Fuzzy Clustering With Partial Supervision in Organization and Classification of Digital Images

Witold Pedrycz, *Fellow, IEEE*, Alberto Amato, Vincenzo Di Lecce, and Vincenzo Piuri, *Fellow, IEEE*

***Abstract*—In a Web-oriented society, organization, retrieval, and classification of digital images have become one of the major endeavors. In this paper, we study the mechanisms of fuzzy clustering and fuzzy clustering with partial supervision in the analysis and classification of images. It is demonstrated that the main features of fuzzy clustering become essential in revealing the structure in a collection of images and supporting their classification. The discussed operational framework of fuzzy clustering is realized by means of fuzzy c-means (FCM). When dealing with the mode of partial supervision, we augment an original objective function guiding the clustering process by an additional component expressing a level of coincidence between the membership degrees produced by the FCM and class allocation supplied by the user(s). The study also contrasts the use of the technology of fuzzy sets in image clustering with other approaches studied in this area. A suite of experiments deals with two collections of images, namely, Columbia object image library (COIL-20) and a database composed of 2000 outdoor images.**

***Index Terms*—Angular spectrum signature, fuzzy c-means (FCM), fuzzy clustering, human-centric systems, image classification, partial supervision, relevance feedback, separability index.**

I. Introduction

ORGANIZING, retrieving, and classifying digital images has become one of the main endeavors in the age of digital information. With the inception of digital media and their intensive sharing over the Internet, building digital albums with effective search capabilities is an important goal. Digital albums are an important category of human-centric systems. Human centricity is profoundly present there—no matter how much the entire process could be automated, there is a pivotal role of a human user. The individual plays a key role both as the user and a designer of the system. Digital albums need to be personalized: there should be a mechanism of adjustment of its structure to the individual needs of the user. As each user could exhibit different preferences as to the same collection of images, this feature becomes important and may very much influence an opinion of the users as to the performance of the system. At the same time, while the relevance feedback is important, there should be efficient implementations of this feedback mechanism, not to mention a comprehensive suite of algorithms that support an automatic process of image annotation and a formation of the overall structure.

Bearing in mind the human-centric aspect of the entire problem, it is noticeable that the potential of fuzzy set technology has been already considered in many instances (cf. [15], [16], and [38]). It is very likely that the efforts within this domain will be continued in the future. Alluding to the observations being made, there are a number of compelling reasons behind further developments of the technology of fuzzy sets in the realm of image classification and interpretation.

1) Classes of images are inherently fuzzy. No matter how much attention has been paid to the specification of the initial taxonomy, images could easily belong to several categories with some nonzero degrees. For instance, for some outdoor categories of images (say, vehicles, landscape, aircrafts), we may have photos that include all of the components (a vehicle on a rural road with some farmland in the background and a small airport with an airplane on a tarmac).
2) Users of digital albums play an active and constructive role in the organization process-–they arrange photos according to their preferences. Any system has to reflect their preferences; namely, there should be some efficient adjustment process coming in the form of some relevance feedback that must override or collaborate with some existing automatic mechanisms of image annotation and organization. This facet of user centricity is pivotal in many intelligent systems where an active and creative role of humans in a well-established communication loop becomes a must.

In this study, to accommodate the relevance feedback while realizing the processes of an automatic development of the categories, we consider fuzzy clustering with partial supervision as originally introduced in [28] and [29]. Given its main features, we show that it becomes a highly relevant vehicle that meets the requirements of the problem formulated here.

The functional facets of the pursuit are schematically outlined in Fig. 1. The phase of feature formation leading to the formal description of images comes with enormous diversity and is very much essential to the ensuing processes of clustering and classification [30], [32], [37]. In this study, we consider one of the existing alternatives of a so-called angular spectrum signature, which has been found to be a viable and effective option leading to a fairly compact image characterization [5], [7], [9]. The process of feature formation is then followed by some standard mechanisms of feature reduction [realized here with the use of the principal component analysis (PCA)]. The organization of images is completed in an unsupervised mode realized

Manuscript received October 12, 2005; revised March 31, 2006 and September 22, 2006; accepted January 14, 2007.

W. Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2G7, Canada and also with the Polish Academy of Sciences, Systems Research Institute, Warsaw PL-01-447, Poland (e-mail: pedrycz@ece.ualberta.ca).

A. Amato and V. Di Lecce are with the Polytechnic of Bari, Bari 70125, Italy (a.amato@poliba.it; dilecce@poliba.it).

V. Piuri is with the Department of Information Technologies, University of Milan, Crema 26013, Italy (e-mail: piuri@dti.unimi.it).

Digital Object Identifier 10.1109/TFUZZ.2008.917287

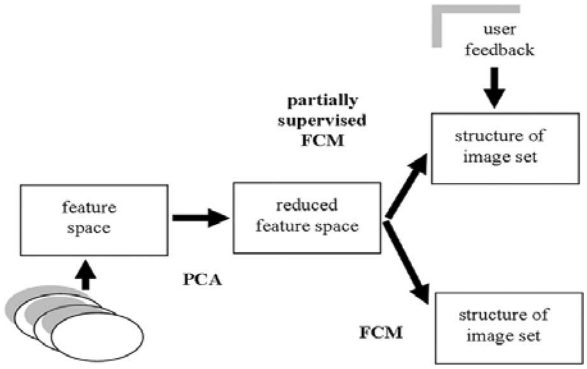


Fig. 1. Schematic overview of the proposed approach to image organization (PCA: principal component analysis; FCM: fuzzy c-means).

by the fuzzy c-means (FCM) clustering that, in turn, becomes augmented by the machinery of partial supervision.

We present a suite of experiments completed for two databases of images such as a Columbia object image library (COIL-20) (see Fig. 3) and 2000 outdoor images belonging to eight categories, that is, airplanes, cars, ships, landscapes, mammals, insects, flowers, and fishes. The first image database was used in some previous studies [16], [22], so it is advantageous when completing some comparative analysis. The images of the second image database were downloaded from several free Internet image galleries such as, e.g., http://www.isolenelmondo.com/wallpaper/wallpaper.html, http://www.sfondifree.it, etc.

The paper is organized as follows. In Section II, we cover a brief overview of some main trends in image classification. This helps us appreciate the diversity of the existing approaches while acknowledging the commonality of the underlying problems. Next, we describe the two databases of images used in this study in Section III. In Section IV, we present a description of the feature space constructed with the use of an angular spectrum signature by showing how such features are formed and how they capture the characteristics of images. Fuzzy clustering with partial supervision is presented in Section V; here we also discuss aspects of classification following clustering and elaborate on the characterization of the clustering results with the use of a separability index that quantifies the ambiguity of clustering results associated with each image. Experimental studies are presented in Section VI, while conclusions are offered in Section VII.

In this study, we will adhere to the standard notation encountered in clustering and pattern recognition. We assume that all images are described by numeric features that are vectors $(x, y, \ldots)$ positioned in some n-dimensional space of real numbers, that is, $\mathbf{x}, \mathbf{y}, \mathbf{R}^n$. By "*N*," we denote the number of images (patterns), while the number of clusters is equal to "*c*."

II. Image Classification: A Brief Overview

Let us review some of the existing trends in image description, clustering, and classification. This also helps put the proposed approach in a certain broader context especially when it comes to the issues of clustering and clustering with supervision.

Content-based image retrieval (CBIR) has emerged as an important area in computer vision and multimedia computing, just to name a few representative examples. Typically, in CBIR systems, we describe images using a collection of some visual features (often referred to as signatures) extracted from the individual images. In the literature, we can encounter an abundance of low-level visual features. Shape, color, and texture are those in common use, no matter whether we are concerned with commercial systems (e.g., query by image content (QBIC) [13]) or those encountered in the academic community (e.g., photo book [30]). In [3], an improved version of color signature was proposed that embeds some local information about the statistical and visual relevance (importance) of each pixel.

With the increasing diffusion of compressed images, there have been substantial attempts to extract their visual features (see [2], [21], [40], and [42]). One could also encounter the use of more sophisticated and abstract features such as wavelets [23], [25], [34], [43]. Visual descriptors coming as a part of the MPEG-7 standard (a standard developed by Moving Picture Experts Group) concentrate on color, texture, shape, etc. [6], [33]. Some other techniques are aimed at the improvement of effectiveness of image retrieval through combining color and texture [11], [44], color and shape [18], and building other aggregates. The study [17] brings a comprehensive overview of the theory, techniques, and applications of CBIR. Another interesting comparative evaluation can be found in [8]. Typically, one encounters the following general scheme:

$$Stimuli \rightarrow Signatures \rightarrow Distance.$$

Stimuli are sought as points in some perceptual space [31], while the notion of similarity between two stimuli is one of the fundamental concepts of the cognitive theories of similarity. For the visualization of the underlying ideas, let us refer to Fig. 2.

Consider two different images from the semantic point of view. Let A_p be the semantic (human-centric) space. In this space, for each image, there is a stimulus. On its basis, it is easy for us to assess the similarity between two or more images. The CBIR systems try to emulate this cognitive chain. Such systems associate a signature (A_s and B_s in Fig. 2) to each image defining a signature space. By endowing the signature space with a distance model, it is possible to define an artificial similarity space in which it is possible to measure the distance between two or more images (denoted here by A_{a-b} and A_{b-a}). In many distance models (such as the one utilized in this work), we assume symmetry, namely, making A_{a-b} equal to A_{b-a}. There could be other constructs, such as, e.g., Tversky's "contrast model" [39] in which the symmetry requirement is not used. Indeed, a central assumption of this model is that the similarity between objects A and B is a function of the features that are common to A and B ("A and B"), those in A but not in B (symbolized as "A–B"), and those in B but not in A (denoted by "B–A"). Based on this concept and several other assumptions, Tversky postulated the following relationship

$$S(\mathrm{A}, \mathrm{B}) = xf(\text{A and B}) - yf(\mathrm{A} - \mathrm{B}) - zf(\mathrm{B} - \mathrm{A})$$

where S is an interval scale of similarity, f is a function of salience of the various features that have been considered, and

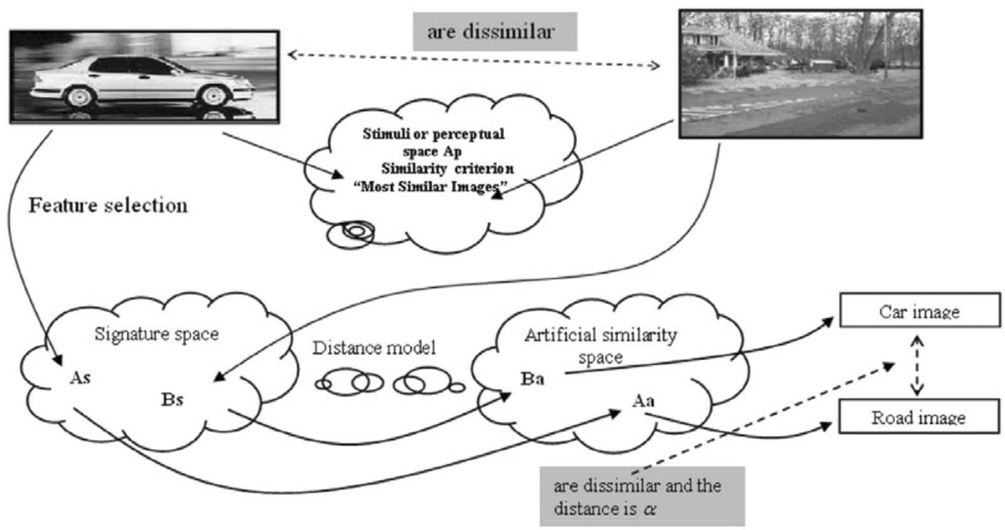


Fig. 2. Relationship diagram between images stimuli →signatures →distance (computable similarity).

Fig. 3. Twenty example images of the Columbia object image library (COIL-20).

x, y and z are weights that underline the relations among the features of the objects A and B.

Following this scheme (in the form of the chain of associations *stimuli→signatures→distance*), several hypotheses have been investigated. In the purely psychological approach based on multidimensional representation, an image is represented as a point in some highly dimensional space [1]. The location of the point is typically determined with the use of some scaling techniques such as, e.g., multidimensional scaling (MDS) [20]. MSD leads to a nonmetric space. The goal is to find a projection space in which the interpoint distance is monotonically related to a human panel response about (di-) similarity. In the purely computational approach, a natural scene is represented by means of a collection of values (*signatures*) explicitly derived from the 2-D image containing basic low-level features. These could be comparable to such features as retinal–brain sensitivity, including shape, colors, and patterns.

In [35], the concepts of semantic and sensory gap are introduced. In this sense, such concepts are correlated with the accuracy that could be obtained by means of CBIR techniques. The semantic gap is defined as the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation. The sensory gap is defined as the difference between the object in the world and the information in a (computational) description derived from a recording of that scene.

A number of image retrieval systems use clustering as a certain preprocessing step. In many studies, such as, e.g., [14] and [24], clustering is used to segment images as an early step in the image understanding process. In other works, the underlying objective is to complete some organization of the entire collection of images into some homogeneous groups. In both approaches, the execution time of the clustering algorithm is a critical point. Many works deal with this problem, and a promising approach can be found in [12].

In general, the number of clusters is specified in advance, or one could consider some mechanisms of adaptive adjustment of their number. In [22], an algorithm was proposed that forms a certain variant of Competitive Agglomeration [15]. Both of the algorithms proposed in [15] and [22] are able to "discover" the number of clusters. The first one is equipped with the concept of a so-called noise cluster, which collects all images that are regarded as ambiguous, and therefore, difficult to categorize. The feature space under consideration captures important aspects of images being used afterward in classification, such as intensity distribution, shape, and texture. The resulting feature space having originally 152 features was further reduced with the use of the standard PCA. The system was tested using the COIL-20 image database. The same image database along with the same feature space was used in [16], in which the proposed algorithm was aimed at an automatic determination of the number of clusters and the weights to be used in the

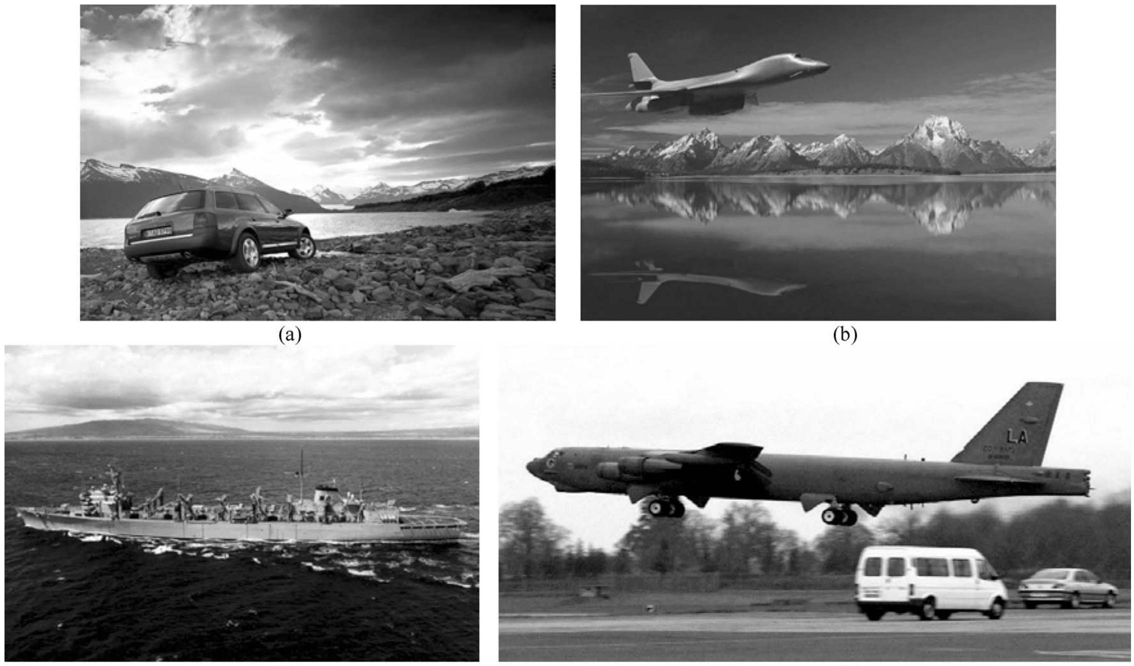
TABLE I
LIST OF SELECTED METHODS AND SYSTEMS USED IN IMAGE CLASIFICATION

Reference	Method and data description	Feature space	Mode of learning	Accuracy
[22]	Adaptive Robust Clustering. Columbia object image library COIL-20. The database contains 1,440 gray images of 20 objects. There are 72 images per each object, with each image taken at pose intervals of 5°	1. Texture (8-D) 2. intensity distribution (16-D) 3. Shape and structure (128-D).	unsupervised	24% of misclassified patterns plus 4% of images classified as noise
[16]	Self-Organization of Oscillators Network (SOON). Columbia object image library COIL-20	1. Texture (8-D) 2. intensity distribution (16-D) 3. Shape and structure (128-D).	unsupervised	38% of misclassified patterns plus 27% of images classified as noise
[41]	Neural networks. Database composed of seven classes with 43 images in each of them. 50% of database used for training.	Images represented as trees. A 12 dimensional visual feature space (color, texture, spatial information) is utilized to represent each node.	supervised (reported are the best results)	Method — Accuracy % Binary Tree — 83.44 Quad-Tree — 88.08 LD-tree (5 seeds) — 78.81 Binary Tree (5 seeds) — 77.48
[36]	Fisher discriminant functions. Brodatz collection with images described by 112 textures	Sub-band energy features composed of 32 elements extracted using four different techniques.	supervised	Method — Accuracy % Wavelet — 92.14 Uniform sub-band — 92.14 4x4 DCT — 85.18 4x4 spatial — 34.65
[38]	Support vector machines (SVMs) and fuzzy sets. Corel image collection.	Color and texture	supervised	The accuracy is evaluated by five human judges. The highest reported accuracy is 64.7%.

distance function. In [41], several standard methods of image partition were analyzed in application to adaptive image categorization. Each image is segmented, and then, represented using various tree structures, including binary trees, quad-trees, and LD-trees. Fisher discriminant functions were used in [36] to classify the Brodatz textures collection. The features adopted in this classification problem were the elements of the subband energy. The system was tested using four different decomposition methods to compute the subband energy, that is, wavelet subband, uniform subband, discrete cosine transform (DCT), and spatial partitioning. The study reported in [38] introduced a two-stage mapping model. In the discussed system, support vector machines (used for color and texture classification) and fuzzy logic (for high level concept inference) work in a collaborative manner. Nine support vector machines (SVMs) were designed to classify colors, while 14 SVMs were utilized for texture classification. The training and testing sets consisted of 16 classes extracted from the Corel image collection (this image database is composed of 17 800 outdoor images belonging to 90 different classes).

A snapshot of the plethora of the existing approaches is covered in Table I. This table presents a series of the most essential features of each of them, such as databases being used, visual features, learning method, and the best performance achieved. It appears evident that the unsupervised learning methods can achieve good results when applied to image databases located in some narrow domain [16], [22]. However, while using image databases embracing a broad spectrum of images [38], [41], the supervision component becomes highly desirable.

Let us note that the predominant objective of Table I is to offer a general, high-level view of the existing approaches and emphasize their evident diversity. Given an acute lack of the standard database to be used as a test bed for CBIR systems, it prevents us from a thorough and fully objective assessment of results reported in the literature. Furthermore,



(a) (b) (c) (d)

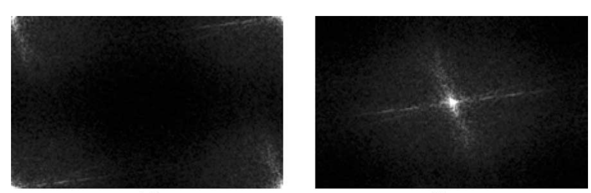
Fig. 4. Examples of images belonging to more than a single class. More than 20% of this database consists of images of this nature. (a) Car landscape. (b) Airplane landscape. (c) Ship landscape. (d) Airplane car (with some landscape in the background).

(a) (b) (c)

Fig. 5. (a) Example image. (b) Visual effect of the application of the 2-D Hamming function to image (a). (c) Angular spectrum of (a); see a detailed description in Section IV.

existing experimental setups used in various references are different and quite often lack pertinent details; hence, a comparison of the numeric values of the classification rates could be quite limited.

Considering the existing image databases of general format, one could refer to the Corel image database. This database is composed of a set of more than 800 photo compact discs (CDs), where each of them consists of 100 images that are grouped along the same category. Unfortunately, Corel (see http://www.corel.com/) does not offer this database any longer but provides differently compiled collections of images of smaller size with sets comprising several thousand images. It seems that this database could solve the problem of the lack of a common test bed for the validation of CBIR systems; however, this is not true. In an interesting study [26], the authors compared different ways of evaluating the performance using a subset of the Corel images and considering the same set of evaluation measures. They show that it is very easy to arrive at different results, even in cases when the same collection of images is being used.



(a) (b)

Fig. 6. (a) Standard representation of the real part of $F(u,v)$ of the image in Fig. 5(a). (b) Representation of $F(u,v)$ in which the zero-frequency component is shifted to the center.

III. Brief Description of Image Databases

In this study, we consider two image databases, that is, the COIL-20 [27] and an image database composed of 2000 outdoor images downloaded from several free Internet image galleries.

The first database is composed of 1440 gray images of 20 objects (refer to Fig. 3). There are 72 images per object where a photo of each image is taken at pose intervals of 5°. This

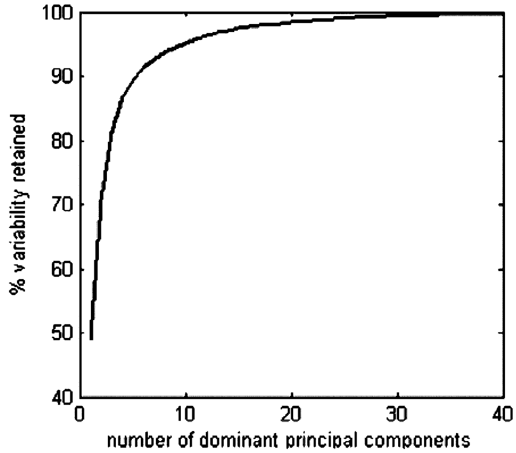


Fig. 7. Total percentage of the variability retained versus the number of principal components used; note a visible saturation effect occurring with the increase of the number of the principal components taken into consideration.

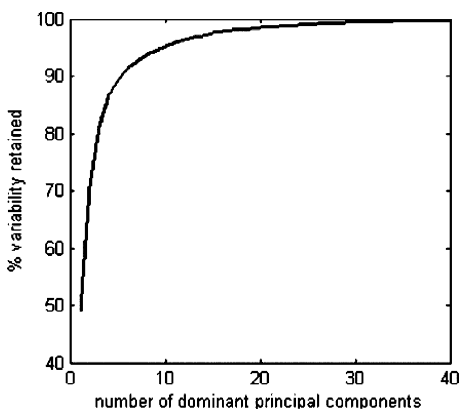


Fig. 8. Percentage of the total variability retained versus the number of principal components being used in the formation of the space.

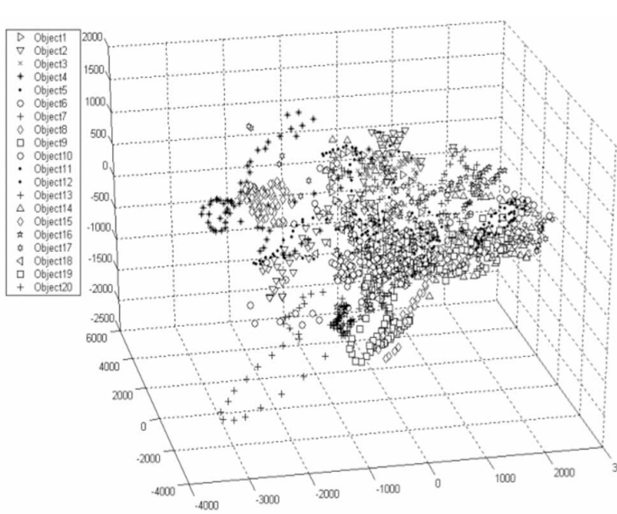


Fig. 9. Representation of the reduced features space, relative to the COIL-20 image database. This reduced space is formed with the first three dominant PCA components (with 79.8% of the total variability retained).

database was often utilized in image categorization (cf. [22] and [41]). Some selected examples of COIL-20 are included in Fig. 3.

The images of the second database were examined by a human, and subsequently, grouped into eight classes such as airplanes, cars, ships, landscapes, mammals, insects, flowers, and fishes. This human classification was used as ground truth to assess the performance of the proposed system. The images of this image database were downloaded from several free Internet image galleries, such as, e.g., http://www.isolenelmondo.com/wallpaper/wallpaper.html, http://www.sfondifree.it, etc. This image database is far more challenging from the classification standpoint than the synthetic images forming the content of the first database. Indeed, quite often, such images are composed of several objects that make their class assignment to a single class highly problematic and perhaps pretty questionable. In this study, to make the classification even more challenging with this regard, the images were chosen in such a way that 20% of them could easily belong to more than a single class. Some selected pictures shown in Fig. 4 are convincing examples of just such conceptual "mixtures" of several categories; the first one is a mix of the "car" category, yet it also exhibits a significant component that could have made it quite well allocated to the "landscape" category. An image in Fig. 4(d) is an interesting mix of the categories of cars and airplanes.

Using the terminology presented in [17], we conclude that the first image database is composed of images positioned in a narrow domain, while those of the second database belong to a broad domain. Let us note that the images belonging to a narrow domain are characterized by a low variability of their pictorial content. Typically, the imaging process is similar for all recordings: the image contains only one object against a clear background. Typically, there is no substantial variation in the level of illumination. Other factors making the images quite diverse are also not present. The first image database is a convincing example of this kind of images. Indeed, each object is recorded from different points of view, introducing a large variability in the visual details, but fully retaining geometrical, physical, and illumination constraints that characterize the pictorial domain. Under such conditions, the semantic content of the image could be well defined and made unique. As the results obtained in [16] and [22] demonstrate, it is possible to achieve good classification results while applying here the CBIR techniques.

On the other hand, the images belonging to a broad domain are characterized by a high variability of their pictorial content even for the same underlying semantics. There are no specific constrains introduced to the imaging process. Each image can contain more than one object and can be recorded under different illumination conditions and at different viewpoints. This situation leads to a significant gap between the visual features and the semantic content of images. The semantic interpretation of these images can be difficult even for a human being, because quite often their semantic content is described only partially and not univocally. The second image database could be a good example of this kind of image; see Fig. 4. Here, the

TABLE II
RELATIONSHIP BETWEEN THE CLUSTERS AND CLASSES

Cluster\Class	ω_1	ω_2	ω_3	ω_4
1	$\mathbf{N_{11}}$	N_{12}	N_{13}	N_{14}
2	N_{21}	N_{22}	N_{23}	$\mathbf{N_{24}}$
3	N_{31}	$\mathbf{N_{32}}$	N_{33}	N_{34}
4	N_{41}	N_{42}	$\mathbf{N_{43}}$	N_{44}

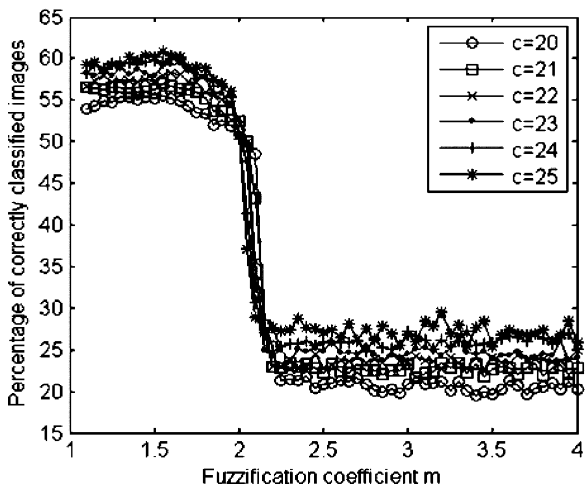


Fig. 10. Classification rate versus the values of the fuzzification coefficient for selected numbers of clusters.

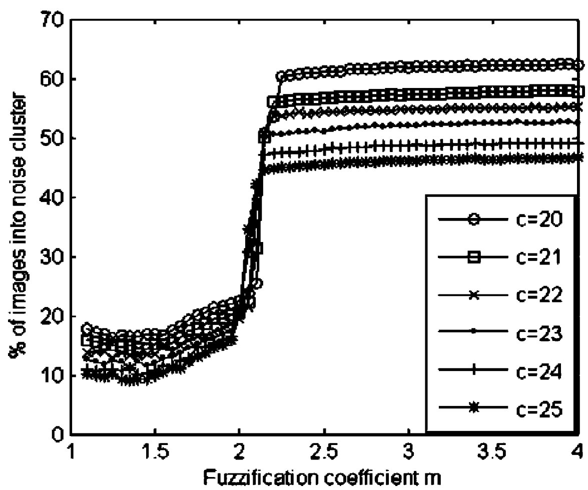


Fig. 11. Percentage of images falling into the noise cluster versus the values of the fuzzification coefficient. This relationship is presented for selected numbers of clusters.

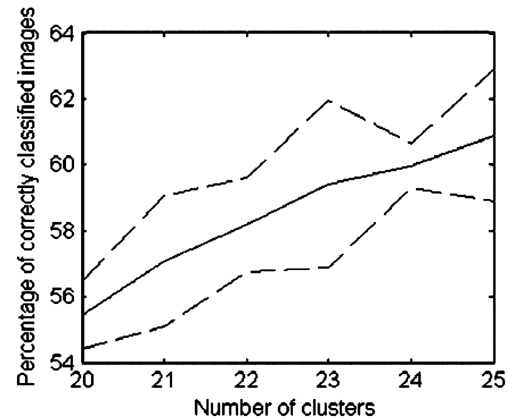


Fig. 12. Classification rate versus the number of clusters; the dotted lines represent a standard deviation of the reported classification rate.

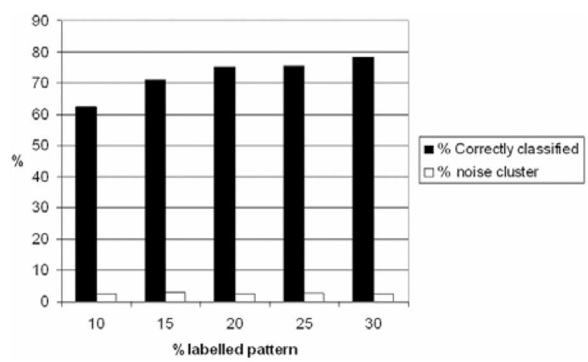


Fig. 13. Classification rate (mean value) along with the size of the noise cluster dimension versus the level of the supervision (expressed by the percentage of the labeled patterns used in the clustering process).

classification accuracy obtained using CBIR techniques and supervised learning is far from 100% [38].

IV. DESIGN OF THE FEATURE SPACE: AN ANGULAR SPECTRUM SIGNATURE AND ITS PROPERTIES

As it has become apparent from the previous investigations, forming a suitable feature space [5], [7], [9] in which images could be described becomes a critical design step that directly impacts the performance of the ensuing classification schemes. In this study, we consider a suite of features called the angular spectrum signature [5]. Interestingly enough, these features were successfully applied to some other problems of image classification including video skimming [9] and image retrieval [7]; in both the cases, they gave rise to good results.

There are two reasons that directly motivate us to treat the angular spectrum signature as a viable and attractive design alternative. They relate directly to the way images are perceived and interpreted:

1) visual properties are mainly related to the largest objects in images; and
2) among various visual descriptors, shape, texture, and orientation play a predominant discriminatory role. In the case of the angular spectrum signature, shape and orientation are well captured.

In essence, finding and describing the distribution of directions of lines present in any image can be done through the analysis of the results of the Fourier transform. Initially, each image is preprocessed applying a windowing function (here a 2-D Hamming function) in order to remove discontinuities at the edges, which otherwise could have complicated the interpretation of frequency spectra. A windowing function is selected in such a way that it goes to zero (or near to zero) at the edges of the image. Here, we confine ourselves to the 2-D Hamming function because it

TABLE III
COMPARATIVE SUMMARY OF RESULTS OBTAINED IN THIS STUDY AND PREVIOUS PAPERS [16] AND [22]

	FCM with noise cluster (c=20, m=1.55)	FCM with noise cluster (c=25, m=1.55)	PS-FCM (10% labelled images, c=20)	PS-FCM (20% labelled images, c=20)	PS-FCM (30% labelled images, c=20)	[16]	[22]
% of correctly classified images	55.5%	61%	62.5%	75.3%	78.4%	35%	72%
% of noise cluster	17.5%	10%	2.5 %	2.4%	2.6%	27%	4%

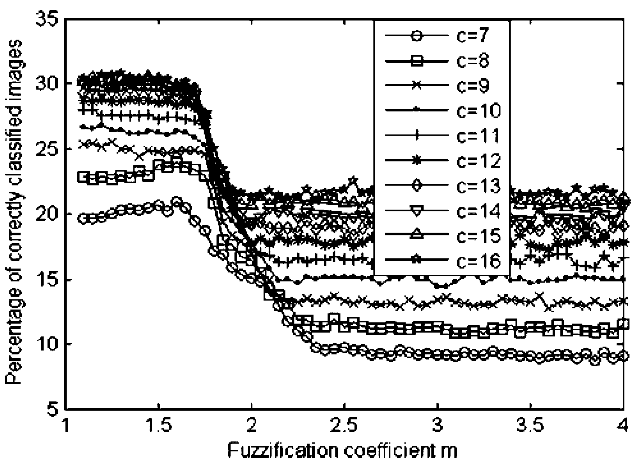


Fig. 14. Classification rate versus the values of the fuzzification coefficient for selected numbers of clusters.

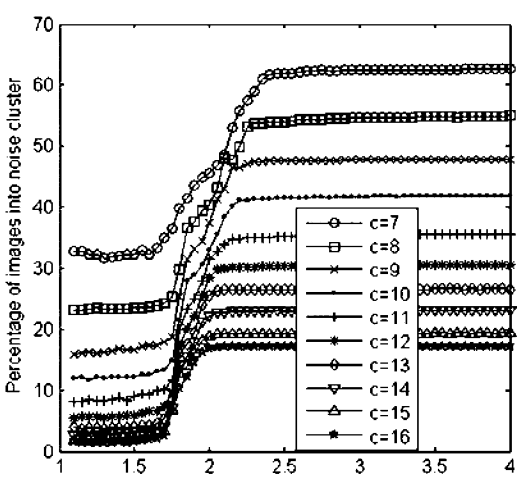


Fig. 15. Percentage of images into noise cluster versus the values of the fuzzification coefficient for selected numbers of clusters.

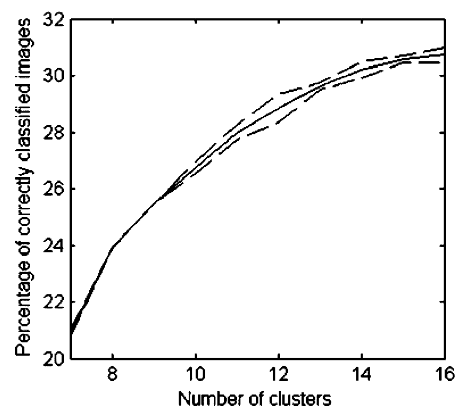


Fig. 16. Classification rate versus the number of clusters; the dotted lines represent a standard deviation of the classification rate.

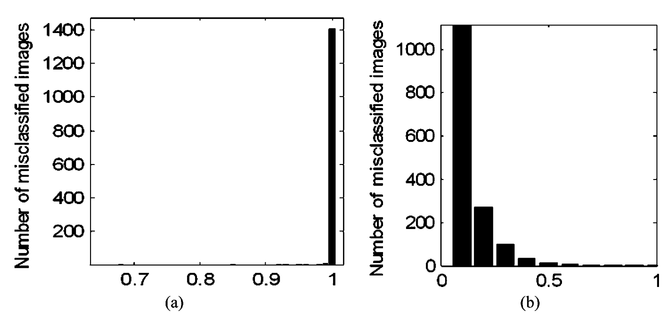


Fig. 17. Histogram of the values of misclassified patterns. (a) $c = 8$ and $m = 1.15$. (b) $c = 8$ and $m = 2.2$.

exhibits a strong radial symmetry starting from the center of the image.

The 1-D Hamming function $w(k)$ comes in the form

$$w(k) = 0.54 - 0.46 \cos \frac{2\pi k}{A - 1} \tag{1}$$

where A is the number of samples and "k" stands for the kth sample.

The 2-D Hamming function, for an image composed of n_1 by n_2 pixels, can be easily formed from the basis of 1-D Hamming functions in the form of an outer product, that is, $w(n_1)^T w(n_2)$ where $w(n_1)$ and $w(n_2)$ are 1-D vectors obtained using (1). The visual effect of applying a 2-D Hamming function to an image is illustrated in Fig. 5(b). In some sense, we could compare it to an effect produced by a low-pass filter. Note that this transform is quite negligible at the center of the image, and its impact increases when moving toward the borders of the image. In this sense, the parts of the image located close to the borders reduce their contributions to the overall image frequency spectrum. In this sense, one could focus the analysis on the central part of the image because, often, in the imaging process, the focus is localized in this part of the scene. The low frequencies

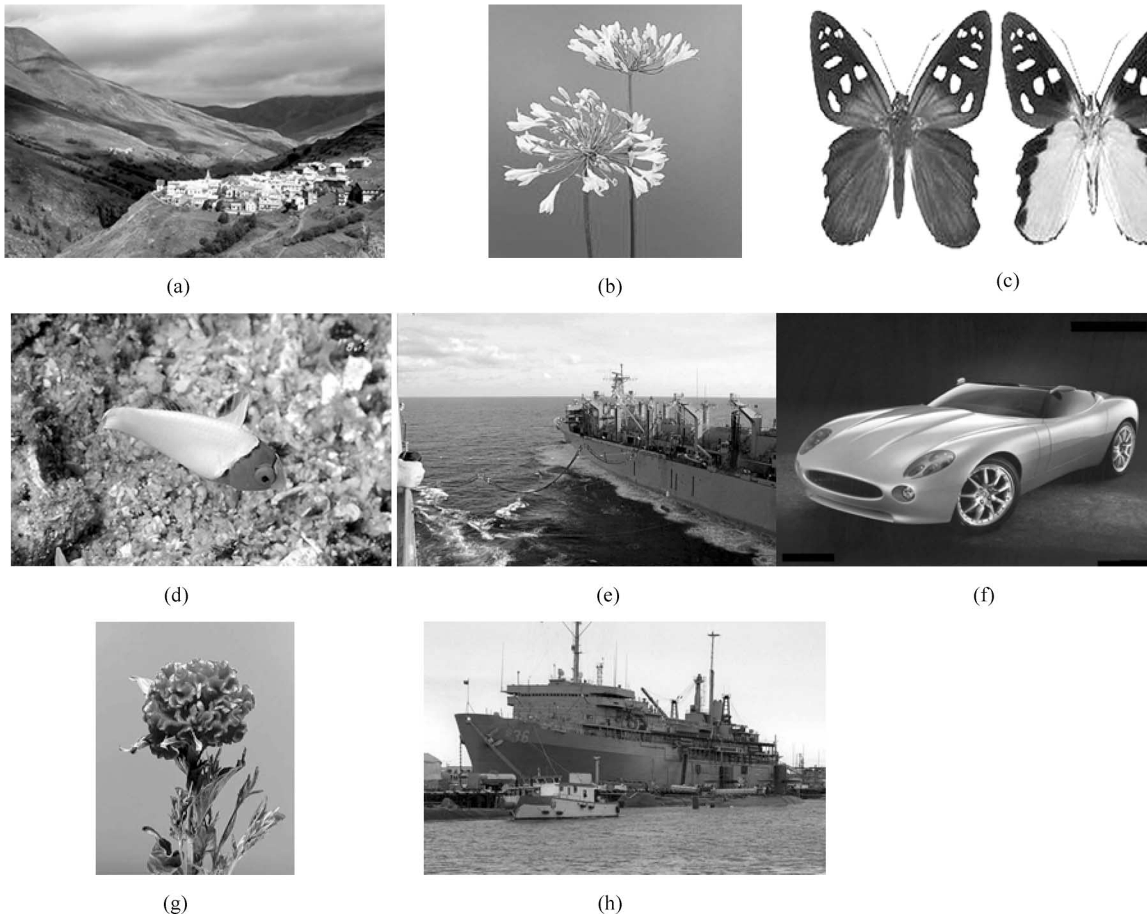


Fig. 18. Images closest to the prototypes constructed by the FCM algorithm with $m = 1.50$ and $c = 8$. Membership values: (a) 0.99, (classified as "car"); (b) 0.99, (classified as "flower"); (c) 0.99, (classified as "insect"); (d) 0.98, (classified as "mammal"); (e) 0.99, (classified as "ship"); (f) 0.99, (classified as "fly"); (g) 0.99, (classified as "mammal"); (h) 0.96, (classified as "fish").

TABLE IV
EXAMPLE OF ASSOCIATION BETWEEN CLUSTERS AND CLASSES FOR $c = 8$ AND $m = 1.50$; THE DOMINANT CLUSTER-CLASS ASSOCIATIONS LINKAGES ARE SHOWN IN BOLDFACE

Cluster\Class	mammals	cars	flowers	fish	airplane	insect	landscape	ship
1	12	**58**	14	7	33	8	32	21
2	34	28	**77**	25	31	12	24	16
3	11	0	4	5	1	**51**	2	0
4	**45**	20	28	40	32	21	22	6
5	11	26	15	9	33	3	45	**95**
6	11	29	17	27	**40**	18	37	33
7	**43**	18	34	40	5	25	22	3
8	15	27	10	**38**	11	13	34	22
Noise	68	44	51	59	64	99	32	54

correspond to the largest and most dominant components occurring in the image. They usually are the most relevant when interpreting the image. One could also assume that they are frequently localized in some central areas (e.g., foreground objects) or are spread across the entire image (e.g., landscapes). Their contribution to the generation of the feature space is almost completely preserved after the application of the 2-D Hamming function.

High frequencies correspond to negligible details and fine-grain texture elements of an image; a reduction of these components means that we focus on the foreground components, located in the image central area, and ignore the details of the peripheral contour. Once the 2-D Hamming function has been applied, we compute the discrete Fourier transform to produce the image spectrum. In general, for an image of "p" by "r" pixels, its Fourier transform $F(u, v)$ comes in the well-known form:

$$F(u,v) = \frac{1}{p \times r} \sum_{i=0}^{p-1} \sum_{1=0}^{r-1} f(i,1)\, e^{-j2\pi(u_j/p + v_1/r)}. \qquad (2)$$

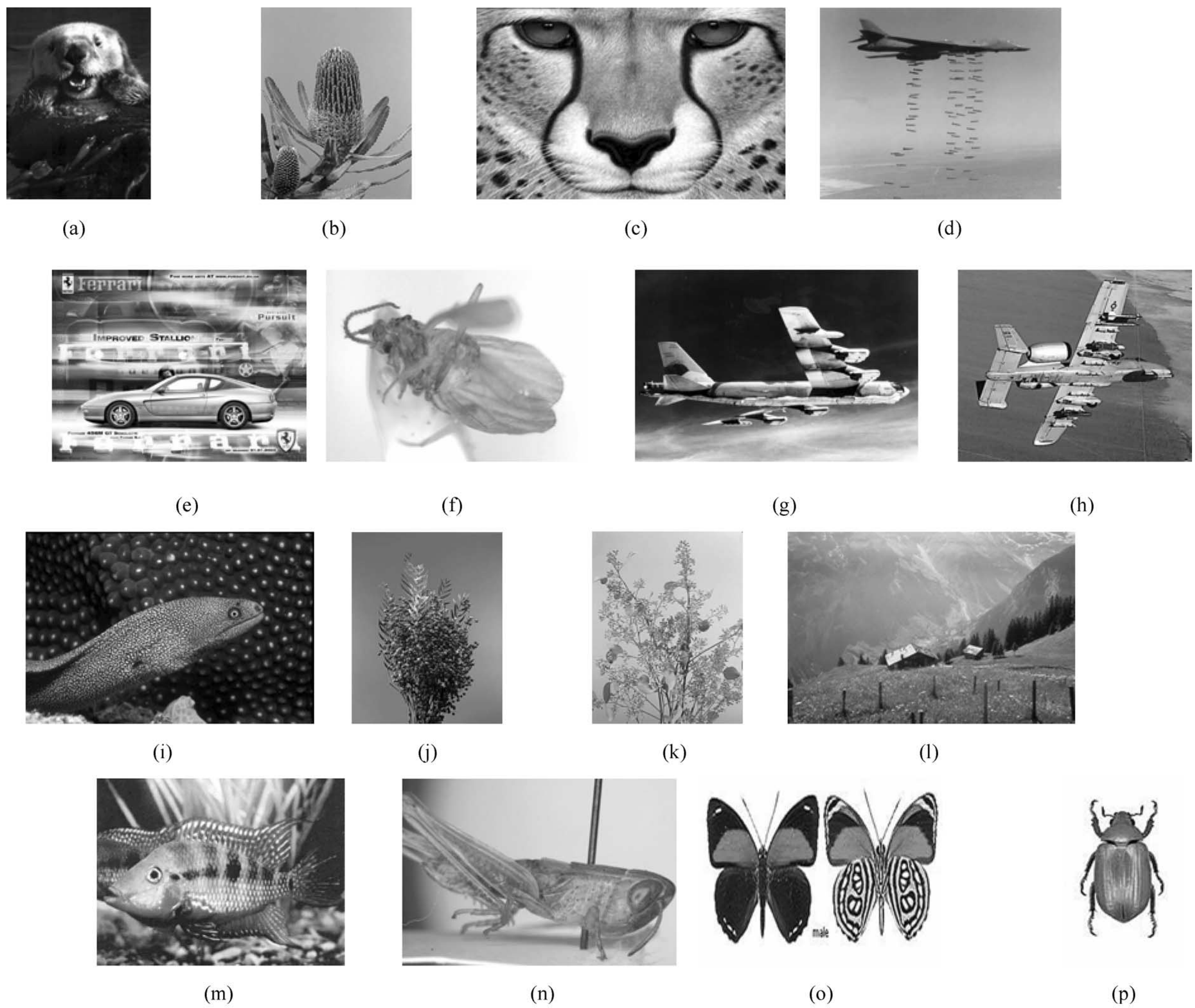


Fig. 19. Closest images to the prototypes constructed by the FCM algorithm with $m = 1.25$ and $c = 16$. Membership values of all these images are equal to 1. The name of the dominant class of the cluster is shown below each image. (a) "mammal," (b) "landscape," (c) "flower," (d) "ship," (e) "insect," (f) "airplane," (g) "insect," (h) "mammal," (i) "ship," (j) "fish," (k) "insect," (l) "mammal," (m) "fish," (n) "fish," (o) "car," (p) "car."

In the sequel, $F(u, v)$ is further processed in order to construct a polar representation of the image spectrum in which the zero-frequency component is shifted to the center (refer to Fig. 6).

This representation of the image spectrum is split into ten uniformly distributed angular sectors, each the size of 18°, starting from $\theta = \pi/2$ and moving down to $\theta = -\pi/2$. For each sector, seven harmonics are evaluated and recorded in a histogram format. The histogram produced in this manner and composed of 7×10 entries forms a feature space of the angular spectrum.

An example of the space is shown in Fig. 5(c). The values of harmonics greater than 7 refer to small elements in the image; their contribution to the feature space becomes very much limited, and hence, they could be easily eliminated. As a result, we consider the feature space as a vector in R^{70}. To eliminate any redundancy and lower computing overhead when processing and classifying the images, we considered a standard mechanism of feature reduction, such as the PCA (see, e.g., [19]). This analysis helps us retain the most essential principal components and learn about the variability of the feature space.

Fig. 7 shows that the percentage of retained variability increases quite rapidly with the growing number of the principal components. As a matter of fact, we needed 17 variables to retain 98% of the total variability of the image. This low-dimensional space will be used to carry out all classification activities.

In the same way, we completed the PCA analysis for the COIL-20 image database (refer to Fig. 8). The saturation effect is also visible. Given the form of the relationship, we have retained 98% variability, which occurs for 16 of the most dominant eigenvalues. Subsequently, this R^{16} space will be used in classification activities.

For illustrative purposes, a PCA-driven reduced space with three variables is shown in Fig. 9. In this case, we retain 79.8% of the total variability; here, some homogeneous groups of objects become distinguishable.

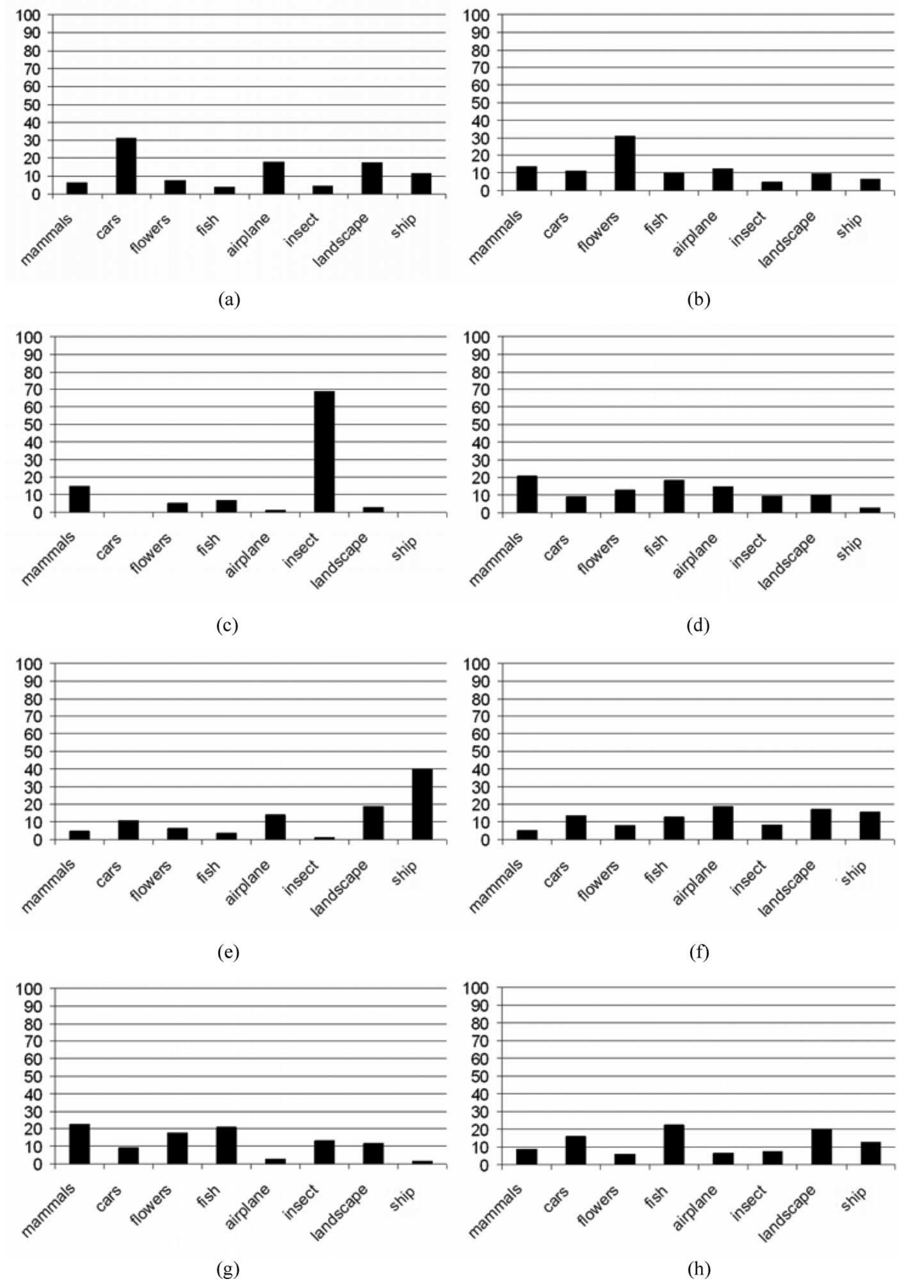


Fig. 20. Images distribution among the clusters produced by the FCM with $c = 8$ and $m = 1.50$. (a) Cars, (b) flower, (c) insect, (d) mammals, (e) ship, (f) airplane, (g) mammals, (h) fish.

V. Fuzzy Clustering and Clustering With Partial Supervision

The organization and classification of images is supported by the mechanisms of fuzzy clustering and fuzzy clustering with partial supervision. The nature of the problem under discussion suggests using the concept of a noise cluster as originally proposed in [10]. We start with fuzzy clustering to investigate to which extent unsupervised learning becomes efficient in revealing the structure in the collections of images. Next, we contrast the performance of the clustering mechanisms with the clustering augmented by some mechanisms of supervision. The supervision mechanism reflects situations when we are provided some guidance that helps discover and describe the structure in the image dataset.

The clustering technique considered here is the FCM. This is a well-known objective function-based clustering whose objective function comes in the form

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{m} \left\| \boldsymbol{x}_k - \boldsymbol{v}_i \right\|^2 . \tag{3}$$

In the previous expression, $\|\cdot\|$ denotes a distance function (Euclidean, Mahalanobis, etc.), $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_c$ are the prototypes (centroids) of the constructed clusters, $U = [u_{ik}]$, $i = 1, 2, \ldots, c; k = 1, 2, \ldots, N$ stands for the partition matrix,

TABLE V
ASSOCIATIONS BETWEEN CLUSTERS AND CLASSES FOR $c = 16$ AND $m = 1.25$; THE DOMINANT CLUSTER-CLASS LINKAGES ARE SHOWN IN BOLDFACE

Cluster\Class	mammals	cars	flowers	fish	airplane	insect	landscape	ship
1	**27**	14	20	22	22	6	10	4
2	14	11	10	24	3	7	**32**	14
3	27	21	**70**	17	23	8	18	13
4	5	28	6	20	38	10	29	**39**
5	24	2	17	11	6	**44**	9	1
6	6	15	8	8	**20**	4	9	15
7	9	0	1	5	0	**44**	2	0
8	**28**	1	10	9	10	24	4	0
9	11	28	16	8	34	2	43	**95**
10	8	12	5	**21**	13	9	11	9
11	8	6	17	2	6	**33**	1	1
12	**38**	18	33	34	4	14	21	2
13	24	20	15	**31**	15	16	16	6
14	9	8	4	**24**	12	14	14	19
15	2	**14**	5	5	11	2	4	14
16	10	**52**	11	2	27	8	27	13
Noise	0	0	2	7	6	5	0	5

Fig. 21. Examples of images belonging to the noise cluster; FCM with $c = 16$ and $m = 1.25$.

while "m" is a fuzzification coefficient greater than 1 whose function is to control a level of fuzziness of the clusters (their membership functions).

The minimization of Q that is carried out with respect to the partition matrix and the prototypes gives rise to the structure in X. The generic optimization scheme involves a sequence of iterations, in which we successively update the values of the partition matrix and the prototypes. The process terminates when some stopping criterion has been met. Typically, we consider

$$\|U - U'\| \max_{i,k} |u_{ik} - u'_{ik}| . \tag{4}$$

The iterative process stops when the value of the previous expression does not exceed some predefined positive threshold ε. U and U' denote partition matrices obtained in the two consecutive iterations.

It becomes apparent that the discovery of the structure in images is completed in a fully unsupervised mode; no supervision is offered and our anticipation is that the objective function is somewhat reflective of the geometry of the data so that the minimization of Q leads to an "acceptable" structure. While this could be the case in many practical situations, in reality we could be offered some guidance in the form of a limited number of hints coming from the user who imposes his point of view at

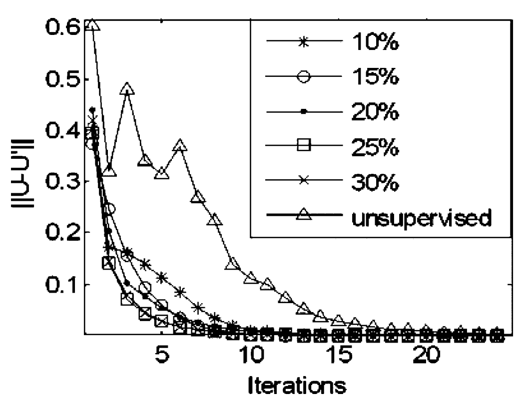


Fig. 22. Values of $||U - U'||$ in successive iterations of the clustering algorithm.

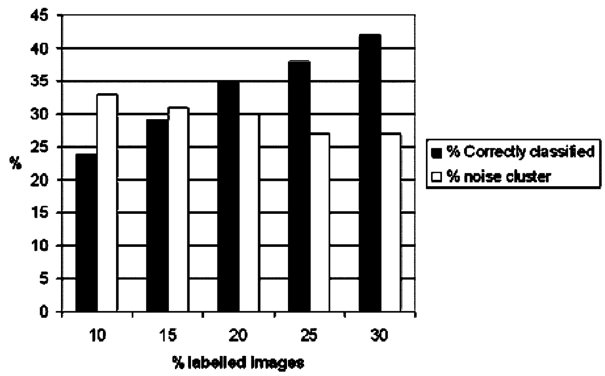


Fig. 23. Classification rate versus the number of labeled images.

the data (here, images) by labeling several selected images. This is an underlying rationale behind the FCM with partial supervision [28]. Let us briefly recall the essence of this algorithmic enhancement.

Denote by $\boldsymbol{X}_1 \subset \boldsymbol{X}$ a (usually small) set of labeled patterns. The cardinality of $\boldsymbol{X}_1$ is equal to M. The elements of this set are denoted by a Boolean vector $\boldsymbol{b}$, $\dim(\boldsymbol{b}) = N$ whose entries are equal to 1 if the corresponding patterns have been labeled and set up as 0 otherwise. The corresponding membership degrees of the labeled patterns are represented in a matrix $F = [f_{ik}]$, $i = 1, 2, \ldots, c; k = 1, 2, \ldots, N$.

The augmented objective function that incorporates the supervision mechanism described before is expressed in the form

$$J = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{m} d_{ik}^{2} + \alpha \sum_{i=1}^{c}\sum_{k=1}^{N} (u_{ik} - b_k f_{ik})^{m} d_{ik}^{2}. \quad (5)$$

The second component of (6) captures the differences between the "true" membership of the images and those coming from the structure produced by the clustering algorithm. We want these two structures to coincide for the labeled patterns. The positive weight (α) balances the effects of the supervised and unsupervised learning. The value of α could be chosen in the form $\alpha \propto N/M$. Some other alternatives of assigning some other values of α could be sought as well. In general, the higher the value of α is, the more essential the impact coming from the supervision part of the scheme.

To minimize the performance index J (which is completed with respect to the partition matrix and the prototypes) while avoiding a tedious process of solving algebraic equations, we set up the value of the fuzzification coefficient (m) to be equal to 2.0. In this case, the partition matrix is updated following the expression

$$u_{ik} = \frac{1}{1+\alpha}\left[\frac{1+\alpha\left(1 - b_k \sum_{l=1}^{c} f_{lk}\right)}{\sum_{l=1}^{c} d_{ik}^{2}/d_{lk}^{2}}\right] + \alpha b_k f_{ik}. \quad (6)$$

The results of the FCM come as a collection of clusters. These clusters might not be homogeneous and could include images belonging to several different categories. This situation is not surprising at all; perhaps, it is rather unrealistic to expect that unsupervised learning could produce a zero classification. Furthermore, some images could be treated as a mixture of several categories (say, landscape–car), as we have explained in Section III. In this way, allocating the nonzero degrees of membership of such images to a number of categories could be quite helpful and fully justifiable. To assess the classification error, let us form a correspondence between the clusters and classes. Given an image x_k, we allocate it to some cluster on a basis of the maximal membership value to this specific cluster, that is

$$i_0 - \arg \max_{i} u_{ik}. \quad (7)$$

Then, counting the class assignment of the images allocated to the given cluster and picking up the dominant class within it and repeating this for all clusters, we establish a relationship between the clusters and classes, as shown in the tabular form in Table II.

The entries of the table marked in boldface indicate the maximal number of images belonging to the class; in this way, we arrive at the correspondence between the clusters and classes; while looking at this table, we note that class ω_2 corresponds to cluster 3. An ideal situation occurs when, for each column, we have exactly a single nonzero entry. Practically, due to overlap between classes, this situation does not occur in practice.

The values of partial membership produced by the FCM are very helpful in identifying and flagging the elements that are the most "unclear" (whose membership is shared between the clusters and the degree of membership to each cluster is equal to $1/c$). To quantify this effect of such ambiguity, we introduce the following separability index

$$\psi_k = 1 - c^{c} \prod_{i=1}^{c} u_{ik} \quad (8)$$

whose values are computed for each image "k." The value of ψ_k equal to zero occurs when all u_{ik} are equal to $1/c$. On the other hand, if one membership degree is equal to 1, ψ_k returns 1.

VI. Experimental Studies

Here we report on a comprehensive suite of experiments of unsupervised and partially supervised clustering in application to the two datasets presented in Section III. The clustering methods used the stopping criterion with the value of ε set up to 10^{-6}. The distance function was the Euclidean one. To explore the possibilities offered by the FCM, we experimented with various values of the fuzzification coefficient as well as looked at the different number of clusters. For each scenario, the experiment was repeated ten times by starting the algorithm from different initial configurations (represented by different partition matrices being initialized in a random manner). This is deemed of importance when interpreting the results and striving for their stability. The scenario of partial supervision was realized by randomly labeling a certain percentage of images (say, 10%, 20%, and 30%).

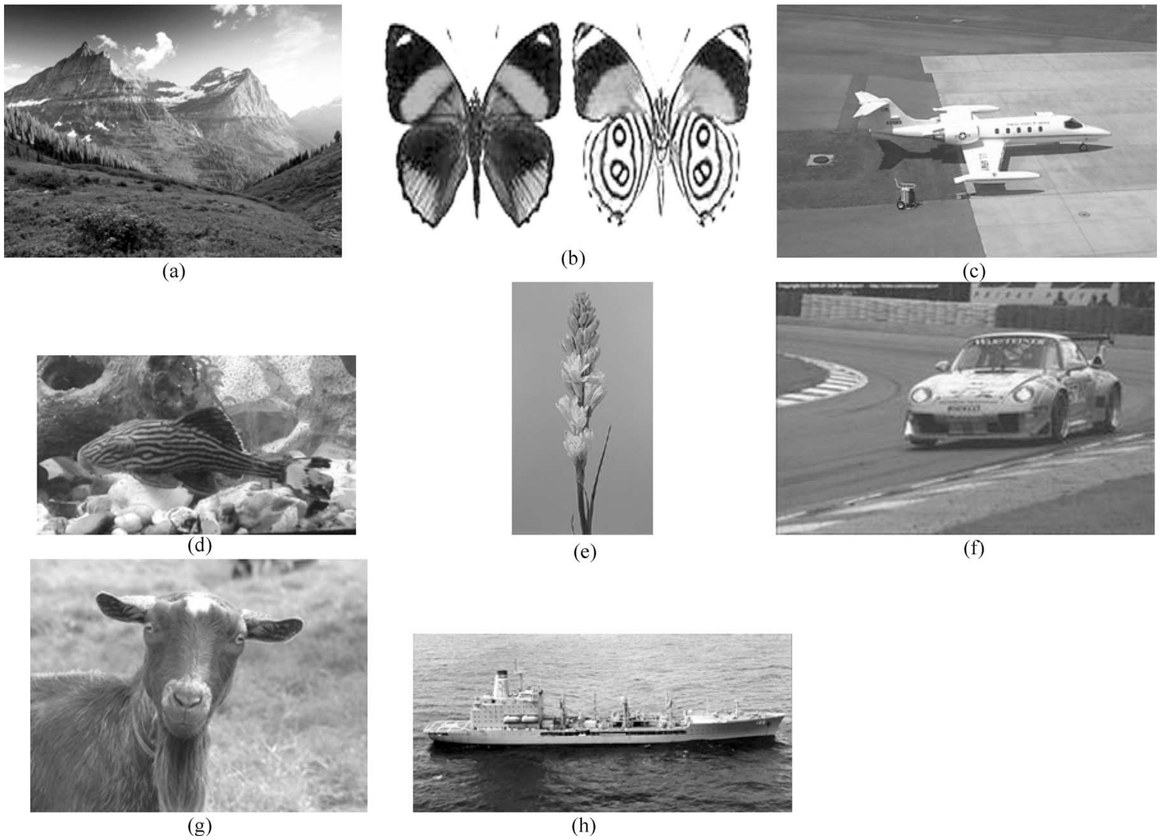


Fig. 24. Images located the closest to the prototypes formed by the supervised FCM ($m = 2$ and 30% of labeled patterns). The membership values: (a) 0.82, (classified as “landscape”), (b) 0.81, (classified as “insect”), (c) 0.82, (classified as “airplane”), (d) 0.83, (classified as “fish”), (e) 0.92, (classified as “flower”), (f) 0.88, (classified as “car”), (g) 0.80, (classified as “mammal”), (h) 0.93, (classified as “ship”).

A. COIL-20 Image Database

In the series of experiments completed for this dataset, we are interested in a number of aspects of the algorithm, in particular, the number of clusters and the related classification error. First, we note that the fuzzification coefficient does exhibit a substantial impact on the quality of the classification (see Figs. 10 and 11).

It is interesting to note that the optimal value of “*m*” is not the one being typically used in various applications of fuzzy clustering (that is 2.0), but the one far lower than that. The highest classification rate (lowest error and smallest noise cluster) occurs when “m” assumes values around 1.5, and this region is preferred regardless of the number of clusters being used.

As expected (see Fig. 12), the increased number of clusters leads to the increase in the values of the classification rate (the rates reported in this figure are obtained for the optimal values of “m”).

What is not surprising is the fact that the values of the classification rate go up quite consistently with the increase of the number of clusters.

Partial supervision helped improve the classification rate. As illustrated in Fig. 13, with the level of 30% labeled images, the clustering led to the classification rate of 78.4%, which represents an increase of 22% over the result obtained without any supervision. It is also quite conspicuous that the noise cluster became smaller than the one formed without partial supervision. This trend is coherent with the structure of the COIL-20 image database, because it is a narrow domain image database. It is well known that in this type of image database, there are few noise images.

Referring to Table I, the results are compared with those presented in [16] and [22] because all of them refer to the same image database and use the noise cluster. The best performance obtained in [22] comes as 72% of correctly classified images, plus 4% of images classified as “noise,” while that obtained in [16] is 35% of correctly classified images plus 27% of the images assigned to the “noise cluster.” The accuracy obtained here is certainly comparable with those obtained in [16] and [22]. In Table III, these results are presented in a more concise format.

B. Outdoor Image Database

We have completed a series of experiments for the outdoor image dataset reporting the results in a similar way, as it was done in the previous suite of experiments for the COIL-20 dataset. In particular, Fig. 14 shows how the classification rate is impacted by the crucial clustering parameters, such as the number of the clusters and the values of the fuzzification coefficient. Given the number of classes (eight), the number of clusters was varied from seven to 16. As to the optimal value of the fuzzification coefficient, the tendency is again quite apparent: regardless of the number of the clusters, its value should be far lower than 2.0, in many cases, very close to 1.5.

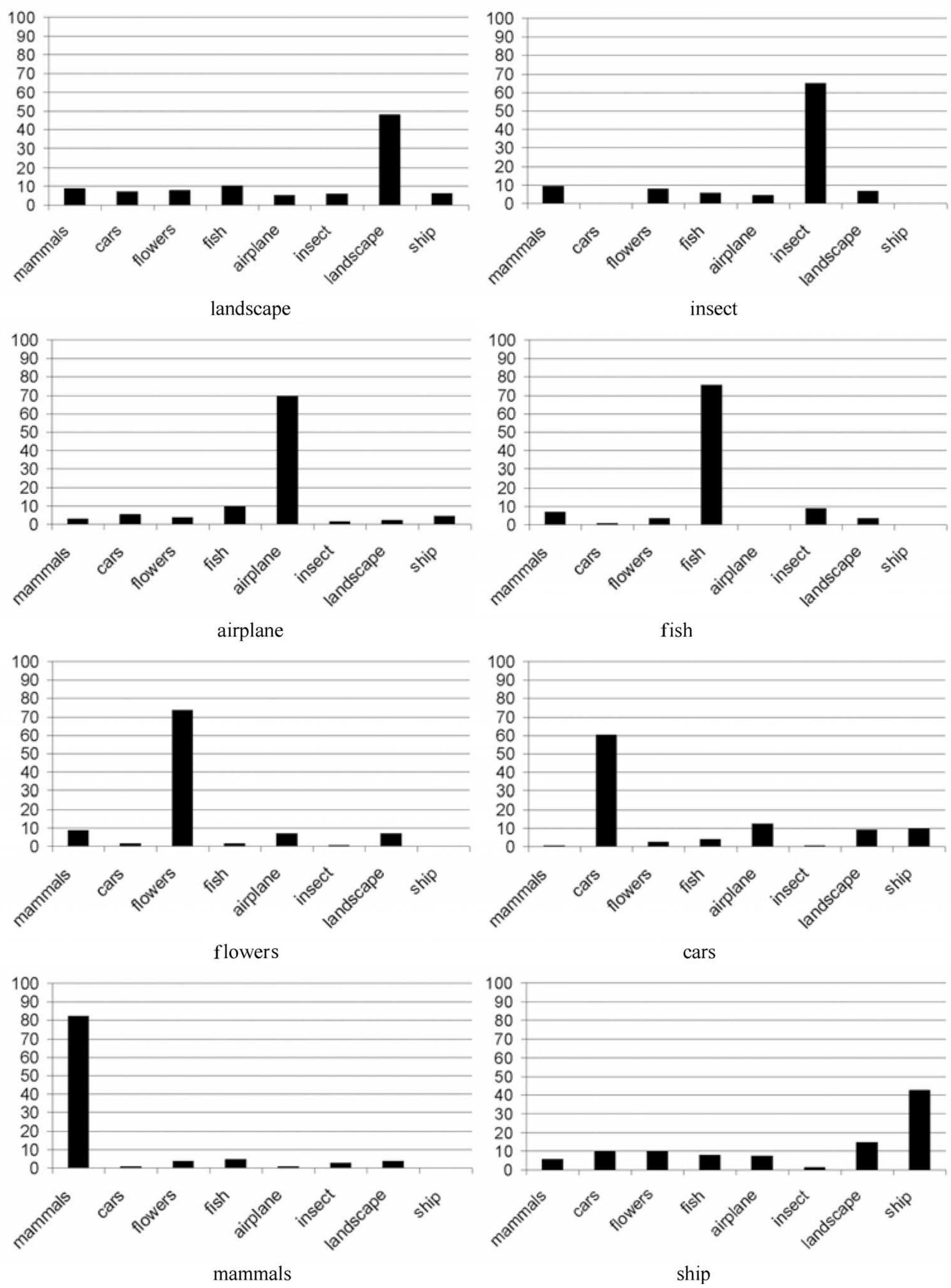


Fig. 25. Image distribution among the clusters for partially supervised FCM with 30% of labeled patterns.

The classification rate improves when the number of clusters goes up, which is quite anticipated (refer to Fig. 16).

The values of the separation coefficient represented in a histogram format (see Fig. 17), reveal that for a relatively broad range of values of the fuzzification coefficient (m), lower values of ψ_k are associated with the higher values of the classification error [see Fig. 17(b)]. This is intuitively appealing as the images whose allocation to clusters is quite "difficult" are those that could be potentially misclassified. For the lower values of "m," [see Fig. 17(a)], this tendency does not hold. The explanation is that for such low values of the fuzzification coefficient, the membership degrees turn out to be either 1 or 0 (with almost no values in between); hence, one should not treat the values of the separability index as being highly indicative of any meaningful tendency.

It is interesting to see what images of the clustering algorithm have been identified to be typical for the entire collection of images. Fig. 18 shows the images that are the closest to the prototypes; these were obtained when using $m = 1.50$ and $c = 8$. Note that the image that is the closest to a prototype is the image belonging to the cluster represented by that prototype and having the highest membership degree in this cluster. Fig. 19 shows the "typical" images that were obtained when running the algorithm for $m = 1.25$ and $c = 16$.

The correspondence between the clusters and the classes is presented in Table IV. The detailed classification results are

Fig. 26. Examples of images belonging to the noise cluster using partially supervised FCM with 30% of labeled images.

displayed in Fig. 20; here, we show an allocation of images belonging to different classes within the individual clusters.

It is noticeable that the clusters differ quite substantially with regard to their homogeneity. For instance, the cluster of insects looks very homogeneous; however, this is not true any longer for the clusters concerning the airplanes and mammals (here, we see a high overlap with the category of landscape; this perhaps explains why there was no landscape cluster in the resulting structure).

The increase in the number of clusters leads to the improvement of the classification rate. The homogeneity of the clusters themselves increased as well when compared to the previous situation of eight clusters. Also, this time, the cluster "insect" looks very homogeneous.

Noticeably, both in Table IV and Table V, there is a big number of landscapes (45 in Table IV and 72 in Table V) in the cluster "ship." It could well be that these two classes overlap to a very significant extent. Furthermore, in Table V, the noise cluster is smaller than the one reported in Table IV. This reduction has been anticipated, considering the definition of noise cluster.

Fig. 21 shows some examples of images belonging to the noise cluster. The situation is not that bad at all, given the fact that the categories overlap, and there are no "pure" classes we have originally defined. In many cases, we observe that the overlap between the landscape and others is quite visible—a phenomenon that is predictable when one agree that the landscape shows up many images classified as "ship," "mammals," etc.

Let us now move on to the use of the partially supervised FCM (here, we used $c = 8$ and $m = 2.0$; these values are essential to the efficient implementation of the supervision mechanism). As before, the experiments were repeated ten times for every scenario. The percentage of the labeled images varied from 10% to 30%. While the convergence of the FCM and its supervised version looks similar (see Fig. 22), the component of supervision helped us avoid some oscillations present when running the FCM itself. The classification rate improves with the increase of the percentage of the labeled images (refer to Fig 23). We stress the relationship between the size of the noise cluster and the mechanism of partial supervision. Using the COIL-20 image database, the number of images belonging to the noise cluster decreases when working with the partial supervision. While using the outdoor image database, this phenomenon is not observed. Furthermore, using this image database, the mean size of the noise cluster remains the same (27%). Analyzing the two image databases, these results are quite intuitive. Indeed, as we mentioned in the previous sections, there are not noise images in COIL-20, while in the outdoor image database, there is 20% of noise images (namely, those are the images belonging to more than a single class). Some examples of images classified as "noise" are presented in Fig. 26.

Fig. 24 shows the images that are located at the closest vicinity of the prototypes being obtained when using 30% of randomly labeled patterns. Similarly, Fig. 25 reports the distribution of the images across the clusters for supervision at the level of 30% (see also Table VI). It is noticeable that each class comes with its own cluster, and that the clusters themselves are more homogeneous than those obtained when not using the mechanisms of supervision.

Fig. 26 shows some examples of images belonging to the noise cluster that were obtained when using partially supervised FCM with 30% of the labeled patterns. Also, at this time, the overlap between the class "landscape" and the other classes is still present.

Fig. 27 presents a histogram of the values of Ψ for misclassified patterns obtained when running the partially supervised FCM ($m = 2$, $c = 4$, and 30% of randomly labeled images). There is a clear trend: lower values of Ψ come with the higher number of misclassified images.

For the reasons indicated in Section II, it is not possible to repeat the comparative evaluation of the results obtained with

TABLE VI
EXAMPLE OF CLUSTERS ASSOCIATION TO CLASSES FOR PARTIALLY SUPERVISED FCM WITH 30% OF LABELED PATTERNS; THE DOMINANT CLUSTER-CLASS ASSOCIATIONS (LINKAGES) ARE SHOWN IN BOLDFACE

Cluster\Class	mammals	cars	flowers	fish	airplane	insect	landscape	ship
1	19	16	17	23	12	13	**105**	14
2	13	0	11	8	6	**88**	9	0
3	4	7	5	13	**91**	2	3	6
4	8	1	4	**84**	0	10	4	0
5	11	2	**95**	2	9	1	9	0
6	1	**116**	5	8	24	1	18	19
7	**85**	1	4	5	1	3	4	0
8	20	35	37	29	27	5	54	**154**
Noise	89	72	72	78	80	127	44	57

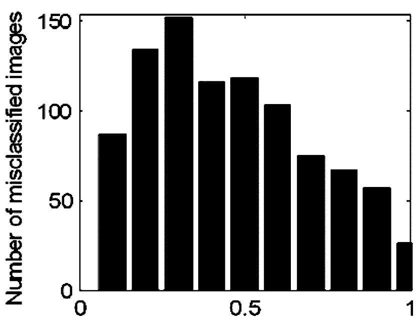


Fig. 27. Histogram of the values of for misclassified images; $c = 8$, $m = 2$, and 30% of labeled patterns.

this image database, as we did for those obtained using the COIL-20.

VII. CONCLUSION

In this study, we have proposed and investigated mechanisms of fuzzy clustering (FCM) and partially supervised clustering in the problems of organization and classification of images. We have demonstrated that the conceptual and algorithmic vehicle of clustering is a viable alternative owing to its abilities to reveal the structure in data and accommodate the mechanisms of relevance/feedback. On a more quantitative side based on the experimental evidence, several interesting points are worth making:

1) the choice of the number of clusters is important, and, in many cases, selecting more clusters than categories (classes) becomes highly beneficial;
2) it has been found that the selection of a suitable value of the fuzzification coefficient impacts the quality of the results; interestingly, this value is far lower than the one commonly encountered (and strangely enough, fully accepted) by the community. The myth of the "magic" value of 2.0 need not be followed, or, at least, it is worth exploring the optimization capabilities associated with this parameter;
3) partial supervision (which, in fact, models the effect of relevance feedback in processes of image organization) contributes to the improvement of the overall organization and classification processes;
4) the ambiguity of clustering (quantified in terms of the separability index) is strongly linked with misclassification (lower values of ψ_k are typically associated with the higher likelihood of classification error). In essence, fuzzy clustering has demonstrated its ability to flag the most "unclear" cases; and
5) the introduction of the concept of "noise cluster" along with the mechanisms of partial supervision augment the capabilities of the proposed approach to reveal the structure of image databases.

The notions of narrow and broad domain are essential when evaluating the performance of the results produced by the clustering procedures. Indeed, using a database belonging to a narrow domain (such as the COIL-20), the proposed system achieves the same performance as that being proposed in the literature when dealing with other systems. On the other hand, when dealing with broad domain image databases, the performance decreases due to the evident existence of the semantic gap.

Furthermore, the study has demonstrated that the description of images with the use of an angular spectrum signature is an efficient way proceeding with description and discrimination of images. The classification results obtained here are comparable with those reported in other studies, in which cases, the images were described in features spaces of a far higher dimensionality (where the use of three-image descriptors is quite typical).

An interesting alternative to exploiting mixed-mode feature spaces (such as, e.g., features and words) existing in some architectures, such as, e.g., Alipr (http://alipr.com), are also worth investigating and will be on the agenda of our future research pursuits.

REFERENCES

[1] A. Amato, T. Delvecchio, and V. Di Lecce, "Silhouettes based evaluation of the effectiveness in image retrieval," presented at the CRETA 6th WSEAS Int. Multiconf. CSCC, Rethymno, Crete Island, Greece, Jul. 2002.

[2] H. J. Bae and S. H. Jung, "Image retrieval using texture based on DCT," in *Proc. Int. Conf. Inf. Commun. Signal Process.*, Singapore, 1997, pp. 1065–1068.

[3] N. Boujemaa, J. Fauqueur, M. Ferecatu, F. Fleuret, V. Gouet, B. Le Saux, and H. Sahbi, "Ikona: Interactive generic and specific image retrieval," in *Proc. Int. 24th Workshop Multimedia Content-Based Indexing Retrieval (MMCBIR)*, Rocquencourt, France, 2001, pp. 25–28.
[4] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, New York: Dover, 1965
[5] A. Celentano and V. Di Lecce, "FFT-based technique for image-signature generation," in *Proc. SPIE Conf. Storage Retrieval Image Video Databases*, San Jose, CA, 1997, pp. 457–466.
[6] S.-F. Chang, T. Sikora, and A. Purl, "Overview of the MPEG-7 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 688–695, Jun. 2001.
[7] V. Di Lecce and A. Guerriero, "A comparative evaluation of retrieval methods for duplicate search in image database," *J. Visual Languages Comput.*, vol. 12, pp. 105–120, 2001.
[8] V. Di Lecce and A. Guerriero, "An evaluation of the effectiveness of image features for image retrieval," *J. Visual Commun. Image Representation*, vol. 10, pp. 351–362, 1999.
[9] V. Di Lecce, A. Guerriero, G. Dimauro, S. Impedovo, G. Pirlo, and A. Salzo, "Image basic features indexing techniques for video skimming," in *Proc. Int. Conf. Image Anal. Process.*, Venice, Italy, 1999, pp. 715–720.
[10] R. N. Davé, "Characterization and detection of noise in clustering," *Pattern Recog. Lett.*, vol. 12, no. 11, pp. 657–664, 1991.
[11] R. Dorairaj and K. R. Namuduri, "Compact combination of MPEG-7 color and texture descriptors for image retrieval," in *Proc. 38th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 2004, vol. 1, pp. 387–391.
[12] S. Eschrich, L. O. J. Ke Hall, and D. B. Goldgof, "Fast accurate fuzzy clustering through data reduction," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 2, pp. 262–270, Apr. 2003.
[13] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system," *IEEE Comput.*, vol. 28, no. 9, pp. 23–32, Sep. 1995.
[14] H. Frigui, "MembershipMap: Data transformation based on granulation and fuzzy membership aggregation," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 885–896, Dec. 2006.
[15] H. Frigui and R. Krishnapuram, "Clustering by competitive agglomeration," *Pattern Recog.*, vol. 30, no. 7, pp. 1223–1232, 1997.
[16] H. Frigui, N. Boujemaa, and S. A. Lim, "Unsupervised clustering and feature discrimination with application to image database categorization," in *Proc. IFSA World Congr. 20th NAFIPS Int. Conf.*, Vancouver, BC, Canada, 2001, pp. 401–406.
[17] 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.
[18] T. Gevers and A. W. M. Smeulders, "PicToSeek: Combining color and shape invariant features for image retrieval," *IEEE Trans. Image Process.*, vol. 9, no. 1, pp. 102–119, Jan. 2000.
[19] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
[20] J. B. Kruskal, "Multi-dimensional scaling by optimizing goodness-of-fit to a non-metric hypothesis," *Psychometrika*, vol. 29, pp. 1–27, 1964.
[21] J. A. Lay and L. Guan, "Image retrieval based on energy histograms of the low frequency DCT coefficients," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Phoenix, AZ, Mar. 1999, vol. 6, pp. 3009–3012.
[22] B. Le Saux and N. Boujemaa, "Unsupervised robust clustering for image database categorization," in *Proc. 16th Int. Conf. Pattern Recog.*, QC, Canada, Aug. 2002, pp. 259–262.
[23] S. Liapis and G. Tziritas, "Color and texture image retrieval using chromaticity histograms and wavelet frames," *IEEE Trans. Multimedia*, vol. 6, no. 5, pp. 676–686, Oct. 2004.
[24] A. W.-C. Liew, S. H. Leung, and W. H. Lau, "Segmentation of color lip images by spatial fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 4, pp. 542–549, Aug. 2003.
[25] E. Loupias, N. Sebe, S. Bres, and J.-M. Jolion, "Wavelet-based salient points for image retrieval," in *Proc. Int. Conf. Image Process.*, Vancouver, BC, Canada, Sep. 2000, vol. 2, pp. 518–521.
[26] H. Muller, S. Marchand-Maillet, and T. Pun, "The truth about Corel—Evaluation in image retrieval," in *Proc. Int. Conf. Image Video Retrieval, Lecture Notes Comput. Sci.,* Springer 2002, vol. 2383, pp. 36–45.
[27] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (COIL-100)," Dep. Comput. Sci., Columbia Univ., Rep. CUCS-006-96, 1996, [Online]. Available: http://www.cs.columbia.edu/CAVE/
[28] W. Pedrycz, "Algorithms of fuzzy clustering with partial supervision," *Pattern Recog. Lett.*, vol. 3, pp. 13–20, 1985.
[29] W. Pedrycz and J. Waletzky, "Fuzzy clustering with partial supervision," *IEEE Trans. Syst., Man, Cybern.*, vol. 27, no. 5, pp. 787–795, Sep. 1997.
[30] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of image databases," *Int. J. Comput. Vis.*, vol. 18, no. 3, pp. 233–254, 1996.
[31] S. Santini and R. Jain, "Similarity measures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 9, pp. 871–883, Sep. 1999.
[32] S. Santini and R. Jain, "The 'El Niño' image database system," in *Proc. IEEE Int. Conf. Multimedia Comput. Syst.*, Florence, Italy, 1999, vol. 1, pp. 524–529.
[33] T. Sikora, "The MPEG-7 visual standard for content description: An overview," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 696–702, Jun. 2001.
[34] P. Skulsujirapa, S. Aramvith, and S. Siddhichai, "Development of digital image retrieval technique using autocorrelogram and wavelet based texture," in *Proc. IEEE Int. MidWest Symp. Circuits Syst.*, Hiroshima, Japan, Jul. 2004, vol. 1, pp. 273–276.
[35] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content based image retrieval at the end of the early years," *IEEE Trans. PAMI*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.
[36] J. R. Smith and S. F. Chang, "Transform features for texture classification and discrimination in large image databases," in *Proc. Int. Conf. Image Process.*, Austin, TX, 1994, pp. 407–411.
[37] J. R. Smith and S. F. Chang, "Visualseek: A fully automated content based image query system," in *Proc. ACM Multimedia*, Boston, MA, Nov. 1996, pp. 87–98.
[38] C. F. Tsai, K. McGarry, and J. Tait, "Using neuro-fuzzy techniques based on a two-stage mapping model for concept-based image database indexing," in *Proc. IEEE 5th Int. Symp. Multimedia Softw. Eng.*, Taichung, Taiwan, 2003, pp. 6–12.
[39] A. Tversky, "Features of similarity," *Psychol. Rev.*, vol. 84, no. 4, pp. 327–352, 1977.
[40] A. Vellaikal and C.-C. J. Kuo, "Joint spatial-spectral indexing for image retrieval," in *Proc. 1996 IEEE Int. Conf. Image Process.*, Lausanne, Switzerland, part 3, pp. 867–870.
[41] Z. Wang, D. Feng, and Z. Chi, "Comparison of image partition methods for adaptive image categorization based on structural image representation," in *Proc. IEEE 8th Int. Conf. Control, Autom., Robot. Vis.*, Kunming, China, 2004, pp. 676–680.
[42] S. S. Yu, J.-R. Liou, and W.-C. Chen, "Computational similarity based on chromatic barycenter algorithm," *IEEE Trans. Consumer Electron.*, vol. 42, no. 2, pp. 216–220, May 1995.
[43] T. Yumin and M. Lixia, "Image retrieval based on multiple features using wavelet," in *Proc. 5th Int. Conf. Comput. Intell. Multimedia Appl.*, Xi'an, China, 2003, pp. 137–143.
[44] D. Zhang, "Improving image retrieval performance by using both color and texture features," in *Proc. 3rd Int. Conf. Image Graph.*, Hong Kong, 2004, pp. 172–175.

Witold Pedrycz (M'88–SM'90–F'99) received the M.Sc., Ph.D., and D.Sci. degrees from the Silesian University of Technology, Gliwice, Poland.

He is currently a Professor and Canada Research Chair (CRC) in Computational Intelligence in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is also with the Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland. He is the author or coauthor of numerous papers and 11 research monographs. His current research interests include computational intelligence, fuzzy modeling, knowledge discovery and data mining, fuzzy control including fuzzy controllers, pattern recognition, knowledge-based neural networks, granular and relational computing, and software engineering.

Dr. Pedrycz has been a member of numerous program committees of IEEE conferences in the area of fuzzy sets and neurocomputing. He is an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS, the IEEE TRANSACTIONS ON NEURAL NETWORKS, and the IEEE TRANSACTIONS ON FUZZY SYSTEMS. He is an Editor-in-Chief of *Information Sciences* and the president of the International Fuzzy Systems Association (IFSA).

Alberto Amato was born in Taranto, Italy, in 1976. He received the B.Sc. degree in environmental engineering from the Polytechnic of Bari, Bari, Italy, in 2003.

He is currently with the Polytechnic of Bari as a Software Designer. His current research interests include clustering algorithms, image retrieval, and knowledge-based systems.

Vincenzo Di Lecce received the Doctoral degree in electric engineering (*cum laude* with honors) from the University of Bari, Bari, Italy, in 1980.

He was a Selenia Spazio Italiana Fellow. During 1988, he was the Technical Chief of the Electronic Calculator Laboratories at the Engineering Faculty, Polytechnic of Bari, where he is currently an Associated Professor of Computer Science. He is the author or coauthor of about 100 papers. His current research interests include image retrieval and analysis, parallel and distributed computing, and the ontological approach to artificial intelligence.

Vincenzo Piuri (S'84–M'86–SM'96–F'01) obtained the Ph.D. degree in computer engineering from the Politecnico di Milano, Milan, Italy, in 1989.

From 1992 to 2000, he was with the Politecnico di Milano as an Associate Professor in operating systems. During 1993–1999, he was a Visiting Professor at the University of Texas, Austin. Since 2000, he has been with the University of Milan, Crema, Italy, as a Full Professor in computer engineering. He is the author or coauthor of more than 200 papers, book chapters, international journals, and proceedings of international conferences. His current research interests include theory and industrial applications of neural networks, intelligent measurement systems, biometrics, distributed and parallel computing systems, computer arithmetic, application-specific processing architectures, digital signal processing architectures, and fault tolerance.

Dr. Piuri is a member of the Association for Computing Machinery (ACM), the International Neural Network Society (INNS), and the Associated Electrical Industries (AEI). He was an Associate Editor of the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT and the IEEE TRANSACTIONS ON NEURAL NETWORKS. He was the Vice-President for publications of the IEEE Instrumentation and Measurement Society, the Vice-President for members activities of the IEEE Neural Networks Society, President of the IEEE Computational Intelligence Society (2006–2007), and member of the Administrative Committee, both of the IEEE Instrumentation and Measurement Society and the IEEE Computational Intelligence Society. He was the recipient of the IEEE Instrumentation and Measurement Society Technical Award 2002 for his contributions to the advancement of computational intelligence theory and practice in measurement systems and industrial applications.