Automatical Definition of Measures from the Combination of Shape Descriptors

J.-P. Salmon, L. Wendling, S. Tabbone
LORIA,
Campus scientifique, BP 239
54 506 Vandœuvre-les-Nancy Cedex France

Abstract

This paper presents a novel approach to combine shape descriptors. Each approach is applied on several clusters of objects. For each cluster and for any descriptor a map is associated directly from the confusion matrix. Such a method allows to determine automatically the better weight associated to the descriptor for the object under consideration. At last, we show that the additive combination of such measures allows to improve the classification.

1. Introduction

There have been relatively few studies of combining the similarity responses of shape descriptors. Generally the "better"[1] descriptor is searched and applied. This, in turn requires efficient ways for recognizing objects [5], or at least for identifying object signatures. The relatively low amount of work in this area is probably due to the huge variety of objects encountered, depending on the type of image or documentation which has to be processed. Generally photometric shape descriptors can be split into two categories [18]: contour-based and region-based approaches with well known advantages and drawbacks. Due to the large variability of the objects the use of invariant descriptors is required for their identification and recognition. Many approaches have been proposed in the early years either based on Fourier descriptors [10, 13], on moments [2, 16], on Hough transform [1, 9] or on angular description [4, 15]. In this paper we focus on basic global methods having low processing time but allowing to keep nice geometrical properties as translation, rotation and scale factor. For a thorough survey of the various techniques and descriptors which may be used, the reader may refer to [8, 11, 18]. Another problem encountered in pattern recognition is the kind of distance calculated during the matching. In this scope, generally the metric is related to the application problem under consideration. Metrics such as Bhattacharya [3], Euclidean distance or EMD [12] are widely used from statistical studies and related to the studied application. Here we used Euclidean distance to compare scalar values and the classical Tanimoto index (min over max) for vector description of an object. This yields interesting results in our application, with only low processing time requirements. The aim of this paper is to assign to each descriptor a recognition map. That is to define automatically a descriptor measure for each object. Such function could be different for another object for a same descriptor. Descriptors used in our study and associated properties are provided in section 2. Simple classification measures are proposed in section 3 to study the behavior of our method. The scheme used to defined measures from the behavior of shape descriptors is presented in section 4. Finally experimental studies and a discussion about the advantages and limitations of our approach are given in sections 5 and 6.

1 We aware that the notion of better usually written – for pattern recognition methods, segmentation, ... – can only refer to the application under consideration even if it is widely used.

2. Set of Basic Descriptors

We focus here on basic operators having low processing time and easy to implement. We draw for each of them their behavior from standard geometric transforms like homotety, rotation or translation.

2.1. Compactness

The compactness is the more common descriptor which can be found in the literature. Its approximation is maximal for a large disk (near one) and minimal for a discrete straight line. An usual formulation is:

$$C = \frac{M_{11}}{P^2} \tag{1}$$

with P the perimeter and M_{11} the surface. Such feature is by definition invariant to basic geometric transforms as rotation, translation and scale factor. Its calculation requires a low processing time. Nevertheless it has a high sensitivity to noise and gives rise to low recognition rates when the number of objects to recognize grows. Such descriptor is rather suitable to split databases in more or less elongated clusters or dedicated for specifical applications having well differentiated shapes. In this paper, objects are classified following the distance of their compactness value.

2.2. Ellipticity Degree

The ellipticity corresponds to the ratio of the big axis and the little axis [16]. Its expression is given by:

$$E = \sqrt{\frac{(M_{20}+M_{02}) + \sqrt{(M_{20}-M_{02})^2 + 4 \cdot M_{11}^2}}{(M_{20}+M_{02}) - \sqrt{(M_{20}-M_{02})^2 + 4 \cdot M_{11}^2}}} \quad (2)$$

Where M_{jk} are the moment of order $j+q$. Such descriptor is invariant to rotation, translation and scale factor. Moreover it is low sensitive to noise as it considers the global shape of the object. Normalized values of distance are compare.

2.3. Angular Signature

We used here a simplified version of the angular signature proposed by T. Bernier *et* Landry[4]. A vector is computed from the centroid $\bar{X}_c(x_c, y_c)$ of the shape as follows:

$$S_A = \left[\frac{d(X_1, \bar{X}_c)}{\sqrt{M_{11}}}, \ldots, \frac{d(X_n, \bar{X}_c)}{\sqrt{M_{11}}}\right] \quad (3)$$

with d the Euclidean distance. S_A corresponds to the successive distances, with a regular angular step, between the centroid X_c and the farer point X_i of the contour normalized by the area of shape and n is the number of processed directions. The method proposed in [4] was adapted to keep fast processing. The derivative of the signature is not studied to take into account the rotation by extracting the maxima. We consider here the maximal diameter of the object. Then a similarity ratio is directly computed from superimposed signatures recentered using their maximal diameter.

2.4. Generic Fourier Descriptors

The polar discrete transform [17] is required to compute such signature. It is similar to Fourier transform but considering the polar image in a polar space as a rectangular image in a Cartesian space. The mathematical expression is:

$$PF(\rho, \phi) = \sum_x \sum_y I(x,y).e^{\left[2j\pi\left(\frac{r(x,y)}{R}\rho + v(x,y)\phi\right)\right]} \quad (4)$$

The radius $r(x,y)$ and the angle $v(x,y)$ are the polar coordinates of the point (x,y) from the centroid frame of the object. I is the intensity function. The parameters ρ and ϕ are bounded: $0 \leq \rho < R$ and $0 \leq \phi < T$ with R and T respectively the radial and angular resolutions. The Fourier descriptor is parameterised by two frequencies: m the radial frequency and n for the angular frequency [17]:

$$GFD = \left\{\frac{|PF(0,0)|}{M_{11}}, \frac{|PF(0,1)|}{|PF(0,0)|}, \ldots, \frac{|PF(m,n)|}{|PF(0,0)|}\right\} \quad (5)$$

GFD is invariant to scaling, translation and rotation and low sensitive to noise effect. The distance between two shapes is directly given by the distance of the associated vectors. We can denote that such method is easy to implement and a full algorithm is provided in [17].

2.5. R-Signature

The Radon transform of a function f, denoted T_{Rf}, is defined as its line integral along a line inclined at an angle θ and at a distance ρ from the frame [6]. A shape measure, called $\mathcal{R}$-signature [15], can be defined from the Radon transform as follows:

$$\mathcal{R}_f(\theta) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho,\theta)d\rho \quad (6)$$

Such signature allows to have an idea of the angular distribution of a shape considering its global aspect and none is centroid. It is by definition invariant to translation. The normalization by the area allows to take into account the scale factor. Circular shifts are performed to keep the rotation property. At last, the similarity measure used to compare two R-signature is based on the classical Tanimoto index.

3. Classification Measures

3.1. Confusion matrix

A confusion matrix allows to evaluate the power of recognition of a descriptor following the current application. Generally a head of cluster if defined assuming that it is a well representation of the cluster. The samples are stored in the matrix by considering the cluster values of the distance reached from all the heads. Such a process requires static objects. Another way consists in adding in the matrix the cluster numerous of the nearest sample. Here, we consider each samples of the database as a query to have a more discriminate matrix. For each query we sort the other objects by decreasing distance (or decreasing similarity ratio). Then we analyse the clusters corresponding to each element of the series achieved. The ranking takes into account all the positioning of the samples of the cluster associated to the query. The confusion matrix is made by corresponding the numerous of clusters found in the initial positions.

3.2. Ranking

The ranking is a well-known measure of retrieval [7] adapted here to show the robustness of the method. Let us consider a cluster c, its ranking value is defined by:

$$r_c = \frac{1}{S} \sum_{j=1,S} \frac{1}{S} \left(\sum_{i=1,S} \delta_{cl(o_c,i)} + \sum_{j=S+1,N} \left(1 - \frac{i-S}{N-1} \right) \right) \quad (7)$$

where $N = C \cdot S$ is the number of objects contained in the database and $cl(o_c, i)$ is the target cluster of the element i of the increasing series of scores (see previously) computed from the object query o_c. δ is set to one when then element i corresponds to the cluster c and 0 otherwise. Values between 0 and 1 can also be integrated if we have *a priori* knowledge of valuated similarity between clusters. When the size of the database grows the evolution of the ranking provides an interesting assessment about the stability of the shape descriptor.

4. Automatical Definition of Measures

A descriptor is generally considered discriminant, under an application, when the associated recognition rates are high that is. A confusion matrix is used here to assess the power of recognition of descriptors. By definition if objects are correctly recognized they should be assigned to the corresponding yarget cluster. Bad recognition remains to a disparate distribution of the objects on the lines of the confusion matrix. Let table 1 be a confusion matrix achieved using the database provided in the experimental section. There exists 9 clusters having 11 samples. The descriptor used is the compactness. The aim is not to study the compactness but to underline the processing step performed for any descriptor considering any confusion matrix.

Cluster	1	2	3	4	5	6	7	8	9
1	**28**	30	0	7	19	0	22	0	15
2	29	**51**	0	7	5	0	20	0	9
3	0	0	**44**	4	0	30	0	42	1
4	7	6	4	**47**	19	0	7	1	30
5	20	6	0	17	**34**	0	15	0	29
6	0	0	33	0	0	**49**	0	39	0
7	20	30	0	7	16	0	**43**	0	5
8	0	0	40	0	0	35	0	**46**	0
9	16	11	0	25	29	0	3	0	**37**

Table 1. Compactness confusion matrix using cluster 6.

Let us consider, for example, the line 6^{th} of the matrix. The values correspond to the distribution of the objects of the cluster number 6 achieved from the calculation of compactness criterion. We can denote that the samples are rather assigned to three clusters (3, 6 and 8). The recognition rate is around 40% (number of correctly classified objects divided by the number of tested objects that is 49/121). Let us now consider a low recognized cluster. For instance the line 5^{th}. In this cluster the distribution is relatively homogeneous between the different clusters and so less discriminate than for the previous example. That is underlined by a graphical representation of the rates provided Figure 1. The dissimilarities are obtained by computing the mean of the distances. The distance, for one cluster, is equal to the difference between the values of the distance of these objects and the one of the searched cluster. We can conclude that

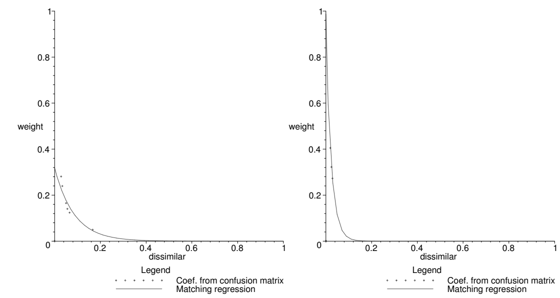

Figure 1. Difference between discriminating descriptor (right) and not (left).

it is important to take into account not only the recognition rate but also the distribution of the errors in order to improve the recognition process. Thus starting from this information one definite a function of confidence (which represents the number of images correctly recognized on the total number of image at a distance given). On the graph 2 we can see the influence of this curve of confidence : the images corresponding to the required class have the smallest distances . The aim is to integrate the behavior of each symbol of the database. Obviously functions may be different for two objects. Then it is important to define different approximations to improve the global recognition. The higher the value of the weight the higher the confidence on the descriptor is. Nevertheless considering one cluster a descriptor can be more discriminating than another one. That is underlined by the value of the weight reached. The score is given by:

$$Score = \sum_{i=1,d} F_{dc}(1 - (\vec{oc} - \vec{ox})) \quad (8)$$

where d is the number of descriptors (here 5). F_{dc} is a weighted function of descriptor d for the cluster c defined from the distance distribution of cluster samples. $\vec{oc}$ and $\vec{ox}$ are respectively the vector of descriptor d for a object of the cluster c and of the cluster searched.

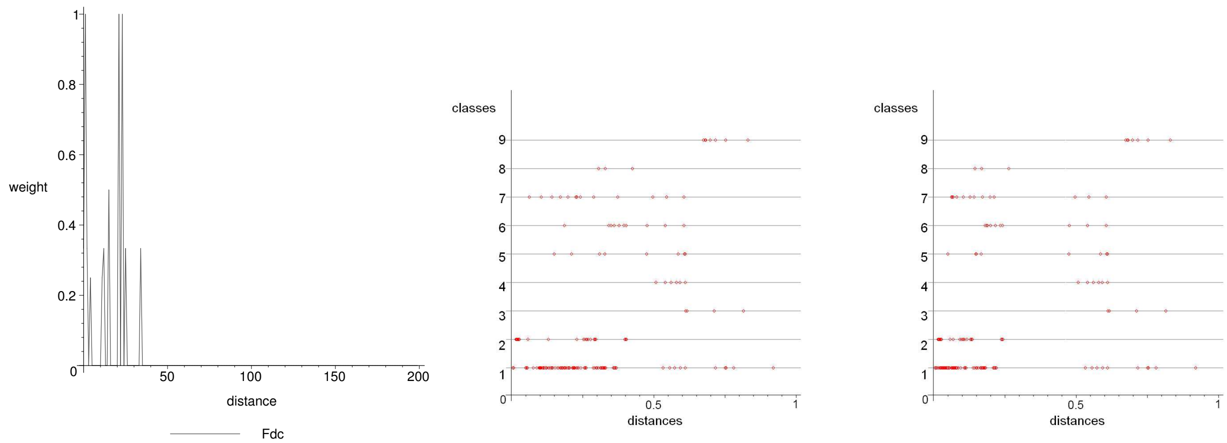

Figure 2. Influence of weighted function F_{dc} (left) on distance between images of cluster 1 and learning database images (before/after).

5. Experimental study

First, a database of D. Sharvit [14] who made it kindly available to us on his Web site: "http://www.lems.brown.edu/vision/researchAreas/SIID/" has been used. Such database consists of nine categories with 11 shapes in each cluster. Some shapes are occluded or distorted and objects are scaled or rotated. We can see (Figure 3) that the proposed approach improves the recognition rates for most of the descriptors considering the whole clusters. Table 2 presents the mean recognition rates. W_s is a normalized weighted sum where the weights are directly defined by the recognition rates reached for all the descriptors applied separately on the database. The results are provided $before$ and $after$ running the approach proposed in section 3. We can remark that the combination of measures improves the recognition.

Reco. Rates	C	E	S_A	G_{FD}	T_{Rf}	W_S
Before	35	43	60	73	54	78
After	39	45	62	73	55	90
Ranking	C	E	S_A	G_{FD}	T_{Rf}	W_S
Before	87	85	88	96	89	97
After	91	91	96	98	92	100

Table 2. Using Sharvit's database.

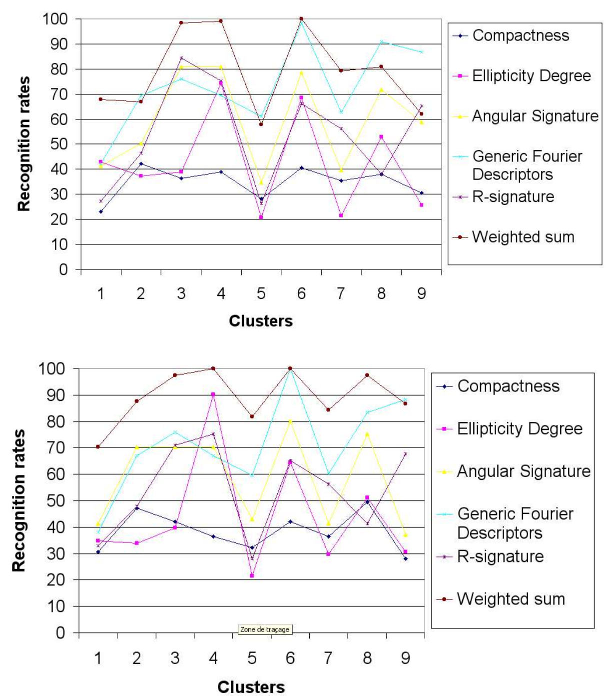

Figure 3. Recognition rates by descriptors.

Another test database consists of dropped initials[2]. Such letters require a post processing in order to be usable within the framework of our application, as provided figure 4. After applying algorithms of dilation and erosion, a set of rules and measures (size, compactness...) allows to choose the good related component to extract. Thus we obtained a database made up of 9 letters declined into 11 noisy samples. We carry out same measures on this database (Table 3) and we could note an consistent improvement us-

2 We would like to thank the Centre d'études Supérieures de la Renaissance for the permission to use their archival documents

Figure 4. Example of extraction of letters.

ing our method. Then we have considered increasing sizes of databases (using a greater database of reference letters merged with the previous one). The results achieved show that our approach remains stable, in this example, when the database grows (Table 4).

Reco. Rates	C	E	S_A	G_{FD}	T_{Rf}	W_S
Before	49	41	70	59	50	55
After	52	39	75	69	50	72
Ranking	C	E	S_A	G_{FD}	T_{Rf}	W_S
Before	85	78	90	89	80	96
After	93	81	97	94	87	100

Table 3. Using database of dropped initials.

Reco. Rates	Size	2×S	3×S	4×S	5×S
Before(W_S)	55	64	58	60	55
After(W_S)	72	88	84	84	75
Ranking	Size	2×S	3×S	4×S	5×S
Before(W_S)	96	96	95	95	90
After(W_S)	100	100	99	99	95

Table 4. Using increasing Sizes (noted S) of databases.

6. Conclusion

In this paper, we have shown that the definition of measures for a set of descriptors can be of great interest for recognizing objects. These results are very promising; however, they still need further validation by processing much larger databases in order to assess the discriminating power and the robustness of the proposed method. In this case, it may be necessary to add some indexing scheme, both for efficiency and for discrimination. We are planning to test the construction of a binary search tree associating each node with the more discriminating feature found.

References

[1] D. H. Ballard. Generalizing the Hough Transform to Detect Arbitrary Shapes. PR, 13(2):111–122, 1981.

[2] R. Bamieh and R. de Figueiredo. A General Moments Invariants/Attributed Graph Method for the Three Dimensional Object Recognition from a Single Image, IEEE Journal of Robotics and Automation, 2:240–242, 1986.

[3] A. Bhattacharya. On a measures of divergence between two statistical populations defined by their probability distributions, Bull. Calculus Math Soc.,99–109, 1943.

[4] T. Bernier and J-A. Landry. A new method for representing and matching shapes of natural objects. Pattern Recognition, 36(8) (2003) 1711–1723.

[5] L. P. Cordella and M. Vento. Symbol recognition in documents: a collection of techniques? International Journal on Document Analysis and Recognition, 3(2):73–88, 2000.

[6] S.R. Deans. Applications of the Radon Transform. New York: Wiley Interscience Publications, (1983).

[7] T. Gevers and A.W.M. Smeulders. Context-based image retrieval by viewpoint-invariant color indexing, Image and Vision Computing, 17(1999) 475–488.

[8] A. K. Jain and R. P. W Duin and J. Mao, Statistical Pattern Recognition: A Review, IEEE Transactions on PAMI, 22(1):4–37, 2000.

[9] A. A. Kassim, T. Tan, and K. H. Tan. A Comparative Study of Efficient Generalised Hough Transform Techniques. Image and Vision Computing, 17:737–748, 1999.

[10] H. Kauppinen, T. Seppnen, and M. Pietikinen. An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification. IEEE Transactions on PAMI, 17(2):201–207, 1995.

[11] S. Loncaric. A Survey of Shape Analysis Techniques. Pattern Recognition, 31(8):983–1001, 1998.

[12] Y. Rubner and S. Tomasi. Perceptual metrics for image database navigation, kluwer, Boston, 2000.

[13] Yong Rui and Alfred She and Thomas S. Huang. A Modified Fourier Descriptor for Shape Matching in MARS, Image Databases and Multimedia Search, ed. S. K. Chang, 8:165–180, 1998.

[14] D. Sharvit. J. Chan. H. Tek and B. Kimia. Symmetry-based Indexing of Image Databases. Journal of Visual Communication and Image Representation, 12, 1998.

[15] S. Tabbone and L. Wendling. Binary shape normalization using the Radon transform, DGCI, LNCS 2886, Naples, Italy, vol. 2886, pp. 184–193, 2003.

[16] R. Teag. Image Analysis via the General Theory of Moments. Journal of the Optical Society of America, 70(8):920–921, 1980.

[17] D. Zhang and G. Lu. Shape-based image retrieval using generic Fourier descriptor, Signal Processing: Image Communication, 17:825–848, 2002.

[18] D. Zhang and G. Lu. Review of shape representation and description techniques, Pattern Recognition, 37(1):1–19, 2004.