INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

A Graph-based Approach for Multiscale Shape Analysis

R.S. Torres *A.X. Falcão* *L. da F. Costa*

Technical Report - IC-03-03 - Relatório Técnico

March - 2003 - Março

Abstract

This paper presents the advantages of computing two recently proposed shape descriptors, multiscale fractal dimension and contour saliences, using the image foresting transform—a graph-based approach to the design of image processing operators. It introduces a robust approach to estimate contour saliences (peaks of high curvature) by exploiting the relation between contour and skeleton. The paper also compares both shape descriptors to fractal dimension, Fourier descriptors, and moment invariants with respect to their invariance to object characteristics that belong to a same class (compact-ability) and to their discriminatory ability to separate objects that belong to distinct classes (separability).

Keywords: Shape analysis, image processing, fracal dimension , shape saliences, Image Foresting Transform , multiscale skeletonization

1 Introduction

In pattern recognition and related areas, shape is an important characteristic to identify and distinguish objects [1]. The shape variations expressed with respect to a given scale, named *multiscale shape representation*, provides even more information about the objects. In this context, shape descriptors are used to encode such representations in signatures (i.e. feature vectors). In practice, objects belong to certain semantic categories, each category defines a class, and the problem consists of grouping the objects that belong to a same class. The main challenge here is to find out "good signatures" to perform such a task successfully.

This paper presents a graph-based approach to create shape descriptors from multiscale shape representations and proposes two alternatives to assess their "goodness" regarding to two important aspects: *compact-ability* and *separability*. The *compact-ability* of a descriptor indicates its invariance to the object variations within the same class, while the *separability* indicates its discriminatory ability between objects that belong to distinct classes.

The proposed graph-based approach uses the framework of the *image foresting transform* (IFT) - a recently introduced technique to the design of image processing operators for image filtering, segmentation, and analysis [2, 3, 4]. The IFT reduces image processing problems to a minimum-cost path forest problem in a graph, whose solution is usually obtained in time proportional to the number of pixels. This paper discusses two recently proposed shape descriptors in the framework of the IFT, *shape saliences* and *multiscale fractal dimension* [5, 6], and compares them with two commonly used shape descriptors, Fourier coefficients [7] and invariant moments [8]. The saliences of a shape are defined as the influence areas of its highest curvature points. This descriptor combines local and global information about a given shape. It allows quantification of the steepness of the curvature at points where the analytical curvature would be infinite [9], and at the same time, takes into account intricate and complex parts of the shape. The present work introduces a more robust method, as compared to the original one [5], to compute salience points along the contour by exploiting its relation to the internal and external skeletons. It also presents a new shape descriptor based on shape saliences and a matching algorithm for this descriptor. The multiscale fractal dimension is a new concept, which copes with many serious drawbacks in current methods for numerical estimation of fractal dimension [5, 10]. The multiscale fractal dimension quantifies the overall complexity or territorial coverage of the shape.

This article starts by presenting an overview of the IFT in Section 2. The IFT is used to obtain two types of shape representation: multiscale contours by exact dilations and multiscale skeletons by label propagation, as described in Section 3. We use the former to estimate multiscale fractal dimension and the later to compute the salience points along the contour in Section 4. The salience values at these points are obtained from the label propagation scheme. Section 5 presents a formal definition of compact-ability and separability, evaluates the proposed shape descriptors, Fourier coefficients, and moment invariants, and discusses the main results of this work. We state the conclusion and discuss our current research on shape descriptors in Section 6.

2 Image Foresting Transform

Many problems in image processing can be interpreted as an image partition problem based on a given set of *root pixels*, where each root defines an *influence zone* consisting of the pixels that are "more closely connected" to that root than to any other, in some appropriate sense [2]. The *image foresting transform* (IFT) reduces such problems to the computation of a *minimum-cost path forest* in a directed graph, whose nodes are the pixels and whose arcs are defined by an *adjacency relation* between pixels.

The cost of a path in this graph is determined by a suitable *path cost function*, which usually depends on local image properties along the path — such as color, gradient, and pixel position. For suitable path-cost functions, one can choose the optimal paths so that their union is an oriented forest, spanning the whole image. The nodes of each rooted tree in the forest are by definition the influence zone of the corresponding root.

The result of the IFT is an *annotated image*, that assigns to each pixel three additional attributes: its predecessor in the optimum path, the cost of that path, and the corresponding root.

2.1 Euclidean IFT

The Euclidean IFT is an IFT where the cost map is the Euclidean distance transform of a set of root pixels [11]. That is, it assigns to each pixel in the cost map its minimum Euclidean distance to the root set. This operator requires a path cost function f_{euc}:

$$f_{euc} = (\sum_{i=1}^{n-1} |x_{t_i} - x_{t_{i+1}}|)^2 + (\sum_{i=1}^{n-1} |y_{t_i} - y_{t_{i+1}}|)^2, \tag{1}$$

where $\pi =< t_1, t_2, ..., t_n >$ is a path in the graph and (x_{t_i}, y_{t_i}) are the (x, y) coordinates of a pixel t_i in the image. The cost map outputs the squared Euclidean distance values, representing the exact dilations of the root set. The root map represents the influence areas of the root pixels, and the optimum-paths connecting each pixel to its corresponding root are encoded in the predecessor map.

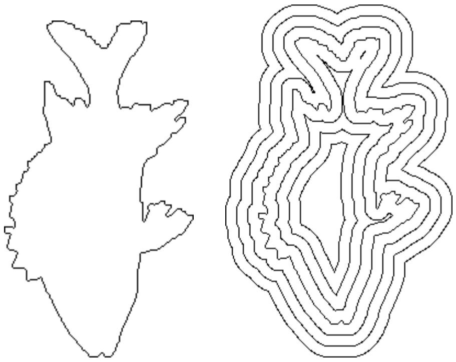
3 Multiscale Shape Representation

A shape can be represented along a range of scales spanning from coarse to fine. If the shape is to be used as an invariant indicator of an object in a scene in which the viewing distance is variable, a multiscale structure is necessary to relate various views, thereby making the representation invariant with respect to the viewing distance [12]. The Euclidean IFT allows efficient computation of multiscale contours by exact dilations and multiscale skeletons by label propagation [11].

Each instance of contour is an oriented, connected and closed curve in the plane which separates the interior and the exterior regions of an object. For a given contour, its internal skeleton is defined as the geometric location of the centers of maximal disks contained in the contour [13]. A similar definition is valid for the external skeleton.

3.1 Multiscale Contours by Exact Dilations

Given a set of points S, represented in terms of their Cartesian coordinates (x, y), its exact Euclidean dilation by a radius r, henceforth represented as S_r, is defined as being the union of all disks of radius r centered at each of the points of S. Observe that this definition is valid for both discrete and continuous objects. Subsequent dilations of a given shape by increasing values of r, create a family of progressively simplified instances of the original shape, as illustrated in figure 1 for a contour.



(a) (b)

Figure 1: (a) A contour of a fish and (b) multiscale contours by exact dilation.

Multiscale contours by exact dilation result from the Euclidean IFT, where the points of the original shape (contour) are taken as root set. Each instance of the multiscale shape is obtained by thresholding the cost map at a given squared Euclidean distance value. The higher the threshold value, the more simplified the shapes become, with smaller details being progressively removed as the threshold increases.

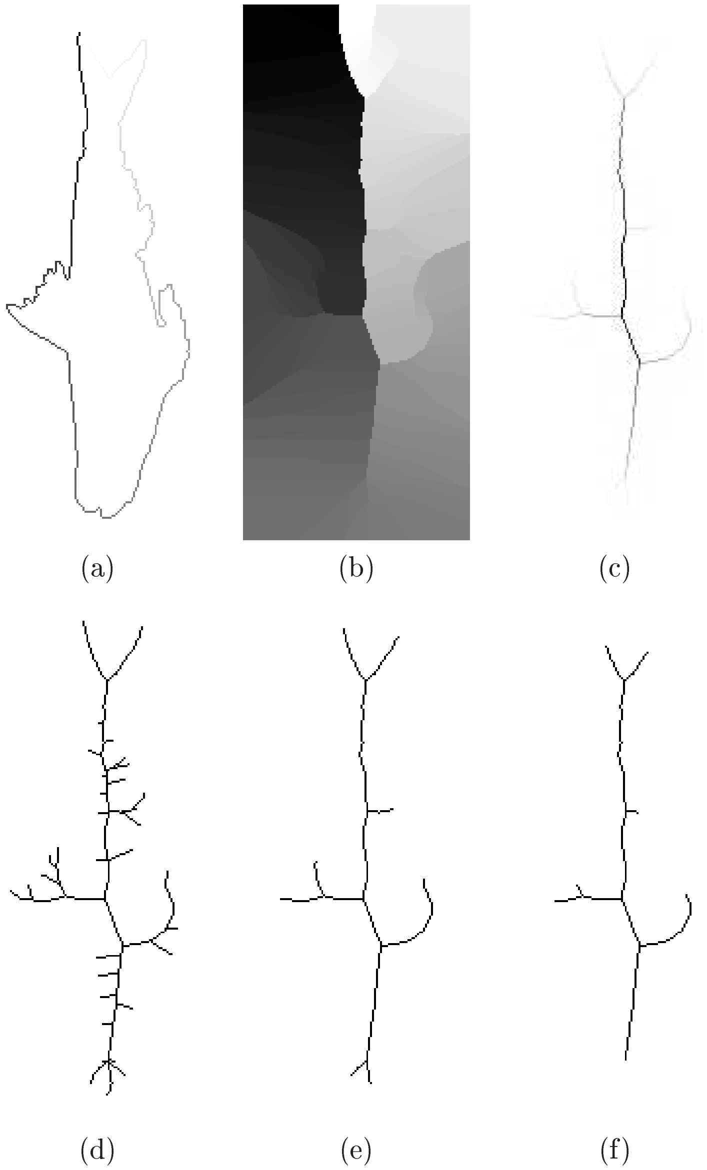


Figure 2: Multiscale skeletonization by label propagation. (a) Labeled contour, (b) label map, (c) difference image, and (d-f) skeletons at three different scales.

3.2 Multiscale Skeletons by Label Propagation

Given a contour with N pixels, the present method aims at assigning to each pixel inside and outside the contour the maximum length of the shortest contour segment between two roots equidistant to the pixel according to the cost map [14, 15, 11]. When a pixel has only one closest root in the contour, the assigned value is zero. This process works as follows.

Each contour pixel is considered a labeled root. They are labeled sequentially by increasing values while circumscribing the contour (Figure 2a). Then the Euclidean IFT generates a label map L in the place of the root map (Figure 2b). A difference image D is created from the label map by computing the following for each pixel p inside and outside the contour (Figure 2c):

$$D(p) = \max_{\forall q \in N_4(p)} \{\min\{\delta(p,q), N - \delta(p,q)\}\}, \tag{2}$$

where $\delta(p,q) = L(q) - L(p)$ and $N_4(p)$ is the set of pixels q that are 4-neighbors of p. The difference image represents the multiscale internal and external skeletons by label propagation. One-pixel wide and connected skeletons can be obtained by thresholding the difference image at subsequent integer values (Figure 2d-f). The higher the threshold value, the more simplified the skeletons become, with smaller details being progressively removed as the threshold increases.

It is important to observe that Equation 2 corrects the original Equation, reported in [15, 11].

4 Shape Descriptors

This section presents the process of creating shape descriptors from the multiscale shape representations presented in Section 3.

4.1 Multiscale Fractal Dimension

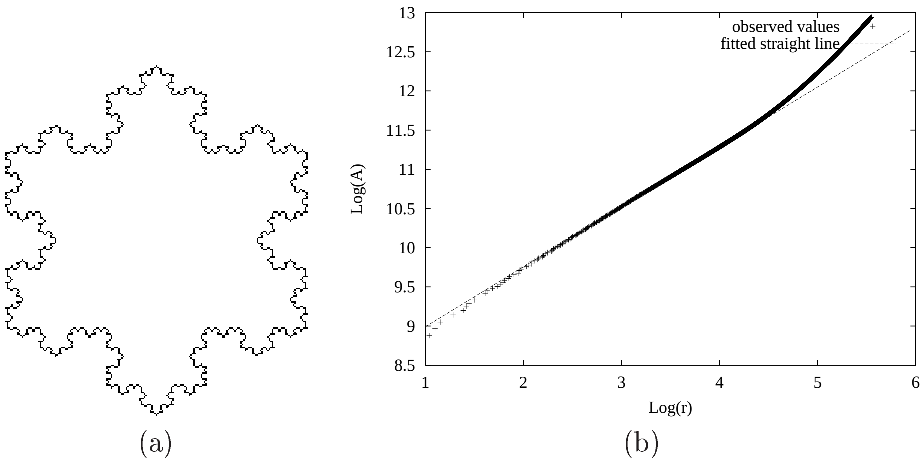
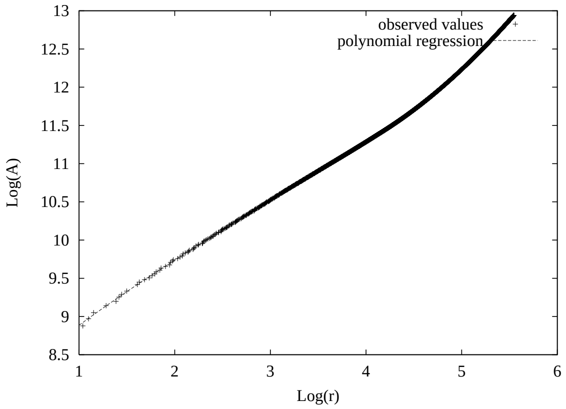
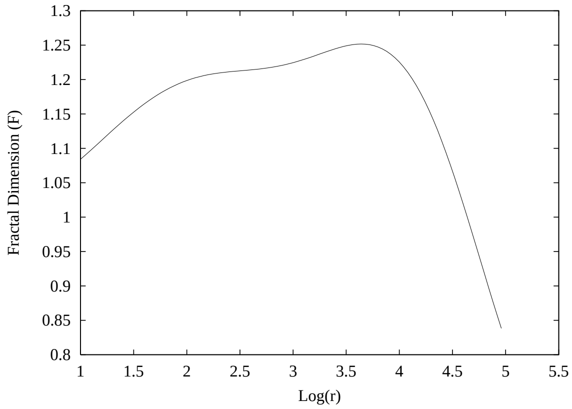


Figure 3: (a) An object similar to the Koch star, whose fractal dimension is known as about 1.26 ($= \frac{log4}{log3}$). (b) The logarithmic area function. By taking 2 minus the inclination of the fitted straight line, the fractal dimension obtained is about 1.23.

While the topological dimension is restricted to integer values, fractal dimension allows fractionary values. Disseminated by Mandelbrot [16], fractal dimension provides an interesting means for characterizing the self-similarity (or self-affinity) of abstract and real



(a)



(b)

Figure 4: (a) The $log \times log$ curve of the areas of each exact dilation radius for Figure 3a. (b) The multiscale fractal dimension of its shape.

objects, being closely related to the concept of power-laws. A particularly intuitive and useful definition of fractal dimension is the Minkowski-Bouligand dimension [17], which is here introduced in terms of the following example. Let the shape under analysis be represented in terms of the set S of the Cartesian coordinates of each of its elements, and let S_r be its dilation by r (see Section 3.1). Let $A(r)$ be the area of the respective dilated version of the shape, i.e. S_r. The Minkowski-Bouligand fractal dimension, hence F, is defined as

$$F = 2 - \lim_{r \to 0} \frac{log(A(r))}{log(r)}. \tag{3}$$

In other words, the fractal dimension descriptor in this case (i.e. considering a two-dimensional space) is a number within $[0, 2]$. It should be borne in mind that F assumes perfect self-similarity of the shape for small spatial scales, i.e. for r close to 0, which is never verified for real data. Indeed, while shapes in nature can exhibit an infinite degree of detail as one moves into the microscopic scales, the self-similarity along these scales is not preserved for an infinite interval. For instance, a fern leaf presents just a few orders (3 or 4) of self-similarity. The situation is even more complicated for experimental data, where the finite resolution of the acquisition device contributes further to limit the small scale detail.

In spite of such limited fractality observed for real objects, the standard numerical procedure for estimating fractal dimensions involves linear interpolating the logarithm curve of the area $(A(r))$ in terms of dilating radius, computing the angular coefficient $(A'(r))$ of this line and taking F as $F(r) = 2 - A'(r)$ (see Figures 3a and 3b). Observe that the area values $A(r)$ for each logarithm of the dilation radius r can be simply obtained by computing the accumulated histogram of the cost map of the Euclidean IFT. Therefore, it is obtained from the multiscale contours by exact dilations (Section 3.1).

Although the deviations of shapes from perfect self-similarity seriously undermine the standard above experimental method, several practical applications of the fractal dimension have been reported in the literature (e.g. [18]). Fractal dimensions have been considered as features useful for expressing the area coverage and the "complexity" of shapes ranging from neurons [10] to heartbeat dynamics [19]. In the particular case of

the Minkwoski-Bouligand dimension, the value of F provides an interesting indication of how much the shape constrains its own dilation. Therefore, simple shapes, such as the point or the straight line impose relatively less constraints to their own dilation and consequently have smaller fractal dimension values than those of an intricated curve in the plane.

In order to address the subjectivity implied by the choice of the interval over which the logarithmic curve is interpolated and to fully take into account the limited self-similarity exhibited by the geometry of real shapes, the concept of *multiscale fractal dimension* was recently reported (e.g. [5]). This approach involves taking into the infinitesimal limit the previous concept of linear interpolation (e.g. [20, 21]), which naturally leads to the estimation of the *derivative* of the logarithmic area function. Therefore, the multiscale fractal dimension becomes a function of the spatial scale rather than a single scalar global value. By expressing the fractality explicitly in terms of the spatial scale, this new measure provides a richer description of the self-similarity of the analyzed shapes along the spatial scales. The derivative function therefore becomes completely independent of the choice of the spatial scale interval adopted for interpolation.

The approach presented here fits a polynomial curve by regression to the logarithmic area function from which the sought derivatives can be immediately obtained. One important advantage of this approach is to be free of the undesirable oscillations often found in the derivative estimation of sampled curves. Note that, the commonly used fractal dimension can be understood as a particular case of the multiscale dimension when the adjusting polynomial is linear. The multiscale fractal dimension is obtained whenever the degree of the polynomial is greater than one. In the examples of this paper, the multiscale fractal dimension is represented by a polynomial of degree nine. In this work the multiscale fractal dimension descriptor is represented by a vector of 50 sample points of this polynomial. The polynomial degree and the vector size were determined through a set of experiments. These experiments showed that vectors containing more than 50 sample points do not improve the results.

Figure 4 illustrates the concept of multiscale fractal dimension with respect to the shape in figure 3a. Observe that the maximum value of the curve in Figure 4b is close to 1.26, which is the actual fractal dimension of the Koch triadic curve (up to three digits).

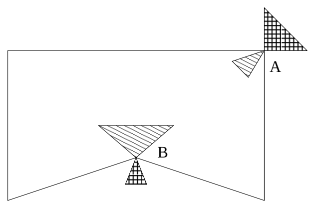


Figure 5: Internal and external influence areas of a convex (A) and a concave (B) point.

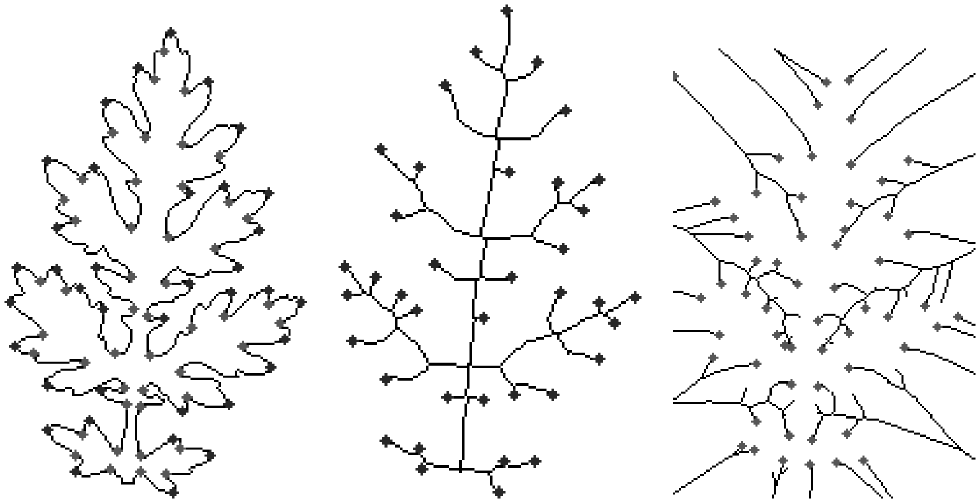


Figure 6: (a) Saliences of the contour of a leaf and (b-c) saliences of its internal and external skeletons.

4.2 Shape Saliences

The storage of the area evolution for each point of the shape also provides perspectives for shape descriptions. The influence areas of higher curvature points, namely *salience points* [5], are expected to be greater than the influence areas of the other points of the shape (see Figure 5). Moreover, in the case of a contour, the influence area of a convex point (point A) is greater outside the contour than inside, and the other way round for a concave point (point B). The influence area of each salience point relates to the aperture angle θ, illustrated in the figure 5, by the formula:

$$Area = \frac{\theta \times R^2}{2}, \tag{4}$$

where R is a dilation radius. Costa et al. [5] propose to estimate the salience points by thresholding the influence areas, computed for low values of R (e.g. 10). This approach, however, misses important salience points in opposite parts of the shape which come close to each other. It has otherwise been particularly effective for skeletons and for simple contours, such as convex polygons, but it fails in finding the salience points of more complex and intricate contours. A robust approach to solve the problem is described next.

For a given contour, multiscale skeletons are first obtained by label propagation. For small scales (low thresholds), each salience point of the internal skeleton corresponds to one convex point of the contour and each salience point of the external skeleton corresponds to one concave point of the contour (see Figure 6). The salience points of the skeletons are found by computing the Euclidean IFT of each skeleton separately, for a small dilation radius ($R = 10$) and using the skeleton points as root pixels. The histogram of the root map gives the influence areas of each skeleton point. The salience points of the skeletons are those with influence area greater than the area threshold obtained by setting $\theta = 70$ in Equation 4.

The Euclidean IFT of the contour allows a natural, simple and direct way to extract the relationship between the contour and its internal and external skeletons. Equation 2 assigns to each pixel inside and outside the contour the maximum length of the shortest contour segment between two roots equidistant to the pixel according to the cost map. Figure 7 illustrates this situation for a salience point c in the skeleton, which is related to a salience point a in the contour. The difference value $D(c)$ is the length of the segment $\overline{dab}$. Suppose b is the root pixel of c in the Euclidean IFT, the point a can be reached from the point c by skipping $\overline{dab}/2$ pixels in the anti-clockwise orientation along the contour starting from b. Similarly, the point a could be found from c through d following the clockwise orientation, when d is the root pixel of c. The method needs only to determine which is the root pixel, either b or d. If the contour pixels are labeled in clockwise orientation, the root pixel of c will be b whenever $(L(q) - L(p)) > N - (L(q) - L(p))$ in Equation 2 for $L(q) = L(d)$ and $L(p) = L(b)$. Otherwise, the root pixel of c will be d for $L(q) = L(b)$ and $L(p) = L(d)$. The same rule is applied for the external skeleton.

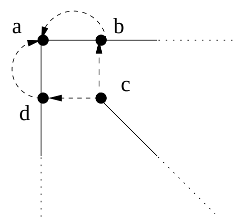
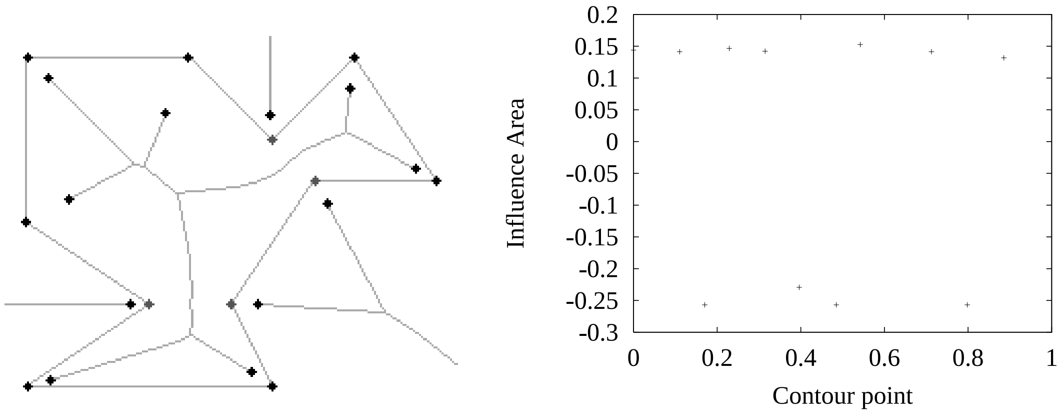


Figure 7: Relationship between skeleton and contour saliences.

Although the influence areas (saliences) of the contour points do not provide a completely robust method to determine the salience points, they encode important local and global information about the contour. The saliences are signed negative for concave points and positive for convex points. An arbitrary point of the contour is taken as reference point and the method computes the relative position of each salience point with respect to the reference point along the contour. Finally, the proposed shape salience descriptor consists of two vectors of the same size: one with the saliences and the other with the relative position of the salience points along the contour. Note that the dimension of these vectors may be different for different contours as well as the reference points. A special algorithm has been designed for matching this descriptor between two contours taking into account these differences. This algorithm is described in Section 5.3.

Figure 8 illustrates the shape salience descriptor for a polygon. The contour of the polygon, its reference point, the internal and external skeletons, and the respective salience points are indicated in Figure 8a. The curve shown in Figure 8b indicates the saliences by the relative position of the points along the contour.



(a) (b)

Figure 8: (a) Contour and skeleton of a polygon, where salience points are indicated by dots. (b) The salience curve of the polygon contour.

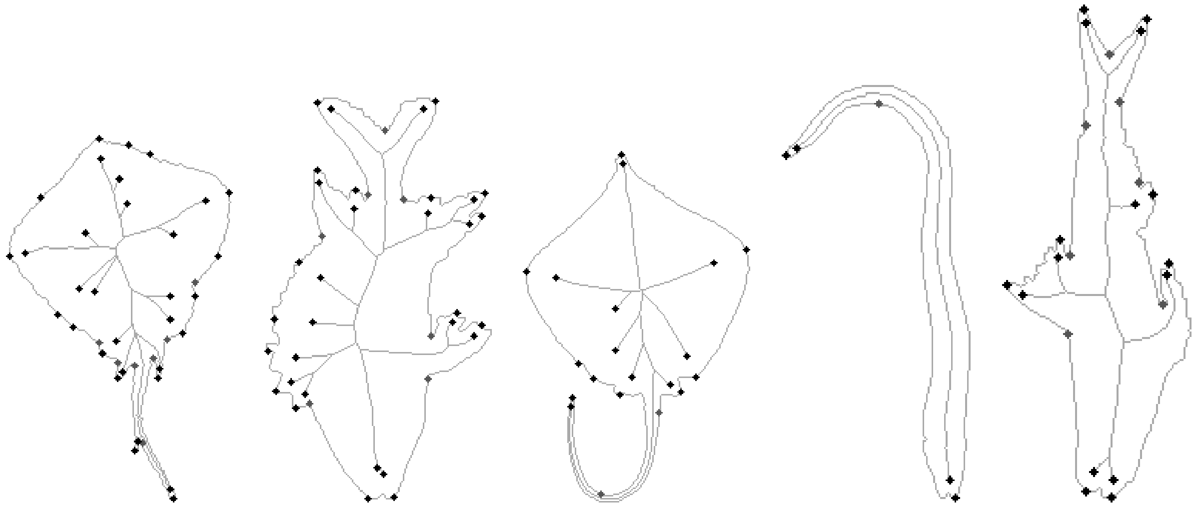
5 Evaluation

In pattern recognition as well as in content-based image retrieval, objects are organized into classes according to some semantic criterion. A "good" object descriptor should represent different classes of objects by separate and compact clusters of points in the feature space. The *compact-ability* of a descriptor indicates its invariance to the object characteristics that belong to the same class, while the *separability* indicates its discriminatory ability between objects that belong to distinct classes. The shape descriptors presented in this paper are evaluated with respect to compact-ability and separability in the context of a specific application. This application aims at designing and implementing an architecture for integrating image and spatial data for biodiversity information management. This architecture has been specified in a generic way, but its implementation is being carried throughout for the specific case of fish species.

One thousand and one hundred fish contours were obtained from the database available at *www.ee.surrey.ac.uk/Research/VSSP/ imagedb/demo.html* for the experiments. Figure 9 shows some examples of fish contours and their respective skeletons.

Since there is no semantic definition of classes for the fish contours in this database, we defined each class as consisting of 10 different manifestations of each contour by rotation and scaling. Then, the problem consists of 1100 classes with 10 shapes each. In this case, compact-ability becomes the invariance to possible rotation and scaling of a given shape, and separability becomes the discriminatory ability of a descriptor among the 1100 classes of the database.

The compact-ability and separability measures require a distance function between objects in the feature space of a given descriptor. The metric $L2$ is adopted for all shape descriptors, except for the shape salience descriptor which requires a special matching algorithm. A precise definition of compact-ability and separability, the matching algorithm for shape saliences, and the experiments are presented in the next sections.



(a) (b) (c) (d) (e)

Figure 9: Fish images used for descriptor evaluation. The concave points were determined through the salience points of the external skeleton, not shown in the figure.

5.1 Compact-ability

Let S be a set (database) of N shapes organized in classes. The compact-ability $\phi_D(C)$ of a descriptor D for a given class C in S is defined as:

$$\phi_D(C) = \frac{\sum_{\forall i,j \in C} \Delta_D(i,j)}{|C|^2}, \tag{5}$$

where $|C|$ is the number of shapes in the class C, $\Delta_D(i,j) = Distance(D_i, D_j)$ and D_i is the value of D for shape i. This measure is normalized ($\overline{\phi}_D(C)$) with respect to the maximum distance between a pair of shapes, considering all shapes of the database (see Equation 6).

$$\overline{\phi}_D(C) = \frac{\phi_D(C)}{\max_{\forall i,j \in S}\{\Delta_D(i,j)\}} \tag{6}$$

The compact-ability of a given descriptor is the average of the normalized compact-abilities of this descriptor over all classes in S, i.e.:

$$\Phi_D = \frac{\sum_{\forall C \in S} \overline{\phi}_D(C)}{|S|}, \tag{7}$$

where $|S|$ is the number of classes in the set S.

5.2 Separability

Let S be a set (database) of N shapes organized in classes. The separability $\psi_D(C)$ of a descriptor D for a given class C is defined as follows. An arbitrary shape r_C is taken as reference for the class C and the distances $\overline{\Delta_D(r_C,i)} = \frac{\Delta_D(r_C,i)}{M}$, where $M = \max_{\forall i \in S, \forall r_C}\{\Delta_D(r_C,i)\}$, is computed for all shapes i in S.

The distance range is quantized in a certain number of values from x to 1.0 with intervals of dx (e.g. $x = 0.02$ and $dx = 0.02$). Let $\eta_{r_C}(x)$ be the number of shapes, whose distance from the reference shape is less than or equal to x ($\overline{\Delta(r_C,i)} \leq x$) and do not belong to the class C ($i \notin C$). For each distance value from x to 1.0, the separability $\psi_D(C)$ of a descriptor D with respect to class C is defined as:

$$\psi_D(C) = 1 - \frac{\eta_{r_C}(x)}{N}, \tag{8}$$

These separability values define a multiscale curve of separability for the class C along x. The separability of a descriptor D is defined as the average of the multiscale separability over all classes, i.e.:

$$\Psi_D = \frac{\sum_{\forall C \in S} \psi_D(C)}{|S|} \tag{9}$$

5.3 Matching Algorithm for Shape Salience

Whenever two shapes of a same object appear in different positions, they should be represented by the same salience points along the contour. Therefore the pairwise comparison between objects using shape saliences requires matching between contours.

The shape salience descriptor considered in the current work preserves the salience values of the points along the contour and their relative position regarding to a reference point. These characteristics encode a lot of information about the shape. The reference point is used only for correction of the relative positions after the matching. The matching algorithm proposed in this paper was based on the matching algorithm proposed to match Curvature Scale Space (CSS) images presented in [22, 23].

Let $S_A = \{(u_{A1}, s_{A1}), \ldots, (u_{An}, s_{An})\}$ and $S_B = \{(u_{B1}, s_{B1}), \ldots, (u_{Bm}, s_{Bm})\}$ be two salience descriptors of shapes A and B, where (u_{Ai}, s_{Ai}) stands for the i^{th} salience value s_{Ai} at the position $u_{Ai} \in [0,1]$ along the contour of the shape A.

1. Create $S'_A = \{(u'_{A1}, s'_{A1}), \ldots, (u'_{An}, s'_{An})\}$ and $S'_B = \{(u'_{B1}, s'_{B1}), \ldots, (u'_{Bm}, s'_{Bm})\}$ by sorting S_A and S_B according to the decreasing order of salience values.

2. Create a list L containing a pair of matching candidates points from S'_A and S'_B. A pair $((u'_{Ai}, s'_{Ai}), (u'_{Bj}, s'_{Bj}))$ belongs to the list L if $|s'_{Ai} - s'_{Bj}| \leq 0.2 s'_{A1}$. A pair $((u'_{Bj}, s'_{Bj}), (u'_{Ai}, s'_{Ai}))$ belongs to the list L if $|s'_{Bj} - s'_{Ai}| \leq 0.2 s'_{B1}$.

3. For each pair of matching candidates in the form $P_{ij} = ((u'_{Ai}, s'_{Ai}), (u'_{Bj}, s'_{Bj}))$ in L, find the shift parameter α as $\alpha = u'_{Ai} - u'_{Bj}$. Shift S_A salience points by α, yielding $S''_A = \{(u''_{A1}, s''_{A1}), (u''_{A2}, s''_{A2}), \ldots, (u''_{An}, s''_{An})\}$.

4. The distance D between S''_A and S_B is given as:

$$D = \sum_{k=1}^{\min\{n,m\}} D_k,$$

Descriptor Id	Descriptor Name
D1	Fractal Dimension
D2	Multiscale Fractal Dimension
D3	Contour Saliences
D4	Fourier Descriptor
D5	Moment Invariants

Table 1: List of evaluated descriptors.

where

$$D_k = \begin{cases} \sqrt{(u''_{Ak} - u_{Bk})^2 + (s''_{Ak} - s_{Bk})^2}, & \textit{if } |u''_{Ak} - u_{Bk}| \leq 0.2 \\ s''_{Ak} + s_{Bk}, & \textit{otherwise.} \end{cases}$$

Finally, if $n \neq m$, it is added to D the height s of the not matched points.

5. Repeat the steps 3 and 4 by considering matching candidate pair in the form $P_{ij} = ((u'_{Bj}, s'_{Bj}), (u'_{Ai}, s'_{Ai}))$ in L.
6. Select the lowest distance D as the distance between S_A and S_B

5.4 Experiments

Clearly, the multiscale fractal dimension is not scaling invariant. In order to allow scaling invariance the shapes are first normalized according to diameter.

Table 1 shows the set of implemented shape descriptors. The proposed descriptors (*D2* and *D3*) are compared with the common implementation of fractal dimension (*D1*), Fourier descriptors (*D4*) and moment invariants (*D5*).

Even though the external skeleton can be used to determine the correct location of a concave salience along a contour (see Figure 9), preliminary results show that the shape salience presents the best behavior if we consider only the convex saliences. We found out that the choice of the best threshold to determine the external skeleton salience is particularly sensitive with respect to the shape segmentation process. In fact, different threshold values were used to estimate the concave saliences in Figure 9.

Moment invariants and Fourier Descriptors have been widely used as shape descriptors [24, 7]. Many versions of these methods have been proposed, but, in this work, we consider conventional implementations. We have implemented the method described in [25, 26] to represent a shape with Fourier Descriptors applied to a contour. Each original object and its transformed versions have been represented by the most significant 126 components. The Euclidean distance has been used to measure the similarity between two Fourier-based representations.

For moment invariants, each object has been represented by a 14 dimensional feature vector, including two sets of normalized moment invariants [8, 27], one from object boundary and the other from solid silhouette. Again, the Euclidean distance has been used to measure the similarity between different shapes represented by their moment invariants.

Descriptor Id	Compact-ability
D1	0.07
D2	0.03
D3	0.30
D4	0.23
D5	0.03

Table 2: Compact-ability values of evaluated descriptors.

5.5 Experimental Results

Initially, we compared the multiscale fractal dimension of a contour and the classical fractal dimension applied to a contour, which is based on a linear interpolation of the logarithm curve of the accumulated area. Figure 10 shows that the multiscale version (*D2*) of the fractal dimension descriptor presents the best separability curve.

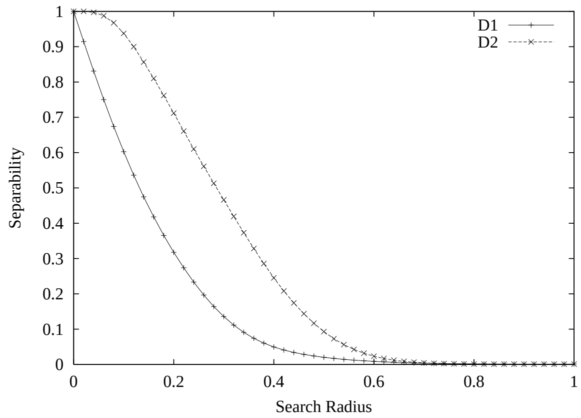


Figure 10: Multiscale separability diagrams for the fractal dimension based descriptors.

Figure 11 shows the separability curves of the proposed descriptors (*D2* and *D3*), the Fourier descriptor (*D4*) and the moment invariants (*D5*). Observe that the contour salience is clearly more "separable" than the other descriptors. The multiscale fractal dimension wins Fourier descriptor for a search radius less than 40% of the maximum distance. Observe that in shape-based image retrieval and pattern recognition context, we are generally interested in low values of search radius. Note also that the Fourier descriptors outperform the invariant moments as expected, if we consider previous results which try to compare these descriptors [23].

Table 2 presents the obtained compact-ability values of the evaluated shape descriptors. It is clear that low values are found for the descriptors based on the fractal dimension concept (*D1* and *D2*) and for moment invariants (*D5*). Note also that the contour salience (*D3*) presents the worst compact-ability value, despite its separability curve being the best one. Although descriptor *D3* has not provided compact clusters, these clusters are better separable in such a feature space as compared to the other descriptors.

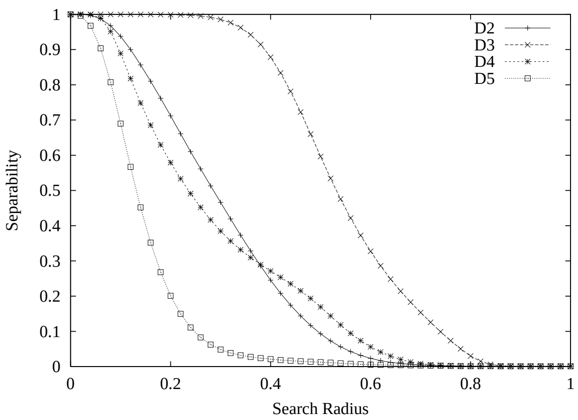


Figure 11: Comparison of the proposed descriptors with the Fourier descriptors and invariant moments.

6 Conclusion

This paper has presented two effective shape descriptors, multiscale fractal dimension and contour saliences, using the framework of the IFT. The IFT provides a more efficient computation of the multiscale shape representations summarized in Section 3 (details are given in [11]), as compared to the original algorithm published in [14]. The presented method to compute multiscale fractal dimension is likely to be more robust than the one published in [5] as the undesirable oscillations commonly found in Fourier-based approaches are eliminated here by the use of polynomial regression. The proposed method to locate salience points along the contour exploits the relation between contour and skeletons, which is naturally obtained via IFT, and is more robust than the one proposed in [5]. The contour salience descriptor, which encodes saliences and relative position of the points along the contour, and the use of a matching algorithm for it are new contributions.

The paper shows for the first time extensive evaluation experiments involving the multiscale fractal dimension and the contour saliences in comparison with two broadly used descriptors, Fourier coefficients [7, 25] and moment invariants [8, 27]. The "goodness" of the proposed descriptors is evident in regard to the defined concepts of compact-ability and separability. Although the underlying ideas of compact-ability and separability are not totally new concepts, the paper has presented an original way to compute them, especially the separability for different scales of search. The experiments have also shown that the proposed descriptors separate clusters of objects in the feature space better than Fractal dimension, Fourier coefficients, and moment invariants, being the contour saliences the best descriptor for this respect.

A larger number of experiments has been carried out considering other shape descriptors based on the multiscale shape representations presented in this paper. Descriptors such as skeleton saliences, skeleton multiscale fractal dimension, multiscale version of common shape measures like area, perimeter and compactness were evaluated. In order to improve the readability of the paper, only the best achieved results have been reported here.

Ongoing developments consider the creation of shape descriptors, which combine the salience features with multiscale fractal dimension information. Another ongoing activity is to include a comparison of the proposed descriptors and the Curvature Scale Space (CSS) descriptor [22, 23]. Moreover, we are currently considering applications in content-based retrieval, using the proposed shape descriptors as effective indexing vectors.

Acknowledgments

The work of R.S. Torres is supported by FAPESP (Proc. 01/02788-7), A.X. Falcão thanks CNPq (Proc. 302966/02-1) and L.F. Costa is grateful to FAPESP (Procs. 96/05497-3 and 99/12765-2) and CNPq (Proc. 301422/92-3) for financial support.

The authors are grateful to Sadegh Abbasi, Farzin Mokhtarian, and Josef Kittler for the fish database.

References

[1] S. Loncaric, A survey of shape analysis techniques, Pattern Recognition 31 (1998) 983–1001.

[2] A. Falcão, J. Stolfi, R. Lotufo, The image foresting transform: Theory, algorithms, and applications, IEEE Trans. on Pattern Analysis and Machine Intelligence In review.

[3] R. Lotufo, A. Falcão, The ordered queue and the optimality of the watershed approaches, in: Mathematical Morphology and its Applications to Image and Signal Processing, Vol. 18, Kluwer Academic, Palo Alto, USA, 2000, pp. 341–350.

[4] A. Falcão, B. S. da Cunha, R. A. Lotufo, Design of connected operators using the image foresting transform, in: Proceedings of SPIE on Medical Imaging, Vol. 4322, San Diego, CA, 2001, pp. 468–479.

[5] L. da F. Costa, A. G. Campos, E. T. M. Manoel, An integrated approach to shape analysis: Results and perspectives, in: International Conference on Quality Control by Artificial Vision, Le Creusot, France, 2001, pp. 23–34.

[6] R. Torres, A. Falcão, L. Costa, Shape description by image foresting transform, in: 14th International Conference on Digital Signal Processing, Santorini, Greece, 2002, pp. 1089–1092.

[7] T. P. Wallace, P. Wintz, An efficient three-dimensional aircraft recognition algorithm using normalised fourier descriptors, Computer Graphics Image Processing 13 (1980) 99–126.

[8] M. K. Hu, Visual pattern recognition by moment invariants, IRE Transactions on Information Theory 8 (1962) 179–187.

[9] F. Attneave, Some informational aspects of visual perception, Psychological Review 61 (1954) 183–193.

[10] L. da F. Costa, E. T. M. Manoel, A shape analysis framework for neuromorphometry, Network 13 (2002) 283–310.

[11] A. Falcão, L. da F. Costa, B. da Cunha, Multiscale skeletons by image foresting transform and its applications to neuromorphometry, Pattern Recognition 35 (7) (2002) 1571–1582.

[12] V. Castelli, L. D. Bergman (Eds.), Image Databses. Search and Retrieval of Digital Imagery, John Wiley Sons, 2002.

[13] R. Kimmel, D. Shaked, N. Kiryati, A. M. Bruckstein, Skeletonization via distance maps and level sets, Computer Vision and Image Understanding: CVIU 62 (3) (1995) 382–391.

[14] L. da F. Costa, L. F. Estrozi, Multiresolution shape representation without border shifting, Electronic Letters 35 (21) (1999) 1829–1830.

[15] A. Falcão, B. S. da Cunha, Multiscale shape representation by image foresting transform, in: Proceedings of SPIE on Medical Imaging, Vol. 4322, San Diego, CA, 2001, pp. 1091–1100.

[16] B. Mandelbrot, The Fractal Geometry of Nature, W. H. Freeman and Co., San Francisco, 1982.

[17] C. Tricot, Curves and Fractal Dimension, Springer-Verlag, 1995.

[18] Y. Y. Tang, Y. Tao, E. C. M. Lam, New method for feature extraction based on fractal behavior, Pattern Recognition 35 (2002) 1071–1081.

[19] L. A. N. Amaral, P. C. Ivanov, N. Aoyagi, I. Hidaka, S. Tomono, A. L. Goldberger, H. E. Stanley, Y. Yamamoto, Behavioral-independent features of complex heartbeat dynamics, Physical Review Letters 86 (26) (2001) 6026–6029.

[20] H. Peitgen, H. Jurgens, D. Saupe, Fractals for the Classroom: Introduction to Fractals and Chaos, Springer Verlag, 1992.

[21] P. Montague, M. Friedlander, Morphogenesis and territorial coverage by isolated mammalian retinal ganglion cells, Journal of Neuroscience 11 (1991) 1440–1457.

[22] S. Abbasi, F. Mokhtarian, J. Kittler, Enhancing css-based rhape retrieval for objects with shallow concavities, Image and Vision Computing 18 (2000) 199–211.

[23] F. Mokhtarian, S. Abbasi, Shape similarity retrieval under a affine transforms, Pattern Recognition 35 (31-41).

[24] A. K. Jain, A. Vailaya, Shape-based retrieval: A case study with trademark image databases, Pattern Recognition 31 (1998) 1369–1390.

[25] R. C. Gonzalez, R. E. Woods, Digital Image Processing, Addison-Wesley, 1992.

[26] B. M. Mehtre, M. S. Kankanhalli, W. F. Lee, Shape measures for content based image retrieval: A comparison, Information Processing and Management 33 (3) (1997) 319–337.

[27] S. A. Dudani, K. J. Breeding, R. B. McGhee, Aircraft identification by moment invariants, IEEE Transactions on Computers c-26 (1) (1977) 39–45.