Biol. Cybern. 70, 99–107 (1993)

Biological Cybernetics

On the use of size functions for shape analysis

A Verri [1], C. Uras [1], P. Frosini [2], M. Ferri [2]

[1] Dipartimento di Fisica dell'Università di Genova, Via Dodecaneso 33, I-16146 Genoa, Italy
[2] Dipartimento di Matematica dell'Università di Bologna, Bologna, Italy

Received: 9 March 1993/Accepted in revised form: 7 June 1993

Abstract. According to a recent mathematical theory a shape can be represented by *size functions*, which convey information on both the topological and metric properties of the viewed shape. In this paper the relevance of the theory of size functions to visual perception is investigated. An algorithm for the computation of the size functions is presented, and many theoretical properties of the theory are demonstrated on real images. It is shown that the representation of shape in terms of size functions (1) can be tailored to suit the invariance of the problem at hand and (2) is stable against small qualitative and quantitative changes of the viewed shape. A distance between size functions is used as a measure of similarity between the representations of two different shapes. The results obtained indicate that size functions are likely to be very useful for object recognition. In particular, they seem to be well suited for the recognition of natural and articulated objects.

1 Introduction

An intriguing property of the human visual system is the capability of recognizing objects independent of their apparent shape in images. The changes in the visual shape can be due to different factors. In the case of rigid and manufactured objects, for example, these changes are due to the object orientation and distance from the viewer. In the case of natural objects, these changes may also be due to the qualitative and quantitative differences between objects which belong to the same "category". Most of the techniques which have been proposed for shape analysis and object recognition appear to be appropriate for some particular and interesting cases, such as polyhedral rigid objects, planar curves, or character recognition, but do not seem to be sufficiently flexible to deal with the general problem.

In a recent series of mathematical papers, studying shape through integer-valued functions, called *size* functions (Frosini 1990, 1991, 1993), has been proposed. The new mathematical 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. Thus, size functions convey information about both the qualitative and quantitative structure of the viewed shape. The aim of this paper is to assess the potential of the theory of size functions in visual perception. An algorithm for the computation of size functions is presented, and the many theoretical properties of the size functions are checked and illustrated on real images. It is shown that the representation of shape through size functions can be tailored to suit the quantitative and qualitative invariant properties of the shape to be studied. Therefore, size functions seem to be suitable for the description and recognition of objects which have similar but not necessarily identical shapes (e.g., natural, articulated, and nonrigid objects).

In Sect. 2, of this paper the approach to shape description through the theory of size functions is introduced through a simple example. An algorithm for the computation of a size function is described in Sect. 3. A distance between the size function of two shapes is defined in Sect. 4. Section 5 deals with the invariant properties which may be incorporated in the theory and illustrates the tolerance of the scheme against changes of a different sort on real images. Section 6 discusses related work. Finally, Sect. 7 summarizes the main open problems, the directions of future research, and the results obtained. The key mathematical concepts and results which are relevant to this work are summarized in the Appendix.

2 Overview of the approach

In this section the concepts of the theory of size functions are illustrated for a simple example. Then, the shape representation which can be obtained from a size function is discussed. Finally, the main theoretical properties of the theory are listed.

2.1 Topological and metric obstructions

Figure 1a shows an example of shape, the letter "w" in sign language performed by one of the authors. By using

Correspondence to: A. Verri

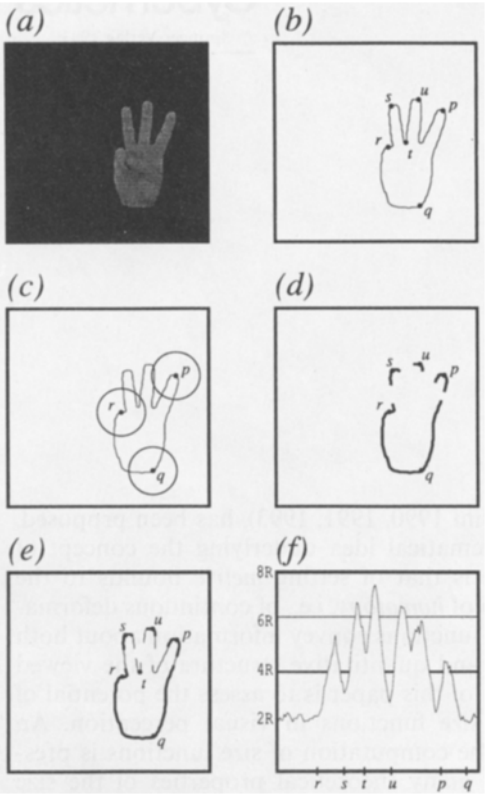

Fig. 1a–f. Topological and metric obstructions. **a** An image of the sign "w" performed by one of the authors. **b** The outline α of the sign of **a** obtained by means of a standard edge detection and contour-following techniques. **c** Computation of the measuring function L at the points p, q, and r. **d** The *thick edges* mark the points with $L \leqslant 4R$. **e** The *thin edges* mark the points with $4R \leqslant L \leqslant 6R$. **f** Plot of the measuring function L over the curve α

standard edge detection and contour-following techniqes, the contour α of Fig. 1b, which corresponds to the outline of the hand of Fig. 1a, can be easily obtained. Let us introduce the key concepts of measuring function and size function on the curve α.

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 transformation H induces an equivalence relation on the points of α, where two points v and w are said to be *H-equivalent* if there exists a continuous trajectory on α which brings v onto w, or if v and w belong to the same arcwise-connected component of α. For example, the points p, q, r, s, t, and u of Fig. 1b are all pairwise H-equivalent. Since, independent of the shape of α, all the points of α fall into one and the same equivalence class, the purely topological concept of H-equivalence is clearly not sufficient to characterize the shape of α. Intuitively, this reflects the absence of "topological obstructions" between the points of α.

In the theory of size functions, this problem is overcome by the notion of *measuring function* (Frosini 1990). The purpose of a measuring function is to generate "metric obstructions" for the transformation H. Let us illustrate the notion of measuring function through a particular example. For each point v of α, let $L = L(v)$ be the length of the portion of α which lies within the circle $c(v)$ of radius R and center v. Figure 1c shows how to compute L at the points p, q, and r. Let $R = D/5$, where D is the diameter of α. It is clear that $L(p)$, $L(q)$, and $L(r)$ can be computed as the sum of the length of the (possibly many) arcs of α which lie within the circles $c(p)$, $c(q)$ and $c(r)$ of Fig. 1c, respectively. The function L, which is defined on the contour α, is an example of measuring function.

Let us now modify the definition of H by means of L. Two points v and w of α are said to be $H(L \leqslant y)$-*equivalent* if $v = w$ or a trajectory exists on α from v to w along which L never exceeds y. Let us call a trajectory along which L never exceeds y and $(L \leqslant y)$-*trajectory*. Intuitively, the points of α with $L > y$ can be thought of as metric obstructions for the $(L \leqslant y)$-trajectories from v to w. This fact is illustrated in Fig. 1d where the points with $L > 4R$ (the metric obstructions) have not been drawn. From the gaps in Fig. 1d, it is easy to conclude that, between p, q, r, s, and u, the points q and r are the only pair of $H(L \leqslant 4R)$-equivalent points.

The notion of $H(L \leqslant y)$-equivalence is essential for the definition of size function (Frosini 1990). The size function $l_L(\alpha; x, y)$, for $x < y$, and x and $y \in \Re^2$, is defined as the number of equivalence classes in which the set of points with $L \leqslant x$ is divided by the $H(L \leqslant y)$-equivalence relation. Let us compute the size function l_L at the point (x, y) with $x = 4R$ and $y = 6R$. The set of points of α with $L \leqslant 4R$ are the thick edges of Fig. 1e. Thus, the size function $l_L(\alpha; 4R, 6R)$ is the number of equivalence classes in which the thick edges of Fig. 1e are divided by the $H(L \leqslant 6R)$-equivalence relation. This amounts to $(L \leqslant 6R)$-trajectories between all the possible pairs of thick edges. The thin edges of Fig. 1e, which are the points with $4R < L \leqslant 6R$, represent the "extra" space which has been made available to the $(L \leqslant 6R)$-trajectories.

It is easy to see that the size function $l_L(\alpha; 4R, 6R)$ equals the number of connected components of the curve of Fig. 1e (ignoring the difference between thick and thin edges) which contains at least one point with $L \leqslant 4R$, i.e., a thick edge. Since of the three connected components of the set of points with $L \leqslant 6R$, the one which contains the point t consists only of thin edges, it follows that $l_L = 2$. Note that the points p, q, r, and u which were not $H(L \leqslant 4)$-equivalent, now belong to the same equivalence class.

An equivalent representation of the connected components of the set of points with $L \leqslant 4R$ under the $(L \leqslant 6R)$-equivalence relation is shown in Fig. 1f, in which L is plotted against the curve $\alpha = \alpha(a)$ with $a \in [0, 1]$ and $\alpha(0) = \alpha(1)$. The thick horizontal line of Fig. 1f, makes it clear that the set of points with $L \leqslant 4R$ consists of four connected components [the leftmost and rightmost components belong to the same component because $\alpha(0) = \alpha(1)$]. The thin horizontal line shows that these components reduce to two when $L \leqslant 6R$.

Before discussing the main properties of the notion of size function, let us show how shape information is represented by means of a size function.

2.2 Shape representation

The size function $l_L = l_L(\alpha; x, y)$ is an integer-valued function of the two real variables x and y. Let us first show that all the relevant information is contained in a region of finite area of the plane (x, y). Let us divide the plane (x, y) in four regions (see Fig. 2a). Region A consists of all the points at the left of the vertical line $x = L'$, B of all the points with $L' < x \leqslant y$ and $y > L''$, C of all the points with $x > L'$ and $y < x$, and T of all the points of the triangle with $L' \leqslant y \leqslant L''$ and $L' \leqslant x \leqslant y$, where L' and L'' are the minimum and maximum of L over α, respectively. Let us now show that in A, B, and C the size function is independent of the shape of the curve α. First, since the set of points with $L < L'$ is the empty set, we have that, for all the points in A, $l_L = 0$. Then, since for $L > L''$ there are no metric obstructions, we have that, for all the points in B, $l_L = 1$. Finally, since for $y < x$ each point of α identifies a different equivalence class, for all the points in C, $l_L = +\infty$.

Thus, all the relevant information is contained in the triangular region T enclosed by the regions A, B, and C. The color-coded representation of the diagram of the size function of Fig. 2a within the triangle T makes it clear that the size function is piecewise constant, nondecreasing along the x-axis, and nonincreasing along the y-axis. These properties follow easily from the definition of size function.

The diagram of Fig. 2a was obtained by computing the size function at a finite number of locations of the triangular region T and sampling the curve α at a finite number of points. The relationship between the size function computed in the discrete and in the ideal continuous case is shown in Fig. 2b. A basic mathematical result (Frosini 1993) ensures that it is possible to divide the discrete estimates of Fig. 2a into two groups. The estimates which belong to the first group are exactly equal to the exact values of the continuous case. These estimates are reproduced unchanged in Fig. 2b. Since the size function is piecewise constant, a striking property of the graph of Fig. 2b is that it reproduces exactly the graph of the size function in the continuous case even at locations in which the discrete estimates were not computed.

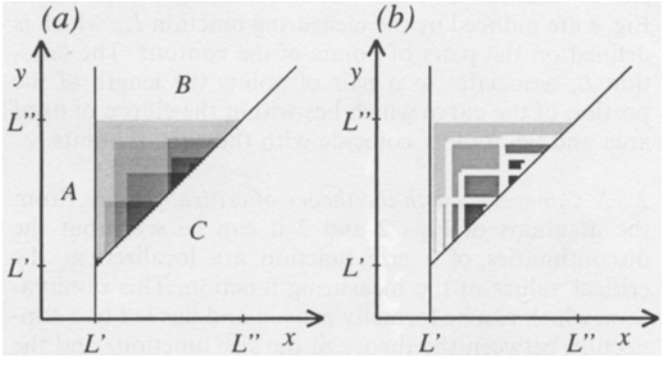

Fig. 2a, b. Color-coded representation of the size function l_L of Fig. 1b. **a** All the relevant information of the size function is contained in the triangular region enclosed by the regions A, B, and C. The color coding is *light gray* for 1, *gray* for 2, *darker gray* for 3, and *black* for 4. **b** The *white stripes*, which have been obtained by thickening the lines of discontinuity of the size function of **a**, represent the only locations in which the estimates of the size function in the discrete approximation do not necessarily coincide with the ideal continuous case

The white stripes within the triangular region T of Fig. 2b identify the locations of the estimates of the second group, in which the values of the size function computed in the discrete and continuous case do not necessarily coincide. The white stripes can be obtained by thickening the lines of discontinuity in the diagram of Fig. 2a of an amount which depends on the coarseness of the sampling and on the modulus of continuity of the measuring function. In the example of Fig. 2b it has been assumed that the maximum distance between the hypothetical continuous curve and the pixels of the contour is smaller than $2\sqrt{2}$ pixels.

Intuitively, the white stripes account for the possible differences between the size functions of two curves which happen to be sampled by the same set of points. Interestingly, these differences are localized in the neighborhood of the lines of discontinuity of the size function computed from the sampling points only. The value of the size function in the continuous case within a white stripe is bounded by the monotonicity constraints along the directions of the coordinate axes, but otherwise arbitrary.

Finally, let us briefly discuss how shape information is represented in a size function across scale. Shape information is not distributed evenly within the triangular region T. It can be seen that the details of the shape are represented in the region near the hypotenuse of the triangle T (where $y \simeq x$). The further from the hypotenuse, the coarser is the information on the viewed shape. This fact is better illustrated in Fig. 3. Figure 3a and b shows a pittosporum and an oak leaf, respectively. The corresponding outlines are shown in Fig. 3c and d. It is evident that at the coarse scales the two contours have a similar shape, whereas at the finer scales, the contour of the oak leaf becomes more "alive". Correspondingly, the size functions of Fig. 3e and f are similar away from the hypotenuse of the triangular region T, whereas the size function of the oak leaf increases near the hypotenuse (see Fig. 3f). The size functions in Fig. 3e and f are induced by the measuring function D_c, which measures the distance of a point of the contour from the center of mass of the contour.

2.3 Basic properties

The notions of measuring function and size function have a number of interesting properties. Let us briefly summarize these properties, some of which will be illustrated in greater detail in Sect. 5. In what follows, α denotes a generic curve of the image plane. First, there is a wide range of possible choices for a measuring function.

2.3.1 Admissible measuring functions. In principle, any continuous real function defined on a curve α can serve as a measuring function (see the Appendix). For example, along with the function L discussed in the

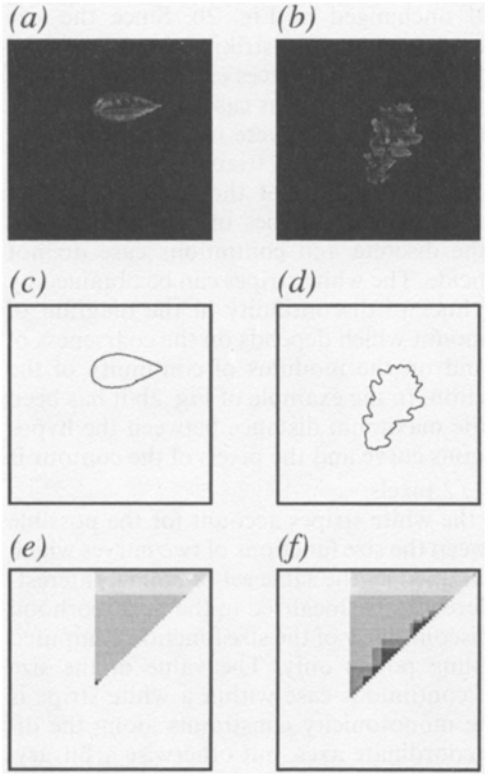

Fig. 3a–f. Shape information across scale. **a** Illustration of an image of a pittosporum leaf. **b** An image of an oak leaf. **c** and **d** Outlines of the leaves of **a** and **b** obtained through standard edge detection and contour-following techniques. **e** and **f** Color-coded representation of the size functions of the contours of **c** and **d** induced by the measuring function D_c, distance of a point of the outline from the center of mass. The diagrams are scaled between 0 and the maximum of D_c over each outline. The color coding is the same as that of Fig. 2

previous section, the curvature, the distance of a point of α from a certain point, such as the center of mass, or the y-coordinate of a point of α with respect to some system of reference are equally good choices of measuring functions. In addition, a measuring function need not be defined on single points of α. For example, a measuring function can be defined on the pairs of points of α. The only differences are the fact that the measuring function cannot be visualized any longer through a one-dimensional plot and that the notion of $H(L \leqslant y)$-*equivalence* must be redefined on $\alpha \times \alpha$, the cartesian product of the curve with itself (see the Appendix). Examples of measuring functions defined on $\alpha \times \alpha$ are the euclidean distance between a pair of points of α, and the ratio between this distance and the length of the shortest arc of α which joins the pair of points.

2.3.2 Inheritance of invariance properties. A fundamental property of the proposed scheme is that the representation of shape through a size function inherits the invariance properties (if any) of the underlying measuring function. These properties may include euclidean invariance (such as invariance for scaling, and translation and rotation over the image plane), or invariance for affine, projective, and perspective transformation. Clearly, the functions L and D_c of Sect. 2 are invariant for translation and rotation. The scaling invariance can always be obtained by scaling the maximum of the measuring function to a fixed value (which is the reason why the diagrams of Fig. 3e and f have similar support even if the oak leaf of Fig. 3b appears to be larger than the pittosporum leaf of Fig. 3a). In general, the invariance properties of the problem at hand may be used to constraint the search for the appropriate measuring function.

2.3.3 Tolerance to qualitative and quantitative changes. A further property of the representation of shape through size functions is robustness against quantitative and qualitative changes. This property derives from the facts that the proposed representation combines topological (i.e., qualitative) aspects of shape with metric (i.e., quantitative) aspects of shape in a redundant fashion, since a size function is piecewise constant. Intuitively, small qualitative and quantitative changes give rise to differences in the size functions over correspondingly small areas of the triangular region of interest. Consequently, the representation of shape in terms of size functions is likely to be suitable for the recognition of objects which are qualitatively and quantitatively similar but not identical.

2.3.4 Dynamics of size functions. A question which arises quite naturally in the attempt to assess the relevance of size functions to shape description and recognition is whether size functions have enough discriminating power to distinguish between similar (and simple) shapes. Let us discuss an example which indicates that size functions can actually be used for the description of simple shapes. Figure 4a, b, c, and d shows the color-coded size functions associated with two "a"s and "e"s, respectively of the sign language performed by two of the authors. The outline of the signs is shown on the lower right of each diagram. It is evident that the outlines of these four signs are much simpler than the outline of the "w" sign of Fig. 1b. Nevertheless, the dynamics of the computed size functions appears to be sufficient to distinguish between an "a" and an "e" sign. The size functions depicted in Fig. 4 are induced by the measuring function L_e, which is defined on the pairs of points of the contour. The function L_e associates to a pair of points the length of the portion of the curve which lies within the ellipse of fixed area and whose foci coincide with the pair of points.

2.3.5 Connection with the theory of critical points. From the diagrams of Figs 2 and 3 it can be seen that the discontinuities of a size function are localized at the critical values of the measuring function. This observation, which can be formally proved and has led to a connection between the theory of the size functions and the theory of critical points (Frosini 1990), provides a different interpretation for the presence of thick white stripes in Fig. 2b. The thickness of a stripe reflects the uncertainty in the localization of the critical points of the measuring function from a discrete sampling of the original curve. Furthermore, and most importantly, the critical values of the measuring function completely determine

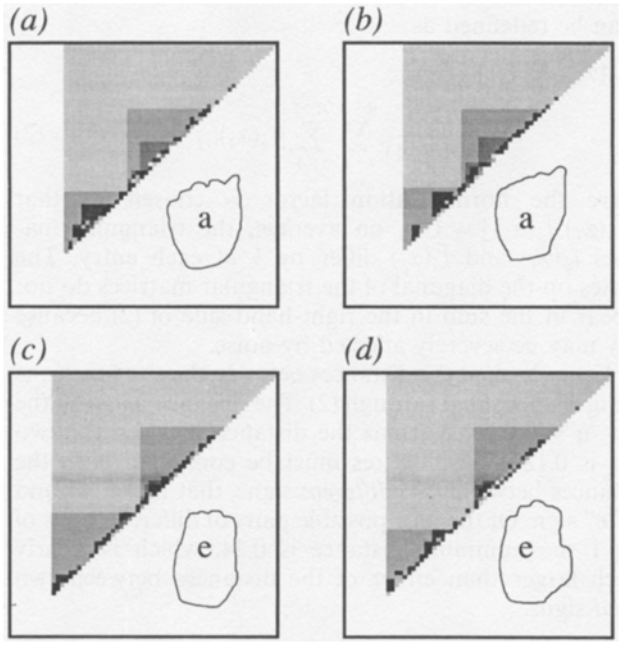

Fig. 4a–d. Color-coded representations of the size functions of the signs whose outlines are displayed in the lower right corner of each diagram. The size functions are induced by the measuring function L_e (see text). The diagrams are scaled between 0 and the maximum of L_e over each outline. The color coding is the same of Fig. 2

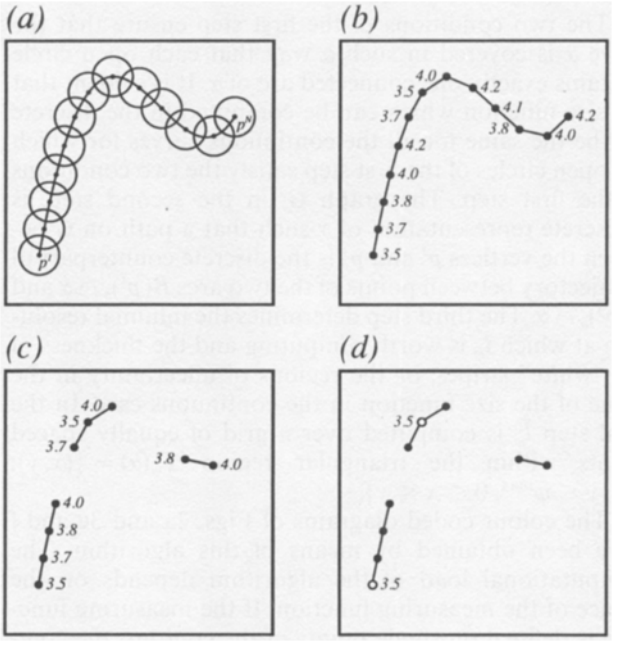

Fig. 5a–d. The algorithm for the discrete computation of a size function. **a** Curve sampling and covering. **b** The graph associated with the sampled curve. The numbers, which correspond to hypothetical values of the measuring function φ at each sampled point, are associated with the corresponding vertex. **c** Subgraph of the graph of **b** induced by the set of vertices with $\varphi \leqslant 4.0$. **d** The vertices of the subgraph of **c** with $\varphi \leqslant 3.6$ are shown as *open circles*. Therefore, the value of the size function of the sampled curve of **a** at the point (3.6, 4) equals 2

the structure of the induced size function. This fact has two important consequences. First, if the sequence of critical values of a measuring function over two different shapes is the same, the size function of the two shapes is the same. Second, if the critical values of a measuring function over two shapes are different, then the induced size functions are different. The implications of these facts for object recognition are evident. Two objects may look similar with respect to a certain size function, i.e., to some aspect of their shape, but then prove to be different with respect to another size function, i.e., to some other aspect of their shape.

2.3.6 Higher-dimensional shapes. Let us conclude this section with a theoretical remark. The theory of size functions is not restricted to the case of curves. In principle (see the Appendix) the shape of a surface of arbitrary dimension can be represented through a size function. For the sake of simplicity, this paper and the present research have been restricted to the analysis of curves on the image plane. Extensions to the two-dimensional case are currently under investigation. The surface to be studied could be the region of the image plane enclosed by a contour (in both binary and gray-valued images) or a "thick" contour (as in character recognition).

3 An algorithm for the computation of size functions

This section describes the implementation of an algorithm for the discrete computation of a size function of a curve of the image plane.

For the sake of simplicity, let us illustrate the implementation in the particular case in which the measuring function φ is defined on single points of a curve α (that is, $k = 1$) with $\varphi \geqslant 0$. 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 continuous and discrete cases, respectively. The algorithm consists of four steps:

1. Sample (or approximate) the curve α at a finite number N of points p^i, $i = 1, \ldots, N$ so that $\alpha \subset \bigcup_{i=1}^{N} B(p^i)_\delta$, and the set $B(p^i)_\delta \cap \alpha$ is nonempty and connected for $i = 1, \ldots, N$ (see Fig. 5a).
2. Define the graph G whose vertices are the points p^i and whose edges link vertices which correspond to adjacent points on α. Compute $\varphi(p^i)$ at each point p^i, $i = 1, \ldots, N$ (see Fig. 5b).
3. Compute the maximum $\varphi^{\max}$ of $\varphi(p^i)$, $i = 1, \ldots, N$ and fix a $\Delta \geqslant \varepsilon_\varphi(\delta)$, where $\varepsilon_\varphi(\delta)$ is the modulus of continuity of φ at δ.
4. For $y = 0$ to $y \leqslant \varphi^{\max}$
   a. Define the subgraph $G_{\varphi \leqslant y}$ of G induced by the set of vertices of G for which $\varphi \leqslant y$ (Fig. 5c).
   b. For $x = 0$ to $x \leqslant y$
      i. Let $\bar{l}_\varphi(\alpha; x, y)$ be the number of connected components of $G_{\varphi \leqslant y}$ which contain at least a vertex p^i such that $\varphi(p^i) \leqslant x$ (Fig. 5d).
      ii. $x \rightarrow x + \Delta$.
   c. $y \rightarrow y + \Delta$.

The two conditions 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. It is evident that the size function which can be computed in the discrete will be the same for all the continuous curves for which the open circles of the first step satisfy the two conditions of the first step. 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 $\bar{l}_\varphi$ is worth computing and the thickness of the "white" stripes, or the regions of uncertainty in the value of the size function in the continuous case. In the final step $\bar{l}_\varphi$ is computed over a grid of equally spaced points within the triangular region $T_\varphi(\alpha) = \{(x, y): 0 \leqslant y \leqslant \varphi^{\max}, 0 \leqslant x \leqslant y\}$.

The colour coded diagrams of Figs. 2a and 3e and f have been obtained by means of this algorithm. The computational load of the algorithm depends on the choice of the measuring function. If the measuring function is defined on single points of the contour, the computation takes less than a second on an SPARC workstation. The computational time may go up to several seconds for a measuring function which is defined on pairs of points of the contour.

4 A distance between size functions

In order to determine quantitatively whether similar shapes are given a similar representation and different shapes are actually distinguishable, a distance between size functions needs to be defined. There are many ways in which a distance between size functions can be defined. Probably the only common requirement to the possible definitions is that the scale-invariant property must be preserved (i.e., the distance between the size functions of the same shape at different scales must always vanish). Let us introduce the simple distance function which will be used throughout the rest of the paper.

Let φ be a measuring function, α_1 and α_2 two curves, and $\varphi^{\max}(\alpha_i)$ the maximum of φ on α_i, for $i = 1, 2$. Without loss of generality it can be assumed that $\varphi \geqslant 0$. Let us scale φ by defining $\hat{\varphi} = \varphi/\varphi^{\max}(\alpha_i)$ on α_i, for $i = 1, 2$. Then, a scale-invariant distance d between the size functions $\bar{l}_{\hat{\varphi}}(\alpha_1)$ and $l_{\hat{\varphi}}(\alpha_2)$ can be defined as (Uras and Verri 1992)

$$d[l_{\hat{\varphi}}(\alpha_1), l_{\hat{\varphi}}(\alpha_2)] = 2 \int_0^1 dy \int_0^y dx \, |l_{\hat{\varphi}}(\alpha_1; x, y) - l_{\hat{\varphi}}(\alpha_2; x, y)| \tag{1}$$

The distance d is simply the L^1 norm of the difference over the triangular region with $0 \leqslant x \leqslant y$ and $0 \leqslant y \leqslant 1$. Similarly, in the discrete case, if $\bar{l}_{\hat{\varphi}}(\alpha_1)$ and $\bar{l}_{\hat{\varphi}}(\alpha_2)$ are 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 $i = 1, \ldots, R-1$ and $j = 1, \ldots, R-i$, then the distance d can be redefined as

$$d[\bar{l}_{\hat{\varphi}}(\alpha_1), \bar{l}_{\hat{\varphi}}(\alpha_2)] = \frac{2}{R(R-1)} \sum_{i=1}^{R-1} \sum_{j=1}^{R-i} |\bar{l}_{\hat{\varphi}}(\alpha_1)_{i,j} - \bar{l}_{\hat{\varphi}}(\alpha_2)_{i,j}| \tag{2}$$

where the normalization factor is chosen so that $d[\bar{l}_{\hat{\varphi}}(\alpha_1), \bar{l}_{\hat{\varphi}}(\alpha_2)] = 1$ 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 do not appear in the sum in the right-hand side of (2) because they may be severely affected by noise.

Let us look at the distances between the size functions of Fig. 4 computed through (2). The distance between the two "a"s is 0.15, whereas the distance between the two "e"s is 0.18. These figures must be compared with the distances between two *different* signs, that is, an "a" and an "e" sign. Of the four possible pairs of different signs of Fig. 1 the minimum distance is 0.34, which is clearly much larger than either of the distances between two *equal* signs.

5 Invariant properties

In this section the invariance properties of the size functions mentioned in Sect. 2 are demonstrated on real images.

5.1 Euclidean invariance

Let us now make some quantitative estimates of the invariance of the representation of shape which can be obtained through a size function. Figure 6a shows the image of an ivy leaf. The size function induced by the measuring function D_c, that is, the distance of a point of the outline from the center of mass of the outline, is shown in Fig. 6b. In principle, the function D_c is clearly invariant for translation and rotation of the shape over the image plane. Figure 6c shows an image of the same leaf translated and rotated while the camera was kept in a fixed position. The size function associated with the outline of the image of Fig. 6c is shown in Fig. 6d. By visual inspection of Fig. 6b and d it can easily be seen that the size functions of the leaves of Fig. 6a and c are very similar. Correspondingly, the distance d between the size functions of Fig. 6b and d, computed by means of (2) is equal to 0.06.

The property of scale invariance is illustrated in Fig. 6e and f. In the image of Fig. 6e, the camera was viewing the same ivy leaf from a further viewpoint. It is evident that the size function of Fig. 6f, which was obtained from the outline of the image of Fig. 6e, is very similar to the size function of Fig. 6b. In this case, the distance between the size functions of Fig. 6b and f, computed by means of (2) is 0.04.

Table 1 summarizes the results obtained in a series of similar experiments. Several images of three leaves of different species (ivy, oak and oleander) were taken by translating and rotating the leaves on the supporting plane and by varying the distance between the viewing

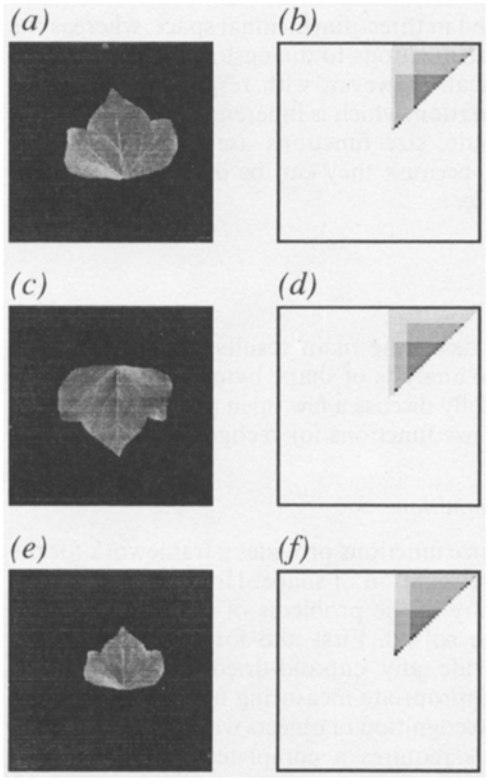

Fig. 6a–f. Invariance for euclidean transformations. **a** An image of an ivy leaf. **b** The size function of the outline of the leaf of **a** induced by the measuring function D_c. **c** An image of the same ivy leaf of **a** translated and rotated over the supporting plane while the camera was kept in a fixed position. **d** The size function of the outline of the leaf of **c**. **e** An image of the same ivy leaf of **a** from a further viewpoint. **f** The size function of the outline of the leaf of **e**. The color coding is as in Fig. 2, and the diagrams are scaled between 0 and the maximum of D_c over each outline

camera and the supporting plane. The euclidean invariance of the computed size functions was checked by computing the distance between the size function of the leaves in the original position and after the translational, rotational, and scaling transformations were performed. As it can be easily inferred from Table 1, the euclidean invariance is always well satisfied and independent of the translational, rotational, and scaling parameters.

5.2 "Ad hoc" invariance

Articulated objects modify their shape according to the changes of some internal parameter. Figure 7a and b shows the same pair of scissors with different openings. For recognition purposes, it would be desirable to be able to represent the shapes of articulated objects independent of their internal parameters. In the present framework, this problem can be solved by looking for an appropriate measuring function. For example, the measuring function D_p, that is, the distance from the pivot, is invariant for different openings (actually, due to self-occlusions, D_p can only be approximately invariant). The size functions induced by D_p and associated with the contours of Fig. 7c and d are shown in Fig. 7e and f,

Table 1. Euclidean invariance

			Leaf		
T	Ω	s	1	2	3
10	0	1.0	0.02	0.03	0.01
15	0	1.0	0.03	0.06	0.00
20	0	1.0	0.05	0.01	0.04
0	30	1.0	0.07	0.04	0.03
0	45	1.0	0.03	0.04	0.06
0	70	1.0	0.03	0.02	0.02
0	0	0.5	0.02	0.05	0.01
0	0	1.5	0.06	0.03	0.00
0	0	2.0	0.06	0.07	0.02
10	30	2.0	0.04	0.01	0.05
15	45	1.5	0.03	0.05	0.02
20	70	0.5	0.05	0.04	0.03

In order to test the invariance of the size function induced by the measuring function "distance from the center of mass" with respect to euclidean transformations, images of an ivy leaf (leaf 1), oak leaf (leaf 2), and oleander leaf (leaf 3) were taken by translating and rotating the leaves on the supporting plane of T centimeters and Ω degrees respectively. Different scaling factors s were obtained by changing the distance between the viewing camera and the supporting plane. The columns leaf 1, leaf 2, and leaf 3 report the distance d between the size function of the ivy, oak, and oleander leaves in the original position and after the translational, rotational, and scaling transformations were performed, respectively

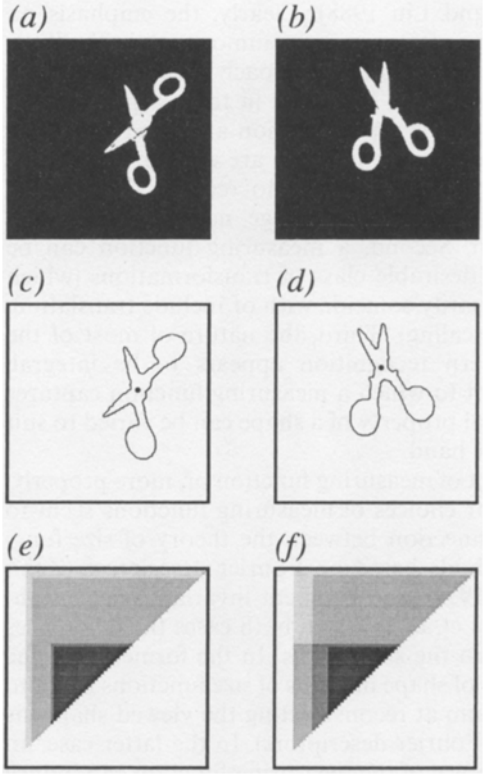

Fig. 7a–f. Invariance ad hoc. **a, b** Two images of the same pair of scissors with different openings. **c, d** Outlines of the shapes of **a** and **b** obtained by means of the same procedures of Fig. 1. **e, f** Color-coded representations of the size functions of **c** and **d** induced by the measuring function D_p, distance of a point of the outline from the pivot. In both **c** and **d**, the pivot was located as the midpoint of the segment whose endpoints are the intersections of the principal inertial axis with the outline. The color coding is as in Fig. 2, and the diagrams are scaled between 0 and the maximum of D_p over each outline

respectively. By using (2) the distance between the size functions of Fig. 7e and f is found to be equal to 0.15. This distance should be compared with the much larger distance ($d = 0.7$) between the size functions of Fig. 7a and b obtained by means of the measuring function distance from the center of mass (which is not invariant for different openings).

6 Related work

In this section similarities and differences between the theory of the size functions and a few classical techniques of shape analysis are briefly discussed.

Invariant pattern recognition studies the representation and recognition of shapes independent of a number of transformations of the viewed shape, translation, rotation, and scale. Methods invariant for rotation (Hsh et al. 1982), rotation and scale (Massone et al. 1985; Caelli and Nagendran 1987), and translation, rotation, and scale (Casasent and Psaltis 1976; Caelli and Liu 1988; Pintsov 1989), for example, can be based on integral transforms [such as Fourier (Hsh et al. 1982), Radon (Pintsov 1989), and Fourier–Mellin transform (Casasent and Psaltis 1976)], log-polar transform (Massone et al. 1985; Caelli and Nagendran 1987), or combination of templates as filters (Caelli and Liu 1988). Clearly, the emphasis on invariant representations is common to both these methods and the presented approach. The major differences can probably be identified in three aspects. First, the invariance of the representation and the representation itself through a size function are weaker. In general, for example, it is not possible to recover the original orientation of a shape in an image, nor the shape, from a size function. Second, a measuring function can be invariant to a desirable class of transformations (which does not necessarily coincide with or include translation, rotation, and scaling). Third, the nature of most of the invariant pattern recognition appears to be integral, while the extent to which a measuring function captures a local or global property of a shape can be varied to suit the problem at hand.

The concept of measuring function or, more properly, some particular choices of measuring functions seem to establish a connection between the theory of size functions and methods based on Fourier descriptors (Zahn and Roskies 1972) and moment-invariant-based techniques (Dudani et al. 1977). In both cases the differences are greater than the similarities. In the former case, the representation of shape in terms of size functions is not to and does not aim at reconstructing the viewed shape (as in the case of Fourier descriptors). In the latter case, an appropriate choice of the measuring function can induce a representation of shape which is based on local properties (and not globally, as in the case of moment-based techniques).

The theory of aspect graphs (Koenderink and Van Doorn 1976) is probably the closer scheme to the theory of the size functions, in the sense that both the schemes describe quantitatively qualitative aspects of shape. An advantage of the concept of aspect graph is that it is naturally defined in three-dimensional space, whereas the relevance of size functions to dimension higher than one is still theoretical. However, with respect to the aspect graph representation, which is inherently continuous and hard to compute, size functions are better suited for digital images, because they can be defined formally in the discrete case.

7 Discussion

Before summarizing the main results which have been obtained in the analysis of shape by means of size functions, let us briefly discuss a few open problems concerning the use of size functions for recognition.

7.1 Object recognition

The theory of size functions provides a framework for the study and representation of shape. However, within this framework many of the problems of object recognition have still to be solved. First and foremost, the theory does not provide any cut-and-dried method for the choice of the appropriate measuring functions. It is very likely that the recognition of objects which belong to very different classes requires a completely different set of measuring functions. At the moment the choice of the appropriate measuring functions is largely guided by empirical principles.

Second, a single measuring function is not sufficient to characterize an object completely. Therefore, since more than one measuring function may be needed, the relationship between two different measuring functions of the same shape and ways to combine the information contained in the induced size functions has to be studied. A possible scheme for object recognition based on size functions, which uses leaves and signs of the sign language as study cases, has been described elsewhere (Verri et al. 1993).

Finally, probably the most important constraint on the choice of the measuring functions comes from the fact that in many images the shape to be analyzed is partially occluded. It is evident that even if size functions do not need to be computed on a closed and connected contour, size functions of a partially occluded contour and of the entire contour will, in general, be different. Intuitively, the measuring functions which are defined locally appear to be more suitable to deal with occlusions. Some preliminary theoretical results indicate a few ways in which the problem of occlusions can be handled within the framework presented, but more theoretical and experimental work needs to be done in this direction.

7.2 Conclusions

In this paper the potential of the theory of size functions to visual perception has been assessed. An algorithm for the computation of size functions from real images has been implemented and used to illustrate a number of theoretical properties of the theory which are likely to be useful for object recognition. Based on the experimental

results presented it can be concluded that the representation of shape in terms of size functions (1) can be tailored to suit the invariance of the problem at hand and (2) is stable against small qualitative and quantitative changes of the viewed shape. In addition a size function can be designed to highlight a particular aspect of the shape of an object, an aspect which can be useful to build similar representations of shape which are similar but quantitatively or qualitatively different. A distance between size functions has been introduced to measure the similarity between the representation of two different shapes. The results obtained indicate that size functions are likely to be very useful for the recognition of objects which have similar but not identical shapes.

Acknowledgements. We wish to thank Steve Omohundro, Jerry Feldman, Jitendra Malik, and Brian Subirana for useful discussions. Terry Caelli and Mario Ferraro pointed us to the literature related to the subject of the paper. Bobby Rao read the manuscript. A.V. was with the International Computer Science Institute, Berkeley, Calif. C.U. was supported by a fellowship from ELSAG-BAILEY S.p.A. This research has been partially supported by the ESPRIT B.R.A. project VIVA and by a grant from the A.S.I. The images of Figs. 1a and 4a and b were taken at the Vision Laboratory of the University of California at Berkeley with the help of Joe Weber.

Appendix

In this Appendix the main concepts of the theory of size functions are formally defined, and two basic theorems are stated without proof. The proof of the theorems can be found in Frosini (1990, 1993).

Let us first establish some basic notation. In what follows, a shape is an n-dimensional, compact, boundaryless, piecewise C^∞ submanifold $\mathcal{M}$ of the euclidean space $E^m (n < m)$. The set of ordered k-tuples p of points p_i of $\mathcal{M}$, $i = 1, \ldots, k$, 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 $d_k(p, q) = \max_{1 \leqslant i \leqslant k} \{d(p_i, q_i)\}$ 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* is any continuous function

$$\varphi : \mathcal{M}^k \to \mathfrak{R}$$

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

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

- $H(0) = p, H(1) = q$
- $\varphi[H(\tau)] \leqslant y \; \forall \tau \in [0, 1]$

We write $p \simeq_{\varphi \leqslant y} q$, if such a metric homotopy exists, or if $p = q$. Now let $\mathcal{M}^k(\varphi \leqslant x)$ be the set of points p in $\mathcal{M}^k$ with $\varphi(p) \leqslant x$ (e.g., the set of points on the continuous lines of Fig. 1d). We have the following definition of size function.

The size function $l_\varphi(\mathcal{M}) : \mathfrak{R}^2 \to N \cup \{+\infty\}$ can be defined as

$$(x, y) \subsetneq \begin{cases} \#(\mathcal{M}^k(\varphi \leqslant x)/\simeq_{\varphi \leqslant y}) \text{ if finite} \\ +\infty \text{ otherwise} \end{cases}$$

Let $\varphi^{\max}$ and $\varphi^{\min}$ be the maximum and minimum of the measuring function φ over $\mathcal{M}^k$ and $T_\varphi(\mathcal{M}) = \{(x, y) : \varphi^{\min} \leqslant y \leqslant \varphi^{\max}, \quad \leqslant \varphi^{\min} \leqslant x \leqslant y\}$. A basic theorem of the theory of size functions ensures that the value of the size function $l_\varphi(\mathcal{M})$ within the triangular region $T_\varphi(\mathcal{M})$ and for many well-behaved φs is always finite and strictly positive.

Let us now consider the point $(\bar{x}, \bar{y})$ within the triangular region $T_\varphi(\mathcal{M})$. Let b, $c \geqslant 0$ and $\bar{x} + \varepsilon_\varphi(\delta) \leqslant \bar{y} - \varepsilon_\varphi(\delta)$, where $\varepsilon_\varphi(\delta)$ is the modulus of continuity of φ at δ. A second theorem ensures that if the size function computed in the discrete at the points $[\bar{x} + \varepsilon_\varphi(\delta), \bar{y} - \varepsilon_\varphi(\delta),]$ and $[(x - b - \varepsilon_\varphi(\delta), \bar{y} + c + \varepsilon_\varphi(\delta)]$ takes on the same value n, then the size function of the underlying continuous curve equals n in the rectangle $\{(x, y) : \bar{x} - b \leqslant x \leqslant \bar{x}, \bar{y} \leqslant y \leqslant \bar{y} + c\}$.

References

Caelli TM, Liu Z (1988) On the minimum number of templates required for shift, rotation, and size invariant pattern recognition. Pattern Recognition 21:205–216

Caelli TM, Nagendran S (1987) Fast edge-only matching techniques for robot pattern recognition. Comput Vision Graphics Image Proc 39:131–143

Casasent D, Psaltis D (1976) Position, rotation, and scale invariant optical correlation. Appl Opt 15:1795–1799

Dudani SA, Breeding KF, McGhee RB (1977) Aircraft identification by moment invariants. IEEE Trans Comput. 26:39–45

Frosini P (1990) A distance for similarity classes of submanifolds of a Euclidean space. Bull Austral Math Soc 42:407–416

Frosini P (1991) Measuring shapes by size functions. In: Proceedings of the SPIE on intelligent robotic systems, Boston, Mas, pp

Frosini P (1993) Discrete computation of size functions. J Combinatorics Inf Syst Sci (in press)

Hsh Y-N, Arsenault HH, April G (1982) Rotation-invariant digital pattern recognition using circular harmonic expansion. Appl Opt 21:4012–4015

Koenderink JJ, Van Doorn AJ (1976) The internal representation of solid shape with respect to vision. Biol Cybern 24:51–59

Massone L, Sandini G, Tagliasco V (1985) Form invariant topological mapping strategy for 2D shape recognition. Comput Vision Graphics Image Proc 30:169–188

Pintsov DA (1989) Invariant pattern recognition, symmetry, and Radon transforms. J Opt Soc Am [A] 6:1544–1554

Uras C, Verri A (1992) Studying shape through size functions. In: Proceedings of the NATO workshop on shape in picture, Driebergen NL, pp

Verri A, Uras C, Frosini P, Ferri M (1993) On the use of size function for shape analysis. In: Proceedings of the IEEE workshop on qualitative vision, New York, pp 89–96

Zahn CT, Roskies RZ (1972) Fourier descriptors for plane closed curves. IEEE Trans Comput 21:269–281