INTERNATIONAL COMPUTER SCIENCE INSTITUTE

1947 Center Street • Suite 600 • Berkeley, California 94704 • 1-510-642-4274 • FAX 1-510-643-7684

Describing and Recognising Shape through Size Functions

Claudio Uras* and Alessandro Verri†

TR-92-057

September 1992

Abstract

According to a recent mathematical theory the intuitive concept of shape can be formalised through functions, named *size functions*, which convey information on both the topological and metric properties of the viewed shape. In this paper the main concepts and results of the theory are first reviewed in a somewhat intuitive fashion. Then, an algorithm for the computation of discrete size functions is presented. Finally, by introducing a suitable distance function, it is shown that size functions can be successfully used for both shape description and recognition from real images.

*C. Uras was supported by a fellowship from ELSAG-Bailey S.p.A.

†A. Verri is on leave from the Dipartimento di Fisica dell'Università di Genova, Italy

1 Introduction

Shape description and recognition are important stages of vision. From the computational perspective, many problems stem from the well known difficulty of dealing with qualitative and quantitative changes in shape within the same scheme.

Studying shape through integer-valued functions, called *size functions* [1, 2, 3], has recently been proposed. The key idea underlying the concept of a size function is that of setting *metric* bounds to the classical notion of *homotopy*, i.e., of continuous deformation. Size functions are very good candidates for shape representation because they *(i)* convey information about both the qualitative and quantitative structure of the viewed shape, *(ii)* can be tailored to suit the invariant properties of the shapes to be studied, and *(iii)* are inherently "stable" against small changes in shape.

The aim of this paper is to assess the potential of the theory of size functions for computer vision. Therefore, after a brief summary of the main concepts of the theory, an algorithm for the computation of size functions in the discrete case is described. Then a simple way to measure distances between size functions is proposed and tested on real images. Finally, the main conclusions which can be drawn from our research are summarised.

2 A Simple Example

First, let us introduce the notion of a size function through a simple example. The aim of this Section is to generate a description of the curve in Fig. 1a which is useful for shape recognition.

As a preliminary step, let us define a transformation H which brings a point of α onto some other point of α without leaving the curve. The arrows in Fig. 1a, for example, help visualise two possible "trajectories" along which H brings the point p onto the point q. The transformation H induces an equivalence relation on the points of α, where two points u and v are said to be *H-equivalent* if there exists a continuous trajectory on α which brings u onto v. Since, independent of the shape of α, all the points fall into one and the same equivalence class, the purely topological concept of H-equivalence is clearly not sufficient to characterise the shape of α. Intuitively, this is due to the absence of "topological obstructions" between points of α.

Let us now change the definition of H-equivalence slightly by introducing "metric obstructions" along the trajectories of H on α. For example, let c be the center of mass of α and $D_c(s)$ denote the distance between c and a point s of α. In Fig. 1b the continuous lines identify the points with $D_c \leq 3$, the dashed lines the points with $3 < D_c \leq 7$, while the points with $D_c > 7$ have not been drawn. The gaps in Fig. 1b make it clear that D_c, which is called a *measuring function*, eventually exceeds 7 (the "metric obstruction") along any trajectory from p to q. This suggests that H should be redefined so that two points u and v are said to be *$H(D_c \leq y)$-equivalent* if a trajectory exists on α from u to v along which D_c never exceeds y. It is evident that

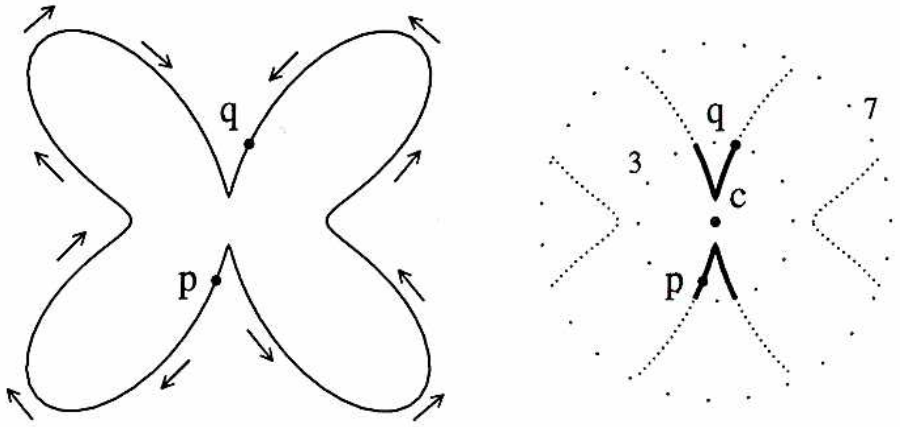

Figure 1: Topological and metric obstructions. (a) Since the curve α has no topological obstruction, the point p can be brought into coincidence with q without leaving α by following either one of the two trajectories indicated by the arrows. (b) If, along either trajectory, D_c, the distance from the center of mass c, cannot be larger than 7 (the metric obstruction), p cannot be brought into coincidence with q any longer. The size function $l_{D_c}(\alpha)$ at the point $(3,7)$ equals 2 because 2 of the 4 connected components of the set of points within the larger circle (the points with $D_c \leq 7$) contain at least a point within the smaller circle (the points with $D_c \leq 3$).

not all the points of α are $H(D_c \leq y)$-equivalent for some value of y and that the number of equivalence classes depends on the shape of α. In Fig. 1b, for example, p is not $H(D_c \leq 7)$-equivalent to q.

The notion of $H(D_c \leq y)$-equivalence is essential for the definition of the size function. For each pair of real numbers (x, y), the *size function* $l_{D_c}(\alpha; x, y)$ induced by the measuring function D_c counts the number of equivalence classes in which the equivalence relation $H(D_c \leq y)$ divides the set of points of α with $D_c \leq x$ (for $y \geq x$). In practice, the size function $l_{D_c}(\alpha; x, y)$ can be computed by counting the number of connected components of the set of points of α with $D_c \leq y$ that contains at least one point with $D_c \leq x$. In the particular example of Fig. 1b we have $l_{D_c}(\alpha; 3, 7) = 2$, because the set of points with $D_c \leq 7$ has 4 connected components, two of which contain at least one point with $D_c \leq 3$ (i.e., a point on a continuous line).

The diagram of $l_{D_c}(\alpha; x, y)$ is shown in Fig. 2a (the horizontal and vertical axis are the x- and y-axis respectively). The size function $l_{D_c}(x, y)$ is a piecewise constant function which, within the triangular region $T = \{(x, y) : 0 \leq x < 10.56..., x \leq y < 10.56...\}$, equals 0 for $0 \leq x < 1$, 2 for $1 \leq x < 4$, and 4 for $4 \leq x < 10.56....$ The numbers 1, 4, and 10.56... are the critical values of the measuring function D_c (see Fig. 2b). The value of l_{D_c} elsewhere is independent of α. In essence, $l_\varphi = 0$ on the left of the vertical axis (there are no points to start with), $l_\varphi = 1$ for $y \geq 10.56...$ and $0 \geq x \geq y$ (all the metric obstructions are removed), and $l_\varphi = \infty$ for $x > 0$ and $y < x$ (each point belongs to a different equivalence class).

3 Main Definitions

Let us now define and comment on the notion of size function in more general terms. We first establish some basic notation.

In this Section a shape is an n-dimensional, compact, boundaryless, piecewise C^∞ submanifold $\mathcal{M}$ of the Euclidean space E^m $(n < m)$ [1]. The set of k-tuples p of points p_i of $\mathcal{M}$, $i = 1, ..., k$, is denoted by $\mathcal{M}^k$ (in the example of Fig. 1, it was simply $k = 1$ and thus $\mathcal{M}^1 = \alpha^1 = \alpha$). If p and q are in $\mathcal{M}^k$, let $d_k(p, q) = \max_{0 \leq i \leq k}\{d(p_i, q_i)\}$ be the distance between p and q, where $d(p_i, q_i)$ is the usual Euclidean distance between p_i and q_i. The important concept of measuring function can now be defined.

A *measuring function* [1] is any continuous function

$$\varphi : \mathcal{M}^k \to \Re.$$

The notion of measuring functions leads to the key concept of metric homotopy [1].

A *metric homotopy* between p and q in $\mathcal{M}^k$ is a continuous function $H : [0, 1] \to \mathcal{M}^k$ such that

- $H(0) = p$, $H(1) = q$ (the usual definition of homotopy);

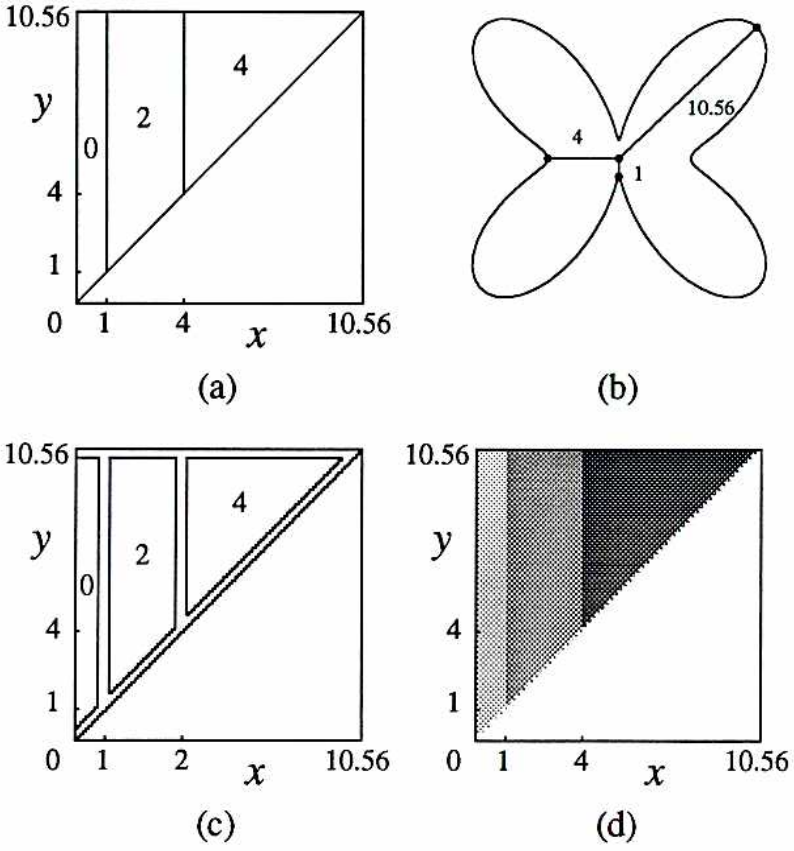

Figure 2: Representing size functions. (a) The size function $l_{D_c}(x,y)$ of the curve α is a piecewise constant function which equals 0, 2, and 4 for $0 \leq x < 1$, $1 \leq x < 4$, and $4 \leq x < 10.56...$ respectively, within the triangular region $T = \{(x,y) : 0 \leq x < 10.56..., x \leq y < 10.56...\}$. (b) The numbers 1, 4, and 10.56... are the local minimum and maximum values of the measuring function D_c. (c) The size function $\bar{l}_{D_c}(x,y)$, output of the algorithm described in the text, consists of thin regions, named "blind stripes", which go all around the regions where $\bar{l}_\varphi$ is known exactly and in which the "true" value of $\bar{l}_\varphi$ is bounded by the monotonicity constraints but otherwise uncertain. (d) However, a grey-value coded representation of $\bar{l}_{D_c}$ shows that the estimated value of $\bar{l}_{D_c}$ within the blind stripes is mostly correct.

- $\varphi[H(\tau)] \leq y \quad \forall\, \tau \in [0,1]$.

We write $p \simeq_{\varphi \leq y} q$, if such a metric homotopy exists. Let now $\mathcal{M}^k(\varphi \leq x)$ be the set of points p in $\mathcal{M}^k$ with $\varphi(p) \leq x$ (e.g. the set of points on the continuous lines of Figs. 1b). We have the following

The size function $l_\varphi(\mathcal{M}) : \Re^2 \rightarrow N \cup \{+\infty\}$ [1] can be defined as

$$(x,y) \hookrightarrow \begin{cases} \#\{\mathcal{M}^{k+1}(\varphi \leq x)/\ \simeq_{\varphi \leq y}\} & \text{if finite,} \\ \\ +\infty \text{ otherwise.} \end{cases}$$

A fundamental theorem of the theory [1] ensures that the value of the size function inside the triangular region $T_\varphi(\mathcal{M}) = \{(x,y) : \varphi^{\min} \leq y \leq \varphi^{\max}, \varphi^{\min} \leq x \leq y\}$, where $\varphi^{\min}$ and $\varphi^{\max}$ are the minimum and maximum value of φ on α respectively, is finite. In what follows, for the sake of simplicity, let us assume that the value of the size function on the boundary of $T_\varphi(\mathcal{M})$ is also finite.

With respect to the previous Section, the definition of size function has been extended in two important directions. First, a size function can be defined on piecewise smooth surfaces of arbitrary dimension. Second, the measuring function does not need to be the distance from the center of mass and can be defined on k-tuples of the shape. In the case of a curve, for example, the curvature, the distance between pairs of points, and the area of the triangle whose vertices lie on the curve, could equally have been used as measuring functions with $k = 1, 2$, and 3 respectively.

A size function $l_\varphi(\mathcal{M}; x, y)$ has a number of general properties. First, by definition, l_φ is nondecreasing in x and nonincreasing in y (see Fig. 2a). Second, l_φ inherits the invariant properties of the measuring function φ (thus, in the previous Section $l_{D_c}(\alpha)$ is invariant for translation and rotation of α on the plane of Fig. 1a). Third, although l_φ can be defined over the entire plane, the relevant information is contained within the triangular region $T_\varphi(\mathcal{M})$.

The relevance of the notion of size function to shape analysis is due to the fact that the main properties of size functions can be extended to the discrete case [2] with little change. In the next Section an algorithm for the discrete computation of size functions will be described and the main properties of the obtained representation discussed. A rigorous account of the mathematical foundations of the theory of size functions in both the continuum and discrete case can be found in ref. [1, 2].

4 Computing Size Functions

This Section describes the implementation of an algorithm for the discrete computation of the size function of a planar curve α and discusses the main properties of size functions in both the continuum and discrete case. For the sake of simplicity, let us restrict the discussion to the case in which the measuring function φ is defined on single points of α (that is, $k = 1$) with $\varphi \geq 0$ (a more general description can be

found in ref. [4]). In addition, let $B(p)_\delta$ be the open circle of center p and radius δ, and l_φ and $\bar{l}_\varphi$ the size function in the continuum and discrete case respectively. The algorithm consists of four steps.

1. Sample (or approximate) the curve α at a finite number N of points p^i, $i = 1, ..., N$, so that *(i)* $\alpha \subset \cup_{i=1}^{N} B(p_i)_\delta$ and *(ii)* the set $B(p_i)_\delta \cap \alpha$ is non-empty and connected for $i = 1, ..., N$. Compute $\varphi(p_i)$, $i = 1, ..., N$.

2. Define the graph G whose vertices are the points p^i and whose edges link vertices which correspond to adjacent points on α.

3. Compute the maximum $\varphi^{\max}$ of $\varphi(p^i)$, $i = 1, ..., N$ and set $\Delta \geq \epsilon_\varphi(\delta)$, where $\epsilon_\varphi(\delta)$ is the modulus of continuity of φ at δ.

4. For $y = 0$ to $y \leq \varphi^{\max}$

   (a) Define the subgraph $G_{\varphi \leq y}$ of G induced by the set of vertices of G for which $\varphi \leq y$.

   (b) For $x = 0$ to until $x \leq y$

      i. Let $\bar{l}_\varphi(\alpha; x, y)$ be the number of connected components of $G_{\varphi \leq y}$ which contain at least a vertex p^i such that $\varphi(p^i) \leq x$.

      ii. $x \rightarrow x + \Delta$.

   (c) $y \rightarrow y + \Delta$.

The conditions (i) and (ii) of the first step ensure that the curve α is covered in such a way that each open circle contains exactly one connected arc of α. The graph G, in the second step, is a discrete representation of α such that a path on G between the vertices p_i and p_j is the discrete counterpart of a trajectory between points of the two arcs $B(p_i)_\delta \cap \alpha$ and $B(p_j)_\delta \cap \alpha$. The third step determines the minimal resolution at which $\bar{l}_{D_c}$ can be computed. In the final step $\bar{l}_{D_c}$ is computed over a grid of equally spaced points within the triangular region $T_\varphi(\alpha) = \{(x, y) : 0 \leq y \leq \varphi^{\max}, 0 \leq x \leq y\}$.

The diagram of $\bar{l}_{D_c}$, that is, the output of the algorithm when α is sampled at 100 points, is shown in Fig. 2c. Somewhat unexpectedly, the diagram of Fig. 2c does not consist of a set of discrete estimates. This is made possible by a fundamental theorem of the theory [2] which ensures that if $\Delta x, \Delta y \geq \epsilon_\varphi(\delta)$ (see the third step of the algorithm) and $\bar{l}_\varphi = n$ at two different points, then, at "every" point in between, $\bar{l}_\varphi = n$ and, most importantly, $l_\varphi = n$. Consequently, there are three areas in the diagram of Fig. 2c where $\bar{l}_{D_c}$ is known to be equal to l_{D_c} and to 0, 2, and 4 respectively with "no margin of error". The differences between $\bar{l}_\varphi$ and l_φ are located nearby the points where $\bar{l}_\varphi$ is not constant. In fact, it can be shown [2] that if $\bar{l}_\varphi$ takes on different values at two adjacent points along either axis, none of the two estimates is *a priori*

equal to l_φ. Thus, there are regions, named "blind stripes", which go all around the locations where $\bar{l}_\varphi$ is known exactly and in which the "true" value of $\bar{l}_\varphi$ is bounded by the monotonicity constraints along the coordinate axes but is otherwise uncertain (see Fig. 2d, however). Intuitively, the width of the blind stripes reflects the coarseness of the sampling stage and the ambiguity of the finite covering of the first step of the algorithm. A finer sampling would narrow down the width of the blind stripes, thereby reducing the uncertainty in the location and value of the discontinuities of l_φ.

5 Experimental Results

Let us present some experimental results on the computation and use of size functions for object recognition from real images. First, a distance between size functions needs to be defined.

5.1 A Distance between Size Functions

Let φ be the measuring function with $\varphi \geq 0$, α_1 and α_2 two planar curves, and $\varphi^{\max}(\alpha_i)$ the maximum of φ on α_i, for $i = 1, 2$. Let us scale φ by defining $\hat{\varphi} = \varphi/\varphi^{\max}(\alpha_i)$ on α_i, for $i = 1, 2$. As a result, $\hat{\varphi}(\alpha_1) = \hat{\varphi}(\alpha_2) = 1$ and a scale-invariant distance D between the size functions $l_\varphi(\alpha_1)$ and $l_\varphi(\alpha_2)$ can be defined simply as

$$D\left(l_{\hat{\varphi}}(\alpha_1), l_{\hat{\varphi}}(\alpha_2)\right) = 2\int_0^1 dy \int_0^y dx |l_{\hat{\varphi}}(\alpha_1; x, y) - l_{\hat{\varphi}}(\alpha_2; x, y)|.$$

Similarly, in the discrete case, $\bar{l}_{\hat{\varphi}}(\alpha_1)$ and $\bar{l}_{\hat{\varphi}}(\alpha_2)$ can be computed at the same fixed resolution R and regarded as triangular matrices $\bar{l}_{\hat{\varphi}}(\alpha_1)_{i,j}$ and $\bar{l}_{\hat{\varphi}}(\alpha_2)_{i,j}$ with $i = 1, ..., R-1$ and $j = 1, ..., R-i$. The distance D can then be redefined as

$$D\left(\bar{l}_{\hat{\varphi}}(\alpha_1), \bar{l}_{\hat{\varphi}}(\alpha_2)\right) = \frac{2}{R(R-1)} \sum_{i=1}^{R-1} \sum_{j=1}^{R-i} | \bar{l}_{\hat{\varphi}}(\alpha_1)_{i,j} - \bar{l}_{\hat{\varphi}}(\alpha_2)_{i,j} | \qquad (1)$$

where the normalisation factor is chosen so that $D = 1$ if, on average, the triangular matrices $\bar{l}_{\hat{\varphi}}(\alpha_1)$ and $\bar{l}_{\hat{\varphi}}(\alpha_2)$ differ by 1 at each entry. The entries on the diagonal of the triangular matrices are not considered because they may be affected by large quantisation errors.

In order to test the algorithm for the computation of size functions of the previous Section and then assess the usefulness of the concept of size functions for object recognition, some experiments on sets of "real" objects were performed. Let us now describe one of these experiments in some detail.

5.2 Leaf Recognition

Fig. 3 shows the images of six leaves from six different plant species (from upper left to lower right: ivy, lemon, oleander, pittosporum, oak, and olive). Each leaf was picked

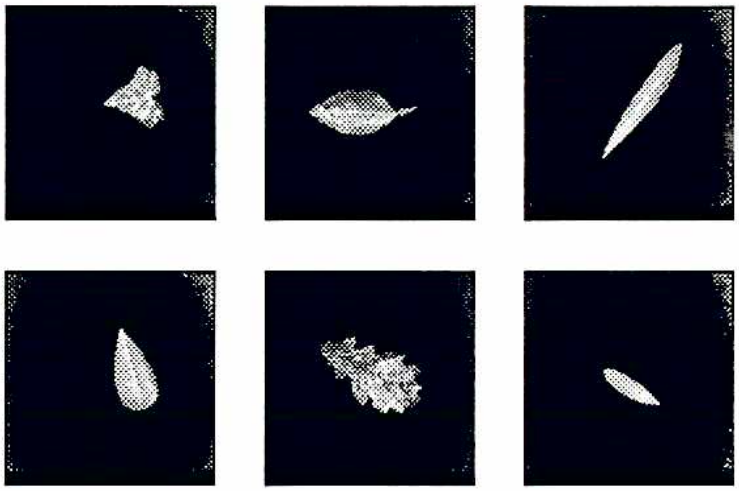

Figure 3: Six images of six different species of leaves. From upper left to lower right: ivy, lemon, oleander, pittosporum, oak, and olive.

from a set of eight leaves of the same species for a total of 48 leaves and one image of each leaf against a dark background was taken. Standard edge detection techniques were applied to extract the silhouette of each leaf [5] and the size function l_{D_c} of each leaf was then computed over a grid of fixed resolution. The distance from the center of mass was always normalised between 0 and 1. In order to test the invariance of l_{D_c} for translation and rotation, the position and orientation of each leaf on a plane nearly parallel to the image plane was varied from image to image.

Fig. 4 shows the grey-value coded "size functions" which have been obtained by averaging the size functions obtained from leaves of the same species (the value at each point of these "size functions" is the average of the value of the size functions at that point). A number of qualitative conclusions can already be drawn by a simple inspection of Fig. 4.

First, the size functions of the ivy leaves appear to be consistently different from all the other size functions. Second, the size functions of oleander and olive leaves are qualitatively similar but quantitatively different (the location of the main discontinuity along the horizontal axis of the oleander size function is further to the left). Third, the difference between the size functions of oak and lemon leaves is localised near the diagonal (which is where the shape "details" can be detected). Fourth, it might not be easy to distinguish between lemon, olive, and pittosporum leaves. Lastly, the size functions of lemon, pittosporum, and oak leaves present a higher degree of variability (larger regions over which the average takes on non-integer values). Quite interestingly, similar conclusions could have been drawn by looking at the original set of leaves.

From the quantitative point of view, the average "size functions" of Fig. 4 have

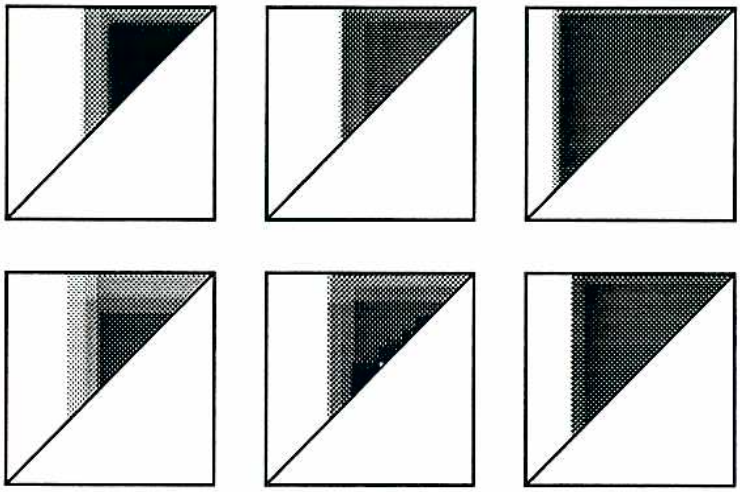

Figure 4: Grey-value coded "size functions" which have been obtained by averaging the size functions obtained from leaves of the same species. From upper left to lower right: ivy, lemon, oleander, pittosporum, oak, and olive.

been used to classify each leaf according to a simple recognition scheme. Each of the "size function" of Fig. 4 has been considered as a "model" for each of the six species and the distance D between each leaf and each model has been computed by using Eq. 1. Finally, each leaf was classified depending on the minimum distance. The described method was able to classify all the leaves correctly with the exception of a lemon leaf which was mistaken for a pittosporum leaf. A more robust technique which combines descriptions from different size functions and correctly classifies each leaf will be described in a forthcoming paper [6].

Similar results (with no classification errors) have been obtained by looking at images of tools such as pliers, screwdrivers, scissors, wrenches, and hammers of different sizes and quantitatively different shapes.

6 Conclusion

The aim of this paper was to assess the relevance of a recent mathematical theory in relation to the problems of shape description and recognition. The theory is based on the concept of size function which combines topological and metric properties of shape.

Our analysis has shown that size functions in the discrete *(i)* can be computed reliably from real images, *(ii)* preserve the invariant properties of the chosen measuring function, *(iii)* can easily be made independent of change-of-scale, and *(iv)* are inherently robust against small qualitative and quantitative changes of shape. In

conclusion, experiments on real images indicate that the shape representation which can be obtained through size functions is likely to be suitable for object recognition.

We thank Massimo Ferri and Patrizio Frosini for many helpful discussions. Patrizio Frosini and Steve Omohundro made valuable comments on the paper. Clive Prestt checked the English.

References

[1] Frosini, P., 1992. Metric homotopies (Submitted for publication).

[2] Frosini, P., 1992. Discrete computation of size functions. *Journal of Combinatorics, Information and System Sciences* (In press).

[3] Frosini, P., 1991. Measuring shapes by size functions. *Proc. of SPIE on Intelligent Robotic Systems*, Boston MA.

[4] Uras, C., 1992. Riconoscimento di forme con strategie metrico-topologiche. Tesi di Laurea in Fisica, Università di Genova.

[5] De Micheli, E., Caprile, B., Ottonello, P., and Torre, V., 1989. Localisation and noise in edge detection. *IEEE Trans. Patt. Anal. Mach. Intell.* **11**, 1106-1117.

[6] Uras, C. and Verri, A., 1992. In preparation.