Point selection: A new comparison scheme for size functions (With an application to monogram recognition)

Massimo Ferri, Patrizio Frosini, Alberto Lovato and Chiara Zambelli

Dip. di Matematica, Università
Piazza Porta S.Donato, 5
I-40127 Bologna ITALY
E-mail: ferri, frosini@dm.unibo.it

Abstract. A new paradigm for the comparison of size functions is presented; it stresses the relevance of the "angular points" of the functions, and gives greater value to the stable ones. A simple example of classification of monograms (88 elements in the training set, 88 in the test set, for 22 classes, with a hit rate of 78%) is given, and a current enhancement of the experiment is described.

1 Introduction

Size function theory (see next Section) is slowly affirming its validity as a mathematical tool for representing and comparing shapes. Its advantages are the capability of formalizing qualitative concepts of shape, and the standard form assumed by size functions in spite of the different criteria adopted by the user. An intrinsic difficulty is the choice of a distance, or of a similarity measure between size functions; another one is the recognition of the relevant features of size functions of a training set.

The present work tries to solve both problems, by defining a "training matrix", a "test matrix", and a particular operation on them. Call *training* (respectively *test*) *size function* the size function of an element of the training (resp. test) set. Through the aforementioned matrices, only those parts — of a test size function — are taken into account, which match a stable feature of the training size functions.

A simple example, which applies this paradigm, completes the paper: classification of monograms.

2 Size functions

Size functions are modular shape descriptors. For extensive references on the subject, see [8, 14]. Here we just recall the main concepts.

Shape — in our opinion — is not simply a quality of a set of points in space or in a plane. It is rather a property of such a set, together with a real function defined on it; e.g, "bumpiness" is the behaviour of the function "distance from the center of mass" on the points of the considered object, "coarseness" is the behaviour of "curvature". We call a real function $\varphi : \mathcal{M} \to \mathbf{R}$ defined on a subset $\mathcal{M}$ of a Euclidean space a *measuring function*.

The *size function* of the pair $(\mathcal{M}, \varphi)$ is a function $\ell_{\mathcal{M}} : \mathbf{R^2} \to \mathbf{N} \cup \{\infty\}$. For each pair $(u, v) \in \mathbf{R^2}$, consider the set of points on which φ is worth $\leq u$. Two such points are then considered to be equivalent if they either coincide, or can be connected in $\mathcal{M}$ by a path, on whose points φ is worth $\leq v$. Then $\ell_{\mathcal{M}}(u, v)$ counts the equivalence classes so obtained.

For an example of the meaning of a size function, see Figure 1a, where a plane curve $\mathcal{M}$ is depicted; φ is taken to be the distance from point B. Fig. 1b represents the corresponding size function: The value $\ell_{\mathcal{M}}(u, v)$ informs us on the number of classes of points on the curve, which have distance $\leq u$ from B and can be joined together by walking on the curve without exceeding distance v. In other words, it is the number of those maximal arcs of $\mathcal{M}$ within distance v from B, which contain at least one point not farther than u from the same point.

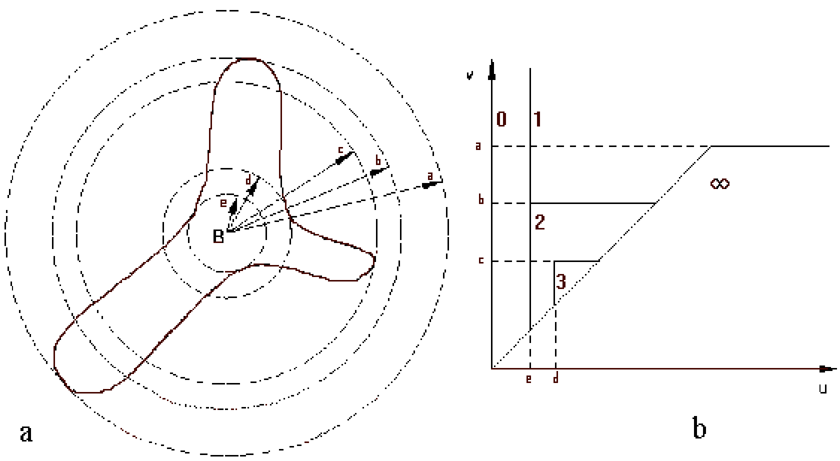

Fig. 1

The size function of a curve (or more generally, of an image) can be thought of as a sort of transform, which stresses only the aspect selected while choosing the measuring function. Modularity is assured by this freedom of choice. Invariance of the measuring function under a transformation implies invariance of the corresponding size function under the same transformation [12]. Of course, in practical cases only a discrete set of points is given (pixels of an image or, in the application we are going to report, sample points of a contour). For technical reference on the computation of size functions, see [6, 7, 13].

Another advantage of modularity is the possibility to use different "viewpoints" (i.e. measuring functions) at the same time; we make them cooperate by getting a fuzzy characteristic function from each, and then taking the average. The measuring functions can be different occurrences of the same function type — as will be the case for the example at the end of this paper — or be totally unrelated, so bringing in different aspects and classification criteria — as we are presently trying on the same problem.

3 Angular points

Whatever the measuring function, the outcoming size functions always have the same "format": it is determined by a finite set of points of $\mathbf{R}^2$ (called *angular points*). They are those intersection points of the discontinuity lines of the size function, at which the vertical value difference of the size function for slightly greater u is bigger than the one for slightly lesser u.

In the example of Fig. 1b, we have (a, d) and (b, c) as angular points. These convey all relevant information about the size function. Note that there are other interesting points. One is (a, ∞); this corresponds with the minimum value of φ; in general, such points correspond to as many connected components of $\mathcal{M}$. Another interesting point is (e, e), corresponding to the maximum value of φ. Anyway, we concentrate on angular points and their multiplicity, i.e. the difference of the vertical differences on the right and on the left of the point.

4 Point selection

A distance between size functions may be defined as a point–by–point difference (in absolute value) of the functions themselves. We used a slightly modified version of this with a fair success in other projects. A progress has been the use of a Hausdorff distance between the sets of angular points of two size functions (see [2]).

In the present work, we concentrate on the fact that angular points don't all have — so to say — the same relevance for the recognition process. We realized this first when dealing with occlusion. Size function are, by their very nature, global invariants. All the same, by comparing the size functions of differently occluded images of the same object, we noted the constant presence of a particular set of angular points. We observed the same phenomenon in those classification problems where the objects are stricltly referrable to a prototype (leaves of the same tree [14], on–line handwritten letters [4], signatures): Some contingent features, e.g. the "tail" of the signature of the same writer, may differ strongly from case to case, so giving angular points in very different positions. Other features don't vary that much, and the corresponding angular points are very stable. This may not be the case with other classification problems, where the objects can be referred to a common description, but not actually to a prototype (e.g., leukocytes [5], free–hand drawings [2], and also the sign language [11]).

Preprocessing of the object may work when the varying feature is cospicuous as a signature tail (even forcing a research group to erase tails before classifying [1]), but variations may be more deeply hidden, and difficult to get rid of. So we have preferred to act on the size functions, instead.

We divide the (u, v)–plane into 32×32 cells. Now, take a fixed measuring function, and the size functions of the members of the training set, belonging to a fixed class. We define the 32×32 *training matrix* of the class. Each angular point of each size function gives a contribution to the matrix, built by a Gaussian mask centered on the cell to which the angular point belongs; the mask has sum one, and a variance which increases with the distance from the diagonal $u = v$; the values are also multiplied by the multiplicity of the angular point. The final training matrix (r^i_j) is the average of the ones coming from the various objects of the training set, belonging to that class. Of course, cells below $u = v$ have null entry, and are inessential. Angular points in the cells just above this diagonal are not taken into account, since they are too much affected by discretization noise.

For each single element of the test set, we build a completely analogous *test matrix* (s^i_j), with the only difference that angular points falling in the leftmost column or in the topmost row are not taken into account.

Given a test matrix, we now compute the likelihood of the corresponding object to belong to the given class. First, we sum

$$p = \sum_{i,j} min\{r^i_j, s^i_j\}.$$

The likelihood is then

$$\frac{p}{\sum_{i,j}(r^i_j + s^i_j) - p}.$$

An entry of the training matrix will be high, if several size functions in the training set have an angular point in the corresponding cell. If the size function of the test set has an angular point in that cell too, then this will give a good contribution to p (less so if it is in a nearby cell). So, the numerator of the fraction evaluates the correspondence of the angular points of the test size functions with the stable ones of the training set.

As for the meaning of the denominator, imagine that the training set is built by just one element, and that each angular point has multiplicity one, and gives a sharp contribution of 1 to its cell (with no Gaussian smoothing); then the denominator would just count the total number of cells occupied by the union of angular points of the test and training size functions.

5 An application: Monogram recognition

The problem of monogram recognition is a nice one for testing our new paradigm, for the following reasons:

– Monograms are "natural" objects, so well suited for analysis by size functions.

– All the same, the elements of the training set can be seen as sort of true "prototypes".
– Variability is sufficient to make superposition methods ineffective.
– Context cannot help recognition.

We have gathered four elements for the training set and four for the test set, from each of 22 subjects. Acquisition was by a hand–held scanner and a PC, from black ink monograms. The only preprocessing is dilation. The program (In Visual Basic and C) extracts the outline (i.e. the outer contour) of the monogram under study; the outline will be the only input to the very classification process. This choice makes us loose much information, as can be seen in Figures 2a and 2b, where the left colums show the monograms of the training sets of "cri" and of "dan" respectively, and the right columns present the relative outlines. This choice was supported by previous experiences on tools [9] and on signatures [10]. A connected curve is a good test, because the analysis of angular points becomes simpler, and this makes it easier for us to understand what has happened. A further, deeper investigation presently under study involves the entire monogram. (More about this later in this section).

Simplification was applied also in the choice of measuring functions. Unlike in other researches, where we strived to differentiate measuring functions, we have adopted five versions of the same one, i.e. distances from points [3]. The five points are the barycentre of the monogram and the middle points of the edges of the horizontal rectangle circumscribed to the monogram. Again, this choice was meant to simplify the analysis of the experiment.

Table 1 reports the results. The first four numerical columns show the placing of the four elements of the test set; e.g., the figure 3 reported in the first column of "dari" means that the first monogram of the test set of this subject has not been recognized, and that there are two other subjects to whom the program would rather ascribe the monogram as more likely. The last column resumes the results as a fraction of hits over total.

We are particularly surprised — and deceived — by the totally negative performance on "stef", whose monograms look easy to recognize to a human eye: see Figure 3a (training left, test right). Conversely, a fairly good 3/4 was reached both by the monograms of "lup" and "monz", which look rather similar to each other (Figures 3b and 3c respectively). The overall hit ratio is 78%.

Since this classification problem turns out to be of interest in itself, we are implementing — at the moment of revising the paper — a more complete analysis which takes the whole monogram into account. This is done by computing a sort of Radó transform of the monogram: For each of four directions we scan the monogram by parallel and adjacent straight lines. For each line we sum the number of black pixel present in a five pixel wide strip. After suitable normalization, this number is assigned to a vertex of a linear graph (otherwise said, to a pixel of an incident segment). Then the size function is computed; comparison is carried out as in the outline case. The two "orthogonal" classifications yield likelyhoods, the mean of which will be the final one.

Subject	# 1	# 2	# 3	# 4	Hit/Tot
amm	1	1	2	1	3/4
clau	1	1	1	1	4/4
cri	1	1	1	1	4/4
cris	1	1	1	2	3/4
dan	1	1	1	1	4/4
dari	3	1	1	1	3/4
eric	1	2	7	12	1/4
fab	1	1	1	1	4/4
gui	1	5	3	4	1/4
luc	1	1	1	1	4/4
luis	1	1	1	1	4/4
lup	7	1	1	1	3/4
mai	1	8	2	1	2/4
mic	1	1	1	1	4/4
moca	1	1	1	1	4/4
moni	1	1	2	1	3/4
monz	1	2	1	1	3/4
pat	1	1	1	1	4/4
pie	1	2	1	1	3/4
ric	1	1	1	1	4/4
rim	1	1	1	1	4/4
stef	18	5	11	4	0/4

Table 1. Monogram recognition

We are confident that use of the whole monogram (not just the outline) and cooperation of different measuring functions, can give competitive results.

6 Conclusions

We have proposed a new paradigm for comparing size functions. It consists on stressing those angular points which are approximately in the same position in the size functions of most elements of the training set. This seems to be more efficient, and also faster, than other comparison methods.

An application to monogram recognition, although designed with big simplifications (use of the mere outline), gives acceptable results.

Future developments will surely be: integration of different measuring functions and use of the whole image in monogram classification with Radó-like measuring functions; possibly, application of the method to retinal images, to faces, and to other natural images for which a fixed (but not strictly geometric) model is given.

Acknowledgements

Research accomplished under support of CNR–GNSAGA, MURST, ASI, and within an agreement with Elsag Bailey.

We wish to thank S. Gallina, P. Donatini, and above all E. Porcellini for their great help.

References

1. Bajaj, R., and Chaudhuri, S., "Signature verification using multiple neural classifiers", *Pattern Recognition* 30 (1997), 1–7.
2. Collina, C., Ferri, M., Frosini, P., and Porcellini, E., "SketchUp: Towards qualitative shape data management", in *Proc. ACCV '98*, Hong Kong (1998).
3. Ferri, M., and Frosini, P., "Range size functions", in *Proc. of the SPIE on Vision Geometry III*, Boston, Mas., Vol. 2356 (1994), 243–251.
4. Ferri, M., Gallina, S., Porcellini, E., and Serena, M., "On–line character and writer recognition by size functions and fuzzy logic", in *Proc. ACCV '95*, Singapore, vol. 3 (1995), 622–626.
5. Ferri, M., Lombardini, S., and Pallotti, C., "Leukocyte classification by size functions", in *Proc. 2nd IEEE WACV*, Sarasota (1994), 223–229.
6. Frosini, P., "Discrete computation of size functions", *J. Combin. Inf. Syst. Sci.* 17 (1992), 232–250.
7. Frosini, P., "Connections between size functions and critical points", *Math. Meth. Appl. Sci.* 19 (1996), 555–569.
8. Frosini, P., "Size theory as a topological tool for computer vision", in: *"From segmentation to interpretation and back: Mathematica methods in computer vision"*, Eds.: T. Moons, E. Pauwels, L. Van Gool (in print).
9. Mokhtarian, F., "Silhouette–based isolated object recognition through curvature scale space", *IEEE Trans. on PAMI* 17, 5 (1995), 539–544.
10. Nouboud, F., "Handwritten signature verification: A global approach", in "Fundamentals in handwriting recognition" Ed.: S. Impedovo, NATO ASI Series F, 124, Springer Verlag (1994), 455–459.
11. Uras, C., and Verri, A., "On the recognition of the alphabet of the sign language through size functions", in *Proc. 12th ICPR*, Jerusalem (1994).
12. Verri, A., and Uras, C., "Invariant size functions", in *Applications of invariance in Computer Vision, Lecture Notes on Computer Science*, J.L. Mundy, A. Zisserman, D. Forsyth eds., vol. 825, Springer–Verlag (1994), 215–234.
13. Verri, A., Uras, C., "Computing size functions from edge maps", *Internat. J. Comput. Vision* (1997).
14. Verri, A., Uras, C., Frosini, P., and Ferri M., "On the use of size functions for shape analysis", *Biol. Cybernetics* 70 (1993), 99–107.
15. Xu, C., and Velastin, S.A., "A comparison between the standard Hough transform and the Mahalanobis distance Hough transform", in: *Proc. ECCV '94*, Ed.: J.–O. Ecklundh, Stockholm, Vol. 1 (1994), 95–100.

Fig. 2a

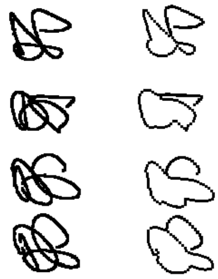

Fig. 2b

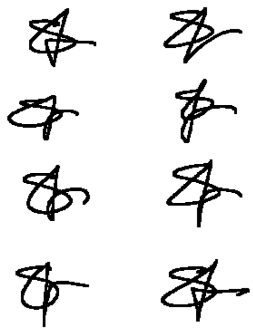

Fig. 3a

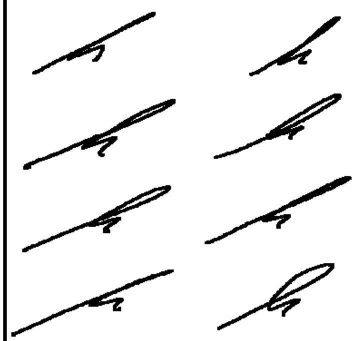

Fig. 3b

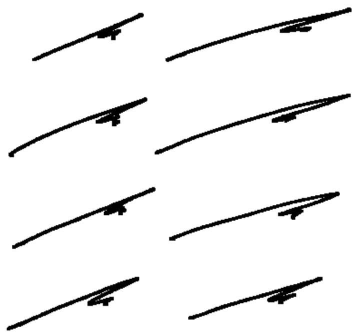

Fig. 3c