Experiences with Shape Classification through Fuzzy *c*-Means Using Geometrical and Moments Descriptors

Ugo Erra[1] and Sabrina Senatore[2]

[1] Università della Basilicata, Dipartimento di Matematica e Informatica,
Viale Dell'Ateneo 10, Macchia Romana - 85100, Potenza, Italy
ugo.erra@unibas.it

[2] Università di Salerno, Dipartimento di Matematica e Informatica
Via Ponte Don Melillo - 84084 Fisciano(SA), Italy
ssenatore@unisa.it

Abstract. Due to the growing diffusion of digital media, most of real world applications have data with multiple modalities, from multiple sources and in multiple formats. The modelling of information coming from multimedia sources represents an important issue for applications which achieve multimedia mining activities. In particular, the last decades have witnessed great interest in image processing by "mining" visual information for objects recognition and retrieval. Some studies have revealed the image disambiguation based on the shape produces better results than features such as color or texture; moreover, the classification of objects extracted from an image database appears more intuitively formulated as a shape classification task.

This paper presents an approach for 2D shapes classification. The approach is based on the combined use of geometrical and moments features extracted by a given collection of images and achieves shape-based classification exploiting fuzzy clustering techniques.

1 Introduction

In the age of digital information, the growing amount of large-scale image repositories in many application domains emphasize the need for effective means for mining and classifying digital image collections.

In general, two different approaches have been applied to allow image retrieval: one based on textual information whereas the other based on image content information. The first retrieval approach consists of attaching textual metadata to each image and then submits a keyword-based query to the database in order to retrieve them [23]. This approach requires an initial annotation activity which often results laborious and time-consuming; moreover, it is a human driven process thus, similar images characteristics can be expressed by different users with different terminologies, affecting the performance of the keyword-based image search.

M. Detyniecki et al. (Eds.): AMR 2010, LNCS 6817, pp. 189–203, 2012.

For these shortcomings, (semi-)automatic approaches have been achieved to process the image in order to get more "objective" *content-based* image properties such as color, texture, and shape. Content-Based Image Retrieval (CBIR) systems involve characterizing an image using a set of features; retrieval or classification is then performed by measuring similarity to a required query image [34] contrasting to the effort needed to annotate images.

Images can be particularly complex to manage; thus, CBIR techniques require the translation of high-level user perceptions into low-level image features. To cope with the so called "semantic gap" problem, these features should be consistent and invariant to remain representative for the images collection in a database. Image indexing is not an issue of string processing (as in the case of standard textual databases), but an n-dimensional vector describes the characteristic of the image [14]. Then, the image retrieval process consists of discovering all the images whose features are similar to the query example image. A direct drawback is that these low-level image features are often too restricted to describe images on a conceptual or semantic level, impacting on the performance of image retrieval approaches.

On the other hand, the CBIR technology tries to address two intrinsic problems: (a) how to mathematically describe an image, and (b) how to assess the similarity between a pair of images based on their abstracted descriptions. Recent methodology development employs statistical and machine learning techniques in various aspects of the CBIR technology. In image classification methods, the approaches are based on learning-based classification and non-parametric classifiers. As been pointed out in [6], despite the large performance gap between non-parametric classifiers and state-of-the-art learning-based, the non-parametric image classification have been considerably under-valued and offer several advantages: (i) can naturally handle a huge number of classes; (ii) avoid overfitting of parameters, which is a central issue in learning based approaches; (iii) require no learning/training phase. As explained later in this paper, our approach could be considered parameter-free, when the number of cluster is known a priori.
The focus of this work is to define an approach for image classification and retrieval based on 2D shape features, exploiting fuzzy clustering techniques. The paper is organized as follows. Section 2 gives a sketched overview of the related works in this area, then as a background, Section 3 focuses on the image processing for the image analysis and features extraction whereas Section 4 introduces the fuzzy clustering algorithm exploited in this approach. Finally, Section 5 describes the experiments and provides the results. Conclusions and future works close the paper.

2 Related Works

Improvements in data storage and image acquisition technologies require new computer-assisted image understanding tools which support the large-scale image and media content datasets and provide assistance in image processing, query and retrieval. CBIR systems address these important issues in computer vision and multimedia computing, supporting effective searching and browsing of large image digital libraries based on automatically derived image features [14]. Some examples of popular CBIR systems are QBIC, Virage, RetrievalWare, Photobook, Chabot, VisualSeek, WebSeek, MARS system, SurfImage, Netra, and CANDID (for additional details about them refer to [28]). Furthermore, a complete and exhaustive survey on CBIR developments and advances is provided in [9]. Almost all of these approaches are based on indexing imagery in a feature space. A feature represents a certain visual property of an image, either globally for the entire image or locally for a small group of pixels. The feature extraction is often considered as a preprocessing step, which represents the inputs to subsequent image analysis tasks. Typical features are color, texture, shape and region.

Also the increasing diffusion of images compression requires challenging techniques to extract visual features [22]: sophisticated global features such as the wavelets [29] and large collection of local image descriptors as SIFT [24].

Some other techniques improve the effectiveness of image retrieval through multi-features combination [15],[33] and then, by measuring similarity to a required query image [34]. Combination of words and features characterize annotated training sets of images [27], which will be used for classification or retrieval. In [13] a hierarchical feature subset selection algorithm for semantic image classification is defined, where the feature subset selection procedure is seamlessly integrated with the underlying classifier training procedure.

The image description and the user's perception of these features evidence the imprecise nature of the retrieval which can benefit by fuzzy techniques. Fuzzy logic is suitable for expressing queries which involve concepts and linguistic expression by means of fuzzy values rather than crisp features values [20].

Applying fuzzy processing techniques to CBIR approaches has been extensively investigated in literature. In general, fuzzy retrieval models offer more flexibility in the representation of the terms' index, preferences among terms in a query and ranked results. In particular CBIR models take advantage by using technique based on fuzzy theory for knowledge representation, for uncertainty management, against traditional information retrieval models based on boolean, vector-based or probabilistic representation. An example is given in [2] where a fuzzy information retrieval model for textual data has been extended to implement a model in image context. In [21], fuzzy logic has been employed to interpret the overall color information of images: according to the human perception, nine classes of colors are defined as features.

A fuzzy color histogram approach in [30] allows the evaluation of similarity through fuzzy logic-based operations. In [8], instead the similarity of two images has been defined by considering the overall similarity between families of fuzzy features. More specifically, each image has been associated to a family of fuzzy features (fuzzy set) representing color, texture, and shape properties. This approach reduces the influence of inaccurate segmentation, compared with other similarity measures based on regions and with crisp-valued feature representations.

Many CBIR approaches exploit clustering for preprocessing activities [12], specifically, fuzzy techniques are widely employed in image classification methods. In [25], a method to calculate image similarity measure using fuzzy partition of the HSI color space has been presented. In particular, the fuzzy c-means (FCM) clustering algorithm [5] has been shown to provide effective partitions for image segmentation on medical images [16], satellite images [18][32], etc. Some extension and modification of FCM are applied to image segmentation in infrared images domain [19]. In [26] a modified version of FCM has been proposed, to solve the problem of large-scale image retrieval and classification, even though the clustering step is performed in lower-dimension space, and image retrieval is only performed in clustered prototypes. Yet, in the most of approaches, the execution time of the clustering algorithm is a critical point, which finds a solution in [19].

3 Design of the Feature Space

The first step toward the shape analysis of a given image involves separating the object (or region) of interest from other non-important image structures by using an image segmentation approach. There are several approaches for the extraction of the shape from a given image based on clustering methods, histogram methods, edge detection, level set methods, graph partitioning methods and so on. In general there is no a general solution and there is always an image where an approach does not yield good result, i.e., if the foreground and background share many similar colors, an approach could give a result with parts of background labelled as foreground object. This is challenging in shapes classification because any approach must take into account this drawback. In our implementation, we adopt the k-means clustering algorithm for image segmentation which is suitable when the foreground and background colors contrast sufficiently with each other.

A shape descriptor is a set of numbers that are extracted from the region of interest in order to describe a given shape feature. Efficient shape features must present some essential properties such as identifiability, invariance, noise resistance, statistically independence and so on.

In this work, we adopt three types of such shape descriptions: geometric description, invariant moments and affine moments. The geometric features discriminate shapes with large difference. They are useful to eliminate false hits and usually are not suitable as single description, in fact they are combined with other shape descriptors to better discriminate shapes. The moment instead, represents a mathematical concept coming from the concept of moment in physics. It is used in computer vision for both contour and region of a shape. In particular, the invariant moments [17] are one of the most popular and widely used contour-based shape descriptors. Affine moments invariants are instead features computed from moments that do not change their value in affine transformation.

In the case of geometric features, let P and A denote the shape perimeter and area, respectively. Note that perimeter and area are invariants respect to translation and rotation but when combined, they are not invariant with respect to scale. The features we adopt are:

- Eccentricity E is the measure of aspect ratio. It is defined as the ratio $E = W_{bb}/H_{bb}$ where W_{bb} and H_{bb} are, respectively, the width and height of minimal bounding rectangle of the shape.
- Rectangularity R represents how rectangular a shape is, i.e. how much it fills its minimum bounding box. It is defines as $R = A/A_{bb}$ where A_{bb} is the area of the minimum bounding rectangle.
- Compactness C is a measure that combines area with perimeter. It is defined as $C = L^2/4\pi A$.
- The value π_{gen} is a measure of the compactness of a shape respect to a circle. It is defined as $\pi_{gen} = P/W_{bb}$.

Among the region-based descriptors, invariant moments m_{pq} are the simplest and is given as:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \qquad p,q = 0,1,2,\ldots$$

where $f(x,y)$ is the intensity function at position (x,y) in a 2D gray level image. In order to obtain translation invariance, the central moments μ_{pq} should be applied:

$$\mu_{pq} = \sum_x \sum_y (x-\overline{x})^p (y-\overline{y})^q f(x,y) \qquad p,q = 0,1,2,\ldots$$

where $\overline{x} = m_{10}/m_{00}$ and $\overline{y} = m_{01}/m_{00}$. Given central moments we are able to compute a set of 7 invariant moments [17], given by:

$$\begin{aligned}
I_1 &= \eta_{20} + \eta_{02} \\
I_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
I_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
I_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
I_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} - \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + \\
&\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \\
I_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] + \\
&\quad 4\eta_{11}^2(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
I_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} - \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + \\
&\quad (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}$$

where $\eta_{pq} = \mu_{pq}^{\gamma}$ and $\gamma = 1 + (p+q)/2$ for $p+q = 2,3,\ldots$. These moments are simple to calculate and they are invariant to translation, rotation and scaling but have an information redundancy drawback since the basis in not orthogonal[7]. From central moments with a little computational effort we are able to obtain also an affine transform invariance which includes the similarity transform and in addition to that stretching and second rotation. We adopt affine moments as defined in [10] and given as:

$$AMI_1 = (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4$$

$$AMI_2 = (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30} + \mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^3\mu_{12}^3)/\mu_{00}^{10}$$
$$AMI_3 = (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2))/\mu_{00}^7$$

All these features are sufficient to characterize the shape of an image. The rationale behind the choice of these moments is that we are interesting in translation, rotation, scale, and projective transform invariance in order that the location, orientation, and scaling of the shape do not affect the extracted features. Further information on these approaches is discussed in [11].

4 Fuzzy Clustering for the Image Arrangement

The clustering algorithms achieve a partitioning of given data into clusters. In general a partition holds two properties: homogeneity within the clusters (data in a cluster must be similar) and homogeneity between clusters (isolation of a cluster from one another: data of different clusters have to be as different as possible).

The data are opportunely translated into a matrix, where each row is a characteristic vector which represents an image. In fact, the images set has been processed to pull out such data matrix, whose rows and columns are respectively the collected images and the relative extracted features. In this study, we are going to apply a fuzzy approach of clustering, the well-known *fuzzy C-Means* (briefly FCM) algorithm [5]. FCM represents the most common fuzzy clustering, particularly useful for flexible data organization. It takes as input a collection of patterns of a universe U (in our case, the collection of images) in form of matrix and produces fuzzy partitions of the given patterns (i.e. images) into (prefixed) c clusters.

The FCM algorithm recognizes spherical clouds of points (clusters) in a multidimensional data space and each cluster is represented by its center point (prototype). This process is completely unsupervised, aimed at identifying some inherent structures in a set of data.

The fuzzy version of clustering produces a more flexible partitioning of data. Precisely, each pattern (in our case, an image) is not associated exclusively to a cluster, but it can belong to more than one. After the fuzzy clustering execution, each pattern has associated a c-dimensional vector, where each cell represents the membership (in the range [0, 1]) of that pattern to each cluster.

Compared to the crisp version, the fuzzy clustering generates a flexible partitioning, more intuitive to interpret: a pattern can have some characteristics that are natively representative of more than one cluster, and the exclusive belonging to one cluster is a too restricted condition. In the fuzzy approach, the membership values better reveal the nature of data set and allow a clearer data analysis. Anyway, it is conceivable to assign a pattern to the cluster, whose membership is the highest.

More formally, each row of the matrix is a vector that represents an image $I \longleftrightarrow \underline{x} = (x_1, x_2, \ldots, x_h)$, where each component of vector is a value computed for a feature. The FCM algorithm aims at minimizing the objective function constituted by the weighted sum of the distances $dist_{i,k}$ between data points $\underline{x}_k = (x_{k,1}, x_{k,2}, \ldots, x_{k,h})$ and the centers (or prototypes) $\underline{v}_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,h})$, according to this formula:

$$Q(U,c) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{i,k}^{m} (dist_{i,k})^2 \tag{1}$$

where $c \geq 2$ is the number of clusters, $u_{i,k} \in [0,1]$ is the membership degree of $\underline{x}_k$ (k=1, ..., n) in the i-th cluster A_i (i=1, ..., c), and $m > 1$ is the fuzzifier, which controls the quantity of fuzziness in the classification process (common choice of fuzzifier is $m = 2$) and finally $dist_{i,k}$ is:

$$dist_{i,k} = dist(\underline{x}_k,\ \underline{v}_i) = \sqrt{(||\underline{x}_k - \underline{v}_i||^2)} \tag{2}$$

just represents the euclidean distance between the data $\underline{x}_k$ and the center $\underline{v}_i$ of the i-th cluster.

In details, $U = (u_{i,k})$ is a $c \times n$ matrix of cluster memberships satisfying some constraints. In particular, M_{fc} is a family of fuzzy partition matrices:

$$M_{fc} = \left\{ U | u_{i,k} \in [0,1]; \sum_{i=1}^{c} u_{i,j} = 1; 0 < \sum_{k=1}^{n} u_{i,j} < n, \forall\, i, j \right\}, \tag{3}$$

and $V = (\underline{v}_1, \ldots, \underline{v}_c)$ is the ordered collections of cluster centers.

In our study, the data matrix is composed of n images, each one with h values, associated to the corresponding features. The FCM algorithm produces a partitioning of this collection into a prefixed number c of clusters.

The algorithm finds an optimal fuzzy partition of the data, which is carried out through an iterative optimization of (1). Main steps are given as follows.

1. Choose the values c, m and a small positive constant ϵ; then, generate randomly a fuzzy c-partition U^0 and set iteration number $t = 0$.
2. Given the membership values $u_{i,k}^{(t)}$, the cluster centers $v_i^{(t)}$, $(i = 1, \ldots, c)$ are calculated by
$$v_i^{(t)} = \frac{\sum_{k=1}^{n} (u_{i,k}^{(t)})^m \quad x_k}{\sum_{k=1}^{n} (u_{i,k}^{(t)})^m} \tag{4}$$
3. Given the new centers $v_i^{(t)}$, update the membership value $u_{i,k}^{(t)}$:
$$u_{i,k}^{(t+1)} = \frac{1}{\sum_{j=1}^{c} \left(\frac{dist_{i,k}^2}{dist_{j,k}^2} \right)^{\frac{1}{m-1}}} \tag{5}$$
4. The process stops when $|U^{(t+1)} - U^{(t)}| < \epsilon$, otherwise go to step 2.

Let us note the only actual parameter of this algorithm is the number c of clusters. In general, this number is not known a priori. Selecting a different number of initial clusters can effectively affect the final partioning of the data. The problem for finding an optimal c is usually called cluster validity [3]. The objective is to find optimal c clusters that can validate the best description of the data structure. Each of these optimal c clusters should be compact and separated from other clusters. In the literature, many heuristic criteria have been proposed for evaluating fuzzy partitions; some of traditional cluster validity indexes, which have been frequently used, are Bezdek's partition coefficient (PC) [4], partition entropy (PE) [3], Xie-Beni's index [31].

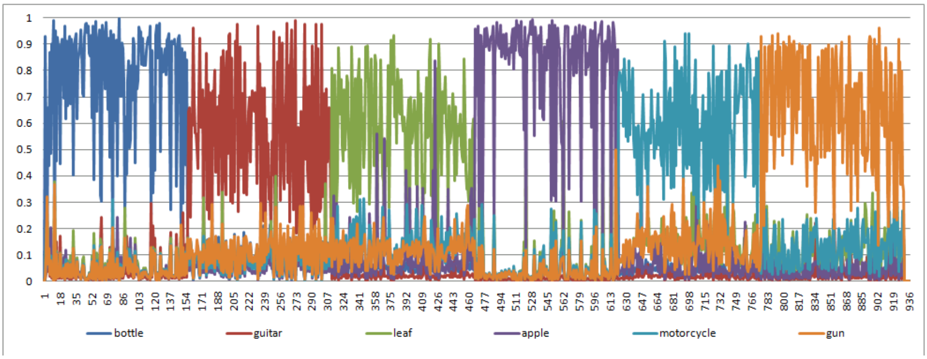

Fig. 1. The membership distribution among six clusters, produced by the FCM with $c = 6$ and $m = 2$

5 Experimental Results

The first experiment exploits a collection of images downloaded through Google images[1]. The testbed consists of a sample of 930 images, composed of six classes of 155 images, ranked as follows: images in the range 1-155 represent bottles; in the range 156-310 there are images of guitars, then the leaves are in the range 311-465, the images of apples cover the range 466-620, the motorcycle images are in 521-775 and finally the last images set consist of guns in the range 776-930.

The test considers all the features presented above: geometrical features (E,R, C, π_{gen}), invariants moments features (I_1, I_2, I_3, I_4, I_5, I_6, I_7), and affine moments features (AMI_1,AMI_2, AMI_3). Then, the FCM algorithm has been executed considering the number of clusters equals to the number of images categories ($c = 6$). The final partitioning is sketched in Figure 1: each line represents the membership distribution in a cluster; in particular, each cluster is in correspondence with a class of images. For instance, the blue line in Figure 1 describes the memberships distribution of a cluster that represents to the class of bottles (first 155 data). Due to the fuzzy approach, the individual image membership can be distributed among all the clusters and assume a value in the range [0,1] according to how it belongs to each cluster. The fuzzy method of clustering reveals more flexibility in the distribution of data: an image can belong to more than one cluster, because it shares similar characteristics with other

[1] The dataset can be downloaded at: `http://www.dmi.unisa.it/people/senatore/www/dati/dataset.rar`

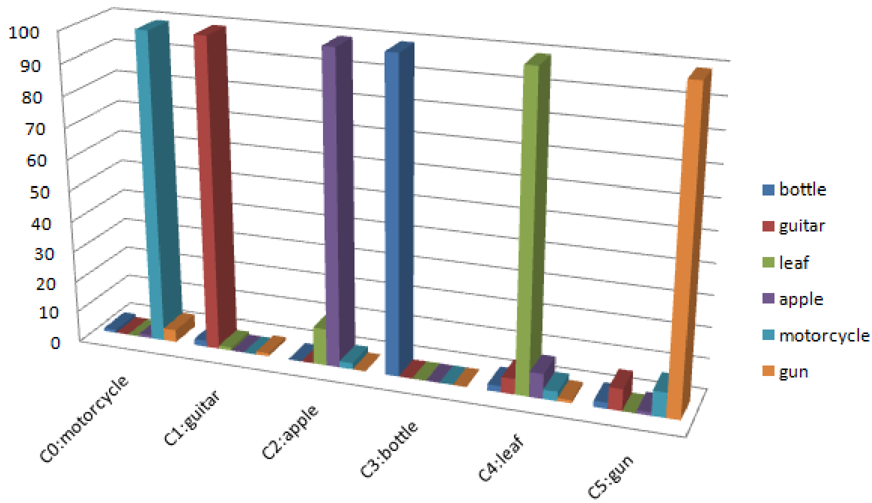

Fig. 2. The membership distribution among six clusters, associated to image classes

images, even though these latter belong to other clusters. It is licit to assume an image belongs to a given cluster, if its membership value for that cluster is the highest one. Figure 2 shows instead, a "synthetic" representation of this image distribution among the clusters, through histogram-based graphs. Each cluster represents a class of images. The clustering results are satisfying, because each class of images is almost completely individuated and associated to a cluster. In particular, in this specific testbed, classification error is quite restrained, as evidenced in Table 1, where the assessment of the clustering results is shown for each class/cluster. Each row provides the name of the class, the *misclassified* images, i.e. those images that have the highest membership in another class, different by the expected one, the *undecided* images, viz. all the images which membership is almost equally distributed among two or more clusters. Then the local *recall* and *precision* that is evaluated for each cluster.
More specifically, in the image retrieval context, the definition of recall and precision can be as follows:

$$Recall = \frac{relevant\ retrieved\ images}{relevant\ images} \tag{6}$$

$$Precision = \frac{relevant\ retrieved\ images}{retrieved\ images} \tag{7}$$

where the *relevant images* are the images which are expected in a certain class, the *retrieval images* are all the (correct and incorrect) images which are returned in that cluster, while the *relevant retrieved images* are just the images that really belong to the right cluster, associated to the correct class. Figure 1 reveals clusters associated to the leaves and motorcycle classes present the lowest membership distribution, even though the most of data are well placed in the cluster. In particular, let us analyze the class of leaves: most of misclassified data appear

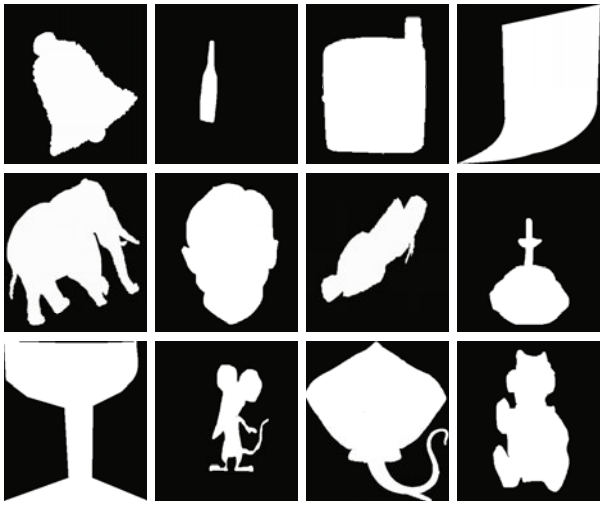

Fig. 3. Sample images representing classes of MPEG-7 CE data set used in the experiment

Fig. 4. Some samples used for the experiments. The entire dataset is composed of 930 images subdivided in six categories; bottles, leafs, guitars, motorcycles, guns, and apples.

in cluster of apples; this is due to the different shapes of leaves: after the image processing, some leaves present rounded shapes that can be easily confused with apples. In fact geometrical feature as Pi and compactness assume values assimilable to those ones of apples. Indeed, the image numbered 424 for instance, is misclassified presenting highest membership value in the apple cluster: its distribution among cluster is [0.027 0.007 0.048 0.836 0.043 0.035], respectively for the clusters associated to the bottle, guitar, leaf, apple, motorcycle and gun classes. It is evident its highest membership value 0.836 in apple cluster versus 0.048 of the right cluster. Anyway, no image of leaves cluster is undecided. In the cluster of motorcycles, instead, two images are undecided: the numbers 746 and 772 with distribution membership [0.057 0.038 0.276 0.044 0.290 0.293] and [0.145 0.170 0.180 0.082 0.208 0.212] respectively. In fact, the highest membership values appear equally distributed among the clusters of motorcycles and guns.

The lowest membership distribution in the cluster of leaves yields worse precision values. The recall is computed on 142 well-classified relevant images, considering all the 155 image of the class. The precision, instead is evaluated as ratio between the 142 well-classified images and all the retrieved images in this class, i.e. 160 images among correct and incorrect ones. Similar considerations can be done for the cluster of guns: here, the retrieved images are 163 even though the well classified images are 149. The overall result emphasizes the efficacy of this approach: the experiment can be considered satisfactory, because presents well-defined classes, composed of most of relevant images.

The next experiment considers a subset of the MPEG-7 Core Experiment Shape-1 dataset, which is frequently used to evaluate shape matching and recognition algorithms. In particular, we have used the MPEG-7 CE Shape-1 Part-B dataset [1], composed of 70 shape categories, each of which has 20 samples with in-plane rotations, articulations, and occlusions. MPEG-7 CE Shape-1 Part-B data set includes 1400 shape samples, 20 for each class. We have used twelve shape classes, considering all the twenty shape samples. We have chosen following twelve classes: bell, bottle, cellular phone, comma, elephant, face, fish, fountain, glasses, rat, ray, teddy, shown in Figure 3. The shape classes are very distinct, but the data set shows substantial within-class variations.

The fuzzy clustering setting considers just 12 clusters, one for each class of the presented images set (totally 240 images) and exploits all the features defined in Section 3. In other words, a 240 $\times$ 14 input matrix is given as input to FCM

Table 1. Class-based evaluation of fuzzy clustering results

Classes	*# Misclassified.*	*# Undecided.*	*Recall. %*	*Precision. %*
bottle	5	2	95	100
guitar	12	0	92	97
leaf	13	0	91	88
apple	9	1	93	91
motorcycle	11	2	91	95
gun	6	0	96	91

Table 2. Confusion matrix relative to a subset of MPEG-7 CE Shape-1 Part-B dataset

		Actual											
		bell	bottle	cellular phone	comma	elephant	face	fish	fountain	glass	rat	ray	teddy
Predicted	bell	20	0	0	0	0	0	0	0	0	0	0	0
	bottle	0	20	0	0	0	0	0	0	0	0	0	0
	cellular phone	0	0	19	0	0	0	0	0	0	0	0	0
	comma	0	0	0	15	0	0	0	0	0	0	7	0
	elephant	0	0	0	0	17	0	0	0	0	0	2	0
	face	0	0	0	0	0	20	0	0	0	0	0	0
	fish	0	0	1	0	0	0	20	0	0	0	0	0
	fountain	0	0	0	0	0	0	0	20	0	0	0	0
	glass	0	0	0	0	0	0	0	0	15	0	0	0
	rat	0	0	0	0	0	0	0	0	5	20	0	0
	ray	0	0	0	5	3	0	0	0	0	0	11	0
	teddy	0	0	0	0	0	0	0	0	0	0	0	20

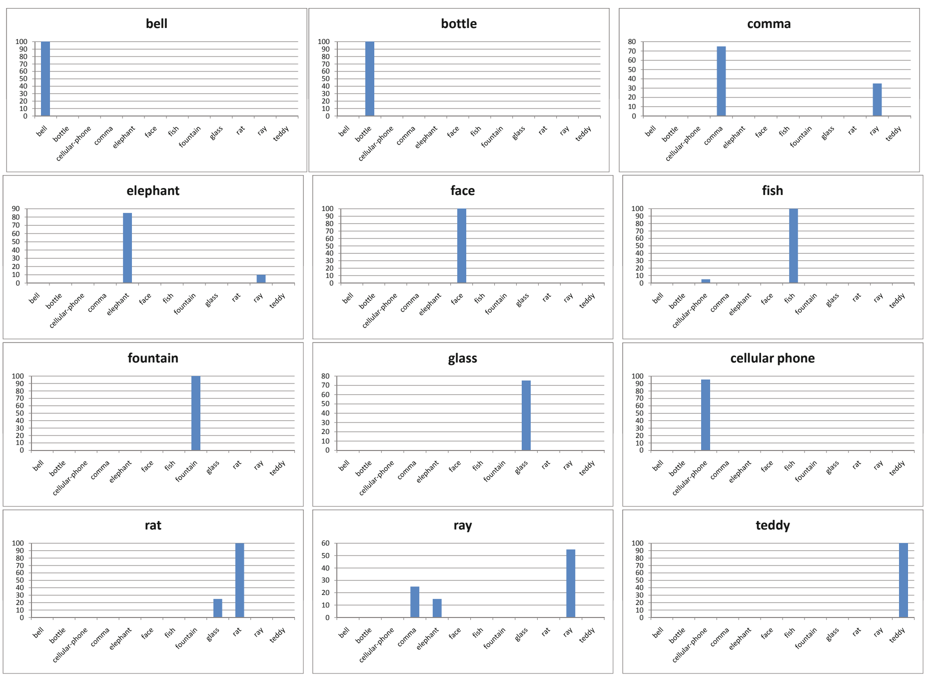

Fig. 5. Images distribution among the clusters produced by the FCM

algorithm. After the loading of these images and the image processing, the clustering phase has been started. Figure 5 shows the detailed classification results, where images belonging to different classes are allocated within the individual clusters. Table 2 synthetizes the results, showing the confusion matrix associated to this experiment. Let us note that many correspondences are revealed between the generated clusters (*predicted*) and the given (*actual*) images classes. In particular, some clusters look very homogeneous; most of them includes averagely about 80% of proper images. Just to give some example, the elements of the classes represented by fish, face, bottle, etc. appear all collocated in each individual cluster (100% of individuals are placed in each of them). This is not true any longer for the clusters concerning the rays and commas, even though we have a low overlap among categories. Finally, some clusters are representative of a specific class, even though elements of other class appear in them (in Figure 5: see the clusters representing the classes of rats, elephants, etc.).

6 Conclusion

The approach achieves an image classification and content-based retrieval. An initial image analysis allows the elicitation of visual features which are exploited to characterize the image through its shape. The fuzzy clustering techniques enable a relaxed distribution of images (compared to the crisp clustering); moreover they are robust respect to an image segmentation approaches based on k-means segmentation which meet some difficulties foreground and background colors do not contrast sufficiently. The effectiveness of this approach is evaluated through Information Retrieval measures, which reveals discrete performance.

This approach exploits a fuzzy clustering technique which, even though requires an a-priori fixed number of clusters, avoids overfitting of parameters and does not require a learning/training phase. In fact, our approach could be also considered non-parametric, if the only parameter, i.e., the number c of clusters is a-priori known. Otherwise, as said, methods based on cluster validity indexes [3], [31] find the optimal c and evaluate the fitness of partitions produced by clustering algorithms. Finally, the approach is robust respect to an image segmentation approach based on k-means segmentation which performs not very well when foreground and background colors do not contrast sufficiently.

Future extensions of this work foresee a development of a GUI-based application which supports the features extraction and the clustering technique. Additional features have been taken into account, particularly, some moments that are invariant to elastic transformations and convolution. We are going to extend the application, designing an visual query interface for the submission of a free hand drawing shape. This way, a ranked list of images whose shape is similar to the sketched one will be returned. Moreover, additional experiments with increased size and comparisons with other classification techniques have been taken into account.

References

1. Shape data for the mpeg-7 core experiment ce-shape-1, `http://www.cis.temple.edu/~latecki/TestData/mpeg7shapeB.tar.gz`
2. Aguilera, A., Suber, A., Martinez, L., Subero, A., Tineo, L.: Fuzzy image retrieval system. In: The 10th IEEE International Conference on Fuzzy Systems, vol. 3, pp. 1247–1250 (2001)
3. Bezdek, J.: Cluster validity with fuzzy sets. J. Cybernet. 3, 58–72 (1974)
4. Bezdek, J.: Numerical taxonomy with fuzzy sets. J. Math. Biol. 1-1, 57–71 (1974)
5. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, Norwell (1981)
6. Boiman, O., Shechtman, E., Irani, M.: In defense of nearest-neighbor based image classification. In: CVPR 2008: Proceedings of the Conference on Computer Vision and Pattern Recognition (2008)
7. Celebi, M.E., Aslandogan, Y.A.: A comparative study of three moment-based shape descriptors. In: ITCC 2005: Proceedings of the International Conference on Information Technology: Coding and Computing, pp. 788–793. IEEE Computer Society, Washington, DC, USA (2005)
8. Chen, Y., Wang, J.: A region-based fuzzy feature matching approach to content-based image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(9), 1252–1267 (2002)
9. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys 40(2), 1–60 (2008)
10. Flusser, J., Suk, T.: Pattern recognition by affine moment invariants. Pattern Recognition 26(1), 167–174 (1993)
11. Flusser, J., Zitova, B., Suk, T.: Moments and Moment Invariants in Pattern Recognition. Wiley, Chichester (2009)
12. Frigui, H.: Membershipmap: Data transformation based on granulation and fuzzy membership aggregation. IEEE Transactions on Fuzzy Systems 14(6), 885–896 (2006)
13. Gao, Y., Fan, J.: Semantic image classification with hierarchical feature subset selection. In: MIR 2005: Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval, pp. 135–142. ACM, New York (2005)
14. Gevers, T., Smeulders, A.W.M.: Image search engines - an overview. In: Medioni, G., Kang, S.B. (eds.) Emerging Topics in Computer Vision. Prentice-Hall, Englewood Cliffs (2004)
15. Ha, J.-Y., Kim, G.-Y., Choi, H.-I.: The content-based image retrieval method using multiple features. In: NCM 2008: Proceedings of the 2008 Fourth International Conference on Networked Computing and Advanced Information Management, pp. 652–657. IEEE Computer Society, Washington, DC, USA (2008)
16. Hall, L., Bensaid, A., Clarke, L., Velthuizen, R., Silbiger, M., Bezdek, J.: A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. IEEE Transactions on Neural Networks 3(5), 672–682 (1992)
17. Hu, M.-K.: Visual pattern recognition by moment invariants. IRE Transactions on Information Theory 8(2), 179–187 (1962)
18. Images, Z.C., Zhang, M., Hall, L.O., Muller-karger, F.E., Goldgof, D.B.: Knowledge-guided classification of coastal. International Journal of Pattern Recognition and Artificial Intelligence 14, 987–1007 (2000)

19. Jingwei, S.E., Eschrich, S., Ke, J., Hall, L.O., Goldgof, D.B.: Fast fuzzy clustering of infrared images. In: Joint 9th IFSA World Congress and 20th NAFIPS International Conference (2001)
20. Krishnapuram, R., Medasani, S., Jung, S.-H., Choi, Y.-S., Balasubramaniam, R.: Content-based image retrieval based on a fuzzy approach. IEEE Transactions on Knowledge and Data Engineering 16(10), 1185–1199 (2004)
21. Kulkami, S., Verma, B., Sharma, P., Selvaraj, H.: Content based image retrieval using a neuro-fuzzy technique, vol. 6, pp. 4304–4308 (July 1999)
22. Lay, J.A., Guan, L.: Image retrieval based on energy histograms of the low frequency dct coefficients. In: ICASSP 1999: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 3009–3012. IEEE Computer Society, Washington, DC, USA (1999)
23. Lieberman, H., Rozenweig, E., Singh, P.: Aria: An agent for annotating and retrieving images. Computer 34, 57–62 (2001)
24. Lowe, D.G.: Object recognition from local scale-invariant features. In: IEEE International Conference on Computer Vision, vol. 2, pp. 1150–1157 (August 1999)
25. Nachtegael, M., der Weken, D.V., Witte, V.D., Schulte, S., Mélange, T., Kerre, E.E.: Color image retrieval using fuzzy similarity measures and fuzzy partitions. In: ICIP (6), pp. 13–16 (2007)
26. Pengyu, L., Kebin, J., Peizhen, Z.: An effective method of image retrieval based on modified fuzzy c-means clustering scheme, vol. 3 (November 2006)
27. Rath, T.M., Manmatha, R., Lavrenko, V.: A search engine for historical manuscript images. In: SIGIR 2004: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 369–376. ACM, New York (2004)
28. Rui, Y., Huang, T.S., Chang, S.-F.: Image retrieval: Current techniques, promising directions, and open issues. Journal of Visual Communication and Image Representation 10(1), 39–62 (1999)
29. Tian, Y., Mei, L.: Image retrieval based on multiple features using wavelet. In: ICCIMA 2003: Proceedings of the 5th International Conference on Computational Intelligence and Multimedia Applications, p. 137. IEEE Computer Society, Washington, DC, USA (2003)
30. Vertan, C., Boujemaa, N.: Embedding fuzzy logic in content based image retrieval, pp. 85 –89 (2000)
31. Xie, X., Beni, G.: A validity measure for fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence 13, 841–847 (1991)
32. Yao, W., Hall, L., Goldgof, D., Muller-Karger, F.: Finding green river in seawifs satellite images. In: International Conference on Pattern Recognition, vol. 2, p. 2307 (2000)
33. Zhang, D.: Improving image retrieval performance by using both color and texture features. In: ICIG 2004: Proceedings of the Third International Conference on Image and Graphics, pp. 172–175. IEEE Computer Society, Washington, DC, USA (2004)
34. Zhang, D., Lu, G.: Evaluation of mpeg-7 shape descriptors against other shape descriptors. Multimedia Syst. 9(1), 15–30 (2003)