Fuzzy shape Classification exploiting Geometrical and Moments Descriptors

Ugo Erra
Dipartimento di Matematica e Informatica,
Università della Basilicata,
Viale Dell'Ateneo 10, Macchia Romana
85100, Potenza, Italy
email: ugo.erra@unibas.it

Sabrina Senatore
Dipartimento di Informatica
Università degli Studi di Salerno
Via Ponte Don Melillo
84084 Fisciano(SA), Italy
email: ssenatore@unisa.it

***Abstract*—In the era of data intensive management and discovery, the volume of images repositories requires effective means for mining and classifying digital image collections. Recent studies have evidenced great interest in image processing by "mining" visual information for objects recognition and retrieval. Particularly, image disambiguation based on the shape produces better results than traditional features such as color or texture. On the other hand, the classification of objects extracted from images appears more intuitively formulated as a shape classification task.**

This work introduces an approach for 2D shapes classification, based on the combined use of geometrical and moments features extracted by a given collection of images. It achieves a shape-based classification exploiting fuzzy clustering techniques, which enable also a query-by-image.

***Index Terms*—fuzzy clustering, image data mining, image retrieval.**

I. Introduction

The evolution of multimedia computing and applications has seen an explosive growth of digital images in many application domains. This development has remarkably increased the need for image recognition management and image retrieval systems that are able to effectively classify a large amount of images and to efficiently retrieve them based on their visual contents. In developing a visual content-based image retrieval system, two different approaches have been applied: one based on textual information, the other based on image content information. The choice of image features, or combinations of image features to be used for image indexing and retrieval purposes is a critical decision for the efficacy of the adopted approach.

The retrieval based on textual information considers attached textual metadata to each image in the database, then a keyword-based query may be submitted to the system to get relevant images [1]. This approach requires a preliminary annotation activity which often is laborious and time-consuming; moreover, it is a human driven process: that means similar images characteristics can be expressed by different terminologies, according to the subjective evaluation and feeling of the user, affecting the performance of the keyword-based image search. 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 instead characterize an image through a set of features; retrieval or classification is then performed by measuring similarity to a required query image [2] contrasting to the effort needed to annotate images. Because images can be particularly complex to manage; CBIR techniques often require a 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.

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 and non-parametric classifiers. As been pointed out in [3], 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.

This work presents an approach for image classification and retrieval based on 2D shape features and extends a previous work [4] for further investigations on shape classification, through fuzzy clustering techniques.

The remainder of the paper is organized as follows. Section II gives a sketched overview of the related works in this area. Section III provides a background on the image processing for the image analysis and features extraction whereas Section IV-A introduces the fuzzy clustering algorithms exploited in this approach. Finally, Section V describes the experiments and provides the results and comparisons. Conclusions and future works close the paper.

II. Related Works

The large repository of digital information needs enhanced image acquisition technologies and computer-assisted image understanding tools in order to get an efficient assistance in image processing, query and retrieval. CBIR aims at efficient retrieval of relevant images form large databases exploiting automatically derived image features [5]. Popular CBIR systems are QBIC, Virage, RetrievalWare, Photobook, Chabot, VisualSeek, WebSeek, MARS system, SurfImage, Netra, and CANDID (for a complete survey, refer to[6]); almost all of these approaches are based on indexing imagery in a feature space. A feature is 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. These features are typically extracted from color, texture, shape and region. Particularly, shapes can be represented by their contours, i.e. a potentially large number of points from their boundaries. Shape descriptors have been used widely as features to characterize an image for classification and image retrieval tasks, such as character recognition [7], word recognition [8] and copyrighted trademark retrieval [7][9]. Like for the texture features, shape features can be used to distinguish between objects where the color feature reveals an unreliable classification due to different objects in image representation can have the same color.

Also the increasing diffusion of images compression requires challenging techniques to extract visual features [10]: sophisticated global features such as the wavelets [11] and large collection of local image descriptors as SIFT [12].

Some other techniques improve the effectiveness of image retrieval through multi-features combination [13],[14] and then, by measuring similarity to a required query image [2]. Combination of words and features characterize annotated training sets of images [8], which will be used for classification or retrieval.

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 linguistic expression by means of fuzzy rather than crisp features values [15].

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 [16] where a fuzzy information retrieval model for textual data has been extended to implement a model in image context. In [17], 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 [18] allows the evaluation of similarity through fuzzy logic-based operations. In [19], 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 [20], specifically, fuzzy techniques are widely employed in image classification methods. In [21], 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 [22] has been shown to provide effective partitions for image segmentation on medical images [23][24], satellite images [25][26], and so on. Some extension and modification of FCM are applied to image segmentation in infrared images domain [27]. In [28] 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 [27].

III. 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 labeled 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 [29] 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 [29], 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[30]. 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 [31] and given as:

$$\begin{aligned}
AMI_1 &= (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4 \\
AMI_2 &= (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30} + \mu_{12}^3 + \\
&\quad 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^3\mu_{12}^3)/\mu_{00}^{10} \\
AMI_3 &= (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \\
&\quad \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2))/\mu_{00}^7
\end{aligned}$$

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 [32].

IV. Unsupervised clustering techniques for the image arrangement

The clustering algorithms achieves 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 separation between clusters (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 (as described previously) to pull out a such data matrix, whose rows and columns are respectively the collected images and the relative extracted features. Two clustering algorithms have been take into account: the well-known fuzzy C-Means (briefly FCM) algorithm [22] which takes as input the whole data set of images and the approach presented in [33] that works with more than one (sub-)set of data, trying to reconcile them in one complete final partitioning.

A. Fuzzy clustering

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 form of matrix and produces fuzzy partitions of the given patterns into (prefixed) c clusters.
The FCM algorithm recognizes spherical clouds of points (clusters) in a multi-dimensional 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:

$$Q(U,c) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{i,k}^{m}(dist(\underline{x}_k,\ \underline{v}_i))^2 \quad (1)$$

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 $dist(\underline{x}_k,\ \underline{v}_i)$ is:

$$dist(\underline{x}_k,\ \underline{v}_i) = \sqrt{(||\underline{x}_k - \underline{v}_i||^2)} \quad (2)$$

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:

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

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).

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 partitioning of the data. The problem for finding an optimal c is usually called cluster validity [34]. 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) [35], partition entropy (PE) [34], Xie-Beni's index [36].

B. Collaborative clustering

The clustering presented in [33] proposes an approach which takes as input two or more subsets of the reference data domain. It achieves a collaborative and proximity-based mechanism of clustering through a collaborative and highly synergistic discovery of structure in data. The collaborative clustering exploits FCM clustering to get a thorough arrangement of data, according to their different typologies of features. In other words, the subsets of features identify collections of descriptors which represent different views of the interest domain, i.e., the structure at different levels of granularity (clusters), providing a scheme of incorporating collaboration. At the end, a collection of homogeneous clusters are given that often better group the data.

More formally, let us consider a collection of patterns (objects) X whose features are described by distinct feature spaces, say $X[1], X[2], \ldots, X[p]$. Let us note X is composed by concatenating all the sub-datasets $X = X[1]X[2] \ldots X[p]$. The approach uses FCM as basic clustering technique and minimizes a objective function given in (1). As said the final outcome of this clustering is a partition matrix that reveals all details about the structure of the data set.

The collaborative clustering achieves two levels of processing:

- numeric: it involves original data sets. FCM clustering is applied to each feature spaces. The approach assumes the number of the clusters by $c_1, c_2, \ldots, c_p$ are different values. The final results are a set of partition matrices $U[1], U[2], \ldots, U[p]$.
- granular: it is based upon membership grades computed through the partition matrices. As the comparison among partition matrices is not possible (they are built with different number of the clusters), the proximity matrices are computed. Let us recall, that for two patterns x_i and x_j the proximity measure returns a number in [0,1] such that it is equal to 1 if $i = j$. Moreover, this measure is symmetric.

  Given a partition matrix $U[k], (k = 1, \ldots, p)$ the proximity between pattern "i" and "j" based on the values of their corresponding entries is [37]:

  $$Prox[k](i,j) = \sum_{l=1}^{c} min(u(l,i), u(l,j)) \quad (4)$$

  The results of final consensus are reached through reconciling proximities based upon the clusters.

Thus, the performance index to optimize is described by the following formula:

$$V = \sum_{k=1}^{p}\sum_{i=1}^{N}\sum_{j=1}^{N} [Prox(U)(i,j) - Prox[k](i,j)]^2 \quad (5)$$

where $Prox(U)$ is the proximity matrices computed by Eq. (4) on the whole data set $X(= X[1]X[2] \ldots X[p])$.

The optimization of V is carried out using a standard gradient-based mechanism, thus the minimization of V proceeds over successive iterations until when the difference between two successive partition matrices is insignificant.

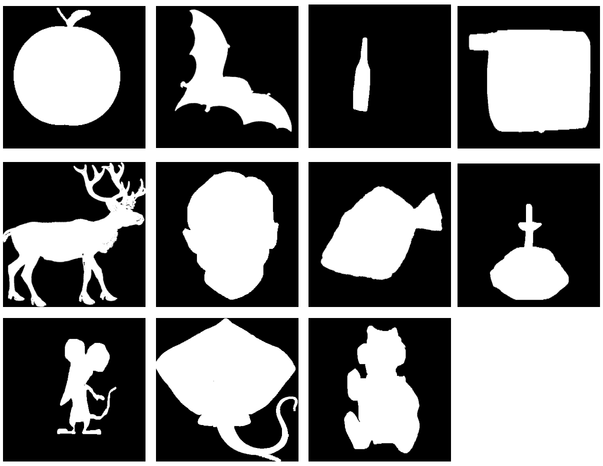

Fig. 1. Sample images representing classes of MPEG-7 CE data set used in the experiment

C. Query by Image

Query by image is a technique that involves providing the CBIR system with an example image that it will then base its search upon. The underlying search approach may vary depending on the application, but result images should all share common descriptors with the provided example. There are essentially two options in providing image: a) the user chooses an image from a given data set. b) the user draws a sketch of the image they are looking for. Both options simplifies the process of describe images with words which can arise in several mismatches during the search.

We explored query by image using as example image a shape draws by the user. For each shape, we compute the shape descriptor associated with it using the approach described in this section. Formally, given a shape descriptor s, a number c of centers v_i, and a fuzziness threshold ϵ, we use the following criteria to obstain the cluster membership. The potential cluster i is found by taking into account the minimal distance center with the shape descriptor, that is

$$i = min_{1 \leq j \leq c} \{j | dist(s, v_j)\}$$

A reference data x point belonging to the cluster i and immediately below the fuzziness threshold is obtained as

$$x_i = min_{1 \leq k \leq n} \{x_k | u_{k,i} \geq \epsilon\}$$

where, n is the size of cluster i. Then, if $dist(v_i, s) \leq dist(v_i, x_i)$, the query is in the space computed as a sphere, with radius equals to $dist(v_i, x_i)$. Then, we return all the images contained in the cluster i sorted based on euclidean distance and whose membership is above the fuzziness threshold. This ranked list represents the answer to the given query.

V. Experimental results

The testing phase consider three datasets. The first 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 [38], composed of 70 shape categories, each of which has 20 samples with in-plane rotations, articulations and occlusions. We have used eleven shape classes, considering all the twenty shape samples. The shape classes are: apple, bat, bottle, cellular phone, deer, face, flatfish, fountain, rat, ray, teddy, shown in Figure 1. The shape classes are very distinct, but the data set shows some within-class variations. All the fourteen features presented above has been exploited in the experiment: 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).

The 220×14 data matrix has been given as input to FCM. Further, we would like to get a direct association between cluster and class (i.e., each cluster is representative of a class), thus, the number of clusters is equal to the number of images categories ($c = 11$).

Then, according to the natural grouping of features, three data matrices have been built: one has the size 220×4,including only the geometrical features; a 220×7 matrix describing the invariant moments and then a 800×3 matrix representing the affine moments. These matrices have been given as an input to the collaborative clustering, setting the number of clusters c_1, c_2, c_3 to 4, 5, 3, respectively. Also in this setting configuration, $c = 11$, i.e., the number of images categories.

In the fuzzy clustering techniques, the membership of an image may be distributed on more than one clusters. Usually, an image (in general, a pattern) is associated to the cluster in which its membership is the highest. In order to emphasize the natural flexibility of the fuzzy clustering methods in the distribution of data, we have processed the dataset on crisp k-means too.

Table I synthesizes the results, showing the confusion matrix associated to this experiment for the three clustering algorithms. Each cell indeed, contains the results computed for k-means, FCM and collaborative clustering algorithms, respectively. A single value zero means all the results are zero. Let us note that many correspondences are revealed between the generated clusters (*predicted*) and the given (*actual*) images classes. The worst case is computed exploiting k-means: some classes are not well-identified; for instance the class *deer* is composed of only eight images, while the most of its images is placed in the class *ray*. Even, class *cellular phone* is completely empty and most of its images is within the class *teddy*. This emphasizes that the k-means is not suitable to cluster these data. Differently, it is evident how the FCM and collaborative provides meaningful results. In fact, some clusters look very homogeneous; most of them includes averagely about 80% of proper images. Just to give some example, on the FCM clustering, the elements of the classes represented by *bat, bottle, deer, face* are well-placed in each individual cluster (100% of individuals appear in the right cluster). This is not true any longer for the cluster concerning the class *ray*, due to some overlaps among categories. Finally, some clusters are representative of specific classes, even though a certain percentage of elements of other classes appears in them (for instance, see the clusters representing the classes *apple, flatfish*, etc.). The classification produces improvements, exploiting the collaborative clustering, where the "uncertain" classes such as *apple* and *cellular phone* are almost completely identified.

Other two experiments have been executed on a collection of images downloaded through Google images[1]. Fig. 2 shows a sample of ten images for each considered category. The first testbed consists of a sample of 800 images, composed of five classes of about 160 images, ranked as follows: images in the range 1-160 represent bottles; in the range 161-320 there are images of guitars, then the images of apples cover the range 321-480, the motorcycles images are in 481-640 and finally the last images set consists of guns in the range 641-800. The dataset has been given as an input to the three clustering algorithms. All the clustering algorithm set the number of cluster equals to the number of images categories ($c = 5$). The k-means and the FCM work on all the features set considering a 800×14 matrix , while the collaborative clustering takes as an input three submatrices: one has the size 800×4, including only the geometrical features; a 800×7 matrix describing the invariant moments and then a 800×3 matrix representing the affine moments. The setting the number of clusters is $c_1 = 4, c_2 = 5, c_3 = 4$.

TABLE II
800DATASET: CLUSTER-BASED EVALUATION OF CLUSTERING RESULTS, FOR K-MEANS, FCM AND COLLABORATIVE CLUSTERING, RESPECTIVELY

Classes	# Misclassified.	# Undecided.	Recall. %	Precision. %
bottle	0-8-7	0-2-2	86-95-95	95-100-100
guitar	0-3-3	0-0-0	95-98-98	95-93-96
apple	0-12-12	0-0-0	89-92-94	97-100-100
motorcycle	0-5-4	0-1-0	91-96-97	77-88-90
gun	0-13-10	0-3-2	91-91-93	93-93-95

TABLE III
965DATASET: CLUSTER-BASED EVALUATION OF CLUSTERING RESULTS, FOR K-MEANS, FCM AND COLLABORATIVE CLUSTERING, RESPECTIVELY

Classes	# Misclassified.	# Undecided.	Recall. %	Precision. %
bottle	21-6-3	0-0-1	86-96-97	93-100-100
guitar	7-3-3	0-0-1	95-98-98	95-96-97
leaf	92-58-45	0-7-5	44-64-72	41-78-81
apple	94-16-14	0-2-0	41-90-91	79-92-94
motorcycle	17-22-17	0-5-4	89-86-90	57-70-78
gun	13-14-10	0-3-3	91-91-93	92-93-96

Table II shows a synthetic view of the results computed for the three algorithms. Each row provides the name of the class, the *misclassified* images, i.e. those images that appear to belong to a cluster, different by the expected one (i.e., they have the highest membership in another class), the *undecided* images, viz. all the images which membership is almost equally distributed among two or more clusters. Recall that, due to the crisp nature of the k-means algorithm, there are no resulting *undecided* images. Then a *recall* and *precision* that is evaluated for each cluster.
In fact, in order to evaluate the recall and precision inside each cluster, we define the *local recall* and the *local precision* 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.

From the analysis of resulting partitioning, let us note that the collaborative clustering tends to get better results in terms of recall and precision. Nevertheless, the cluster associated to the class *motorcycle* presents the lowest membership distribution, even though the most of images are well placed in the right cluster.

Last test considers the same dataset of 800 images, but adds 165 images representing *leaf*. We have choice deliberately a set composed of very heterogeneous image samples of leaf, in order to evaluate the performance of FCM and collaborative clustering algorithms in this case study. Table III shows the results for the three clustering algorithms. The introduction of the set of leaf makes the resulting partition more confused: most of misclassified data appear 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, values computed for geometrical features such as Pi and compactness are close to those ones of apples. The best performance are obtained exploiting the collaborative clustering even though the recall and precision values, computed for the class *leaf* are quite low, anyway.

VI. CONCLUSION

The work presents an image classification and content-based retrieval. After an initial phase of image analysis for extracting visual shape-based features, fuzzy clustering techniques have been applied. These techniques provide 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 recall and precision, revealing discrete performance.

Let us observe that the use of fuzzy clustering techniques which, even though requires an a-priori fixed number of clusters, avoid overfitting of parameters and does not require a learning/training phase. Moreover an image retrieval technique has been proposed through a query-by-image. As future extensions of this work we would investigates in this direction as well as introduce additional features particularly, some moments that are invariant to elastic transformations and convolution.

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

TABLE I

CONFUSION MATRIX RELATIVE TO A SUBSET OF MPEG-7 CE SHAPE-1 PART-B DATASET FOR K-MEANS, FCM AND COLLABORATIVE CLUSTERING ALGORITHMS, RESPECTIVELY.

		ACTUAL										
		apple	bat	bottle	cell_phone	deer	face	flatfish	fountain	rat	ray	teddy
PREDICTED	apple	17-18-19	0	0	2-1-0	0	0-0-0	0	0	0	0	0
	bat	0	18-20-20	0	0	2-0-0	0	0	0	0	0	0
	bottle	0	0	11-20-20	0	0	0	0	0	0	0	0
	cell_phone	0-0-0	0	0	0-19-20	0	0	0	0	0	0	0
	deer	0	0	0	0	8-20-20	0	0	0	0	0	0
	face	3-0-0	0	0	0	0	20-20-20	2-2-0	0	1-0-0	0	0
	flatfish	0	0	0	0	0	0	17-18-20	0	0	8-2-2	0
	fountain	0	0	0	5-0-0	0	0	0	20-20-20	0	0	0
	rat	0	2-0-0	9-0-0	0	0	0	0	0	19-20-20	2-7-7	0
	ray	0	0	0	0	10-0-0	0	1-0-0	0	0	10-11-11	0
	teddy	0-2-1	0	0	13-0-0	0	0	0	0	0	0	20-20-20

Fig. 2. 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.

REFERENCES

[1] H. Lieberman, E. Rozenweig, and P. Singh, "Aria: An agent for annotating and retrieving images," *Computer*, vol. 34, pp. 57–62, 2001.

[2] D. Zhang and G. Lu, "Evaluation of mpeg-7 shape descriptors against other shape descriptors," *Multimedia Syst.*, vol. 9, no. 1, pp. 15–30, 2003.

[3] O. Boiman, E. Shechtman, and M. Irani, "In defense of nearest-neighbor based image classification," in *CVPR'08: Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2008.

[4] U. Erra and S. Senatore, "Experiences with shape classification using fuzzy c-means with geometrical and moments description," in *Proceedings of 8th International Workshop on Adaptive Multimedia Retrieval*, ser. Lecture Notes in Computer Science (In press). Springer Berlin / Heidelberg, 2010.

[5] T. Gevers and A. W. M. Smeulders, "Image search engines - an overview," in *Emerging Topics in Computer Vision*. G. Medioni and S. B. Kang, Eds. Englewood, Cliffs, NJ: Prentice-Hall, 2004.

[6] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, no. 2, pp. 1–60, 2008.

[7] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, 2002.

[8] T. M. Rath, R. Manmatha, and V. Lavrenko, "A search engine for historical manuscript images," in *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: ACM, 2004, pp. 369–376.

[9] A. Jain and A. Vailaya, "Shape-based retrieval: A case study with trademark image databases," *Pattern Recognition*, vol. 31, pp. 1369–1390, 1998.

[10] J. A. Lay and L. Guan, "Image retrieval based on energy histograms of the low frequency dct coefficients," in *ICASSP '99: Proceedings of the Acoustics, Speech, and Signal Processing, 1999. on 1999 IEEE International Conference*. Washington, DC, USA: IEEE Computer Society, 1999, pp. 3009–3012.

[11] Y. Tian and L. Mei, "Image retrieval based on multiple features using wavelet," in *ICCIMA '03: Proceedings of the 5th International Conference on Computational Intelligence and Multimedia Applications*. Washington, DC, USA: IEEE Computer Society, 2003, p. 137.

[12] D. G. Lowe, "Object recognition from local scale-invariant features," *Computer Vision, IEEE International Conference on*, vol. 2, pp. 1150–1157 vol.2, August 1999. [Online]. Available: http://dx.doi.org/10.1109/ICCV.1999.790410

[13] J.-Y. Ha, G.-Y. Kim, and H.-I. Choi, "The content-based image retrieval method using multiple features," in *NCM '08: Proceedings of the 2008 Fourth International Conference on Networked Computing and Advanced*

Information Management. Washington, DC, USA: IEEE Computer Society, 2008, pp. 652–657.
[14] D. Zhang, "Improving image retrieval performance by using both color and texture features," in *ICIG '04: Proceedings of the Third International Conference on Image and Graphics*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 172–175.
[15] R. Krishnapuram, S. Medasani, S.-H. Jung, Y.-S. Choi, and R. Balasubramaniam, "Content-based image retrieval based on a fuzzy approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 10, pp. 1185–1199, 2004.
[16] A. Aguilera, A. Suber, L. Martinez, A. Subero, and L. Tineo, "Fuzzy image retrieval system," in *The 10th IEEE International Conference on Fuzzy Systems*, vol. 3, 2001, pp. 1247 –1250.
[17] S. Kulkami, B. Verma, P. Sharma, and H. Selvaraj, "Content based image retrieval using a neuro-fuzzy technique," in *Neural Networks, 1999. IJCNN '99. International Joint Conference on*, vol. 6, jul 1999, pp. 4304 –4308 vol.6.
[18] C. Vertan and N. Boujemaa, "Embedding fuzzy logic in content based image retrieval," in *Fuzzy Information Processing Society, 2000. NAFIPS. 19th International Conference of the North American*, 2000, pp. 85 –89.
[19] Y. Chen and J. Wang, "A region-based fuzzy feature matching approach to content-based image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1252 – 1267, sep 2002.
[20] H. Frigui, "Membershipmap: Data transformation based on granulation and fuzzy membership aggregation," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 885 –896, dec. 2006.
[21] M. Nachtegael, D. V. der Weken, V. D. Witte, S. Schulte, T. Mélange, and E. E. Kerre, "Color image retrieval using fuzzy similarity measures and fuzzy partitions," in *ICIP (6)*, 2007, pp. 13–16.
[22] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers, 1981.
[23] L. Hall, A. Bensaid, L. Clarke, R. Velthuizen, M. Silbiger, and J. Bezdek, "A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 672 –682, sep 1992.
[24] H. Zhou, G. Schaefer, A. Sadka, and M. E. Celebi, "Anisotropic mean shift based fuzzy c-means segmentation of skin lesions," in *Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology*, ser. CSTST '08. New York, NY, USA: ACM, 2008, pp. 438–443.
[25] Z. C. Images, M. Zhang, L. O. Hall, F. E. Muller-karger, and D. B. Goldgof, "Knowledge-guided classification of coastal," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, pp. 200–0, 2000.
[26] W. Yao, L. Hall, D. Goldgof, and F. Muller-Karger, "Finding green river in seawifs satellite images," *International Conference on Pattern Recognition*, vol. 2, p. 2307, 2000.
[27] S. E. Jingwei, S. Eschrich, J. Ke, L. O. Hall, and D. B. Goldgof, "Fast fuzzy clustering of infrared images," in *Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, 2001.
[28] L. Pengyu, J. Kebin, and Z. Peizhen, "An effective method of image retrieval based on modified fuzzy c-means clustering scheme," in *Signal Processing, 2006 8th International Conference on*, vol. 3, nov 2006.
[29] M.-K. Hu, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 179 –187, february 1962.
[30] M. E. Celebi and Y. A. Aslandogan, "A comparative study of three moment-based shape descriptors," in *ITCC '05: Proceedings of the International Conference on Information Technology: Coding and Computing*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 788–793.
[31] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognition*, vol. 26, no. 1, pp. 167–174, 1993.
[32] J. Flusser, B. Zitova, and T. Suk, *Moments and Moment Invariants in Pattern Recognition*. Wiley, December 2009.
[33] V. Loia, W. Pedrycz, and S. Senatore, "Semantic web content analysis: A study in proximity-based collaborative clustering," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 6, pp. 1294–1312, 2007.
[34] J. C. Bezdek, "Cluster validity with fuzzy sets," *Journal of Cybernetics*, vol. 3, p. 5872, 1974.
[35] ——, "Numerical taxonomy with fuzzy sets," *Journal of Mathematical Biology*, vol. 1-1, pp. 57–71, 1974.
[36] X. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 841–847, 1991.
[37] J. C. Bezdek and J. D. Harris, "Fuzzy partitions and relations: an axiomatic basis for clustering," *Fuzzy Sets and Systems*, vol. 1, no. 2, pp. 111 – 127, 1978.
[38] "Shape data for the mpeg-7 core experiment ce-shape-1." http://www.cis.temple.edu/ latecki/TestData/mpeg7shapeB.tar.gz.