Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

PERGAMON

Pattern Recognition 37 (2004) 1163–1174

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

A graph-based approach for multiscale shape analysis

R. da S. Torres[a], A.X. Falcão[a,*], L. da F. Costa[b]

[a] *Institute of Computing, University of Campinas, Av. Albert Einstein, 1251, CEP 13084-851, Campinas, SP, Brazil*
[b] *Institute of Physics at São Carlos, University of São Paulo, C.P. 369, CEP 13560-970, São Carlos, SP, Brazil*

Received 8 January 2003; received in revised form 10 October 2003; accepted 10 October 2003

Abstract

This paper presents two shape descriptors, multiscale fractal dimension and contour saliences, using a graph-based approach— the image foresting transform. It introduces a robust approach to locate contour saliences from the relation between contour and skeleton. The contour salience descriptor consists of a vector, with salience location and value along the contour, and a matching algorithm. We compare both descriptors with fractal dimension, Fourier descriptors, moment invariants, Curvature Scale Space and Beam Angle Statistics regarding to their invariance to object characteristics that belong to a same class (compact-ability) and to their ability to separate objects of distinct classes (separability).

Keywords: Shape analysis; Image processing; Fractal dimension; Shape saliences; Image foresting transform; Multiscale skeletonization; Shape exact dilation

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*, provide even more information about the objects. In this context, shape descriptors have been used to encode such representations into 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 the advantages of computing two recently proposed shape descriptors, multiscale fractal dimension and contour saliences [2,3], using the *image foresting transform* (IFT)— a graph-based approach to the design of image processing operators [4–8]. In this case, the shape descriptors are obtained from the multiscale shape representations created by the IFT. The multiscale fractal dimension [2,9] is a new concept, which copes with many serious drawbacks in current methods [10,11] for numerical estimation of fractal dimension. The multiscale fractal dimension of a shape is computed based on the Euclidean distance transform (EDT) of its pixels. The EDT of these pixels is also related to their geometric Voronoi diagram [12], where each pixel defines an influence zone (*discrete Voronoi region*) composed by its closest image pixels. The saliences of a shape are computed based on the areas of the discrete Voronoi regions of its higher curvature pixels within a narrow band around the shape [2]. This approach allows the quantification of the curvature values at points (center of pixels) where the analytical curvature would be infinite. The IFT provides the simultaneous computation of the EDT and the discrete Voronoi regions in time usually proportional to the number of pixels [7], being more efficient than the method proposed in Ref. [2]. The present paper also introduces improvements in the multiscale fractal dimension and contour saliences computations. The original approach for multiscale fractal dimension suffers from undesirable oscillations on the fractal curve, and the location of higher curvature points along the contour for saliences computation

* Corresponding author. Tel.: +55-19-37885881; fax: +55-19-37885847.
E-mail addresses: rtorres@ic.unicamp.br (R. da S. Torres), afalcao@ic.unicamp.br (A.X. Falcão), luciano@if.sc.usp.br (L. da F. Costa).

0031-3203/$30.00 © 2003 Published by Elsevier Ltd on behalf of Pattern Recognition Society.
doi:10.1016/j.patcog.2003.10.007

is very sensitive in the case of intricate and complex shapes. The oscillation problem is solved using polynomial regression. The relation between the salience points of the contour and of its internal and external skeletons— an important concept introduced in Ref. [13]— is used to locate the higher curvature points along the contour, considerably improving the robustness of the contour saliences computation. This relation is obtained in a direct way using the IFT framework. The contour saliences descriptor is also redefined to include point location and salience value along the contour and a special distance metric, which make it possible to reach high effectiveness in shape recognition.

The proposed descriptors are compared with single fractal dimension, two classical (Fourier descriptors [14] and moment invariants [15]), and two recently published shape descriptors (curvature scale space (CSS) [16,17] and beam angle statistics (BAS) [18,19]) in regarding to the following aspects: *compact-ability* and *separability*. The compact-ability of a descriptor indicates its invariance to the object variations within a same class, while the separability indicates its discriminatory ability between objects that belong to distinct classes. In other words, a descriptor is considered "good" when it creates compact clusters far away from each other, for all classes in the corresponding feature space. This condition should be sufficient for the success of any suitable classification method.

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 locate the salience points along the contour in Section 4. Section 5 gives a formal definition of compact-ability and separability, evaluates the proposed shape descriptors, and discusses the main results of this work. We present the conclusion and our current research on shape descriptors in Section 6.

2. Image foresting transform

The *image foresting transform* (IFT) is a recent approach to the design of image processing operators based on connectivity [4–8]. The IFT reduces image partition problems based on a given seed set 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. A path in this graph is a sequence of adjacent pixels. The cost of a path 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, the IFT assigns to each image pixel a minimum-cost path from the seed set, such that the union of those optimum paths form an oriented forest spanning the whole image. The nodes of each rooted tree in the forest are composed by pixels that are "more closely connected" to its root pixel than to any other seed in some appropriate sense. The IFT assigns to each pixel three attributes: its predecessor in the optimum path (predecessor map P), the cost of that path (cost map C), and the corresponding root (root map R) or some label associated with it (label map L).

For given set S of seed pixels, the IFT can provide the simultaneous computation of the Euclidean distance transform in the cost map C and of the discrete Voronoi regions in the root map R [4]. This operator asks for an Euclidean adjacency relation A and a path-cost function f_{euc} defined for any path $\pi = \langle p_1, p_2, \ldots, p_n \rangle$ in the graph as

$$q \in A(p) \Rightarrow (x_q - x_p)^2 + (y_q - y_p)^2 < \rho^2, \tag{1}$$

$$f_{euc}(\pi) = \begin{cases} (x_{p_n} - x_{p_1})^2 + (y_{p_n} - y_{p_1})^2 & \text{if } p_1 \in S, \\ +\infty & \text{otherwise,} \end{cases} \tag{2}$$

where ρ is the adjacency radius and (x_{p_i}, y_{p_i}) are the (x, y) coordinates of a pixel p_i in the image. Note that, the main idea is to find for every image pixel p_n a path $P^*(p_n)$ from a seed pixel $p_1 \in S$, such that $f_{euc}(P^*(p_n))$ is minimum. The exact Euclidean distance transform will depend on the appropriate choice of ρ, as demonstrated in Ref. [4]. However, for most practical situations involving 8-connected curves, such as contours and skeletons, $\rho = \sqrt{2}$ is enough [7]. Algorithm 1 below presents an IFT procedure with f_{euc}.

Algorithm 1.

Input: An image I, a set S of seed pixels in I, and an Euclidean adjacency relation A;
Output: An optimum-path forest P, and the corresponding cost map C and root map R.
Auxiliary Data structures: A priority queue Q.
(1) For all pixels p of the image I, set $C(p) \leftarrow +\infty$;
(2) For all $p \in S$, set $P(p) \leftarrow nil$, $R(p) \leftarrow p$, $C(p) \leftarrow 0$, and insert p in Q;
(3) While Q is not empty, do
  3.1 Remove from Q a pixel $p = (x_p, y_p)$ such that $C(p)$ is minimum;
  3.2 For each pixel $q = (x_q, y_q)$ such that $q \in A(p)$ and $C(q) > C(p)$, do
    3.2.1 Set $C' \leftarrow (x_q - x_{R(p)})^2 + (y_q - y_{R(p)})^2$, where $R(p) = (x_{R(p)}, y_{R(p)})$ is the root pixel of p;
    3.2.2 If $C' < C(q)$, then
      3.2.2.1. If $C(q) \neq +\infty$, then remove q from Q.
      3.2.2.2. Set $P(q) \leftarrow p$, $C(q) \leftarrow C'$, $R(q) \leftarrow R(p)$, and insert q in Q.

Note that, the IFT algorithm is essentially Dijkstra's shortest-path algorithm [20–23], slightly modified to multiple sources and general path-cost functions. Its correctness

for weaker conditions that are applied to only optimum paths in the graph is presented in [4].

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 [24]. The IFT with f_{euc} allows efficient computation of *multiscale contours by exact dilations* and *multiscale skeletons by label propagation* [7].

3.1. Multiscale contours by exact dilations

Given a set S of points, 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 in 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 Fig. 1 for a contour.

Multiscale contours by exact dilation result from Algorithm 1, where the pixels of the original shape (contour) are taken as the seed set S. Each instance of the multiscale shape is obtained by thresholding the cost map C 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.

3.2. Multiscale skeletons by label propagation

Given a contour with N pixels, its internal skeleton is defined as the geometric location of the centers of maximal disks contained in the contour [25]. A similar definition is valid for the external skeleton.

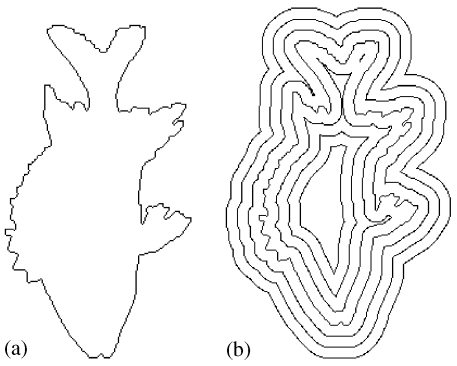

Fig. 1. (a) A contour of a fish and (b) multiscale contours by exact dilation.

Algorithm 1 applied to the contour creates a root map R. Multiscale skeletons can be computed from R if each contour pixel p (root) is assigned to a subsequent label value $\lambda(p)$, varying from 1 to N, while circumscribing the contour (Fig. 2a). A label map L can be created by computing $L(R(p))$ to each image pixel p (Fig. 2b). A more efficient way, however, is to propagate the labels of the contour pixels during Algorithm 1. In this case, the labeling function λ is used in step (2), when the contour pixels are inserted in Q, and the label map L is created similarly and simultaneously to the root map R. A difference image D results from the label map L by computing the following for each pixel p inside and outside the contour (Fig. 2c):

$$D(p) = \max_{\forall q \in A_4(p)} \{\min\{\delta(p,q), N - \delta(p,q)\}\}, \tag{3}$$

where $\delta(p,q) = L(q) - L(p)$ and $A_4(p)$ 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 [7,8,26]. One-pixel wide and connected skeletons can be obtained by thresholding the difference image at subsequent integer values (Fig. 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 Eq. (3) corrects the original equation, reported in Refs. [7,8], as pointed out in Ref. [27].

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

While the topological dimension is restricted to integer values, fractal dimension allows fractionary values. Disseminated by Mandelbrot [28], fractal dimension provides an interesting means for characterizing the self-similarity (or self-affinity) of abstract and real objects, being closely related to the concept of power-laws. A particularly intuitive and useful definition of fractal dimension is the Minkowski–Bouligand dimension [29], 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

$$F = 2 - \lim_{r \to 0} \frac{\log(A(r))}{\log(r)}. \tag{4}$$

In other words, the fractal dimension descriptor in this case (i.e. considering a two-dimensional space) is a number

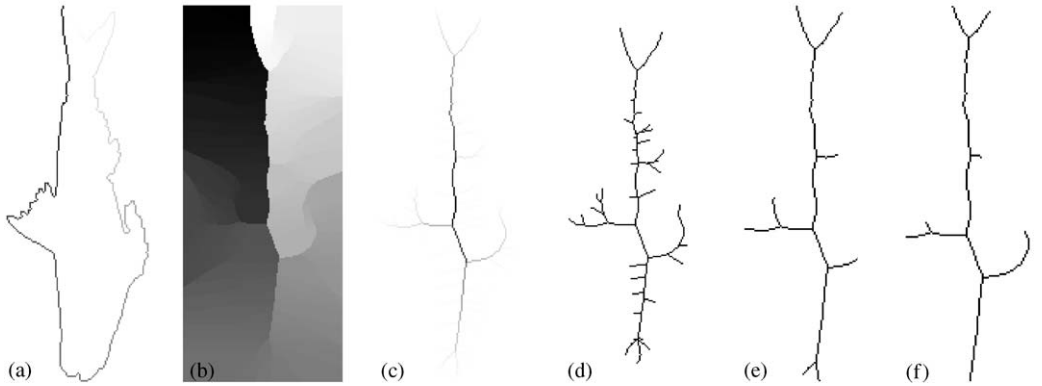

Fig. 2. Multiscale skeletonization by label propagation. (a) Labeled contour, (b) label map, (c) difference image, and (d–f) skeletons at three different scales.

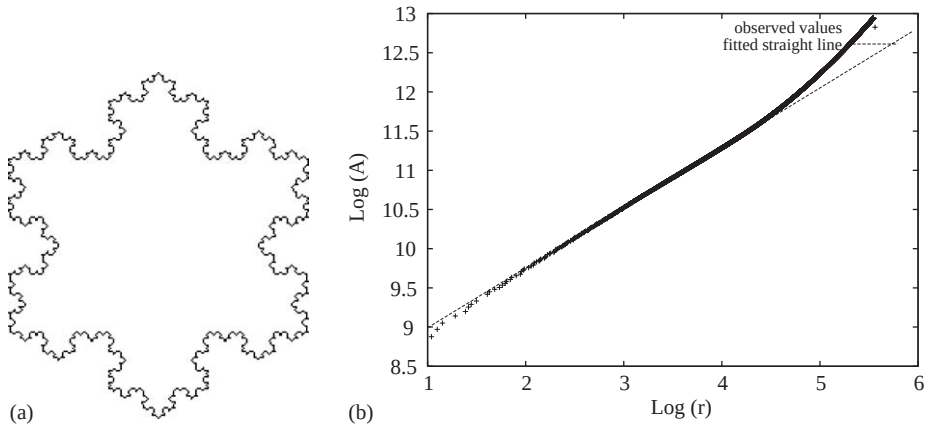

Fig. 3. (a) An object similar to the Koch star, whose fractal dimension is known as about 1.26 (=log4/log3). (b) The logarithmic area function. By taking 2 minus the inclination of the fitted straight line, the fractal dimension obtained is about 1.23.

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 Figs. 3a and b). 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 IFT with f_{euc}. 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 aforementioned experimental method, several practical applications of the fractal dimension have been reported in the literature (e.g. Ref. [30]). Fractal dimensions have been considered as features useful for expressing the area coverage and the "complexity" of shapes ranging from neurons [9] to heartbeat dynamics [31]. In the particular case of the Minkowski–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 intricate 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 [2]. This approach involves taking into the infinitesimal limit the previous concept of linear interpolation [10,11], 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. Two multiscale fractal vectors are compared using the L_2 metric.

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

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* [2], are expected to be greater than the influence areas of the other points of the shape. 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 around for a concave point (point B, see Fig. 5). The influence area A of each salience point relates to the aperture angle θ, illustrated in Fig. 5, by the formula:

$$Area = \frac{\theta \times r^2}{2}, \tag{5}$$

where r is a dilation radius. Costa et al. [2] proposed to estimate the salience points by thresholding the influence areas, computed for low values of r (e.g. $r = 10$). The influence area A of each pixel belonging to a shape (contour or skeleton) can be simply obtained from the histogram of the root map R created by Algorithm 1, restricted to pixels p where $C(p) \leqslant r^2$. 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 this problem for contours is described next.

For a given contour, multiscale internal and external skeletons are first obtained by label propagation as described in Section 3.2. For small scales (e.g. 5% of the maximum label difference $N - 1$), 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 [13]

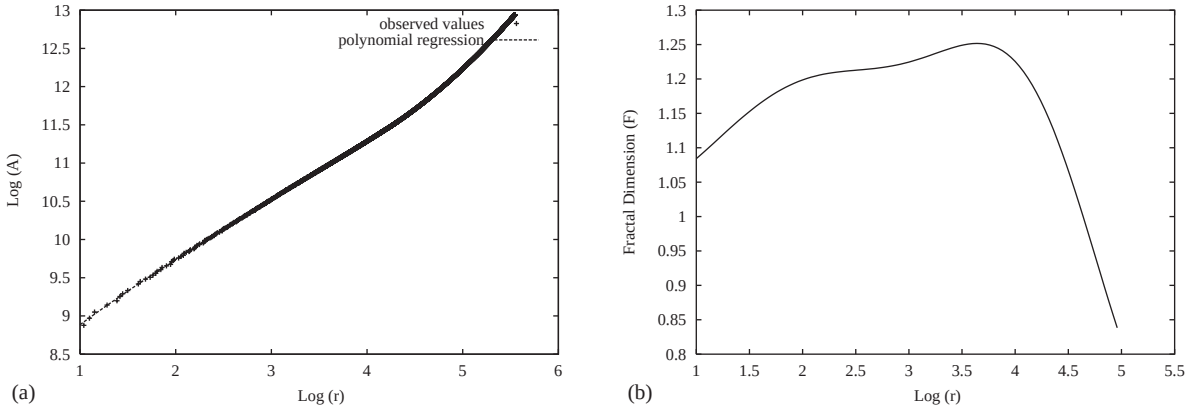

Fig. 4. (a) The log × log curve of the areas of each exact dilation radius for Fig. 3a. (b) The multiscale fractal dimension of its contour.

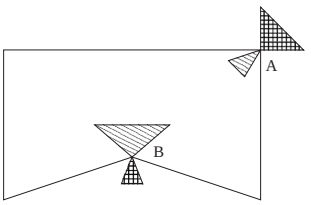

Fig. 5. Internal and external influence areas of a convex (A) and a concave (B) point.

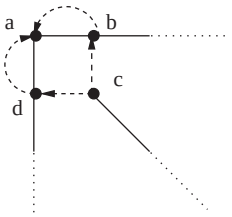

Fig. 7. Relation between skeleton and contour saliences.

(see Fig. 6). Let L, R_i, R_e be the label and root maps resulting from Algorithm 1 applied to the contour with label propagation, to the internal skeleton, and to the external skeleton, respectively. The influence areas of each point of the skeletons are determined based on the histogram of R_i and R_e, restricted to pixels within a narrow band around the skeletons (e.g. $r = 10$). The salience points of the skeletons are those with influence area greater than an area threshold obtained by setting $\theta = 70$ in Eq. (5). In order to locate the salience points along the contour from the salience points of the skeletons, the algorithm uses the label map L as follows. Note that, Eq. (3) essentially assigns to each pixel inside and outside the contour the maximum length of the shortest contour segment between two roots equidistant to the pixel. Fig. 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 the root pixel of c is b, 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 can 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 $\delta(p,q) > (N - \delta(p,q))$ in Eq. (3) 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.

The location and the influence area of the salience points along the contour represent important local and global information for shape analysis. The influence areas (salience values) are obtained from the histogram of L restricted to a narrow band around the contour (e.g. $r = 10$). They are signed negative for concave points and positive for convex points. An arbitrary point of the contour is taken as reference point and the algorithm computes the relative position of each salience point with respect to the reference point along the contour. Finally, a *contour saliences descriptor* is defined as two vectors of the same size: one with the salience values 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.

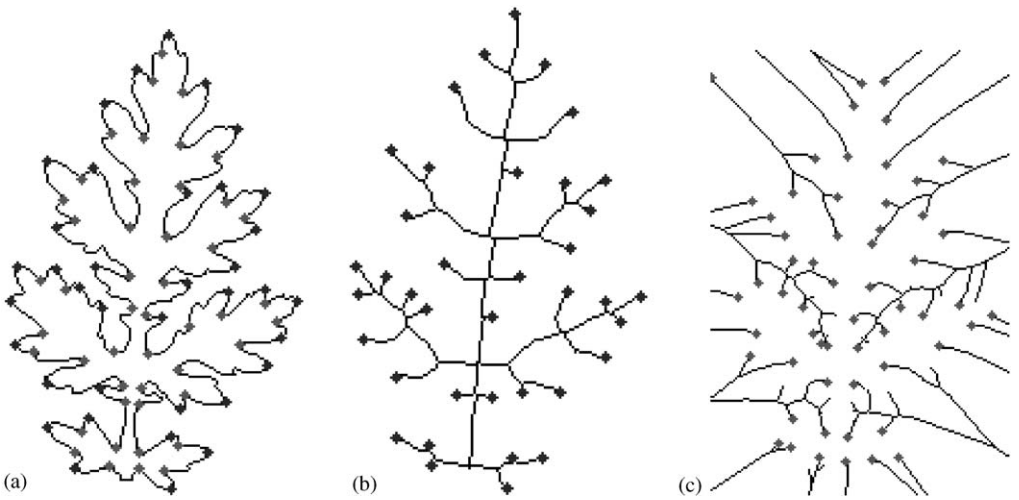

Fig. 6. (a) Saliences of the contour of a leaf and (b–c) saliences of its internal and external skeletons.

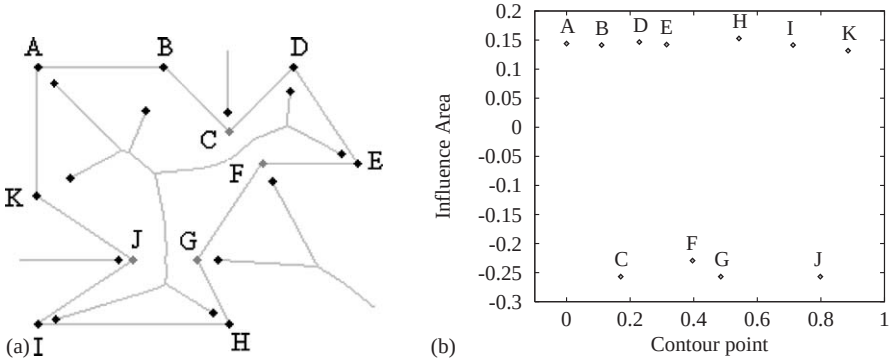

Fig. 8. (a) Contour and skeletons of a polygon, where salience points are indicated by dots. (b) The salience values of the vertices of the polygon by their relative position along the contour.

Fig. 8 illustrates the contour saliences descriptor for a polygon. The contour of the polygon, its reference point (A), the internal and external skeletons, and the respective salience points are indicated in Fig. 8a. Fig. 8b indicates the salience values of the vertices of the polygon by their relative position along the contour.

5. Evaluation

For classification purposes, a descriptor is considered more effective than another one when it increases the number of correctly classified objects. These objects are organized into classes according to some semantic criterion. A "good" shape descriptor should represent different classes of objects by compact clusters of points separated from each other in the corresponding feature space. These aspects ask for two concepts: *compact-ability* and *separability*. The compact-ability of a descriptor indicates its invariance to the object characteristics that belong to a same class, while the separability indicates its discriminatory ability between objects that belong to distinct classes. They evaluate the "goodness" of a description independent of the classification method. Moreover, the separability determines the effectiveness of the descriptor independent of the compact-ability. However, the compact-ability gives an idea of how the separability may be affected if the number of classes increases.

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 [32] for the experiments. Fig. 9 shows some examples of fish contours and their respective skeletons together with the respective salience points.

Since there is no semantic definition of classes for the fish contours in this database, each class is defined as consisting of 10 different manifestations of each contour by rotation and scaling. Then, the problem consists of 1100 classes with 10 shapes each, totalizing 11,000 contours. 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.

A precise definition of compact-ability and separability, the matching algorithm for contour saliences, and the experiments are presented in the next sections.

5.1. Compact-ability

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

$$\phi_D(C) = 1 - \frac{\sum_{\forall i,j \in C} \Delta_D(i,j)}{|C|^2 \max_{\forall i,j \in C}\{\Delta_D(i,j)\}}, \tag{6}$$

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. Note that this measure is normalized with respect to the maximum distance between a pair of shapes, considering all shapes in the class C.

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

$$\Phi_D = \frac{\sum_{\forall C \in \Sigma} \phi_D(C)}{|\Sigma|}, \tag{7}$$

where $|\Sigma|$ is the number of classes in the set Σ.

5.2. Separability

Let Σ be a set (database) of κ 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 $\overline{\Delta_D(r_C, i)} =$

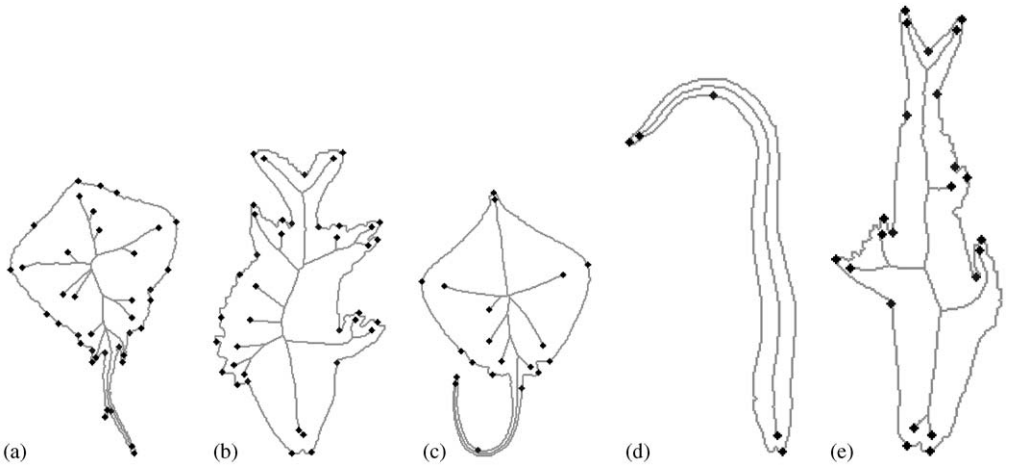

Fig. 9. Fish images used for descriptor evaluation. The concave points were determined through the salience points of the external skeleton, not shown in the figure.

$\Delta_D(r_C, i)/M$, where $M = \max_{\forall i \in \Sigma, \forall r_C} \{\Delta_D(r_C, i)\}$, is computed for all shapes i in Σ.

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

$$\psi_{\mathcal{D}}(C) = 1 - \frac{\eta_{r_C}(x)}{\kappa}. \tag{8}$$

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.:

$$\Psi_D = \frac{\sum_{\forall C \in \Sigma} \psi_{\mathcal{D}}(C)}{|\Sigma|}. \tag{9}$$

5.3. Matching algorithm for contour saliences

Whenever two contours 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 contour saliences requires matching between contours.

The contour saliences 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 is based on the matching algorithm proposed to match CSS images presented in Refs. [16,17].

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 contours 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 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 $|s'_{Ai} - s'_{Bj}| \leqslant 0.2 s'_{A1}$. A pair $((u'_{Bj}, s'_{Bj}), (u'_{Ai}, s'_{Ai}))$ belongs to the list L if $|s'_{Bj} - s'_{Ai}| \leqslant 0.2 s'_{B1}$.
(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

$$d = \sum_{k=1}^{\min\{n,m\}} d_k,$$

where

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

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 scale invariant. In order to reduce this problem the contours have been first normalized according to their diameter. Even though the area thresholding method can be applied to locate the salience points of an external skeleton. These points may not correspond to relevant concave salience points along the contour. The reason is that the external skeleton may present spurious branches due to rotation and scaling of the contour, and the salience points of those branches should not be considered. In fact, they can be eliminated for distinct contours by varying the area threshold in Eq. (5) (see Fig. 9). However, a fixed area threshold (i.e. $\theta = 70$ and $r = 10$ in Eq. (5)) may affect the performance of the contour saliences descriptor with the concave points are considered. Therefore, the experiments used only the convex salience points along the contour.

Table 1 shows the set of implemented shape descriptors. The proposed descriptors (*D2* and *D3*) are compared with the single fractal dimension (*D1*), two classical descriptors (Fourier descriptors (*D4*) and moment invariants (*D5*)) and two recently published shape descriptors (CSS (*D6*) and BAS (*D7*)). Many versions of these methods have been presented, but this work considers their conventional implementations.

Fourier descriptors: The Fourier descriptors of a contour consist of a feature vector with the 126 most significant coefficients of its Fourier Transform using the method described in Refs. [33,34]. The Euclidean distance was used to measure the similarity between two Fourier-descriptors vectors.

Moment invariants: For Moment Invariants, each object was represented by a 14-dimensional feature vector, including two sets of normalized Moment Invariants [15,35], one from the object contour and another from its solid silhouette. Again, the Euclidean distance was used to measure the similarity between different shapes represented by their Moment Invariants.

CSS descriptor: The CSS descriptor is a shape descriptor, adopted in MPEG-7 standard [36], which represents a multiscale organization of the curvature zero-crossing points of a planar curve. The extraction algorithm of the CSS descriptor is described in Refs. [16,17]. A special matching algorithm is necessary to compare two CSS descriptors (e.g. the algorithm presented in Section 5.3). The experiments used a C version of the Matlab prototype presented in Ref. [37].

BAS: The BAS [18,19] is a novel shape descriptor which has been compared with several others [17,38–42]. In Ref. [18], it is shown that BAS functions with 40 and 60 samples outperform all of them. The experiments of the present paper used the BAS descriptor with 60 samples. Basically, the BAS descriptor is based on the *beams* originated from a contour pixel. A beam is defined as the set of lines connecting a contour pixel to the rest of the pixels along the contour. At each contour pixel, the angle between a pair of lines is calculated, and then the shape descriptor is defined by using the third-order statistics of all the beam angles in a set of neighborhood systems. The BAS algorithm is presented in Refs. [18,19]. The similarity between two BAS moment functions is measured by an optimal correspondent subsequence (OCS) algorithm as shown in Ref. [18].

Table 1
List of evaluated descriptors

Descriptor id	Descriptor name
D1	Fractal dimension
D2	Multiscale fractal dimension
D3	Contour saliences
D4	Fourier descriptors
D5	Moment invariants
D6	Curvature scale space (CSS)
D7	Beam angle statistics (BAS)

5.5. Experimental results

Initially, the multiscale fractal dimension of a contour (*D2*) was compared with its single fractal dimension (*D1*). Fig. 10 shows that the multiscale version of the fractal dimension descriptor presents the best separability curve.

Fig. 11 shows the separability curves of the proposed descriptors (*D2* and *D3*), against the Fourier descriptors (*D4*), the moment invariants (*D5*), the CSS (*D6*) and the BAS (*D7*). Observe that the Contour Saliences (*D3*), CSS (*D6*) and BAS (*D7*) present equivalent performance for search radii less than 18% of their maximum distance. From this point on, the BAS's separability curve (*D7*) decreases quickly, being worse than the separability of the Multiscale Fractal Dimension (*D2*) and of the Fourier descriptors (*D4*) for search radii above 25%. This behavior indicates that

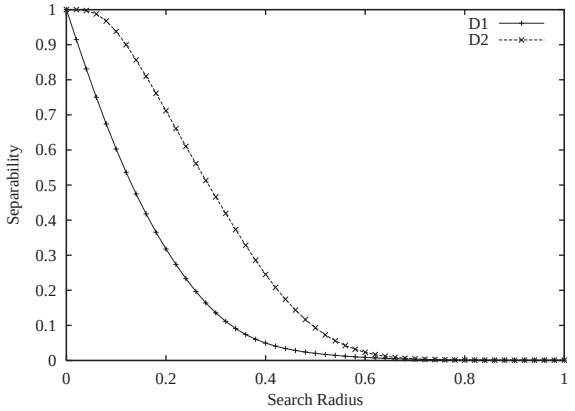

Fig. 10. Multiscale separability diagrams for the shape descriptors based on fractal dimension.

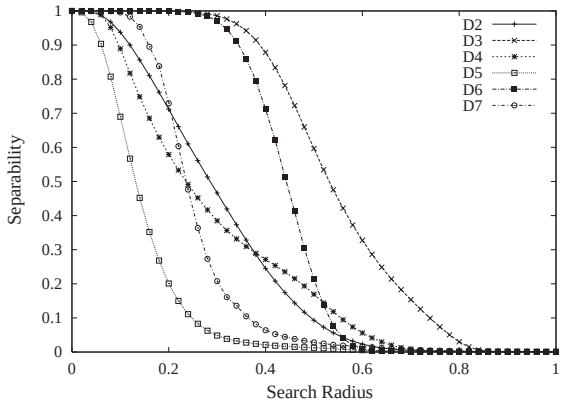

Fig. 11. Comparison of the proposed descriptors with the Fourier descriptors, invariant moments, CSS, and BAS.

Table 2
Compact-ability values of the evaluated descriptors

Descriptor id	Compact-ability
D1	0.93
D2	0.97
D3	0.70
D4	0.77
D5	0.97
D6	0.73
D7	0.95

the BAS descriptor is neither robust nor effective for search radii greater than 25%. The Multiscale Fractal Dimension (*D2*) has a better separability curve than Fourier descriptors (*D4*), moment invariants (*D5*) and BAS (*D7*) for search radii between 25% and 40%. Its performance, however, decreases, being lower than Fourier descriptors (*D4*) for search radii greater than 40%. The most relevant result is that the Contour Saliences descriptor (*D3*) has the best separability curve considering all search radii. Although its performance is equivalent to the famous CSS descriptor (*D6*) for search radii less than 30%, it is more robust and more effective for higher search radii.

Table 2 presents the compact-ability values of the evaluated shape descriptors. The higher values were found for single fractal dimension (*D1*), multiscale fractal dimension (*D2*), moment invariants (*D5*), and BAS (*D7*), while the contour saliences (*D3*) presented the lowest value. Fortunately, the compact-ability 0.70 of *D3* cannot be considered sufficiently low to interfere in its separability, even considering a database with 1100 classes. According to these experiments, *D3* is more effective than the others.

6. Conclusion

This paper has presented two effective shape descriptors, multiscale fractal dimension and contour saliences, using the framework of the IFT. The presented method to compute multiscale fractal dimension is more efficient [7] and robust than the one published in Ref. [2], since the undesirable oscillations commonly found in Fourier-based approaches have been eliminated here by the use of polynomial regression. The location of salience points along a contour was computed in a direct way using the IFT framework to exploit the relation between the contour and its skeletons [13]. This method is more robust and efficient than the approach presented in Ref. [2]. Moreover, the paper redefines the contour salience descriptor to include point location and salience value along the contour and a special distance metric.

The multiscale fractal dimension and the contour saliences were also extensively evaluated for the first time, using a database with 1100 classes and 11,000 contours. Their "goodness" (compact-ability and separability) have been showed by using as references the single fractal dimension, two classical (Fourier descriptors [14,33] and moment invariants [15,35]) and two recently proposed shape descriptors (CSS [16,17] and BAS [18,19]). The underlying ideas of compact-ability and separability may not be totally new concepts, however this paper has presented an original way to compute them, especially the multiscale separability.

The experiments showed that the contour saliences descriptor was the most effective (with the best separability curve). This is certainly a breakthrough result considering that the experiments have taken into account recent descriptors and a database with 1100 classes. The Multiscale Fractal Dimension is competitive with BAS and Fourier Descriptors in terms of separability and compact-ability, but it is less effective than the CSS and the Contour Saliences. This may indicate that the normalization procedure was not effective to make the Multiscale Fractal Dimension totally scale independent. In Ref. [18], the BAS descriptor was shown to be more effective than CSS for the MPEG-7 Core Experiments shape-1. The experiments with separability showed the opposite result. Note that one cannot say that a descriptor is better than another without taking into account several application domains.

Current work concerns to solve the scale-dependency problem of the multiscale fractal dimension and to incorporate concave points in the composition of the contour saliences descriptor. In view of that, we are investigating a special distance metric for the multiscale fractal dimension and an automatic area thresholding method to avoid salience points of the spurious branches of the external skeleton. We are also interested in validating the proposed descriptors for other application domains. In special, we are currently considering applications in content-based image retrieval.

Acknowledgements

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, and to Nafiz Arica and Fatos Vural for the BAS source code.

References

[1] S. Loncaric, A survey of shape analysis techniques, Pattern Recognition 31 (1998) 983–1001.
[2] 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.
[3] R. Torres, A.X. Falcão, L. Costa, Shape description by image foresting transform, in: 14th International Conference on Digital Signal Processing, Santorini, Greece, 2002, pp. 1089–1092.
[4] A.X. Falcão, J. Stolfi, R.A. Lotufo, The image foresting transform: theory, algorithms, and applications, IEEE Trans. Pattern Anal. Mach. Intell., 26 (1) (2004) 19–29. Source code available at http://www.ic.unicamp.br/∼afalcao/ift.html.
[5] R.A. Lotufo, A.X. Falcão, The ordered queue and the optimality of the watershed approaches, in: J. Goutsias, L. Vincent, D.S. Bloomberg (Eds.), Mathematical Morphology and its Applications to Image and Signal Processing, Vol. 18, Kluwer Academic, Palo Alto, USA, 2000, pp. 341–350.
[6] A.X. Falcão, B.S. da Cunha, R.A. Lotufo, Design of connected operators using the image foresting transform, in: M. Sonka, K. Hanson (Eds.), Proceedings of SPIE on Medical Imaging, San Diego, CA, 2001, pp. 468–479.
[7] A.X. Falcão, L. da F. Costa, B.S. da Cunha, Multiscale skeletons by image foresting transform and its applications to neuromorphometry, Pattern Recognition 35 (7) (2002) 1571–1582.
[8] A.X. Falcão, B.S. da Cunha, Multiscale shape representation by image foresting transform, in: M. Sonka, K. Hanson (Eds.), Proceedings of SPIE on Medical Imaging, Vol. 4322, San Diego, CA, 2001, pp. 1091–1100.
[9] L. da F. Costa, E.T.M. Manoel, A shape analysis framework for neuromorphometry, Network 13 (2002) 283–310.
[10] H. Peitgen, H. Jurgens, D. Saupe, Fractals for the Classroom: Introduction to Fractals and Chaos, Springer, Berlin, 1992.
[11] P. Montague, M. Friedlander, Morphogenesis and territorial coverage by isolated mammalian retinal ganglion cells, J. Neurosci. 11 (1991) 1440–1457.
[12] F.P. Preparata, M.I. Shamos, Computational Geometry: An Introduction, Springer, Berlin, 1985.
[13] M. Leyton, Symmetry-curvature duality, Comput. Vision Graphics Image Process. 38 (1987) 327–341.
[14] T.P. Wallace, P. Wintz, An efficient three-dimensional aircraft recognition algorithm using normalised fourier descriptors, Comput. Graphics Image Process. 13 (1980) 99–126.
[15] M.K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Inf. Theory 8 (1962) 179–187.
[16] S. Abbasi, F. Mokhtarian, J. Kittler, Enhancing CSS-based shape retrieval for objects with shallow concavities, Image Vision Comput. 18 (2002) 199–211.
[17] F. Mokhtarian, S. Abbasi, Shape similarity retrieval under affine transforms, Pattern Recognition 35 (1) (2002) 31–41.
[18] N. Arica, F.T.Y. Vural, BAS: a perceptual shape descriptor based on the beam angle statistics, Pattern Recognition Lett. (2003) 1627–1639.
[19] N. Arica, F.T.Y. Vural, A perceptual shape descriptor, in: International Conference on Pattern Recognition, 2002, pp. 375–378.
[20] E. Moore, The shortest path through a maze, in: Proceedings of International Symposium on the Theory of Switching, Harvard University, 1959, pp. 285–292.
[21] R. Bellman, On a routing problem, Quart. Appl. Math. 16 (1958) 87–90.
[22] E. Dijkstra, A note on two problems in connexion with graphs, Numer. Math. 1 (1959) 269–271.
[23] A. Frieze, Minimum paths in directed graphs, Oper. Res. Quart. 28 (2,i) (1977) 339–346.
[24] V. Castelli, L.D. Bergman (Eds.), Image Databases. Search and Retrieval of Digital Imagery, Wiley, New York, 2002.
[25] R. Kimmel, D. Shaked, N. Kiryati, A.M. Bruckstein, Skeletonization via distance maps and level sets, Comput. Vision Image Understanding 62 (3) (1995) 382–391.
[26] L. da F. Costa, L.F. Estrozi, Multiresolution shape representation without border shifting, Electron. Lett. 35 (21) (1999) 1829–1830, http://cyvision.if.sc.usp.br/msskeletons.
[27] A.X. Falcão, L. da F. Costa, B.S. da Cunha, Erratum to multiscale skeletons by image foresting transform and its applications to neuromorphometry, [Pattern Recognition 35 (7) (2002) 1571–1582], Pattern Recognition 36 (12) (2003) 3013.
[28] B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman and Co., San Francisco, 1982.
[29] C. Tricot, Curves and Fractal Dimension, Springer, Berlin, 1995.
[30] Y.Y. Tang, Y. Tao, E.C.M. Lam, New method for feature extraction based on fractal behavior, Pattern Recognition 35 (2002) 1071–1081.
[31] 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, Phys. Rev. Lett. 86 (26) (2001) 6026–6029.
[32] www.ee.surrey.ac.uk/research/vssp/imagedb/demo.html (2003).
[33] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Addison-Wesley, Reading, MA, 1992.
[34] B.M. Mehtre, M.S. Kankanhalli, W.F. Lee, Shape measures for content based image retrieval: a comparison, Inform. Process. Manage. 33 (3) (1997) 319–337.
[35] S.A. Dudani, K.J. Breeding, R.B. McGhee, Aircraft identification by moment invariants, IEEE Trans. Comput. c-26 (1) (1977) 39–45.
[36] M. Bober, MPEG-7 visual shape descriptors, IEEE Trans. Circuits Syst. Video Technol. 11 (6) (2001) 716–719.
[37] C.Y. Ming, Shape-based image retrieval in iconic image databases, Master's Thesis, Chinese University of Hong Kong, June 1999.
[38] G. Chuang, C.-C. Kuo, Wavelet descriptor of planar curves: theory and applications, IEEE Trans. Pattern Anal. Mach. Intell. 5 (1) (1996) 56–70.
[39] L.J. Latecki, R. Lakamper, Shape similarity measure based on correspondence of visual parts, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1185–1190.

[40] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Trans. Pattern Anal. Mach. Intell. 24 (24) (2002) 509–522.
[41] A. Khotanzan, Y. H. Hong, Invariant image recognition by zernike moments, IEEE Trans. Pattern Anal. Mach. Intell. 12 (5) (1990) 487–489.
[42] L.-J. Lin, S.-Y. Kung, Coding and comparison of DAGs as a novel neural structure with applications to on-line handwriting recognition, IEEE Trans. Signal Process. 45 (11) (1997) 2701–2708.

About the Author—RICARDO DA SILVA TORRES received a B.Sc. in Computer Engineering from the University of Campinas, Brazil, in 2000. Currently he attends a Ph.D. program in Computer Science at the same university.

About the Author—ALEXANDRE XAVIER FALCÃO received a B.Sc. in Electrical Engineering from the University of Pernambuco, Brazil, in 1988. In 1993, he received a M.Sc. in Electrical Engineering from the University of Campinas, Brazil. He got his Ph.D. in Electrical Engineering from the University of Campinas in 1996. He has been Professor at the University of Campinas since 1998, and his research interests include image analysis, 3D visualization, image retrieval, and medical imaging.

About the Author—LUCIANO DA FONTOURA COSTA got a B.Sc. in Electronic Engineering from the University of São Paulo, Brazil (1984), and a Ph.D. degree from King's College, University of London (1992). He has been with the Institute of Physics of the University of São Paulo in São Carlos since 1985, where he started and has since coordinated the Cybernetic Vision Research Group. Luciano is co-author of the book Shape Analysis and Classification, CRC Press, 2001.