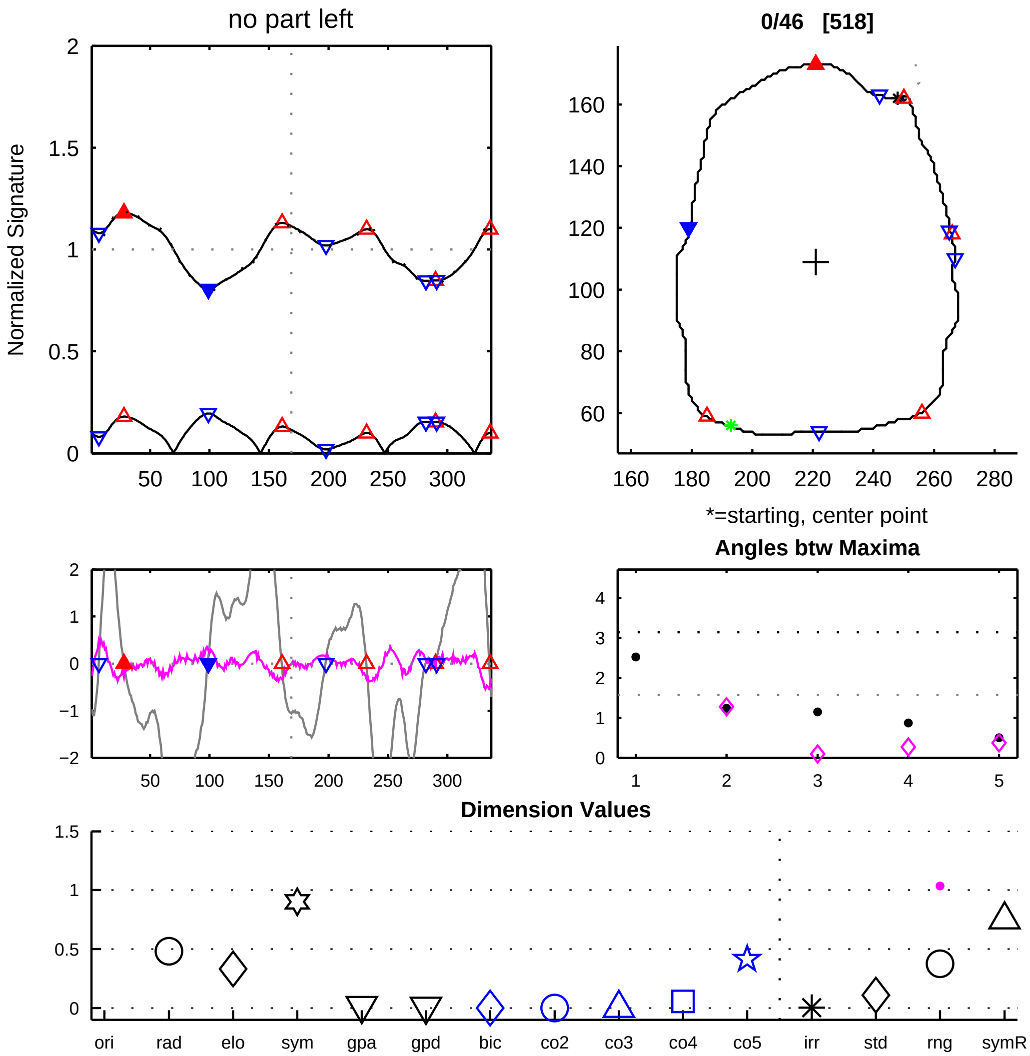

Figure 1: Radial signature, $R(v)$, and parameterization. **Top right**: curve with starting and center points marked as asterisk; upward and downward pointing triangles: maxima and minima, resp. plus sign: pole. **Top left**: $R(v)$; x-axis=arc length variable v [pixels]. **Left center**: 1st and 2nd derivative of $R(v)$. **Angles btw Maxima**: sorted, included angles between directional angles of the maxima (direction from pole). **Dimension Values** (aspects): ori=orientation; rad=avg. radius; elo=elongation; sym=symmetry; gpa=angular gap; gpd=spatial gap; bic=curved; co2, co3, co4, co5=corner biases. The remaining values are shown for illustration only - they are not used as parameters per se: irr=irregularity; std=std($R(v)$); rng=rng($R(v)$);

Fourier Descriptor (FD). The Fourier descriptors for an (unfiltered) radial signature were generated as specified in [1, 2]. The Discrete Fourier Transform was taken for the shape's full boundary length. Its coefficients FD_k ($k = 0, 1, .., N-1$) are normalized by the first coefficient (DC component) $\mathrm{FD}_l/\mathrm{FD}_0$ ($l = 1, .., N-1$), later referred to simply as the Fourier descriptors.

Distance Measures. For the vector search (subsection 4.1), the distance measure was a Gaussian Radial-Basis function (and so actually a similarity measure). For the Fourier descriptors, the distance is taken by the city-block distance. When the shape vectors and Fourier descriptors were combined (as in shape description no. 4, see subsection 4.2), the two individual distances were summed whereby the one for the Fourier descriptors was multiplied by a factor of 8.

4. Results

Three types of evaluations were performed. One is the visual inspection of sortings of the shape vector (subsection 4.1). It serves the purpose to check whether the low-dimensional space is approximately homogenous. The second type of evaluation is the task of shape retrieval (subsection 4.2). The shape sorting is carried out on half the number of images (30000) of the Corel collection, because many of those images contain many small shapes. The MPEG7 collection is less suitable for this, as it contains only few and mostly large shapes. The MPEG7 collection is however suitable for comparing to the performance of other shape descriptions, that were tested on that collection with a shape retrieval task (e.g. [5, 6, 7, 1]). The third type of evaluation was a typical classification task (subsection 4.3).

4.1. Low-dimensional vector search

Figure 2 demonstrates the sorting of shape vectors $\mathbf{s}$, carried out for individual dimensions (row-wise). For each row, a 'base' geometry was manually selected and one dimension was increased in five increments. For each increment the most similar (least distant) curve was selected and plotted. The corresponding dimension value of the found vector is given in the upper right of each subplot. This search does not guarantee the optimal, expected selection for two reasons: one is the limited amount of curve geometries in an image collection; the other is the multi-dimensionality of the space. To counteract the latter, the weights for the base geometry were set to a low value, whereas the weights for the incrementing dimension was set high (vice versa for the radial-basis function). Still, for some selections an actual increment cannot be found, see for instance in the second last row, in which two times the same segment pair appears.

In the top row the elongation dimension was increased, whereby the weight for structural biases is set low, thus selecting any shape. The actual elongation value of the most similar shape is noted in the upper right and corresponds to the unnormalized value. In the second row, an attempt was made to specifically select shapes with a curved bias value, that is selections that are most similar to a bicorn, D-shape and so on; the effect of the dimension change is however only limitedly evident due to the rare occurrence of such shapes. In the 3rd row, three-corner shapes with increasing elongation value are selected, which favors triangles initially and 'distorted' star shapes for higher values (in this particular collection). In the 4th row, squares with increasing elongations value were selected, which resulted in a similar change in shape qualities as for the three-corner shape selection. The 5th row shows an analogous selection for five-corner shapes. The last row is an example of a 'morphing' from circles to pentagons.

4.2. Retrieval

The retrieval performance was tested on set B of the MPEG7 image collection (1400 images total: 70 classes per 20 shapes). We tested four shape representations, which we implemented ourselves and generate precision-recall curves for them. Those in turn are compared to three precision-recall curves which we obtained from the literature [1, 2]. Figure 3 summarizes the seven curves. The ones from the literature are thin stippled.

The precision-recall curve is calculated as explained in [1, 2]. For a desired number of first k relevant retrievals (relevant = same class), the precision is the ratio of k and the rank order of the last of the k relevant shapes. For instance, for a recall of ten relevant shapes (50 percent recall), and a rank order of 15 for the tenth relevant shape, the precision is two third.

The first tested description consisted of only 8 dimensions, namely the above shape

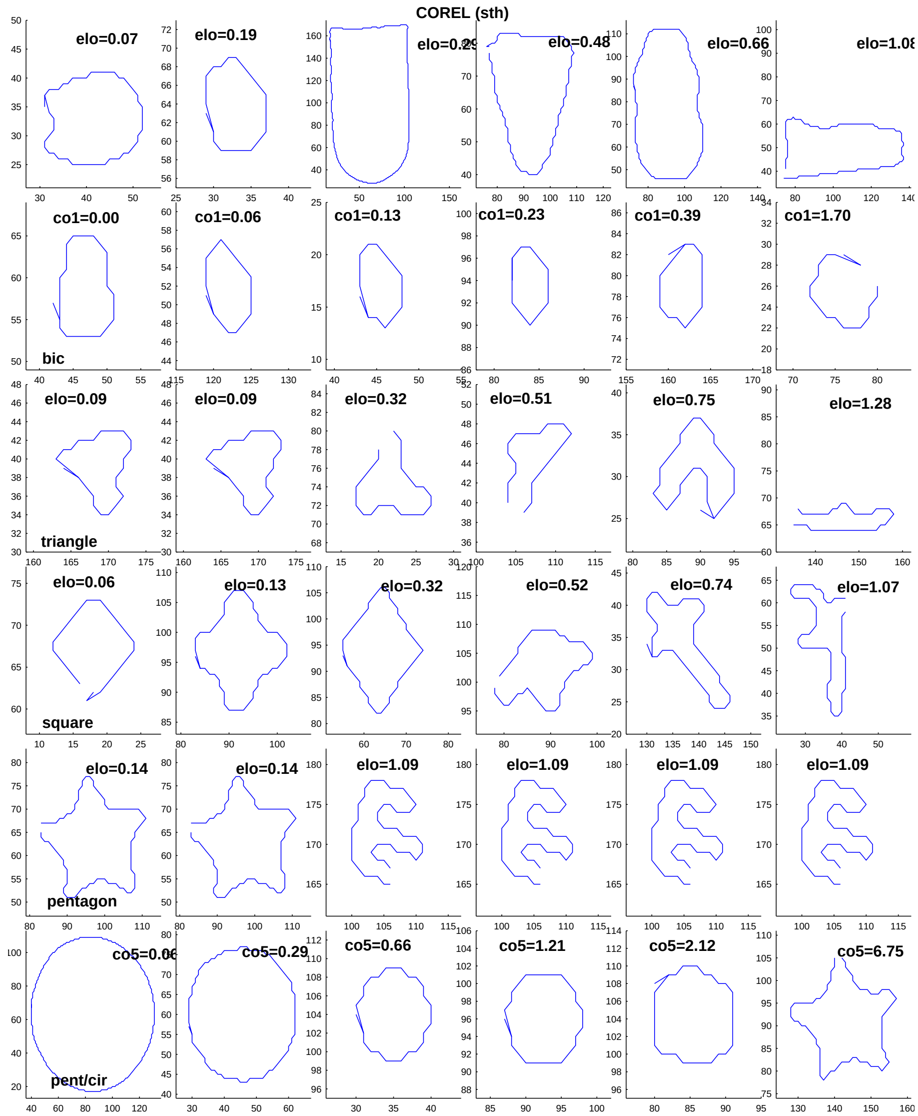

Figure 2: Sorting shape vectors along different dimensions (aspects). In each row one dimension value was systematically increased (in the upper right of each subplot the actual and not the preset value is displayed). Top row: increasing (inc.) elongation (η) value. 2nd row: inc. η value for a two-corner polygon 2; 3rd row: inc. η value for a three-corner polygon. 4th row: inc. η value for a five-corner polygon. 5th row: inc. η value for a circular shape.

vector without the dimensions for the gaps, for the orientation $(r, \eta, y, \breve{s}_\frown, \breve{s}_2, \breve{s}_3, \breve{s}_4, \breve{s}_5)$. The orientation dimension is omitted to achieve rotation independence and the gap dimensions are zero as the shapes are all closed in the MPEG7 collection. The precision-recall curve for this representation is labeled 'RS geo simple'. Its course is already at a level of that of a well performing signature, e.g. the chord-length signature as tested in Zhang and Lu, labeled 'CL-FD (Z&L)' (data obtained from figure 11c, p.44 in [1]).

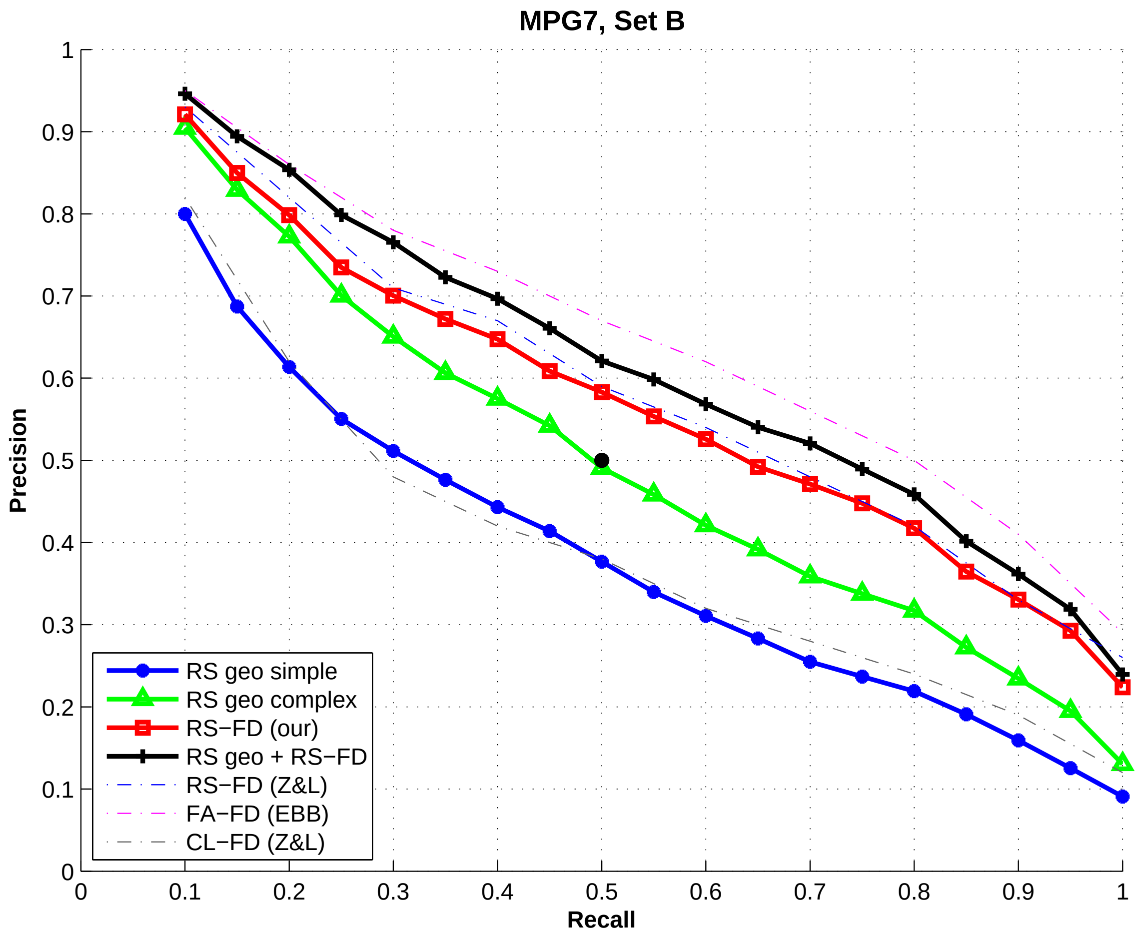

Figure 3: Shape retrieval with the MPEG7 image collection (set B). RS geo simple: 8 geometric parameters from $R(v)$. RS geo complex: 18 geometric parameters from $R(v)$. RS-FD (our): Fourier descriptors (20 coefficients) from $R(v)$(our implementation). RS geo + RS-FD: combination of the latter two. RS-FD (Z&L): Zhang and Lu's Fourier descriptors obtained from radial signature. FS-FD (EBB) El-ghazal et al.'s farthest distance. CL-FD (Z&L): chord-length signature (from Zhang and Lu, 2005).

The second description was the 18-dimensional vector as explained previously (labeled 'RS geo complex'). The performance increases substantially and the precision-recall curve now runs almost diagonally crossing value x=0.5 and y=0.5.

The third description aimed at replicating the results by Zhang and Lu [1], who used a combination of the radial signature with the Fourier descriptors: their data are plotted as

'RS-FD (Z&L)'. We did however not yield exactly the same level of performance (20 Fourier descriptors; labeled 'RS-FD (our)'): our curve is lower by a value of ca. 0.2 for some recall rates, likely due to small differences in exact implementation.

The fourth description was a combination of our second and third shape descriptions (38 dimensions; labeled 'RS geo + RS-FD'). The performance further increases, but not proportional to the number of parameters used.

For further comparison, the performance for the best-performing signature is also shown, which is the farthest-distance signature (combined with Fourier descriptors) by El-ghazal et al. (curve taken from figure 8, p. 580 in [2]).

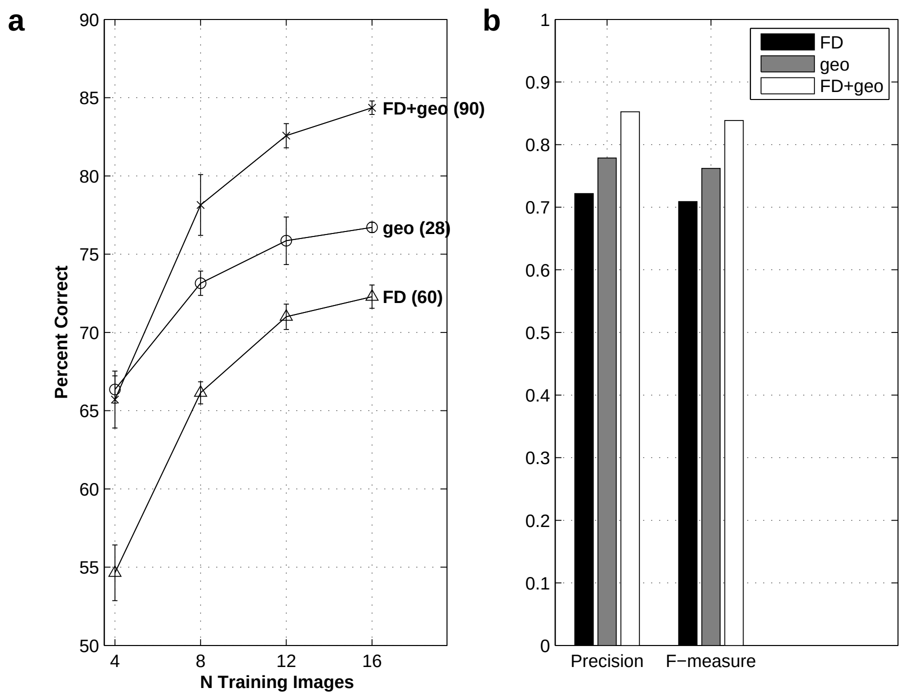

Figure 4: Classification performance with a Linear Discriminant Analysis. a. Learning curve for geometric parameters ('geo') and Fourier descriptors ('FD') and for a combination of both ('FD+geo'). b. Precision and F-measure for the three shape descriptions.

4.3. Classification

A Linear Discriminant classifier was trained for different amounts of images and correct classification was determined (figure 4a). The learning curve for the Fourier descriptors reaches ca. 73 percent for 16 training images but is clearly lower than the one for the geometric descriptors (ca. 77 percent), even though twice as many dimensions were used (60 and 28 respectively). When the two descriptors were combined, the performance reached 84 percent. Precision and F-measure is also determined, with F1 = 2 · Precision · Recall/(Precision + Recall). For each measure, the average value of all one-against-others discriminations is taken for 16 training images (figure 4b): the values reflect the level of the classification measure.

5. Discussion

It was shown that powerful shape representations can be developed exploiting the extrema information of the radial signature alone, without the use of the Fourier transform. Five geometric biases were developed that corresponded to the number of corners of a polygon, or to the number of peaks of a star shape. And while those parameters lead to only moderate shape retrieval results on the MPEG7 collection in comparison to the best-performing Fourier-based methods (figure 3), the representation may well suffice for improving image classification, because scenes often consist of a set of smaller shapes that are loosely cyclic. A key-advantage of the representation is that it can be generated very quickly, because it essentially corresponds to computing the signature and its derivatives only.

When the geometric parameters were combined with the Fourier description, the performance almost reached the level of the best-performing boundary-based description - for the first 20 percent of recall, the level is even the same. This best-performing description is based on a combination of the farthest-point signature with the Fourier description by El-ghazal et al. [2]. The computation of that signature however, is relatively expensive as for each boundary point, the farthest distance is computed. The clear advantage of our method is its short processing duration.

In a classification task in contrast, the geometric parameters showed clearly better performance than the Fourier descriptors (figure 4) even for half the number of dimensions. The strength of the Fourier transform is its capability to deal with arbitrary curves, but shapes often contain a certain degree of regularity and symmetry [8], which is better captured by an explicit geometric as presented here. The classification performance is below the one of many other systems that achieve more than 95 percent (see table 3 for summary in [9]), but the key advantage of our system is the short computation duration.

The long-term goal is to exploit the low-dimensional descriptor $\mathbf{s}$ for (general) image classification, see [4] for a beginning. Scenes often contain texture and shapes whose rapid description can be essential to a fast categorization.

References

[1] D. Zhang, G. Lu, Study and evaluation of different fourier methods for image retrieval, Image and Vision Computing 23 (2005) 33–49.

[2] A. El-ghazal, O. Basir, S. Belkasim, Farthest point distance: a new shape signature for fourier descriptors, Signal Processing: Image Communication 24 (2009) 572–586.

[3] M. Peura, J. Iivarinen, Efficiency of simple shape descriptors, in: Third International Workshop on Visual Form, Capri, IT, 1997, pp. 443–451.

[4] C. Rasche, An approach to the parameterization of structure for fast categorization, International Journal of Computer Vision 87 (2010) 337–356.

[5] L. Latecki, R. Lakamper, T. Eckhardt, Shape descriptors for non-rigid shapes with a single closed contour, in: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, Vol. 1, 2000, pp. 424 – 429.

[6] B. Super, Learning chance probability functions for shape retrieval or classification, in: Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04., 2004.

[7] B. Super, Retrieval from shape databases using chance probability functions and fixed correspondence, International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI) 8 (2006) 1117–1137.

[8] J. Gielis, A generic geometric transformation that unifies a wide range of natural and abstract shapes, American Journal of Botany 90 (2003) 333–338.

[9] M. R. Daliri, V. Torre, Classification of silhouettes using contour fragments, Computer Vision and Image Understanding 113 (2009) 1017–1025.