Image Retrieval through Abstract Shape Indication

A. Brucale, F. Cesari, M. d'Amico, M. Ferri, P. Frosini,
L. Gualandri, M. Guerra, A. Lovato, I. Pace *
Dip. di Matematica,
Università di Bologna

September 19, 2000

Abstract

A new system for image retrieval is presented. The query is indicated by means of three pictures, which should span the abstract shape concept that the user has in mind. The search is accomplished by using a set of size functions, and giving them different weights, computed on the base of the three input images. The system has been tested on a database of 2976 synthetic images .

1 Introduction

This is a preliminary report on a new type of qualitative image retrieval system. The query is expressed by means of three images that the user either uploads or draws directly. The user is requested to propose three "extreme" instances of the shape he/she has in mind. See Figure 1 for some samples from the database of 2976 images we have used.

In what is this system innovative? The novelty consists in the fact that the user should indicate his/her "abstract" idea of the queried shape to the system, through the three images. What the three pictures should have in common, is not to be superimposable [7, 1, 6], nor to have a similar colour distribution [8], and not even to share some parts of a feature vector. The system then outputs a list of images with decreasing pertinence. See Figure 2 for an example.

The importance of this work is in the shift from quantitative to qualitative, from rigid to flexible, from template matching to shape understanding, in the recognition and retrieval of images.

2 How it works

The mathematical core of the system is a "team" of *size functions* [5, 9]. These are transforms of images — or also of other types of signal — seen as subsets of a Euclidean space. They depend on the choice of so called *measuring functions* (continuous real maps defined on the set). The definition of particular measuring functions allows to isolate those aspects of the image shape, which are of interest for the particular application goal. (Of course, a discrete version of the theory is used in the concrete applications.)

Here is how the system should extrapolate the abstract shape concept from the three images. For each measuring function embedded in the program, the system computes the size functions of the three images, and evaluates their reciprocal distances. Then a comparison is done with the probability that these dis-

*Address: Piazza di Porta S. Donato 5, I–40126 BOLOGNA Italy. E–mail: `ferri@dm.unibo.it`

Query	T	F	H
Asterisks	40/50	29/30	10/10
Crosses	25/38	13/15	10/10
Double arrows	39/48	25/43	10/10
Segments	20/23	18/23	10/10
Horiz. segments	10/12	8/6	8/8
Hands	28/29	18/26	10/10
Notes	31/45	30/43	10/10
Little men	43/48	22/26	10/10
Upside–d. l. men	32/37	19/27	10/10
Waves	13/44	13/44	10/10
Sea fauna	20/31	9/8	9/9
Quadrilaterals	27/30	28/31	10/10
Suns	35/46	35/17	10/10
Spirals	31/34	24/35	10/10
Stars	38/67	34/34	10/10
Tongs	22/24	17/3	10/8

Table 1: Query with database images.

Query	T	F	H
Asterisks	33/36	23/19	10/10
Crosses	5/26	1/12	1/10
Arrows	32/40	1/19	9/10
Double arrows	18/27	1/18	1/10
Notes	14/18	7/1	8/1
Little men	14/24	1/9	1/9
Waves	8/34	1/31	1/10
Sea fauna	3/6	1/1	0/3
Squares	18/23	12/3	10/9
Quadrilaterals	37/48	12/5	10/7
Suns	40/51	5/19	9/10
Spirals	32/54	12/51	10/10
Stars	1/17	1/4	0/7
Tongs	9/11	2/1	2/1

Table 2: Query with hand–drawn images.

tances occur in a random triple. The lower the probability, the higher the weight that is given to the measuring function. The classifiers, corresponding to the measuring functions, co–operate — with contributions depending on these weights — in the determination of a pertinence factor. On the base of the pertinence factor a set of images is extracted; finally, this set is sorted by a much finer comparison of the size functions.

In other words, the user puts a common feature into the three pictures, which for all the rest should be as different as possible. Then, those measuring functions which are best fit to recognize that feature, are stressed in the comparison of the three input images (or, better said: of their size functions) with the database.

3 The database

As a preliminary version, we have chosen to build a database of simple black–on–white silhouettes. It consists of 2976 images belonging to 18 classes, plus 62 classes corresponding to alphabet characters, plus an extra "jamming" set of 197 unrelated images. This is the database to which tables 1 and 2 refer. There, T = Total # of hits, F = Position of First wrong, H = # of hits in top 10.

The images are first transformed into their size functions, one for each measuring function. These, in turn, are codified by complex polynomials [4], or better said, by their coefficient vectors. These are organized in an M–tree [2, 3].

4 The experiment

We have tested the system by experiments of two kinds: Input of pictures taken from the database itself, and input of hand-drawn sketches.

We have always input triples of objects belonging to the same class. Whenever possible, we input a triple as generic as possible (e.g. with respect to orientation; see upper Figure 3) and a triple with some more restricted features (e.g. a particular orientation; see lower Figure 3). Tables 1 and 2 contains the worst and best results of a first set of experiments.

As can be seen, the results of the queries are extremely satisfactory when the input pictures come from the database (Table 1). The hit ratio is sensibly lower — as was to be expected — with the hand–drawn inputs (Table 2). This we want to overcome by using a wider set of measuring functions.

We have considered to be a "hit" also some outputs which don't formally belong to the class of the input pictures: E.g., if three (not necessarily vertical) crosses are drawn as input, the output of an "X" character has been considered as a valid answer.

Computing time is still a concern: a query takes 10 to 15 seconds in average on a Pentium III based PC with a still nonoptimized code.

5 Conclusions and future developments

The use of size functions with adaptable weights has proven reliable for making the system catch the "abstract" shape category, that the user has conveyed by the three input images.

Presently, we are working at a database of clip art. Segmentation is an essential issue, and we are facing it with the help of the Mumford–Shah functional in its implementation within the Megawave package by Cérémade.

About the aforementioned speed problem, we are exploring different distances to be applied directly to the trees, which store the whole information of the size functions.

The next version of the program will also foresee an interactive query process: The user will be enabled to give scores to the output images.

Acknowledgements

This research was accomplished under support of INdAM–GNSAGA and of the University of Bologna, funds for selected research topics.

References

[1] R. Basri, D.W. Jacobs, *Recognition using region correspondences*, Int. J. Comp. Vision 25 (1997), 145–166.

[2] P. Ciaccia, M. Patella, P. Zezula, *M-tree: An efficient access method for similarity search in metric spaces*, Proc. 23rd Intl. Conf. On Very Large Data Bases (VLDB '97), Athens, Greece (1997), p. 426–435.

[3] F. Cesari, "Uso dell'albero metrico M-tree e delle funzioni di taglia nel riconoscimento mediante immagini in database estesi", Tesi di Laurea in Ing. Informatica, Università di Bologna (2000).

[4] M. Ferri, C. Landi, *Representing size functions by complex polynomials*, Proc. Math. Met. in Pattern Rec. 9, Moskow (1999).

[5] P. Frosini, C. Landi, *Size Theory as a Topological Tool for Computer Vision*, Pattern Recognition and Image Analysis 9 (1999), 596–603.

[6] M. Hagedoorn, R.C. Veltkamp, *Reliable and efficient pattern matching using an affine invariant metric*, Int. J. Comp. Vision 31 (1999), 203–225.

[7] J. Ross Beveridge, E.M. Riseman, *How easy is matching 2D line models using local search?*, IEEE Trans. PAMI 19 (1997), 564–579.

[8] C. Schmid, R. Mohr, *Local grayvalue invariants for image retrieval*, IEEE Trans. PAMI 19 (1997), 530–535.

[9] A. Verri, C. Uras, P. Frosini, M. Ferri, *On the use of size functions for shape analysis*, Biol. Cybern. 70 (1993), 99–107.

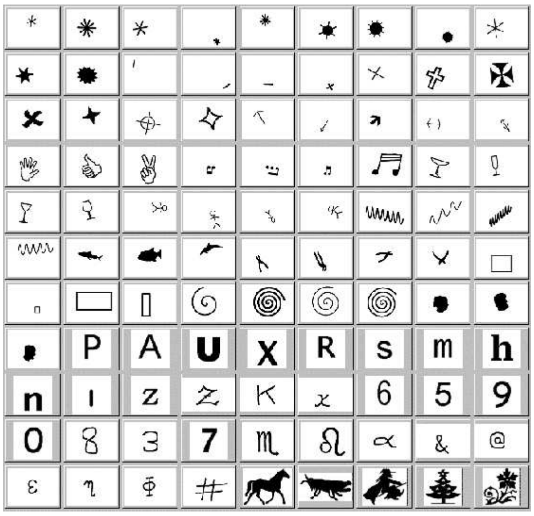

Figure 1: *Samples from the database.*

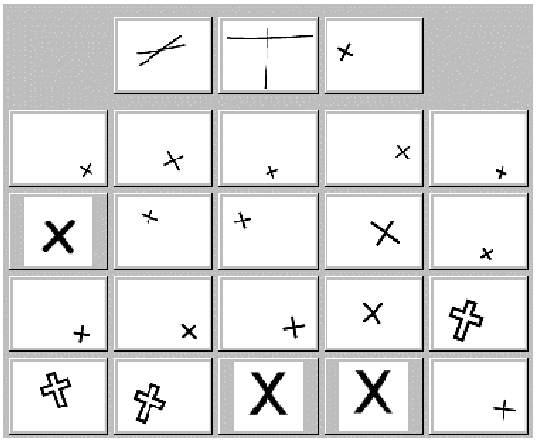

Figure 2: *A hand–drawn query and the output (first 20 images).*

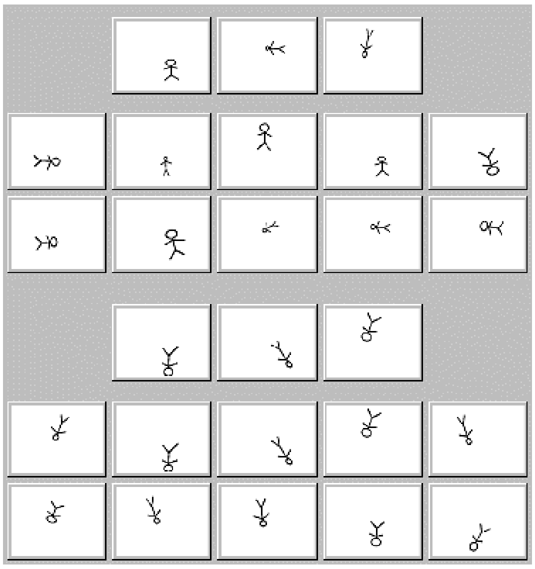

Figure 3: *Two queries with images from the database, and the output (first 10 images).*